Projekt 1: Schätzung der lokalen Orientierungen eines Bildes

(a) Ein skalares Bild $u(x)$ sei nahe des Punktes x_0 ungefähr in Richtung des (normierten) Vektors $v_* \in \mathbb{R}^2$ orientiert. Begründen Sie das Modell

$$v_* \cdot \nabla u(x) \approx 0 \qquad (x \approx x_0), \tag{1}$$

indem Sie eine Schraffur betrachten, deren Höhenlinien überall *exakt* in Richtung von v_* ausgerichtet sind. Ordnen Sie jetzt jedem Punkt x_0 diejenige lokale Orientierung $v_* = v(x_0)$ zu, die (1) im Least-Squares-Sinne *optimal* löst. Sie sollten einen Ausdruck der Form

$$v_* = \operatorname{argmin}_{|v|=1} \int_{\mathbb{R}^2} |v \cdot \nabla(x)|^2 \, K_\rho(x_0 - x)\, dx$$

mit einem geeigneten Lokalisierungskern K_ρ erhalten. Zeigen Sie schließlich, dass dieses v_* Eigenvektor zum minimalen Eigenwert der symmetrischen, positiv-semidefiniten (warum?) 2×2 Matrix (der sog. „Strukturtensor“ von W. Förstner (1986))

$$K_\rho \star (\nabla u \nabla u^T)(x_0)$$

ist. Interpretieren Sie die Bedeutung beider Eigenwerte und definieren Sie ein geeignetes Maß der „Kohärenzstärke“ der erhaltenen lokalen Orientierung.

(b) Wenden Sie die Orientierungsschätzung auf das Beispielbild eines Fingerabdrucks aus der Vorlesung an. Begründen Sie anhand eines geeigneten Modellbeispiels, warum bei derartig verrauschten Bildern eine *Vorglättung* $u \mapsto u_\sigma = K_\sigma \star u$ unverzichtbar ist. Studieren Sie die Robustheit der Schätzung bezüglich unterschiedlich starkem Rauschen und geben Sie Empfehlungen zur Wahl der Parameter. Wie verhält es sich mit der Kohäherenzstärke?

(c) Verallgemeinern Sie die Orientierungsschätzung auf den Fall vektorwertiger Bilder (z.B. Farbbilder). Gesucht ist *eine* gemeinsame Richtung $v_* = v(x_0)$ für alle Kanäle. Präzisieren Sie Ihre Verallgemeinerung für den Fall von Farbkanälen u_1, u_2, u_3, denen ein Grauwertbild in der Form $u_{grau} = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3$ zugeordnet ist. Wie unterscheidet sich das Ergebnis von einer Richtungsschätzung für u_{grau}?

(d) Verallgemeinern Sie die Orientierungsschätzung auf den Fall zweier (transparent) überlagerter skalarer Bilder $u = \lambda u_1 + (1-\lambda)u_2$, $\lambda \in (0,1)$. Gesucht sind also diesmal für gegebenes u *zwei* Richtungsvektoren $v_1(x_0)$ und $v_2(x_0)$, die zu den (nicht explizit gegebenen) Teilbildern u_1 und u_2 gehören. Wenden Sie Ihre Schätzung auf den Fall eines verrauschten Schachbrettmusters an. Was würde in diesem Fall hingegen die Schätzung einer einzigen Richtung liefern?

Literatur (nur falls Sie für (d) keine eigene Lösung finden):

T. Aach, C. Mota, I. Stuke, M. Mühlich, E. Barth, *Analysis of Superimposed Oriented Patterns*, IEEE Trans. Imag. Proc. 15, 3690–3700 (2006).

Projekt 2: Sind natürliche Bilder von beschränkter Variation?

Der Raum BV der Funktionen beschränkter Variation wurde lange als *das* Modell für Bilder angesehen. Eine sorgfältige statistische Analyse *natürlicher* Bilder führte Y. Gousseau und J.-M. Morel in dem bemerkenswerten Aufsatz „Are Natural Images of Bounded Variation?“ (SIAM J. Math. Anal. 33, 634–648, 2001) aber zur Gegenthese $u \notin BV$. Die prinzipielle Schwierigkeit der wissenschaftlichen Untermauerung einer solchen Gegenthese besteht darin, aus einem gegebenen digitalen Bild (das natürlich vordergründig in BV liegt) auf die quantitative Struktur der kleinen Skalen des „eigentlichen“, natürlichen Bildes zu extrapolieren. Gousseau und Morel gelang das, indem sie die Statistik der Flächen gewisser Objekte betrachteten: Bezeichnet $f(a)$ die Anzahl solcher Objekte mit einer Fläche von a Pixeln, so beobachtet man mit überraschender Genauigkeit

$$f(a) \approx Ca^{-\alpha}. \tag{2}$$

Die Analyse einer Datenbank natürlicher Bilder zeigt, dass für die meisten Bilder $\alpha \approx 2$ gilt. Für $\alpha > 3/2$ konnten Gousseau und Morel aber *beweisen*, dass $u \notin BV$.

(a) Skizzieren Sie die wesentliche Struktur des Beweises des zentralen Theorems 4.5 von Gousseau und Morel. Gehen Sie insbesondere auf die Bedeutung von Coarea-Formel und isoperimetrischer Ungleichung ein.

(b) Reproduzieren Sie (qualitativ) die Figur 2 aus dem Paper (das Airportbild findet sich auf unserer Webseite; der Befehl `bwconncomp` aus der Image Processing Toolbox von Matlab hilft bei der Berechnung von $f(a)$). Beachten Sie hierzu, dass Gousseau und Morel nur die Flächen bis zu jenem kleinstmöglichen A_{max} betrachten, für welches es keine Komponente mit $A_{max}+1$ Pixeln gibt. Größere a verhalten sich völlig erratisch und spielen für die Extrapolation $a \to 0$ der Verteilung von Mikrostrukturen auch ganz offensichtlich keine Rolle.

(c) Reproduzieren Sie (qualitativ) Tabelle 1 des Papers (der Least-Squares-Fehler E ist mit der Wurzel der Anzahl der Terme zu skalieren). Die Datenbank mit den 4167 Fotos von J. van Hateren finden Sie im Web unter der Adresse

`www.kyb.mpg.de/bethge/vanhateren/index.php`.

Dort wird auch erklärt, wie sich die Bilder in Matlab einlesen lassen. Beachten Sie bitte, dass es sich bei den Bildern um RAWs handelt, deren u-Wert proportional zur Lichtleistung ist. (Ohne γ-Korrektur erscheinen die Bilder am Schirm daher viel zu dunkel). Untersuchen Sie den Einfluss anderer Bildnormierungen wie etwa der Histogramm-Äquilibrierung (Matlab: `histeq`) auf die Statistik (was erwarten Sie aufgrund einer theoretischen Vorüberlegung, etwa anhand eines Modells?).

(d) Suchen Sie nach Ausreißern mit besonders kleinem $\alpha < 3/2$. Wie unterscheiden sich diese Beispiele von den restlichen Bildern der Datenbank? Ist für die Ausreißer das Modell (2) noch hinreichend genau? Wenn ja, ist BV hier ein plausibler Raum? Wiederholen Sie die Datenbankanalyse für ein paar Themengebiete aus `www.flickr.com`, für die Sie entweder viel (z.B. Landschaften) oder wenig (z.B. synthetische Bilder oder Architektur) Mikrostruktur erwarten. Welche α beobachten Sie hier?