03-MB-709.03

Echtzeitbildverarbeitung (10)

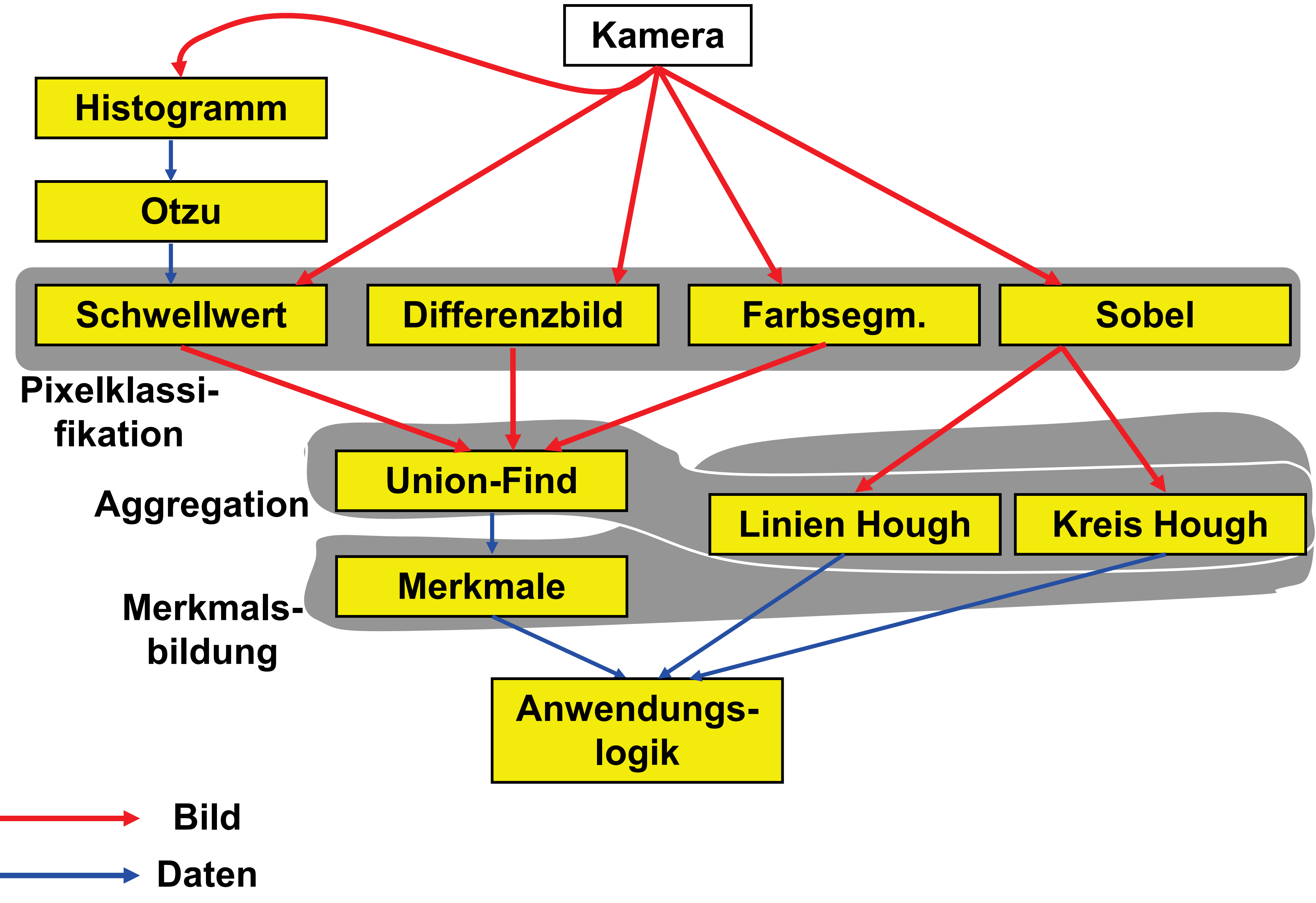
Prof. Dr. Udo Frese

Kameragleichung in 3D

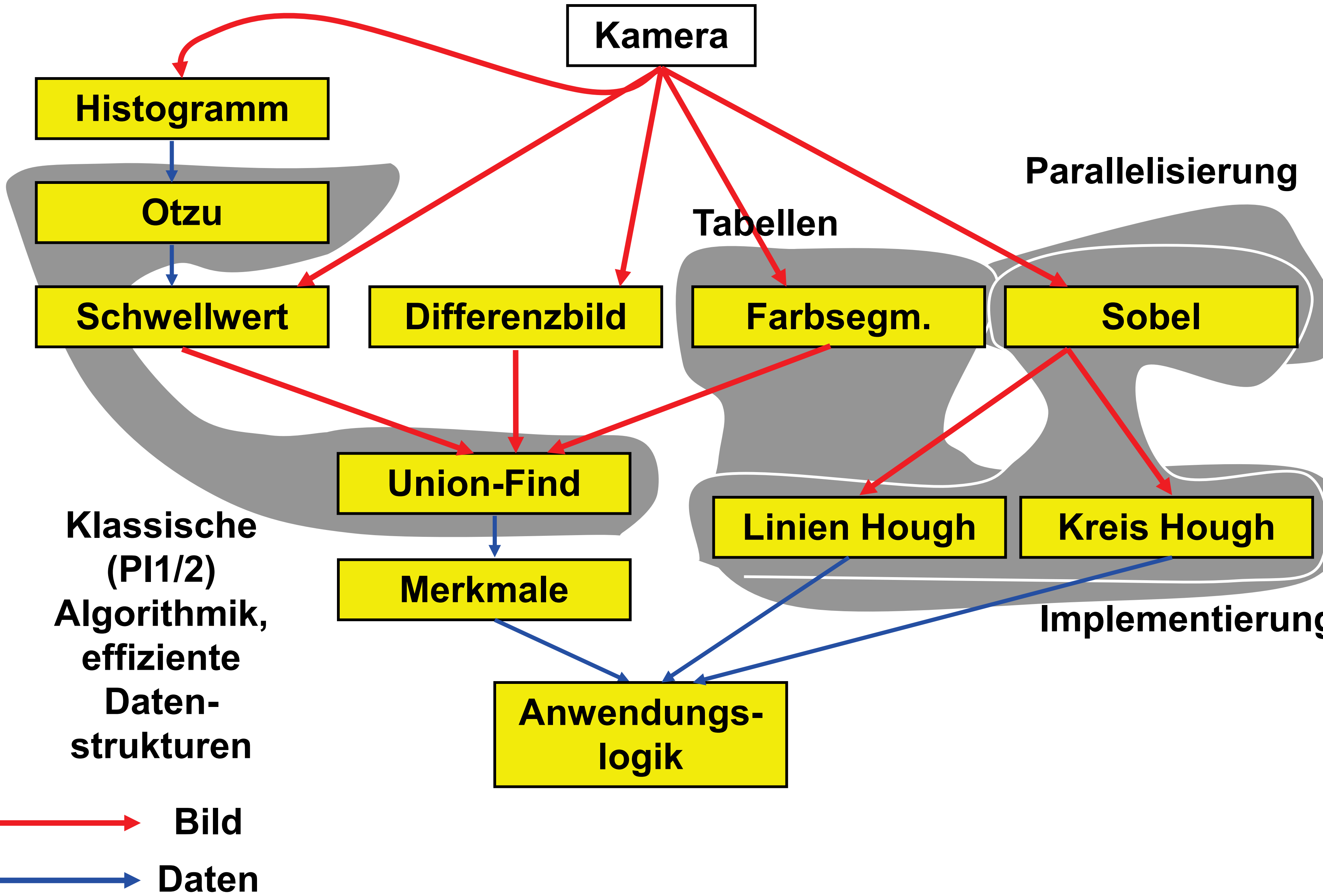
Geometrische Rekonstruktion

RANSAC

Schema: Datengetriebene (bottom up) Bildverarbeitung



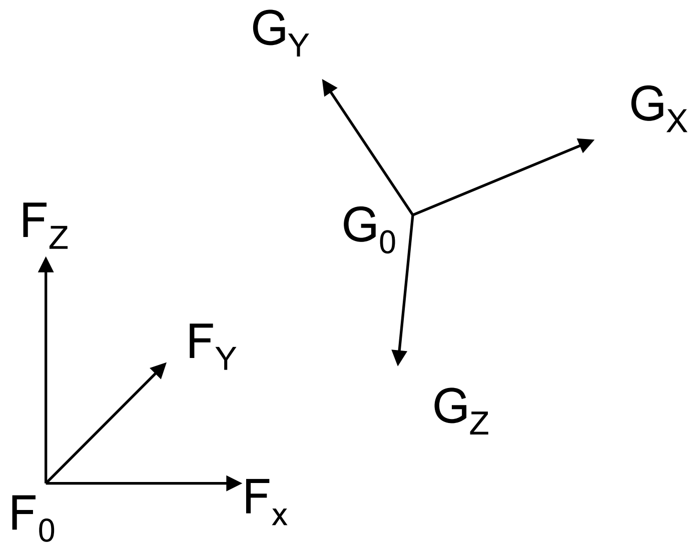
Techniken zur Beschleunigung



Was bisher geschah

Homogene Koordinaten

- 4. Komponente, entweder 0 (freier Vektor) oder 1 (Ortsvektor).
- Koordinatentransformationen als 4×4 Matrix.
- Verkettung mit Matrizenmultiplikation
- Umkehrung mit Matrixinvertierung
- Matrix immer GInF nennen, dann Konsistenz ablesbar.
- Spalten 1-3 von GInF sind die Achsen G_x, G_y, G_z von G in Koordinaten von F
- Spalte 4 ist der Ursprung G_0 von G in Koordinaten von F



Kameragleichung in 3D

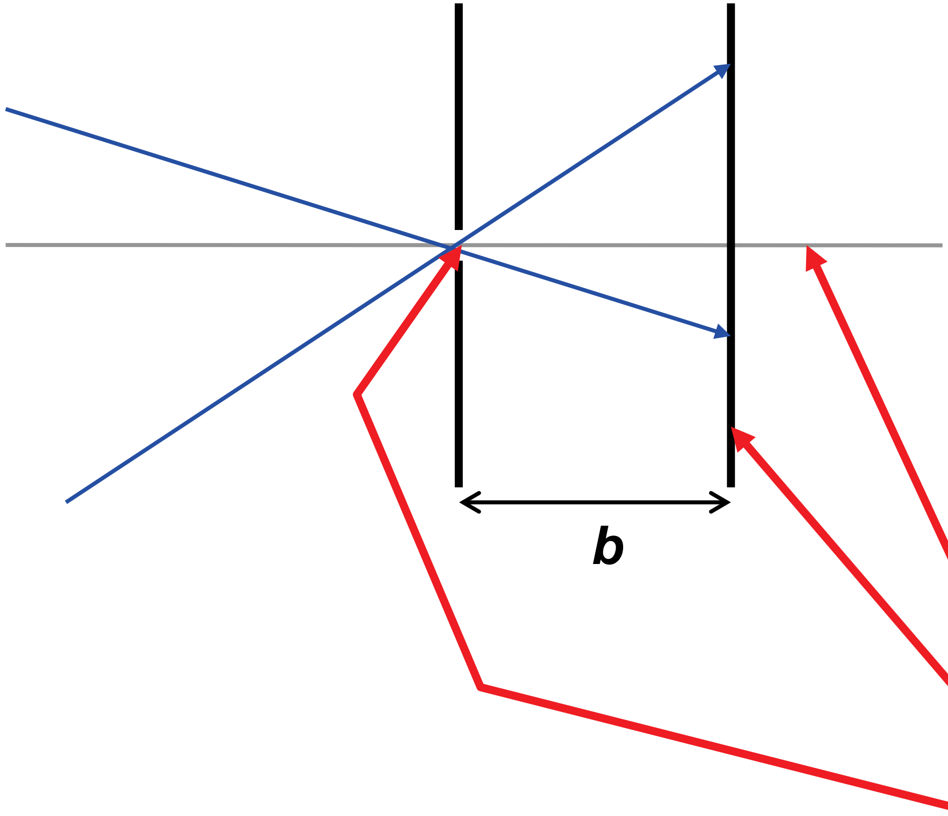
Kameragleichung in 3D

Lochkamera

- **Definiert Abbildung**
  - Eingabe: p_W Punkt in Weltkoordinaten $\in R^3$
  - Eingabe: CInW Kamerapose in Weltkoordinaten $\in R^{4\times 4}$
  - Ausgabe: p_I Punkt in Bildkoordinaten $\in R^2$
  - p: $R^3 \times R^{4\times 4} \rightarrow R^2$:($p_W$, C2W) $\rightarrow$ (p_I),
- **Vier Schritte**
  - Welt- zu Kamerakoordinaten
  - Perspektive
  - Verzerrung
  - Skalierung
- **Welt zu Kamerakoordinaten**
  - Multiplikation mit $CInW^{-1}$

$$\begin{pmatrix} x_C \\ y_C \\ z_C \\ 1 \end{pmatrix} = CInW^{-1} \cdot \begin{pmatrix} x_W \\ y_W \\ z_W \\ 1 \end{pmatrix}$$

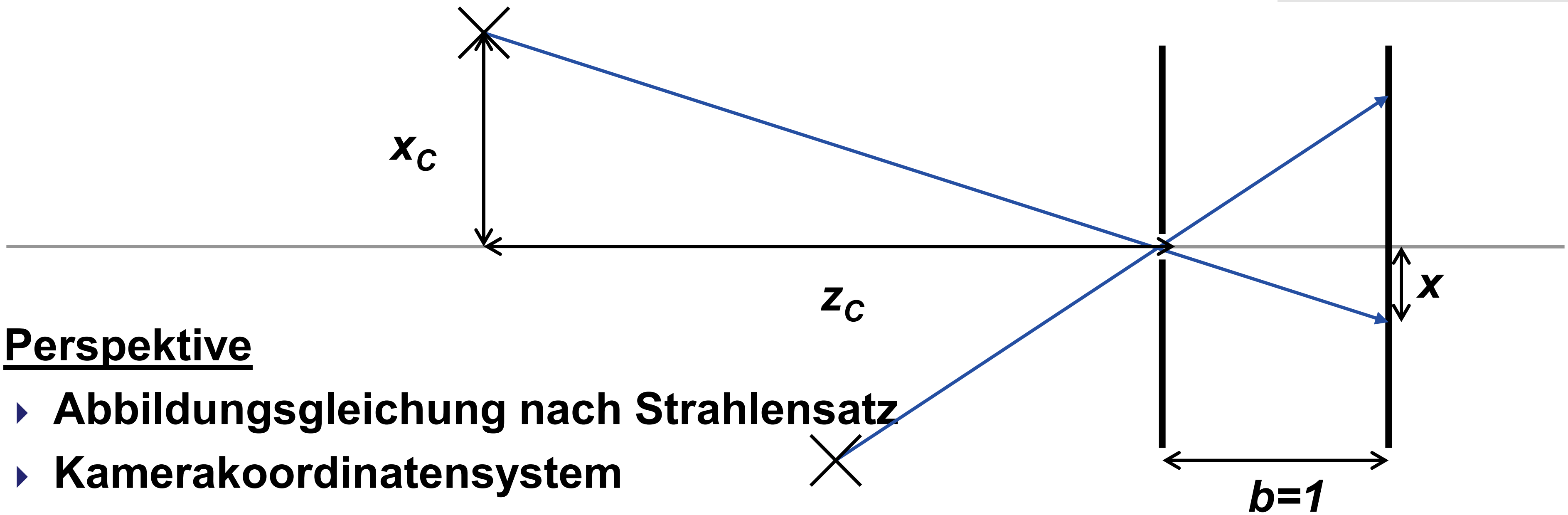
Kameragleichung in 3D



Lochkamera (Camera obscura)

- Lichtstrahlen fallen durch Loch in der Vorderseite
- auf Rückseite entsteht das (umgedrehte) Bild
- definiert abstrakte perspektivische Abbildung
- Abstand ist die Bildweite *b* *(ähnlich Brennweite f)*
- Optische Achse
- Bildebene
- Optisches Zentrum der Kamera
- Setze b=1
- Effekt des wirklichen b später bei Skalierung

Kameragleichung in 3D



Perspektive

- **Abbildungsgleichung nach Strahlensatz**
- **Kamerakoordinatensystem**
  - x_C zeigt in Bild-rechts Richtung
  - y_C zeigt in Bild-unten Richtung
  - z_c zeigt in die Tiefe (optische Achse)
- **(x_c, y_c, z_c) Punkt in Kamerakoordinaten**
- **(x, y) Abbild im *physikalischen* Bild**
- **Perspektive *p***

$$\begin{pmatrix} x \\ y \end{pmatrix} = p \begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix} = \begin{pmatrix} x_C / z_C \\ y_C / z_C \end{pmatrix}$$

Kameragleichung in 3D

Verzerrung

- besonders weitwinklige oder billige Objektive verzerren das Bild.
- Geraden im Bild werden krumm
- Linsen radialsymmetrisch $\Rightarrow$ Verzerrung radialsymmetrisch.
- exaktes Modell ist schwierig.
- einfaches phänomenologisches Modell mit Parameter κ_1, κ_2 (Zhang)

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = d_{\kappa_1,\kappa_2}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \left(1 + \kappa_1\left(x^2 + y^2\right) + \kappa_2\left(x^2 + y^2\right)^2\right)$$

Kameragleichung in 3D

Skalierung

- Pixelgröße und Bildweite b (Brennweite f) skalieren Bild
- zusammengefasst in α
- optische Achse trifft Bild in (u_0, v_0)
- (u_0, v_0) ungefähr in Bildmitte aber nicht exakt (±30 Pixel)

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \alpha \cdot d_\kappa \left(p \begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix} \right)$$

Kameragleichung in 3D

- **Gesamtformel**
  - W steht für „im Weltsystem"

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \alpha \cdot d_{\kappa_1,\kappa_2}\left(p\left(CInW^{-1} \cdot p_W\right)\right)$$

- **Zusammenfassung:**
  - Welt nach Kamera umrechnen ($CInW^{-1}$)
  - Perspektive
  - Verzerrung (κ_1, κ_2)
  - Skalierung (α) und Verschiebung (u_0,v_0)
- **Wichtig: Diese Formel erklären können (für Leben und Prüfung)**

Kameragleichung in 3D

- **Frage an das Auditorium:**
  Schätzt die Parameter
  - Verschiebung (u_0,v_0)
  - Skalierung (α),
  - Radialverzerrung (κ_1, κ_2=0);
- **Objektiv hat +/- 45° Öffnungswinkel**

1024

768

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \alpha \cdot d_{\kappa_1,\kappa_2}\left(p\left(CInW^{-1} \cdot p_W\right)\right)$$

$$p\begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix} = \begin{pmatrix} x_C / z_C \\ y_C / z_C \end{pmatrix}$$

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = d_{\kappa_1,\kappa_2}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \left(1 + \kappa_1\left(x^2 + y^2\right) + \kappa_2\left(x^2 + y^2\right)^2\right)$$

Kameragleichung in 3D

1024

768

- **Frage an das Auditorium:**
  Schätzt die Parameter
  - Verschiebung (u_0, v_0) = **(512, 384)**
  - Skalierung (α), ≈ ***512***
  - Radialverzerrung $(\kappa_1, \kappa_2^{=0})$; $\kappa_1 \approx$ **-0.05**
- **Objektiv hat +/- 45° Öffnungswinkel**

$$\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}u_0\\v_0\end{pmatrix}+\alpha\cdot d_{\kappa_1,\kappa_2}\left(p\left(CInW^{-1}\cdot p_W\right)\right)$$

$$p\begin{pmatrix}x_C\\y_C\\z_C\end{pmatrix}=\begin{pmatrix}x_C/z_C\\y_C/z_C\end{pmatrix}$$

$$\begin{pmatrix}\tilde{x}\\\tilde{y}\end{pmatrix}=d_{\kappa_1,\kappa_2}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}\cdot\left(1+\kappa_1\left(x^2+y^2\right)+\kappa_2\left(x^2+y^2\right)^2\right)$$

Geometrische Rekonstruktion

*Multiple view geometry in computer vision
Von Richard Hartley, Andrew Zisserman

Geometrische Rekonstruktion

Fragestellung

- **„Welche geometrische Information lässt sich aus den perspektivischen Bildern von *n* Punkten gewinnen?“**
  - über die Punkte
  - über die Kamera
  - 2D → 3D, Bild → Szene
- **betrachte zuerst Struktur der Information**
- **betrachte dann verschiedene Fälle**

Geometrische Rekonstruktion

Was misst eine Kamera?

- **Eine kalibrierte (α und κ_1 κ_2 bekannt) Kamera ist ein Winkelmessgerät.**
- **Zu jedem Pixel gehört ein Strahl**
  - in bekannter Richtung
  - relativ zur Kamera
  - vom optischen Zentrum
  - Berechnung: Kameragleichung rückwärts
  - entspricht zwei Winkeln (horizontal, vertikal), relativ zur Kamera
- **Zwei Pixel haben einen Winkel unabhängig von Kameraorientierung**
  - zwei Strahlen
  - Winkel der Strahlen

Geometrische Rekonstruktion

Fundamentale Eigenschaft der Perspektive:

- **Distanzen lassen sich nicht rein aus Bildgrößen ausrechnen.**
- **mindestens eine Länge muss bekannt sein.**
- **so genannter Skalenmasstab**
- **Begründung: Würde man die ganze Szene inkl. Kamera skalieren, gäbe es das gleiche Bild**
- **Theorem:**
  - Aus mehreren Kamerabildern einer Szene
  - läßt sich Szene geometrisch rekonstruieren
  - bis auf Lage des Koordinatensystems und Skalierung.
  - (unter einigen Bedingungen)

Geometrische Rekonstruktion

- **Frage an das Auditorium:**
  - Liefert die Brennweite (z.B. 8mm) einen Skalenmaßstab?
  - Liefert die Blende (Objektivdurchmesser) einen Skalenmaßstab?

Geometrische Rekonstruktion

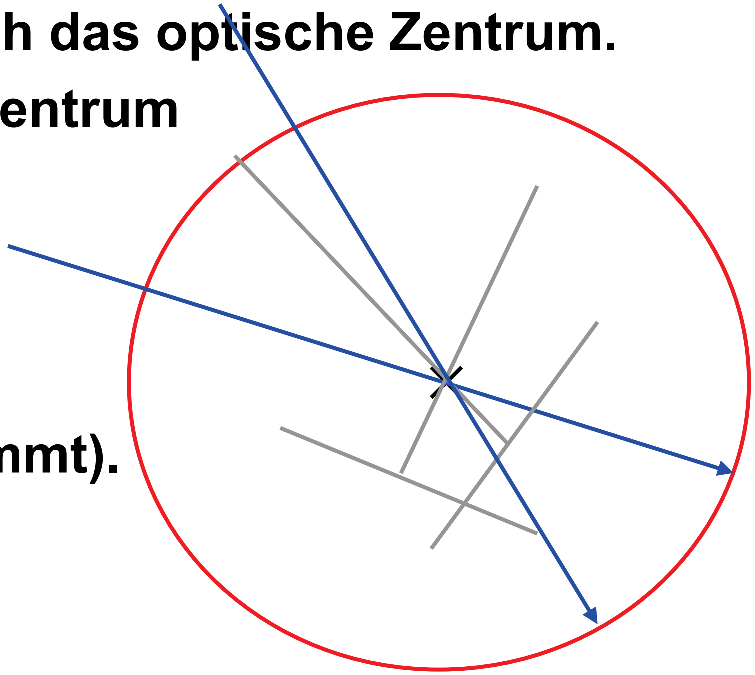
- **Frage an das Auditorium:**
  - Liefert die Brennweite (z.B. 8mm) einen Skalenmaßstab?
  - Liefert die Blende (Objektivdurchmesser) einen Skalenmaßstab?

- **Brennweite: Nein, man kann die Brennweite nicht sehen. Das Bild bleibt das selbe, wenn man die Szene (Objekte und Kamerapose) skaliert, aber die Kamera selbst beibehält.**
- **Blende: Ja, z.B. bei Scharfstellung auf unendlich wird die Schärfe größer, wenn die Szene größer wird. Das unscharfe Bild eines Lichtpunktes entspricht der Blende an der Stelle des Punktes. Damit definiert die Blende einen Skalenmaßstab.**
- **Außer für Mikroskopie unpraktikabel, dort erlaubt es aber 3D Bilder zu generieren.**

Geometrische Rekonstruktion

Kugelbild

- Strahlen durch optisches Zentrum auf Kugel projizieren
- technisch nicht durchführbar.
- theoretisch: Helligkeit aller Strahlen durch das optische Zentrum.
- alle Kamerabilder mit selben optischen Zentrum (variabel bzgl. Orientierung, α, κ) können daraus berechnet werden.
- Pixel $\rightarrow$ Strahl $\rightarrow$ Punkt auf Kugel
- $\Rightarrow$ drehen der Kamera bringt keine neue Information (außer wo neues ins Bild kommt).
- dagegen, bewegen der Kamera bringt Information über die Tiefe.



Geometrische Rekonstruktion

Freiheitsgrade (DOF)

- **Anzahl der reellen Zahlen zur Beschreibung eines Objektes**
  - für jede Zahl +1 DOF
  - für jede Nebenbedingung -1 DOF
  - für jede Mehrdeutigkeit -1 DOF
- **Grobe Abschätzung, ob bestimmte Größen ausreichen, andere Größen zu bestimmen.**

Geometrische Rekonstruktion

Freiheitsgrade

- **Anzahl der reellen Zahlen zur Beschreibung eines Objektes**
  - für jede Zahl +1 DOF
  - für jede Nebenbedingung -1 DOF
  - für jede Mehrdeutigkeit -1 DOF
- **Grobe Abschätzung, ob bestimmte Größen ausreichen, andere Größen zu bestimmen.**
- **Frage an das Auditorium: Wie viele Freiheitsgrade hat**
  - ein Punkt im Raum?
  - eine Gerade im Raum?
  - ein Kreis in der Ebene?
  - ein Kreis im Raum?

Geometrische Rekonstruktion

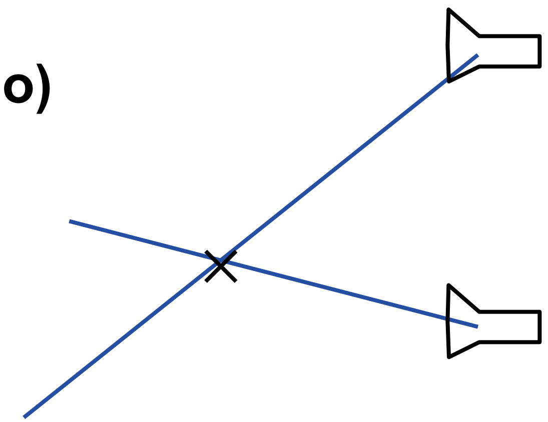
Freiheitsgrade

- **Anzahl der reellen Zahlen zur Beschreibung eines Objektes**
  - für jede Zahl +1 DOF
  - für jede Nebenbedingung -1 DOF
  - für jede Mehrdeutigkeit -1 DOF
- **Abschätzung, ob Größen ausreichen, andere Größen zu bestimmen.**
- **Frage an das Auditorium: Wie viele Freiheitsgrade hat**
  - ein Punkt im Raum? $3 \times 1 = 3$
  - eine Gerade im Raum? $2 \times 3 - 2 = 4$
  - ein Kreis in der Ebene? $2+1 = 3$
  - ein Kreis im Raum? $3+1+(3-1) = 6$

Geometrische Rekonstruktion

<u>unbekannte Punkte, bekannte Kameraposen:</u>

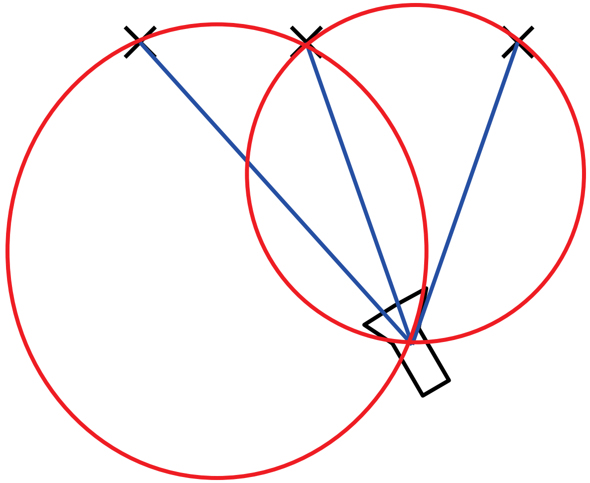
- **1 Kamerapose (2 DOF → 2/3 DOF):**
  - Punkt liegt auf Geraden.
- **2 Kameraposen (4 DOF → 3 DOF):**
  - volle Position.
- **2 Kameras mit bekannter Relativpose (Stereo) (4 DOF → 3 DOF):**
  - Position relativ zur Kamera.



Geometrische Rekonstruktion

bekannte Punkte, unbekannte Kamera (2D) :

- **1 Punkt (1 DOF $\rightarrow$ 1/3 DOF):**
  - Orientierung als Funktion der Position
- **2 Punkte (Abstand bekannt) (2 DOF $\rightarrow$ 2/3 DOF):**
  - Kamera auf Kreis durch beide Punkte (Peripheriewinkelsatz)
  - Orientierung als Funktion der Position
- **3 Punkte (Bekanntes Dreieck) (3 DOF $\rightarrow$ 3 DOF):**
  - Volle Position und Orientierung.



Geometrische Rekonstruktion

<u>bekannte Punkte, unbekannte Kamera (3D) :</u>

- **1 Punkt (2 DOF $\rightarrow$ 2/6 DOF):**
  - keine Information über Position.
  - bei bekannter Position Orientierung aber mit unbekannter Drehung um Strahl zum Punkt.
- **2 Punkte (Abstand bekannt) (4 DOF $\rightarrow$ 4/6 DOF):**
  - bei bekannter Position vollständige Orientierung (4 DOF $\rightarrow$ 3 DOF).
  - unbekannte Drehung um Verbindungsgeraden (1 DOF).
  - Position in der Kamera-Punkt-Punkt Ebene auf einem Kreis durch beide Punkte (1 DOF)
- **3 Punkte (bekanntes Dreieck) (6 DOF $\rightarrow$ 6 DOF)**
  - vollständige Position und Orientierung mit bis zu 4 Lösungen.
- **4 Punkte (bekanntes Viereck) (8 DOF $\rightarrow$ 7 DOF)**
  - vollständige Pose und f (Routine in Übungen vorgegeben)

Geometrische Rekonstruktion

bekannte Punkte, unbekannte Kamera (3D) :

- **viele Punkte (z.B. Gitter, Übung) (n-DOF)**
  - vollständige Pose, α, κ_1, κ_2, u_0,v_0, (Kalibrierung)
- **beachten bei Kalibrierung**
  - Ohne Verzerrung: Kalibrierkörper soll möglichst das ganze Bild abdecken, sonst moderat sinkende Genauigkeit
  - mit Verzerrung: Kalibrierkörper muss (fast) das ganze Bild abdecken.
  - es müssen Punkte unterschiedlicher Tiefe in der Szene sein, weil sonst Abstand und Brennweite nicht unterschieden werden können.
  - Bildmitte (u_0,v_0) und Orientierung sind nur bei weitwinkligen Kameras unterscheidbar

Geometrische Rekonstruktion

Bekannte Geraden, unbekannte Kamera

- **Schneidende Geraden:**
  - Schnittpunkte in 3D/2D bilden und diese behandeln.
- **Parallele Geraden:**
  - schneiden sich im Bild (Fluchtpunkt).
  - Fluchtpunkt ist Bild der Parallelen durchs optischen Zentrum
  - $\Rightarrow$ Fluchtpunkt unabhängig von Translation der Kamera / Geraden,
  - Fluchtpunkt nur abhängig von Orientierung der Kamera / Richtung der Geraden.

Geometrische Rekonstruktion

Unbekannte Punkte, unbekannte Kamerapose, viele Bilder

- **Posen und Punkte bis auf Skalenfaktor**
- **Structure from Motion**
- **Kalibrierung (wenn genug Punkte auf den Bildern)**

Video 1

Zusammenfassung

- **Geometrische Rekonstruktion**
  - Kalibrierte Kamera ist ein Winkelmessgerät, eine Länge benötigt
  - Meist ist es so, wie die Zahl der Freiheitsgrade suggeriert.
  - Fluchtpunkt von parallelen Geraden hängt nur von Richtung ab.
  - Zwei Winkel zu Landmarken beschränkt die Position auf einen Kreis.
- **Kameragleichung in 3D (Parameter DOF)**
  - Transformation in Kamerakoordinaten $C2W^{-1}$ (6DOF)
  - Perspektive *p* (0 DOF)
  - Verzerrung *d* (üblich: 0-2DOF)
  - Skalierung/Offset (1/3DOF)

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \alpha \cdot d_{\kappa_1, \kappa_2}\left(p\left(CInW^{-1} \cdot p_W\right)\right)$$