

Correspondence Problems in Computer Vision WS20/21

Gedächtnis Protokoll

Prüfer: Prof. Bruhn Beisitz: Lukas Mehl

Datum: 05.03.2021 Dauer: 30 Minuten

Die Prüfung konnte auf Englisch oder Deutsch durchgeführt werden (Englische Fachbegriffe waren kein Problem). Unterteilt wurde die Prüfung in vier Themenblöcke. Man durfte sich aussuchen, mit welchem Block angefangen wird. Geschrieben wurde an der Tafel.

1 Optical Flow

- (a) Was ist das Ziel von optischem Fluss?
- (b) Welche Annahme haben wir getroffen, um unsere Pixel wiederzufinden?
- (c) Wie sieht das ohne Linearisierung aus?
- (d) Welches Problem haben wir dabei?
- (e) Was können wir dennoch berechnen?
- (f) In welche Richtung zeigt der Normalflow?
- (g) Schreiben Sie das Energiefunktional von Horn & Schunck auf.
- (h) Allgemein, was müssen wir mit dem Funktional machen, um unsere Lösung zu erhalten?
- (i) Wie können wir das berechnen?
- (j) Schreiben Sie die allgemeinen Euler-Lagrange Gleichungen auf.
- (k) Was haben wir zusätzlich zu diesen zwei Gleichungen? Schreiben Sie diese allgemein auf.
- (l) Wie sehen die EL-Gleichungen zweiter Ordnung aus?
- (m) Was müssen wir machen, um diese PDEs zu lösen?
- (n) Sind die entstandenen Formeln linear oder nicht linear?
- (o) Wie groß ist das System/Wie viele Gleichungen hat es?
- (p) Wie gehen wir vor, wenn wir die BCCE nicht linearisieren?
- (q) Wie sieht das Warping aus?
- (r) Welches Bild genau wird in diesem Schritt gewarpt?

2 Stereo

- (a) Schreiben Sie den Epipolar Constraint auf.
- (b) Wie viele Freiheitsgrade hat die Matrix \mathcal{F} ?
- (c) Erklären Sie den Epipolar Constraint anhand eines Bildes.
- (d) Wenn wir alles über unser System wissen, und \mathcal{F} nicht schätzen müssen, wie berechnet man dann \mathcal{F} ?
- (e) Erklären Sie die Bestandteile der Formel.
- (f) Welche Matrix sorgt dafür, dass wir nicht vollen Rang haben?
- (g) Wie viele Freiheitsgrade hat A_{int}

3 Medical Image Registration

- (a) Welches Problem hat MIR?
- (b) Wie umgehen wir dieses Problem, welches Maß wählen wir?
- (c) Was ist Mutual Information?
- (d) Was genau sind $p(a)$ und $p(b)$? Auf welches Bild beziehen sie sich?
- (e) Wie sieht die Formel für *Curvature Based Regularisation* aus?
- (f) Warum machen wir das/Was ist der Effekt von dieser Regularisierung?
- (g) $(\Delta u)^2$ erlaubt nicht zu 100% affinen Flow. Wie können wir komplett affinen Flow erhalten?
(Tipp: Affiner Flow ist gegeben durch $s_1 xy + ax + by + c$, $\partial x \partial x$ eliminiert $s_1 xy$, wie bleibt es erhalten?)

4 Particle Image Velocimetry

- (a) Welche Regularisierung hatten wir bei PIV diskutiert? Schreiben Sie die Formeln auf.
- (b) Was beschreiben diese Konstrukte?
- (c) Wie sieht der Flow aus, den wir erhalten?
- (d) Welche Veränderung wurde getroffen, um mehr turbulenten Flow zu erlauben?
- (e) Wie sieht das am Beispiel von div aus?
- (f) Warum haben wir den Incompressibility constraint angenommen?
- (g) Was versuchen wir mit Navier Stokes vorauszusagen, div oder curl?

Antworten

Die Antworten hier sind nicht als Musterlösung zu verstehen. Keine Garantie auf Vollständigkeit oder Korrektheit.

1 Optical Flow

- (a) Was ist das Ziel von optischem Fluss?

Gegeben sind zwei Bilder (meist von der gleichen Szene). Das Ziel ist es, ein Flow Field zu bestimmen, also für jeden Pixel im ersten Bild zu sagen, wohin er sich im zweiten Bild verschoben hat.

- (b) Welche Annahme haben wir getroffen, um unsere Pixel wiederzufinden?

Die Grey-Value-Constancy assumption.

- (c) Wie sieht das ohne Linearisierung aus?

$$f(x, y, t) = f(x + u, y + v, t + 1) \iff f(x, y, t) - f(x + u, y + v, t + 1) = 0$$

- (d) Welches Problem haben wir dabei?

Wir haben zwei Unbekannte (u, v) aber nur eine Gleichung, um diese zu bestimmen. Dies bezeichnet man als Aperture Problem.

- (e) Was können wir dennoch berechnen?

Den Normalflow.

- (f) In welche Richtung zeigt der Normalflow?

Der Normalflow ist gegeben durch

$$\begin{pmatrix} u \\ v \end{pmatrix}_n = \frac{-f_t \nabla f}{|\nabla f| |\nabla f|}$$

Wir projizieren also auf den Image Gradient ∇f , der Flow geht in dessen Richtung.

- (g) Schreiben Sie das Energiefunktional von Horn & Schunck auf.

$$E(u, v) = \int_{\Omega} (f_x u + f_y v + f_t)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2) dx dy$$

- (h) Allgemein, was müssen wir mit dem Funktional machen, um unsere Lösung zu erhalten?

Wir müssen das Funktional minimieren.

- (i) Wie können wir das berechnen?

Wir müssen die Euler-Lagrange Gleichungen lösen.

- (j) Schreiben Sie die allgemeinen Euler-Lagrange Gleichungen auf.

$$\begin{aligned} 0 &\stackrel{!}{=} F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} \\ 0 &\stackrel{!}{=} F_v - \frac{\partial}{\partial x} F_{v_x} - \frac{\partial}{\partial y} F_{v_y} \end{aligned}$$

- (k) Was haben wir zusätzlich zu diesen zwei Gleichungen? Schreiben Sie diese allgemein auf.

Wir benötigen noch die Boundary Conditions:

$$\mathbf{n}^\top \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} = 0$$

$$\mathbf{n}^\top \begin{pmatrix} F_{v_x} \\ F_{v_y} \end{pmatrix} = 0$$

- (l) Wie sehen die EL-Gleichungen zweiter Ordnung aus?

$$0 \stackrel{!}{=} F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} + \frac{\partial^2}{\partial x \partial x} F_{u_{xx}} + \frac{\partial^2}{\partial x \partial y} F_{u_{xy}} + \frac{\partial^2}{\partial y \partial x} F_{u_{yx}} + \frac{\partial^2}{\partial y \partial y} F_{u_{yy}}$$

$$0 \stackrel{!}{=} F_v - \frac{\partial}{\partial x} F_{v_x} - \frac{\partial}{\partial y} F_{v_y} + \frac{\partial^2}{\partial x \partial x} F_{v_{xx}} + \frac{\partial^2}{\partial x \partial y} F_{v_{xy}} + \frac{\partial^2}{\partial y \partial x} F_{v_{yx}} + \frac{\partial^2}{\partial y \partial y} F_{v_{yy}}$$

- (m) Was müssen wir machen, um diese PDEs zu lösen?

Wir müssen sie diskretisieren.

- (n) Sind die entstandenen Formeln linear oder nicht linear?

Es entsteht ein lineares Gleichungssystem.

- (o) Wie groß ist das System/Wie viele Gleichungen hat es?

Das Gleichungssystem hat eine Größe von 2-mal die Anzahl der Pixel. u und v für jeden Pixel.

- (p) Wie gehen wir vor, wenn wir die BCCE nicht linearisieren?

Anstatt die BCCE zu linearisieren, linearisieren wir in der Numerik. Anstatt den Flow direkt zu berechnen, starten wir mit einer groben Version des Bildes. Für diese berechnen wir den Flow. Da dieser Flow natürlich nur grob ist, ist das Ergebnis nur ein Inkrement unseres Flows. Um den verbliebenen Flow zu berechnen, übertragen wir ihn auf eine feinere Stufe. Das heißt, wir passen das zweite Bild mit dem bisher berechneten Flow an (Warping). Anschließend berechnen wir die Differenz zwischen dem etwas feineren ersten und dem gewarpten zweiten Bild. Das wiederholt sich, bis wir auf der feinsten Stufe angekommen sind.

- (q) Wie sieht das Warping aus?

Beim Warping haben wir das (zweite) Bild und einen Flow gegeben. Das Ergebnis, also das gewarpte Bild, ist anfangs leer. Nun wird für jeden Pixel dieses leeren Bildes geschaut, an welcher Position dieser Pixel mit dem Flow landet. Da der Flow nicht ausschließlich aus ganzen Zahlen besteht, kann die Position zwischen den Pixeln liegen. Um das sinnvoll auszugleichen, nutzen wir bilineare Interpolation und schreiben diesen Wert in das Ergebnis-Bild zurück. So entsteht das gewarpte Bild.

- (r) Welches Bild genau wird in diesem Schritt gewarpt?

Wir warpen das zweite Bild und berechnen den verbliebenen Flow zwischen dem ersten und dem gewarpten zweiten Bild.

2 Stereo

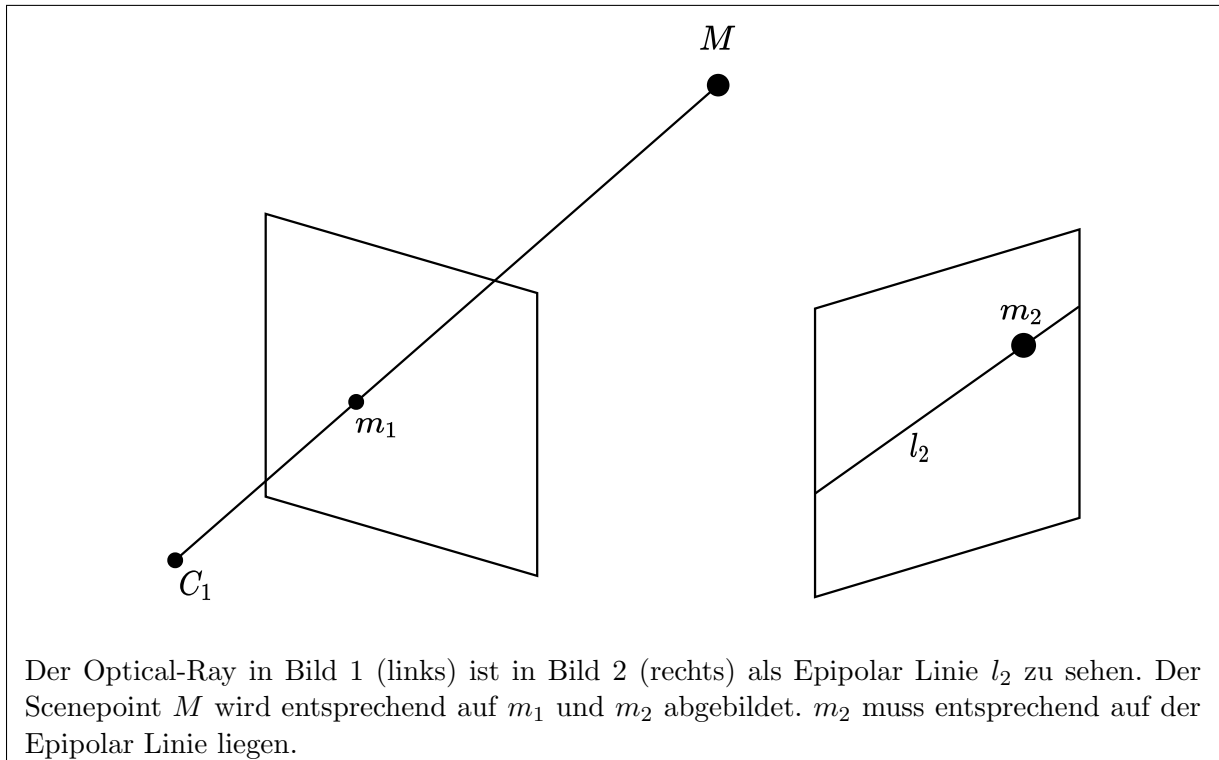
- (a) Schreiben Sie den Epipolar Constraint auf.

$$\tilde{m}_2^\top \mathcal{F} \tilde{m}_1 = 0$$

- (b) Wie viele Freiheitsgrade hat die Matrix \mathcal{F} ?

\mathcal{F} hat 7 Freiheitsgrade (-Scale, -Rang).

- (c) Erklären Sie den Epipolar Constraint anhand eines Bildes.



- (d) Wenn wir alles über unser System wissen, und \mathcal{F} nicht schätzen müssen, wie berechnet man dann \mathcal{F} ?

$$\mathcal{F} = A_{int2}^{-T} [t]_{\times} \mathcal{R} A_{int1}^{-1}$$

- (e) Erklären Sie die Bestandteile der Formel.

A_{int} ist die Intrinsische Matrix. Sie beschreibt die Transformation von Bild- zu Pixel-Koordinaten.

$[t]_{\times}$ ist eine Matrix, die ein Kreuzprodukt mit dem Translationsvektor t realisiert. t ist dabei die Verschiebung zwischen den beiden Kameras.

\mathcal{R} ist eine Rotationsmatrix, die die Rotation zwischen den Kameras beschreibt.

- (f) Welche Matrix sorgt dafür, dass wir nicht vollen Rang haben?

$[t]_{\times}$ hat lediglich Rang 2.

- (g) Wie viele Freiheitsgrade hat A_{int}

A_{int} hat 5 Freiheitsgrade. 2 für den Shift des Ursprungs. 2 für die Pixeldichte in x - und y -Richtung. Und einen für die Schiefe des Koordinatensystems.

3 Medical Image Registration

- (a) Welches Problem hat MIR?

Das Problem ist, dass wir zwei Bilder haben, deren Grauwerte Unterschiedliches aussagen. Ein weißer Pixel im einen Bild gehört vielleicht zu einem schwarzen im anderen. Dennoch wollen wir ein Verformungsfeld zwischen den beiden Bildern.

- (b) Wie umgehen wir dieses Problem, welches Maß wählen wir?

Wir nutzen die Mutual Information.

- (c) Was ist Mutual Information?

Mutual Information ähnelt der Kullback-Leibler Divergenz. Diese misst den Abstand zweier Verteilungen. Wählt man für die zwei Verteilungen die gemessene Joint Probability $p(a, b)$ und die Probability für den unabhängigen Fall $p(a) \cdot p(b)$, so erhält man die Mutual Information.

- (d) Was genau sind $p(a)$ und $p(b)$? Auf welches Bild beziehen sie sich?

$p(a)$ ist die Wahrscheinlichkeit, an einer beliebigen Stelle im ersten Bild den Grauwert a anzutreffen. $p(b)$ entsprechend den Wert b im zweiten Bild. $p(a, b)$ beschreibt die Wahrscheinlichkeit, an einer Stelle im ersten Bild den Wert a zu finden und im zweiten Bild an der gleichen Stelle den Wert b .

- (e) Wie sieht die Formel für *Curvature Based Regularisation* aus?

$$\alpha((\Delta u)^2 + (\Delta v)^2)$$

- (f) Warum machen wir das/Was ist der Effekt von dieser Regularisierung?

Damit erlauben wir affinen Flow.

- (g) $(\Delta u)^2$ erlaubt nicht zu 100% affinen Flow. Wie können wir komplett affinen Flow erhalten?

(Tipp: Affiner Flow ist gegeben durch $s_1 xy + ax + by + c$, $\partial x \partial x$ eliminiert $s_1 xy$, wie bleibt es erhalten?)

Wir brauchen eine Regularisierung, die einmal nach x und einmal nach y ableitet. Ein Möglichkeit dafür wären alle zweiten Ableitungen, diese finden sich z.B. in der Hesse-Matrix.

4 Particle Image Velocimetry

- (a) Welche Regularisierung hatten wir bei PIV diskutiert? Schreiben Sie die Formeln auf.

Div-Curl-Regularisierung.

$$\left(\operatorname{div} \begin{pmatrix} u \\ v \end{pmatrix} \right)^2 = (u_x + v_y)^2$$
$$\left(\operatorname{curl} \begin{pmatrix} u \\ v \end{pmatrix} \right)^2 = (v_x - u_y)^2$$

- (b) Was beschreiben diese Konstrukte?

Eine hohe Divergence bedeutet, dass Partikel auftauchen oder verschwinden. Curl beschreibt eine Rotation der Partikel ähnlich einem Strudel. Die Regularisierung verhindert, dass solche Phänomene in unserem Flow auftauchen.

- (c) Wie sieht der Flow aus, den wir erhalten?

Wir erhalten laminaren Flow, also Flow, bei dem die Partikel sich parallel bewegen.

- (d) Welche Veränderung wurde getroffen, um mehr turbulenten Flow zu erlauben?

Wir haben nicht direkt div und curl bestraft, sondern deren Ableitungen. Die Annahme ist, dass div und curl sich smooth ändern.

(e) Wie sieht das am Beispiel von div aus?

$$\left| \nabla \operatorname{div} \begin{pmatrix} u \\ v \end{pmatrix} \right|^2$$

(f) Warum haben wir den Incompressibility constraint angenommen?

Flüssigkeiten lassen sich schwer komprimieren. Es ist deshalb sinnvoll, Dinge wie Divergenz im Flow Field zu bestrafen.

(g) Was versuchen wir mit Navier Stokes vorauszusagen, div oder curl?

curl

Anmerkungen

Prof. Bruhn ist ein sehr angenehmer Prüfer, der einem auch hilft, sollte man nicht sofort die Antwort wissen (siehe Tipp). Die Atmosphäre war entsprechend entspannt. Manchmal will er die Dinge gar nicht im Detail erklärt haben, sondern ihm reicht ein Stichwort. Bsp. Aufgabe 3 c), hier war ihm wichtig, dass es sich um ein Abstandsmaß handelt.