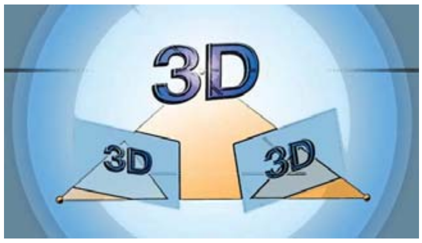
Martin Kleinsteuber: Computer Vision

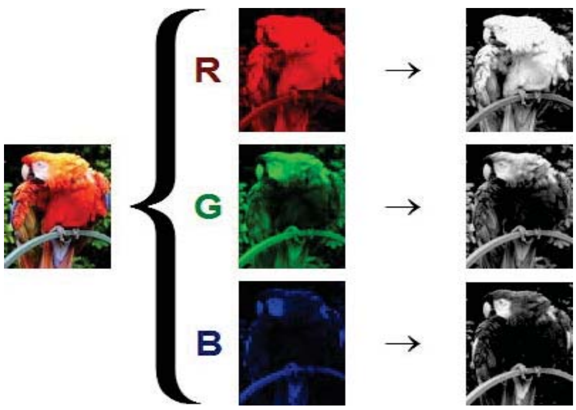
Kap. 1 – Wissenswertes über Bilder

1. Darstellung von Bildern



Darstellung von Bildern

Vom Farbbild zum Intensitätsbild



- Farbbilder bestehen aus mehreren Kanälen
- In diesem Kurs ausschließlich Graustufenbilder

Darstellung von Bildern

Kontinuierliche und diskrete Darstellung

- **Kontinuierliche** Darstellung als Funktion zweier Veränderlicher (zum Herleiten von Algorithmen)

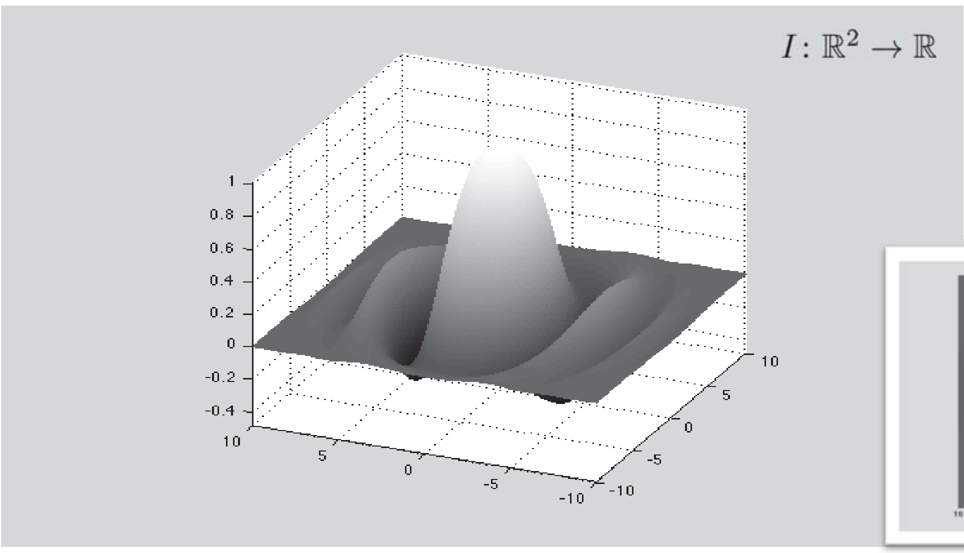
  $I\colon \mathbb{R}^2 \supset \Omega \to \mathbb{R}, \quad (x,y) \mapsto I(x,y)$
- Häufige Annahmen
  - I differenzierbar
  - Ω einfach zusammenhängend und beschränkt

- **Diskrete** Darstellung als Matrix $I \in \mathbb{R}^{m \times n}$
  Eintrag $I_{k,l}$ entspricht dem Intensitätswert
- Skalierung typischerweise zwischen [0, 255] oder [0, 1]

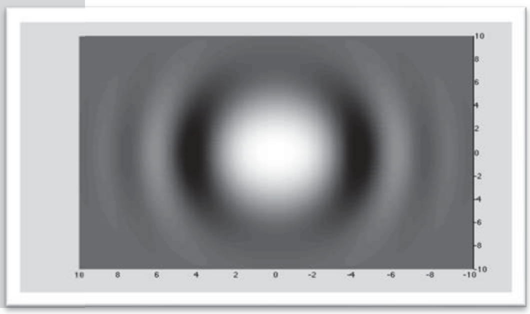
VGA: 480 x 640 Pixel (ca. 0.3 Megapixel)
HD: 720 x 1280 Pixel (ca. 1.0 Megapixel)
FHD: 1080 x 1920 Pixel (ca. 2.1 Megapixel)

Darstellung von Bildern

Graph einer Funktion

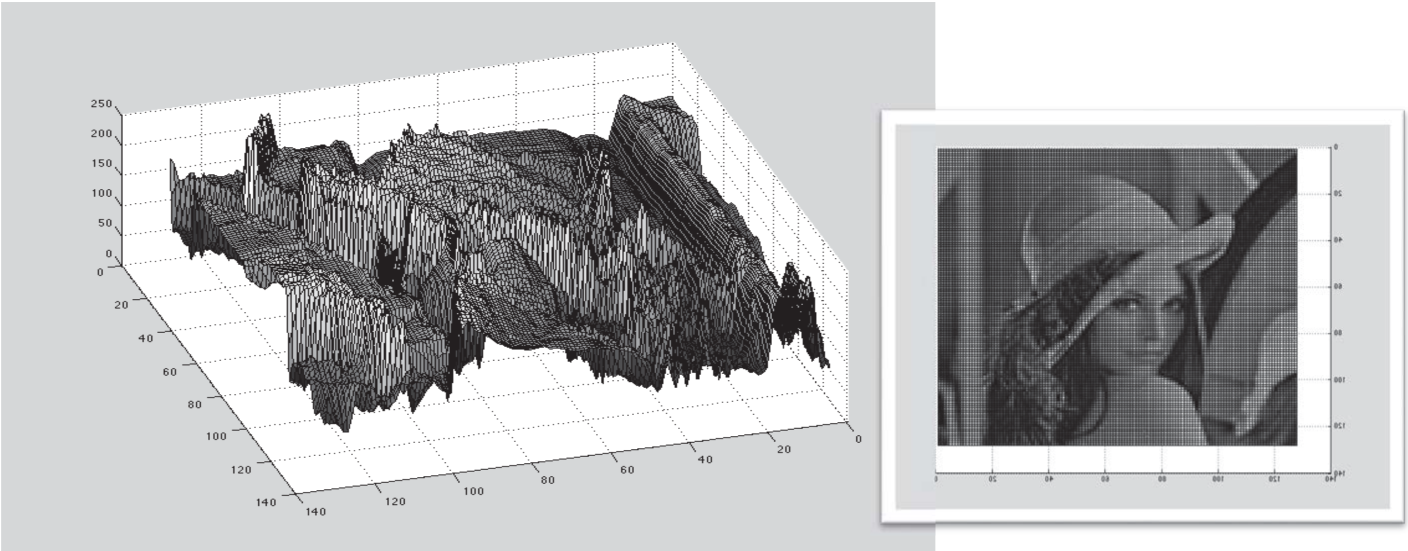


Draufsicht



Darstellung von Bildern

Graph eines Fotos



Darstellung von Bildern

Diskretes Abtasten

- Abtasten eines eindimensionalen Signals

$$S\{f(x)\} = (\dots, f(x-1), f(x), f(x+1), \dots)$$

- Abtasten eines Bildes

$$S\{I(x,y)\} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \\ \dots & I(x-1,y-1) & I(x-1,y) & I(x-1,y+1) & \dots \\ \dots & I(x,y-1) & I(x,y) & I(x,y+1) & \dots \\ \dots & I(x+1,y-1) & I(x+1,y) & I(x+1,y+1) & \dots \\ & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Darstellung von Bildern

Diskrete Darstellung/Matrixdarstellung

- Annahme: Ursprung links oben
- Matrixeintrag ist

$$I_{k,l} = S\{I(0,0)\}_{kl}$$

Columns 1 through 41

163	163	162	161	160	160	159	159	159	165	169	173	176	170	148	126	103
162	162	161	161	160	159	159	159	160	166	171	173	174	167	147	126	102
161	161	160	160	159	159	159	158	162	168	173	173	171	163	145	126	101
160	160	160	159	159	159	159	159	164	170	174	172	169	161	143	126	99
160	160	160	160	160	160	160	160	167	171	172	169	167	160	143	125	99
160	160	161	161	161	162	162	162	170	170	168	166	166	162	144	124	99
161	161	162	162	163	164	164	165	172	169	163	162	167	164	145	123	100
162	162	162	163	164	165	166	166	174	168	161	160	167	166	146	123	100
158	160	162	164	167	170	171	170	164	167	164	160	164	166	147	121	97
162	162	162	163	167	170	171	170	163	166	164	160	164	165	146	121	97
166	163	162	165	168	169	168	168	161	165	164	160	163	164	145	120	97
167	163	164	169	170	165	161	160	159	165	165	161	163	163	144	120	96
166	164	167	173	170	157	150	151	158	165	167	163	164	163	144	120	94
166	165	170	174	164	145	138	143	157	166	169	165	165	163	145	122	91
169	167	171	171	154	132	129	140	158	168	172	168	167	164	146	123	89
172	169	171	167	146	123	124	140	158	169	173	169	168	165	147	124	87
175	173	171	156	124	101	115	143	163	172	173	167	165	161	143	120	86
171	172	166	144	110	93	113	144	163	171	174	168	165	161	143	120	86
169	167	155	126	94	86	113	144	162	171	174	168	166	162	142	119	87
171	159	137	107	86	89	116	142	161	171	174	169	166	162	142	118	87
170	146	116	95	88	97	120	140	161	171	174	170	167	162	142	117	88
159	129	101	93	95	102	120	140	160	170	175	170	168	163	141	117	88
139	111	93	97	98	97	116	144	159	170	175	171	169	163	141	116	89
123	99	91	101	99	91	111	148	159	170	175	171	169	163	141	116	89
96	99	97	99	101	97	111	141	157	167	171	167	166	162	143	119	90
99	101	100	102	103	99	113	143	157	167	171	167	166	162	143	119	90
101	103	102	103	105	100	114	144	157	167	171	167	166	162	143	119	89
101	103	101	103	104	99	112	142	157	167	171	167	166	162	143	119	89
99	102	99	100	101	96	109	139	157	167	171	167	166	162	143	119	88
99	101	99	100	100	95	108	137	157	167	171	167	166	162	143	119	88
102	104	101	101	102	96	109	138	157	167	171	167	166	162	143	119	87
104	106	103	104	104	98	111	140	157	167	171	167	166	162	143	119	87
103	108	105	103	106	105	115	137	156	166	171	170	172	168	145	118	87
104	109	106	104	107	106	115	138	157	166	172	171	172	168	145	118	87
105	110	107	105	108	107	117	139	157	167	172	171	173	169	146	119	88
106	111	107	106	109	107	117	140	158	168	173	172	174	170	147	120	88
106	111	107	106	109	107	117	140	159	169	174	174	175	171	148	121	89
105	110	107	105	108	107	117	139	160	170	175	175	176	172	149	122	90
104	109	106	104	107	106	115	138	161	171	176	175	177	173	150	123	91
103	108	105	103	106	105	115	137	162	172	177	176	177	173	150	123	91
103	108	104	103	107	107	119	143	162	171	175	173	175	172	150	124	92
104	108	105	103	107	108	120	143	162	171	175	173	175	172	150	124	91
104	109	105	104	108	109	120	144	162	171	176	174	175	172	150	124	91
105	110	106	105	109	110	121	145	163	172	176	174	176	173	151	125	90
106	111	108	106	110	111	122	146	164	173	177	175	176	173	152	125	89
107	112	109	107	111	112	123	147	164	173	177	175	177	174	152	126	88
108	113	109	108	112	112	124	148	164	173	178	176	177	174	152	126	88
109	113	110	108	112	113	125	148	165	174	178	176	178	175	153	126	87
117	114	107	110	115	111	120	144	159	170	177	175	176	173	153	129	91
118	115	108	110	114	109	118	143	159	170	177	175	176	173	153	129	91
119	116	109	110	112	107	115	140	159	170	177	175	176	173	153	129	89
119	117	111	110	111	103	112	138	159	170	177	175	176	173	153	129	89
118	117	111	109	109	101	110	137	159	170	177	175	176	173	153	129	89
116	116	111	109	107	99	109	137	159	170	177	175	176	173	153	129	89
113	114	110	108	106	98	109	138	159	170	177	175	176	173	153	129	91

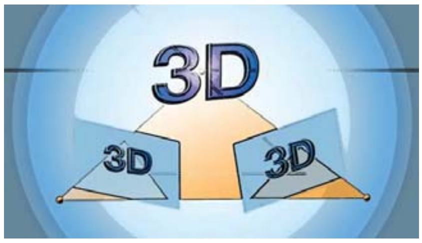
Zusammenfassung

- Bilder in Grautönen
- Bilder als Matrizen
- Bilder als glatte Funktionen

Martin Kleinsteuber: Computer Vision

Kap. 1 – Wissenswertes über Bilder

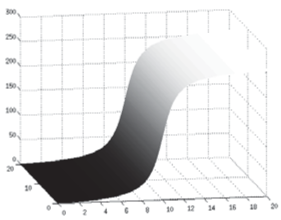
2. Bildgradient

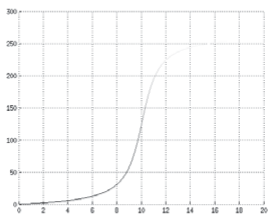


Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Der Gradient eines Bildes





Kanten sind starke lokale Änderungen der Intensität

Lokale Änderungen werden durch den Gradienten beschrieben

$$\nabla I(x,y) = \begin{bmatrix} \frac{d}{dx} I(x,y) \\ \frac{d}{dy} I(x,y) \end{bmatrix}$$

Video

Gradient eines Bildes

Wie schätzt man den Gradienten?

- Gegeben ist das Bild in diskreter Form $I \in \mathbb{R}^{m \times n}$
- Naiver Ansatz:

$$\frac{d}{dx} I(x,y) \approx I(x+1,y) - I(x,y)$$
$$\frac{d}{dy} I(x,y) \approx I(x,y+1) - I(x,y)$$

Video

Diskretes und kontinuierliches Signal

Interpolation

- Vom diskreten Signal $f[x] = S\{f(x)\}$ zum kontinuierlichen Signal $f(x)$

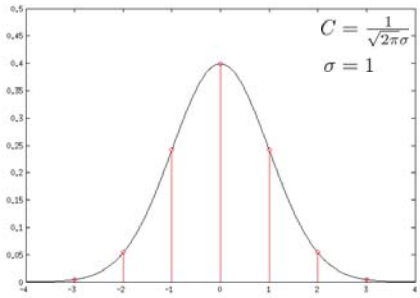
- Interpoliertes Signal ist Faltung der Abtastwerte mit dem Interpolationsfilter

$$f(x) \approx \sum_{k=-\infty}^{\infty} f[k]h(x-k) =: f[x] * h(x)$$

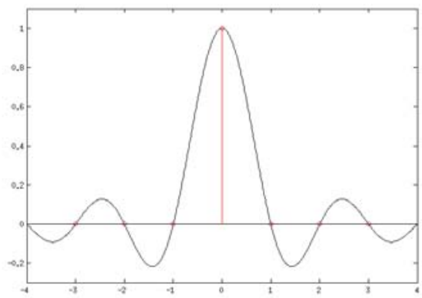
Video

Diskretes und kontinuierliches Signal
Interpolationsfilter

- Diskretes Signal: $f[x] = S\{f(x)\}$
- Kontinuierliches Signal: $f(x) \approx f[x] * h(x)$
- Gaußfilter: $h(x) = g(x)$
- Ideales Interpolationsfilter: $h(x) = \mathrm{sinc}(x)$
- Damit gilt: $f[x] * h(x) = f(x)$



$$g(x) = Ce^{\frac{-x^2}{2\sigma^2}}$$

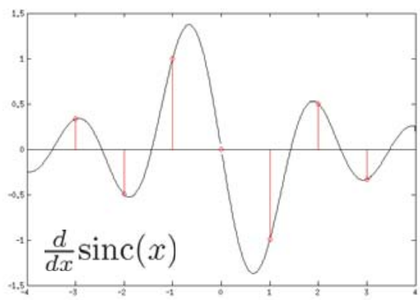


$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}, \quad \mathrm{sinc}(0) := 1$$

Die diskrete Ableitung
Mit Hilfe des rekonstruierten Signals

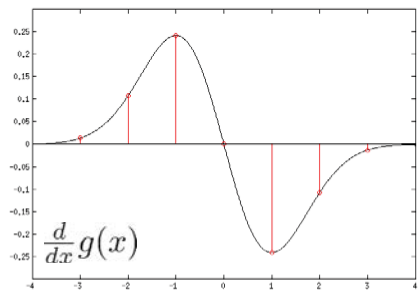
- Algorithmisch
  1. Rekonstruktion des kontinuierlichen Signals
  2. Ableitung des kontinuierlichen Signals
  3. Abtastung der Ableitung

- Sinc-Funktion
  - Langsames Abklingen



- Herleitung:
  - $f'(x) \approx \frac{d}{dx}(f[x] * h(x))$
    $= f[x] * h'(x)$
  - $f'[x] = f[x] * h'[x]$
    $$= \sum_{k=-\infty}^{\infty} f[x-k]h'[k]$$

- Gaußfilter
  - Schnelles Abklingen

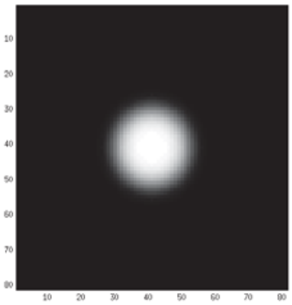


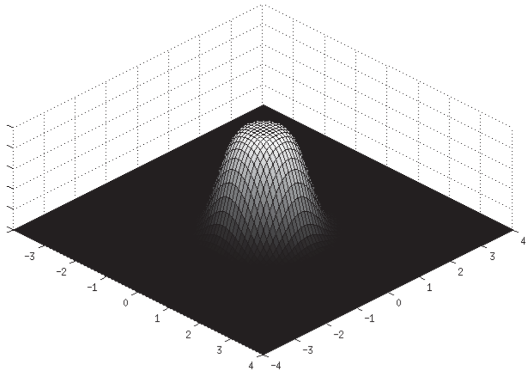
Zweidimensionale Rekonstruktion

Separables 2D-Gaußfilter

- 2D-Rekonstruktion: $I(x,y) \approx I[x,y] * h(x,y) = \sum\limits_{k=-\infty}^{\infty} \sum\limits_{l=-\infty}^{\infty} I[k,l]g(x-k)g(y-l)$

$$h(x,y) := g(x)g(y)$$



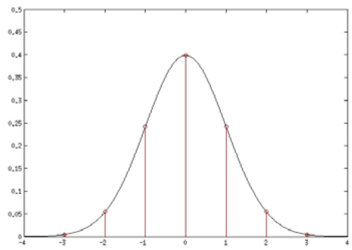


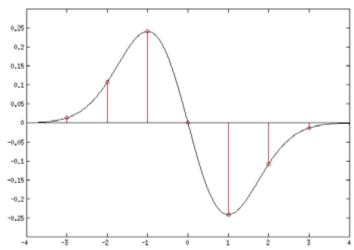
Video

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen
Technische Universität München

Zweidimensionale Ableitung

Ausnutzen der Separabilität





- Ableitung in x-Richtung

$$\begin{aligned}\frac{d}{dx}I(x,y) &\approx I[x,y] * \left(\frac{d}{dx}h(x,y)\right)\\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} I[k,l]g'(x-k)g(y-l)\end{aligned}$$

$$\begin{aligned}S\{\tfrac{d}{dx}I(x,y)\} &= I[x,y] * g'[x] * g[y]\\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} I[x-k,y-l]g'[k]g[l]\end{aligned}$$

Video

Endliche Approximation des Gaußfilters
Normierung des endlichen Filters

- In der Praxis wird die unendliche Summe durch wenige Summanden approximiert
- Wie wählt man eine geeignete Gewichtung C des Gaußfilters $g(x) = Ce^{\frac{-x^2}{2\sigma^2}}$?
- Interpoliertes Signal: $f(x) \approx \sum_{k=-\infty}^{\infty} f[k]g(x-k)$
- Abgetastetes interpoliertes Signal: $f[x] \approx \sum_{k=-\infty}^{\infty} f[x-k]g[k]$
- Approximation durch endliche Summe:

$$f[x] \approx \sum_{k=-n}^{n} f[x-k]g[k]$$

Video

Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

Endliche Approximation des Gaußfilters
Normierung des endlichen Filters

- Die endliche Approximation von $f[x]$ ist eine gewichtete Summe der Werte $f[x-n], \ldots, f[x+n]$ mit den Gewichten $g[n], \ldots, g[-n]$
- Normierungskonstante C so gewählt, dass sich alle Gewichte zu 1 addieren
- Wähle $C = \dfrac{1}{\sum_{-n \le k \le n} e^{\frac{-k^2}{2\sigma^2}}}$

Video

Sobel-Filter

Herleitung

- Approximation von $S\{\frac{d}{dx}I(x,y)\} = I[x,y] * g'[x] * g[y] = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} I[x-k, y-l]g'[k]g[l]$

  durch endliche Summe $\sum_{k=-1,0,1}\sum_{l=-1,0,1} I[x-k, y-l]g'[k]g[l]$
- Daraus folgt der Normierungsfaktor $C = \dfrac{1}{1+2e^{-\frac{1}{2\sigma^2}}}$
- Für die Wahl $\sigma = \sqrt{\frac{1}{2\log 2}}$ ergeben sich somit die Werte

$$g[-1] = \tfrac{1}{4};\ g[0] = \tfrac{1}{2};\ g[1] = \tfrac{1}{4}$$
$$g'[-1] = \tfrac{1}{2}\log 2;\ g'[0] = 0;\ g'[1] = -\tfrac{1}{2}\log 2 \quad (\tfrac{1}{2}\log 2 \approx 0.35)$$

Video

Sobel-Filter

Herleitung

- Aus praktischen Gründen sind ganzzahlige Filterkoeffizienten erwünscht
- Für das Detektieren von Intensitätsunterschieden ist ein Vielfaches des Gradienten ausreichend

$$\frac{1}{8}\log 2 \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

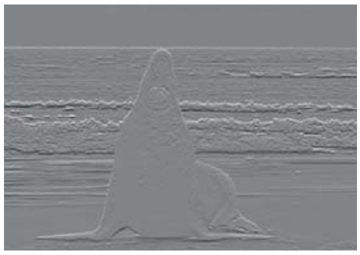
Ganzzahlige Approximation →

1	0	-1
2	0	-2
1	0	-1

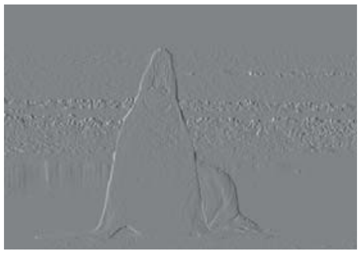
Horizontales Sobel-Filter

Beispiel

Sobel-Filterung



1	2	1
0	0	0
-1	-2	-1



1	0	-1
2	0	-2
1	0	-1

Zusammenfassung

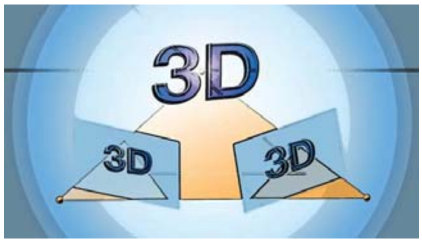
Bildgradient

- Der Bildgradient ist ein wichtiges Werkzeug für die Bestimmung von lokalen Intensitätsänderungen
- Diskrete Ableitung wird durch Differenzieren des interpolierten Signals berechnet
- Sobel-Filter sind ganzzahlige Approximation eines Vielfachen des Gradienten

Martin Kleinsteuber: Computer Vision

Kap. 1 – Wissenswertes über Bilder

3. Merkmalspunkte – Ecken und Kanten



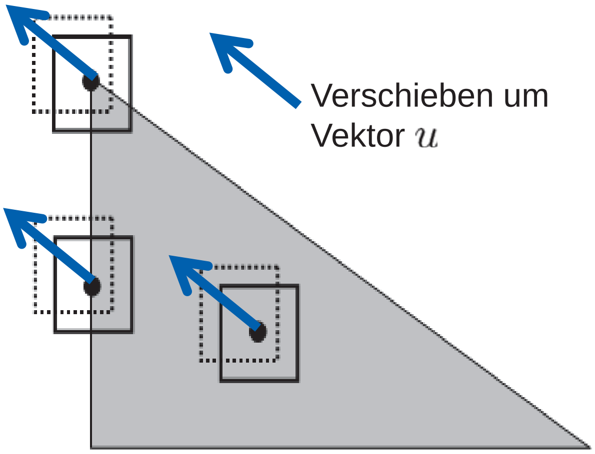
Ecken und Kanten...

...liefern markante Bildmerkmale

- Bestimmung von Konturen
- Berechnungen von Bewegungen in Bildsequenzen
- Schätzen von Kamerabewegung
- Registrierung von Bildern
- 3D-Rekonstruktion

Harris Ecken- und Kantendetektor

Änderung des Bildsegments in Abhängigkeit der Verschiebung



- Ecke: Verschiebung in jede Richtung bewirkt Änderung
- Kante: Verschiebung in jede bis auf genau eine Richtung bewirkt Änderung
- Homogene Fläche: Keine Änderung, egal in welche Richtung

Harris Ecken- und Kantendetektor

Formelle Beschreibung der Änderung

- Position im Bild: $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad I(x) = I(x_1, x_2)$
- Verschiebungsrichtung: $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Änderung des Bildsegments:

$$S(u) = \int_W \Big(I(x+u) - I(x)\Big)^2 dx$$

- Differenzierbarkeit von I:

$$\lim_{u \to 0} \frac{I(x+u) - I(x) - \nabla I(x)^\top u}{\|u\|} = 0$$

Harris Ecken- und Kantendetektor

Approximation der Änderung

- Folgerung aus Differenzierbarkeit: $I(x+u) - I(x) = \nabla I(x)^\top u + o(\|u\|)$
- Restterm $o(\|u\|)$ mit der Eigenschaft $\lim\limits_{u \to 0} o(\|u\|)/\|u\| = 0$
- Approximation für kleine Verschiebungen: $I(x+u) - I(x) \approx \nabla I(x)^\top u$
- Approximation der Änderung im Bildsegment:

$$S(u) = \int_W \Big(I(x+u) - I(x) \Big)^2 dx \approx \int_W \Big(\nabla I(x)^\top u \Big)^2 dx$$

Harris Ecken- und Kantendetektor

Die Harris-Matrix

- Ausmultiplizieren des Integrals:

$$\int_W \Big(\nabla I(x)^\top u \Big)^2 dx = u^\top \Big(\int_W \nabla I(x) \nabla I(x)^\top dx \Big) u$$

- Harris-Matrix: $G(x) = \int_W \nabla I(x) \nabla I(x)^\top dx$

$$\nabla I(x) \nabla I(x)^\top = \begin{bmatrix} (\frac{\partial}{\partial x_1} I(x))^2 & \frac{\partial}{\partial x_1} I(x) \frac{\partial}{\partial x_2} I(x) \\ \frac{\partial}{\partial x_2} I(x) \frac{\partial}{\partial x_1} I(x) & (\frac{\partial}{\partial x_2} I(x))^2 \end{bmatrix}$$

- Approximative Änderung des Bildsegments:

$$S(u) \approx u^\top G(x) u$$

Exkurs: Lineare Algebra

Positiv definite und positiv semidefinite Matrizen

- Definition. Eine reelle symmetrische Matrix $A = A^\top$ heißt
  - positiv definit, falls $x^\top Ax > 0, \quad x \neq 0$
  - positiv semidefinit, falls $x^\top Ax \geq 0$
- Beispiele
  - Die Null-Matrix ist positiv semidefinit, aber nicht positiv definit.
  - Die Einheitsmatrix ist positiv definit.
  - Jede positiv definite Matrix ist auch positiv semidefinit.
  - $G(x)$ ist positiv semi-definit. Warum?

Exkurs: Lineare Algebra

Positiv definite und positiv semidefinite Matrizen

- Satz. Für reelle symmetrische Matrizen $A = A^\top$ ist gleichbedeutend:
  - A positiv (semi-) definit
  - Alle Eigenwerte von A sind größer als Null (größer oder gleich Null)
- Eigenwertzerlegung reeller symmetrischer Matrizen.
  Jede reelle symmetrische Matrix $A = A^\top$ kann zerlegt werden in ein Produkt $A = V\Lambda V^\top$ mit $VV^\top = I$ und einer Diagonalmatrix Λ, auf deren Diagonale die Eigenwerte von A stehen. Die Spalten von V sind die zugehörigen Eigenvektoren.

Harris Ecken- und Kantendetektor
Eigenwertzerlegung

- Eigenwertzerlegung der Harris-Matrix:

$$G(x) = \int_W \nabla I(x) \nabla I(x)^\top dx = V \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} V^\top$$

mit $VV^\top = I_2$ und den Eigenwerten $\lambda_1 \geq \lambda_2 \geq 0.$

- Änderung in Abhängigkeit der Eigenvektoren: $V = [v_1, v_2]$

$$S(u) \approx u^\top G(x) u = \lambda_1 (u^\top v_1)^2 + \lambda_2 (u^\top v_2)^2$$

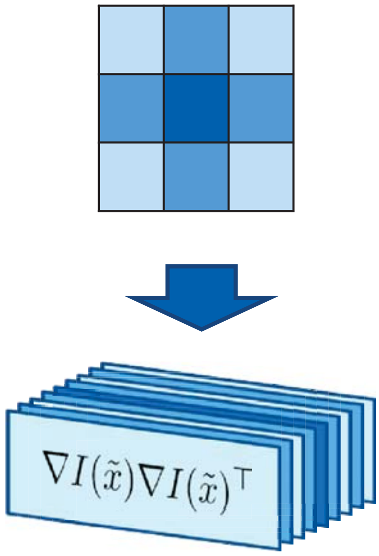
Harris Ecken- und Kantendetektor
Art des Merkmals in Abhängigkeit der Eigenwerte

$$S(u) \approx u^\top G(x) u = \lambda_1 (u^\top v_1)^2 + \lambda_2 (u^\top v_2)^2$$

- Beide Eigenwerte positiv
  - $S(u) > 0$ für alle u (Änderung in jede Richtung)
  - Untersuchtes Bildsegment enthält eine Ecke
- Ein Eigenwert positiv, ein Eigenwert gleich null
  - $S(u) \begin{cases} = 0, & \text{falls } u = r v_2 \\ > 0, & \text{sonst} \end{cases}$ (Keine Änderung nur in Richtung des Eigenvektors zum Eigenwert 0)
  - Untersuchtes Bildsegment enthält eine Kante
- Beide Eigenwerte gleich null
  - $S(u) = 0$ für alle u (Keine Änderung, egal in welche Richtung)
  - Untersuchtes Bildsegment ist eine homogene Fläche

Praktische Realisierung des Harris-Detektors

Berechnung der Harris-Matrix



- Approximiere $G(x)$ durch endliche Summe

$$G(x) = \int_W \nabla I(x)\nabla I(x)^\top dx \approx \sum_{\tilde{x}\in W(x)} \nabla I(\tilde{x})\nabla I(\tilde{x})^\top$$

- Gewichtete Summe in Abhängigkeit der Position von $\tilde{x}$

$$G(x) \approx \sum_{\tilde{x}\in W(x)} w(\tilde{x})\nabla I(\tilde{x})\nabla I(\tilde{x})^\top$$

- Gewichte $\omega(\tilde{x}) > 0$ betonen Einfluss der zentralen Pixel

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Praktische Realisierung des Harris-Detektors

Eigenwerte

- In der Realität nehmen Eigenwerte nie genau den Wert Null an, z.B. auf Grund von Rauschen, diskreter Abtastung und numerischen Ungenauigkeiten

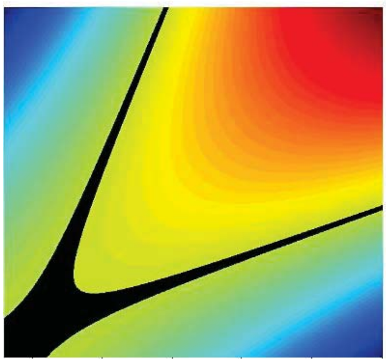
- Charakteristik in der Praxis
  - Ecke: zwei große Eigenwerte
  - Kante: ein großer Eigenwert, ein kleiner Eigenwert
  - Homogene Fläche: zwei kleine Eigenwerte

- Entscheidung mittels empirischer Schwellwerte

Exkurs: Lineare Algebra
Eigenwerte, Determinante, Spur

- Zusammenhang von Eigenwerten, Determinante und Spur einer Matrix.
  Seien $\lambda_1, \ldots, \lambda_n$ Eigenwerte einer Matrix G.
  Dann gilt
  - $\det G = \prod_{i=1}^{n} \lambda_i$ (Determinante ist das Produkt der Eigenwerte)
  - $\operatorname{tr} G = \sum_{i=1}^{n} \lambda_i$ (Spur ist die Summe der Eigenwerte)

Praktische Realisierung des Harris-Detektors
Ein einfaches Kriterium für Ecken und Kanten



- Betrachte die Größe $H := \det(G) - k(\operatorname{tr}(G))^2$

  $$H = (1-2k)\lambda_1\lambda_2 - k(\lambda_1^2 + \lambda_2^2)$$

- Ecke (beide Eigenwerte groß)
  - H größer als ein positiver Schwellwert
- Kante (ein Eigenwert groß, ein Eigenwert klein)
  - H kleiner als ein negativer Schwellwert
- Homogene Fläche (beide Eigenwerte klein)
  - H betragsmäßig klein

Zusammenfassung

Harris-Detektor zur Bestimmung von Merkmalspunkten

- Auswertung der (approximierten) Harris-Matrix

$$G(x) \approx \sum_{\tilde{x} \in W(x)} w(\tilde{x}) \nabla I(\tilde{x}) \nabla I(\tilde{x})^\top$$

- Eigenwertzerlegung von $G(x)$ liefert auch Info über Richtung etwaiger Kanten
- Effiziente Implementierung mit Hilfe des Ausdrucks

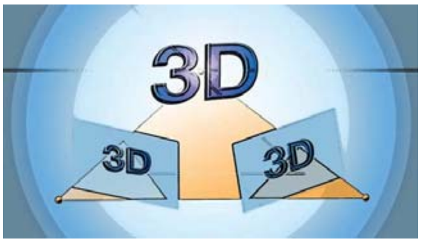
$$H := \det(G) - k(\mathrm{tr}(G))^2$$

- Entscheidung mittels Schwellwerten
  - Ecke: $0 < \tau_+ < H$
  - Kante: $H < \tau_- < 0$
  - Homogene Fläche: $\tau_- < H < \tau_+$

Martin Kleinsteuber: Computer Vision

Kap. 1 – Wissenswertes über Bilder

4. Korrespondenzschätzung für Merkmalspunkte



Korrespondenzschätzung

Problemstellung

- Gegeben sind zwei Bilder $I_1 : \Omega_1 \to \mathbb{R}, I_2 : \Omega_2 \to \mathbb{R}$ derselben 3D-Szene
- Finde Paare von Bildpunkten $(x^{(i)}, y^{(i)}) \in \Omega_1 \times \Omega_2$, die zu gleichen 3D-Punkten korrespondieren.

Korrespondenzschätzung

Problemstellung

- In dieser Session: Korrespondenzen für Merkmalspunkte in I_1 und I_2
- Habe Merkmalspunkte $\{x_1, \ldots x_n\} \subset \Omega_1$ und $\{y_1, \ldots y_n\} \subset \Omega_2$
- Finde passende Paare von Merkmalspunkten

Naive Lösung des Problems

Sum of squared differences (SSD)

- Betrachte Bildausschnitte V_i um x_i und W_i um y_i in Matrixdarstellung und vergleiche die Intensitäten

d(,)

Sum of Squared Differences (SSD)

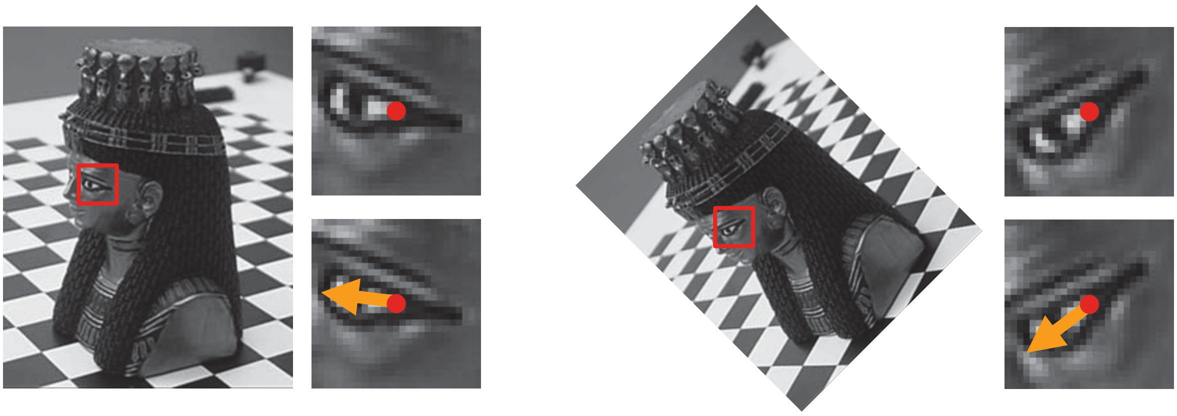
Formelle Beschreibung

- Ein Kriterium: $d(V,W) = \|V - W\|_F^2$
- Dabei ist $\|A\|_F^2 = \sum_{kl} A_{kl}^2$ die quadrierte Frobeniusnorm
- Finde zu V_i das W_j mit $j = \arg \min_{k=1,\ldots,n} d(V_i, W_k)$
- Annahme: Wenn W_j zu V_i passt, dann auch umgekehrt

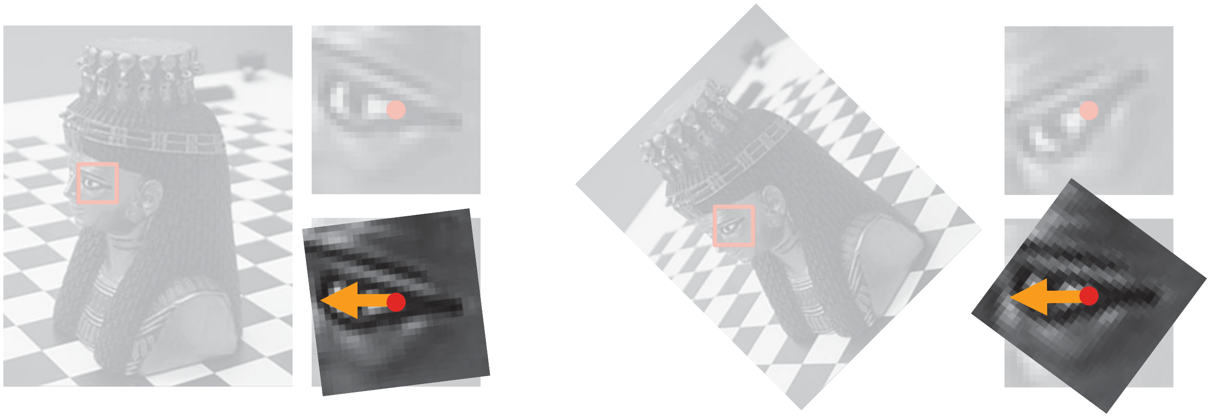
Schwachpunkte der SSD-Methode
Änderungen der Beleuchtung oder Drehungen

- Normierung von Intensität und Orientierung benötigt!

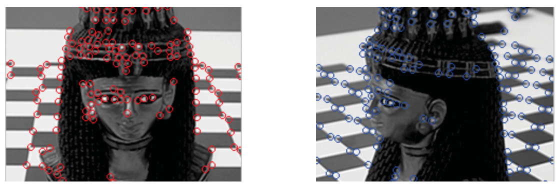
Rotationsnormierung
mittels Gradientenrichtung



Rotationsnormierung
mittels Gradientenrichtung



Rotationsnormierung
mittels Gradientenrichtung



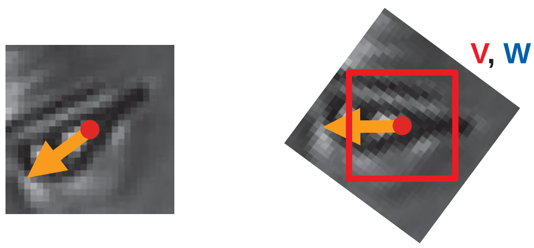
- Vorverarbeitung:

  1. Bestimme Gradienten in allen Merkmalspunkten

  2. Rotiere Regionen um Merkmalspunkte so, dass Gradient in eine Richtung zeigt.

  3. Extrahiere V, W aus rotierten Regionen

$$\theta = \arctan\left(\frac{\frac{\partial I}{\partial x_2}}{\frac{\partial I}{\partial x_1}}\right)$$



Bias-and-Gain-Modell

Modellierung von Kontrast und Helligkeit

- Skalierung der Intensitätswerte (Gain) mit α
- Verschiebung der Intensitätswerte (Bias) mit β

- Gain-Modell: $W \approx \alpha V$
- Bias-Modell: $W \approx V + \beta \mathbb{1}\mathbb{1}^\top$

$$\mathbb{1} = (1, \dots 1)^\top$$

- Bias-and-Gain Modell: $W \approx \alpha V + \beta\, \mathbb{1}\mathbb{1}^\top$

Skalierung bewirkt Kontraständerung

Verschiebung bewirkt Helligkeitsänderung

Bias-and-Gain-Modell

Berechnung des Mittelwerts

- Mittelwertbildung der Intensität

$$\begin{aligned}
\overline{W} &= \tfrac{1}{N}\left(\mathbb{1}\mathbb{1}^\top W \mathbb{1}\mathbb{1}^\top\right) \\
&\approx \tfrac{1}{N}\left(\mathbb{1}\mathbb{1}^\top \left(\alpha V + \beta \mathbb{1}\mathbb{1}^\top\right) \mathbb{1}\mathbb{1}^\top\right) \\
&= \alpha \tfrac{1}{N}\left(\mathbb{1}\mathbb{1}^\top V \mathbb{1}\mathbb{1}^\top\right) + \beta \mathbb{1}\mathbb{1}^\top \\
&= \alpha \overline{V} + \beta \mathbb{1}\mathbb{1}^\top
\end{aligned}$$

- Subtraktion der Mittelwertmatrix

$$\begin{aligned}
W - \overline{W} &\approx \alpha V + \beta \mathbb{1}\mathbb{1}^\top - \left(\alpha \overline{V} + \beta \mathbb{1}\mathbb{1}^\top\right) \\
&= \alpha \left(V - \overline{V}\right)
\end{aligned}$$

Bias-and-Gain-Modell

Berechnung der Standardabweichung

- Standardabweichung der Intensität

$$
\begin{aligned}
\sigma(W) &= \sqrt{\tfrac{1}{N-1}\|W-\overline{W}\|_F^2} \\
&= \sqrt{\tfrac{1}{N-1}\operatorname{tr}\left(\left(W-\overline{W}\right)^\top\left(W-\overline{W}\right)\right)} \\
&\approx \sqrt{\tfrac{1}{N-1}\operatorname{tr}\left(\alpha\left(V-\overline{V}\right)^\top\alpha\left(V-\overline{V}\right)\right)} \\
&= \alpha\,\sigma(V)
\end{aligned}
$$

Bias-and-Gain-Modell

Kompensation von Bias und Gain

- Normalisierung der Bildsegmente durch
  1. Subtraktion des Mittelwertes
  2. Division durch Standardabweichung

$$
\begin{aligned}
W_n &:= \frac{1}{\sigma(W)}(W-\overline{W}) \\
&\approx \frac{1}{\alpha\,\sigma(V)}\left(\alpha\left(V-\overline{V}\right)\right) \\
&= \frac{1}{\sigma(V)}(V-\overline{V}) \\
&=: V_n
\end{aligned}
$$

Normalized Cross Correlation (NCC)

Herleitung aus SSD

- SSD von zwei normalisierten Bildsegmenten
  $\|V_n - W_n\|_F^2 = 2(N-1) - 2\mathrm{tr}(W_n^\top V_n)$
- Die Normalized Cross Correlation der beiden Bildsegmente ist definiert als $\frac{1}{N-1}\mathrm{tr}(W_n^\top V_n)$
- Es gilt $-1 \leq \mathrm{NCC} \leq 1$
- Zwei normalisierte Bildsegmente sind sich ähnlich, wenn
  - SSD klein (wenig Unterschiede)
  - NCC nahe bei +1 (hohe Korrelation)

Zusammenfassung

Korrespondenzschätzung von Merkmalspunkten

- Finde Merkmale in Bild 1 und Bild 2
- Kompensiere Rotation durch Ausrichten des Gradienten für jeden Merkmalspunkt
- Extrahiere Bildsegment um jeden Merkmalspunkt
- Beleuchtungskompensation durch Normierung der Bildsegmente
- Vergleiche die normalisierten Bildsegmente durch SSD oder NCC

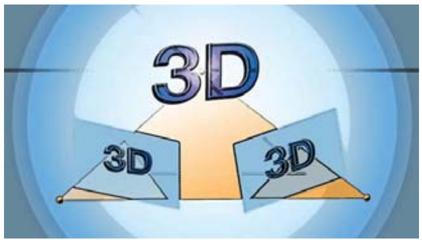
Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Martin Kleinsteuber: Computer Vision

Kapitel 2 – Bildentstehung

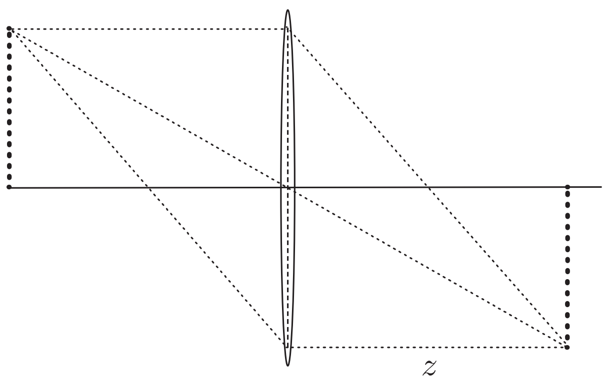
1. Das Lochkameramodell



Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

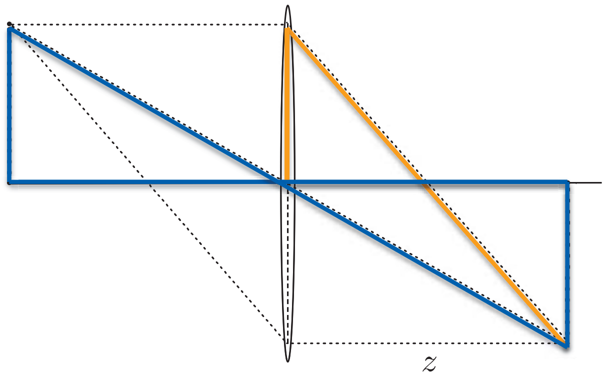
Technische Universität München

Abbildung durch eine dünne Linse



- Strahlen parallel zur optischen Achse werden so gebrochen, dass sie durch die Brennebene gehen
- Strahlen durch das optische Zentrum werden nicht abgelenkt

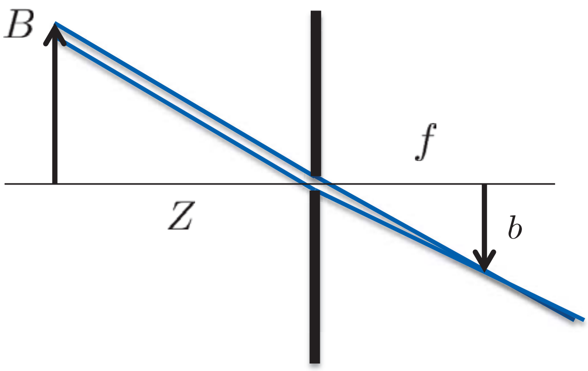
Abbildung durch eine dünne Linse
Herleitung der Abbildungsgleichung



- $$\left|\frac{b}{B}\right| = \frac{z-f}{f}.$$
- $$\left|\frac{z}{Z}\right| = \left|\frac{b}{B}\right|$$
- Gleichung für dünne Linsen

$$\frac{1}{|Z|} + \frac{1}{z} = \frac{1}{f}$$

Abbildung durch eine Lochkamera
Idealisierte Annahmen

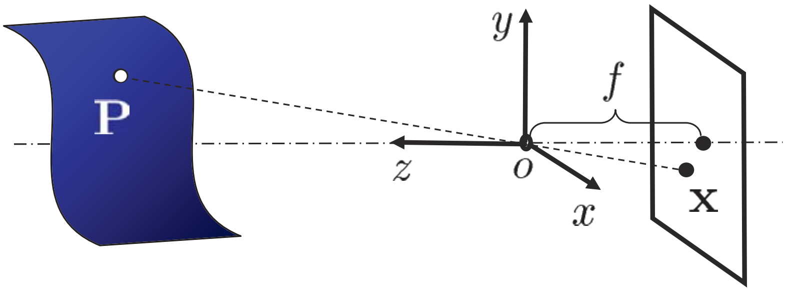


- Sehr kleine Öffnung vor der Linse
- Beliebig großer Blickwinkel
- Bild wird scharf auf Brennebene abgebildet
- Es gilt:

$$b = -\frac{fB}{Z}$$

Abbildung durch eine Lochkamera

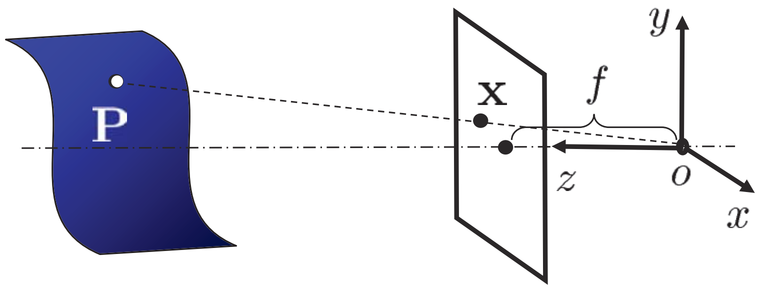
Abbildung eines Punktes im Raum auf die Brennebene



- Koordinaten von P bzgl. optischem Zentrum der Kamera seien (X, Y, Z)
- Dann sind die Koordinaten des Bildpunktes $\left(-\frac{fX}{Z}, -\frac{fY}{Z}, -f\right)$

Frontales Lochkameramodell

Projektion eines Punktes

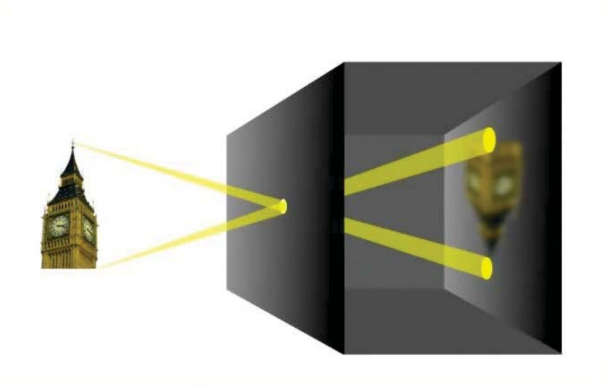


- Koordinaten des Bildpunktes beim frontalen Lochkameramodell $\left(\frac{fX}{Z}, \frac{fY}{Z}, f\right)$
- Ideale perspektivische Projektion:

$$\pi : \mathbb{R}^3 \setminus \{(X,Y)\text{-Ebene}\} \to \mathbb{R}^2, \qquad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \mapsto \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Abbildung durch eine Lochkammera

Idealisierte Annahmen



Quelle: Wikipedia.de, 2013

- Dünne Linse
- Kleine Blende
- Beliebig großer Blickwinkel
- Perspektivische Projektion:

$$\pi : \mathbb{R}^3 \setminus \{(X,Y)\text{-Ebene}\} \to \mathbb{R}^2, \qquad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \mapsto \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Zusammenfassung

- Lochkameramodell
- Ideale perspektivische Projektion:
  die Abbildung der Koordinaten des 3D-Punktes auf die 2D-Koordinaten in der Brennebene

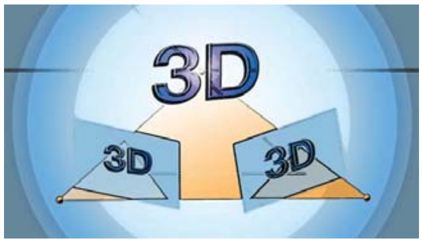
Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Martin Kleinsteuber: Computer Vision

Kapitel 2 – Bildentstehung

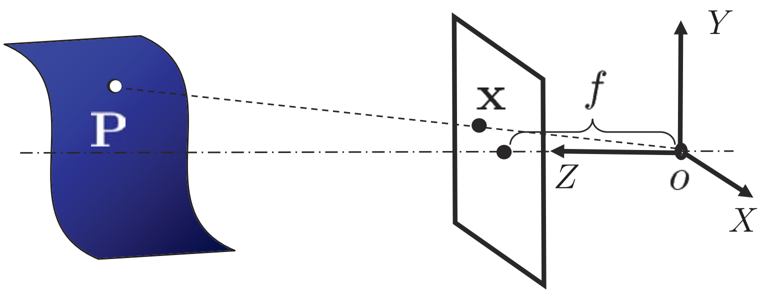
2. Homogene Koordinaten



Wiederholung: Lochkameramodell

Bildpunkte und Geraden

- Alle Punkte auf einer Geraden durch das optische Zentrum werden auf denselben Bildpunkt abgebildet
- Umgekehrt existiert zu jedem Bildpunkt genau eine Gerade



Der projektive Raum

- Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ nennen wir zueinander **äquivalent**, falls ein $\lambda \neq 0$ existiert mit $\mathbf{x} = \lambda \mathbf{y}$. In diesem Fall schreiben wir

$$\mathbf{x} \sim \mathbf{y}$$

- Eine Gerade durch $\mathbf{x} \neq 0$ können wir nun beschreiben als **Äquivalenzklasse**

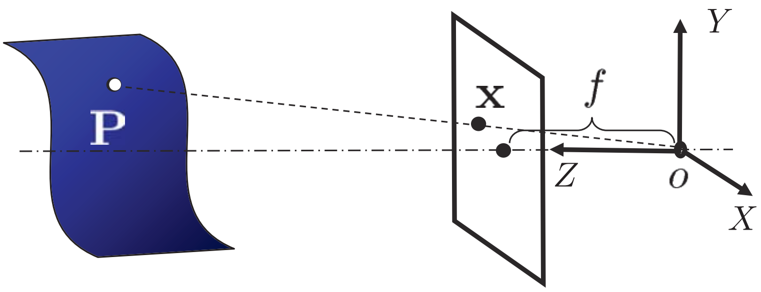
$$[\mathbf{x}] := \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} \sim \mathbf{x}\}$$

- Die Menge aller Geraden im $\mathbb{R}^{n+1}$ heißt **projektiver Raum**

$$\mathbb{P}_n = \{[\mathbf{x}] \mid \mathbf{x} \in \mathbb{R}^{n+1} \setminus \{0\}\}$$

Homogene Koordinaten

- Normiert man die Längeneinheit auf die Brennweite, ist die Bildebene gegeben durch



$$BE = \left\{ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3 \mid Z = 1 \right\}$$

- Der Vektor $\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$ heißt die **homogenen Koordinaten** von $\begin{bmatrix} X \\ Y \end{bmatrix}$

Homogene Koordinaten

- Allgemeine Definition: $\mathbf{x} := (X_1, \ldots, X_n)^\top \in \mathbb{R}^n$
  Dann heißt

$$\mathbf{x}^{(\mathrm{hom})} := (X_1, \ldots, X_n, 1)^\top \in \mathbb{R}^{n+1}$$

  die homogenen Koordinaten von $\mathbf{x}$

Zusammenfassung

- Äquivalente Vektoren unterscheiden sich nur durch Multiplikation mit einer Zahl ungleich Null.

- Homogene Koordinaten eines Vektors erhält man durch Hinzufügen einer weiteren Koordinate mit dem Wert Eins.

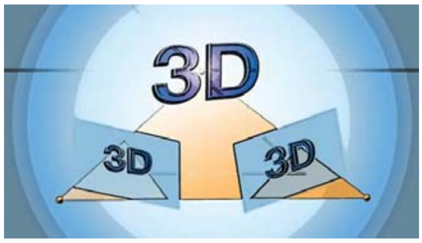
Fachgebiet Geometrische Optimierung und Maschinelles Lernen

Technische Universität München

Martin Kleinsteuber: Computer Vision

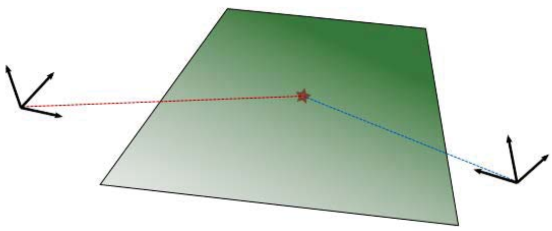
Kapitel 2 – Bildentstehung

3. Euklidische Bewegungen

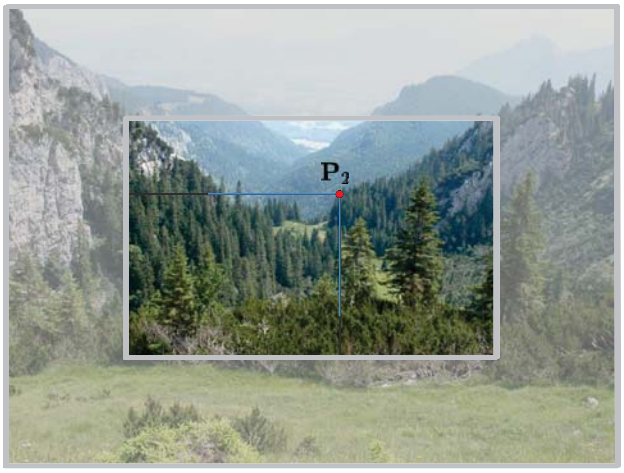


Motivation

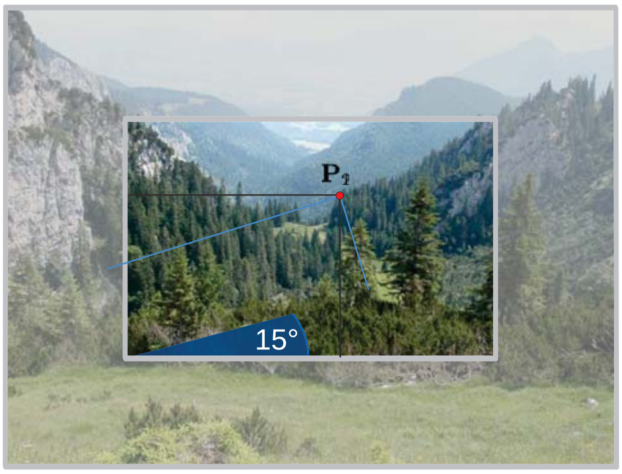
- Zwei Bilder einer 3D-Szene aus unterschiedlichen Positionen
- Positionswechsel bestehen aus Rotation der Kamera und anschließender Translation
- Beschreibung der Kamerabewegung in Form von Koordinatenänderung eines festen Raumpunktes



Visualisierung
Translation



Visualisierung
Rotation



Euklidische Bewegungen
Allgemein

- Die Matrizen $O(n) := \{O \in \mathbb{R}^{n\times n} \mid O^\top O = I_n\}$ heißen orthogonale Matrizen
- Rotationen im $\mathbb{R}^n$ werden beschrieben durch die speziellen orthogonalen Matrizen
$$SO(n) := \{R \in \mathbb{R}^{n\times n} \mid R^\top R = I_n, \det(R) = 1\}$$
- Euklidische Bewegungen durch die Koordinatenänderung
$$g_{R,T} : \mathbb{R}^n \to \mathbb{R}^n, \qquad \mathbf{P} \mapsto R\mathbf{P} + T$$
- Seien $\mathbf{P}_1$ die Koordinaten eines Punktes bzgl. CF 1 und $\mathbf{P}_2$ die Koordinaten des selben Punktes bzgl. des bewegten CF 2, dann ist $\mathbf{P}_2 = R\mathbf{P}_1 + T$

Euklidische Bewegungen
In homogenen Koordinaten

- In homogenen Koordinaten als Matrix-Vektor-Multiplikation

$$M = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \quad , R \in SO(n)\, , T \in \mathbb{R}^n$$

$$\mathbf{P}_2^{(\mathrm{hom})} = M\,\mathbf{P}_1^{(\mathrm{hom})}$$

- Spezielle Euklidische Gruppe

$$SE(n) = \left\{ M = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \middle| R \in SO(n)\, , T \in \mathbb{R}^n \right\} \subset \mathbb{R}^{(n+1)\times(n+1)}$$

Euklidische Bewegungen

Eigenschaften

- Verknüpfungen zweier eukl. Bewegungen ist wieder eine eukl. Bewegung

$$M_1 M_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1 R_2 & R_1 T_2 + T_1 \\ 0 & 1 \end{bmatrix}$$

- Euklidische Bewegungen sind invertierbar:

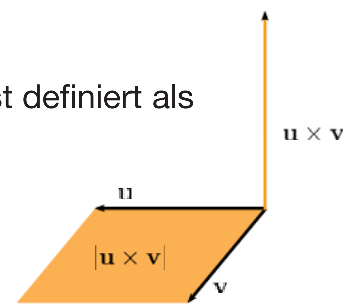
$$M^{-1} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^\top & -R^\top T \\ 0 & 1 \end{bmatrix}$$

Das Kreuzprodukt

Definition

- Das Kreuzprodukt zwischen den Vektoren $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ ist definiert als

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} \in \mathbb{R}^3$$



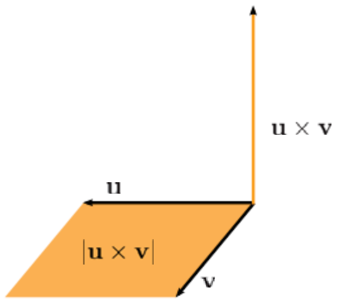
- Es gilt: $\langle \mathbf{u} \times \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u} \times \mathbf{v}, \mathbf{v} \rangle = 0 \quad \text{und} \quad \mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$

Das Kreuzprodukt

Matrix-Vektor-Multiplikation

- Das Kreuzprodukt lässt sich schreiben als

$$\mathbf{u} \times \mathbf{v} = \hat{\mathbf{u}}\mathbf{v}, \quad \text{mit} \quad \hat{\mathbf{u}} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$



- Es gilt (für A invertierbar):

$$\widehat{A\mathbf{v}} = \det(A) A^{-\top} \hat{\mathbf{v}} A^{-1}$$

Euklidische Bewegungen

Eigenschaften Orthogonaler Transformationen

- Orthogonale Transformationen erhalten Skalarprodukt $\langle \mathbf{u}, \mathbf{v} \rangle = \langle O\mathbf{u}, O\mathbf{v} \rangle$
- Euklidische Transformationen erhalten Abstand
- Spezielle orthogonale Transformationen erhalten das Kreuzprodukt

$$R\mathbf{u} \times R\mathbf{v} = R(\mathbf{u} \times \mathbf{v})$$

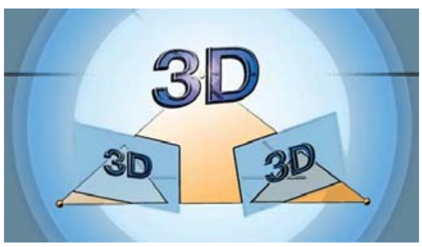
Zusammenfassung

- Kamerabewegungen durch Koordinatenänderung eines festen Raumpunktes
- Koordinatenänderung durch euklidische Bewegungen
- Rotationen durch spezielle orthogonale Matrizen
- Euklidische Bewegung in homogenen Koordinaten durch Matrix-Vektor-Multiplikation

Martin Kleinsteuber: Computer Vision

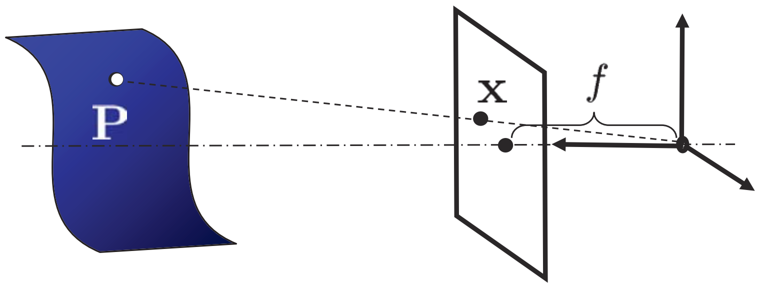
Kapitel 2 – Bildentstehung

4. Perspektivische Projektion mit kalibrierter Kamera



Perspektivische Projektion

Lochkameramodell



- Abbildung eines Punktes auf den Bildpunkt $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$

Perspektivische Projektion

- Die Abhängigkeit der homogenen Koordinaten erhalten wir aus

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Perspektivische Projektion

Mit Brennweitenmatrix und generischer Projektionsmatrix

- Definiere $K_f := \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Pi_0 := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
- Betrachte 3D-Punkt $\mathbf{P}^{(\mathrm{hom})} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$ und $\mathbf{x}^{(\mathrm{hom})} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
- Die Perspektivische Projektion ist somit

$$\mathbf{x}^{(\mathrm{hom})} \sim K_f \Pi_0 \mathbf{P}^{(\mathrm{hom})}$$

Die ideale Kamera

Abbildung von Weltkoordinaten auf Bildkoordinaten

- Transformation der homogenen Koordinaten $\mathbf{P}^{(\mathrm{hom})}$ bei euklidischer Bewegung der Kamera

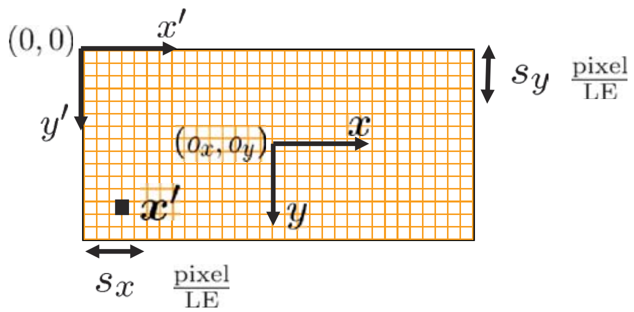
$$\mathbf{P}^{(\mathrm{hom})} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mathbf{P_0}^{(\mathrm{hom})} \quad R \in SO(3), \quad T \in \mathbb{R}^3$$

- Perspektivische Projektion mit euklidischer Transformation

$$\mathbf{x}^{(\mathrm{hom})} \sim K_f \Pi_0 \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mathbf{P_0}^{(\mathrm{hom})}$$

Sensorparameter

Transformation von Bildkoordinaten in Pixelkoordinaten



1. Spezifizieren der Längeneinheiten (LE)

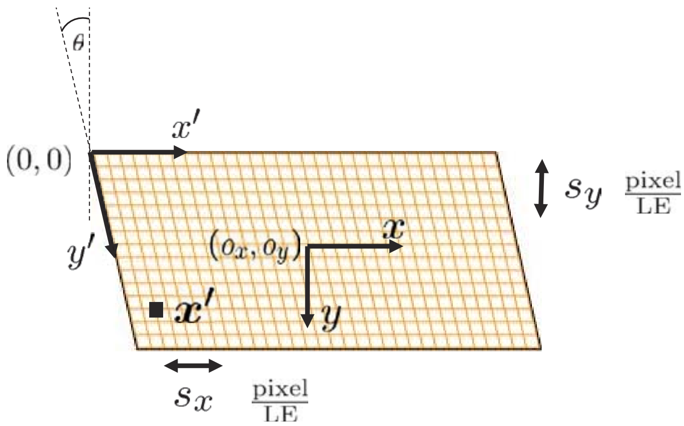
- $x_s = s_x x$, gleiches mit $y_s = s_y y$
- $s_x = s_y$ bedeutet quadratische Pixel

2. Justieren des Ursprungs

- Pixelkoordinaten $x' = x_s + o_x$, gleiches mit $y' = y_s + o_y$

Sensorparameter

Lineare Transformation von Bildkoordinaten in Pixelkoordinaten



3. Einführen eines Scherungsfaktors s_θ

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{:=K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\mathbf{x}' = K_s \mathbf{x}$$

$$\mathbf{x} = K_s^{-1} \mathbf{x}'$$

Kalibrierungsmatrix

Zusammenführen von Brennweite und Sensorparametern

- Pixelkoordinaten
- Perspektivische Projektion

$$\mathbf{x}' = K_s \mathbf{x} \qquad \mathbf{x} \sim K_f \Pi_0 \mathbf{P}$$

$$\mathbf{x}' \sim K_s K_f \Pi_0 \mathbf{P} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{:=K} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Kalibrierungsmatrix (intrinsische Kameraparameter)

$$K = K_s K_f = \begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

Zusammenfassung

- Abbildung von Weltkoordinaten auf Pixelkoordinaten durch

$$\mathbf{x}' \sim K \Pi_0 \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mathbf{P}_0$$

- Intrinsische Parameter K:
- Extrinsische Parameter: Position der Kamera
- Umrechnung von idealen Bildkoordinaten und Pixelkoordinaten

$$\mathbf{x}' = K_s \mathbf{x}$$
$$\mathbf{x} = K_s^{-1} \mathbf{x}'$$

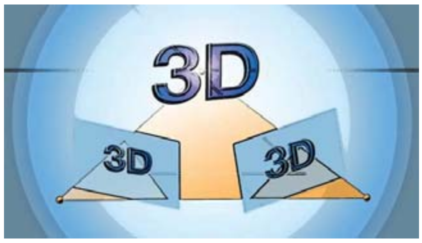
Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

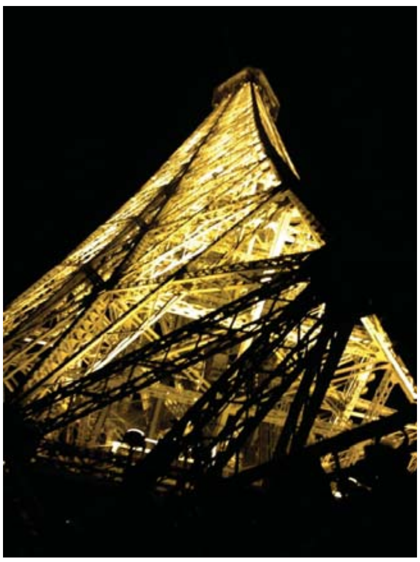
Martin Kleinsteuber: Computer Vision

Kapitel 2 – Bildentstehung

5. Bild, Urbild und Cobild



Motivation



Wiederholung

Projektion von Geraden

- Gerade im Raum in homogenen Koordinaten

$$L^{(\text{hom})} = \left\{ \mathbf{P}_0^{(\text{hom})} + \lambda \left[v_1, v_2, v_3, 0 \right]^\top \mid \lambda \in \mathbb{R} \right\}$$

- Perspektivische Projektion einer Geraden – Beispiel

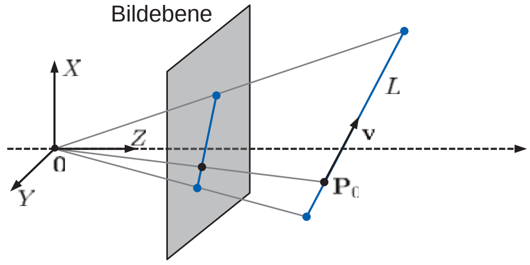


Bild und Urbild

Definitionen

- Das Bild eines Punktes bzw. einer Geraden ist deren perspektivische Projektion
$\Pi_0 \mathbf{P}^{(\text{hom})}$ bzw. $\Pi_0 L^{(\text{hom})}$

- Das Urbild eines Punktes $\mathbf{P}$ bzw. einer Geraden L sind alle Punkte im Raum, die auf den gleichen Bildpunkt bzw. auf die gleiche Gerade in der BE projiziert werden.

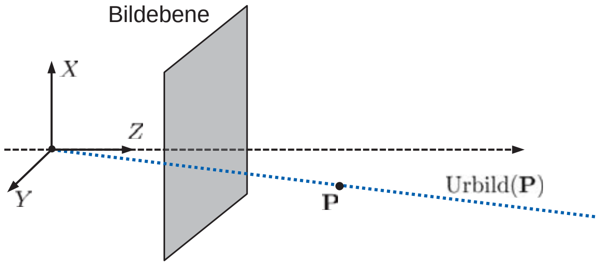
$$\text{Urbild}(\mathbf{P}) = \{ \mathbf{Q} \in \mathbb{R}^3 \mid \Pi_0 \mathbf{Q}^{(\text{hom})} \sim \Pi_0 \mathbf{P}^{(\text{hom})} \}$$

$$\text{Urbild}(L) = \bigcup_{\mathbf{P} \in L} \text{Urbild}(\mathbf{P})$$

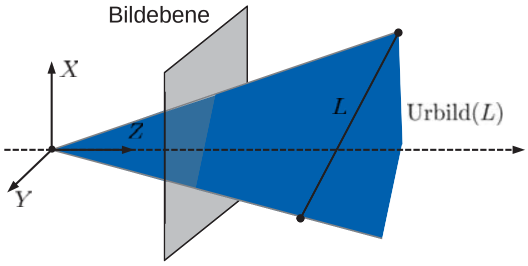
Bild und Urbild

Eigenschaften

- Urbilder von Punkten sind Geraden durch den Ursprung

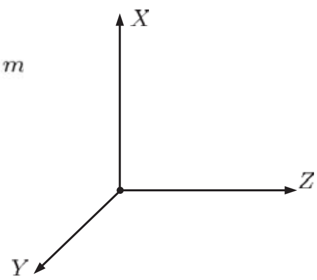


- Urbilder von Geraden sind Ebenen durch den Ursprung



Exkurs: Lineare Algebra

Erzeugnis und Orthogonales Komplement



- Erzeugnis von Spaltenvektoren $\mathbf{a}_i$ einer Matrix $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m] \in \mathbb{R}^{n \times m}$

$$\mathrm{span}(\mathbf{A}) = \left\{ \sum_{i=1}^{m} \lambda_i \mathbf{a}_i \,|\, \lambda_i \in \mathbb{R} \right\}$$

- Das Erzeugnis ist ein Untervektorraum vom $\mathbb{R}^n$

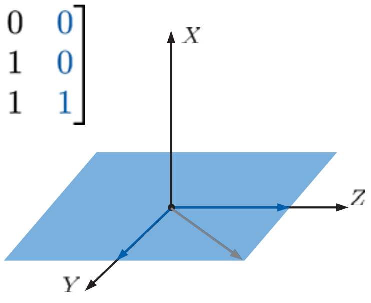
- Orthogonales Komplement eines Untervektorraums

$$\mathrm{span}(\mathbf{A})^{\perp} = \{ \mathbf{v} \in \mathbb{R}^n \,|\, \langle \mathbf{a}_i, \mathbf{v} \rangle = 0, i = 1, \dots, m \}$$

Exkurs: Lineare Algebra

Dimension, Kern und Rang

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

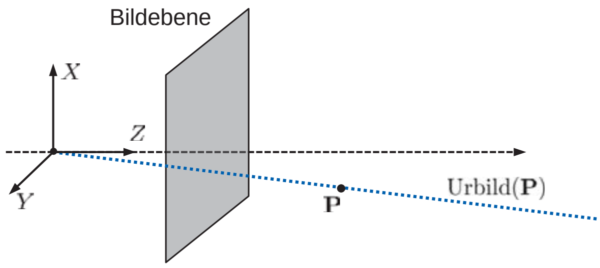


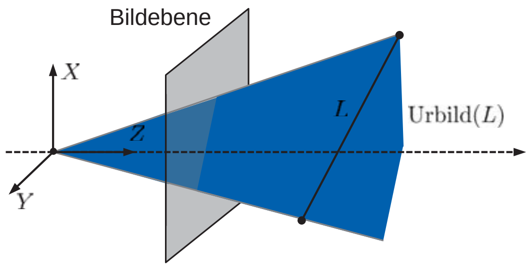
- Dimension eines Untervektorraums:
  Anzahl der Elemente eines minimalen Erzeugendensystems
- Rang einer Matrix ist Dimension des Erzeugnisses der Spalten
- Kern einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$
  $$\mathrm{Ker}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}\mathbf{x} = 0\}$$
- $\mathrm{Ker}(\mathbf{A})$ ist Untervektorraum im $\mathbb{R}^m$
- Dimensionssatz:
  $$\mathbf{A} \in \mathbb{R}^{n \times m}$$
  $$\mathrm{Rang}(\mathbf{A}) + \dim \mathrm{Ker}(\mathbf{A}) = m$$

Cobild

Definition

- Das Cobild von Punkten oder Geraden ist
  das orthogonale Komplement des Urbildes



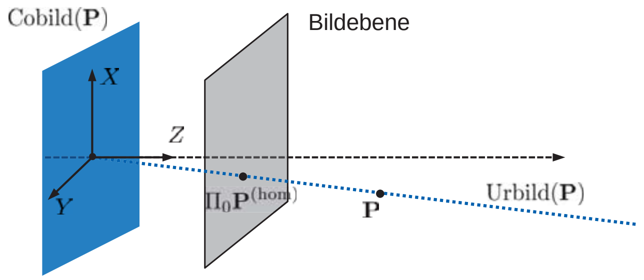


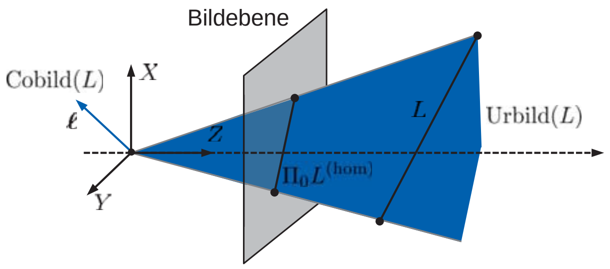
Bild, Urbild und Cobild

Zusammenhänge

- Äquivalente Darstellung von Bild, Urbild und Cobild

	Bild	Urbild	Cobild
Punkt	$\mathrm{span}(\mathbf{P}) \cap \mathrm{BE}$	$\mathrm{span}(\mathbf{P})$	$\mathrm{span}(\hat{\mathbf{P}})$
Linie	$\mathrm{span}(\hat{\boldsymbol{\ell}}) \cap \mathrm{BE}$	$\mathrm{span}(\hat{\boldsymbol{\ell}})$	$\mathrm{span}(\boldsymbol{\ell})$





Cobild

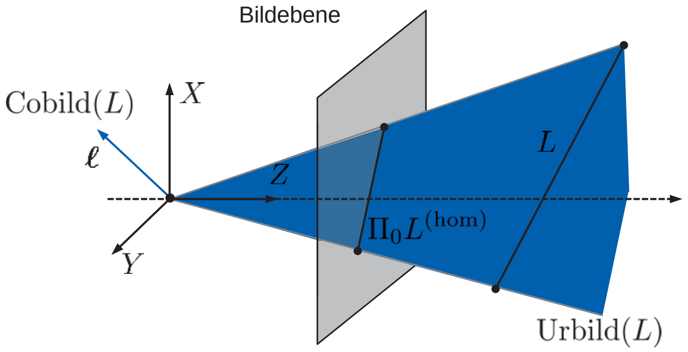
Nützliche Eigenschaften

- Sei L eine Gerade im Raum mit $\boldsymbol{\ell} \in \mathrm{Cobild}(L)$ und sei $\mathbf{x}$ das Bild eines Punktes auf dieser Linie. Dann gilt:

$$\mathbf{x}^\top \boldsymbol{\ell} = \boldsymbol{\ell}^\top \mathbf{x} = 0$$

- Seien $\mathbf{x}_1$ und $\mathbf{x}_2$ Bilder zweier Punkte im Raum. Dann gilt für das Cobild $\boldsymbol{\ell}$ der Verbindungsgeraden:

$$\boldsymbol{\ell} \sim \mathbf{x}_1 \times \mathbf{x}_2$$

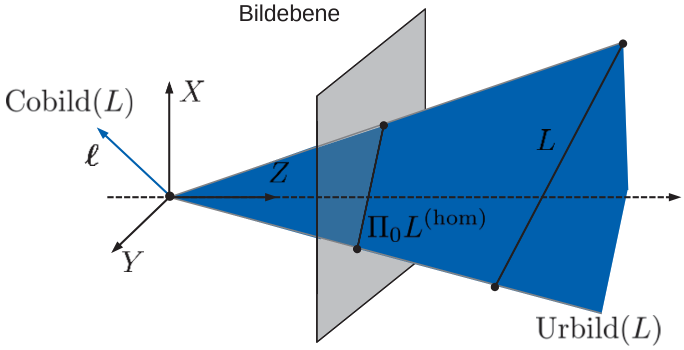


Cobild

Nützliche Eigenschaften

- Seien $\boldsymbol{\ell}_1$ und $\boldsymbol{\ell}_2$ die Cobilder zweier Geraden. Dann gilt für den Schnittpunkt $\mathbf{x}$ der Bilder dieser Geraden:

$$\mathbf{x} \sim \boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2$$



Kollinearität von Bildpunkten

Untersuchung mit Hilfe des Ranges

- Die Bildpunkte $\mathbf{x}_1, \ldots, \mathbf{x}_n$ liegen genau dann auf einer Linie (sind kollinear), wenn

$$\mathrm{Rang}([\mathbf{x}_1, \ldots, \mathbf{x}_n]) \leq 2$$

Allgemein sind für $\mathbf{A} \in \mathbb{R}^{n \times n}$ äquivalent:

- $\det \mathbf{A} = 0$
- $\mathrm{Rang}(\mathbf{A}) < n$

Kollinearität von Bildpunkten

Untersuchung mit Hilfe des Ranges

- Die Bildpunkte $\mathbf{x}_1, \ldots, \mathbf{x}_n$ liegen genau dann auf einer Linie (sind kollinear), wenn

$$\mathrm{Rang}([\mathbf{x}_1, \ldots, \mathbf{x}_n]) \leq 2$$

Allgemein sind für $\mathbf{A} \in \mathbb{R}^{n \times n}$ äquivalent:

- $\det \mathbf{A} = 0$
- $\mathrm{Rang}(\mathbf{A}) < n$

- Drei Bildpunkte sind genau dann kollinear, wenn

$$\det [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = 0$$

Eigenwerte und Eigenvektoren

Diagonalisierbarkeit von Matrizen

- Eigenwerte und Eigenvektoren einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v} \qquad \text{falls} \quad \exists\, \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$$

- Eigenvektoren sind nur bis auf Skalierung bestimmt

$$\alpha \mathbf{A}\mathbf{v} = \alpha \lambda \mathbf{v}$$

- Manche Matrizen $\mathbf{A} \in \mathbb{R}^{n \times n}$ sind diagonalisierbar, d.h.

$$\exists\, \mathbf{S} \in \mathbb{R}^{n \times n}, \mathbf{S} \text{ invertierbar} \qquad \mathbf{D}_A = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$$

- Reelle symmetrische Matrizen sind diagonalisierbar und haben zueinander orthogonale Eigenvektoren, d.h.

$$\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]\,,\ \mathbf{v}_i \text{ EV von } \mathbf{A} \qquad \mathbf{D}_A = \mathbf{V}^\top \mathbf{A}\mathbf{V}$$

Positiv (semi-)definite Matrizen

Orthogonale Diagonalisierbarkeit

- Eine symmetrische Matrix $\mathbf{A} = \mathbf{A}^\top$ heißt positiv (semi-)definit, wenn $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0,\ \mathbf{x} \neq \mathbf{0}$
- Die Eigenwerte positiv (semi-)definiter Matrizen sind positiv (nicht negativ)
- Also können positiv (semi-)definite Matrizen orthogonal diagonalisiert werden, d.h.

$$\lambda_i \geq 0,\ \forall \lambda_i \,\mathrm{EW\ von}\ \mathbf{A}$$

$$\mathbf{D}_A = \mathbf{V}^\top \mathbf{A} \mathbf{V} \qquad \mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]\,,\ \mathbf{v}_i\,\mathrm{EV\ von}\ \mathbf{A}$$

Kollinearität von Bildpunkten

Untersuchung mit Hilfe der Eigenwerte

- Die Bildpunkte $\mathbf{x}_1, \ldots, \mathbf{x}_n$ liegen genau dann auf einer Linie (sind kollinear), wenn

$$\mathrm{Rang}([\mathbf{x}_1, \ldots, \mathbf{x}_n]) \leq 2$$

- Drei Bildpunkte sind genau dann kollinear, wenn

$$\det\,[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = 0$$

- Die Bildpunkte $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sind genau dann kollinear, wenn der kleinste Eigenwert von

$$\mathbf{M} = \sum_{i=1}^{n} \omega_i \mathbf{x}_i \mathbf{x}_i^\top$$

gleich null ist $\forall \omega_i > 0$

Aspekte bei der praktischen Umsetzung

Ungenauigkeiten durch Diskretisierung / Rauschen

- In der Praxis sind die Bedingungen
  - $\det[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = 0$
  - Kleinster Eigenwert von $\mathbf{M} = \sum_{i=1}^{n} \omega_i \mathbf{x}_i \mathbf{x}_i^\top$ ist gleich null

  nicht erfüllt

- Verwendung von Schwellwerten

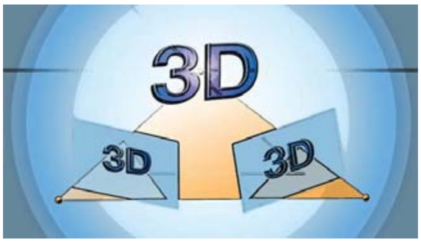
Zusammenfassung

- Urbilder von Punkten und Geraden sind Untervektorräume
- Cobild ist das orthogonale Komplement des Urbildes
- Darstellung von Linien im Bild mittels Cobildes
- Kriterien zur Kollinearität von Punkten
- Bild, Urbild und Cobild sind nützlicher Formalismus zum Erklären von einfachen geometrischen Zusammenhängen von Punkten und Geraden im Raum und auf der Bildebene

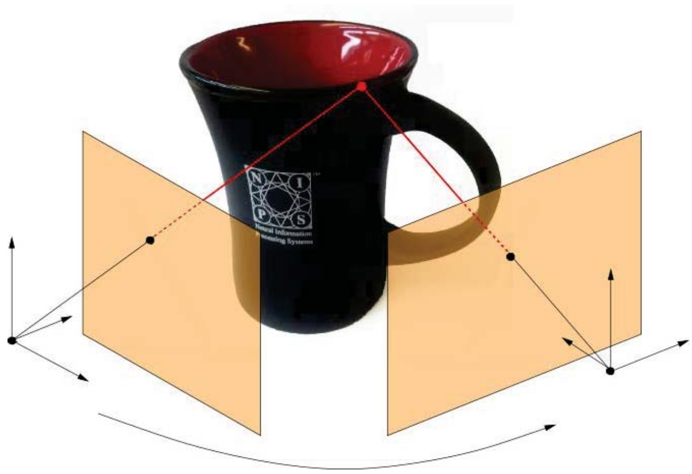
Martin Kleinsteuber: Computer Vision

Kapitel 3 – Epipolargeometrie

1. Epipolargleichung



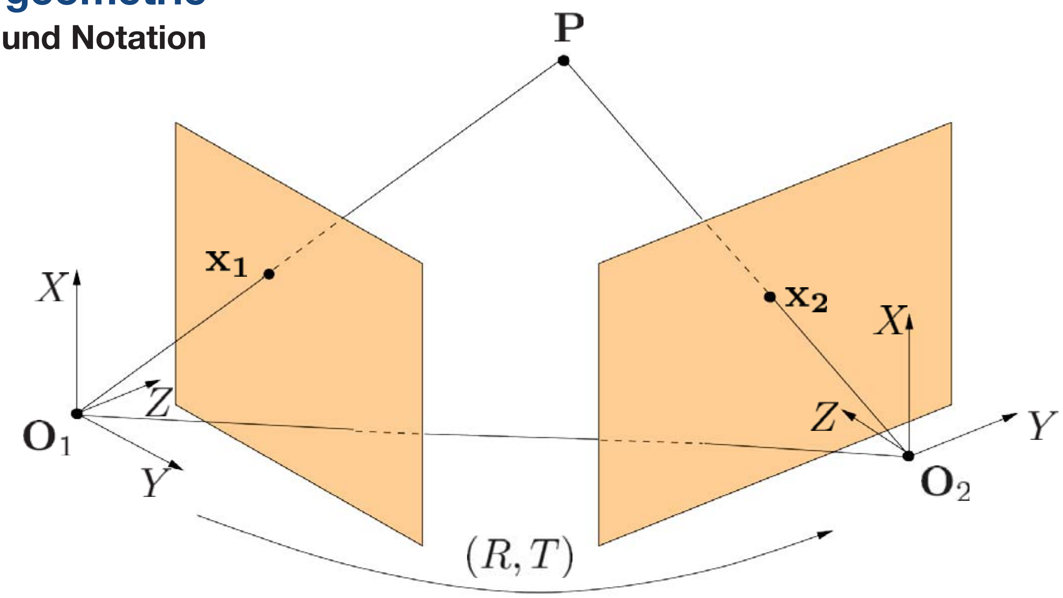
Motivation



- Ziel: Zusammenhang zwischen korrespondierenden Bildpunkten in Abhängkeit der euklidischen Bewegung der Kamera beschreiben

Epipolargeometrie

Überblick und Notation



Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

Euklidische Bewegung

Umrechnung von Kamera 1 in Kamera 2

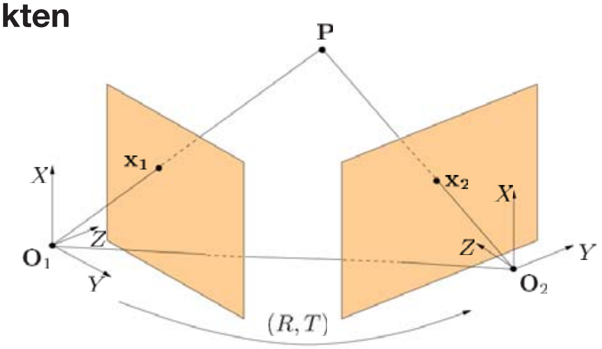
- $\mathbf{P_1}$ sind die Koordinaten des Punktes $\mathbf{P}$ in Kamerasystem 1
- $\mathbf{P_2}$ sind die Koordinaten des Punktes $\mathbf{P}$ in Kamerasystem 2
- Euklidische Bewegung der Kameras beschrieben durch die Koordinatentransformation

$$\mathbf{P}_2 = R\mathbf{P}_1 + T$$

mit $R \in SO(3), T \in \mathbb{R}^3$

Perspektivische Projektion

Zusammenhang zwischen Punkten und Bildpunkten



- Unter Annahme einer idealen Kamera gilt

$$\lambda_i \mathbf{x_i} = \mathbf{P_i}, \quad i = 1, 2, \quad \lambda \in \mathbb{R}$$

- Einsetzen in Euklidische Bewegung

$$\lambda_2 \mathbf{x_2} = R\lambda_1 \mathbf{x_1} + T$$

- Problem: Die Skalierungsfaktoren λ_i sind im Allgemeinen unbekannt

Die Epipolargleichung

Formeller Zusammenhang zwischen korrespondierenden Bildpunkten

$$\lambda_2 \mathbf{x_2} = R\lambda_1 \mathbf{x_1} + T$$

- Die Matrix $E = \hat{T}R$ heißt *essentielle Matrix* zur euklidischen Bewegung (R, T)
- Epipolargleichung $\quad \mathbf{x_2}^\top E \mathbf{x_1} = 0$

Eigenschaften der essentiellen Matrix
Singulärwertzerlegung

- Jede Matrix $A \in \mathbb{R}^{n \times m}$ lässt sich schreiben als Produkt
$$A = U\tilde{\Sigma}V^\top$$
wobei $U \in O(n), V \in O(m)$ und
- $n \leq m: \ \tilde{\Sigma} = \left[\begin{array}{c|c} \Sigma & 0 \end{array}\right]$
- $m < n: \ \tilde{\Sigma} = \left[\begin{array}{c} \Sigma \\ \hline 0 \end{array}\right]$
- $\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{\min\{n,m\}} \end{bmatrix} \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq 0$

Eigenschaften der essentiellen Matrix
Singulärwertzerlegung

- Singulärwerte einer Matrix sind eindeutig bestimmt. U und V in der Regel nicht.

- Zusammenhang mit der Eigenwertzerlegung von $AA^\top$ und $A^\top A$:

Eigenschaften der essentiellen Matrix

Charakterisierung durch Singulärwertzerlegung

- Eine Matrix E ist genau dann eine essentielle Matrix, also von der Form $E = \hat{T}R$ mit schiefsymmetrischem $\hat{T}$ und Rotationsmatrix R, wenn für die Singulärwertzerlegung von E gilt:

$$E = U \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} V^\top$$

Zusammenfassung

- Epipolargleichung beschreibt Zusammenhang zwischen korrespondierenden Bildpunkten aus unterschiedlichen Aufnahmen
- Die essentielle Matrix enthält die Information aus der euklidischen Bewegung
- Essentielle Matrizen sind genau die mit zwei gleichen Singulärwerten und einem Singulärwert gleich null

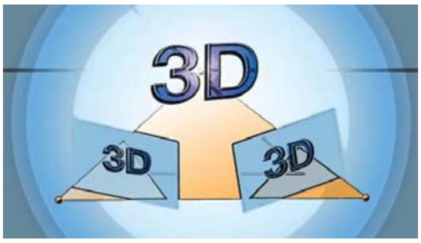
Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Martin Kleinsteuber: Computer Vision

Kapitel 3 – Epipolargeometrie

2. Epipole und Epipolarlinien

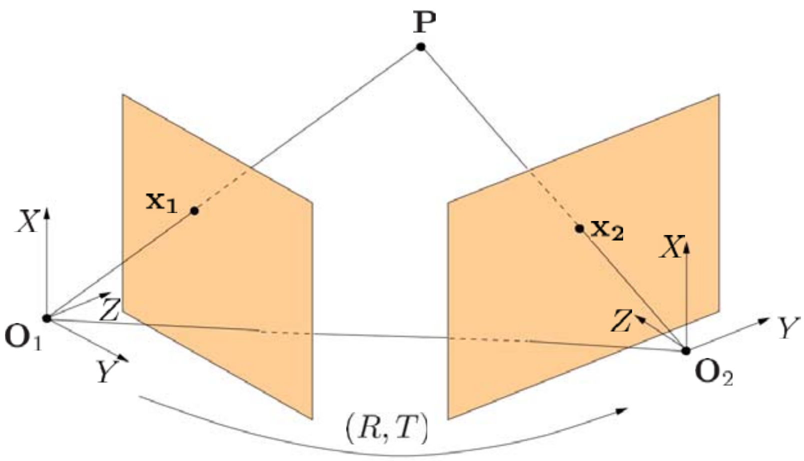


Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Epipolargeometrie

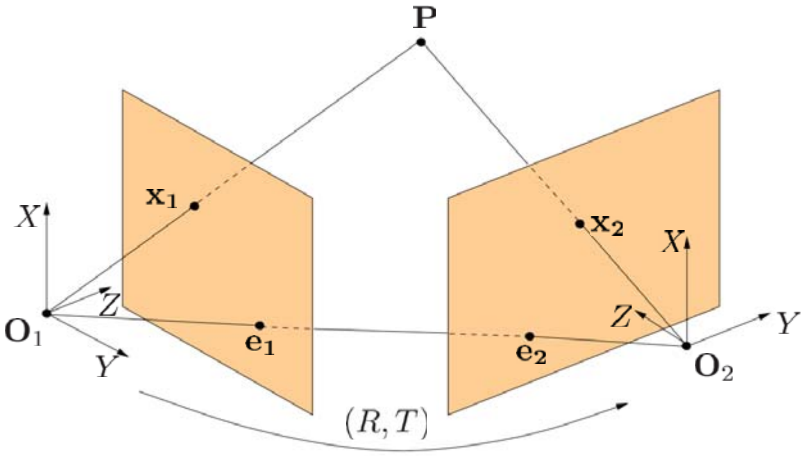
Wiederholung Epipolargleichung



Epipolargeometrie

Definition der Epipole

- Die perspektivischen Projektionen der optischen Zentren in das jeweils andere Kamerasystem heißen **Epipole**



Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

Epipole

Eigenschaften

- Aus Geometrie der eukl. Bewegung: $\mathbf{e_1} \sim R^\top \mathbf{T}, \quad \mathbf{e_2} \sim \mathbf{T}$

- $\mathbf{e_1}$ liegt im Kern von E: $E\mathbf{e_1} = 0$
- $\mathbf{e_2}$ liegt im Kern von $E^\top$: $E^\top \mathbf{e_2} = 0$

Die Nullräume von E
aus der Singulärwertzerlegung

- $E = \begin{bmatrix} \mathbf{u_1} & \mathbf{u_2} & \mathbf{u_3} \end{bmatrix} \begin{bmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v_1}^\top \\ \mathbf{v_2}^\top \\ \mathbf{v_3}^\top \end{bmatrix}$

- $\mathbf{e_1} \sim \mathbf{v_3} \Rightarrow E\mathbf{e_1} = 0$

- Das Urbild von $\mathbf{e_1}$ ist äquivalent zum dritten rechtsseitigen Singulärvektor von E

Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

Die Nullräume von E$^\top$
aus der Singulärwertzerlegung

- $E^\top = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} & \mathbf{v_3} \end{bmatrix} \begin{bmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u_1}^\top \\ \mathbf{u_2}^\top \\ \mathbf{u_3}^\top \end{bmatrix}$

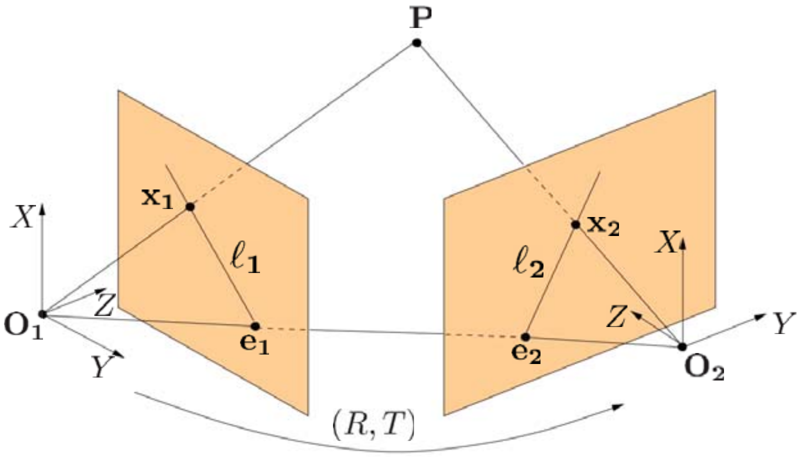
- $\mathbf{e_2} \sim \mathbf{u_3} \Rightarrow E^\top \mathbf{e_2} = 0$

- Das Urbild von $\mathbf{e_2}$ ist äquivalent zum dritten linksseitigen Singulärvektor von E

Epipolargeometrie

Definition Epipolarebene und Epipolarlinie

- Die Ebene, die durch $\mathbf{O_1}, \mathbf{O_2}$ und **P** aufgespannt wird, heißt **Epipolarebene von P**
- Der Schnitt der Epipolarebene mit der Bildebene heißt **Epipolarlinie**



Epipolarlinien

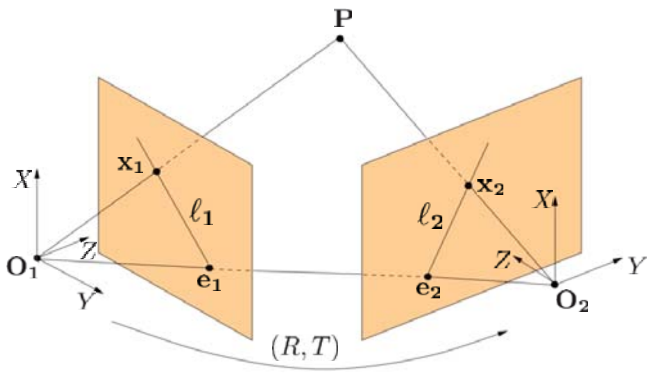
Geometrische Interpretation

- Die Epipolarlinie ist das Bild, das vom Urbild eines Punktes im anderen Kamerasystem erzeugt wird.

- Die Epipolarebene wird von den Ortsvektoren des Bildpunktes und des Epipols aufgespannt: $\operatorname{span}(\mathbf{x_i}, \mathbf{e_i})$

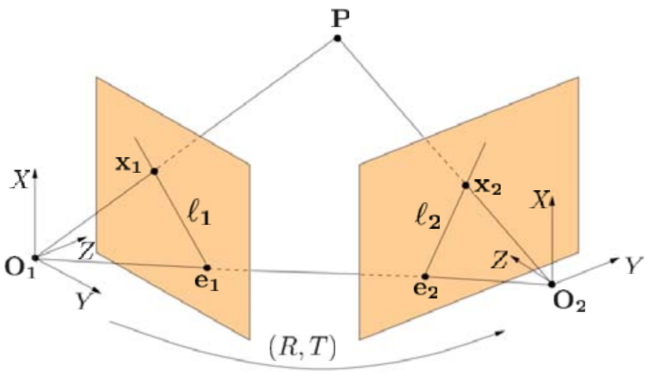
- Epipolarlinie wird über Cobild identifiziert

$$\ell_{\mathbf{i}} \sim \mathbf{e_i} \times \mathbf{x_i}$$



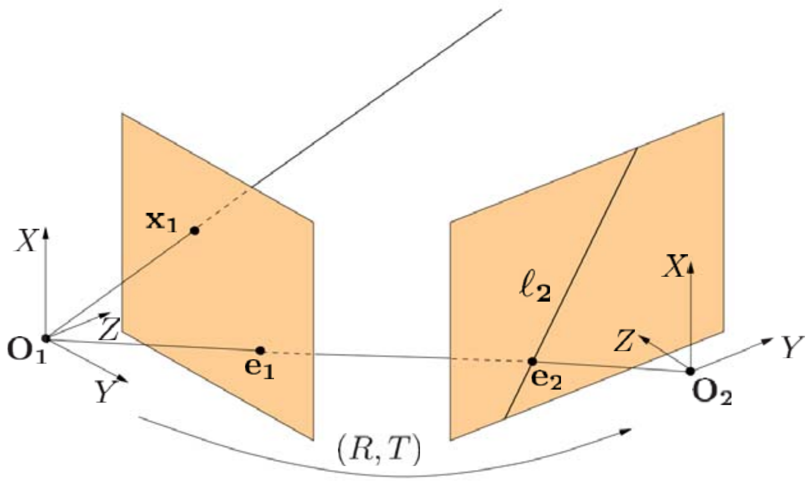
Epipolarlinien
Eigenschaften

- $\ell_i{}^\top \mathbf{x_i} = 0, \quad \ell_i{}^\top \mathbf{e_i} = 0$
- $\ell_1 \sim E^T \mathbf{x_2}, \quad \ell_2 \sim E \mathbf{x_1}$



Korrespondenzsuche
mit Hilfe der essentiellen Matrix

- Gegeben: E und $\mathbf{x_1}$
- Berechne $\ell_2 \sim E\mathbf{x_1}$
- Bestimme das Bild der Epipolarlinie
- Suche (z.B. mit NCC) entlang des Bildes nach $\mathbf{x_2}$



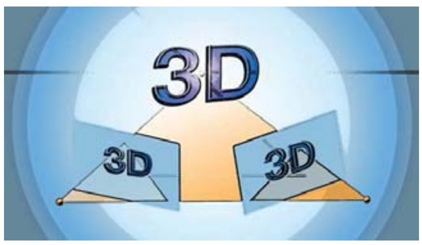
Zusammenfassung

- Berechnung der Epipole und Epipolarlinien mit Hilfe der essentiellen Matrix
- Urbilder der Epipole aus Singulärwertzerlegung der essentiellen Matrix
- Epipolarlinie mittels $\ell_2 \sim E\mathbf{x}_1$
- Vereinfachte Suche nach Korrespondenzen mit Epipolarlinie

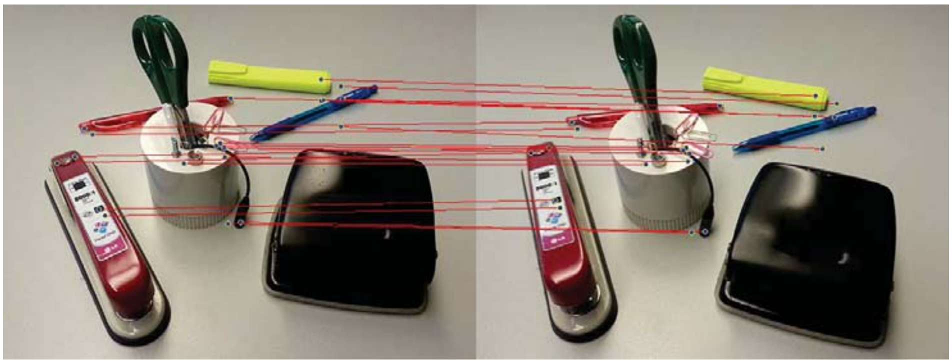
Martin Kleinsteuber: Computer Vision

Kapitel 3 – Epipolargeometrie

3. Der 8-Punkt-Algorithmus



Motivation



(R,T)

- Euklidische Bewegung nicht bekannt
- Korrespondenzen von Merkmalspunkten bekannt
- Wie schätzt man die essentielle Matrix?

Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

8-Punkt-Algorithmus zur Schätzung der Essentiellen Matrix

Motivation / Voraussetzungen

- Gegeben: n Korrespondenzpunktpaare $(\mathbf{x}_1^j, \mathbf{x}_2^j)$
- Idealerweise erfüllen alle KP die Epipolargleichung
  $\mathbf{x}_2^{j\top} E \mathbf{x}_1^j = 0$
- Ziel: Berechne die essentielle Matrix E aus den geschätzten KP
  $E \in \mathbb{R}^{3\times 3}$, somit 9 Unbekannte
- Skalierungsinvarianz: Wenn E Lösung ist, dann auch
  $\lambda E, \quad \lambda \in \mathbb{R}$
- Benötige 8 unabhängige Gleichungen

Vektorisierte Epipolargleichung

- Bislang: ${\mathbf{x}_2^j}^\top E \mathbf{x}_1^j = 0$

- Für ein ideales Korrespondenzpunktpaar $(\mathbf{x}_1^j, \mathbf{x}_2^j)$ gilt somit

$${\mathbf{a}^j}^\top \mathbf{E}^s = 0$$

Kronecker-Produkt $\otimes$

$$\mathbf{x}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \qquad \mathbf{a} := \mathbf{x}_1 \otimes \mathbf{x}_2 = \begin{bmatrix} x_1x_2 \\ x_1y_2 \\ x_1z_2 \\ y_1x_2 \\ y_1y_2 \\ y_1z_2 \\ z_1x_2 \\ z_1y_2 \\ z_1z_2 \end{bmatrix} \in \mathbb{R}^9$$

Vektorisieren („stacking“)

$$E = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \quad \text{ergibt} \quad \mathbf{E}^s = \begin{bmatrix} e_{11} \\ e_{21} \\ e_{31} \\ e_{12} \\ e_{22} \\ e_{32} \\ e_{13} \\ e_{23} \\ e_{33} \end{bmatrix}$$

Der 8-Punkt-Algorithmus

Gleichungssystem zur Bestimmung von E

- Aus n Korrespondenzpunktpaaren ergibt sich das homogene lineare Gleichungssystem $A\mathbf{E}^s = 0$, mit $A := \begin{bmatrix} {\mathbf{a}^1}^\top \\ {\mathbf{a}^2}^\top \\ \vdots \\ {\mathbf{a}^n}^\top \end{bmatrix} \in \mathbb{R}^{n \times 9}$
- Für generisch verteilte ideale KP gilt $\dim(\mathrm{kern}(A)) = 1$ und somit $\mathrm{rk}(A) = 8$

- Durch Diskretisierungsfehler gilt in der Realität

$$\mathrm{rk}(A) = 9$$

- Für $n > 8$ hat das homogene LGS keine nicht-triviale Lösung

Exkurs: Lineare Algebra
Homogene Gleichungssysteme

- Idee: Statt $A\mathbf{x} = 0$ wird Minimierungsproblem gelöst

  $$\min_{\mathbf{x}} \|A\mathbf{x}\|_2^2$$

- Da Lösungen skalierungsinvariant sind, kann die Suche beschränkt werden auf

  $$\mathbf{x} : \; \|\mathbf{x}\|_2 = 1$$

- Also finde $\mathbf{x}$ mit $\|\mathbf{x}\|_2 = 1$, welches $\|A\mathbf{x}\|_2^2$ minimiert.

Exkurs: Lineare Algebra
Lösung des Minimierungsproblems mittels SVD

- Singulärwertzerlegung $A = U\Sigma V^\top$

$$= \mathbf{x}^\top V \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix} V^\top \mathbf{x}$$

- $\mathbf{v}_n = \arg \min_{\|\mathbf{x}\|_2 = 1} \|A\mathbf{x}\|_2^2$

Der 8-Punkt-Algorithmus

Bisherige Herleitung

- Bilde Matrix A aus $n \geq 8$ generisch gelegenen Korrespondenzpunktpaaren
- Löse Minimierungsproblem $\mathbf{G}^s = \arg \min_{\|\mathbf{E}^s\|_2=1} \|A\mathbf{E}^s\|_2^2$:

  Singulärwertzerlegung $A = U_A \Sigma_A V_A^\top$ liefert die Lösung

  $\mathbf{G}^s = \mathbf{v}_9$ (9. Spalte von V_A)
- Umsortieren der Einträge von $\mathbf{G}^s$ führt zu $G \in \mathbb{R}^{3\times 3}$

Der 8-Punkt-Algorithmus

Von der Lösung des Minimierungsproblems zur Essentiellen Matrix

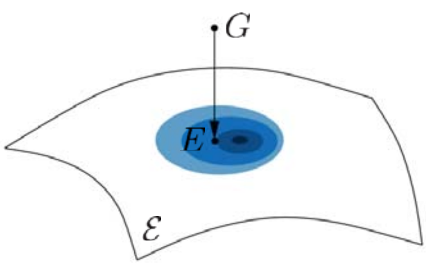
- G ist in der Regel keine essentielle Matrix

$$G = U_G \Sigma_G V_G^\top \qquad \Sigma_G = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix}$$

- Finde die „nächste“ essentielle Matrix zu G

$$E = \arg \min_{E \in \mathcal{E}} \|E - G\|_F^2$$

- Projektion auf den Raum der essentiellen Matrizen $\mathcal{E}$



Der 8-Punkt-Algorithmus

Projektion auf die nächste Essentielle Matrix

- $E = \arg\min_{E \in \mathcal{E}} \|E - G\|_F^2$
- $E = U_G \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} V_G^\top$
- $\sigma = \dfrac{\sigma_1 + \sigma_2}{2}$

- E kann nur bis auf Skalierung geschätzt werden
- In der Praxis daher üblicherweise

$$E = U_G \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} V_G^\top$$

Zusammenfassung

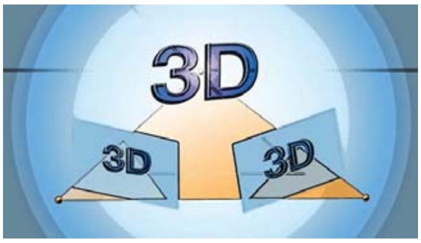
8-Punkt-Algorithmus

- Lineares Gleichungssystem aus Epipolarbedingungen
- Koeffizientenmatrix aus Kroneckerprodukt der KP-Paare
- Lösung ist der 9. rechtsseitige Singulärvektor
- Projektion auf normierte essentielle Matrizen
  - Singulärwertzerlegung der umsortieren Lösung
  - Zu-Eins-Setzen der ersten beiden Singulärwerte
  - Zu-Null-Setzen des kleinsten Singulärwerts

Martin Kleinsteuber: Computer Vision

Kapitel 3 – Epipolargeometrie

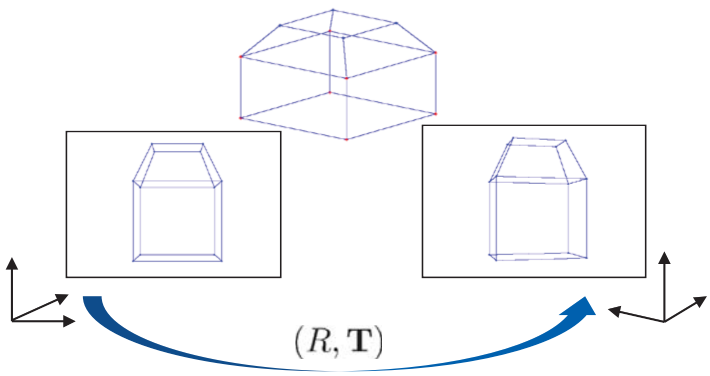
4. 3D-Rekonstruktion



3D-Rekonstruktion

Motivation und Annahmen

- Ziel: Schätze die euklidische Bewegung $(R, \mathbf{T})$ aus der essentiellen Matrix und rekonstruiere 3D-Punkte aus Punktkorrespondenzen $(\mathbf{x}_1^j, \mathbf{x}_2^j)$



Rekonstruktion von Rotation und Translation
mit Singulärwertzerlegung der essentiellen Matrix

- Benötige $E = U\Sigma V^\top, \quad U, V \in \mathrm{SO}(3)$ (Rotationsmatrizen mit Determinante 1)
- SVD nicht eindeutig, U und V nicht zwangsläufig Rotationsmatrizen
- Beispiel: $det(U) = -1$

$$E = U \qquad \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} V^\top$$

Fachgebiet Geometrische Optimierung und Maschinelles Lernen
Technische Universität München

Rekonstruktion von Rotation und Translation
mit Singulärwertzerlegung der essentiellen Matrix

- Gegeben: $E = U\Sigma V^\top, \quad U, V \in \mathrm{SO}(3)$

- Man kann zeigen, dass mit den Matrizen

$R_Z(\pm\frac{\pi}{2}) = \begin{bmatrix} 0 & \mp 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ gilt:

$$R = U R_Z^\top(\pm\tfrac{\pi}{2}) V^\top$$
$$\hat{T} = U R_Z(\pm\tfrac{\pi}{2}) \Sigma U^\top$$

Rekonstruktion von Rotation und Translation
Mehrdeutigkeiten

- 8-Punkt-Algorithmus liefert $\pm E$
- Je zwei Lösungen für $(R, \hat{T})$
- Insgesamt 4 euklidische Transformationen, welche die Korrespondenzen erklären
- Nur eine Transformation geometrisch plausibel
- $\mathbf{T}$ aus $\hat{T}$ nur bis auf Skalierung bestimmbar

Rekonstruktion der 3D-Koordinaten
aus geschätzter euklidischer Transformation

- Ziel: Schätze die Tiefe der Raumpunkte,
  z.B. λ_1^j in Kamerasystem 1

- $\lambda_2^j \mathbf{x}_2^j = \lambda_1^j R \mathbf{x}_1^j + \gamma \mathbf{T}, \quad j = 1, 2, \ldots, n$

- Jedes Korrespondenzpunktpaar liefert

$$\underbrace{\begin{bmatrix} \widehat{\mathbf{x}_2^j} R \mathbf{x}_1^j & \widehat{\mathbf{x}_2^j} \mathbf{T} \end{bmatrix}}_{:= M^j} \begin{bmatrix} \lambda_1^j \\ \gamma \end{bmatrix} = 0$$

Rekonstruktion der 3D-Koordinaten

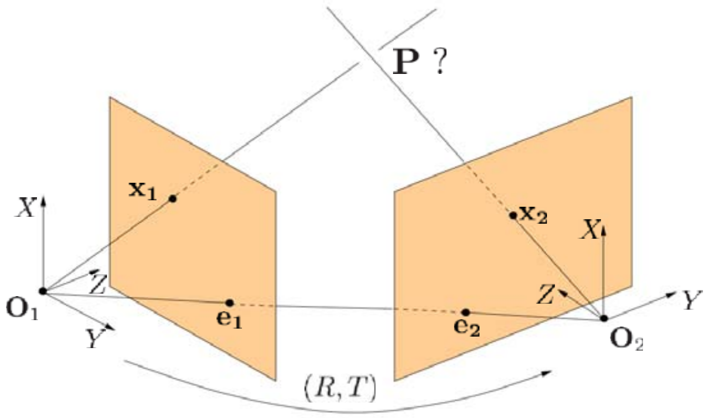
bis auf Skalierung möglich

- Ideale KP erfüllen $$M\mathbf{d} := \begin{bmatrix} \widehat{\mathbf{x}_2^1} R\mathbf{x}_1^1 & & & & \widehat{\mathbf{x}_2^1}\mathbf{T} \\ & \widehat{\mathbf{x}_2^2} R\mathbf{x}_1^2 & & & \widehat{\mathbf{x}_2^2}\mathbf{T} \\ & & \ddots & & \vdots \\ & & & \widehat{\mathbf{x}_2^n} R\mathbf{x}_1^n & \widehat{\mathbf{x}_2^n}\mathbf{T} \end{bmatrix} \begin{bmatrix} \lambda_1^1 \\ \lambda_1^2 \\ \vdots \\ \lambda_1^n \\ \gamma \end{bmatrix} = 0$$
- In der Praxis: Löse Minimierungsproblem $\min_{\|\mathbf{d}\|_2=1} \|M\mathbf{d}\|_2^2$ über SVD von M
- Inhärente Skalierungsinvarianz: Ohne weiteres Wissen über die Szene keine Unterscheidung zwischen Skalierung des Objektes und Distanz der Kamera
- Geschätzte 3D-Koordinaten von $\mathbf{P}$ bzgl. Kamerasystems 1 sind $\mathbf{P}_1^j = \lambda_1^j \mathbf{x}_1^j$

Rekonstruktion der 3D-Koordinaten

Schwierigkeiten und Lösungsansätze

- Fehlerhafte Korrespondenzschätzung
  - Lösung z.B. über RanSaC-Methode
- Fehlerhafte 3D-Rekonstruktion durch Diskretisierungsfehler
  - Zusätzliche Schätzung von λ_2^j
  - Robuste Triangulationsverfahren, die 8PA als Initialisierung verwenden.



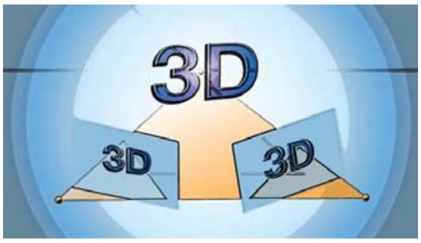
Zusammenfassung

- Zwei euklidische Transformationen zu einer essentiellen Matrix
- Vier euklidische Transformationen nach 8-Punkt-Algorithmus
- Physikalisch richtige durch Überprüfen der Positivität der Tiefen
- Rekonstruktion der 3D-Raumkoordinaten mittels euklidischer Transformation

Martin Kleinsteuber: Computer Vision

Kapitel 3 – Epipolargeometrie

5. Die Fundamentalmatrix



Motivation

- Epipolargleichung für kalibrierte Kameras $\mathbf{x_2}^\top E \mathbf{x_1} = 0$
- Kann man eine ähnliche Beziehung zwischen Pixelkoordinaten im unkalibrierten Fall finden?
- Beziehung zwischen kalibrierten und unkalibrierten Koordinaten ergibt sich aus der Matrix K

$$\mathbf{x}' \sim K\Pi_0 \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mathbf{P}$$

$$\mathbf{x} = K^{-1}\mathbf{x}' \sim \Pi_0 \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mathbf{P}$$

$$K = K_s K_f = \begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

Epipolargleichung

Unkalibrierter Fall

- Epipolargleichung für kalibrierte Kameras $\mathbf{x_2}^\top E \mathbf{x_1} = 0$

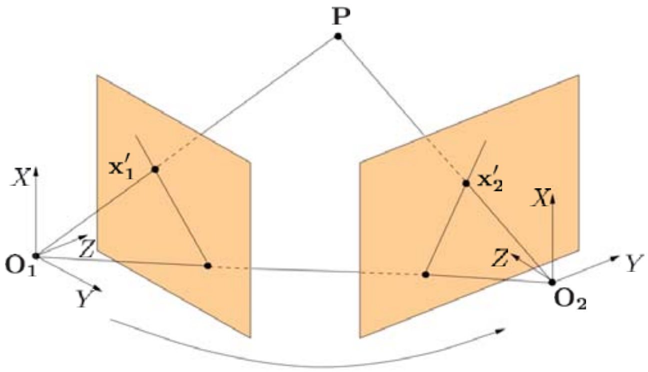
- Unkalibrierte Version der Epipolargleichung:

$$\mathbf{x_2'}^\top K^{-\top} E K^{-1} \mathbf{x_1'} = 0$$

- Die Matrix $F := K^{-\top} E K^{-1}$ heißt Fundamentalmatrix.

Epipolargeometrie

Unkalibrierter Fall



(KRK^{-1}, KT)

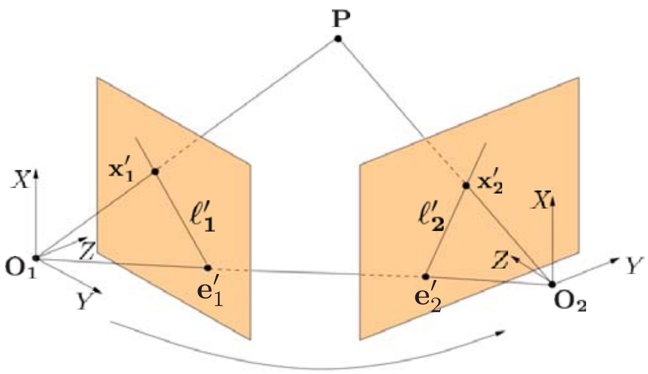
- Euklidische Transformation $\lambda_2 \mathbf{x_2} = R\lambda_1 \mathbf{x_1} + T$

- Es gilt $\mathbf{x_2'}^\top \hat{T}' KRK^{-1} \mathbf{x_1'} = 0$
- Mit $\hat{T}' \sim K^{-\top} \hat{T} K^{-1}$ folgt $F \sim \hat{T}' KRK^{-1}$

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen
Technische Universität München

Eigenschaften der Fundamentalmatrix

Epipole und Epipolarlinien



- Der Korrespondenzpunkt $\mathbf{x}_2'$ liegt auf der Linie $\ell_2' \sim F\mathbf{x_1}'$ und umgekehrt: $\mathbf{x}_1'$ liegt auf $\ell_1' \sim F^\top \mathbf{x}_2'$

- Die Epipole in Pixelkoordinaten:

$$\mathbf{e}_2'^\top F = 0, \quad F\mathbf{e}_1' = 0$$

Eigenschaften der Fundamentalmatrix

Singulärwertzerlegung der Fundamentalmatrix

- Eine Matrix ist genau dann eine Fundamentalmatrix wenn für ihre Singulärwertzerlegung gilt:

$$F = U\Sigma V^\top \qquad \text{mit} \quad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix}$$

- Fundamentalmatrix kann über den 8-Punkt-Algorithmus geschätzt werden

Eigenschaften der Fundamentalmatrix

8-Punkt-Algorithmus zum Schätzen der Fundamentalmatrix

- Gegeben: n Korrespondenzpunktpaare in Pixelkoordinaten $(\mathbf{x}_1'^{\,j}, \mathbf{x}_2'^{\,j}),\ j = 1, \ldots, n \quad (n \geq 8)$
- Bilde Koeffizientenmatrix $A := \begin{bmatrix} \mathbf{a}^{1\,\top} \\ \vdots \\ \mathbf{a}^{n\,\top} \end{bmatrix} \in \mathbb{R}^{n\times 9}$ mit $\mathbf{a}^\mathbf{j} = \mathbf{x}_1'^{\,j} \otimes \mathbf{x}_2'^{\,j}$
- Extrahiere rechtsseitigen Singulärvektor von A zum kleinsten Singulärwert und konstruiere daraus eine Matrix $G \in \mathbb{R}^{3\times 3}$
- Projiziere auf die Menge der Fundamentalmatrizen:
  - SVD von $G = U_G \operatorname{diag}\{\sigma_1, \sigma_2, \sigma_3\} V_G^\top$
  - Schätzung der Fundamentalmatrix $F = U_G \operatorname{diag}\{\sigma_1, \sigma_2, 0\} V_G^\top$

Kalibrierter vs. Unkalibrierter Fall

Überblick & Diskussion

Kalibrierte Kamera

- Epipolargleichung $\mathbf{x}_2^\top E \mathbf{x}_1 = 0$
- Essentielle Matrix $E = \hat{T} R$
- Epipole und Epipolarlinien

$$E\mathbf{e_1} = 0, \quad E^\top \mathbf{e_2} = 0$$
$$\ell_2 \sim E\mathbf{x_1}, \quad \ell_1 \sim E^\top \mathbf{x_2}$$

- 3D-Rekonstruktion möglich

Unkalibrierte Kamera

- Epipolargleichung $\mathbf{x}_2'^\top F \mathbf{x}_1' = 0$
- Fundamentalmatrix $F = K^{-\top} \hat{T} R K^{-1}$
- Epipole und Epipolarlinien

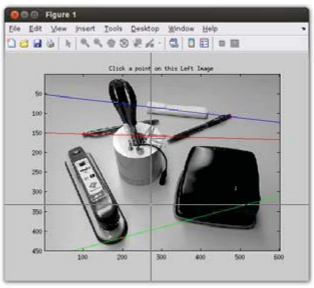
$$F\mathbf{e}_1' = 0, \quad F^\top \mathbf{e}_2' = 0$$
$$\ell_2' \sim F\mathbf{x}_1', \quad \ell_1' \sim F^\top \mathbf{x}_2'$$

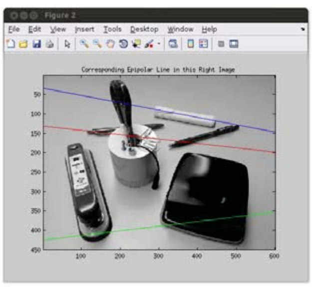
- 3D-Rekonstruktion nicht ohne Weiteres möglich

MATLAB-Demonstration

8-Punkt-Algorithmus und Epipolarlinien

- Mitul Saha & Rohit Singh, Stanford University
- http://ai.stanford.edu/~mitul/cs223b/fm.html





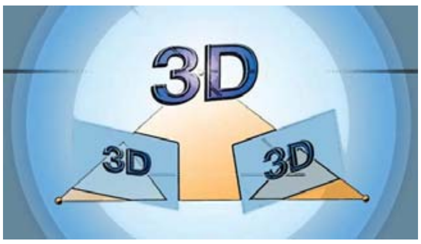
Zusammenfassung

- Epipolargleichung für unkalibrierte Kamera
- Fundamentalmatrix beinhaltet Information der intrinsischen Kameraparameter und der euklidischen Bewegung
- Schätzen der Fundamentalmatrix mit dem 8-Punkt-Algorithmus und Korrespondenzpunkten in Pixelkoordinaten

Martin Kleinsteuber: Computer Vision

Kapitel 4 – Planare Szenen

1. Die planare Epipolargleichung



Motivation

Planare Szene (Luftbild)

- Ziel 1: Formeller Zusammenhang zwischen Korrespondenzen
- Ziel 2: Schätze euklidische Bewegung der Kamera mit Korrespondenzen in einer planaren Szene

Planare Szene

Ebenengleichung

- Ebene in Kamerasystem 1

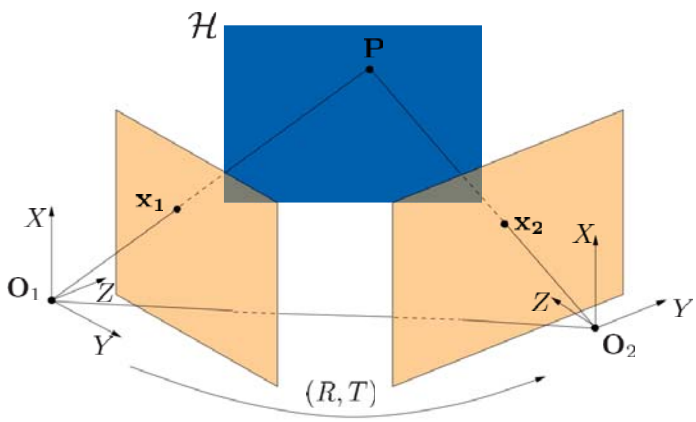
  $\mathcal{H} =: \{\mathbf{X} \in \mathbb{R}^3 \,|\, \mathbf{n}^\top \mathbf{X} = d\}$

  $\mathbf{n}$: Normierter Normalenvektor

  d: Abstand von $\mathcal{H}$ zu $\mathbf{O}_1$

- Ebenengleichung mit Hilfe von $\mathbf{P}_1$

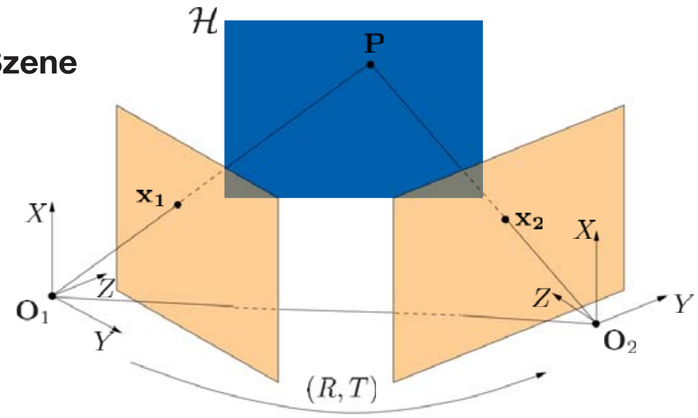
  $\mathbf{n}^\top \mathbf{P}_1 = n_1 X + n_2 Y + n_3 Z = d$



Planare Epipolargleichung

Zusammenhang zwischen KP in planarer Szene

- Euklidische Bewegung $\mathbf{P}_2 = R\mathbf{P}_1 + T$
- Ebenengleichung $\mathbf{n}^\top \mathbf{P}_1 = d$



- H heißt planare Homographiematrix zur Ebene $\mathcal{H}$ und zur euklidischen Bewegung (R, T)
- Die Gleichung $\mathbf{x}_2 \sim H\mathbf{x}_1$ heißt planare Epipolargleichung

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen
Technische Universität München

Eigenschaften der Homographiematrix

Charakterisierung durch Singulärwertzerlegung

- Eine Matrix H ist genau dann Homographiematrix, wenn für die Singulärwertzerlegung von H gilt:

$$H = U \begin{bmatrix} \sigma_1 & & \\ & 1 & \\ & & \sigma_3 \end{bmatrix} V^\top$$

Scheitern des 8-Punkt-Algorithmus

im planaren Fall

- Planare Epipolargleichung $\mathbf{x}_2 \sim H\mathbf{x}_1$
- $\mathbf{x}_2$ orthogonal zu $\mathbf{u} \times H\mathbf{x}_1$ für alle $\mathbf{u} \in \mathbb{R}^3$
- Für ideale KP hat Koeffizientenmatrix $A := \begin{bmatrix} \mathbf{a}^{1\top} \\ \mathbf{a}^{2\top} \\ \vdots \\ \mathbf{a}^{n\top} \end{bmatrix} \in \mathbb{R}^{n\times 9}, \quad \mathbf{a}^j := \mathbf{x}_1^j \otimes \mathbf{x}_2^j$

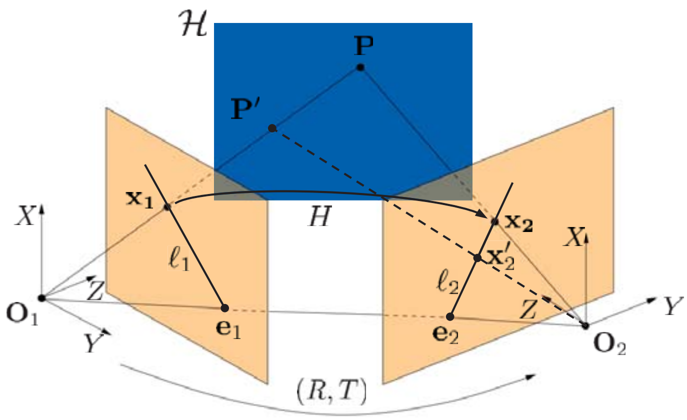
$\dim(\mathrm{Ker}(A)) \geq 3$ und $\mathrm{Rang}(A) \leq 6$

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen

Technische Universität München

Planare Epipolargeometrie

Schlussfolgerungen



- Für jedes Bild $\mathbf{x}_1$ eines Punktes in der Ebene existiert ein eindeutiger KP $\mathbf{x}_2$
- Epipolarlinien durch Homographie bestimmt

$\boldsymbol{\ell}_2 \sim \hat{\mathbf{x}}_2 H\mathbf{x}_1, \quad \boldsymbol{\ell}_1 \sim H^\top \boldsymbol{\ell}_2$

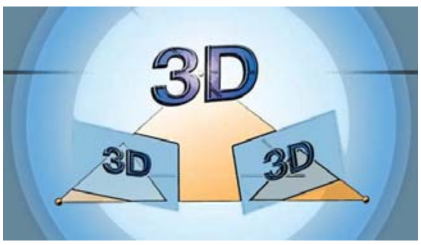
Zusammenfassung

- Im planaren Fall liegen 3D-Raumpunkte auf einer Ebene in der Szene
- Planare Epipolargleichung beschreibt formellen Zusammenhang der Bildpunkte
- Eindeutige Bestimmung von Korrespondenzen mit Hilfe der Homographiematrix
- Homographiematrix erlaubt die Berechnung von Epipolarlinien auch für Punkte außerhalb der Ebene

Martin Kleinsteuber: Computer Vision

Kapitel 4 – Planare Szenen

2. Der 4-Punkt-Algorithmus



Motivation

planare Szene

- Wie schätzt man die Homographiematrix aus planaren KP?
- Wie extrahiert man daraus die euklidische Bewegung?
- Wie hängen die Homographiematrix und die essentielle Matrix zusammen?

4-Punkt-Algorithmus zur Schätzung der Homographiematrix

Motivation / Voraussetzungen

- Gegeben: n Korrespondenzpunktpaare $(\mathbf{x}_1^j, \mathbf{x}_2^j)$ von Punkten in einer Ebene
- Idealerweise erfüllen alle KP die planare Epipolargleichung

  $\hat{\mathbf{x}}_2^j H \mathbf{x}_1^j = 0$
- Ziel: Berechne die Homographiematrix H aus den geschätzten KP

  $H \in \mathbb{R}^{3\times3}$, somit 9 Unbekannte
- Skalierungsinvarianz: Wenn H Lösung ist, dann auch

  $\lambda H, \quad \lambda \in \mathbb{R}$

Vektorisierte planare Epipolargleichung

Kronecker-Produkt $\otimes$

$$\mathbf{x}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \qquad \hat{\mathbf{x}}_2 = \begin{bmatrix} 0 & -z_2 & y_2 \\ z_2 & 0 & -x_2 \\ -y_2 & x_2 & 0 \end{bmatrix}$$

$$B := \mathbf{x}_1 \otimes \hat{\mathbf{x}}_2 = \begin{bmatrix} x_1 \hat{\mathbf{x}}_2 \\ y_1 \hat{\mathbf{x}}_2 \\ z_1 \hat{\mathbf{x}}_2 \end{bmatrix} \in \mathbb{R}^{9\times 3}$$

- Bislang: $\hat{\mathbf{x}}_2^j H \mathbf{x}_1^j = 0$
- $\operatorname{Rang}(\hat{\mathbf{x}}_2) = 2$, also i.A. auch $\operatorname{Rang}(B) = 2$
- Für jedes ideale Korrespondenzpunktpaar $(\mathbf{x}_1^j, \mathbf{x}_2^j)$ enthält das homogene LGS $B^{j\top} \mathbf{H}^s = 0$ in der Regel zwei unabhängige Gleichungen
- Benötige mind. 4 allgemein liegende Korrespondenzpunktpaare

Der 4-Punkt-Algorithmus

- Aus n Korrespondenzpunktpaaren ergibt sich das homogene lineare Gleichungssystem $A\mathbf{H}^s = 0$, mit $A := \begin{bmatrix} B^{1\top} \\ B^{2\top} \\ \vdots \\ B^{n\top} \end{bmatrix} \in \mathbb{R}^{3n\times 9}$
- Löse Minimierungsproblem
  $\mathbf{H}_\mathbf{L}^\mathbf{s} = \arg\min_{\|\mathbf{H}^\mathbf{s}\|_2 = 1} \|A\mathbf{H}^\mathbf{s}\|_2^2$ mittels SVD von A
- $H_L := \lambda H$ sind umsortierte Einträge des rechtsseitigen Singulärvektors zum kleinsten Singulärwert, Schätzung von H bis auf Skalierung
- Aus den Eigenschaften der Homographiematrix
  $|\lambda| = \sigma_2(H_L)$ und somit $H = \frac{1}{\sigma_2(H_L)} H_L$

Der 4-Punkt-Algorithmus

Vorzeichen der geschätzten Homographiematrix

- Schätzung von H bis auf Vorzeichen, $\pm H$ erfüllen planare Epipolargleichung
- Ausnutzen der Positivitätsbedingung der Tiefen führt zu Kriterium $\mathbf{x}_2^\top H \mathbf{x}_1 > 0$

- Wähle Vorzeichen von H entsprechend dieses Kriteriums

3D-Rekonstruktion aus Homographiematrix

Herleitung

- Zu einer Homographiematrix $H = (R + \frac{1}{d}\mathbf{T}\mathbf{n}^\top)$ gibt es höchstens zwei physikalisch mögliche Zerlegungen in die Parameter $\{R, \frac{1}{d}\mathbf{T}, \mathbf{n}\}$
- Eigenwertzerlegung $H^\top H = V\Sigma^2 V^\top, \quad V = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} \in \mathrm{SO}(3)$
- Definiere

$$\mathbf{u}_1 := \frac{\sqrt{1-\sigma_3^2}\mathbf{v}_1 + \sqrt{\sigma_1^2 - 1}\mathbf{v}_3}{\sqrt{\sigma_1^2 - \sigma_3^2}}, \quad \mathbf{u}_2 := \frac{\sqrt{1-\sigma_3^2}\mathbf{v}_1 - \sqrt{\sigma_1^2 - 1}\mathbf{v}_3}{\sqrt{\sigma_1^2 - \sigma_3^2}}$$

- Definiere

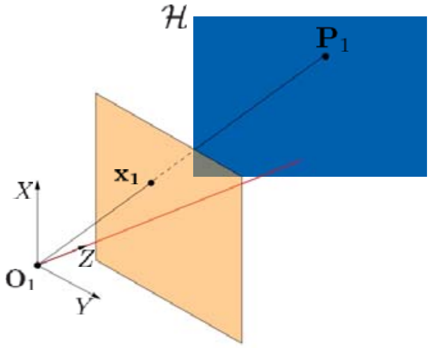
$$U_1 := \begin{bmatrix}\mathbf{v}_2 & \mathbf{u}_1 & \hat{\mathbf{v}}_2\mathbf{u}_1\end{bmatrix}, \quad W_1 := \begin{bmatrix}H\mathbf{v}_2 & H\mathbf{u}_1 & \widehat{H\mathbf{v}_2}H\mathbf{u}_1\end{bmatrix}$$

$$U_2 := \begin{bmatrix}\mathbf{v}_2 & \mathbf{u}_2 & \hat{\mathbf{v}}_2\mathbf{u}_2\end{bmatrix}, \quad W_2 := \begin{bmatrix}H\mathbf{v}_2 & H\mathbf{u}_2 & \widehat{H\mathbf{v}_2}H\mathbf{u}_2\end{bmatrix}$$

3D-Rekonstruktion aus Homographiematrix

4 Lösungen

Lösung 1: $R_1 = W_1 U_1^\top$, $\mathbf{n}_1 = \mathbf{v}_2 \mathbf{u}_1$, $\frac{1}{d}\mathbf{T}_1 = (H - R_1)\mathbf{n}_1$	Lösung 3: $R_3 = R_1$, $\mathbf{n}_3 = -\mathbf{n}_1$, $\frac{1}{d}\mathbf{T}_3 = -\frac{1}{d}\mathbf{T}_1$
Lösung 2: $R_2 = W_2 U_2^\top$, $\mathbf{n}_2 = \mathbf{v}_2 \mathbf{u}_2$, $\frac{1}{d}\mathbf{T}_2 = (H - R_2)\mathbf{n}_2$	Lösung 4: $R_4 = R_2$, $\mathbf{n}_4 = -\mathbf{n}_2$, $\frac{1}{d}\mathbf{T}_4 = -\frac{1}{d}\mathbf{T}_2$



Homographiematrix und essentielle Matrix

Zusammenhänge

- Betrachte $E = \hat{\mathbf{T}}R$ und $H = R + \mathbf{T}\mathbf{u}^\top$, $\quad R \in \mathbb{R}^{3\times 3}$, $\mathbf{T}, \mathbf{u} \in \mathbb{R}^3$, $\quad \|\mathbf{T}\| = 1$
- Dann gilt
  - $E = \hat{\mathbf{T}}H$
  - $H^\top E + E^\top H = 0$
  - $H = \hat{\mathbf{T}}^\top E + \mathbf{T}\mathbf{v}^\top$ für ein bestimmtes $\mathbf{v} \in \mathbb{R}^3$

Homographiematrix und essentielle Matrix

Berechnung der essentiellen Matrix aus der Homographiematrix

- Homographiematrix H bekannt
- Habe außerdem zwei KP-Paare $(\mathbf{x}_1^j, \mathbf{x}_2^j),\ j = 1, 2$, deren 3D-Punkte nicht auf der Ebene zur Homographiematrix liegen
- Epipolarlinien $\ell_2^j \sim \hat{\mathbf{x}}_2^j H \mathbf{x}_1^j$ schneiden sich im Epipol $\mathbf{e}_2 \sim \mathbf{T}$
- Dann gilt $E = \hat{\mathbf{T}} H$ mit $\mathbf{T} \sim \hat{\ell}_2^1 \ell_2^2$ und $\|\mathbf{T}\|_2 = 1$

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen
Technische Universität München

Homographiematrix und essentielle Matrix

Berechnung der Homographiematrix aus der essentiellen Matrix

- Essentielle Matrix E bekannt
- Habe außerdem drei KP-Paare $(\mathbf{x}_1^j, \mathbf{x}_2^j),\ j = 1, 2, 3$, deren 3D-Punkte eine Ebene im Raum definieren
- Dann gilt $H = \hat{\mathbf{T}}^\top E + \mathbf{T}\mathbf{v}^\top$ mit $\mathbf{v} \in \mathbb{R}^3$, welches das Gleichungssystem $\hat{\mathbf{x}}_2^j \left(\hat{\mathbf{T}}^\top E + \mathbf{T}\mathbf{v}^\top \right) \mathbf{x}_1^j = 0, \quad j = 1, 2, 3$ löst

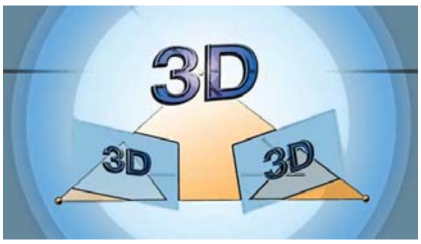
Zusammenfassung

- 4-Punkt-Algorithmus zur Bestimmung der Homographiematrix
- Zwei physikalisch plausible Lösungen für die Parameter der 3D-Rekonstruktion
- Mit Hilfe von weiteren Korrespondenzpunkten lässt sich die essentielle Matrix aus der Homographie bestimmen und umgekehrt

Martin Kleinsteuber: Computer Vision

Kapitel 4 – Planare Szenen

3. Kamerakalibrierung



Motivation

- Bestimme Kalibrierungsmatrix K aus mehreren Ansichten eines Schachbretts
- Nutze dabei Kenntnis über die Geometrie des Schachbrettes

Fachgebiet
Geometrische Optimierung
und Maschinelles Lernen
Technische Universität München

Wiederholung

Perspektivische Projektion

- Pixelkoordinaten

$$\mathbf{x}' = K_s \mathbf{x}$$

- Perspektivische Projektion

$$\mathbf{x} \sim K_f \Pi_0 \mathbf{P}$$

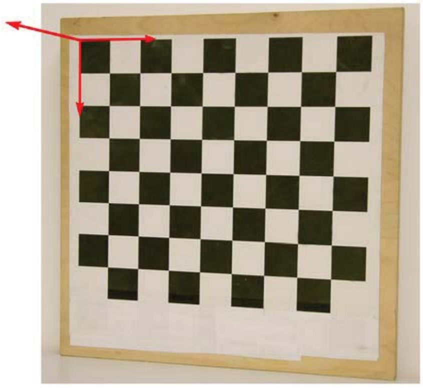
$$\mathbf{x}' \sim K_s K_f \Pi_0 \mathbf{P} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{:=K} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Mit euklidischer Transformation von $\mathbf{P}$

$$\mathbf{x}' \sim K \Pi_0 \begin{bmatrix} R & \mathbf{T} \\ 0 & 1 \end{bmatrix} \mathbf{P}$$

Video

Kalibrierung
Ansatz



- Schachbrett liegt im Ursprung der Weltkoordinaten
- Z-Achse der Weltkoordinaten senkrecht zum Schachbrett
- Dann gilt für einen Punkt $\mathbf{P} = \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$ auf dem Schachbrett

$$\mathbf{x}' \sim K\Pi_0 \begin{bmatrix} R & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{T} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Video

Kalibrierung
Schätzen der Homographie

- Die Matrix $H := K \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{T} \end{bmatrix}$ ist bis auf Skalierung eine Homographiematrix, die homogene Koordinaten vom Schachbrett auf homogene, unkalibrierte Koordinaten in der Bildebene abbildet:

$$\hat{\mathbf{x}}' H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = 0$$

- Schätze H bis auf Skalierung mittels des 4-Punkt-Algorithmus

  durch gegebene Korrespondenzpunktpaare $\left(\mathbf{x}'^{j}, \begin{bmatrix} X^j \\ Y^j \\ 1 \end{bmatrix} \right)$

Video

Kalibrierung
Bedingungen an die Kalibrierungsmatrix

- Es gilt $K^{-1}\begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2\end{bmatrix} \sim \begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2\end{bmatrix}$
- Durch Orthogonalität und Normierung von $\mathbf{r}_1$ und $\mathbf{r}_2$ erhalten wir die Bedingungen:
  - $\mathbf{h}_1^\top K^{-\top} K^{-1} \mathbf{h}_2 = 0$
  - $\mathbf{h}_1^\top K^{-\top} K^{-1} \mathbf{h}_1 = \mathbf{h}_2^\top K^{-\top} K^{-1} \mathbf{h}_2$
- Die Matrix $K^{-\top} K^{-1}$ ist symmetrisch:

$$B := K^{-\top} K^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$

Video

Kalibrierung
Schätzen der Kalibrierungsmatrix

- Das Produkt $\mathbf{a}^\top B \mathbf{c}$ mit $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ lässt sich

  umformulieren zu $\mathbf{a}^\top B \mathbf{c} = \mathbf{v}_{(\mathbf{a},\mathbf{c})}^\top \mathbf{b}$ mit $\mathbf{v}_{(\mathbf{a},\mathbf{c})} = \begin{bmatrix} a_1 c_1 \\ a_1 c_2 + a_2 c_1 \\ a_2 c_2 \\ a_3 c_1 + a_1 c_3 \\ a_3 c_2 + a_2 c_3 \\ a_3 c_3 \end{bmatrix}$

  und $\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^\top \in \mathbb{R}^6$

Video

Kalibrierung

Schätzen der Kalibrierungsmatrix

- Für jede Ansicht liefern die Bedingungen
  - $\mathbf{h}_1^\top B\, \mathbf{h}_2 = 0$
  - $\mathbf{h}_1^\top B\, \mathbf{h}_1 = \mathbf{h}_2^\top B\, \mathbf{h}_2$

  zwei Gleichungen

$$V^j \mathbf{b} = \begin{bmatrix} \mathbf{v}_{(\mathbf{h}_1,\mathbf{h}_2)}^\top \\ (\mathbf{v}_{(\mathbf{h}_1,\mathbf{h}_1)} - \mathbf{v}_{(\mathbf{h}_2,\mathbf{h}_2)})^\top \end{bmatrix} \mathbf{b} = \mathbf{0}$$

- Für n Ansichten des Schachbretts ergibt sich das Gleichungssystem $V\mathbf{b} = \mathbf{0}$ mit $V \in \mathbb{R}^{2n\times 6}$
- Wir benötigen also mindestens 3 Ansichten, um $\mathbf{b}$ bis auf Skalierung zu schätzen

Video

Kalibrierung

Schritte zur Berechnung der Kalibrierungsmatrix

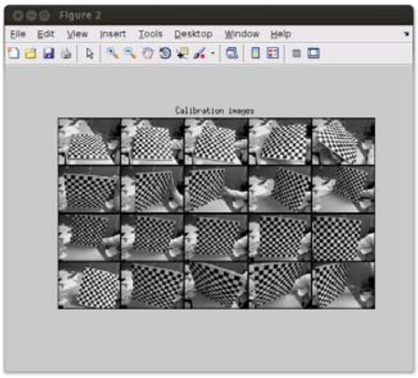
- Löse LGS $V\mathbf{b} = \mathbf{0}$ mittels Singulärwertzerlegung von V
- Erstelle symmetrische Matrix $\tilde{B}$ aus dem rechtsseitigen Singulärvektor zum kleinsten Singulärwert von V
- Wähle positiv definites $B = \pm\tilde{B}$
- Zerlege B mittels Cholesky-Faktorisierung in das Produkt $B = \tilde{K}^\top \tilde{K}$, wobei $\tilde{K}$ eine obere Dreiecksmatrix ist
- Dann gilt $K \sim \tilde{K}^{-1}$
- Normiere die Kalibrierungsmatrix an Hand ihres (3,3)-Eintrags

Video

MATLAB-Demonstration

Toolbox zur Kamerakalibrierung

- Jean-Yves Bouget, CalTech
- http://www.vision.caltech.edu/bouguetj/calib_doc/index.html



Video

Zusammenfassung

- Mit mindestens drei Aufnahmen eines Schachbretts lässt sich die Kalibrierungsmatrix bestimmen
- Erhalte pro Aufnahme eine Homographie zwischen 3D-Punkten auf dem Schachbrett und zugehörigen unkalibrierten Bildpunkten
- Jede Ansicht liefert zwei Gleichungen für $B = K^{-\top} K^{-1}$
- Extrahiere die Kalibrierungsmatrix mittels Cholesky-Faktorisierung