Computer Vision: Stereo

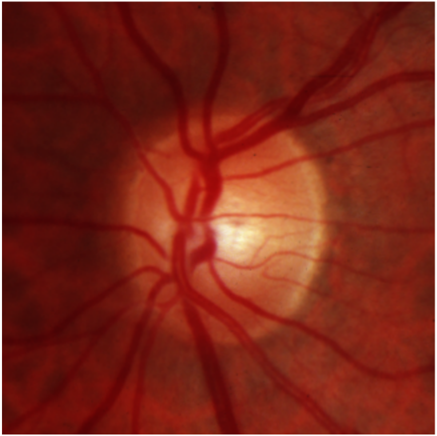
D. Schlesinger – TUD/INF/KI/IS

– Geometrisches Stereo vs. andere Methoden
– Geometrische Grundlagen
– Einfache diskrete Verfahren
– Diskrete Energieminimierung
– Statistische Modelle
– Optische Flüsse

Geometrisches Stereo vs. andere Methoden

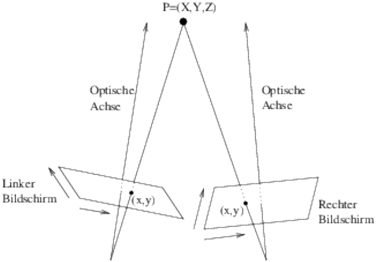
Shape from Texture

Shape from Shading



Stereo

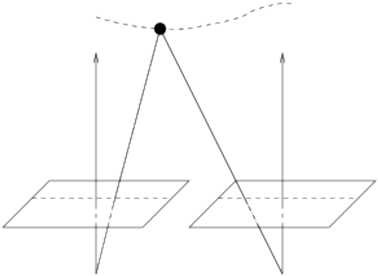
(Zur Erinnerung)



Allgemeine Situation:

$$x_l F x_r = 0$$

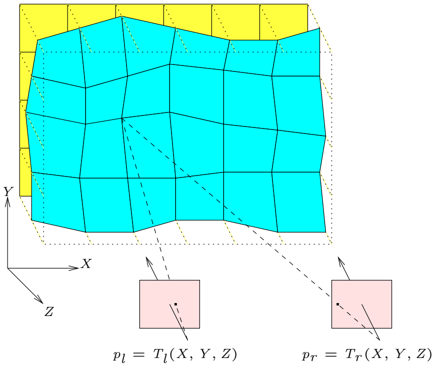
(Bedingungen der Epipolargeometrie)

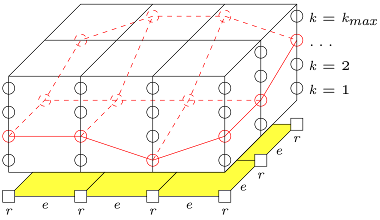


Paralleles (rektifiziertes) Stereo:

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

Tiefenkarte (diskret)





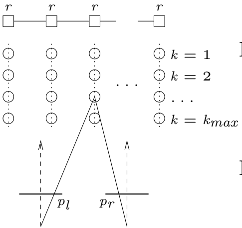
Definitionsbereich ist ein Graph $V = (R, E)$, $r \in R$ ein „Pixel“, $R \subset \mathbb{Z}^2$

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

Tiefenkarte ist eine Abbildung $y : R \to K$,
d.h. jeder Position $r \in R$ wird ein Tiefenwert $k \in K$ zugeordnet.

Jedes Paar (r, k) repräsentiert einen Punkt $(X, Y, Z) \in \mathbb{R}^3$
$\rightarrow$ wird auf das entsprechende Korrespondenzpaar $(p_l \in \mathbb{R}^2, p_r \in \mathbb{R}^2)$ abgebildet

Ähnlichkeitsmaße



Quadratische Differenz der Farbwerte:

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

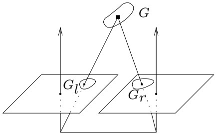
Rauschen unterdrücken:

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

Farbtransformationen erlauben:

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

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



Geometrische Transformationen erlauben:

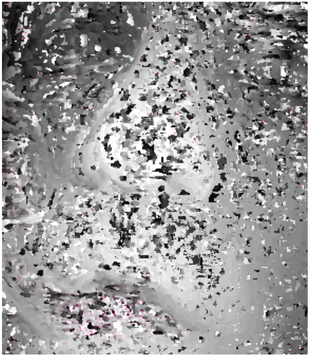
$$A(p_l, p_r) =$$

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

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

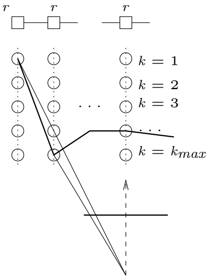
– sehr schnell (durch Integralbild)

– kann für die Schätzung „undichter“ Tiefenkarten verwendet werden

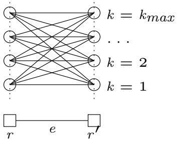
– als eine Initialisierung für komplizierter Verfahren oft gut geeignet

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

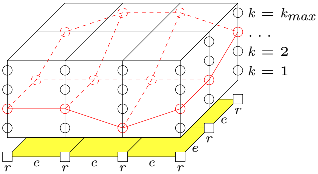
$$y^* = \arg\min_y \left[\sum_{i=1}^{n} q_i(y_i) + \sum_{i=2}^{n} g(y_{i-1}, y_i) \right]$$

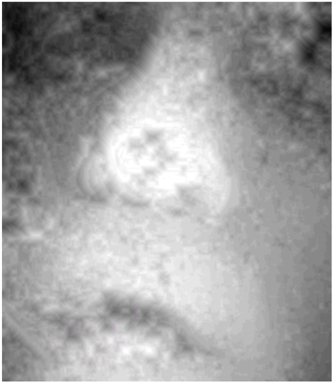
→ Dynamische Programmierung

[Gimel'farb, sehr lange her]

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

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





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

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

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

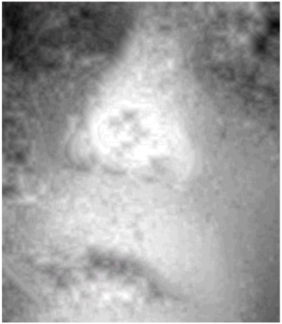
[Boykov, Kolmogorov, Veksler, Zabih, um 2001]

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]

Optische Flüsse

Idee [Slesareva, Bruhn, Weickert, 2005]: der Optische Fluss $d(x)$ so einzuschränken, dass er die Bedingungen der (bekannten) Epipolargeometrie $x^T \cdot F \cdot (x + d(x))$ erfüllt.

Energiefunktional:

$$E(p) = E_D(p) + \beta E_S(p) \to \min_p$$

$$E_D(p) = \int_\Omega \Psi_D\Big(|g_r(x + d(p)) - g_l(x)|^2 + \alpha |\nabla g_r(x + d(p)) - \nabla g_l(x)|^2\Big)\, dxdy$$

$$E_S(p) = \int_\Omega \Psi_S\Big(|\nabla p|^2\Big)\, dxdy$$

x – Position im linken Bild, q_l, g_r – Bilder, d – Flussfeld,
p – die Komponente des Flussfeldes entlang den entsprechenden Epipolarlinien,
$p(x)$ ist (im Gegensatz zu $d(x)$) eine eindimensionale Größe.

Einsetzen, ableiten, mit Standardmethoden lösen ...

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