

Image Analysis And Computer Vision
Theory

Christian Rossi

Academic Year 2023-2024

Abstract

The topics of the course are:

- Introduction.
- Camera sensors: transduction, optics, geometry, distortion
- Basics on Projective geometry: modelling basic primitives (points, lines, planes, conic sections, quadric surfaces) and projective spatial transformations and projections.
- Camera geometry, and single view analysis: calibration, image rectification, localization of 3D models.
- Multi-view analysis: 3D shape reconstruction, self-calibration, 3D scene understanding.
- Linear filters and convolutions, space-invariant filters, Fourier Transform, sampling and aliasing.
- Nonlinear filters: image morphology and morphology operators (dilate, erode, open, close), median filters.
- Edge detection and feature detection techniques. Feature matching and feature tracking along image sequences.
- Inferring parametric models from noisy data (including outliers), contour segmentation, clustering, Hough Transform, Ransac (random sample consensus).
- Applications: object tracking, object recognition, classification.

Contents

1	Optical Sensors	2
1.1	Photocamera	2
1.2	Light rays deviation	3
1.3	Focalization of parallel light rays	4
1.4	Path of a light ray	5
1.5	Pin-hole camera	5
1.6	From real world to 2D images	6
1.7	Perspective and vanishing point	7
2	Planar Projective Geometry	8
2.1	Introduction	8
2.2	Points	8
2.3	Lines	10

Chapter 1

Optical Sensors

1.1 Photocamera

Definition

The *photocamera* is an optical sensor; this means that produces data using electric transducers. It uses an optical system that select the direction of the incoming light at each element of its screen made with millions of photosensitive elements. Most of the actual cameras can capture up to 30-60 frames per second.

For simplicity, we suppose that the optical system of a photocamera is a single lens that is:

- Spherical: the lens is obtained by intersecting two spheres.
- Thin: the distance between the center of the two spheres is almost identical to the sum of the diameter of them.
- Small angles: the light forms small inclination angles with the optical axis.

This simplifies the computation of the path of the ray crossing the lens. In fact, the refraction of the light when crossing a border between two media is given by the Snell's law:

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

where:

- θ_1 and θ_2 are the angles between the normal at the surface and the direction of the light ray, respectively before and after crossing the border.

- n_1 and n_2 are the refraction indexes of the two materials.

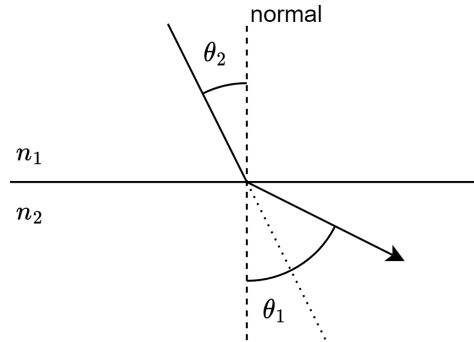


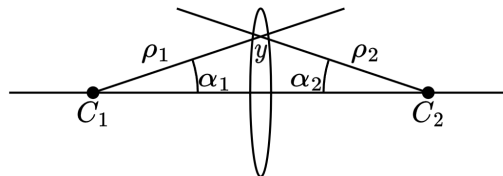
Figure 1.1: Snell's law

Definition

The *optical axis* is the straight line that connects the center of the two spheres that are used to form the lens.

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

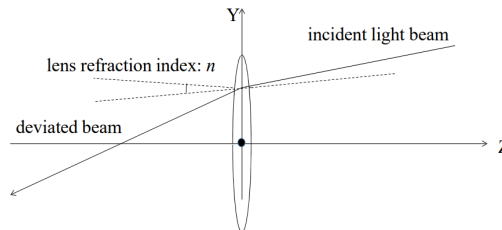
$$\alpha_1 = \frac{y_1}{\rho_1} \quad \alpha_2 = -\frac{y_2}{\rho_2}$$



Since we have a simplified lens, it is possible to say that:

$$y_1 = y_2 = y$$

1.2 Light rays deviation



Given a lens with refraction n we have that the following equations are valid:

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

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

where:

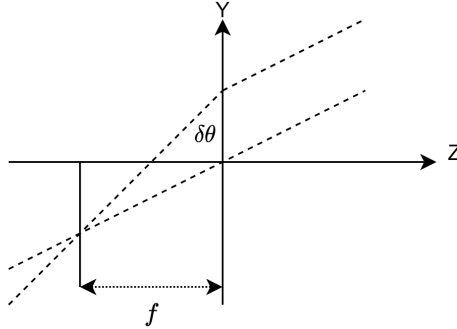
- θ is the angle before the lens.
- θ' is the angle in the lens (not visible in the image).
- θ'' is the angle after the lens.

Comparing the two equations it is possible to find the difference between the input angle θ and the output angle θ'' that is:

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

It is possible to see that the first term $(n - 1)$ is due to the matter of the lens and the second $\left(\frac{1}{\rho_1} + \frac{1}{\rho_2} \right)$ depend on the curvature of the surface.

1.3 Focalization of parallel light rays



In the image we have one ray that passes through the center of the lens and the other that passes in another point, but it is parallel to the first one. So we have that:

- $Y = 0$: the deviation of the ray is null and proceed without being deviated.

$$\bullet Y = f \cdot \delta\theta \Rightarrow f = \frac{1}{(n-1) \left(\frac{1}{\rho_1} + \frac{1}{\rho_2} \right)}.$$

This means that all the rays that proceed in parallel meet in one common point on the focal point Z , with a distance from the Y axis equal to:

$$Z = -f$$

1.4 Path of a light ray

To find the projection of a light ray crossing a lens in any position we need to:

1. Draw a line parallel to the selected ray and that passes through the center of the lens.
2. Intersect the line with the focal plane.
3. The ray will go from the point in which it crosses the lens to the point found on the focal plane.

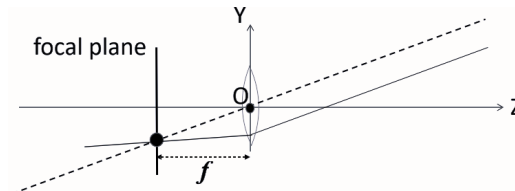


Figure 1.2: Path of a light ray through a lens

1.5 Pin-hole camera

To have a focussed image we need that every ray hits the focal plane of the camera in exactly one pixel. To obtain this we need that the distance between the lens and the source of the ray $Z(P)$ must be much greater than the lens aperture a (at least $1000\times$). In this way we can place the screen at distance Z from the lens. If all the rays respect this constraint they will be all parallel for the lens and the image will be on focus. The camera that we have defined so far is called pin-hole camera and needs:

1. Thin spherical lens.

2. Small angles.
3. $Z(P) \gg a$.
4. $Z = f$.

1.6 From real world to 2D images

The images are in a 2D plane while the real world is in three dimensions. So, a picture have less information than the original subject in the real world. In fact, the space projection of space in a 2D image is a perspective projection, that is:

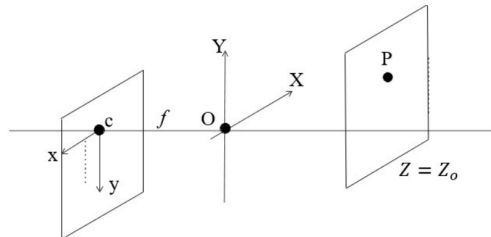
- Nonlinear.
- Not shape-preserving.
- Not length-ratio preserving.

Due to the triangle equality we have that:

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

The only way to avoid this information loss is to take a picture of a planar scene on plane parallel to image plane, so we need to have

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



In this case the only difference between reality and projection will be the down scaling, while the other dimensions are preserved:

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

1.7 Perspective and vanishing point

If we decide to choose all the lines parallel to the direction parameters $\begin{bmatrix} \alpha & \beta & 1 \end{bmatrix}$ we obtain the following system:

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

All these lines can be projected in the 2D image with the triangle equality in the following way:

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

We now want to find the image of the point at the infinity of these lines, so we obtain the point:

$$[f\alpha f\beta]$$

that is independent of X_0 and Y_0 . So, we have that all parallel lines have the same image if their points at the infinity.

Definition

The image of the point at the infinity of the lines is called *vanishing point*.

So, we have that all the lines that are parallel in the real world will project onto concurrent lines in the image.

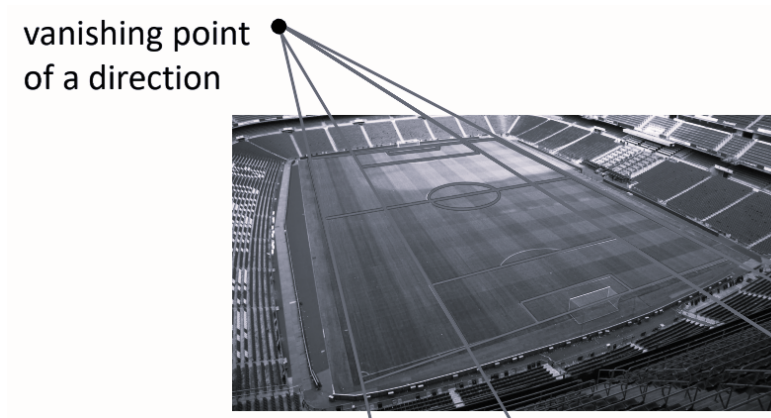


Figure 1.3: Example of vanishing point

Chapter 2

Planar Projective Geometry

2.1 Introduction

The elements needed to define a planar geometry are: point, lines, conics and dual conics. The possible transformations in this type of geometry are: projectivities, affinities, similarities and isometries.

2.2 Points

The points can be defined in cartesian coordinates by defining a euclidean plane with his origin. In this way every point is defined unambiguously with two cartesian coordinates (x, y) .

To analyze images it is better to use homogeneous coordinates. To define this type of coordinates we need to construct a 3D space with x, y, w axis. So, now to define a point we have to assign three values. This means that every number can be represented in infinite ways by changing the value of w . The correlation between cartesian and homogeneous coordinates is the following:

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

So we have that a vector $\bar{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$ and all its nonzero multiples $\lambda \begin{bmatrix} x \\ y \\ w \end{bmatrix}$,

including $\begin{bmatrix} x/w \\ y/w \\ 1 \end{bmatrix}$, represent the point of cartesian coordinates $\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x/w \\ y/w \end{bmatrix}$

on the euclidean plane. This type of representation has the homogeneity

property: any vector \bar{x} is equivalent to all its nonzero multiples $\lambda\bar{x}, \lambda \neq 0$, since they represent the same point. The null vector does not represent any point.

Definition

The *projective plane* is defined as:

$$\mathbb{P}^2 = \{[x \ y \ w]^T \in \mathbb{R}^3\} - \{[0 \ 0 \ 0]^T\}$$

Example: The origin of the plane is defined as:

$$[0 \ 0 \ 1]^T$$

A generic point in heterogeneous coordinates can be transformed into a couple of cartesian coordinate with a simple division. The point:

$$[0 \ 8 \ 4]^T$$

in cartesian coordinate is equal to:

$$[x/w \ y/w]^T = [0/4 \ 8/4] = [0 \ 4]$$

Consider the point $\bar{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$ and let w slowly drop to 0 starting from $w = 1$.

As w decreases, the point will move along a constant direction $[x \ y]$, with increasing distance from the origin. As w tends to 0, this points tends to the infinity along the direction $[x \ y]$.

Definition

The point

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

is called the *point at the infinity along the direction* $[x \ y]$.

Points at the infinity, who represent directions, are not part of the euclidean plane: they are extra points, well-defined within the projective plane. So, in general we have that the projective plane is the euclidean plane with also the points at the infinity.

2.3 Lines

In the euclidean plane a line is defined as:

$$aX + bY + c = 0$$

In the heterogeneous plane the lines are defined as:

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

This equation can be also represented using two vectors called respectively \bar{l}^T and \bar{x} :

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

where the vector $\bar{l} = [a \ b \ c]^T$ and all its nonzero multiples represent a line. This type of representation has the homogeneity property: any vector \bar{l} is equivalent to all its nonzero multiples $\lambda\bar{l}$, $\lambda \neq 0$, since they represent the same line. a, b, c are called homogeneous parameters of the line.

As for numbers, there is an infinite number of equivalent representations for a single line, namely all nonzero multiples of the unit normal vector. The null vector does not represent any lines.

Definition

The *projective dual plane* is defined as:

$$\mathbb{P}^2 = \{[a \ b \ c]^T \in \mathbb{R}^3\} - \{[0 \ 0 \ 0]^T\}$$

There are three important remarks:

1. If the third parameter is null, $\bar{l} = [a \ b \ 0]^T$, then the line goes through point $[0 \ 0]$.
2. Within the euclidean plane, direction $[a \ b]$ is normal to the line $\bar{l} = [a \ b \ c]^T$.
3. Two lines $\bar{l} = [a \ b \ c]^T$ and $\bar{l}' = [a' \ b' \ c']^T$ are parallel: their common direction is $[b, -a]$.

Example: The cartesian axes are defined as:

$$\bar{l}_x = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \bar{l}_y = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The incidence relation of a line $\bar{l}^T \bar{x} = 0$ is defined when:

- The point \bar{x} is on the line \bar{l} or
- The line \bar{l} goes through the point \bar{x} .

Definition

The line

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

is called the *line at the infinity* $\bar{l}_\infty = [0 \ 0 \ 1]^T$.

Since the dot product is commutative, also the incidence relation is commutative, so the sentences "point \bar{x} is on line \bar{l} " and "point \bar{l} is on line \bar{x} " have the same algebraic meanings. This is called the principle of duality between points and lines. Given the lines \bar{l}_1^T and \bar{l}_2^T , the intersection can be found by imposing:

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

that is equal to find the right null space of the first column vector.

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

The system found is underdetermined, so there is only one intersection point between two lines (that can be represented in infinite ways in homogeneous plane). Note that the vector \bar{x} is orthogonal to both lines, so in the 2D projective geometry can be found with the cross product:

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

Example: Suppose we have two parallel lines $\bar{l}_1 = [a \ b \ c_1]^T$ and $\bar{l}_2 = [a \ b \ c_2]^T$. The point that is common to both points can be found with the system:

$$\begin{cases} ax + by + c_1w = 0 \\ ax + by + c_2w = 0 \end{cases}$$

So the solution will be $\bar{x} = [b \ -a \ 0]^T$ that is the point at the infinity along the direction of both lines.

The line through two points can be found with the dual of the previous problem, that is:

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

that in 2D become:

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

Property: The point \bar{x} given by the linear combination $\bar{x} = \alpha\bar{x}_1 + \beta\bar{x}_2$ of two points \bar{x}_1 and \bar{x}_2 is on the line \bar{l} through \bar{x}_1 and \bar{x}_2 .

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

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

■

This proves that there is also a duality between colinear and concurrent.

Theorem (*duality principle*)

For any true sentence containing the words: point, line, is on and goes through there is a dual sentence (also true) obtained by substituting, in the previous one, each occurrence of:

- Point \Leftrightarrow line.
- Is on \Leftrightarrow goes through.
- Colinear \Leftrightarrow concurrent.

Within the euclidean plane, $[a \ b]$ is the direction normal to the line $\bar{l} = [a \ b \ c]^T$. Since we have that the angle between two lines is the angle between their normal and that the angle between two vectors is equal to:

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

we have that the angle between two lines $\bar{l}_1 = [a_1 \ b_1 \ c_1]^T$ and $\bar{l}_2 = [a_2 \ b_2 \ c_2]^T$ is the angle between their normal directions $[a_1 \ b_1]$ and $[a_2 \ b_2]$, namely:

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

The ratio of length is an invariant