Chapter 3
Perspective Projection

Multiple View Geometry
Summer 2018

Prof. Daniel Cremers
Chair for Computer Vision and Artificial Intelligence
Departments of Informatics & Mathematics
Technische Universität München

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Overview

1. Historic Remarks
2. Mathematical Representation
3. Intrinsic Parameters
4. Spherical Projection
5. Radial Distortion
6. Preimage and Coimage
7. Projective Geometry

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Some Historic Remarks

The study of the image formation process has a long history. The earliest formulations of the geometry of image formation can be traced back to Euclid (4th century B.C.). Examples of a partially correct perspective projection are visible in the frescoes and mosaics of Pompeii (1 B.C.).

These skills seem to have been lost with the fall of the Roman empire. Correct perspective projection emerged again around 1000 years later in early Renaissance art.

Among the proponents of perspective projection are the Renaissance artists Brunelleschi, Donatello and Alberti. The first treatise on the projection process, “Della Pittura” (1435) was published by Leon Battista Alberti).

Apart from the geometry of image formation, the study of the interaction of light with matter was propagated by artists like Leonardo da Vinci in the 1500s and by Renaissance painters such as Caravaggio and Raphael.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

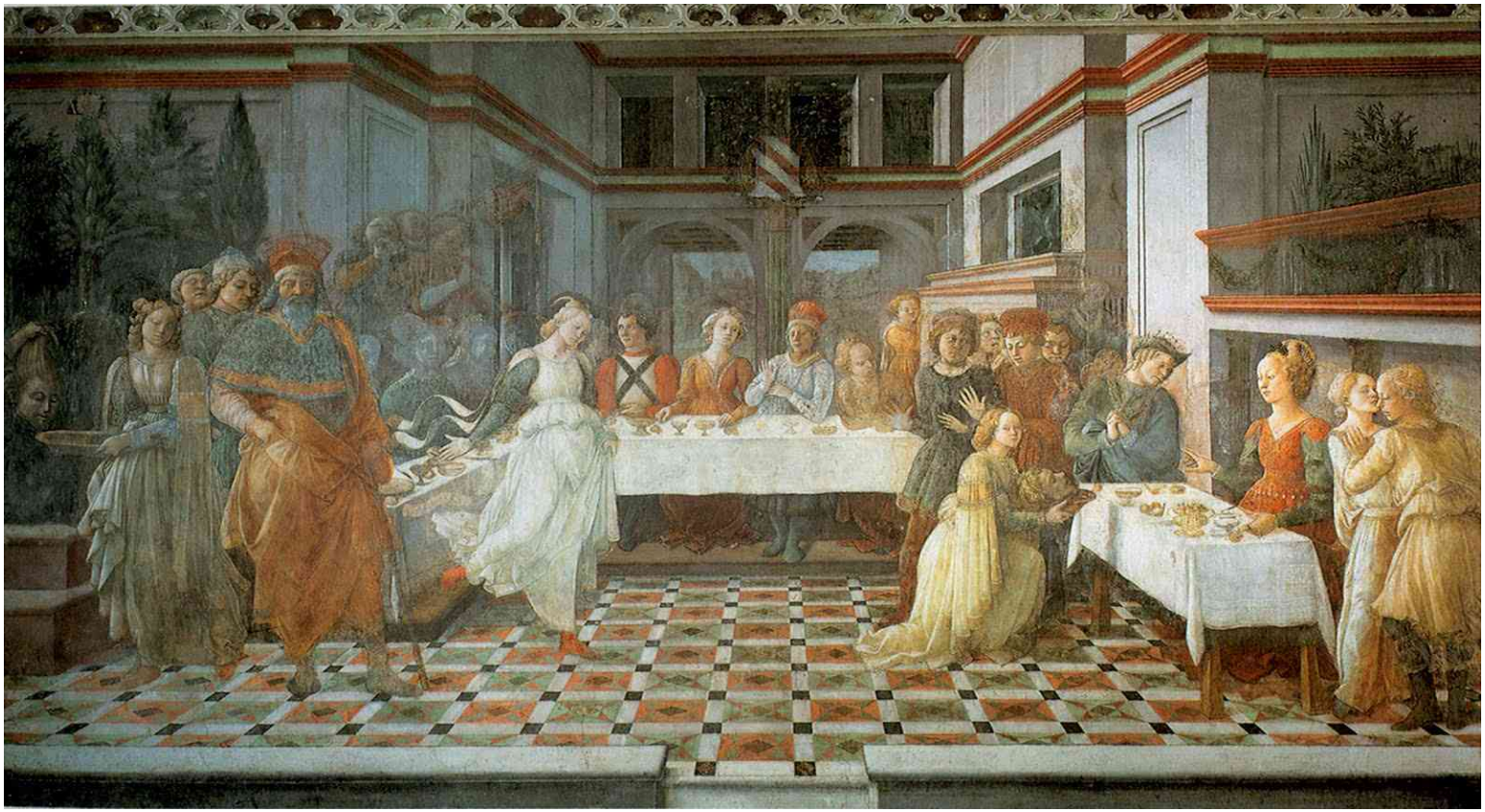
Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Perspective Projection in Art



Filippo Lippi, “The Feast of Herod: Salome’s Dance.”
Fresco, Cappella Maggiore, Duomo, Prato, Italy, c.1460-1464.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Perspective Projection in Art

Raphael, The School of Athens (1509)

Historic Remarks

Mathematical Representation

Intrinsic Parameters

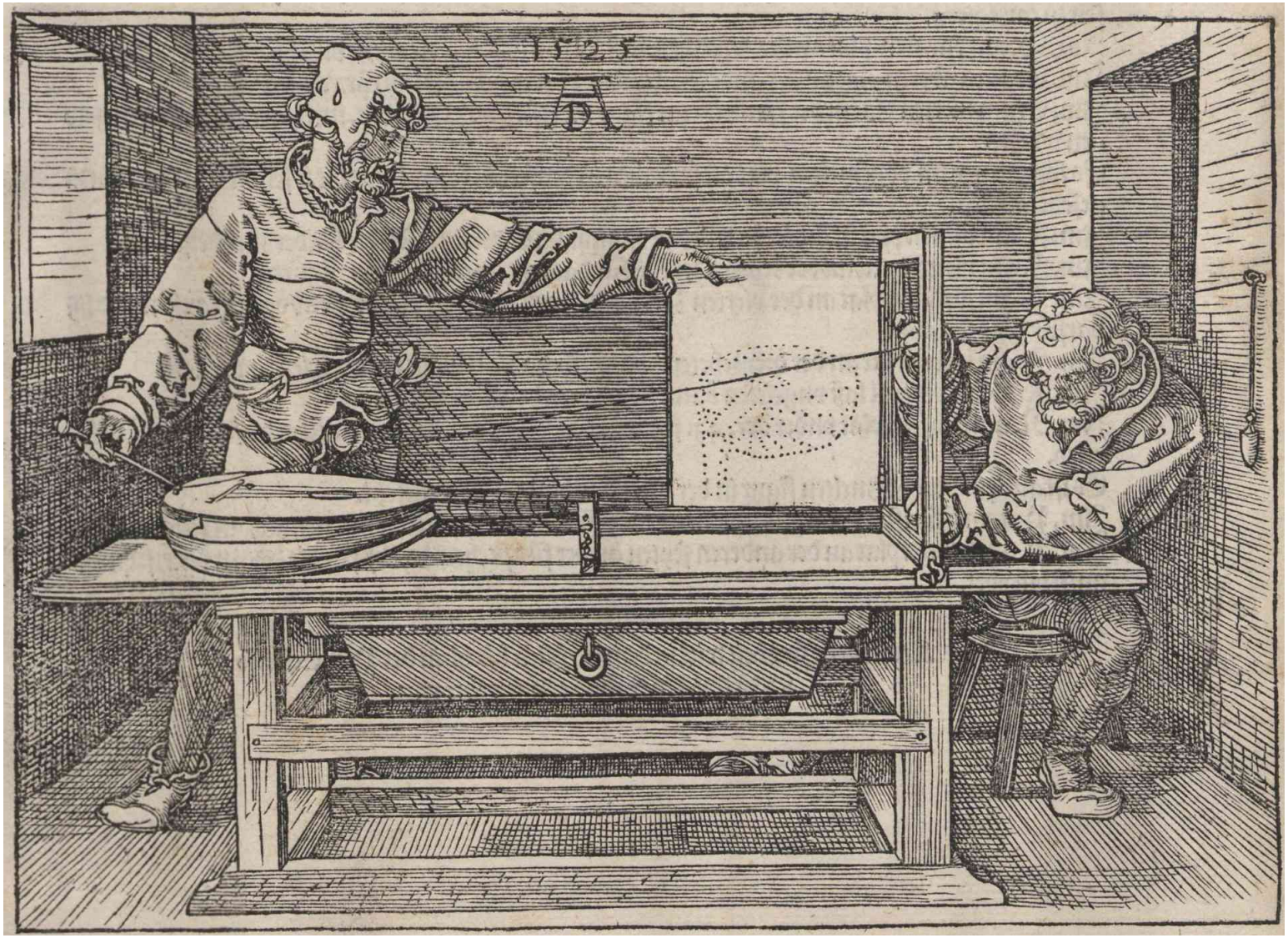
Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Perspective Projection in Art



Dürer's machine (1525)

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Perspective Projection in Art

Satire by Hogarth 1753

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Perspective Projection in Art

M.C. Escher, Another World 1947

Escher, Belvedere 1958

Historic Remarks

Mathematical Representation

Intrinsic Parameters

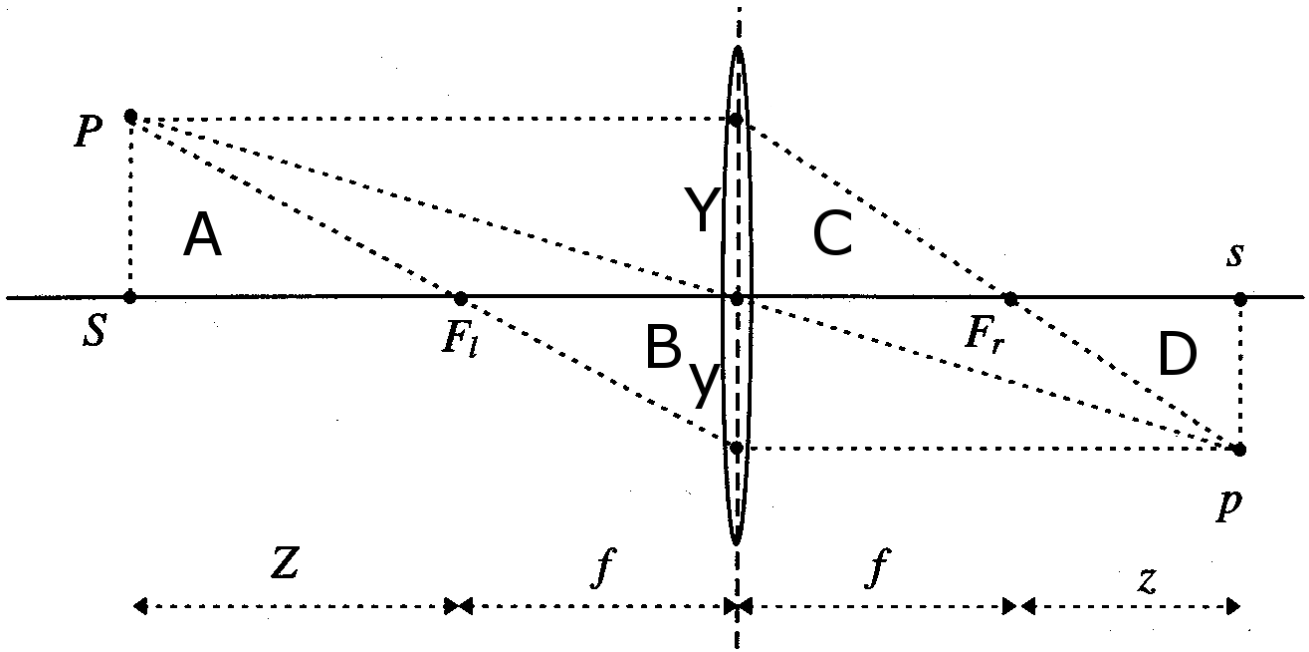
Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Mathematics of Perspective Projection



The above drawing shows the perspective projection of a point P (observed through a thin lens) to its image p.

The point P has coordinates $\boldsymbol{X} = (X, Y, Z) \in \mathbb{R}^3$ relative to the reference frame centered at the optical center, where the z-axis is the optical axis (of the lens).

Historic Remarks

Mathematical Representation

Intrinsic Parameters

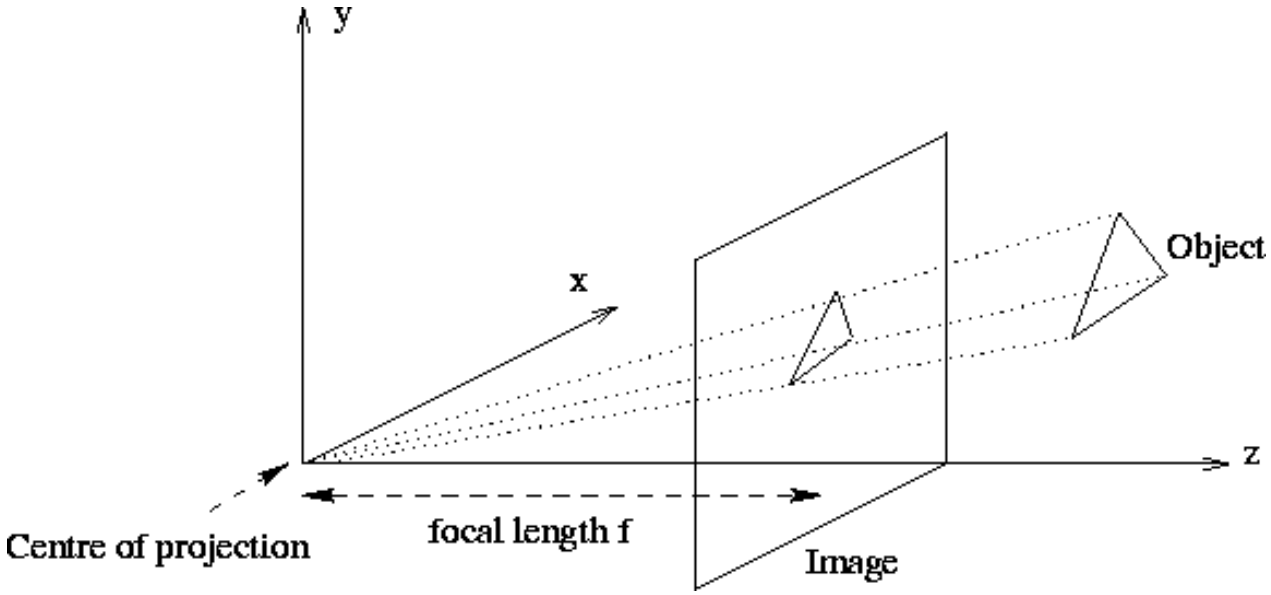
Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Mathematics of Perspective Projection



To simplify equations, one flips the signs of x- and y-axes, which amounts to considering the image plane to be in front of the center of projection (rather than behind it). The perspective transformation π is therefore given by

$$\pi : \mathbb{R}^3 \to \mathbb{R}^2; \quad \boldsymbol{X} \mapsto x = \pi(\boldsymbol{X}) = \begin{pmatrix} f\frac{X}{Z} \\ f\frac{Y}{Z} \end{pmatrix}.$$

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

An Ideal Perspective Camera

In homogeneous coordinates, the perspective transformation is given by:

$$Z\boldsymbol{x} = Z\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = K_f \, \Pi_0 \, \boldsymbol{X}.$$

where we have introduced the two matrices

$$K_f \equiv \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Pi_0 \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The matrix Π_0 is referred to as the standard projection matrix. Assuming Z to be a constant $\lambda > 0$, we obtain:

$$\lambda\boldsymbol{x} = K_f \, \Pi_0 \, \boldsymbol{X}.$$

Historic Remarks
Mathematical Representation
Intrinsic Parameters
Spherical Projection
Radial Distortion
Preimage and Coimage
Projective Geometry

An Ideal Perspective Camera

From the previous lectures, we know that due to the rigid motion of the camera, the point $\boldsymbol{X}$ in camera coordinates is given as a function of the point in world coordinates $\boldsymbol{X}_0$ by:

$$\boldsymbol{X} = R\boldsymbol{X}_0 + T,$$

or in homogeneous coordinates $\boldsymbol{X} = (X, Y, Z, 1)^\top$:

$$\boldsymbol{X} = g\boldsymbol{X}_0 = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \boldsymbol{X}_0.$$

In total, the transformation from world coordinates to image coordinates is therefore given by

$$\lambda \boldsymbol{x} = K_f \, \Pi_0 \, g \, \boldsymbol{X}_0.$$

If the focal length f is known, it can be normalized to 1 (by changing the units of the image coordinates), such that:

$$\boxed{\lambda \, \boldsymbol{x} = \Pi_0 \, \boldsymbol{X} = \Pi_0 \, g \, \boldsymbol{X}_0.}$$

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Intrinsic Camera Parameters

If the camera is not centered at the optical center, we have an additional translation o_x, o_y and if pixel coordinates do not have unit scale, we need to introduce an additional scaling in x- and y-direction by s_x and s_y. If the pixels are not rectangular, we have a skew factor s_θ.
The pixel coordinates $(x', y', 1)$ as a function of homogeneous camera coordinates $\boldsymbol{X}$ are then given by:

$$\lambda \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{pmatrix}}_{\equiv K_s} \underbrace{\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\equiv K_f} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\equiv \Pi_0} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

After the perspective projection Π_0 (with focal length 1), we have an additional transformation which depends on the (intrinsic) camera parameters. This can be expressed by the intrinsic parameter matrix $K = K_s K_f$.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

The Intrinsic Parameter Matrix

All intrinsic camera parameters therefore enter the intrinsic parameter matrix

$$K \equiv K_s K_f = \begin{pmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{pmatrix}.$$

As a function of the world coordinates $\boldsymbol{X}_0$, we therefore have:

$$\boxed{\lambda \boldsymbol{x}' = K \Pi_0 \boldsymbol{X} = K \Pi_0 g \boldsymbol{X}_0 \equiv \Pi \boldsymbol{X}_0.}$$

The 3×4 matrix $\Pi \equiv K \Pi_0 g = (KR, KT)$ is called a general projection matrix.
Although the above equation looks like a linear one, we still have the scale parameter λ. Dividing by λ gives:

$$x' = \frac{\pi_1^\top \boldsymbol{X}_0}{\pi_3^\top \boldsymbol{X}_0}, \quad y' = \frac{\pi_2^\top \boldsymbol{X}_0}{\pi_3^\top \boldsymbol{X}_0}, \quad z' = 1,$$

where $\pi_1^\top, \pi_2^\top, \pi_3^\top \in \mathbb{R}^4$ are the three rows of the projection matrix Π.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

The Intrinsic Parameter Matrix

The entries of the intrinsic parameter matrix

$$K = \begin{pmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{pmatrix},$$

can be interpreted as follows:

o_x:	x-coordinate of principal point in pixels,
o_y:	y-coordinate of principal point in pixels,
$fs_x = \alpha_x$:	size of unit length in horizontal pixels,
$fs_y = \alpha_y$:	size of unit length in vertical pixels,
α_x/α_y:	aspect ratio σ,
fs_θ:	skew of the pixel, often close to zero.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Spherical Perspective Projection

The perspective pinhole camera introduced above considers a planar imaging surface. Instead, one can consider a spherical projection surface given by the unit sphere $\mathbb{S}^2 \equiv \{x \in \mathbb{R}^3 \,|\, |x| = 1\}$. The spherical projection π_s of a 3D point $\boldsymbol{X}$ is given by:

$$\pi_s : \mathbb{R}^3 \;\to\; \mathbb{S}^2; \quad \boldsymbol{X} \mapsto \boldsymbol{x} = \frac{\boldsymbol{X}}{|\boldsymbol{X}|}.$$

The pixel coordinates $\boldsymbol{x}'$ as a function of the world coordinates $\boldsymbol{X}_0$ are:

$$\lambda\, \boldsymbol{x}' = K\, \Pi_0\, g\, \boldsymbol{X}_0,$$

except that the scalar factor is now $\lambda = |\boldsymbol{X}| = \sqrt{X^2 + Y^2 + Z^2}$. One often writes $\boldsymbol{x} \sim \boldsymbol{y}$ for homogeneous vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ if they are equal up to a scalar factor. Then we can write:

$$\boldsymbol{x}' \sim \Pi\, \boldsymbol{X}_0 = K\, \Pi_0\, g\, \boldsymbol{X}_0.$$

This property holds for any imaging surface, as long as the ray between $\boldsymbol{X}$ and the origin intersects the imaging surface.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Radial Distortion

bookshelf with regular lens

bookshelf with short focal lens

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Radial Distortion

The intrinsic parameters in the matrix K model linear distortions in the transformation to pixel coordinates. In practice, however, one can also encounter significant distortions along the radial axis, in particular if a wide field of view is used or if one uses cheaper cameras such as webcams. A simple effective model for such distortions is:

$$x = x_d(1 + a_1 r^2 + a_2 r^4)), \quad y = y_d(1 + a_1 r^2 + a_2 r^4)),$$

where $\boldsymbol{x}_d \equiv (x_d, y_d)$ is the distorted point, $r^2 = x_d^2 + y_d^2$. If a calibration rig is available, the distortion parameters a_1 and a_2 can be estimated.

Alternatively, one can estimate a distortion model directly from the images. A more general model (Devernay and Faugeras 1995) is

$$\boldsymbol{x} = c + f(r)(\boldsymbol{x}_d - c), \quad \text{with } f(r) = 1 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4,$$

Here, $r = |\boldsymbol{x}_d - c|$ is the distance to an arbitrary center of distortion c and the distortion correction factor $f(r)$ is an arbitrary 4-th order expression. Parameters are computed from distortions of straight lines or simultaneously with the 3D reconstruction (Zhang '96, Stein '97, Fitzgibbon '01).

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Preimage of Points and Lines

The perspective transformation introduced above allows to define images for arbitrary geometric entities by simply transforming all points of the entity. However, due to the unknown scale factor, each point is mapped not to a single point $\boldsymbol{x}$, but to an equivalence class of points $\boldsymbol{y} \sim \boldsymbol{x}$. It is therefore useful to study how lines are transformed.

A line L in 3-D is characterized by a base point $\boldsymbol{X}_0 = (X_0, Y_0, Z_0, 1)^\top \in \mathbb{R}^4$ and a vector $\boldsymbol{V} = (V_1, V_2, V_3, 0)^\top \in \mathbb{R}^4$:

$$\boldsymbol{X} = \boldsymbol{X}_0 + \mu \boldsymbol{V}, \quad \mu \in \mathbb{R}.$$

The image of the line L is given by

$$\boldsymbol{x} \sim \Pi_0 \boldsymbol{X} = \Pi_0(\boldsymbol{X}_0 + \mu \boldsymbol{V}) = \Pi_0 \boldsymbol{X}_0 + \mu \Pi_0 \boldsymbol{V}.$$

All points $\boldsymbol{x}$ treated as vectors from the origin o span a 2-D subspace P. The intersection of this plane P with the image plane gives the image of the line. P is called the preimage of the line.

A preimage of a point or a line in the image plane is the largest set of 3D points that give rise to an image equal to the given point or line.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

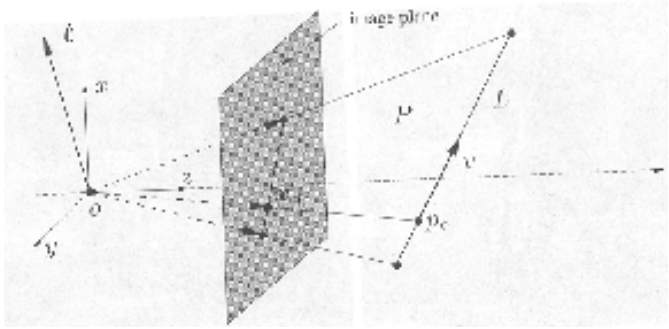
Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Preimage and Coimage



Preimage P of a line L

Preimages can be defined for curves or other more complicated geometric structures. In the case of points and lines, however, the preimage is a subspace of $\mathbb{R}^3$. This subspace can also be represented by its orthogonal complement, i.e. the normal vector in the case of a plane. This complement is called the coimage. The coimage of a point or a line is the subspace in $\mathbb{R}^3$ that is the (unique) orthogonal complement of its preimage. Image, preimage and coimage are equivalent because they uniquely determine oneanother:

$$\text{image} = \text{preimage} \cap \text{image plane}, \qquad \text{preimage} = \text{span}(\text{image}),$$

$$\text{preimage} = \text{coimage}^{\perp}, \qquad \text{coimage} = \text{preimage}^{\perp}.$$

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Preimage and Coimage of Points and Lines

In the case of the line L, the preimage is a $2D$ subspace, characterized by the $1D$ coimage given by the span of its normal vector $\ell \in \mathbb{R}^3$. All points of the preimage, and hence all points $\boldsymbol{x}$ of the image of L are orthogonal to ℓ:

$$\ell^\top \boldsymbol{x} = 0.$$

The space of all vectors orthogonal to ℓ is spanned by the row vectors of $\widehat{\ell}$, thus we have:

$$P = \text{span}(\widehat{\ell}).$$

In the case that $\boldsymbol{x}$ is the image of a point p, the preimage is a line and the coimage is the plane orthogonal to $\boldsymbol{x}$, i.e. it is spanned by the rows of the matrix $\widehat{x}$.

In summary we have the following table:

	Image	Preimage	Coimage
Point	$\text{span}(\boldsymbol{x}) \cap$ im. plane	$\text{span}(\boldsymbol{x}) \subset \mathbb{R}^3$	$\text{span}(\widehat{\boldsymbol{x}}) \subset \mathbb{R}^3$
Line	$\text{span}(\widehat{\ell}) \cap$ im. plane	$\text{span}(\widehat{\ell}) \subset \mathbb{R}^3$	$\text{span}(\ell) \subset \mathbb{R}^3$

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Summary

In this part of the lecture, we studied the perspective projection which takes us from the 3D (4D) camera coordinates to 2D camera image coordinates and pixel coordinates. In homogeneous coordinates, we have the transformations:

4D World coordinates $\xrightarrow{g \in SE(3)}$ 4D Camera coordinates $\xrightarrow{K_f\ \Pi_0}$ 3D image coordinates $\xrightarrow{K_s}$ 3D pixel coordinates.

In particular, we can summarize the (intrinsic) camera parameters in the matrix

$$K = K_s\, K_f.$$

The full transformation from world coordinates $\boldsymbol{X}_0$ to pixel coordinates $\boldsymbol{x}'$ is given by:

$$\lambda \boldsymbol{x}' = K\, \Pi_0\, g\, \boldsymbol{X}_0.$$

Moreover, for the images of points and lines we introduced the notions of preimage (maximal point set which is consistent with a given image) and coimage (its orthogonal complement). Both can be used equivalently to the image.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Projective Geometry

In order to formally write transformations by linear operations, we made extensive use of homogeneous coordinates to represent a $3D$ point as a $4D$-vector $(X, Y, Z, 1)$ with the last coordinate fixed to 1. This normalization is not always necessary: One can represent $3D$ points by a general $4D$ vector

$$\boldsymbol{X} = (XW, YW, ZW, W) \in \mathbb{R}^4,$$

remembering that merely the direction of this vector is of importance. We therefore identify the point in homogeneous coordinates with the line connecting it with the origin. This leads to the definition of projective coordinates.
An n-dimensional projective space $\mathbb{P}^n$ is the set of all one-dimensional subspaces (i.e. lines through the origin) of the vector space $\mathbb{R}^{n+1}$. A point $p \in \mathbb{P}^n$ can then be assigned homogeneous coordinates $\boldsymbol{X} = (x_1, \ldots, x_{n+1})^\top$, among which at least one x is nonzero. For any nonzero $\lambda \in \mathbb{R}$, the coordinates $\boldsymbol{Y} = (\lambda x_1, \ldots, \lambda x_{n+1})^\top$ represent the same point p.

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry

Projective Geometry

If the two coordinate vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ differ by a scalar factor, then they are said to be equivalent:

$$\boldsymbol{X} \sim \boldsymbol{Y}.$$

The point p is represented by the equivalence class of all multiples of $\boldsymbol{X}$. Since all points are represented by lines through the origin, there exist two alternative representations for the two-dimensional projective space $\mathbb{P}^2$:

1. One can represent each point as a point on the 2D-sphere $\mathbb{S}^2$, where any antipodal points represent the same line.
2. One can represent each point p either as a point on the plane of $\mathbb{R}^2$ (homogeneous coordinates) modeling all points with non-zero z-component, or as a point on the circle $\mathbb{S}^1$ (again identifying antipodal points) which is equivalent to $\mathbb{P}^1$.

Both representations hold for the n-dimensional projective space $\mathbb{P}^n$, which can be either seen as a an nD-sphere $\mathbb{S}^n$ or as $\mathbb{R}^n$ with $\mathbb{P}^{n-1}$ attached (to model lines at infinity).

Historic Remarks

Mathematical Representation

Intrinsic Parameters

Spherical Projection

Radial Distortion

Preimage and Coimage

Projective Geometry