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

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 intersecating two spheres.
- Thin: the distance between the center of the two spheres is almost identical to the sum of the diameter of them.

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

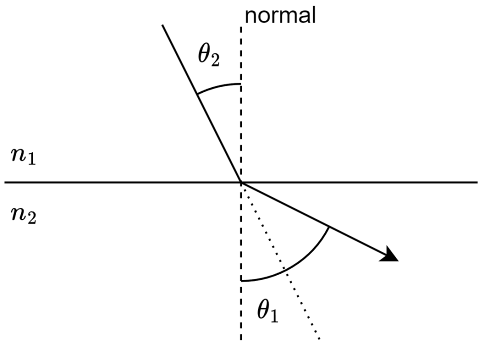


Figure 1.1: Snell's law

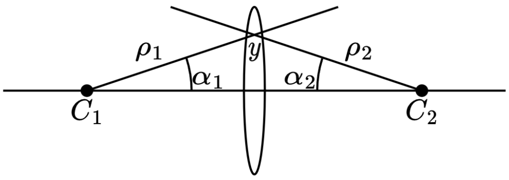
Other than that, we also have to hypothesize that the light forms small inclination angles with the optical axis.

Definition

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

The angles of a ray passing trough the centres of the spheres can be calculated as follows:

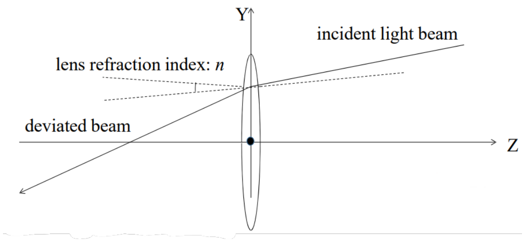
$$\alpha_1 = \frac{y_1}{\rho_1} \qquad \alpha_2 = -\frac{y_2}{\rho_2}$$



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 equeations 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:

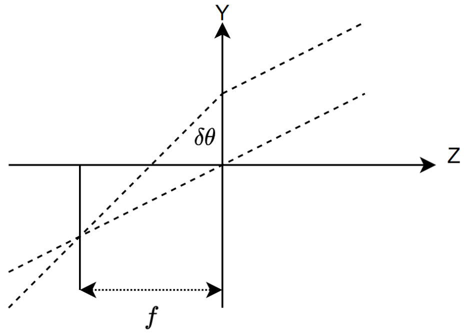
- θ is the incident angle (light on the lens).
- θ' is the angle in the lens (not visible in the image).
- θ'' is the angle after the lens.

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 trough the centre of the lens and the other that passes in another point but it is parallel to the first one. So we have that:

- $Y = 0$, so we have that the deviation of the ray is null and proceed without being deviated.
- $Y = f \cdot \delta\theta \Rightarrow f = \dfrac{1}{(n-1)\left(\dfrac{1}{\rho_1} + \dfrac{1}{\rho_2}\right)}$.

This means that all the rays that proceed in parallel meets 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.

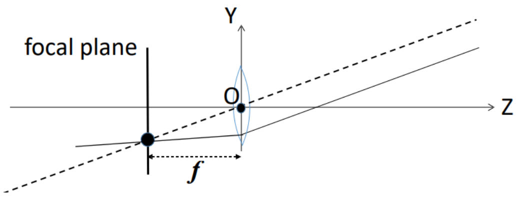


Figure 1.2: Path of a light ray through a lens

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 they 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:

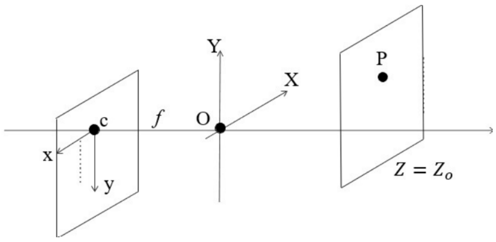
- Nonlinear.
- Not shape-preserving.
- Not length-ratio preserving.

Due to the triangle equality we have that:

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

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:

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

1.7 Perspective and 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:

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

So we have that a vector $\overline{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$ 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 $\overline{x}$ is equivalent to all its nonzero multiples $\lambda\overline{x}, \lambda \neq 0$, since they represent the same point. Since the null vector does not represent any point we have that:

Definition

The *projective plane* is defined as:

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

Example : The origin of the plane is defined as:

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

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

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

in cartesian coordinate is equal to:

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

Consider the point $\overline{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix}$ and let w slowly drop to 0 starting from $w = 1$. As w decreases, the point will move along a constant direction $\begin{bmatrix} x & y \end{bmatrix}$, with increasing distance from the origin. As w tends to 0, this points tends to the infinity along the direction $\begin{bmatrix} x & y \end{bmatrix}$.

Definition

The point

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

is called the *point at the infinity along the direction* $\begin{bmatrix} x & y \end{bmatrix}$.

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:

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

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

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

where the vector $\bar{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$ 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:

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

There are three important remarks:

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

Example : The cartesian axes are defined as:

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

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

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

Definition

The line

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

is called the *point at the infinity* $\bar{l}_\infty = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$.