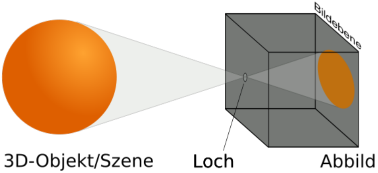
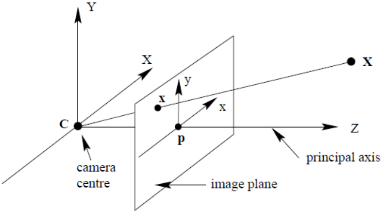
Bildverarbeitung: Epipolargeometrie

D. Schlesinger – TUD/INF/KI/IS

Lochkamera Modell





C – Projektionszentrum, Optische Achse, Bildebene,
P – Hauptpunkt (optische Achse kreuzt die Bildebene), x – Bildpunkt, X – Weltpunkt

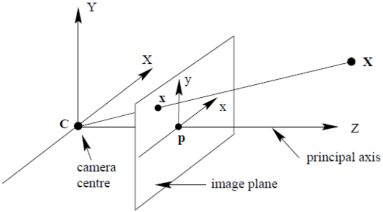
3D-Koordinatensystem: Ursrung im Kamerazentrum, Z-Achse – optische Achse

Weltpunkt $X = [X_1, X_2, Z]$
x ist ein Punkt im 3D, d.h. $x = [x_1, x_2, f]$ mit Brennweite f

Projektive Abbildung: $$x_1 = X_1 \cdot f/Z$$ $$x_2 = X_2 \cdot f/Z$$

Nicht eineindeutig – einem x entsprechen alle X, die auf dem Strahl $\lambda(x - C) = \lambda x$ liegen.

Homogene Koordinaten



Ein „Vektor" $h \in \mathbb{R}^3$, d.h. $h = [h_1, h_2, h_3]$ beschreibt die Menge der Strahlen im $\mathbb{R}^3$ mit $X = \lambda \cdot [h_1/h_3, h_2/h_3, 1]$

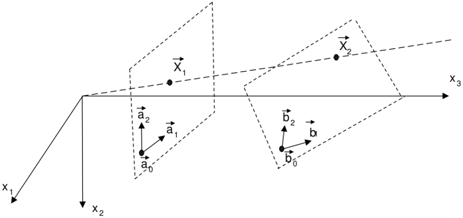
Topologisch gesehen ist diese Menge zweidimensional – redundante Beschreibung.

Die Mende der Weltpunkte im $\mathbb{R}^3$ wird auf Teilmengen partitioniert.
Zwei Punkte X und X' gehöhren einer Teilmenge (Äquivalenzklasse, Strahl), wenn $X_1/Z = X'_1/Z'$ und $X_2/Z = X'_2/Z'$

In homogenen Koordinaten gibt es keine projektive Abbildung mehr, denn X und x sind bereits in einer Äquivalenzklasse – ein Strahl.

Abbildung „Homogene Koordinaten → Bildkoordinaten": $\begin{array}{l} x_1 = h_1 \cdot f/h_3 \\ x_2 = h_2 \cdot f/h_3 \end{array}$

Wie wird eine Ebene im $\mathbb{R}^3$ auf eine andere Ebene im $\mathbb{R}^3$ (z.B. Bildschirm) Abgebildet?



Sei $[\lambda_1, \lambda_2]$ Koordinaten eines Punktes im Koordinatensystem der abzubildenden Ebene.

Die Lage der abzubildenden Ebene im $\mathbb{R}^3$ wird durch: $b_0 \in \mathbb{R}^3$ (ein Punkt in der Ebene) und $b_1, b_2 \in \mathbb{R}^3$ (Basis, der die Ebene aufspannt) angegeben.

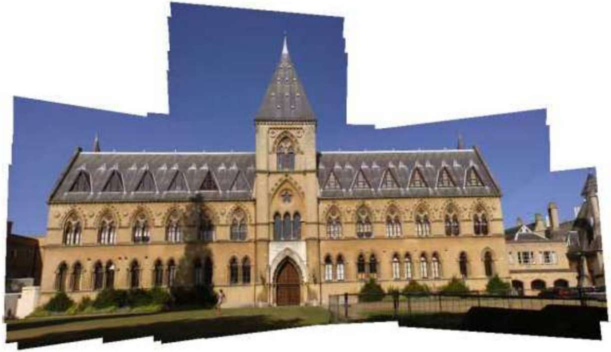
Die 3D-Punkte der Ebene sind somit $b_0 + \lambda_1 b_1 + \lambda_2 b_2$, d.h. $X = H \cdot [\lambda_1, \lambda_2, 1]$.

Das sind zugleich die homogenen Koordinaten.

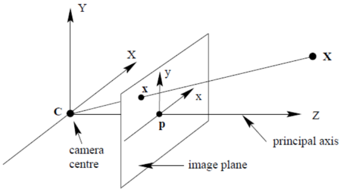
Homographie ist eine lineare Transformation in homogenen Koordinaten.

Homographie – Mosaik

Viele Ebenen (Bildschirme) werden auf eine abgebildet.



Kamerakalibrierung



Kameraparameter (keine Verzehrung, quadratische Pixel):

– intrinsische Parameter – Brennweite f, Hauptpunkt $[P_x, P_y]$

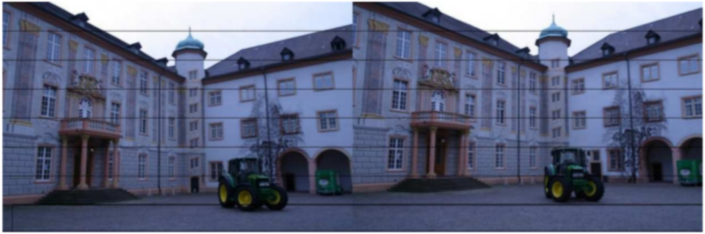
– extrinsische Parameter – Position des Kamerakoordinatensystem in der Welt (3 Winkel in Form der 3×3 Rotationsmatrix R und 3D-Verschiebung t)

$$x = PX = \begin{bmatrix} f & 0 & P_x & 0 \\ 0 & f & P_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

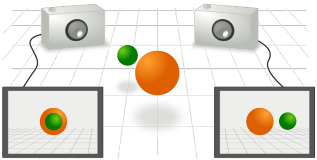
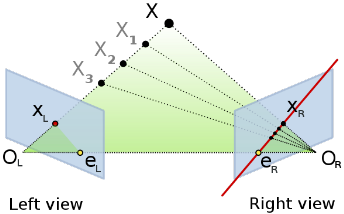
Kamerakalibrierung:
Gegeben sei die Lernstichprobe $\big((X, x) \dots\big)$, Gesucht werden die Kameraparameter.

Bestimmung der Lage eines Objektes:
Gegeben sind die Kameraparameter und Abbilder x, Gesucht werden die X.

Homographie – Rektifizierung



Epipolargeometrie

Sei $t = O_l - O_r$ der Verschiebungsvektor.

$x_l + t$ liegt in der (grünen) Ebene π (Koordinatensystem der linken Kamera wird um t verschoben)

$(x_l+t) \times t = (x_l+t)[t]_\times$ steht senkrecht zur π ($[t]_\times$ ist der Kreuzprodukt in der Matrixform)

Derselbe Vektor in der Koordinatensystem der rechten Kamera ist $a = (x_l + t)[t]_\times R$

x_r liegt in π, d.h. $\langle a, x_r \rangle = 0$

Schließlich: $(x_l + t)[t]_\times R x_r = x_l[t]_\times R x_r + t[t]_\times R x_r = x_l[t]_\times R x_r = x_l E x_r = 0$
mit **Essential Matrix** $E = [t]_\times R$