

Bildverarbeitung: Stereo

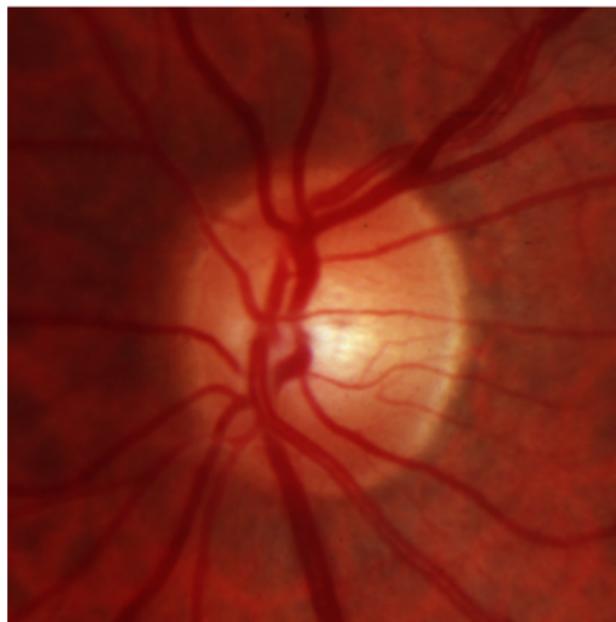
Geometrisches Stereo vs. andere Methoden



Shape from Texture



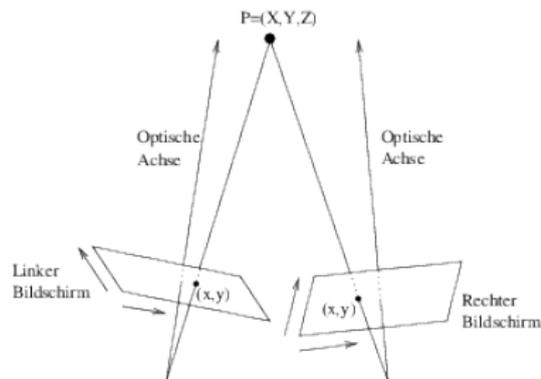
Shape from Shading



Stereo

Heute:

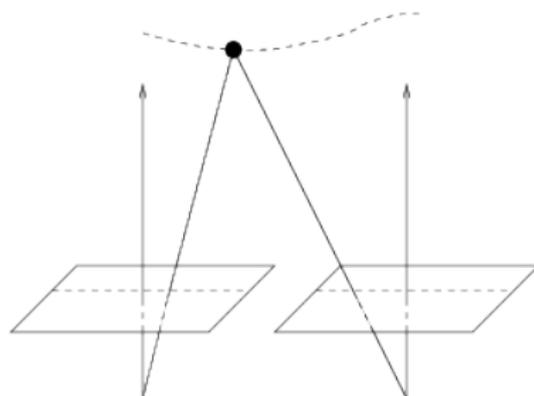
Stereo = Epipolargeometrie + Ähnlichkeitsmaße + Diskrete Energieminimierung



Allgemeine Situation:

$$x_l F x_r = 0$$

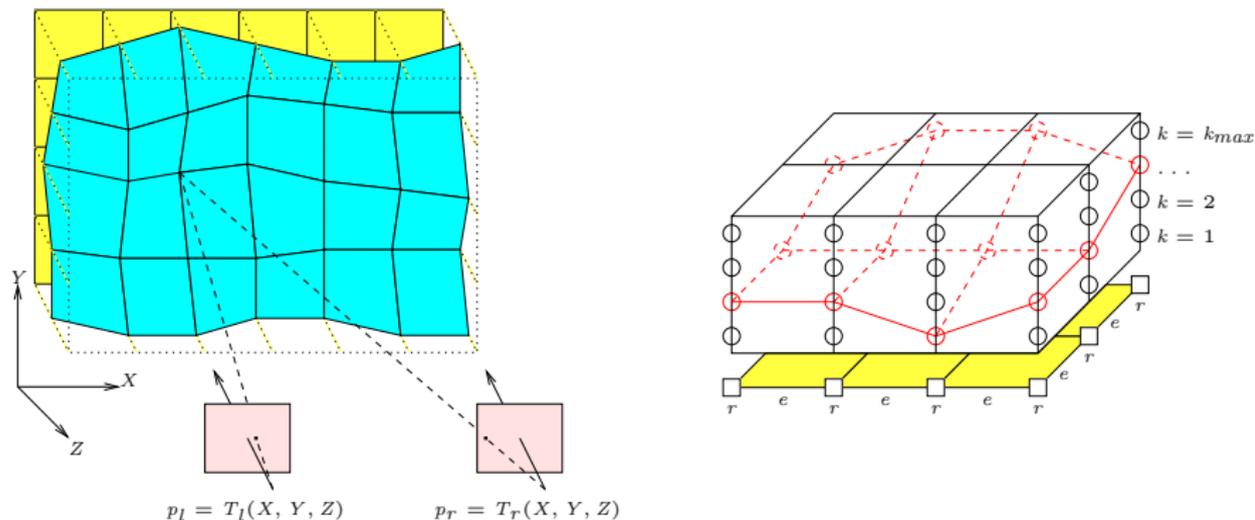
(Bedingungen der Epipolargeometrie)



Paralleles (rektifiziertes) Stereo:

Die Menge der Korrespondenzpaare ist (x_l, x_r, y) , d.h. $\in \mathbb{R}^3$

Repräsentation der Tiefenkarte (diskret)



Definitionsbereich ist ein Graph $V = (R, E)$, $r \in R$ ein „Pixel“, $R \subset \mathbb{Z}^2$

Wertebereich ist eine (diskrete) Menge K der Label (Tiefenwerte)

Tiefenkarte ist eine Abbildung $y : R \rightarrow K$,

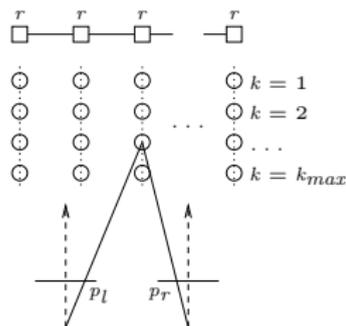
d.h. jeder Position $r \in R$ wird ein Tiefenwert $k \in K$ zugeordnet.

Jedes Paar (r, k) repräsentiert einen Punkt $(X, Y, Z) \in \mathbb{R}^3$

→ wird auf das entsprechende Korrespondenzpaar $(p_l \in \mathbb{R}^2, p_r \in \mathbb{R}^2)$ abgebildet

Quadratische Differenz der Farbwerte:

$$A(p_l, p_r) = (I_l(p_l) - I_r(p_r))^2$$



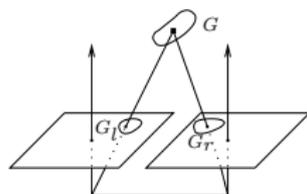
Rauschen unterdrücken:

$$A(p_l, p_r) = \sum_{\Delta p \in F} [I_l(p_l + \Delta p) - I_r(p_r + \Delta p)]^2$$

Farbtransformationen erlauben:

$$A(p_l, p_r) = \min_{C_v} \sum_{\Delta p \in F} [I_l(p_l + \Delta p) + C_v - I_r(p_r + \Delta p)]^2$$

$$A(p_l, p_r) = \min_{C_v, C_s} \sum_{\Delta p \in F} [I_l(p_l + \Delta p) \cdot C_s + C_v - I_r(p_r + \Delta p)]^2$$



Geometrische Transformationen erlauben:

$$A(p_l, p_r) = \min_{C_v, C_s, Tr} \sum_{\Delta p \in F} [I_l(p_l + \Delta p) \cdot C_s + C_v - I_r(Tr(p_r + \Delta p))]^2$$

Block Matching

Gar kein a-priori Modell – unabhängige Entscheidungen für alle r :

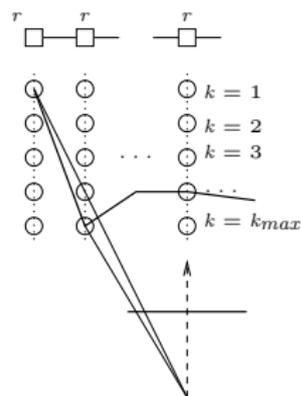
$$y(r) = \arg \min_k A(r, k) \quad \forall r$$



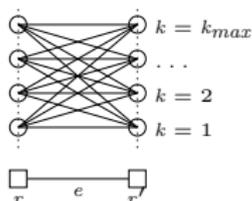
- sehr einfach
- sehr schnell (durch Integralbild)
- kann für die Schätzung „undichter“ Tiefenkarten verwendet werden
- als eine Initialisierung für komplizierter Verfahren oft gut geeignet

Zeilenweise Ansätze

Bestimmte Kombinationen der Label in den (horizontal) benachbarten Knoten sind gar nicht möglich:



Eine Funktion definieren, die alle Labelpaare bewertet:



Für jede Zeile des Definitionsbereiches

$$y^* = \arg \min_y \left[\sum_{i=1}^n q_i(y_i) + \sum_{i=2}^n g(y_{i-1}, y_i) \right]$$

→ Dynamische Programmierung

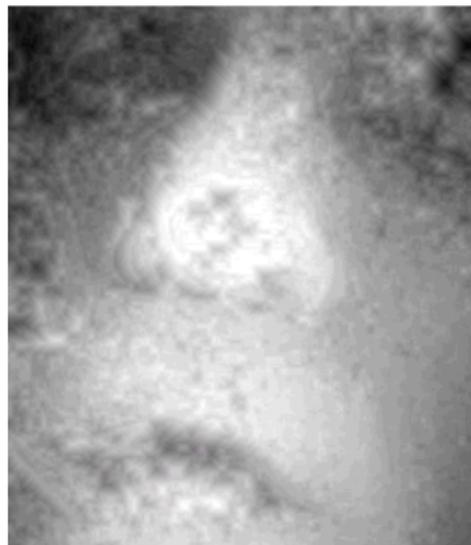
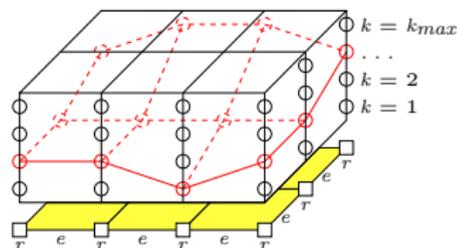


[Gimel'farb, sehr lange her]

Energieminimierung

Es gibt Funktionen g für die Bewertung der Labelpaare sowohl in der horizontalen als auch in der vertikalen Richtung

$$y^* = \arg \min_y \left[\sum_{r \in R} q_r(y_r) + \sum_{rr' \in E} g(y_r, y_{r'}) \right]$$



$$g(k, k') = \begin{cases} 0 & \text{wenn } |k - k'| \leq \delta \\ \infty & \text{sonst} \end{cases}$$

$$g(k, k') = c \cdot (k - k')^2$$

$$g(k, k') = \begin{cases} 0 & \text{wenn } k = k' \\ a > 0 & \text{sonst} \end{cases}$$

[Boykov, Kolmogorov, Veksler, Zabih, um 2001]

Die a-posteriori Wahrscheinlichkeitsverteilung der Labellings (Tiefenkarten):

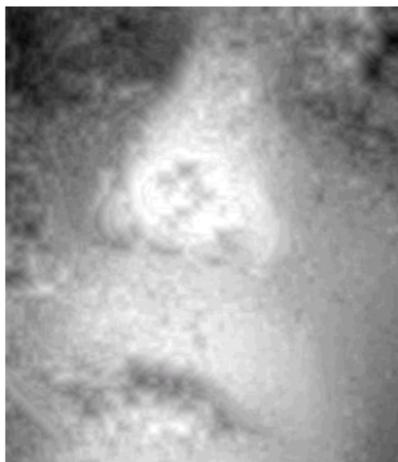
$$p(y) \sim \exp \left[\sum_r q_r(y_r) + \sum_{rr'} g(y_r, y_{r'}) \right]$$

Maximum a-posteriori Entscheidung:

$$y^* = \arg \min_y \left[\sum_r q_r(y_r) + \sum_{rr'} g(y_r, y_{r'}) \right]$$

Minimum mean square error:

$$y_r^* = \sum_k k \cdot p(y_r = k) \quad \forall r$$



[Schlesinger, 2003]

„Einfache“ Variante:

- 1) Bilder rektifizieren (SIFT, RANSAC ... , Homographien anwenden)
- 2) Rektifiziertes Stereo lösen

„Kompliziertere“ Variante: Sowohl die Tiefenkarte als auch die Epipolargeometrie sind Unbekannte in einem einheitlichen Modell.

Iterieren:

- 1) Korrespondenzpaare \rightarrow Epipolargeometrie (z.B. mit 8-Punkte Algorithmus)
 - 2) Epipolargeometrie \rightarrow Rektifiziertes Stereo, Abweichungen zulassen!!!
-

Kontinuierliche Ansätze:

Idee: der Optische Fluss (Bewegungsfeld) so einzuschränken, dass er die Bedingungen der (bekannten) Epipolargeometrie erfüllt.
Slesareva, Bruhn, Weickert, 2005

Erweiterung: die Fundamentalmatrix F ist unbekannt, man schätze sie auch
Valgaerts, Bruhn, Mainberger, Weickert, 2010