Image Analysis And Computer Vision

Christian Rossi

Academic Year 2024-2025

Abstract

The course begins with an introduction to camera sensors, including their transduction, optics, geometry, and distortion characteristics. It then covers the basics of projective geometry, focusing on modeling fundamental primitives such as points, lines, planes, conic sections, and quadric surfaces, as well as understanding projective spatial transformations and projections.

The course continues with an exploration of camera geometry and single-view analysis, addressing topics like calibration, image rectification, and the localization of 3D models. This is followed by a study of multi-view analysis techniques, which includes 3D shape reconstruction, self-calibration, and 3D scene understanding.

Students will also learn about linear filters and convolutions, including space-invariant filters, the Fourier Transform, and issues related to sampling and aliasing. Nonlinear filters are discussed as well, with a focus on image morphology and operations such as dilation, erosion, opening, and closing, as well as median filters.

The course further explores edge detection and feature detection techniques, along with feature matching and tracking in image sequences. It addresses methods for inferring parametric models from noisy data and outliers, including contour segmentation, clustering, the Hough Transform, and RANSAC (random sample consensus).

Finally, the course applies these concepts to practical problems such as object tracking, recognition, and classification.

Contents

CHAPTER 1

Optical sensors

1.1 Camera

Definition (*Camera*)**.** A camera is an optical sensor that generates data using electric transducers. It features an optical system designed to direct incoming light to its millions of photosensitive elements. Modern cameras are typically capable of recording 30 to 60 frames per second.

For simplicity, we will consider the optical system of a camera as a single lens with the following characteristics:

- *Spherical*: the lens is formed by the intersection of two spherical surfaces.
- *Thin*: the distance between the centers of the two spheres is nearly equal to the sum of their radii.
- *Small angles*: the light rays make only slight angles with respect to the optical axis.

These assumptions simplify the calculations involved in determining the path of a light ray as it passes through the lens. Specifically, the refraction of light at the boundary between two media is described by Snell's law:

$$\frac{\sin\theta_2}{\sin\theta_1} = \frac{n_1}{n_2}$$

Here:

- θ_1 and θ_2 are the angles between the normal at the surface and the direction of the light ray before and after crossing the boundary, respectively.
- n_1 and n_2 are the refractive indices of the two materials.

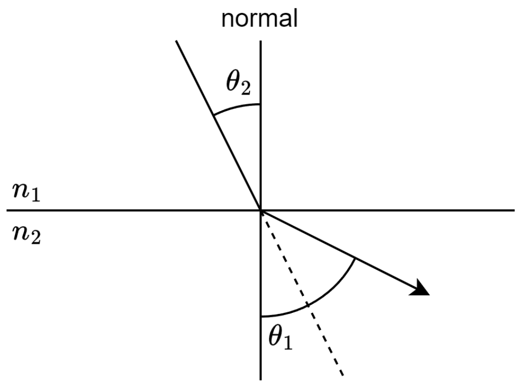
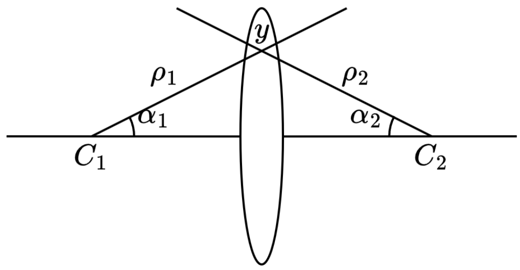


Figure 1.1: Snell's law

Definition (*Optical axis*)**.** The optical axis is the straight line that connects the centers of the two spheres that form the lens.

The angles of a ray passing through the centers of the spheres can be expressed as follows:

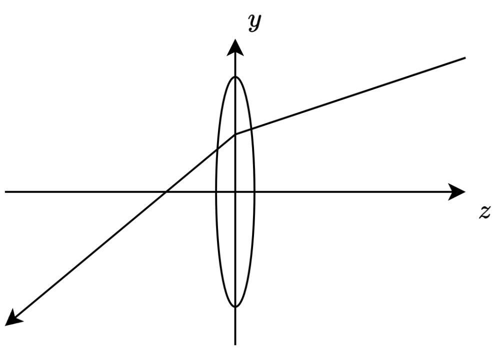
$$\alpha_1 = \frac{y_1}{\rho_1} \qquad \alpha_2 = -\frac{y_2}{\rho_2}$$



In this context, with the simplified lens, it is reasonable to assume:

$$y_1 = y_2 = y$$

1.2 Light rays deviation



For a lens with a refractive index n, the following equations apply:

$$\frac{\theta - \alpha_1}{\theta' - \alpha_1} \Rightarrow \frac{\sin(\theta - \alpha_1)}{\sin(\theta' - \alpha_1)} = n$$

$$\frac{\theta'' - \alpha_2}{\theta' - \alpha_2} \Rightarrow \frac{\sin(\theta'' - \alpha_2)}{\sin(\theta' - \alpha_2)} = n$$

Here:

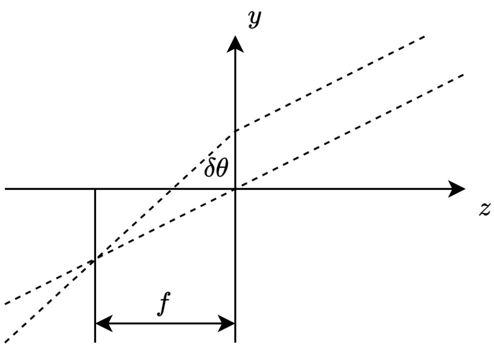
- θ is the angle of the incoming ray before entering the lens.
- θ' is the angle of the ray within the lens (not visible in the image).
- θ'' is the angle of the ray after exiting the lens.

By comparing these two equations, we can express the difference between the input angle θ and the output angle θ'' as:

$$\delta\theta = y(n-1)\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right)$$

Here, the term $n-1$ reflects the influence of the lens material, while the term $\frac{1}{\rho_1} + \frac{1}{\rho_2}$ is determined by the curvature of the lens surfaces.

1.3 Focalization of parallel light rays



In the image, we observe two rays: one passing through the center of the lens and another that remains parallel to the first ray but passes through a different point. From this, we can make the following observations:

- When $Y = 0$, the ray experiences no deviation and continues straight without deflection.
- Using the relationship $Y = f \cdot \delta\theta$, we can express the focal length of the lens as follows:

$$f = \frac{1}{(n-1)\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right)}$$

This indicates that all parallel rays converge at a common point known as the focal point, denoted as Z. The distance from the focal point to the y axis is given by:

$$Z = -f$$

1.4 Path of a light ray

To determine the trajectory of a light ray as it passes through a lens at any given position, you can follow these steps:

1. Draw a line parallel to the selected ray, passing through the center of the lens.
2. Identify the intersection point of this line with the focal plane.
3. The ray will travel from the point where it crosses the lens to the point on the focal plane.

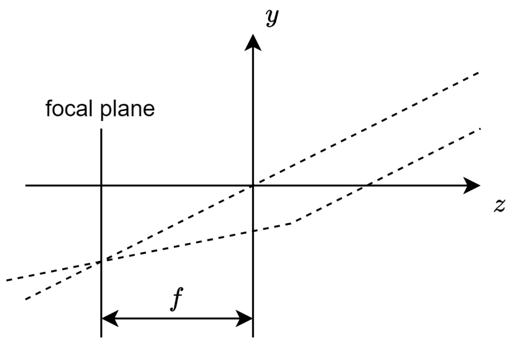


Figure 1.2: Path of a light ray through a lens

1.5 Pin-hole camera

To achieve a sharply focused image, it's crucial that each light ray converges precisely onto a single pixel on the camera's focal plane. To ensure this, the following conditions must be met:

- The distance between the lens and the light source, denoted as $Z(P)$, should be significantly greater than the lens aperture a ideally at least 1000 times larger.
- By positioning the screen at a distance Z from the lens, all rays can maintain parallel trajectories as they pass through the lens, resulting in a well-focused image.

The camera described is commonly known as a pin-hole camera, and it requires the following characteristics:

1. A thin spherical lens.
2. Utilization of small angles.
3. Ensuring that $Z(P) \gg a$.
4. Maintaining $Z = f$, where f represents the focal length.

1.6 From real world to two-dimensional images

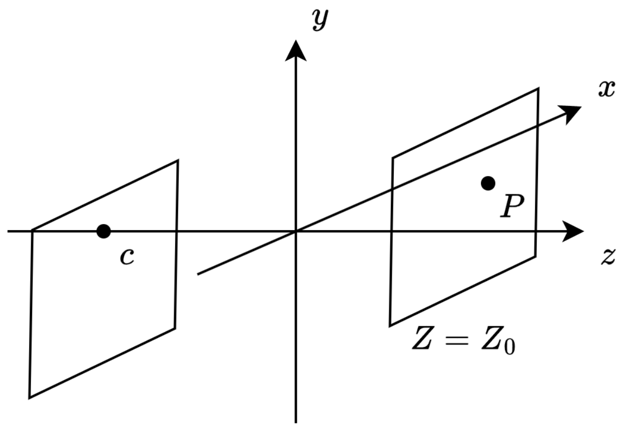
Images exist on a 2D plane, while the real world is three-dimensional, leading to a reduction of information compared to the original subject. This reduction is a result of perspective projection, which exhibits the following characteristics: nonlinearity, lack of shape preservation, and failure to maintain length ratios.

Using the triangle equality, we can express this perspective projection as:

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

To minimize information loss, one effective approach is to capture an image of a planar scene on a plane that is parallel to the image plane. This requires that:

$$Z = Z_0 = \text{constant}$$



In this scenario, the only difference between reality and the projection is a uniform downscaling, while other dimensions are preserved, yielding:

$$x = f\frac{X}{Z_0} = kX \qquad y = f\frac{Y}{Z_0} = kY$$

1.7 Perspective and vanishing point

When considering all lines parallel to the direction parameters $\begin{bmatrix}\alpha & \beta & 1\end{bmatrix}$, we can establish the following system of equations:

$$\begin{cases} X = X_0 + \alpha Z \\ Y = Y_0 + \beta Z \\ Z = 1 \cdot Z \end{cases}$$

To project these lines onto the 2D image using the triangle equality, we derive the following expressions:

$$x = f\frac{X}{Z} = f\frac{X_0 + \alpha Z}{Z} = f\alpha + \frac{X_0}{Z} \qquad y = f\frac{Y}{Z} = f\frac{Y_0 + \beta Z}{Z} = f\beta + \frac{Y_0}{Z}$$

Next, we find the image of the point at infinity along these lines, which results in the point:

$$\begin{bmatrix} f\alpha & f\beta \end{bmatrix}$$

Remarkably, this image point is independent of the values of X_0 and Y_0. Therefore, all parallel lines share the same image of their points at infinity.

Definition (*Vanishing point*)**.** The image of the point at infinity of the lines is known as the vanishing point.

Consequently, we observe that all parallel lines in the real world project onto converging lines in the image.

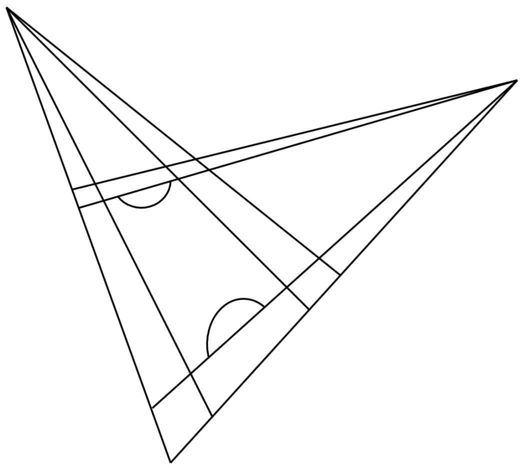


Figure 1.3: Vanishing point

CHAPTER 2

Two-dimensional planar projective geometry

2.1 Introduction

In the realm of planar geometry, the foundational elements consist of points, lines, conics, and dual conics. The transformations allowed within this geometry include projectivities, affinities, similarities, and isometries.

2.2 Points

To define points in Cartesian coordinates, we establish a Euclidean plane with a designated origin. Each point is uniquely represented by a pair of Cartesian coordinates, $\begin{bmatrix} x & y \end{bmatrix}$.

For image analysis, it is advantageous to use homogeneous coordinates. This involves constructing a 3D space with axes labeled $\begin{bmatrix} x & y & w \end{bmatrix}$. To represent a point, we assign three values, which allows for an infinite number of representations by varying the value of w. The relationship between Cartesian and homogeneous coordinates can be expressed as follows:

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

Consequently, a vector $\mathbf{x} = \begin{bmatrix} x & y & w \end{bmatrix}^T$ and all its nonzero multiples, including $\begin{bmatrix} \frac{x}{w} & \frac{y}{w} & 1 \end{bmatrix}^T$, represent the same point in Cartesian coordinates $\begin{bmatrix} X & Y \end{bmatrix}^T = \begin{bmatrix} \frac{x}{w} & \frac{y}{w} \end{bmatrix}^T$ on the Euclidean plane.

Property 2.2.1 (Homogeneity)**.** Any vector $\mathbf{x}$ is equivalent to all its nonzero multiples $\lambda\mathbf{x}$, where $\lambda \neq 0$, as they denote the same point.

The null vector does not represent any point.

Definition (*Projective plane*)**.** We define the projective plane as:

$$\mathbb{P}^2 = \left\{ \begin{bmatrix} x & y & w \end{bmatrix}^T \in \mathbb{R}^3 \right\} \setminus \left\{ \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \right\}$$

Example:
The origin of the plane is defined as:

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$$

A generic point in homogeneous coordinates can easily be transformed into Cartesian coordinates by simple division. For instance, the point:

$$\begin{bmatrix} 0 & 8 & 4 \end{bmatrix}^T$$

in Cartesian coordinates is:

$$\begin{bmatrix} \frac{x}{w} & \frac{y}{w} \end{bmatrix}^T = \begin{bmatrix} \frac{0}{4} & \frac{8}{4} \end{bmatrix} = \begin{bmatrix} 0 & 4 \end{bmatrix}$$

Consider a point $\mathbf{x} = \begin{bmatrix} x & y & w \end{bmatrix}^T$, and let w slowly decrease from $w = 1$. As w decreases, the point moves in a constant direction $\begin{bmatrix} x & y \end{bmatrix}$, distancing itself from the origin. As w approaches 0, the point tends toward infinity along the direction $\begin{bmatrix} x & y \end{bmatrix}$.

Definition (*Point at the infinity along the direction*)**.** We define the point at the infinity along the direction $\begin{bmatrix} x & y \end{bmatrix}$ as:

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

Points at infinity, representing directions, exist outside the Euclidean plane and are well-defined within the projective plane. Thus, the projective plane encompasses not only the Euclidean plane but also these points at infinity.

2.3 Lines

In the Euclidean plane, a line is typically defined by the equation:

$$aX + bY + c = 0$$

In the homogeneous plane, lines are represented as:

$$a\frac{x}{w} + b\frac{y}{w} + c = 0 \Longrightarrow ax + by + cw = 0$$

This equation can also be expressed using two vectors, denoted as $\mathbf{l}^T$ and $\mathbf{x}$, as follows:

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

Here, the vector $\mathbf{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$ epresents a line, with all its nonzero multiples also representing the same line.

Property 2.3.1 (Homogeneity)**.** Any vector $\mathbf{l}$ is equivalent to all its nonzero multiples, denoted as $\lambda\mathbf{l}$ (where $\lambda \neq 0$), since they denote the same line.

The coefficients a, b, and c are known as the homogeneous parameters of the line.

Similar to numbers, there are multiple equivalent representations for a single line, specifically all nonzero multiples of the unit normal vector. However, the null vector does not represent any lines.

Definition (*Projective dual plane*)**.** The projective dual plane is defined as:

$$\mathbb{P}^2 = \left\{ \begin{bmatrix} a & b & c \end{bmatrix}^T \in \mathbb{R}^3 \right\} \setminus \left\{ \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \right\}$$

Property 2.3.2. If the third parameter is zero, denoted as $\mathbf{l} = \begin{bmatrix} a & b & 0 \end{bmatrix}^T$, then the line passes through the point $\begin{bmatrix} 0 & 0 \end{bmatrix}$.

Property 2.3.3. In the Euclidean plane, the direction $\begin{bmatrix} a & b \end{bmatrix}$ is perpendicular to the line represented by $\mathbf{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$.

Property 2.3.4. Two lines, $\mathbf{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$ and $\mathbf{l} = \begin{bmatrix} a & b & c' \end{bmatrix}^T$, are considered parallel if they share the same direction, represented by $\begin{bmatrix} b & -a \end{bmatrix}$.

> **Example:**
> The Cartesian axes are defined as:
>
> $$\mathbf{l}_x = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$$
>
> $$\mathbf{l}_y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$$

In this context, the incidence relation of a line $\mathbf{l}^T\mathbf{x} = 0$ is defined when the point $\mathbf{x}$ lies on the line $\mathbf{l}$ or when the line $\mathbf{l}$ goes through the point $\mathbf{x}$.

Definition (*Line at the infinity*)**.** The line

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} = w = 0$$

is called the line at the infinity, denoted as $\mathbf{l}_\infty = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$.

The principle of duality between points and lines states that the incidence relation is commutative, as the dot product is commutative.

To find the intersection of two lines, $\mathbf{l}_1$ and $\mathbf{l}_2$, the following condition is imposed:

$$\begin{bmatrix} \mathbf{l}_1^T \\ \mathbf{l}_2^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This equation leads to finding the right null space of the matrix formed by stacking the line vectors:

$$\mathbf{x} = \mathrm{RNS}\left(\begin{bmatrix} \mathbf{l}_1^T \\ \mathbf{l}_2^T \end{bmatrix} \right)$$

The system is under-determined, meaning there is only one intersection point between the two lines, which can be represented in multiple ways using homogeneous coordinates. In 2D projective geometry, the vector $\mathbf{x}$, orthogonal to both lines, can be found using the cross product:

$$\mathbf{x} = \mathbf{l}_1 \times \mathbf{l}_2$$

Example:
Suppose we have two parallel lines $\mathbf{l}_1 = \begin{bmatrix} a & b & c_1 \end{bmatrix}^T$ and $\mathbf{l}_2 = \begin{bmatrix} a & b & c_2 \end{bmatrix}^T$. The point common to both lines can be found using the system:

$$\begin{cases} ax + by + c_1 w = 0 \\ ax + by + c_2 w = 0 \end{cases}$$

The solution is $\mathbf{x} = \begin{bmatrix} b & -a & 0 \end{bmatrix}^T$, which represents the point at infinity in the direction of both lines.

The line passing through two points can be found using the dual of the previous problem, expressed as:

$$\begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \end{bmatrix} \mathbf{l} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In 2D, this simplifies to:

$$\mathbf{l} = \mathbf{x}_1 \times \mathbf{x}_2$$

Property 2.3.5. A point $\mathbf{x}$, which is a linear combination $\mathbf{x} = \alpha \mathbf{x}_1 + \beta \mathbf{x}_2$ of two points $\mathbf{x}_1$ and $\mathbf{x}_2$, lies on the line $\mathbf{l}$ through $\mathbf{x}_1$ and $\mathbf{x}_2$.

Proof. The line $\mathbf{l}$ passing through both points satisfies $\mathbf{l}^T \mathbf{x}_1 = 0$ and $\mathbf{l}^T \mathbf{x}_2 = 0$. By adding α times the first equation to β times the second one, we obtain:

$$0 = \mathbf{l}^T \left(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2 \right) = \mathbf{l}^T \mathbf{x} = 0$$

□

This establishes the duality between collinearity and concurrence.

Theorem 2.3.1. *For any true sentence containing the terms: point, line, is on, goes through, co-linear, and concurrent, there exists a dual statement (also true) obtained by making the following replacements:*

- *Point* $\Leftrightarrow$ *line.*
- *Is on* $\Leftrightarrow$ *goes through.*
- *Co-linear* $\Leftrightarrow$ *concurrent.*

In the Euclidean plane, the direction normal to the line $\mathbf{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$ is represented by $\begin{bmatrix} a & b \end{bmatrix}$. The relationship between lines can be understood by recognizing that the angle between two lines is equal to the angle between their normal vectors. The formula for the angle between two vectors is:

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| \, |\mathbf{v}|}$$

This applies to the angle between two lines $\mathbf{l}_1 = \begin{bmatrix} a_1 & b_1 & c_1 \end{bmatrix}^T$ and $\mathbf{l}_2 = \begin{bmatrix} a_2 & b_2 & c_2 \end{bmatrix}^T$. In this context, it is the angle between their respective normal directions $\begin{bmatrix} a_1 & b_1 \end{bmatrix}$ and $\begin{bmatrix} a_2 & b_2 \end{bmatrix}$, calculated as:

$$\cos\theta = \frac{a_1 a_2 + b_1 b_2}{\sqrt{\left(a_1^2 + b_1^2\right)\left(a_2^2 + b_2^2\right)}}$$

2.3.1 Cross ratio

Now, consider a line with four points related as follows:

$$\mathbf{x}_1 = \alpha_1 \mathbf{y} + \beta_1 z \qquad \mathbf{x}_2 = \alpha_2 \mathbf{y} + \beta_2 \mathbf{z}$$

The cross ratio is given by:

$$CR_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{z}} = \frac{\left(\dfrac{c-a}{c-b}\right)}{\left(\dfrac{d-a}{d-b}\right)} = \frac{\left(\dfrac{\beta_1}{\alpha_1}\right)}{\left(\dfrac{\beta_2}{\alpha_2}\right)} = \frac{\beta_1\alpha_2}{\beta_2\alpha_1}$$

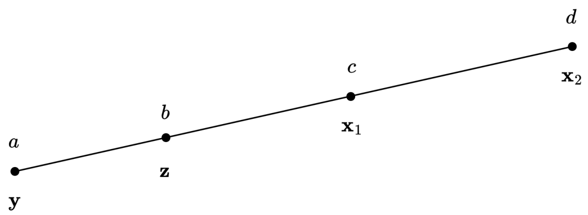
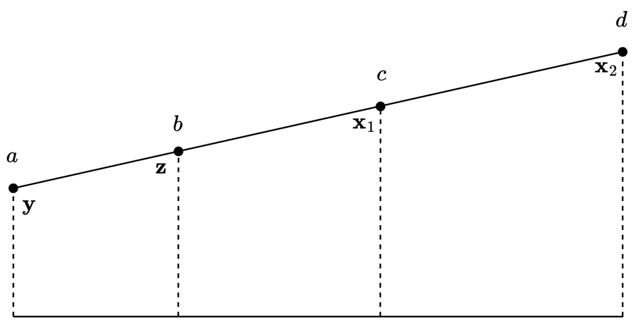


Figure 2.1: Line with the point of previous problem

Proof. Since the abscissae are proportional, the abscissae can be replaced by the x-coordinate, as shown in the figure.



The relation:

$$CR_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{z}} = \frac{\left(\dfrac{c-a}{c-b}\right)}{\left(\dfrac{d-a}{d-b}\right)}$$

still holds. If we consider $\mathbf{y} = \begin{bmatrix} y & * & v \end{bmatrix}^T$ and $\mathbf{z} = \begin{bmatrix} z & * & w \end{bmatrix}^T$, we can determine that:

$$\mathbf{x}_1 = \begin{bmatrix} \alpha_1 y + \beta_1 z \\ * \\ \alpha_1 v + \beta_1 w \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} \alpha_2 y + \beta_2 z \\ * \\ \alpha_2 v + \beta_2 w \end{bmatrix}$$

The difference between the x coordinates of $\mathbf{x}_1$ and $\mathbf{y}$ is:

$$c - a = \frac{\beta_1(zv - yw)}{(\alpha_1 y + \beta_1 z)v}$$

Similarly, the difference between the x coordinates of $\mathbf{x}_1$ and $\mathbf{z}$ is:

$$c - b = \frac{-\alpha_1(zv - yw)}{(\alpha_1 y + \beta_1 z)w}$$

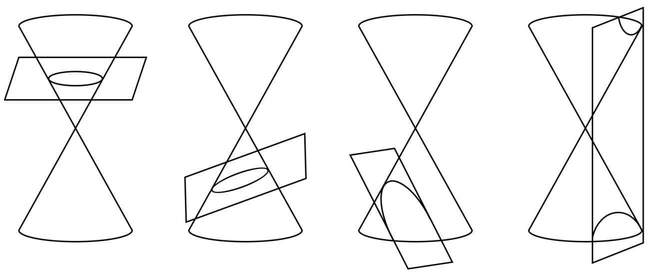
Substituting these expressions yields:

$$\frac{c-a}{c-b} = -\frac{\beta_1 w}{\alpha_1 v} \qquad \frac{d-a}{d-b} = -\frac{\beta_2 w}{\alpha_2 v}$$

□

2.4 Conics

Conics are the geometric shapes formed by the intersection of a cone and a plane. These shapes include circles, ellipses, parabolas, and hyperbolas, as illustrated in the figure below:



Definition (*Conic*)**.** A point $\mathbf{x}$ lies on a conic $\mathbf{C}$ if it satisfies the homogeneous quadratic equation $\mathbf{x}^T C\mathbf{x} = 0$, where $\mathbf{C}$ is a symmetric matrix, a standard convention in defining conics.

Conics are curves that can be described by second-degree equations in the plane. In Euclidean coordinates, a conic is expressed as:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

In homogeneous coordinates, this equation becomes:

$$ax^2 + bxy + cy^2 + dxw + eyw + fw^2 = 0$$

Alternatively, conics can be represented in matrix form as:

$$\mathbf{x}^T \begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & f \end{bmatrix} \mathbf{x} = 0$$

Conics have five degrees of freedom, meaning that five points are sufficient to uniquely determine a conic.

> **Example:**
> A circle in Cartesian coordinates is represented as:
>
> $$(x - x_0)^2 + (y - y_0)^2 - r^2 = 0$$

In homogeneous coordinates, it is given by:

$$\begin{bmatrix} x & y & w \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ -x_0 & -y_0 & x_0^2 + y_0^2 - r^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} = 0$$

When a conic is represented by a quadratic equation and a line by a linear equation, their intersection results in a second-degree equation in the point $\mathbf{x}$. Therefore, a line and a conic always intersect at two points, which can be categorized as follows:

- *Real and distinct*: two separate real points.
- *Real and coincident*: one real point, a repeated or double root.
- *Complex and distinct*: two distinct complex conjugate points.
- *Complex and coincident*: one complex point, a repeated or double root.

This is due to the fundamental theorem of algebra, which guarantees two solutions to a second-degree equation when considering complex numbers.

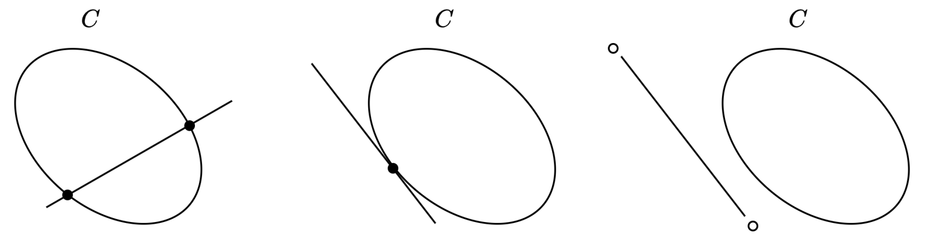


Figure 2.2: Intersection with two real roots, two coincident roots and two imaginary roots

The intersection of the line at infinity with a conic yields the following results:

- *Parabola*: two coincident solutions, representing the point at infinity along the axis.
- *Ellipse*: two complex-conjugate solutions, indicating no real solutions.
- *Hyperbola*: two real and distinct solutions, corresponding to the asymptotes.

2.4.1 Circular points

Example:
When a circle is intersected with the line at infinity, the system becomes:

$$\begin{cases} x^2 - 2x_0 w + x_0^2 w^2 + y^2 - 2y_0 w + y_0^2 w^2 - r^2 w^2 = 0 \\ w = 0 \end{cases}$$

This simplifies to:

$$x^2 + y^2 = 0$$

Here, the parameters of the circle (center and radius) disappear, indicating that the two intersection points are the same for all circles.

Definition (*Circular points*)**.** The two intersection points obtained by intersecting any circle with the line at infinity are called the circular points. These points are:

$$\mathbf{I} = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} \qquad \mathbf{J} = \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}$$

2.4.2 Polar line

Definition (*Polar line*)**.** Given a point $\mathbf{y}$ and a conic $\mathbf{C}$, the line $\mathbf{l} = \mathbf{C}\mathbf{y}$ is called the polar line of the point $\mathbf{y}$ with respect to the conic $\mathbf{C}$.

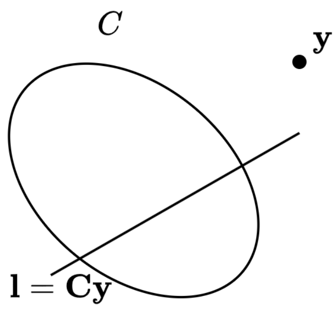


Figure 2.3: Polar line

2.4.3 Harmonic tuples

Definition (*Harmonic tuple*)**.** A 4-tuple of collinear points $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ is called a harmonic tuple if their cross ratio is equal to -1.

A harmonic cross ratio value is shared by other collinear 4-tuples. A common example is:

$$(\mathbf{y}, \mathbf{z}, \text{midPoint}(\mathbf{y}, \mathbf{z}), P(\infty))$$

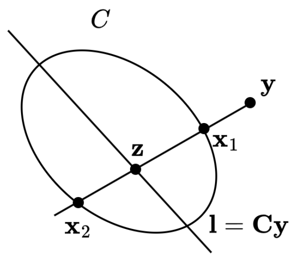
Moreover, if $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ is harmonic, then $(\mathbf{c}, \mathbf{d}, \mathbf{a}, \mathbf{b})$ is also harmonic.

Definition (*Conjugate points*)**.** In a harmonic tuple $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$, the points $\mathbf{a}$ and $\mathbf{b}$ are said to be conjugate with respect to $\mathbf{c}$ and $\mathbf{d}$.

Since the cross ratio of a harmonic tuple is negative, conjugate points $\mathbf{a}$ and $\mathbf{b}$ lie on opposite sides of the segment formed by $\mathbf{c}$ and $\mathbf{d}$, with one point inside and the other outside the segment.

2.4.4 Polar line and harmonic tuples

Given any point $\mathbf{z}$ on the polar line $\mathbf{l} = \mathbf{C}\mathbf{y}$, consider the line passing through the points $\mathbf{y}$ and $\mathbf{z}$. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ represent the points at which this line intersects the conic $\mathbf{C}$.



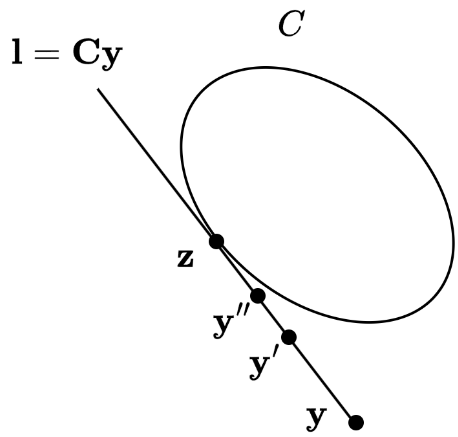
Theorem 2.4.1. *If* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are the points where the line passing through* $\mathbf{y}$ *and* $\mathbf{z}$ *intersects the conic* $\mathbf{C}$*, then* $\mathbf{y}$ *and* $\mathbf{z}$ *are conjugate with respect to* $\mathbf{x}_1$ *and* $\mathbf{x}_2$*.*

The polar line $\mathbf{l} = \mathbf{C}\mathbf{y}$ contains all points that are conjugate to $\mathbf{y}$ with respect to the conic $\mathbf{C}$. Specifically, it includes points that are conjugate with respect to the intersection points of $\mathbf{C}$ with any line passing through $\mathbf{y}$.

2.4.5 Polar line and tangency points

As the line through $\mathbf{y}$ approaches tangency with the conic $\mathbf{C}$, the points $\mathbf{x}_1$ and $\mathbf{x}_2$ merge into the point of tangency. Consequently, the conjugate point $\mathbf{z}$ also coincides with the tangency point, applying to any line tangent to $\mathbf{C}$ from point $\mathbf{y}$.

Therefore, the polar line $\mathbf{l} = \mathbf{C}\mathbf{y}$ passes through the tangency points where lines from $\mathbf{y}$ meet the conic $\mathbf{C}$. If a point $\mathbf{z}$ lies on the conic, $\mathbf{y}$ is one of its conjugates with respect to the same conic. The tangent line $\mathbf{lz}$ to the conic at point $\mathbf{z}$ contains points conjugate to $\mathbf{z}$, making $\mathbf{lz}$ the polar line of $\mathbf{z}$ with respect to the conic.



In the illustration, the polar line $\mathbf{lz} = \mathbf{C}\mathbf{z}$ for point $\mathbf{z}$ on the conic $\mathbf{C}$ corresponds to the tangent line to the conic at point $\mathbf{z}$.

Example:

Consider a circle with radius r centered at the origin and the point $\mathbf{y} = \begin{bmatrix} x & 0 & 1 \end{bmatrix}^T$. The equation for the polar line is:

$$\mathbf{l} = \mathbf{C}\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ 0 \\ -r^2 \end{bmatrix}$$

The Cartesian equation of the polar line becomes:

$$\mathbf{x}x - r^2 = 0 \rightarrow x = \frac{r^2}{\mathbf{x}}$$

This represents a vertical line.

From this, we deduce that the polar of a point $\mathbf{p}$ with respect to a circle is a line perpendicular to the line segment connecting the center of the circle to $\mathbf{p}$.

Example:

For a circle with radius r centered at the origin and the point $\mathbf{y} = \begin{bmatrix} x & 0 & 0 \end{bmatrix}^T$, the polar line equation is:

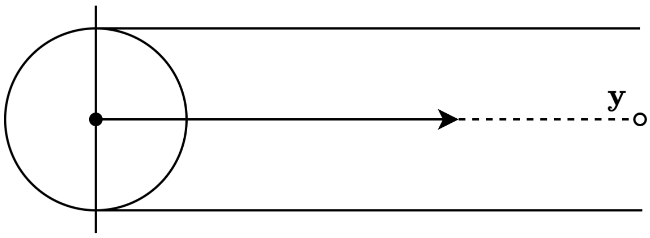
$$\mathbf{l} = \mathbf{C}\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ 0 \\ 0 \end{bmatrix}$$

The Cartesian equation becomes:

$$\mathbf{x}x = 0 \rightarrow \mathbf{x} = 0$$

This equation describes the diameter of the circle perpendicular to the direction of $\mathbf{y}$.

Parallel tangent lines from a point at infinity will have tangency points lying along a diameter that is perpendicular to the direction of the tangents.



Example:

For a circle with radius r centered at the origin and the point $\mathbf{y} = \begin{bmatrix} x & 0 & 0 \end{bmatrix}^T$, the polar line equation is:

$$\mathbf{l} = \mathbf{C}\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -r^2 \end{bmatrix}$$

The Cartesian equation becomes:

$$-r^2 w = 0 \rightarrow \mathbf{x} = 0$$

This equation describes the line at infinity.

Property 2.4.1. The polar line of any point at infinity is a diameter.

Property 2.4.2. Any diameter passes through the center of the circle.

Property 2.4.3. The center is conjugate to every point at infinity.

Property 2.4.4. All points at infinity are conjugate to the center.

Property 2.4.5. The polar of the center is the line that includes all points at infinity.

Property 2.4.6. The polar line of the center is the line at infinity.

2.4.6 Degenerate conics

Definition (*Non-degerate conic*)**.** A non-degenerate conic has a non-singular matrix $\mathbf{C}$, meaning:

$$\text{rank}(\mathbf{C}) = 3$$

Definition (*Degerate conic*)**.** A degenerate conic has a singular matrix $\mathbf{C}$, meaning:

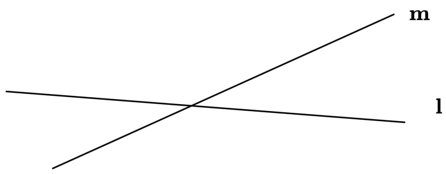
$$\text{rank}(\mathbf{C}) < 3$$

There are two types of degenerate conic:

- *Rank 2*: when $\text{rank}(\mathbf{C}) = 2$, $\mathbf{C}$ can be written as:

$$\mathbf{C} = \mathbf{l}\mathbf{m}^T + \mathbf{m}\mathbf{l}^T$$

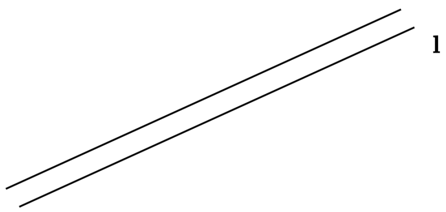
Here, $\mathbf{x}$ satisfies $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$ when either $\mathbf{x}^T\mathbf{l} = 0$ or $\mathbf{m}^T\mathbf{x} = 0$, meaning $\mathbf{x}$ lies on the union of lines $\mathbf{l}$ and $\mathbf{m}$.



- *Rank 1*: when $\text{rank}(\mathbf{C}) = 1$, $\mathbf{C}$ can be written as:

$$\mathbf{C} = \mathbf{l}\mathbf{l}^T$$

The conic consists of points $\mathbf{x}$ that satisfy $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$, meaning $\mathbf{x}$ lies on the repeated line $\mathbf{l}$.



2.5 Dual conics

Definition (*Dual conic*)**.** A dual conic is a set of lines $\mathbf{l}$ that satisfy equation:

$$\mathbf{l}^T\mathbf{C}^*\mathbf{l} = 0$$

where $\mathbf{C}^*$ is a 3×3 symmetric matrix.

Definition (*Non degenrate dual conic*)**.** A non degenerate dual conic is a dual conic whose matrix $\mathbf{C}^*$ is non-singular:

$$\text{rank}(\mathbf{C}^*) = 3$$

Consider a non-degenerate conic, denoted as $\mathbf{C}$, and the collection of all lines $\mathbf{l}$ that are tangents to it. For each point $\mathbf{c}$ on the conic $\mathbf{C}$, there exists a line $\mathbf{l}$ that is tangent to $\mathbf{C}$. Since $\mathbf{l}$ is the polar line of $\mathbf{x}$ with respect to $\mathbf{C}$, we can express it as $\mathbf{l} = \mathbf{C}\mathbf{c}$. Consequently, we can represent $\mathbf{x}$ as:

$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{l}$$

Moreover, given that $\mathbf{C}$ is a symmetric matrix, we have:

$$\mathbf{x}^T = \mathbf{l}^T\mathbf{l}^{-T} = \mathbf{l}^T\mathbf{C}^{-1}$$

Now, considering that the point $\mathbf{x}$ lies on the conic $\mathbf{C}$, we have:

$$\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$$

By substituting the previously derived expressions, we arrive at:

$$\mathbf{l}^T\mathbf{C}^{-1}\mathbf{l} = 0$$

This equation represents a quadratic homogeneous equation on $\mathbf{l}$. Therefore, we can conclude that for the dual conic holds $\mathbf{C}^* = \mathbf{C}^{-1}$. We can also note that a non-degenerate dual conic $\mathbf{C}^*$ is the collection of lines that are tangent to a non-degenerate conic $\mathbf{C}$.

2.5.1 Degenerate dual conics

Definition (*Degerate dual conic*)**.** A degenerate dual conic is a conic where the matrix $\mathbf{C}^*$ is singular:

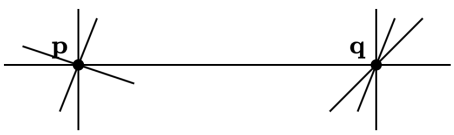
$$\text{rank}(\mathbf{C}^*) < 3$$

There are two possible scenarios to consider:

- When $\text{rank}(\mathbf{C}^*) = 2$, any symmetric 3×3 matrix $\mathbf{C}^*$ can be expressed as:

$$\mathbf{C}^* = \mathbf{p}\mathbf{q}^T + \mathbf{q}\mathbf{p}^T$$

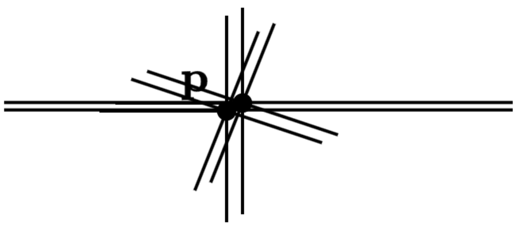
  In this case, the conic represents the line $\mathbf{l}$ passing through point $\mathbf{p}$ or the line $\mathbf{l}$ passing through point $\mathbf{q}$.



- When $\text{rank}(\mathbf{C}^*) = 1$, any symmetric 3×3 matrix $\mathbf{C}^*$ can be expressed as:

$$\mathbf{C}^* = \mathbf{p}\mathbf{p}^T$$

  In this situation, the conic corresponds to the line $\mathbf{l}$ going through point $\mathbf{p}$ repeated twice.



Definition (*Conic dual to the circular points*)**.** The degenerate dual conic $\mathbf{C}^* = \mathbf{p}\mathbf{q}^T + \mathbf{q}\mathbf{p}^T$ going through two circular point $\mathbf{p}$ and $\mathbf{q}$ is known as the conic dual to the circular points, and it can be expressed as:

$$\mathbf{C}^*_\infty = \mathbf{I}\mathbf{J}^T + \mathbf{J}\mathbf{I}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

2.6 Transformations

Definition (*Projective mapping*)**.** A projective mapping between a projective plane $\mathbb{P}^2$ and another projective plane $\mathbb{P}'^2$ is an invertible mapping which preserves co-linearity:

$$h : \mathbb{P}^2 \to \mathbb{P}'^2, x' = h(\mathbf{x}), \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \text{ are colinear}$$

$$\Leftrightarrow$$

$$\mathbf{x}'_1 = h(\mathbf{x}_1), \mathbf{x}'_2 = h(\mathbf{x}_2), \mathbf{x}'_3 = h(\mathbf{x}_3) \text{ are colinear}$$

Projective mapping is also called projectivity or homography.

Theorem 2.6.1. *A mapping $h : \mathbb{P}^2 \to \mathbb{P}'^2$ is projective if and only if there exists an invertible 3×3 matrix $\mathbf{H}$ such that for any point in $\mathbb{P}^2$ represented by the vector $\mathbf{x}$, is $h(\mathbf{x}) = \mathbf{H}\mathbf{x}$, where:*

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h13 \\ h_{21} & h_{22} & h23 \\ h_{31} & h_{32} & h33 \end{bmatrix}$$

Projective mappings are linear when expressed in homogeneous coordinates, but they do not exhibit linearity when represented in Cartesian coordinates.

According to the theorem, if we have $h(\mathbf{x}) = \mathbf{x}' = \mathbf{H}\mathbf{x}$, then multiplying the matrix $\mathbf{H}$ by any nonzero scalar λ still satisfies the relation for the same points, giving us $\mathbf{x}' = \lambda\mathbf{H}\mathbf{x}$. Therefore, any nonzero scalar multiple of the matrix $\mathbf{H}$ represents the same projective mapping as $\mathbf{H}$. As a result, we can conclude that $\mathbf{H}$ is a homogeneous matrix. Despite having nine entries, it possesses only eight degrees of freedom, specifically the ratios between its elements. Consequently, we can estimate $\mathbf{H}$ using just four point correspondences. Each point correspondence, expressed as $\mathbf{x}' = \mathbf{H}\mathbf{x}$, provides two independent equations in this estimation process.

Definition (*Homography*)**.** A homography transforms various geometric entities as follows:

1. It maps a point $\mathbf{x}$ to a point $\mathbf{x}'$, where the transformation is expressed as:

$$\mathbf{x} \to \mathbf{H}\mathbf{x} = \mathbf{x}'$$

2. It maps a line $\mathbf{l}$ to a line $\mathbf{l}'$, and this transformation is represented as:

$$\mathbf{l} \to \mathbf{H}^{-T}\mathbf{l} = \mathbf{l}'$$

3. It maps a conic $\mathbf{C}$ to a conic $\mathbf{C}'$, and the transformation is given by:

$$\mathbf{C} \to \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1} = \mathbf{C}'$$

4. It maps a dual conic $\mathbf{C}^*$ to a dual conic $C^{*\prime}$, with the transformation being:

$$\mathbf{C}^* \rightarrow \mathbf{H}\mathbf{C}^*\mathbf{H}^T = \mathbf{C}^{*\prime}$$

Proof of mapping two. To transform the equation of the line in terms of $\mathbf{x}$, given by $\mathbf{l}^T\mathbf{x} = 0$, into a constraint on $\mathbf{x}' = \mathbf{H}\mathbf{x}$, we combine the two equations, resulting in a linear equation on $\mathbf{x}'$:

$$\mathbf{l}'^T\mathbf{x}' = 0$$

Here, $\mathbf{l}'^T = \mathbf{l}^T\mathbf{H}^{-1}$. Thus, we have:

$$\mathbf{l}' = \mathbf{H}^{-T}\mathbf{l}$$

□

Proof of mapping three. To transform the equation of the conic in terms of $\mathbf{x}$, given by $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$, into a constraint on $\mathbf{x}' = \mathbf{H}\mathbf{x}$, we have $\mathbf{x} = \mathbf{H}^{-1}\mathbf{x}'$ and $\mathbf{x}^T = \mathbf{x}'^T\mathbf{H}^{-T}$. Combining these three equations, we obtain a linear equation on $\mathbf{x}'$:

$$\mathbf{x}'^T\mathbf{C}'\mathbf{x}' = 0$$

Hence, we have:

$$\mathbf{C}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}$$

□

Proof of mapping four. For the transformation of a dual conic, we apply the same idea, yielding:

$$\mathbf{C}^{*\prime} = \mathbf{H}\mathbf{C}^*\mathbf{H}^T$$

□

The point-line incidence is preserved.

Proof. Let $\mathbf{x}$ be a point on the line $\mathbf{l}$. This is expressed as $\mathbf{l}^T\mathbf{x} = 0$. When we apply the projective transformation $\mathbf{H}$ to both $\mathbf{x}$ and $\mathbf{l}$, resulting in $\mathbf{H}\mathbf{x} = \mathbf{x}'$ and $\mathbf{H}^{-1}\mathbf{l} = \mathbf{l}'$, they remain incident if $\mathbf{l}'^T\mathbf{x}' = 0$:

$$\mathbf{l}'^T\mathbf{x}' = \mathbf{l}^T\mathbf{H}^{-1}\mathbf{x}' = \mathbf{l}^T\mathbf{H}^{-1}\mathbf{H}\mathbf{x} = \mathbf{l}^T x = 0$$

□

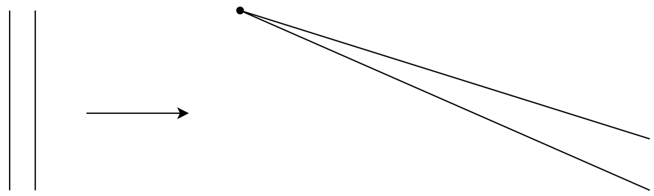
2.6.1 Vanishing points and vanishing line

The point that is common to both parallel lines $\mathbf{l}_1 = \begin{bmatrix} a & b & c_1 \end{bmatrix}^T$ and $\mathbf{l}_2 = \begin{bmatrix} a & b & c_2 \end{bmatrix}^T$ is the point $\mathbf{x} = \begin{bmatrix} b & -a & 0 \end{bmatrix}^T$. This point is situated at infinity along the direction of both lines. When seeking the common point of the infinite lines $\mathbf{l}_i$, we find that they all share the same point:

$$\mathbf{x}_\infty = \begin{bmatrix} b & -a & 0 \end{bmatrix}^T$$

Hence, it becomes apparent that all these lines converge at $\begin{bmatrix} b & -a & 0 \end{bmatrix}^T$.

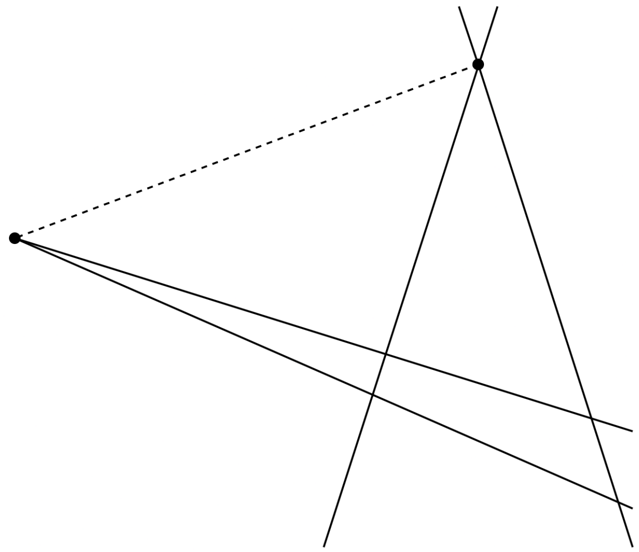
If we apply a projective transformation to all the aforementioned parallel lines $\mathbf{l}_i$, we obtain the transformed lines $\mathbf{l}'_i$. The common point $\mathbf{x}_\infty$, shared by all lines $\mathbf{l}_i$, is mapped to a point $\mathbf{x}'_\infty$ which belongs to each of the lines $\mathbf{l}'_i$.



Therefore, we can assert that all lines $\mathbf{l}'_i$ intersect at the point $\mathbf{x}'_\infty = \mathbf{H}\mathbf{x}_\infty$, referred to as the vanishing point associated with the direction $(b, -a)$ of the parallel lines.

Theorem 2.6.2. *The image of a set of parallel lines $\mathbf{l}_i$ is a set of lines $\mathbf{l}'_i$ concurrent at a common point $\mathbf{x}'$ known as the vanishing point of the direction of lines $\mathbf{l}_i$.*

By applying a projective transformation to the line at infinity $\mathbf{l}_\infty$, we obtain a line $\mathbf{l}'_\infty$. This line intersects the image all the points at the infinity $\mathbf{x}_\infty$ from the original plane. Consequently, the vanishing line $\mathbf{l}'_\infty$ can be determined from two vanishing points.



2.6.2 Polarity

Polarity remains unaltered in the presence of projective mappings. The polar line $\mathbf{l} = \mathbf{C}\mathbf{x}$ corresponding to a point $\mathbf{x}$ with respect to a conic $\mathbf{C}$ gets mapped to the polar line of the transformed point $\mathbf{x}' = \mathbf{H}\mathbf{x}$ with respect to the transformed conic:

$$\mathbf{C}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}$$

Proof. This property holds because:

$$\mathbf{C}'\mathbf{x}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}\mathbf{H}\mathbf{x} = \mathbf{H}^{-T}\mathbf{C}\mathbf{x} = \mathbf{H}^{-T}\mathbf{l} = \mathbf{l}'$$

Therefore, the polar line of the transformed point aligns with the polar line of the original point. □

In conclusion, as polarity remains intact under projective mappings, conjugacy is similarly preserved, and the relationship $\mathrm{CR} = -1$ is also upheld.

2.6.3 Cross ratio

Given a line defined by four points with the following relationships:

$$\mathbf{x}_1 = \propto_1 \mathbf{y} + \beta_1 \mathbf{z}$$

$$\mathbf{x}_2 = \propto_2 \mathbf{y} + \beta_2 \mathbf{z}$$

The cross ratio is expressed as:

$$\mathrm{CR}_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{z}} = \frac{\frac{\beta_1}{\alpha_1}}{\frac{\beta_2}{\alpha_2}}$$

Upon applying a projective transformation $\mathbf{H}$ to these four points:

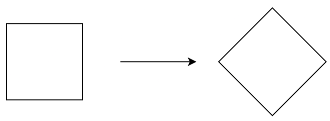
$$\begin{cases} \mathbf{y}' = \mathbf{H}\mathbf{y} \\ \mathbf{z}' = \mathbf{H}\mathbf{z} \\ \mathbf{x}_1' = \mathbf{H}\mathbf{x}_1 \propto_1 \mathbf{y}' + \beta_1 \mathbf{z}' \\ \mathbf{x}_2' = \mathbf{H}\mathbf{x}_2 \propto_2 \mathbf{y}' + \beta_2 \mathbf{z}' \end{cases}$$

The coefficients of the linear combination remain the same. Hence, the cross ratio is conserved, maintaining its original value:

$$\mathrm{CR}_{\mathbf{x}_1',\mathbf{x}_2',\mathbf{y}',\mathbf{z}'} = \frac{\frac{\beta_1}{\alpha_1}}{\frac{\beta_2}{\alpha_2}} = \mathrm{CR}_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{z}}$$

2.6.4 Isometries

Isometries possess three degrees of freedom, which include translation denoted as t and the rotation angle represented by ϑ. Consequently, the invariants of this transformation encompass lengths, distances, and areas.



Definition. The *orthogonal matrix* $\mathbf{R}_\perp$ is defined as follows:

$$\mathbf{R}_\perp^{-1} = \mathbf{R}_\perp^T$$

Hence, the matrix $\mathbf{H}_I$ for isometries takes the following form:

$$\mathbf{H}_I = \begin{bmatrix} \cos\vartheta & -\sin\vartheta & t_x \\ \sin\vartheta & \cos\vartheta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Here, $\begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix} = \mathbf{R}_\perp$

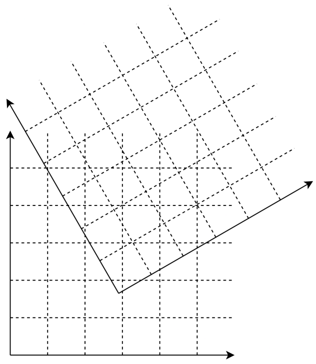
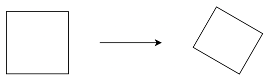


Figure 2.4: Isometry

2.6.5 Similarities

Similarities are characterized by four degrees of freedom, encompassing the translation, denoted as t; the scale, represented by s; and the rotation angle, expressed as ϑ. Consequently, the invariants of this transformation encompass the ratio of lengths and angles. Furthermore, the circular points $\mathbf{I}$ and $\mathbf{J}$ remain invariant throughout this transformation.



Hence, the matrix $\mathbf{H}_S$ for similarities is as follows:

$$\mathbf{H}_I = \begin{bmatrix} s\cos\vartheta & -s\sin\vartheta & t_x \\ s\sin\vartheta & s\cos\vartheta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Here, $\begin{bmatrix} s\cos\vartheta & -s\sin\vartheta \\ s\sin\vartheta & s\cos\vartheta \end{bmatrix} = s\mathbf{R}_\perp$

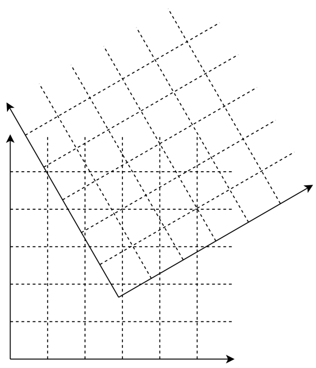
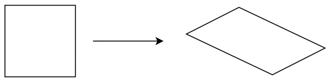


Figure 2.5: Similarity

2.6.6 Affinities

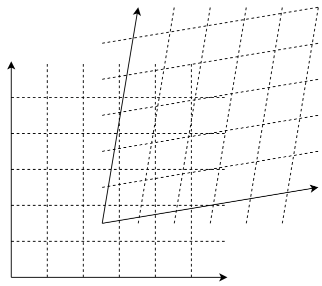
Affinities exhibit six degrees of freedom, consisting of the sub-matrix $\mathbf{A}$ and the translation component. As a result, the invariants of this transformation encompass parallelism, the ratio of parallel lengths, and the ratio of areas. The matrix $\mathbf{A}$ is defined as a 2×2 matrix with a rank of two. Additionally, the line at infinity, denoted as $\mathbf{l}_\infty$, remains invariant throughout the transformation.



Hence, the matrix $\mathbf{H}_A$ for affinities takes the following form:

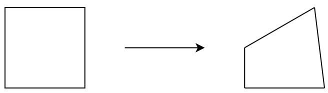
$$\mathbf{H}_I = \begin{bmatrix} a_{11} & a_{21} & t_x \\ a_{12} & a_{21} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Here, $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{21} \end{bmatrix} = \mathbf{A}$



2.6.7 Projectivities

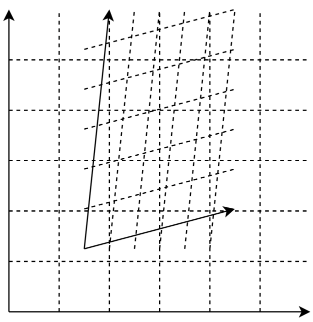
Projectivities possess eight degrees of freedom, encompassing the sub-matrix $\mathbf{A}$, the vector $\mathbf{v}$, and the translation component. Therefore, the invariants of this transformation include co-linearity, incidence, and the order of contact. The matrix $\mathbf{A}$ is defined as a 2×2 matrix with a rank of two. Furthermore, the cross ratio remains invariant throughout this transformation.



Hence, the matrix $\mathbf{H}_P$ for projectivities takes the following form:

$$\mathbf{H}_I = \begin{bmatrix} a_{11} & a_{21} & t_x \\ a_{12} & a_{21} & t_y \\ v_1 & v_2 & 1 \end{bmatrix}$$

Here, $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{21} \end{bmatrix} = \mathbf{A}$



CHAPTER 3

Two-dimensional reconstruction

3.1 Introduction

Recovering a model of an unknown planar scene from a single image, where the image is a projective transformation of the scene, presents a challenging problem. This transformation is represented by the equation $\mathbf{x}'_i = \mathbf{H}\mathbf{x}_i$, where $\mathbf{x}_i$ denotes the scene points and $\mathbf{H}$ is the transformation matrix. The difficulty arises from the fact that while the scene points $\mathbf{x}_i$ are known, the transformation matrix $\mathbf{H}$ is unknown, making a direct inversion of the mapping impossible.

The complexity of this task stems from the large number of unknown variables in the transformation matrix $\mathbf{H}$. In its general form, this problem is underdetermined, meaning that without additional constraints or simplifications, it cannot be uniquely solved. To address this challenge, two primary strategies are typically employed:

1. *Reducing the number of unknowns*: in many practical cases, the goal is not to recover the exact original configuration of the scene but rather to retrieve its overall geometric structure, a process known as shape reconstruction. By focusing on the shape rather than the full projective transformation, the number of unknowns can be reduced from eight to four. In this case, the transformation matrix $\mathbf{H}$ simplifies to:

$$\mathbf{H} = \begin{bmatrix} s\cos\vartheta & -s\sin\vartheta & t_x \\ s\sin\vartheta & s\cos\vartheta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

2. *Adding constraints*: another approach is to incorporate additional information to constrain the reconstruction. This extra information often involves parameters that remain invariant under the specific type of transformation being sought, but vary under more general classes of transformations. By leveraging these invariants, we can narrow down the possible solutions and recover a more accurate model of the scene.

Reconstruction methods can be broadly classified into two categories:

- *Affine reconstruction*: in this approach, the reconstructed scene is related to the original through an affine transformation, preserving parallelism but not necessarily angles or distances.

- *Shape reconstruction*: this method seeks to recover the overall shape of the scene using a similarity transformation, which preserves both angles and relative distances, providing a more faithful representation of the scene's geometry while simplifying the problem.

3.2 Affine reconstruction

Theorem 3.2.1. *A projective transformation* $\mathbf{H}$ *that maps the line at infinity* $\mathbf{l}_\infty$ *onto itself implies that* $\mathbf{H}$ *is affine.*

Proof. A point at infinity, represented as $\mathbf{x}_\infty = \begin{bmatrix} x & y & 0 \end{bmatrix}^T$, is mapped to another point $\mathbf{x}' = \mathbf{H}\mathbf{x}_\infty$ to remain a point at infinity, its third coordinate must be zero. This condition can be expressed as:

$$\begin{bmatrix} v_1 & v_2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = 0$$

Which simplifies to:

$$\begin{bmatrix} v_1 & v_2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Thus, the matrix $\mathbf{H}$ has the structure of an affine transformation, confirming that $\mathbf{H}$ is affine. □

In an image produced by a general projective transformation of a scene, the line at infinity $\mathbf{l}'_\infty$ in the image will not coincide with the original line at infinity $\mathbf{l}_\infty$. However, this difference can be exploited by using $\mathbf{l}'_\infty$ as additional information. By applying a corrective projective transformation $\mathbf{H}_{\mathrm{AR}}$ that maps $\mathbf{l}'_\infty$ back to $\mathbf{l}_\infty$, a modified image is obtained. In this new image, the line at infinity $\mathbf{l}_\infty$ is preserved.

According to the theorem, this resulting transformation produces a model that is an affine mapping of the original scene. Therefore, the modified image is an affine reconstruction of the scene.

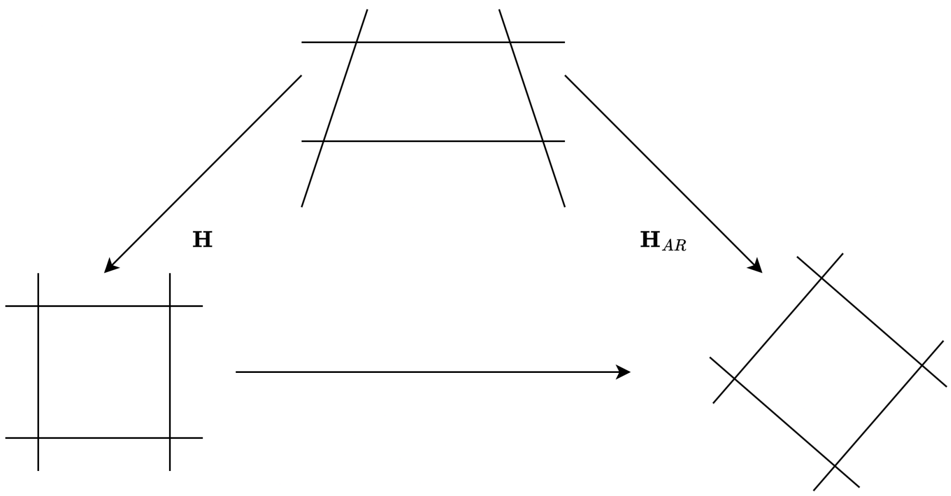


Figure 3.1: Affine transformation

The key challenges in this approach are:

- Determining the projective transformation $\mathbf{H}_{\mathrm{AR}}$ that maps $\mathbf{l}'_\infty$ back to $\mathbf{l}_\infty$.
- Identifying the vanishing line in the image, which corresponds to $\mathbf{l}'_\infty$.

3.2.1 Projective transformation determination

o find the corrective projective transformation $\mathbf{H}_{\text{AR}}$ that restores $\mathbf{l}'_\infty$ to $\mathbf{l}_\infty$, the mapping must satisfy the condition that any point $\mathbf{x}' \in \mathbf{l}'_\infty$ is mapped to a point at infinity. Mathematically, this can be written as:

$$\mathbf{H}_{\text{AR}}\mathbf{x}' = \begin{bmatrix} * \\ * \\ 0 \end{bmatrix}$$

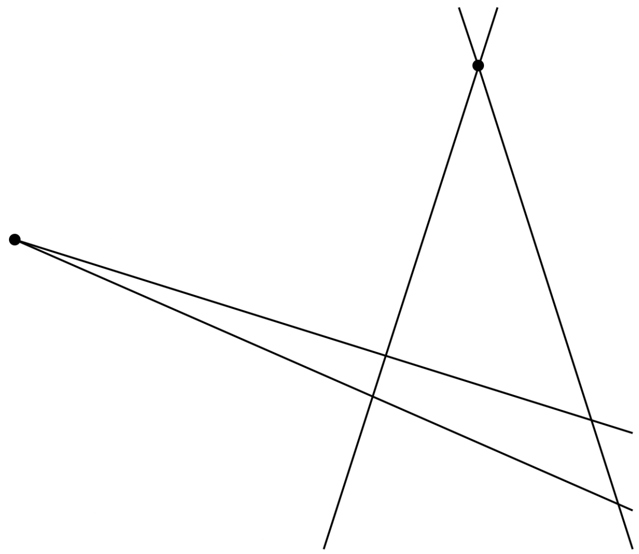
The transformation $\mathbf{H}_{\text{AR}}$ can be represented in matrix form as:

$$\mathbf{H}_{\text{AR}} = \begin{bmatrix} * & * & * \\ * & * & * \\ & \mathbf{l}'^{T}_\infty & \end{bmatrix}$$

which ensures that any point $\mathbf{x}' \in \mathbf{l}'_\infty$ is correctly mapped to the line at infinity.

3.2.2 Vanishing line identification

To identify the vanishing line $\mathbf{l}'_\infty$, additional geometric information can be used, such as the images of parallel lines in the scene. These parallel lines intersect at points on the vanishing line in the projective image.



By leveraging such information, the vanishing line can be accurately determined, enabling the projective transformation $\mathbf{H}_{\text{AR}}$ to be applied for an affine reconstruction of the scene.

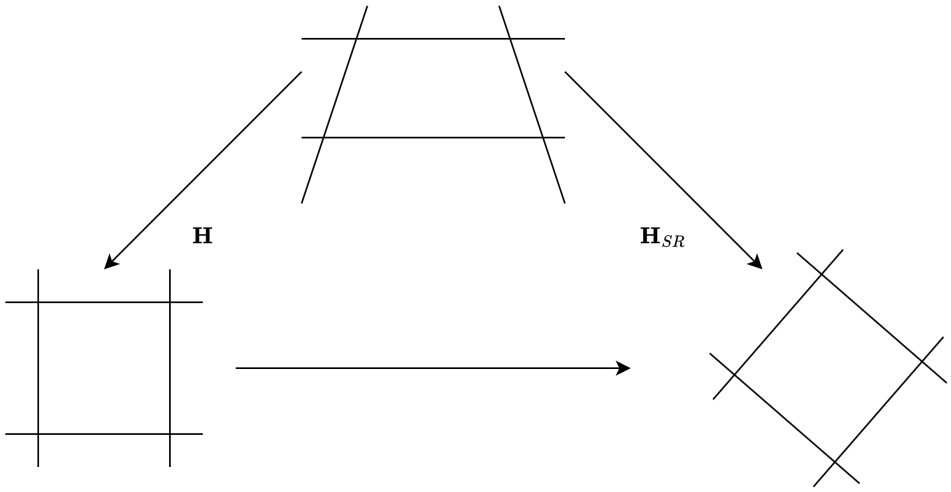
3.3 Shape reconstruction

Theorem 3.3.1. *A projective transformation* $\mathbf{H}$ *that maps the circular points* $\mathbf{I}$ *and* $\mathbf{J}$ *onto themselves implies that* $\mathbf{H}$ *is a similarity transformation.*

Proof. When a similarity transformation matrix $\mathbf{H}_{\text{S}}$ is applied to the circular point $\mathbf{I}$, it produces a scalar multiple of $\mathbf{I}$. The same holds true for the other circular point, $\mathbf{J}$. Since both $\mathbf{I}$ and $\mathbf{J}$ remain unchanged under this transformation, $\mathbf{H}_{\text{S}}$ is indeed a similarity transformation. □

In a general projective mapping of the original scene, the images of the circular points, denoted as $(\mathbf{I}', \mathbf{J}')$, do not coincide with the original circular points $\mathbf{I}$ and $\mathbf{J}$. To perform shape reconstruction, we apply a corrective projective transformation $\mathbf{H}_{\mathrm{SR}}$ that maps $\mathbf{I}'$ and $\mathbf{J}'$ back to $\mathbf{I}$ and $\mathbf{J}$, respectively. This results in a modified image where the circular points are restored to their original positions.

According to the theorem, this new transformation results in a similarity transformation of the original scene. Hence, the reconstructed model is a shape reconstruction, maintaining the overall proportions and geometry of the original scene.



The main challenges in this approach are:

- Finding the projective transformation $\mathbf{H}_{\mathrm{SR}}$ that restores the points $\mathbf{I}'$ and $\mathbf{J}'$ to $\mathbf{I}$ and $\mathbf{J}$.
- Determining the vanishing line in the image to aid in finding the circular points.

3.3.1 Projective transformation determination

Finding the projective transformation $\mathbf{H}_{\mathrm{SR}}$ that maps $\mathbf{I}'$ and $\mathbf{J}'$ back to $\mathbf{I}$ and $\mathbf{J}$ is equivalent to solving for one of the infinitely many matrices that satisfy:

$$\begin{cases} \mathbf{H}_{\mathrm{SR}}\mathbf{I}' = \mathbf{I} \\ \mathbf{H}_{\mathrm{SR}}\mathbf{J}' = \mathbf{J} \end{cases}$$

This task is non-trivial, but it can be simplified by leveraging additional geometric information, such as the degenerate conic dual to the circular points.

The degenerate conic dual to $\mathbf{I}', \mathbf{J}'$ is given by:

$$\mathbf{C}'^{*}_{\infty} = \mathbf{I}'\mathbf{J}'^{T} + \mathbf{J}'\mathbf{I}'^{T}$$

This conic is the image of the original conic dual to the circular points $\mathbf{I}$ and $\mathbf{IJ}$, denoted by:

$$\mathbf{C}^{*}_{\infty} = \mathbf{I}\mathbf{J}^{T} + \mathbf{J}\mathbf{I}^{T}$$

Since $\mathbf{C}'^{*}_{\infty}$ is the projective image of $\mathbf{C}^{*}_{\infty}$, any projective transformation $\mathbf{H}_{\mathrm{SR}}$ that restores $\mathbf{I}'$ and $\mathbf{J}'$ to $\mathbf{I}$ and $\mathbf{J}$ will also restore $\mathbf{C}'^{*}_{\infty}$ to $\mathbf{C}^{*}_{\infty}$. Using the transformation rule for dual conics, we have:

$$\mathbf{C}^{*}_{\infty} = \mathbf{H}_{\mathrm{SR}}\mathbf{C}'^{*}_{\infty}\mathbf{H}_{\mathrm{SR}}^{T}$$

Reversing this relationship gives:

$$\mathbf{C}'^{*}_{\infty} = \mathbf{H}_{\mathrm{SR}}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{H}_{\mathrm{SR}}^{-T}$$

By applying singular value decomposition (SVD) to the equation above, we find that $\mathbf{H}_{\mathrm{SR}}^{-1}$ and $\mathbf{H}_{\mathrm{SR}}^{-T}$ are orthogonal matrices. This leads to:

$$\mathrm{SVD}\left(\mathbf{C}'^{*}_{\infty}\right) = \mathbf{U}_{\perp} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{U}_{\perp}^{T}$$

Thus, the solution for $\mathbf{H}_{\mathrm{SR}}$ is:

$$\mathbf{H}_{\mathrm{SR}} = \mathbf{U}_{\perp}^{-1} = \mathbf{U}_{\perp}^{T}$$

To ensure proper image rectification and scaling, the matrix $\mathbf{H}_{\mathrm{SR}}$ can be adjusted to:

$$\mathbf{H}_{\mathrm{SR}} = \begin{bmatrix} \frac{1}{\sqrt{a}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{b}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{U}^{T}$$

This transformation not only maps the circular points back to their original positions but also ensures that the final image is a faithful similarity reconstruction of the scene.

3.3.2 Vanishing line identification

To determine the vanishing line, one can leverage additional information from the observed scene. This information helps establish several key constraints:

1. *Known angles between lines*: when the angles between lines in the scene are known, they provide useful constraints on the vanishing line. The angle between two lines depends on the angle between their normal vectors and is independent of parameters c_1 and c_2. This relationship can be expressed mathematically as:

   $$\cos\vartheta = \frac{a_1 a_2 + b_1 b_2}{\sqrt{(a_1^2 + b_1^2)(a_2^2 + b_2^2)}}$$

   Here, a_1, b_1, a_2, and b_2 are coefficients of the normal vectors of the lines. This expression can be rewritten as:

   $$\cos\vartheta = \frac{\mathbf{l}^{T}\mathbf{C}^{*}_{\infty}\mathbf{m}}{\sqrt{(\mathbf{l}^{T}\mathbf{C}^{*}_{\infty}\mathbf{l})(\mathbf{m}^{T}\mathbf{C}^{*}_{\infty}\mathbf{m})}}$$

   The equation can be further simplified using the relationship $\mathbf{C}^{*}_{\infty} = \mathbf{H}^{-1}\mathbf{C}^{*\prime}_{\infty}\mathbf{H}^{-T}$, allowing us to express $\mathbf{l}^{T}\mathbf{C}^{*}_{\infty}\mathbf{m}$ as $\mathbf{l}'^{T}\mathbf{C}^{*\prime}_{\infty}\mathbf{m}'$. Thus, the equation becomes:

   $$\cos\vartheta = \frac{\mathbf{l}'^{T}\mathbf{C}^{*\prime}_{\infty}\mathbf{m}'}{(\mathbf{l}'^{T}\mathbf{C}^{*\prime}_{\infty}\mathbf{l}')(\mathbf{m}'^{T}\mathbf{C}^{*\prime}_{\infty}\mathbf{m}')}$$

   In this case, $\mathbf{m}'$ and $\mathbf{l}'$ are derived from the image. Since the angle ϑ is known, this equation provides a linear constraint on $\mathbf{C}^{*\prime}_{\infty}$. When the lines are perpendicular ($\cos\vartheta = 0$), this constraint is particularly useful. The unknown matrix $\mathbf{C}'_{\infty}$ is symmetric, homogeneous, and singular, providing four independent constraints.

2. *Known objects shapes*: f the shape of objects in the scene is known, the reconstruction matrix $\mathbf{H}_{\text{SR}}$ can be determined. This transformation matrix is defined as:

$$\mathbf{H}_{\text{SR}} = \begin{bmatrix} \frac{1}{\sqrt{a}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{b}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{U}^T$$

   The Euclidean-reconstructed image can then be calculated as $\mathbf{M}_S = \mathbf{H}_{\text{SR}} \cdot \text{image}$.

3. *Combinations of constraints*: it is also possible to combine known angles between lines and object shapes to derive additional constraints on the vanishing line.

4. *Observation of rigid planar motion*: when observing rigid planar motion, which follows a similarity transformation, circular points remain invariant. The object in motion has three degrees of freedom, and both the center of rotation and the rotation angle can be derived. Given a transformation matrix $\mathbf{H}$, its eigenvectors provide important insights:

   - Eigenvectors of $\mathbf{H}$ correspond to fixed points.
   - Eigenvectors of $\mathbf{H}^{-1}$ correspond to fixed lines.

   Specifically:

   - Complex eigenvalues correspond to eigenvectors $\mathbf{I}'$ and $\mathbf{J}'$, where the phase represents the rotation angle.
   - Real eigenvalues correspond to the eigenvector $\mathbf{O}'$, representing the center of rotation.

   The eigenvectors associated with the complex eigenvalues $e^{i\theta}$, and $-e^{i\theta}$ represent the images of the circular points. The eigenvector corresponding to the eigenvalue 1 gives the image of the center of rotation, O, while θ corresponds to the rotation angle. Using the relationship $\mathbf{C}'^*_\infty = \mathbf{I}'\mathbf{J}'^T + \mathbf{J}'\mathbf{I}'^T$, the singular value decomposition (SVD) can be applied to obtain $\text{SVD}(\mathbf{C}'^*_\infty) = \mathbf{U}\mathbf{C}^*_\infty\mathbf{U}^T$, where $\mathbf{U}^T$ is the rectification matrix. There are two primary methods to address this:

   - *Direct method*: first compute $\mathbf{C}^{*\prime}_\infty$, and then derive the rectification matrix $\mathbf{H}_{\text{rect}}$.
   - *Stratified method*: first perform an affine reconstruction from projective to affine geometry, and then perform shape reconstruction from affine to metric geometry.

   The stratified method can sometimes reduce numerical errors, providing more accurate results.

3.4 Accuracy issues

Achieving accurate image rectification is influenced by several factors. Here, we identify key challenges and strategies for improving accuracy:

1. *Noise and numerical errors*: noise and numerical errors in the input data can significantly impact the rectification process, leading to inaccuracies. Preprocessing techniques, such as noise filtering and data normalization, are essential to mitigate these issues and improve stability in the rectification calculations.

2. *Insufficient line information*: selecting appropriate lines to determine vanishing points is critical. Lines that are too close to each other may result in poorly defined vanishing points, introducing inaccuracies into the rectification. It is important to select lines that are sufficiently spaced to ensure precise vanishing point estimation.

3. *Vanishing points near infinity*: when a vanishing point is nearly at infinity, achieving accurate affine rectification becomes challenging. To address this issue, consider the following approaches:

   - Draw two lines in the scene perpendicular to the original lines in the image. If these lines are not parallel, adjust their intersections to make them parallel. This adjustment helps bring the vanishing point closer, making it easier to compute an affine transformation.

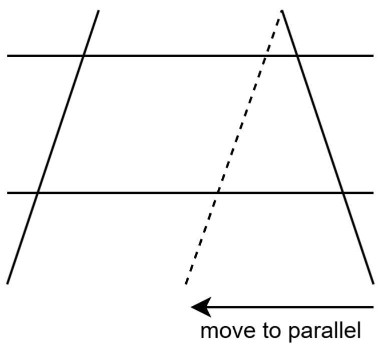


Figure 3.2: Aligning lines to bring the vanishing point closer

After adjusting, apply affine reconstruction to obtain more accurate results.

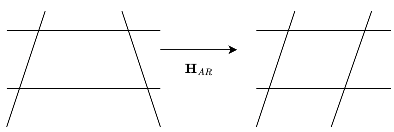


Figure 3.3: Applying affine reconstruction

   - For scenes with multiple sets of parallel lines, select one line from each set and randomly choose an additional pair of perpendicular lines. Using these four lines, compute the matrix product $\mathbf{K}\mathbf{K}^T$, where $\mathbf{K}$ represents the camera calibration matrix. Derive $\mathbf{K}$ through Cholesky factorization and use it in the rectifying transformation:

$$\mathbf{H}_{\text{rect}} = \begin{bmatrix} \mathbf{K} & \mathbf{t} \\ 0 & 1 \end{bmatrix}^{-1}$$

where $\mathbf{t}$ is a translation vector. This transformation corrects perspective distortions.

The accuracy of image rectification is crucial for various computer vision and image processing applications, and addressing these issues is essential for obtaining reliable results.

CHAPTER 4

Three-dimensional space projective geometry

4.1 Introduction

In spatial geometry, fundamental elements such as points, planes, quadrics, and dual quadrics serve as the building blocks for defining geometric relationships and structures. These elements are manipulated through various types of transformations, each preserving different geometric properties. The principal transformations include: projectivities, affinities, similarities, and isometries.

4.2 Points

In spatial geometry, points are typically represented in a Euclidean coordinate system by defining a reference origin and specifying each point's location with three Cartesian coordinates (x, y, z). This system provides an unambiguous means of defining each point's position within three-dimensional space.

For image analysis, however, it is often more convenient to use homogeneous coordinates. Homogeneous coordinates introduce an additional coordinate, w, allowing points to be represented as follows:

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

In this form, any non-zero scalar multiple of $\mathbf{X}$, denoted as $\lambda\mathbf{X}$ for $\lambda \neq 0$, represents the same point. This property, known as homogeneity, is foundational in projective geometry as it enables the representation of points at infinity, which cannot be represented in a traditional Cartesian system. Notably, the null vector does not correspond to any point.

Definition (*Projective space*)**.** The projective space $\mathbb{P}^3$ is defined as:

$$\mathbb{P}^3 = \left\{ \begin{bmatrix} x & y & z & w \end{bmatrix}^T \in \mathbb{R}^4 \right\} - \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T \right\}$$

Property 4.2.1. A point $\mathbf{X}$, formed by a linear combination $\mathbf{X} = \alpha\mathbf{X}_1 + \beta\mathbf{X}_2$ of two points $\mathbf{X}_1$ and $\mathbf{X}_2$, lies on the line $\mathbf{l}$ that passes through both $\mathbf{X}_1$ and $\mathbf{X}_2$.

4.3 Planes

In homogeneous coordinates, planes in space can be represented by the vector:

$$\boldsymbol{\pi} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Here, (a, b, c) defines the direction normal to the plane. The perpendicular distance from the origin to the plane is given by:

$$\text{distance} = -\frac{d}{\sqrt{a^2 + b^2 + c^2}}$$

Similar to points, plane coordinates in homogeneous form exhibit the property of homogeneity. Thus, any non-zero scalar multiple $\lambda\boldsymbol{\pi}$, where $\lambda \neq 0$, represents the same plane. This makes a, b, c, and d the homogeneous parameters of the plane. As with points, a single plane has infinitely many equivalent representations, all scaled versions of its normal vector. The null vector, however, does not represent any plane, and if $d = 0$, the plane passes through the origin.

To determine whether a point $\mathbf{X}$ lies on a plane $\boldsymbol{\pi}$ or if a plane passes through a point, the following equation must be satisfied:

$$\begin{cases} ax + by + cz + dw = 0 \\ \boldsymbol{\pi}^T\mathbf{X} = \mathbf{X}^T\boldsymbol{\pi} = 0 \end{cases}$$

Definition (*Plane at the infinity*)**.** The plane at infinity, $\boldsymbol{\pi}_\infty$, is defined as:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = w = 0$$

This plane has an undefined normal direction and represents points at infinity.

Theorem 4.3.1 (Duality principle)**.** *For any true statement involving terms like point, plane, is on, and goes through, there exists a corresponding dual statement that is also true, derived by making the following substitutions:*

- *Point* $\Leftrightarrow$ *plane.*
- *Is on* $\Leftrightarrow$ *goes through.*

Point as intersection of planes A point can be described as the intersection of three distinct planes:

$$\begin{cases} \boldsymbol{\pi}_1^T\mathbf{X} = 0 \\ \boldsymbol{\pi}_2^T\mathbf{X} = 0 \\ \boldsymbol{\pi}_3^T\mathbf{X} = 0 \end{cases} \implies \begin{bmatrix} \boldsymbol{\pi}_1^T \\ \boldsymbol{\pi}_2^T \\ \boldsymbol{\pi}_3^T \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This setup finds the Right Null Space of the matrix of planes:

$$\mathbf{X} = \text{RNS}\left(\begin{bmatrix}\boldsymbol{\pi}_1^T \\ \boldsymbol{\pi}_2^T \\ \boldsymbol{\pi}_3^T\end{bmatrix}\right)$$

Yielding a solution vector representing the intersection point, along with all of its scalar multiples.

Plane trough tree points A plane passing through three points $\mathbf{X}_1^T$, $\mathbf{X}_2^T$, and $\mathbf{X}_3^T$ can be determined by finding the Right Null Space of the matrix of points:

$$\mathbf{X} = \text{RNS}\left(\begin{bmatrix}\mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \mathbf{X}_3^T\end{bmatrix}\right)$$

Property 4.3.1. The plane $\boldsymbol{\pi}$, defined by the linear combination $\boldsymbol{\pi} = \alpha\boldsymbol{\pi}_1 + \beta\boldsymbol{\pi}_2$ of two planes $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$, passes through the line $\mathbf{l}^*$, which lies on both $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$.

4.4 Lines

In space geometry, lines act as intermediate entities between points and planes. While they are considered primitive in planar geometry, in three-dimensional space they are defined in terms of points and planes. Lines are self-dual elements, meaning that statements about lines retain their validity when duality is applied, reflecting the symmetry in their relationships with points and planes.

Lines can be defined in various ways:

- *Intersection of two planes*: a line can be obtained as the intersection of two distinct planes, $\boldsymbol{\pi}_1^T$ and $\boldsymbol{\pi}_2^T$. This yields:

  $$\mathbf{X} = \text{RNS}\left(\begin{bmatrix}\boldsymbol{\pi}_1^T \\ \boldsymbol{\pi}_2^T\end{bmatrix}\right)$$

  Here, $\mathbf{X}$ represents a 2D solution set of points within both planes, reduced to a 1D set of points due to homogeneity, parameterized by a single variable along the line.

- *Passing throug two points*: a line can also be defined by two distinct points, $\mathbf{X}_1$ and $\mathbf{X}_2$. This produces:

  $$\mathbf{X} = \text{RNS}\left(\begin{bmatrix}\mathbf{X}_1^T \\ \mathbf{X}_2^T\end{bmatrix}\right)$$

  This results in a 2D set of solution vectors (representing all planes containing the line), which reduces to a 1D set due to homogeneity, where each plane in the set is determined by a rotation angle.

- *Linear combination of two points*: A line $\mathbf{l}$ can also be represented as the set of all points that are linear combinations of two given points $\mathbf{X}_1$ and $\mathbf{X}_2$:

  $$\mathbf{X} = \alpha\mathbf{X}_1 + \beta\mathbf{X}_2$$

  Here, α and β are scalars. This defines the line in terms of its point-based representation.

- *Linear combination of two planes*: alternatively, a line $\mathbf{l}^*$ can be defined as the set of all planes that are linear combinations of two given planes $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$:

$$\boldsymbol{\pi} = \alpha\boldsymbol{\pi}_1 + \beta\boldsymbol{\pi}_2$$

  Here, α and β are scalars. This representation defines the line by the set of planes containing it.

Theorem 4.4.1 (Duality Principle)**.** *For any true statement involving the terms point, line, plane, is on, and goes through, there exists a corresponding dual statement that is also true, derived by making the following substitutions:*

- *Point* $\Leftrightarrow$ *plane.*
- *Is on* $\Leftrightarrow$ *goes through.*
- *Line* $\Leftrightarrow$ *line.*

Each plane $\boldsymbol{\pi}$ has its own line at the infinity $\mathbf{l}_\infty(\boldsymbol{\pi})$ and also its own circular points $\mathbf{I}\boldsymbol{\pi}$ and $\mathbf{J}\boldsymbol{\pi}$ parallel planes share the same $\mathbf{l}_\infty$ and the same circular points $\mathbf{I}$ and $\mathbf{J}$.

4.5 Quadrics

Definition (*Quadric*)**.** A point $\mathbf{X}$ lies on a quadric $\mathbf{Q}$ if it satisfies a homogeneous quadratic equation, given by:

$$\mathbf{X}^T\mathbf{Q}\mathbf{X} = \mathbf{0}$$

Here, $\mathbf{Q}$ is a 4×4 symmetric matrix.

The matrix $\mathbf{Q}$ is homogeneous, meaning that for any scalar λ, $\lambda\mathbf{Q}$ represents the same quadric as $\mathbf{Q}$ This implies that the matrix $\mathbf{Q}$ has 9 degrees of freedom. In general, 9 points in a general position are sufficient to define a unique quadric surface.

Intersection with a plane A quadric can intersect with a plane in various ways, depending on the geometric properties of both. One important concept in projective geometry is the absolute conic.

Proposition. The absolute conic $\boldsymbol{\Omega}_\infty$ contains the circular points $\mathbf{I}_\infty, \mathbf{J}_\infty$ of any plane $\boldsymbol{\pi}$.

These circular points are a fundamental extension in the context of projective geometry, connecting the concept of a plane's geometry with the quadric surfaces.

4.5.1 Degenrate quadrics

When the matrix $\mathbf{Q}$ is degenrate, the quadric takes on special forms. Specifically, the rank of $\mathbf{Q}$ reveals the type of degeneration:

1. If $\text{rank}(\mathbf{Q}) = 1$, the quadric corresponds to a repeated plane, and $\mathbf{Q}$ can be factored as $\mathbf{Q} = \mathbf{A}\mathbf{A}^T$, where $\mathbf{A}$ is a vector defining the plane.
2. If $\text{rank}(\mathbf{Q}) = 2$, the quadric is the intersection of two distinct planes, and $\mathbf{Q}$ can be written as $\mathbf{Q} = \mathbf{A}\mathbf{B}^T + \mathbf{B}\mathbf{A}^T$, where $\mathbf{A}$ and $\mathbf{B}$ define the two planes.

3. If $\text{rank}(\mathbf{Q}) = 3$, the quadric is a cone. The vertex of the cone can be found as the right null space (RNS) of $\mathbf{Q}$.

A special case of a cone occurs when the vertex is at infinity, leading to the geometric object known as a cylinder.

4.6 Transformations

Definition (*Projective Mapping*)**.** A projective mapping between a projective space $\mathbb{P}^3$ and another projective space $\mathbb{P}'^3$ is an invertible mapping that preserves collinearity of points.

Theorem 4.6.1. *A mapping $h : \mathbb{P}^3 \to \mathbb{P}'^3$ is projective if and only if there exists an invertible 4×4 matrix $\mathbf{H}$ such that for any point $\mathbf{X}$ in $\mathbb{P}^3$, we have:*

$$h(\mathbf{X}) = \mathbf{H}\mathbf{X}$$

Projective mappings are linear in homogeneous coordinates, but not linear in Cartesian coordinates. From the theorem, we have the relation:

$$h(\mathbf{X}) = \mathbf{X}' = \mathbf{H}\mathbf{X}$$

Thus, if we multiply the matrix $\mathbf{H}$ by any nonzero scalar λ, the relation still holds for the same points:

$$\mathbf{X}' = \lambda\mathbf{H}\mathbf{X}$$

Therefore, any nonzero scalar multiple of $\mathbf{H}$ represents the same projective mapping as $\mathbf{H}$. This implies that $\mathbf{H}$ is a homogeneous matrix: despite having 16 entries, $\mathbf{H}$ only has 15 degrees of freedom, corresponding to the ratios between its elements.

A homography transforms each point $\mathbf{X}$ into a point $\mathbf{X}'$ such that:

$$\mathbf{X}' = \mathbf{H}\mathbf{X}$$

Additionally, a homography transforms each plane $\boldsymbol{\pi}$ into a plane $\boldsymbol{\pi}'$ such that:

$$\boldsymbol{\pi}' = \mathbf{H}^{-T}\boldsymbol{\pi}$$

A homography transforms each quadric $\mathbf{Q}$ into a quadric $\mathbf{Q}'$ such that:

$$\mathbf{Q}' = \mathbf{H}^{-T}\mathbf{Q}\mathbf{H}^{-1}$$

Similarly, a homography transforms each dual quadric $\mathbf{Q}^*$ into a dual quadric $\mathbf{Q}^{*'}$ such that:

$$\mathbf{Q}^{*'} = \mathbf{H}\mathbf{Q}^*\mathbf{H}^T$$

Cross-ratios are invariant under projective mappings.

4.6.1 Isometries

An isometry is a transformation that preserves distances. The matrix representation of an isometry $\mathbf{H}_I$ is given by:

$$\mathbf{H}_I = \begin{bmatrix} \mathbf{R}_\perp & & \mathbf{t} & \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\mathbf{R}_\perp$ is a 3×3 orthogonal matrix, meaning that:

$$\mathbf{R}_\perp^{-1} = \mathbf{R}_\perp^T \quad \text{and} \quad \det(\mathbf{R}_\perp) = \pm 1$$

The sign of the determinant indicates whether the transformation is a rigid displacement ($\det(\mathbf{R}_\perp) = 1$) or includes a reflection ($\det(\mathbf{R}_\perp) = -1$).

Isometries have 6 degrees of freedom: 3 for translation ($\mathbf{t}$) and 3 for rotation (Euler angles θ, ϕ, ψ).

Invariants of isometries include lengths, distances, and areas, which means that the shape and size of objects are preserved, as well as their relative positions.

4.6.2 Similarities

A similarity transformation includes both rigid motion and scaling. The matrix representation of a similarity $\mathbf{H}_S$ is:

$$\mathbf{H}_S = \begin{bmatrix} s\mathbf{R}_\perp & & \mathbf{t} & \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where s is a scalar that represents the scale factor and $\mathbf{R}_\perp$ is a 3×3 orthogonal matrix. The matrix $\mathbf{R}_\perp$ satisfies:

$$\mathbf{R}_\perp^{-1} = \mathbf{R}_\perp^T \quad \text{and} \quad \det(\mathbf{R}_\perp) = \pm 1$$

Similarities have 7 degrees of freedom: 3 for rigid displacement (translation and rotation) and 1 for scaling.

Invariants of similarities include the ratio of lengths and angles, which preserve the shape (but not the size) of objects. Additionally, the absolute conic Ω_∞ and the absolute dual quadric $\mathbf{Q}^*_\infty$ remain invariant under similarity transformations.

4.6.3 Affinities

An affinity is a transformation that preserves parallelism and the ratio of distances along parallel lines. The matrix representation of an affinity $\mathbf{H}_A$ is:

$$\mathbf{H}_A = \begin{bmatrix} \mathbf{A} & & \mathbf{t} & \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\mathbf{A}$ is any invertible 3×3 matrix, and $\mathbf{t}$ is a translation vector. Affinities have 12 degrees of freedom: 9 for the matrix $\mathbf{A}$ and 3 for the translation vector $\mathbf{t}$.

Invariants of affinities include parallelism, the ratio of lengths along parallel lines, and the ratio of areas. A key invariant is the plane at infinity $\boldsymbol{\pi}_\infty$, which remains unchanged under affine transformations.

4.6.4 Projectivities

A projectivity is a transformation that preserves collinearity and incidence relations. The matrix representation of a projectivity $\mathbf{H}_P$ is:

$$\mathbf{H}_P = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & 1 \end{bmatrix}$$

where $\mathbf{A}$ is any invertible 3×3 matrix, $\mathbf{t}$ is a translation vector, and $\mathbf{v}$ is a vector that represents the relation between the homogeneous coordinates of the points.

Projectivities have 15 degrees of freedom: 9 for the matrix $\mathbf{A}$, 3 for the translation $\mathbf{t}$, and 3 for the vector $\mathbf{v}$.

Invariants of projectivities include collinearity, incidence, and the order of contact (such as crossing, tangency, and inflections). Additionally, the 1D, 2D, and 3D cross ratios are invariant under projective transformations.

CHAPTER 5

Single view geometry

5.1 Camera projective model

A pinhole camera follows a perspective projection model that is inherently nonlinear, neither shape-preserving nor length-ratio preserving. Despite this, it maintains the collinearity property: collinear points in the 3D scene are projected onto collinear points in the 2D image. This implies a linear transformation between homogeneous coordinates, though the mapping is not invertible due to the loss of one dimension when capturing the image.

A 3D point $\mathbf{X}$ in homogeneous coordinates is projected onto an image point $\mathbf{u}$ through the camera projection matrix $\mathbf{P}$:

$$\mathbf{u} = \mathbf{P}\mathbf{X} = \begin{bmatrix}\mathbf{M} & \mathbf{m}\end{bmatrix}\mathbf{X}$$

The back-projection of an image point $\mathbf{u}$ through a camera with projection matrix $\mathbf{P} = \begin{bmatrix}\mathbf{M} & \mathbf{m}\end{bmatrix}$ forms a straight line passing through the camera center $\mathbf{O} = \mathrm{RNS}(\mathbf{P})$, with direction given by: $\mathbf{d} = \mathbf{M}^{-1}\mathbf{u}$.

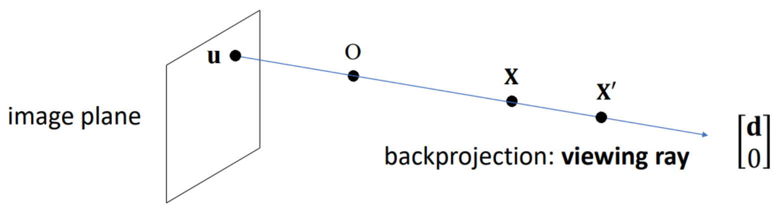


Figure 5.1: Projection from 3D scene to 2D image plane

The optical center $\mathbf{o}$ in world coordinates is: $-\mathbf{M}^{-1}\mathbf{m}$.

5.1.1 Camera intrinsic parameters

The homogeneous coordinates of an image point are represented as:

$$\mathbf{p} = \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$$

Camera reference frame The camera reference frame is centered at the optical center $\mathbf{O}$, with the Z-axis perpendicular to the image plane. The X and Y axes are aligned with pixel rows and columns, respectively, and distances are measured in focal length units f.

Image reference frame The image reference frame is centered at the principal point $\mathbf{C}$. The coordinates (U, V) are parallel to pixel rows and columns but have an orientation opposite to the camera frame's (X, Y). Measurements along the U and V axes are also expressed in units of focal length f.

The homogeneous pixel coordinates of an image point are:

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

To convert between image and pixel coordinates:

- A unit increment in U corresponds to an increase of fX pixels in x.
- A unit increment in V corresponds to an increase of fY pixels in y.

The pixel coordinates of the principal point $\mathbf{C}$ are (U_0, V_0). The relationship between pixel and image coordinates is described by the calibration matrix $\mathbf{K}$:

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & U_0 \\ 0 & f_y & V_0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p} = \mathbf{K} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{X}_{\text{cam}}$$

Here, $\mathbf{K}$ is the intrinsic calibration matrix.

5.1.2 World and camera coordinate systems

In general, the world coordinate system is distinct from the camera coordinate system. The transformation between world and camera coordinates is given by:

$$\mathbf{u} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \mathbf{X}_{\text{world}}$$

Here, $\mathbf{K}$ is the intrinsic calibration matrix, $\mathbf{R}$ is the rotation matrix, $\mathbf{t}$ is the translation vector, and $\mathbf{X}_{\text{world}}$ is the point in world coordinates.

The camera projection matrix $\mathbf{P}$ is then given by:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$$

Thus, the projection of a world point $\mathbf{X}_{\text{world}}$ onto the image plane is:

$$\mathbf{u} = \mathbf{P}\mathbf{X}_{\text{world}} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \mathbf{X}_{\text{world}}$$

This can be rewritten as:

$$\mathbf{u} = \begin{bmatrix} \mathbf{M} & \mathbf{m} \end{bmatrix} \mathbf{X}_{\text{world}}$$

Here, $\mathbf{M} = \mathbf{K}\mathbf{R}$ encapsulates intrinsic and rotational transformations, and $\mathbf{m} = \mathbf{K}\mathbf{t}$ represents intrinsic scaling and translation.

5.2 Calibration matrix

The calibration matrix, also known as the intrinsic matrix, is defined as:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & U_0 \\ 0 & f_y & V_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Here, f_x and f_y are the focal lengths in pixel units along the x- and y-axes, respectively, and (U_0, V_0) represents the principal point.

The intrinsic parameters are properties of the camera itself and remain constant regardless of its position or orientation in space. Key invariant factors include:

- The principal point coordinates U_0 and V_0, which define the relative position of the camera center with respect to the image plane.
- The pixel aspect ratio, defined as $a = \frac{f_x}{f_y}$. For most natural cameras with square pixels, a is equal to 1.

Considering the full projection model, the camera projection matrix is given by:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$$

Here, $\mathbf{R}$ is the rotation matrix, and $\mathbf{t}$ is the translation vector. These extrinsic parameters describe the camera's pose relative to the world coordinate system, determining how the 3D scene is mapped onto the image plane.

5.3 Camera calibration

Camera calibration consists of estimating both intrinsic and extrinsic parameters to accurately model the imaging process.

Intrinsic calibration The intrinsic calibration determines the camera's internal characteristics, specifically the calibration matrix $\mathbf{K}$, which includes focal lengths, the principal point, and, in some cases, the skew factor s (used in older cameras to account for non-square pixels or timing differences in sampling):

$$\mathbf{K} = \begin{bmatrix} f_x & s & U_0 \\ s & f_y & V_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Extrinsic calibration The extrinsic calibration estimates the camera's pose relative to the world coordinate system. This includes:

- The rotation matrix $\mathbf{R}$, which describes the camera's orientation.
- The translation vector $\mathbf{t}$, which defines the camera's position.

Unlike intrinsic parameters, extrinsic parameters change as the camera moves, affecting the transformation between world and image coordinates.

5.3.1 Properties

Vanishing point A vanishing point $\mathbf{v}$ represents the image of a point at infinity along a given direction $\mathbf{d}$:

$$\mathbf{U_v} = \mathbf{Md}$$

Since the back-projection of the image point $\mathbf{v}$ is given by:

$$\mathbf{M}^{-1}\mathbf{u_v} = \mathbf{MM}^{-1}\mathbf{d} = \mathbf{d}$$

We can state the following theorem.

Theorem 5.3.1. *The viewing ray associated with the vanishing point* $\mathbf{v}$ *of a direction* $\mathbf{d}$ *is parallel to* $\mathbf{d}$.

Viewing rays angle Given two image points $\mathbf{u}_1$ and $\mathbf{u}_2$, the angle θ between the directions of the lines connecting the camera center $\mathbf{C}$ to these points can be computed using the calibration matrix $\mathbf{K}$:

$$\cos\theta = \frac{\mathbf{u}_1^T\left(\mathbf{K}^{-T}\mathbf{K}^{-1}\right)\mathbf{u}_2}{\sqrt{\left(\mathbf{u}_1^T\left(\mathbf{K}^{-T}\mathbf{K}^{-1}\right)\mathbf{u}_1\right)\left(\mathbf{u}_2^T\left(\mathbf{K}^{-T}\mathbf{K}^{-1}\right)\mathbf{u}_2\right)}}$$

Since these angles are computed using only the intrinsic parameters, they are invariant to the absolute position of the camera. By defining the image of the absolute conic as:

$$\boldsymbol{\omega} = \left(\mathbf{KK}^T\right)^{-1}$$

We obtain the following property.

Property 5.3.1. The directions $\mathbf{d}_1$ and $\mathbf{d}_2$ of the viewing rays corresponding to two image points $\mathbf{u}_1$ and $\mathbf{u}_1$ form an angle given by:

$$\cos\theta = \frac{\mathbf{u}_1^T\boldsymbol{\omega}\mathbf{u}_2}{\sqrt{\left(\mathbf{u}_1^T\boldsymbol{\omega}\mathbf{u}_1\right)\left(\mathbf{u}_2^T\boldsymbol{\omega}\mathbf{u}_2\right)}}$$

Circular points in space Each plane $\boldsymbol{\pi}$ has an associated pair of circular points, denoted as $\mathbf{I}_\pi$, $\mathbf{J}_\pi$. Notably, parallel planes share the same circular points.

To transform a reference XY plane into a generic plane $\boldsymbol{\pi}$, we apply an isometric transformation $\mathbf{T}_{\boldsymbol{\pi}}$:

$$\mathbf{T}_{\boldsymbol{\pi}} = \begin{bmatrix} & \mathbf{R}_{\boldsymbol{\pi}\mathbf{w}} & & \mathbf{o}_{\boldsymbol{\pi}} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Which results in the following mappings for the circular points:

$$\mathbf{I}_{\boldsymbol{\pi}} = \mathbf{T}_{\boldsymbol{\pi}}\mathbf{I}_{XY} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\pi}\mathbf{w}}\mathbf{I} \\ \mathbf{0} \end{bmatrix} \qquad \mathbf{J}_{\boldsymbol{\pi}} = \mathbf{T}_{\boldsymbol{\pi}}\mathbf{J}_{XY} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\pi}\mathbf{w}}\mathbf{J} \\ \mathbf{0} \end{bmatrix}$$

Plane and image homography Consider the homography $\mathbf{H}$ mapping points from a plane $\boldsymbol{\pi}$ to their corresponding image points.

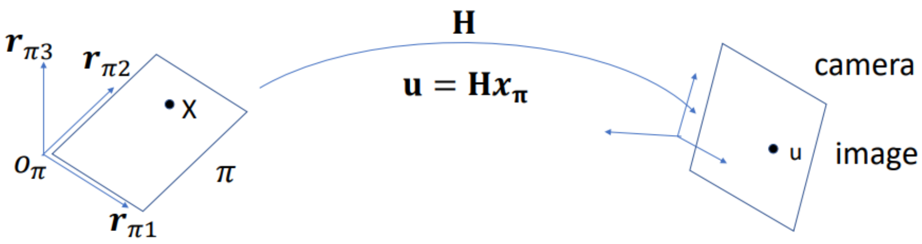


Figure 5.2: Plane homography

The plane $\boldsymbol{\pi}$ serves as a reference for determining its relative pose with respect to the camera, characterized by a rotation $\mathbf{R}_\pi$ and translation $\mathbf{t}_\pi$. The camera reference frame is taken as the world reference frame, with the projection matrix:

$$\mathbf{P} = \begin{bmatrix}\mathbf{K} & \mathbf{0}\end{bmatrix}$$

For a point $\mathbf{X}_\pi$ on the plane, the homography is given by:

$$\mathbf{H} = \mathbf{K}\begin{bmatrix}\mathbf{r}_{\pi 1} & \mathbf{r}_{\pi 2} & \mathbf{o}_\pi\end{bmatrix}$$

Property 5.3.2. For any plane $\boldsymbol{\pi}$, the images of its circular points satisfy:

$$\mathbf{I}'_\pi \boldsymbol{\omega} \mathbf{I}_\pi = \mathbf{0} \qquad \mathbf{J}'_\pi \boldsymbol{\omega} \mathbf{J}_\pi = \mathbf{0}$$

5.3.2 Calibration constraints

To determine the calibration matrix $\mathbf{K}$, we can estimate the image of the absolute conic $\boldsymbol{\omega}$ and apply Cholesky factorization to its inverse.

The constraints on $\boldsymbol{\omega}$ can be derived from various geometric properties, including:

- Known angles between directions, by analyzing their vanishing points.
- Self-orthogonality, by observing the images of circular points.
- Known planar shapes, by analyzing their image projections.

The main constraints that can be used to estimate $\boldsymbol{\omega}$ are:

- Known angles between directions $\mathbf{d}_i$ and their vanishing points. If the directions $\mathbf{d}_i$ and their corresponding vanishing points $\mathbf{v}_i$ are known, where:

  $$\mathbf{v}_i = \mathbf{KRd}_i$$

  Then, the angle constraint between two directions is given by:

  $$\cos\theta = \frac{\mathbf{v}_1^T \boldsymbol{\omega} \mathbf{v}_2}{\sqrt{(\mathbf{v}_1^T \boldsymbol{\omega} \mathbf{v}_1)\,(\mathbf{v}_2^T \boldsymbol{\omega} \mathbf{v}_2)}}$$

- Images of the circular points. The images of the circular points provide two independent equations, one for the real part and one for the imaginary part:

  $$\mathbf{I}'_\pi \boldsymbol{\omega} \mathbf{I}'_\pi = \mathbf{0}$$

- Known planar shapes and the homography $\mathbf{H}$ between image and scene. If a homography $\mathbf{H}$ between the image and the scene is known, then we obtain the following two constraints:

  $$\mathbf{h}_1 \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \qquad \mathbf{h}_1 \boldsymbol{\omega} \mathbf{h}_1 - \mathbf{h}_2 \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0}$$

5.3.3 Zhang method

When performing camera calibration using images of multiple known planar shapes, we can apply Zhang's method to estimate the intrinsic matrix $\mathbf{K}$. For each homography: $\mathbf{H} = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3\end{bmatrix}$ we impose the following constraints on the image of the absolute conic $\boldsymbol{\omega}$:

$$\begin{cases} \mathbf{h}_1 \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{h}_1 \boldsymbol{\omega} \mathbf{h}_1 - \mathbf{h}_2 \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \end{cases}$$

These equations are homogeneous in $\boldsymbol{\omega}$, meaning that at least three homographies are required to fully determine $\boldsymbol{\omega}$. Once $\boldsymbol{\omega}$ is estimated, we apply Cholesky factorization to its inverse to obtain the intrinsic calibration matrix $\mathbf{K}$.

5.3.4 Camera calibration with camera position

When the camera center position $\mathbf{O}$ is known, we can perform extrinsic and intrinsic calibration using a single image of a known planar shape. The calibration process follows these steps:

1. Compute the homography $\mathbf{H}$ between known points on the plane $\boldsymbol{\pi}$ and their corresponding image points.

2. Define the transformation matrix $\mathbf{M_o}$ that relates any point $\mathbf{x}$ on $\boldsymbol{\pi}$ to the direction $\mathbf{d}$ of a ray from the camera center $\mathbf{O}$:

$$\mathbf{M_o}^{-1} = \begin{bmatrix} 1 & 0 & -x_\mathbf{o} \\ 0 & 1 & -y_\mathbf{o} \\ 0 & 0 & -z_\mathbf{o} \end{bmatrix}$$

3. Compute the projection matrix $\mathbf{M}$ relating any image point $\mathbf{u}$ to the direction $\mathbf{d}$ of its viewing ray:

$$\mathbf{d} = \mathbf{M_o}^{-1} \mathbf{H}^{-1} \mathbf{u} = \mathbf{M}^{-1} \mathbf{u}$$

4. Perform a QR decomposition on $\mathbf{M}^{-1}$. The decomposition $\mathbf{M}^{-1} = \mathbf{R}^{-1}\mathbf{K}^{-1}$ allows us to extract: $\mathbf{K}$, the intrinsic calibration matrix, and $\mathbf{R}^{-1}$, the rotation matrix from the world reference (attached to $\boldsymbol{\pi}$) to the camera.

5.3.5 Camera calibration with unknown scene

When the shape of the scene is unknown but we have some images of it, we can perform camera calibration using a method that is generally less precise than the Zhang method. The process is as follows:

1. Capture images of the scene using a camera with a constant intrinsic matrix $\mathbf{K}$.

2. Compute the homographies between different images of the scene.

3. Formulate the system of equations:

$$\begin{cases} \mathbf{I}'^T \boldsymbol{\omega} \mathbf{I}' = \mathbf{0} \\ \mathbf{I}'^T \mathbf{H}'^T \boldsymbol{\omega} \mathbf{H}' \mathbf{I}' = \mathbf{0} \end{cases}$$

4. Solve for $\boldsymbol{\omega}$ and $\mathbf{I}'$.

5. Compute the inverse matrix $\boldsymbol{\omega}^{-1}$ and perform a Cholesky factorization to find $\mathbf{K}$.

5.3.6 Camera calibration with natural camera

Calibration of natural cameras (where $f_X = f_Y = f$) can be performed using the vanishing points of mutually orthogonal directions.

In this case, the image of the absolute conic $\boldsymbol{\omega}$ matrix has the following form:

$$\boldsymbol{\omega} = \begin{bmatrix} 1 & 0 & -U_0 \\ * & 1 & -V_0 \\ * & * & f^2 + U_0^2 + V_0^2 \end{bmatrix}$$

This matrix can be determined by solving the following system of three equations involving the vanishing points $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$:

$$\begin{cases} \mathbf{v}_1^T \boldsymbol{\omega} \mathbf{v}_2 = \mathbf{0} \\ \mathbf{v}_2^T \boldsymbol{\omega} \mathbf{v}_3 = \mathbf{0} \\ \mathbf{v}_3^T \boldsymbol{\omega} \mathbf{v}_1 = \mathbf{0} \end{cases}$$

5.3.7 Calibration from reconstructed shape and normal vanishing point

In this calibration method, we use a planar face with a known (reconstructed) shape and a vanishing point of the normal to the plane. The image of the absolute conic matrix $\boldsymbol{\omega}$ in this case is:

$$\boldsymbol{\omega} = \begin{bmatrix} a^2 & 0 & -u_0 a^2 \\ * & 1 & -v_0 \\ * & * & f_Y^2 + a^2 u_0^2 + v_0^2 \end{bmatrix}$$

Here, a represents the pixel aspect ratio. To fully define this matrix, we need four equations:

$$\begin{cases} \mathbf{h}_1^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{v}^T \boldsymbol{\omega} \mathbf{h}_4 = \mathbf{0} \\ \mathbf{v}^T \boldsymbol{\omega} \mathbf{h}_1 = \mathbf{0} \\ \mathbf{v}^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \end{cases}$$

These equations allow us to solve for $\boldsymbol{\omega}$. Once we have $\boldsymbol{\omega}$, we can use Cholesky factorization on its inverse to obtain the camera calibration matrix.

Direct method The direct method is independent of the chosen pairs of mutually orthogonal vanishing points. It works from:

1. The reconstructed homography $\mathbf{H}_R$ from the given image to the rectified image.
2. The image of the line at infinity $\mathbf{l}'_\infty$, or the vanishing point $\mathbf{v}$ along the direction orthogonal to the face.

Consider the plane-to-image homography $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} = \mathbf{H}_R^{-1}$. Additionally, the line at infinity $\mathbf{l}'_\infty = \mathbf{v}_1 \times \mathbf{v}_2$. From this, we can extract the following constraints:

$$\begin{cases} \mathbf{h}_1^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{h}_1^T \boldsymbol{\omega} \mathbf{h}_1 - \mathbf{h}_2^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{l}'_\infty = \boldsymbol{\omega} \mathbf{v} \end{cases}$$

The last term of the system gives two equations.

Alternatively, we can impose the following constraints:

$$\begin{cases} \mathbf{h}_1^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{h}_1^T \boldsymbol{\omega} \mathbf{h}_1 - \mathbf{h}_2^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \\ \mathbf{v}^T \boldsymbol{\omega} \mathbf{h}_1 = \mathbf{0} \\ \mathbf{v}^T \boldsymbol{\omega} \mathbf{h}_2 = \mathbf{0} \end{cases}$$

5.4 Calibrated camera rectification

Consider a calibrated camera capturing a single image of a planar object. If two vanishing points can be identified, the line at infinity $\mathbf{l}'_\infty$ passing through them can also be determined. The plane-to-image homography $\mathbf{H}$ describes the mapping of points from the plane to the image. The inverse transformation maps image points back onto the plane:

$$\mathbf{H}_R = \mathbf{H}^{-1}$$

Given that the homography is expressed as:

$$\mathbf{H} = \mathbf{K} \begin{bmatrix} \mathbf{r_{\pi 1}} & \mathbf{r_{\pi 2}} & \mathbf{o_\pi} \end{bmatrix}$$

It follows that $\mathbf{H}$ depends on both the camera calibration matrix $\mathbf{K}$ and the relative pose of the plane $\boldsymbol{\pi}$ with respect to the camera.

Since the only constrained element is the normal direction $\mathbf{n_\pi}$ of the plane, we are free to define a convenient reference frame on $\boldsymbol{\pi}$. The normal is given by:

$$\mathbf{n_\pi} = \mathbf{K}^T \mathbf{l}'_\infty$$

To simplify rectification, we choose $\mathbf{r_{\pi 1}}$ and $\mathbf{r_{\pi 2}}$ as orthogonal vectors, both perpendicular to $\mathbf{n_\pi}$ and normalize them. Setting $\mathbf{o_\pi} = \mathbf{n_\pi}$, the rectifying homography becomes:

$$\mathbf{H}_R = \mathbf{R}_{\boldsymbol{\pi}}^T \mathbf{K}^{-1}$$

5.4.1 Rectification with unknown planar scene

In the case of reconstructing an unknown planar scene from two calibrated images, image-to-image homographies are utilized. The following constraints hold:

$$\begin{cases} \mathbf{I}'^T \boldsymbol{\omega} \mathbf{I}' = \mathbf{0} \\ \mathbf{I}'^T \mathbf{H}'^T \boldsymbol{\omega} \mathbf{H}' \mathbf{I}' = \mathbf{0} \end{cases}$$

Here, $\boldsymbol{\omega}$ is known, and the unknowns are the complex coordinates $\mathbf{I}'$. At least two images are required, as each equation provides two constraints (real and imaginary parts). Geometrically, this corresponds to the intersection of two conics, resulting in two pairs of imaged circular points. The selection of the correct solution can be guided by reprojection error minimization or an additional third image.

The steps for the rectification are:

1. Extract the image of circular points and compute the conic dual to the circular points:

$$\mathbf{C}_\infty^{*\prime} = \mathbf{I}' \mathbf{J}'^T + \mathbf{J}' \mathbf{I}'^T$$

2. Apply Singular Value Decomposition to $\mathbf{C}_\infty^{*\prime}$:

$$\mathrm{SVD}(\mathbf{C}_\infty^{*\prime}) = \mathbf{U} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{U}^T$$

3. Construct the rectification matrix from SVD output $\mathbf{U}$:

$$\mathbf{H}_{SR} = \begin{bmatrix} \frac{1}{\sqrt{a}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{b}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{U}^T$$

4. Apply the rectification transformation to the image:

$$\mathbf{M}_S = \mathbf{H}_{SR} \cdot \text{image}$$

5.5 Localization

The localization of a known planar shape from a calibrated image consists of determining the relative position and orientation of the object with respect to the camera. This requires estimating the homography $\mathbf{H}$ that maps points from the plane $\boldsymbol{\pi}$ to their corresponding image points.

The plane $\boldsymbol{\pi}$ has a relative pose with respect to the camera, which is described by a rotation $\mathbf{R}_{\boldsymbol{\pi}}$ and a translation $\mathbf{0}_{\boldsymbol{\pi}}$. Assuming the world reference frame is aligned with the camera frame, the projection matrix is $\mathbf{P} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix}$. From this, the object's pose relative to the camera can be determined using the homography:

$$\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} = \mathbf{K}^{-1}\mathbf{H}$$

Here $\mathbf{r}_1$ and $\mathbf{r}_2$ are the first two columns of the rotation matrix, and the third column is given by:

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

If the coordinates of the points $\mathbf{x}_{\boldsymbol{\pi}\mathbf{j}}$ on the plane are known only up to scale, the translation vector $\mathbf{o}_{\boldsymbol{\pi}}$ can also only be determined up to scale. This means that its direction is known, but its magnitude remains uncertain.

Knowing the shape of the object allows estimation of its image orientation and the viewing direction. Knowing both the shape and size of the object enables full determination of its image position, orientation, and viewing direction.

CHAPTER 6

Multi view geometry

6.1 Introduction

When we project a 3D scene onto a 2D image, there's an inherent ambiguity because different depths can't be distinguished in the image. This means that it can be difficult to determine the exact 3D positions of points just from a single image.

However, using multiple views of the same scene can help resolve this ambiguity. By taking images from different perspectives, we get different information about the scene that can be used to better understand its 3D structure.

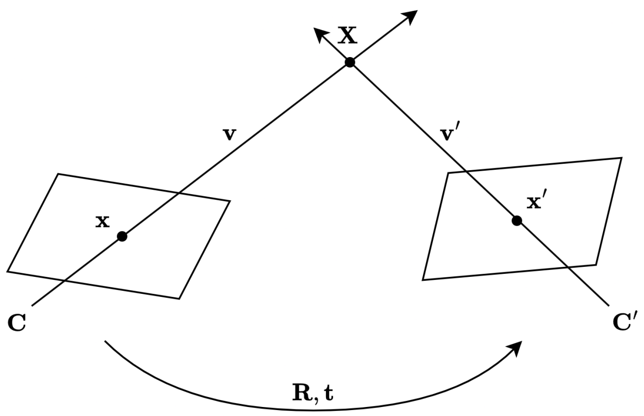


Figure 6.1: Two view geometry

In a 3D scenario, we can obtain multiple views of a point $\mathbf{X}$ from different cameras, and there are several ways to compute its position:

- *Stereo vision*: we know the mapping between the image points $\mathbf{x}$ and $\mathbf{x}'$, as well as the rotation and translation between the two cameras. The calibration matrix for both cameras is also known. To find the 3D point $\mathbf{X}$, we compute the viewing rays $\mathbf{v}$ or $\mathbf{v}'$ from the calibration matrices $\mathbf{K}$ or $\mathbf{K}'$, and the image points $\mathbf{x}$ or $\mathbf{x}'$. Then, we use triangulation to determine the 3D location:

$$\mathbf{X} = \mathbf{v} \cap \mathbf{v}'$$

- *Calibrated structure from motion*: we know the mapping between the image points $\mathbf{x}$ and $\mathbf{x}'$, and the calibration matrices for both cameras are known, but the rotation and translation between the two images are unknown. To find the 3D point $\mathbf{X}$, as well as the rotation matrix $\mathbb{R}$ and translation vector $\approx$, we use the epipolar constraint to estimate $\mathbb{R}$ and $\approx$. Afterward, we compute the viewing rays $\mathbf{v}$ and $\mathbf{v}'$ and apply triangulation:

$$\mathbf{X} = \mathbf{v} \cap \mathbf{v}'$$

- *Uncalibrated structure from motion*: we know the mapping between $\mathbf{x}$ to $\mathbf{x}'$, the rotation, translation between the images, and the calibration matrices for the cameras are unknown. To estimate the 3D point $\mathbf{X}$, along with the camera calibration matrices $\mathbf{K}$, $\mathbf{K}'$, $\mathbb{R}$, and $\approx$, we can use the epipolar constraint and partial information about the scene or cameras. Once we have these estimates, we compute the viewing rays $\mathbf{v}$ and $\mathbf{v}'$ and use triangulation:

$$\mathbf{X} = \mathbf{v} \cap \mathbf{v}'$$

6.2 Epipolar geometry

Epipolar geometry describes the relationship between two views of the same 3D scene. In the diagram below, the scene point, its projections, and the camera centers are coplanar, meaning the cameras' optical centers ($\mathbf{C}$ and $\mathbf{C}'$), image points ($\mathbf{x}$ and $\mathbf{x}'$), and the 3D scene point ($\mathbf{X}$) all lie on the same plane.

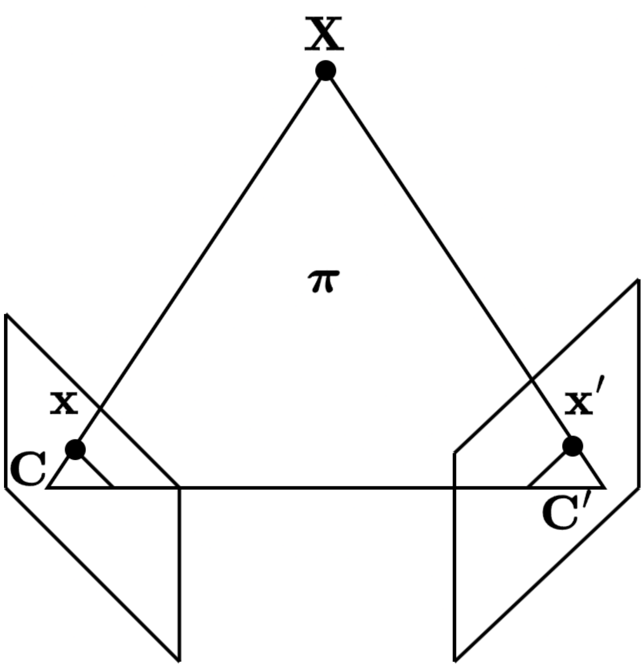


Figure 6.2: Epipolar geometry

The epipolar constraint states that the viewing rays corresponding to matching image points must intersect in 3D space. However, if only the camera centers and one image pointare known, the possible position of the 3D scene point lies somewhere along the viewing ray associated with the point.

Epipolar line The epipolar line is where the image of the 3D point varies along a line $\mathbf{l}'$ in the second image. Specifically, $\mathbf{l}'$ is the image of the viewing ray, the line connecting the camera center to the 3D point. The epipolar lines in both images correspond to each other and are related geometrically. The viewing ray in the 3D space is a line through the first camera center. Therefore, its image in the second camera image plane, $\mathbf{l}'$, always passes through the epipole $\mathbf{e}'$, which is the projection of the first camera center in the second image.

Epipolar constraint All points on the epipolar plane $\boldsymbol{\pi}$ project onto corresponding epipolar lines $\mathbf{l}$ and $\mathbf{l}'$ in the two image planes. The family of coaxial planes $\boldsymbol{\pi}$ where the baseline between the two cameras serves as the axis, intersects the image planes at the epipolar lines $\mathbf{l}$ and $\mathbf{l}'$.

The epipolar lines $\mathbf{l}$ and $\mathbf{l}'$ converge at the epipoles $\mathbf{e}$ and $\mathbf{e}'$, respectively. These epipoles represent:

- The intersection of the baseline with the image plane.
- The projection of the camera centers in the other image.
- The vanishing point of the relative motion between the two cameras.

Definition (*Epipolar plane*)**.** A plane containing the baseline between the two cameras. It defines a 1D family of planes that intersect the image planes in epipolar lines.

Definition (*Epipolar line*)**.** The intersection of an epipolar plane with the image plane. These lines come in corresponding pairs in each image and always pass through the epipoles.

Definition (*Epipolar points*)**.** Points where the baseline intersects the image planes. They are the projections of the camera centers in the opposite image and also serve as the vanishing points of the camera's relative motion direction.

6.2.1 Fundamental matrix

The fundamental matrix provides an algebraic representation of the epipolar geometry between two views of a scene. It defines how points in one image correspond to lines in the other. This relationship can be described as follows:

$$\mathbf{x} \mapsto \mathbf{l}' = \mathbf{F}\mathbf{x}$$

In this equation, $\mathbf{F}$ is the fundamental matrix, and it maps points from one image to corresponding epipolar lines in the other image. This mapping is a form of correlation.

Mathematically, this is represented by the following expression:

$$\mathbf{F} = \left[\mathbf{e}'\right]_\times \mathbf{M}'\mathbf{M}^{-1}\mathbf{x}$$

Here, $\left[\mathbf{e}'\right]_\times$ is a skew-symmetric matrix used to compute the cross-product via matrix multiplication:

$$\left[\mathbf{e}'\right]_\times = \begin{bmatrix} 0 & -e'_z & -e'_y \\ e'_z & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{bmatrix}$$

This matrix is singular, meaning it doesn't have an inverse.

The fundamental matrix satisfies a key condition: for any pair of corresponding points $(\mathbf{x}, \mathbf{x}')$ in two images, the following equation holds:

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = \mathbf{0}$$

Here, $\mathbf{x}' T \mathbf{l}' = 0$. Even if the cameras are not calibrated and their projection matrices $\mathbf{P}$ and $\mathbf{P}'$ are unknown, the fundamental matrix $\mathbf{F}$ can still be computed by using point correspondences between the two images.

Properties The fundamental matrix has several important properties:

- *Uniqueness*: $\mathbf{F} = \left[\mathbf{e}'\right]_\times \mathbf{M}'\mathbf{M}^{-1}$ is the only 3×3 matrix with rank two that satisfies $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$ for all corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$.
- *Transpose property*: if $\mathbf{F}$ is the fundamental matrix for the camera pair $(\mathbf{P}, \mathbf{P}')$, then $\mathbf{F}^T$ is the fundamental matrix for the reversed camera pair $(\mathbf{P}', \mathbf{P})$.
- *Epipolar lines*: the epipolar line corresponding to a point $\mathbf{x}$ in the first image is given by $\mathbf{l} = \mathbf{F}\mathbf{x}$, and the epipolar line corresponding to a point $\mathbf{x}'$ in the second image is $\mathbf{l}' = \mathbf{F}^T \mathbf{x}'$.
- *Epipoles*: the epipoles are the points where all the epipolar lines converge. They lie on the null space of $\mathbf{F}$. Thus, $\mathbf{e}'^T \mathbf{F} = 0$ and $\mathbf{F}\mathbf{e} = 0$, where $\mathbf{e}$ and $\mathbf{e}'$ are the epipoles in the first and second image, respectively.
- *Degrees of freedom*: the fundamental matrix has 7 degrees of freedom, as it is a 3×3 matrix with 9 elements, but constrained by the rank-2 condition and the homogeneous coordinate system (removing 2 degrees of freedom).
- *Projective mapping*: the fundamental matrix represents a projective mapping from a point in one image to a line in the other image. This is not a proper correlation because the mapping is not invertible.

Computation To compute the fundamental matrix $\mathbf{F}$ from a set of corresponding image points, we solve the following system of equations:

$$\mathbf{F}\mathbf{x}_i = 0$$

These equations are linear in the elements of $\mathbf{F}$, and the matrix can be computed using various methods. The most common approaches are:

- Using 8 point pairs, which provides a linear system of equations.
- Using 7 point pairs, which gives a non-linear system.
- Using 8 or more point pairs and solving the system using least squares optimization.

6.3 Triangulation

In triangulation, the 3D scene point $\mathbf{X}$ is determined by the intersection of two viewing rays:

- One viewing ray is associated with image point $\mathbf{x}$ and originates from camera $\mathbf{P}$.

- The other viewing ray is associated with image point $\mathbf{x}'$ and originates from camera $\mathbf{P}'$.

This process finds the 3D point $\mathbf{X}$ by intersecting the rays in space.

Theorem 6.3.1. *A fundamental matrix* $\mathbf{F}_{12}$ *is compatible with the camera pairs* $(\mathbf{P}_1, \mathbf{P}_2)$ *and* $(\mathbf{P}'_1, \mathbf{P}'_2)$ *if and only if the camera pairs are projectively related. Specifically, there exists an invertible* 4×4 *matrix* $\mathbf{H}$ *such that:*

$$\mathbf{P}'_1 = \mathbf{P}_1\mathbf{H}^{-1} \qquad \mathbf{P}'_2 = \mathbf{P}_2\mathbf{H}^{-1}$$

This means that the camera matrices of two projectively related camera pairs are connected through a transformation matrix $\mathbf{H}$. Thus, the images captured by both pairs are the same, even if the cameras themselves are modified, as long as they remain related by a projective transformation.

The fundamental matrix $\mathbf{F}_{12}$ remains invariant for two camera pairs that are projectively related. Specifically:

- Any set of images of a scene can be transformed through a projective transformation, leading to identical images if the camera matrices are adjusted accordingly.
- If two camera pairs are projectively related, they will share the same fundamental matrix $\mathbf{F}_{12}$.

If the cameras $\mathbf{P}_1 = \begin{bmatrix}\mathbf{I} \mid \mathbf{0}\end{bmatrix}$ and $\mathbf{P}_2 = \begin{bmatrix}\mathbf{A} \mid \mathbf{a}\end{bmatrix}$ are compatible with the fundamental matrix $\mathbf{F}_{12}$, then the modified camera matrices also cameras $\mathbf{P}'_1 = \begin{bmatrix}\mathbf{I} \mid \mathbf{0}\end{bmatrix}$ and $\mathbf{P}'_2 = \begin{bmatrix}\mathbf{A} + \mathbf{a}\mathbf{v}^T \mid \lambda\mathbf{a}\end{bmatrix}$, for any vector $\mathbf{v}$ and scalar λ, are also compatible with the same fundamental matrix $\mathbf{F}_{12}$.