Diagnostic Medical Image Processing

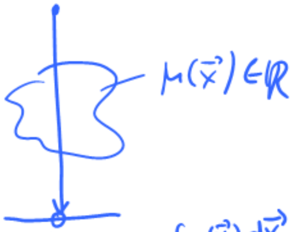
Projection Models and Homogeneous Coordinates

10/11

Joachim Hornegger, Dietrich Paulus, Markus Kowarschik

Lehrstuhl für Mustererkennung (Informatik 5)

Friedrich-Alexander-Universität Erlangen-Nürnberg



$$I = I_0 e^{-\int_L \mu(\vec{x})\, d\vec{x}}$$

Diagnostic Medical Image Processing

Motivation

We have seen so in the previous section that the solution of the reconstruction problem requires detailed knowledge about projection rays.

Thus the questions that we have to consider in detail are:

- How can we characterize the projection rays mathematically?
- How can we characterize different projection geometries?
- What is the mechanical setup for the calibration of projection parameters?
- How can we estimate the camera parameters?
- How can we compute the path of X-rays?
- How reliable are the estimates?

$\vec{p} + \lambda \vec{g}$

Projections

X-ray projection geometry is best modeled by a perspective projection.
All X-ray beams intersect at focal point of the X-ray tube.

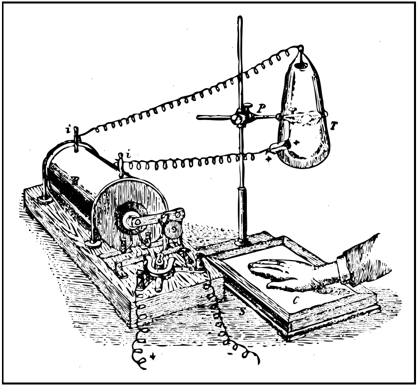


Figure: Conventional Röntgen scheme using photographic paper

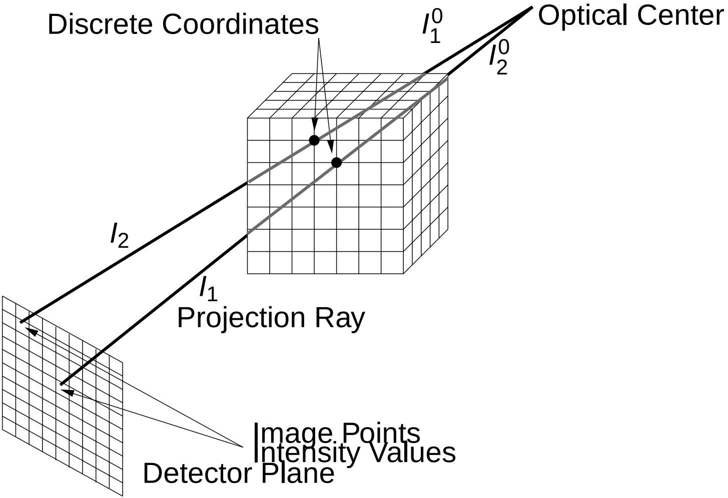


Figure: 3–D/2–D Projection

Diagnostic Medical Image Processing

Projection Geometries

In the following discussion we assume that the image plane is in a fixed position in the 3-D space:

- The 2-D image plane is parallel to the (x, y)-plane of the 3-D coordinate system.
- The distance of the image plane to the origin of the 3-D coordinate system is $z = f$ and constant.

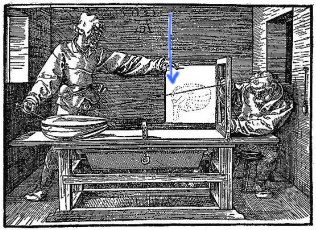


Figure: Illustration of the perspective projection

Parallel / Orthographic projection

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \left[\begin{array}{c} x \\ y \\ \hline 0 \end{array}\right]$$

trivial

Projection Models

In computer vision and graphics different projection models are used:

1. **Orthographic projection:**
   The projected point results from the cancelation of the z components

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

   Obviously this is a linear mapping and can be written in homogeneous coordinates:

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

Projection Models

2. **Weak perspective projection:**
   Weak perspective projection is a scaled orthographic projection, i.e. the coordinates (x, y) are scaled by a scaling factor k.

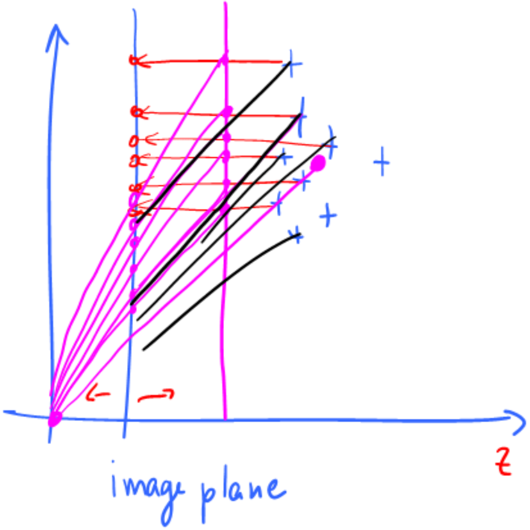
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} k \cdot x \\ k \cdot y \end{pmatrix}$$

   This is again a linear mapping and can be written in homogeneous coordinates:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

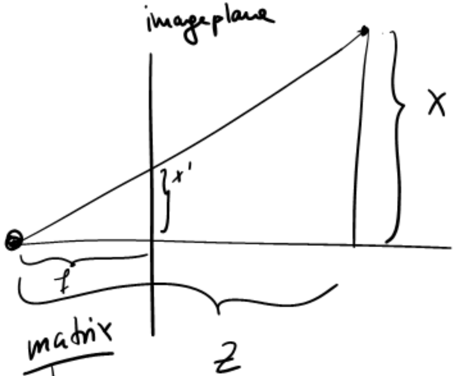
Projection Models

3. **Para-perspective projection:**
   - If multiple points are projected, an auxiliary plane through the points' centroid and parallel to the image plane is defined.
   - Then a parallel projection is applied to map all points onto this auxiliary plane, where the projection direction is parallel to the vector defining the centroid.
   - The points on the auxiliary plane are mapped by perspective projection into the image plane, i.e. we perform a scaled orthographic projection.
   - The para-perspective projection is an affine mapping.



z
image plane

Perspective Projection



$$\frac{x'}{f} = \frac{x}{z}$$

$$x' = \frac{x}{z} \cdot f$$

↑

NONLINEAR in "z" !

matrix

↓

$$M \begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} \frac{x}{z} f \\ \frac{y}{z} \cdot f \end{pmatrix}$$

Projection Models

4. **Perspective projection:**
   The projected point is the intersection of the line connecting point and optical center (focal point) with the detector plane.
   The **nonlinear mapping** of points is given by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} f \cdot x/z \\ f \cdot y/z \end{pmatrix}$$

   where f is the distance of the image plane to the origin.

Observation: The projection model of X-ray systems can be approximated by perspective projection.

Illustration of the Different Projection Models

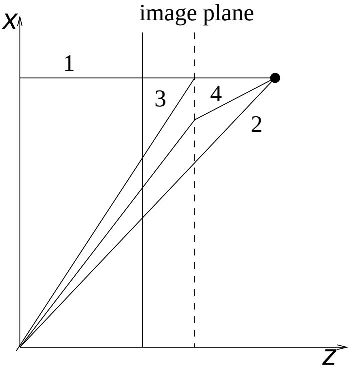


Figure: Projection models in computer vision and graphics: orthographic (1), perspective (2), weak perspective (3), para-perspective (4)

Illustration of the Different Projection Models

Remarks on projection models:

- The projections (1) and (3) can be expressed in terms of a linear mapping in 3-D,
- projection (4) is defined by an affine mapping, and
- the perspective projection (2) is non-linear.

Too bad: Perspective projection model is the model we are required to deal with.

Diagnostic Medical Image Processing

$$\underbrace{\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{Projection matrix}} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f \cdot x \\ f \cdot y \\ z \end{pmatrix} \rightsquigarrow \begin{pmatrix} \frac{fx}{z} \\ \frac{fy}{z} \end{pmatrix}$$

Homogeneous Coordinates

$$\vec{y} = A\vec{x} + \vec{t} = \left(\begin{array}{c|c} A & \vec{t} \\ \hline 0 & 1 \end{array}\right)\begin{pmatrix}\vec{x}\\ 1\end{pmatrix} = \begin{pmatrix}\vec{y}\\ 1\end{pmatrix}$$

By a simple trick, we now extend the vectors by an additional component that allows us to write

- affine as linear mappings, and ✓
- the perspective projection as a linear mapping.

Let us first consider the 2-D case:
We extend $\mathbb{R}^2$ by a third coordinate to create the projective space $\mathbb{P}^2$

Homogeneous Coordinates

Definition

A two-dimensional point in Cartesian coordinates $\boldsymbol{p} = (x, y)^\mathsf{T} \in \mathbb{R}^2$ is represented by $(wx, wy, w)^\mathsf{T} \in \mathbf{P}^2$ in **homogeneous coordinates**, where $w \in \mathbb{R}$ is an arbitrary real valued constant.

Note:

- Homogeneous coordinates have a singularity for $w = 0$.
- A vector $(a, b, c)^\mathsf{T}$ in homogeneous coordinates can be transformed in a 2-D vector by dividing the first two components a and b with the third component $c \neq 0$, i.e.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a/c \\ b/c \end{pmatrix}.$$

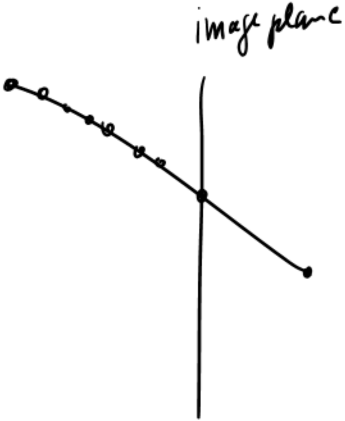


image plane

Homogeneous Coordinates

- A 2-D point $(x, y)^{\mathrm{T}}$ in Cartesian coordinates corresponds to a line in 3-D:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto w \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \text{where} \quad w \in \mathbb{R}.$$

- There exists an infinite number of homogeneous points that correspond to one and the same 2-D point.
- The representation has a singularity for $w \to 0$.

Homogeneous Coordinates

$\tilde{p}$

Definition

A point in homogeneous coordinates is denoted by $\widetilde{\boldsymbol{p}} = (wx, wy, w)^{\mathrm{T}}$.

Definition

The point $\widetilde{\boldsymbol{p}} = (u, v, 0)^{\mathrm{T}}$ is called **ideal point** or **point at infinity**.

Homogeneous Coordinates

$\tilde{p} \sim \tilde{p}'$

We now define an equivalence relation:

Definition

We call two homogeneous points $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}$ equivalent, if $\tilde{\boldsymbol{p}} = \lambda \tilde{\boldsymbol{q}}$ where $\lambda \in \mathbb{R} \backslash \{0\}$. This equivalence is denoted by $\tilde{\boldsymbol{p}} \widetilde{=} \tilde{\boldsymbol{q}}$.

Example

The homogeneous points $\tilde{\boldsymbol{p}} = (2, 3, 1)^{\mathsf{T}}$ and $\tilde{\boldsymbol{q}} = (4, 6, 2)^{\mathsf{T}}$ are equivalent by $\tilde{\boldsymbol{p}} \widetilde{=} \tilde{\boldsymbol{q}}$ as both project to the same point $(2, 3)^{\mathsf{T}} \in \mathbb{R}^2$. They are not equal considered as vectors in $\mathbb{R}^3$, i.e. $\tilde{\boldsymbol{p}} \neq \tilde{\boldsymbol{q}}$.

Homogeneous Coordinates

Let us now consider lines in 2-D. The equation defining a line in $\mathbb{R}^2$ is

Hesse normal form:

$$ax + by + c = 0 \quad . \tag{1}$$

$$\vec{n}^T \vec{x} = d \Leftrightarrow \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = n_1 x_1 + n_2 x_2 - d = 0 \Leftrightarrow (n_1 \; n_2 \; -d) \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = 0$$

- This equation can be multiplied by an arbitrary but non-zero factor $w \in \mathbb{R}$, and it still represents the same line.
- Each vector $(a, b, c) \in \mathbb{R}^3$ represents a line, and

$$ax + by + c = (w \cdot a)x + (w \cdot b)y + (w \cdot c) = 0 \tag{2}$$

  holds for each non-zero w.
- In terms of homogeneous coordinates we can state that each 2-D line can be represented by the vector $\boldsymbol{l} = (a, b, c)^T$.

Homogeneous Coordinates

- A point $\tilde{\boldsymbol{p}}$ (thus represented in homogeneous coordinates) lies on the line $\boldsymbol{l}$ if

$$\boldsymbol{l}^{\mathrm{T}} \cdot \tilde{\boldsymbol{p}} = 0. \tag{3}$$

- Intersection of lines: Two lines $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ intersect in point $\tilde{\boldsymbol{p}}$ if

$$\underline{\boldsymbol{l}_1^{\mathrm{T}} \cdot \tilde{\boldsymbol{p}} = \boldsymbol{l}_2^{\mathrm{T}} \cdot \tilde{\boldsymbol{p}} = 0} \tag{4}$$

EXAM!

and thus

EL: 50

$$\tilde{\boldsymbol{p}} = \boldsymbol{l}_1 \times \boldsymbol{l}_2 \tag{5}$$

Homogeneous Coordinates

Definition

The set of ideal points lie on the line at infinity $l_\infty{}^T = (0, 0, 1)$:

$$(0, 0, 1)(x_1, x_2, 0)^T = 0 \tag{6}$$

Exercise: Do parallel lines intersect in $\mathbb{P}^2$? Where?

The concept of homogeneous coordinates can be transferred to higher dimensional spaces. We will not continue to look into the details of this theory. Interested students are referred to the literature on perspective geometry (see for instance Hartley's book).

Diagnostic Medical Image Processing

Incorporation of rotations and translation

$$\vec{x} \in \mathbb{R}^3$$

$$\vec{x}' = R\vec{x} + \vec{t}$$

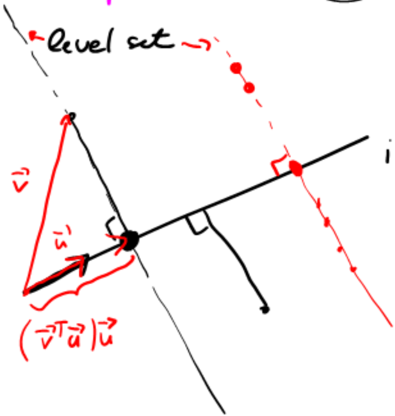
$$\tilde{\vec{x}} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \curvearrowright \begin{pmatrix} \vec{x}' \\ 1 \end{pmatrix} = \tilde{\vec{x}}' = \left(\begin{array}{c|c} R & \vec{t} \\ \hline 0 & 1 \end{array}\right) \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} R\vec{x} + \vec{t} \\ 1 \end{pmatrix}$$

$$R = \left(\vec{r}_1 \ \vec{r}_2 \ \vec{r}_3 \right)$$

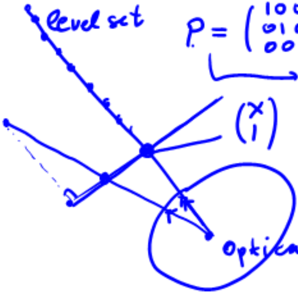
$$R \cdot \vec{x} = x_1 \cdot \vec{r}_1 + x_2 \vec{r}_2 + x_3 \vec{r}_3$$

Perspective, parallel projection



$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{pmatrix} x/1 \\ y/1 \end{pmatrix}$$



$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P\vec{x} = P\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cong \begin{pmatrix} x/z \\ y/z \end{pmatrix}$$

Orthographic Projection

We will now formulate projections from 3-D to 2-D using homogeneous coordinates. The *orthographic projection* in homogeneous coordinates is defined by:

$$\widetilde{\boldsymbol{p}} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{7}$$

This mapping from $\mathbb{P}^3 \to \mathbb{P}^2$ can be simply written in matrix form as

$$\widetilde{\boldsymbol{p}}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \widetilde{\boldsymbol{p}} \quad . \tag{8}$$

Weak Perspective Projection

The *weak perspective projection* in homogeneous coordinates is defined by:

$$\widetilde{\boldsymbol{p}} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} kx \\ ky \\ 1 \end{pmatrix} \tag{9}$$

where $k \in \mathbb{R}$ is a scaling factor.
This mapping from $\mathbb{P}^3 \rightarrow \mathbb{P}^2$ can be simply written in matrix form as:

$$\widetilde{\boldsymbol{p}}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/k \end{pmatrix} \widetilde{\boldsymbol{p}} \quad . \tag{10}$$

Perspective Projection

Using homogeneous coordinates perspective projection becomes a *linear* mapping:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} fx/z \\ fy/z \end{pmatrix} \mapsto \begin{pmatrix} fx \\ fy \\ z \end{pmatrix} \tag{11}$$

We get the following linear mapping from $\mathbb{P}^3 \to \mathbb{P}^2$:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} fx \\ fy \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\boldsymbol{P}} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{12}$$

Extrinsic Camera Parameters

So far we have described the projection of a 3-D point into the image plane. We have not yet considered the motion of the position and orientation of the acquisition device:

- The X-ray source can be translated in 3-D.
- The X-ray source can be rotated in 3-D

Definition

Extrinsic parameters characterize the **pose**, i.e position and orientation of the camera with respect to a world coordinate system. The position is defined by the 3-D translation vector, the orientation by three rotation angles.

Extrinsic Camera Parameters

Figure: C-arm device in different positions and orientations that can be characterized by the extrinsic parameters of the acquisition device.

Extrinsic Camera Parameters

Mathematical characterization:
Rotation and translation of a 3-D point can be expressed by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \boldsymbol{t} \qquad (13)$$

where $\boldsymbol{R} \in \mathbb{R}^{3\times 3}$ denotes a rotation matrix (with its known properties) and $\boldsymbol{t} \in \mathbb{R}^3$ represents the translation in Euclidean space. This is an affine mapping.

Extrinsic Camera Parameters

Using homogeneous coordinates we can rewrite the affine as a linear mapping:

$$\begin{pmatrix} wx' \\ wy' \\ wz' \\ w \end{pmatrix} = \boldsymbol{D} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \left(\begin{array}{ccc|c} & \boldsymbol{R} & & \boldsymbol{t} \\ \hline 0 & 0 & 0 & 1 \end{array} \right) \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{14}$$

Problem: How does the rotation matrix look like?
Solution: As we know already, the columns of the linear mapping are the images of the base vectors of the original coordinate system.

Extrinsic Camera Parameters

Example

Rotation around the *x*-axis
We have the following images of base vectors that define the columns of the rotation matrix:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{15}$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ \cos\varphi \\ \sin\varphi \end{pmatrix} \tag{16}$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ -\sin\varphi \\ \cos\varphi \end{pmatrix} \tag{17}$$

Extrinsic Camera Parameters

Example

Thus the rotation around the x-axis is defined by the matrix:

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \tag{18}$$

Exercise: Compute the composed rotation matrix that includes rotations first around the x–, then y–, and finally around the z–axis.

Intrinsic Camera Parameters

Besides the position and orientation of the acquisition device, we do have also a mapping of the projected points in the ideal image plane to the used detector.

Definition

Intrinsic parameters define the mapping of 2-D coordinates from the ideal image plane to the 2-D detector coordinates.

Intrinsic Camera Parameters

- Intrinsic parameters do not change (usually), if the camera moves.
- Origin of the detector coordinate system does not coincide with the intersection of optical axis and the ideal image plane in general.
- The coordinate axes of the detector are not necessarily orthogonal, but intersect with a skew angle Θ.
- The pixels in the detector coordinate system are not necessarily squared pixels, but scaled by k_x and k_y.
- There might exist a radial distortion due to the camera optics (not considered here).

Intrinsic Camera Parameters

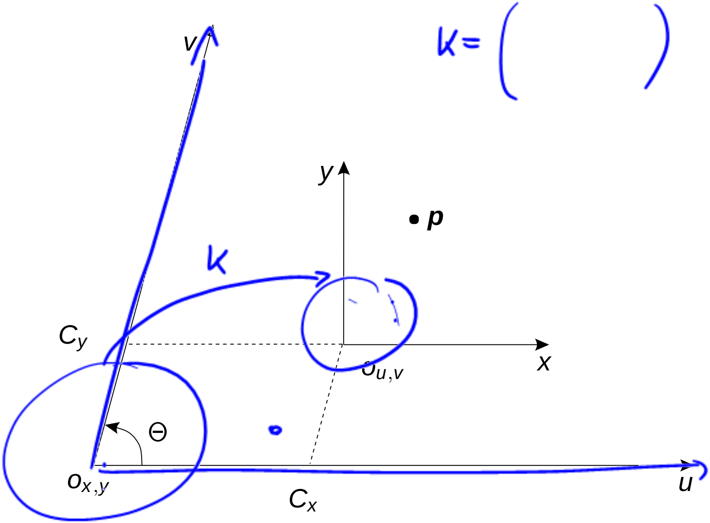


Figure: (u, v) detector and (x, y) image coordinate system

Intrinsic Camera Parameters

1. (x, y) ideal image coordinate system:
   - ideal coordinate system used so far in all formulas
   - ideal image coordinate system with origin $o_{x,y}$
2. (u, v) detector coordinate system:
   - real image matrix of measurements
   - Θ: skew angle between axes,
   - k_x, k_y: scaling of u and v axis with respect to units in (x, y) system
   - (C_x, C_y): offset of origins of coordinate systems.

Intrinsic Camera Parameters

Transformation between u/v– and x/y–coordinate system
At first, we consider the images of base vectors of the detector coordinate system in the image coordinate system:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{k_x} \\ 0 \end{pmatrix} \tag{19}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{k_y} \cos\theta \\ \frac{1}{k_y} \sin\theta \end{pmatrix} \tag{20}$$

The required transform from the (x, y) to the (u, v) coordinate system is given by the inverse of the above matrix:

$$\boldsymbol{T} = \begin{pmatrix} \frac{1}{k_x} & \frac{1}{k_y} \cos\Theta \\ 0 & \frac{1}{k_y} \sin\Theta \end{pmatrix}^{-1} = \begin{pmatrix} k_x & -k_x \frac{\cos\Theta}{\sin\Theta} \\ 0 & \frac{k_y}{\sin\Theta} \end{pmatrix} \tag{21}$$

Intrinsic Camera Parameters

The complete mapping of (x, y)- to (u, v)-coordinates in Euclidean space is thus given by:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} k_x & -k_x \frac{\cos\Theta}{\sin\Theta} \\ 0 & \frac{k_y}{\sin\Theta} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} C_x \\ C_y \end{pmatrix} \qquad (22)$$

Using homogeneous coordinates we get the matrix that maps the ideal image coordinates to detector coordinates that includes the intrinsic parameters:

$$\boxed{\boldsymbol{K} = \left(\begin{array}{cc|c} \multicolumn{2}{c}{\multirow{2}{*}{$\boldsymbol{T}$}} & -C_\mathrm{x} \\ & & -C_\mathrm{y} \\ 0 & 0 & 1 \end{array} \right) .} \qquad (23)$$

intrinsic parameters

The complete projection reads as follows:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{c|c} R & \vec{t} \\ \hline 0 & 1 \end{array} \right) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$\uparrow$

3x4 projection matrix

Constraint:

Seaman's way: $p_{34} = 1$

Alternative: $\|P\|_F = 1$

$$= \sum_{ij} p_{ij}^2 = 1$$

Modalities

$$-\ln \frac{I}{I_0} = \int_L \mu(\vec{x})\, d\vec{x}$$

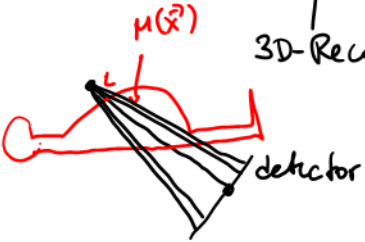
DMIP

Pre-Processing

- MR: Bias Correct.
- $KL(p\|q) = \sum_x p \log \frac{p}{q}$

3D-Reconstruction

X-Ray



$$I = I_0\, e^{-\int_L \mu(\vec{x})\, d\vec{x}}$$

$P \in \mathbb{R}^{3 \times 4}$

$$P \tilde{\vec{X}} = \tilde{\vec{x}}' \qquad \text{where} \quad \vec{X} \in \mathbb{R}^3, \ \vec{x}' \in \mathbb{R}^2$$

$$P = K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{c|c} R & \vec{t} \\ \hline 0 & 1 \end{array}\right) \tilde{\vec{X}}$$

$$= \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & \boxed{p_{34}} \end{pmatrix}$$

1) Seaman's approach:

$$p_{34} = 1 = 1$$

2) Poor Man's approach:

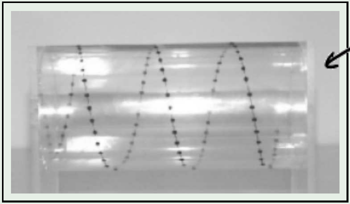
$$\|P\|_F = \sum_{ij} p_{ij}^2 = 1$$

Intrinsic Camera Parameters

Example

Let us assume we have a so-called calibration pattern that is manufactured precisely and the centroids of 3-D metal spheres is known in world coordinates:

$$X = \{\boldsymbol{x}_i = (x_{i,0}, x_{i,1}, x_{i,2})^{\mathsf{T}}; i = 1, 2, \ldots, N\}$$



Calibration Pattern 1994, GE

Figure: Calibration pattern used for the computation of the projection geometry of a C-arm

Calibration Pattern

Example

A few remarks on the calibration pattern that is used for today's C-arm systems:

- 108 steel spheres
- large spheres represent a logical one
- small spheres represent a logical zero
- 8 Bit binary encoding
- position of spheres is known in the World Coordinate System
- World Coordinate System is attached to the phantom

Let us estimate the projection matrix P:

Given: $\vec{x}_i \in \mathbb{R}^3$: points in the calibration pattern

$\vec{x}_i' \in \mathbb{R}^2$: corresponding points in the image

$$\forall i: \quad \boxed{P\tilde{\vec{x}}_i \simeq \tilde{\vec{x}}_i'}$$

$$\sum_i \| P\tilde{\vec{x}}_i - \tilde{\vec{x}}_i' \|^2 \longrightarrow \min$$

"Martin's idea":

$$P = \begin{pmatrix} \vec{p}_1^{\,T} \\ \vec{p}_2^{\,T} \\ \vec{p}_3^{\,T} \end{pmatrix}$$

id.

$$(\mathrm{I}) \quad \frac{\tilde{\vec{x}}_{i,1}'}{\tilde{\vec{x}}_{i,3}'} = \frac{\vec{p}_1^{\,T} \cdot \tilde{\vec{x}}_i}{\vec{p}_3^{\,T} \tilde{\vec{x}}_i}$$

$$(\mathrm{II}) \quad \frac{\tilde{\vec{x}}_{i,2}'}{\tilde{\vec{x}}_{i,3}'} = \frac{\vec{p}_2^{\,T} \tilde{\vec{x}}_i}{\vec{p}_3^{\,T} \tilde{\vec{x}}_i}$$

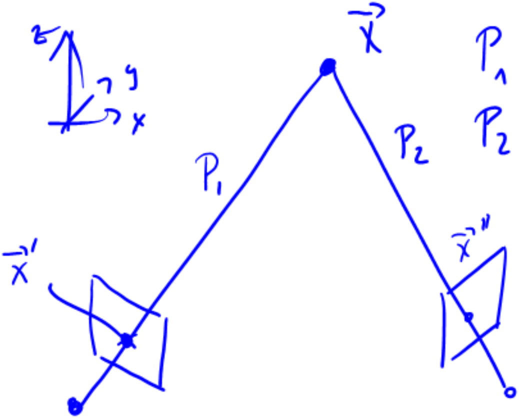
$$(\mathrm{I})\quad \frac{\tilde{\vec{x}}'_{i,1}}{\tilde{\vec{x}}'_{i,3}} = \frac{\vec{p}_1^{\,T} \cdot \tilde{\vec{x}}_i}{\vec{p}_3^{\,T}\, \tilde{\vec{x}}_i} \quad \Big|\ \vec{p}_3^{\,T}\tilde{\vec{x}}_i$$

$$(\mathrm{II})\quad \frac{\tilde{\vec{x}}'_{i,2}}{\tilde{\vec{x}}'_{i,3}} = \frac{\vec{p}_2^{\,T}\tilde{\vec{x}}_i}{\vec{p}_3^{\,T}\tilde{\vec{x}}_i}$$

$$\Rightarrow\quad \frac{\tilde{\vec{x}}'_{i,1}}{\tilde{\vec{x}}'_{i,3}} \cdot \vec{p}_3^{\,T}\tilde{\vec{x}}_i = \underbrace{\vec{p}_1^{\,T}\tilde{\vec{x}}_i}_{(p_{11},\, p_{12},\, p_{13},\, p_{14})^T}$$

is LINEAR in

$$\Rightarrow\quad M \cdot \begin{pmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{34} \end{pmatrix} \equiv \vec{0} \iff M \cdot \vec{p} \equiv \vec{0}$$



$$P_1 \tilde{\vec{X}} = \tilde{\vec{x}}'$$

$$P_2 \tilde{\vec{X}} = \tilde{\vec{x}}''$$

Intrinsic Camera Parameters

Example

Now we further assume that the extrinsic parameters are known (no rotation, no translation) and that we have to estimate the source-detector-distance f of the perspective projection model and the intrinsic parameters of the detector.

$$\tag{24}$$

Intrinsic Camera Parameters

The list of unknown parameters is as follows

- Coordinates of principal point $\boldsymbol{C}$
- Width d_x and height d_y of each sensor element
- Focal length f, or $k_\mathrm{x} = f/d_\mathrm{x}$, $k_\mathrm{y} = f/d_\mathrm{y}$.
- Position of the camera (expressed by $\boldsymbol{R}$ and $\boldsymbol{t}$)
- Skew $s = -k_\mathrm{x} \tan \Theta$

Complete Projection

The total perspective transformation is

$$\rho \widetilde{\boldsymbol{q}} = \boldsymbol{P} \cdot \widetilde{\boldsymbol{p}} = \boldsymbol{K} \cdot \boldsymbol{P}_{\text{proj}} \cdot \boldsymbol{D} \cdot \widetilde{\boldsymbol{p}} \quad . \tag{25}$$

where

- $\boldsymbol{D}$: extrinsic camera parameters
  - position and orientation of camera with respect to the world coordinate system
- $\boldsymbol{P}_{\text{proj}}$: projection model matrix, perspective projection
- $\boldsymbol{K}$: intrinsic camera parameters.
  - optical and geometric characteristics of the camera,
  - do *not* change with camera movement.

Calibration of X-ray Systems

Definition

The estimation of the projection parameters is called *calibration*.

Let us first consider a simplified case for calibration, where we consider the following problems in detail:

- Design a calibration pattern where the 3–D geometry is known exactly.
- Capture a X–ray image of the calibration pattern.
- Compute the corresponding 2–D/3–D points.
- Compute f using a least square estimator.

Calibration Patterns

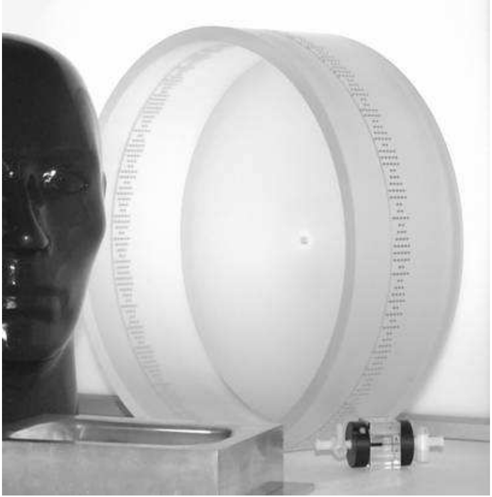


Figure: Calibration Pattern used for C-Arm Calibration

Calibration Patterns

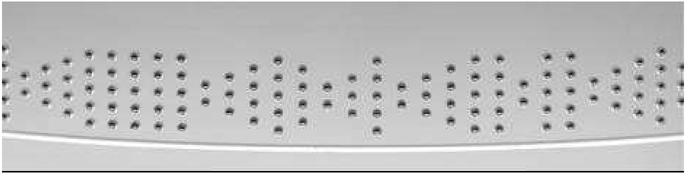


Figure: Observed 2–D Point Features

Calibration Patterns

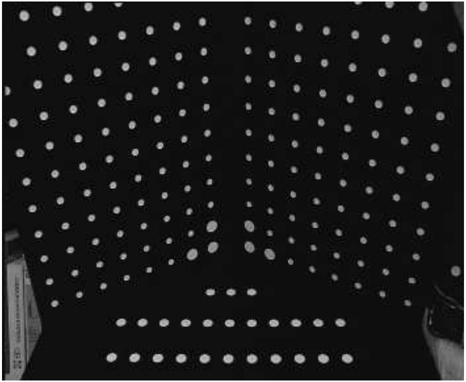


Figure: Calibration object with ground truth 3D points.

Calibration of X-ray Systems

Let $(p_{i,0}, p_{i,1}, p_{i,2})^{\mathsf{T}}$ $(i = 1, \ldots, N)$ be the set of 3–D points of the calibration pattern, and let $(q_{i,0}, q_{i,1})^{\mathsf{T}}$ be the set of 2-D points. The projection is defined by:

$$\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_{i,0} \\ p_{i,1} \\ p_{i,2} \\ 1 \end{pmatrix} \stackrel{\sim}{=} \begin{pmatrix} q_{i,0} \\ q_{i,1} \\ 1 \end{pmatrix} \tag{26}$$

Calibration of X-ray Systems

For each point we get the pair of equations:

$$\frac{fp_{i,0}}{p_{i,2}} = q_{i,0} \tag{27}$$

$$\frac{fp_{i,1}}{p_{i,2}} = q_{i,1} \tag{28}$$

Due to image noise and segmentation errors, the equations have to be replaced by the following minimization problem:

$$\widehat{f} = \operatorname{argmin}_f \sum_{i=1}^{N} \sum_{k=0}^{1} \left(\frac{fp_{i,k}}{p_{i,2}} - q_{i,k} \right)^2 \tag{29}$$

Calibration of X-ray Systems

We compute the zero crossings of the partial derivative regarding the unknown parameter f and get the following estimator for f:

$$\widehat{f} = \frac{\sum_{i,k} \frac{q_{i,k} p_{i,k}}{p_{i,2}}}{\sum_{i,k} \left(\frac{p_{i,k}}{p_{i,2}} \right)^2} \tag{30}$$

Different Objective Functions

- We can multiply (27) and (28) with the denominator and get:

$$fp_{i,0} = p_{i,2}q_{i,0} \tag{31}$$

$$fp_{i,1} = p_{i,2}q_{i,1} \tag{32}$$

The resulting objective function is:

$$\widehat{f} = \operatorname{argmin}_F \sum_{i=1}^{N} \sum_{k=0}^{1} \left(fp_{i,k} - p_{i,2}q_{i,k}\right)^2 \tag{33}$$

Different Objective Functions

- There is another option to rewrite (27) and (28):

$$f = \frac{p_{i,2}q_{i,0}}{p_{i,0}} \tag{34}$$

$$f = \frac{p_{i,2}q_{i,1}}{p_{i,1}} \tag{35}$$

Using this identity, the resulting objective function is:

$$\widehat{f} = \operatorname{argmin}_f \sum_{i=1}^{N} \sum_{k=0}^{1} \left(f - \frac{p_{i,2}q_{i,k}}{p_{i,k}} \right)^2 \tag{36}$$

Different Objective Functions

Question: Which objective function is the best one? Should we optimize (29), (33) or (36)?

Exercise: Compute the estimators for (29), (33), and (36) and analyze the sensitivity of estimates by adding noise to input data.

Rule of thumb: Always optimize differences in the image space resp. in space of observations.

Camera Calibration

Let

- $X = \{\boldsymbol{x}_i = (x_{i,0}, x_{i,1}, x_{i,2})^{\mathrm{T}}; i = 1, \ldots, N\}$ be the set of 3–D points of the calibration pattern, and
- $Y = \{\boldsymbol{y}_i = (y_{i,0}, y_{i,1})^{\mathrm{T}}; i = 1, \ldots, N\}$ be the set of 2-D observations.

Looking at the set of all corresponding points $\{(\boldsymbol{x}_i, \boldsymbol{y}_i); i = 1, \ldots, N\}$, we get the N homogeneous equations:

$$\boldsymbol{P} \cdot \tilde{\boldsymbol{x}}_i = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} \end{pmatrix} \begin{pmatrix} x_{i,0} \\ x_{i,1} \\ x_{i,2} \\ 1 \end{pmatrix} \hat{=} \begin{pmatrix} y_{i,0} \\ y_{i,1} \\ 1 \end{pmatrix} \quad (37)$$

where $i = 1, \ldots, N$ and the unknowns are the components of the projection matrices.

Camera Calibration

Using the definition of homogeneous coordinates, we get the following $2N$ equations

$$\frac{p_{1,1}x_{i,0} + p_{1,2}x_{i,1} + p_{1,3}x_{i,2} + p_{1,4}}{p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4}} = y_{i,0} \qquad (38)$$

$$\frac{p_{2,1}x_{i,0} + p_{2,2}x_{i,1} + p_{2,3}x_{i,2} + p_{2,4}}{p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4}} = y_{i,1} \qquad (39)$$

that are nonlinear in the components of the projection matrix $\boldsymbol{P}$.

Note: The points in the image plane are computed by applying segmentation methods on real images. Segmentation errors and noise will be present, and the equations will not be fulfilled exactly.

Camera Calibration

We apply the idea of least-square-estimation, and estimate the projection matrix according to the non-linear optimization problem:

$$\hat{\boldsymbol{P}} = \underset{\boldsymbol{P}}{\operatorname{argmin}} \sum_{i=1}^{N} \left(\frac{p_{1,1}x_{i,0} + p_{1,2}x_{i,1} + p_{1,3}x_{i,2} + p_{1,4}}{p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4}} - y_{i,0} \right)^2 + \sum_{i=1}^{N} \left(\frac{p_{2,1}x_{i,0} + p_{2,2}x_{i,1} + p_{2,3}x_{i,2} + p_{2,4}}{p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4}} - y_{i,1} \right)^2$$

This nonlinear optimization problem is hard to solve. Numerial optimization requires a good initializion usually.

c

Camera Calibration

A linear method to estimate the projection matrix results from the multiplication of the equations with the denominator:

$$
\begin{aligned}
p_{1,1}x_{i,0} + p_{1,2}x_{i,1} + p_{1,3}x_{i,2} + p_{1,4} = \\
= (p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4})y_{i,0} \qquad (40)\\
p_{2,1}x_{i,0} + p_{2,2}x_{i,1} + p_{2,3}x_{i,2} + p_{2,4} = \\
= (p_{3,1}x_{i,0} + p_{3,2}x_{i,1} + p_{3,3}x_{i,2} + p_{3,4})y_{i,1} \qquad (41)
\end{aligned}
$$

Observations:

- The above equations are linear in the components of the projection matrix $\boldsymbol{P}$.
- The above equations can be rewritten in matrix form, where the measurement matrix $\boldsymbol{M}$ will include the information on the 3-D calibration points and the measured 2-D points according the above equations.

Camera Calibration

Camera calibration thus reduces to the nullspace computation of the measurement matrix $\boldsymbol{M}$:

$$\boldsymbol{M} \begin{pmatrix} p_{1,1} \\ p_{1,2} \\ \vdots \\ p_{3,3} \\ p_{3,4} \end{pmatrix} = \mathbf{0} \rightarrow \hat{p} \tag{42}$$

where

$$\boldsymbol{M} = \begin{pmatrix} x_{1,0} & x_{1,1} & x_{1,2} & 1 & 0 & 0 & 0 & 0 & x_{1,0}y_{1,0} & x_{1,1}y_{1,0} & x_{1,2}y_{1,0} & y_{1,0} \\ \vdots & & & & & \ddots & & & & & & \vdots \\ x_{N,0} & x_{N,1} & x_{N,2} & 1 & 0 & 0 & 0 & 0 & x_{N,0}y_{N,0} & x_{N,1}y_{N,0} & x_{N,2}y_{N,0} & y_{N,0} \\ 0 & 0 & 0 & 0 & x_{1,0} & x_{1,1} & x_{1,2} & 1 & x_{1,0}y_{1,1} & x_{1,1}y_{1,1} & x_{1,2}y_{1,1} & y_{1,1} \\ \vdots & & & & & \ddots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & x_{N,0} & x_{N,1} & x_{N,2} & 1 & x_{N,0}y_{N,1} & x_{N,1}y_{N,1} & x_{N,2}y_{N,1} & y_{N,1} \end{pmatrix}$$

Camera Calibration

Observations:

- The calibration problem is reduced to the computation of the nullspace of the measurement matrix $\boldsymbol{M}$.
- We know how to compute the nullspace of $\boldsymbol{M}$ using SVD.
- The rank of $\boldsymbol{M}$ is 11.

Camera Calibration

$(M\vec{p})^T(M\vec{p}) = \vec{p}^T M^T M \vec{p}$

The estimation problem can also be reduced to an eigenvalue/eigenvector problem:

$$||\boldsymbol{Mp}||^2 \quad \rightarrow \quad \text{min}, \quad \text{subject to} \quad ||\boldsymbol{p}||^2 = 1. \tag{43}$$

$\vec{p}^T \vec{p}' - 1 = 0$

Lagrange multiplier method:

$$\boldsymbol{p}^T\boldsymbol{M}^T\boldsymbol{Mp} - \lambda(\boldsymbol{p}^T\boldsymbol{p} - 1) \quad \rightarrow \quad \text{min} \tag{44}$$

Lagrangian

Camera Calibration

Now we compute the zero crossing of the gradient regarding the components of $\boldsymbol{P}$:

$$2\boldsymbol{M}^{\mathsf{T}}\boldsymbol{M}\boldsymbol{p} - 2\lambda\boldsymbol{p} = \boldsymbol{0} \tag{45}$$

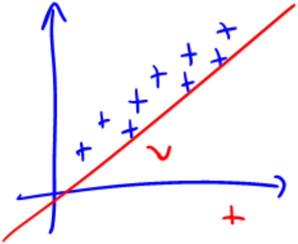
and thus we obtain:

$$\boldsymbol{M}^{\mathsf{T}}\boldsymbol{M}\boldsymbol{p} = \lambda\boldsymbol{p} \tag{46}$$

Conclusions:

- The components of the projection matrix $\boldsymbol{P}$ result from the eigenvector belonging to the smallest eigenvalue.
- The linear estimate of $\boldsymbol{P}$ is an excellent initialization for the nonlinear least square estimate of the projection matrix.

1 Projection Models and Homogeneous Coordinates



RANSAC — RANdom SAmple Consensus

Problem in calibration: inaccuracies in observations and outliers.
There are 2 types of outliers:

- badly localized points,
- wrong correspondences.

Outliers in Linear Regression

$\vec{n}^T\vec{x} - d = 0$

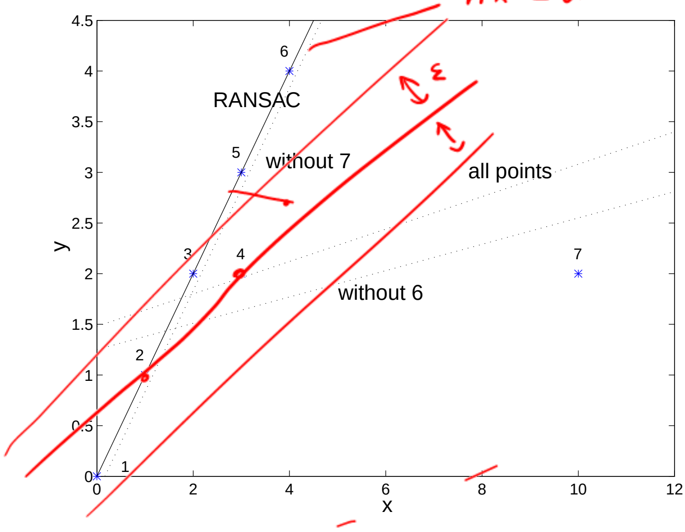


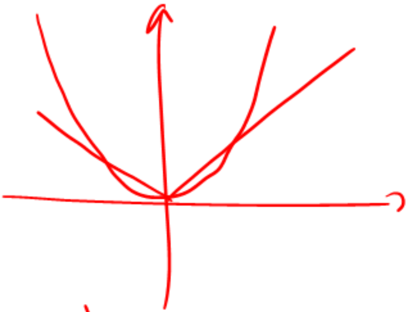
Figure: Example: Influence of an outlier in linear regression (least-squares method)

RANSAC Algorithm

1. draw samples uniformly and at random from the input data set
2. cardinality of sample set:
   smallest size sufficient to estimate the model parameters
3. compute the model parameters for each element the sample data
4. evaluate the quality of the hypothetical models on the full data set
5. cost function for the evaluation of the quality of the model
6. inliers: data points which agree with the model within an error tolerance
7. The hypothesis which gets the most support from the data set: best estimate.

$$\sum_i (mx_i + t - y_i)^2$$

$$\sum |mx_i + t - y_i|$$



Diagnostic Medical Image Processing

Bootstrapping

Problem: Set of sample points is given and no additional points can be acquired. How can we estimate the bias and the variance?
Solution: A bootstrap data set is generated by randomly selecting a subset of the available data. This selection process is repeated independently *n*–times, and we get a sequence of estimates:

$$\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \dots, \widehat{\theta}_n \tag{47}$$

Bootstrap

Bootstrap mean estimate

$$\bar{\widehat{\theta}} = \frac{1}{n}\sum_{i=1}^{n}\widehat{\theta}_i \tag{48}$$

Bootstrap variance estimate

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\widehat{\theta}_i - \bar{\widehat{\theta}})^2 \tag{49}$$

Bootstrap bias estimate

$$b = \bar{\widehat{\theta}} - \theta^* \tag{50}$$

Sampling Density

Point estimation method:

- given measurement matrix $\boldsymbol{W}$
- objective function $h(\boldsymbol{W};\theta)$
- estimation of the parameter vector:

$$\widehat{\theta} \quad = \quad \operatorname{argmin}_{\theta} h(\boldsymbol{W};\theta) \tag{51}$$

Observation: Repeated experiments will lead to different values of $\boldsymbol{W}$ and thus different estimates for θ. The estimates can be considered as random variables and thus underlie a density function — the *sampling density*.

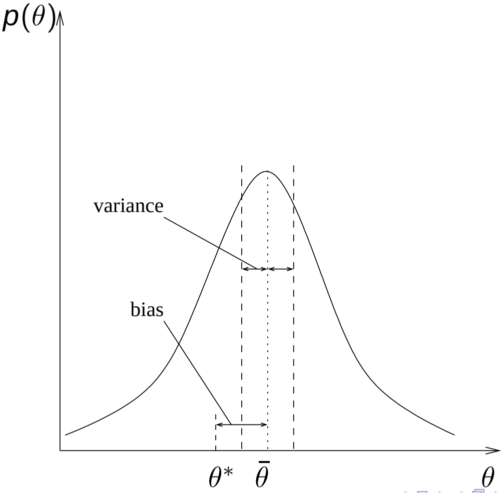


Figure: Sampling density $p(\theta)$ of parameter θ and the bias/variance, where $\bar{\theta}$

Bias

- **bias of an estimator:** difference between the expected value and the true value of the parameter:

$$b = \bar{\theta} - \theta^* \tag{52}$$

- **unbiased estimator:** bias vanishes, i. e.

$$\bar{\theta} = \theta^* \tag{53}$$

- totally unbiasedness is mostly unobtainable
- if an estimator is unbiased for some parameter θ it is generally biased for nontrivial functions of θ. Even if $E(\widehat{\theta}) = \theta^*$ it is not necessarily true that $E(\widehat{\theta}^2) = \theta^{*2}$.
- bias is a measure of the systematic error of an estimator

Variance

- **variance of an estimator:** measures the randomness of the estimate
- **Cramer–Rao bound:** theoretical lower bound on the attainable variance of an estimator
- **Efficiency:** An estimate is called efficient if its variance is the lowest theoretically attainable.

Bias–Variance–Tradeoff

$$
\begin{aligned}
& E((\theta^* - \widehat{\theta})(\theta^* - \widehat{\theta})^T) \\
= \; & E(((\theta^* - \bar{\theta}) - (\widehat{\theta} - \bar{\theta}))((\theta^* - \bar{\theta}) - (\widehat{\theta} - \bar{\theta}))^T) \\
= \; & E((\theta^* - \bar{\theta})(\theta^* - \bar{\theta})^T) + E((\widehat{\theta} - \bar{\theta})(\widehat{\theta} - \bar{\theta})^T) - \underbrace{2 \cdot E((\theta^* - \bar{\theta})(\widehat{\theta} - \bar{\theta})^T)}_{=0} \\
= \; & \underbrace{E((\theta^* - \bar{\theta})(\theta^* - \bar{\theta})^T)}_{\text{bias}^2} + \underbrace{E((\widehat{\theta} - \bar{\theta})(\widehat{\theta} - \bar{\theta})^T)}_{\text{variance}}
\end{aligned}
$$

Diagnostic Medical Image Processing

Take Home Messages

- Projection models
- Homogeneous Coordinates
- Projection matrices
- Camera calibration using least-squares
- RANSAC
- Bias-variance-tradeoff

Further Readings

- R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, Cambridge, 2004.
- O. Faugeras. Three-Dimensional Computer Vision: a Geometric Viewpoint. MIT Press, 1993.