1 Einführung

Das 20. Jahrhundert markiert in weiten Teilen der Gesellschaft den vollständigen Übergang von akustisch zu visuell dominierter Wahrnehmung [1]. Prägend dafür ist die (Weiter-)Entwicklung von Fotografie, Film, Fernsehen (jeweils zunächst analog, dann digital) und zuletzt kleinerer Geräte mit Display.

Analog dazu, wenn auch mit ein wenig Verspätung, etabliert sich der Übergang auch in der Wissenschaft, in der sich ein neues Feld der *Imaging Sciences* (noch ohne etablierte deutsche Übersetzung) quer über verschiedenste Disziplinen gebildet hat. Besonders stark vertreten ist Imaging in der Medizin und Biologie, verschiedenste Verfahren zur Bildgebung wurden in den letzten vier Jahrzehnten entwickelt und teilweise klinisch nutzbar gemacht. Andererseits funktionieren diese Verfahren zu großen Teilen noch weniger automatisch als Film und Fotografie. Insbesondere bei der Bildgebung in vivo sind viele Fragen noch offen bzw. es besteht in den meisten Fällen noch deutliches Verbesserungspotential. Dies betrifft u.a. alle Bereiche der Naturwissenschaft, etwa die Physik bei der Konstruktion von Detektoren oder Lasern, die Chemie zum Design von Tracern in der molekularen Bildgebung, und nicht zuletzt die Informatik und Mathematik zur Rekonstruktion und automatischen Analyse der Bilder. Mit den letzten Aspekten werden wir uns in dieser Vorlesung beschäftigen, wobei wir uns auf sogenannte Variationsmethoden konzentrieren werden, d.h. Probleme in denen wir verbesserte Bilder durch Minimierung geeigneter Energiefunktionale erhalten wollen.

[1] Franz Hillenkamp.

1.1 Mathematische Konzepte von Bildern

Wir beginnen mit einer geeigneten Definition eines Bildes bzw. der wichtigsten Strukturteile eines Bildes. Wir werden zur Vereinheitlichung keinen speziellen Unterschied zwischen klassischen Bildern (2D) und Volumenbildern (3D) machen, wir werden selbst Filme (2+1D oder 3+1D) als Bilder bezeichnen. Die Bilddimension werden wir mit $d \in \{2, 3, 4\}$ bezeichnen.

Bei der mathematischen Modellierung eines Bildes unterscheiden wir zwischen zwei Versionen

- Das (idealisierte) kontinuierliche Bild als Funktion $u : \Omega \to \mathbb{R}$, wobei $\Omega \subset \mathbb{R}^d$ das Gebiet des Bildes (meist ein Rechteck oder Würfel) ist.
- Das digitale Bild als Matrix (Tensor) $U \in \mathbb{R}^{N_1 \times N_2 \times \ldots \times N_d}$.

Zwischen der kontinuierlichen und digitalen Version der Bilder besteht ein offensichtlicher Zusammenhang, der auch durch die Visualisierung von Bildern definiert ist. Dabei unterteilt man das Gebiet Ω in $N_1 \times N_2 \times \ldots \times N_d$ kleine Rechtecke (Pixel) oder Würfel (Voxel), die in den Grauwerten definiert durch die Einträge von U eingefärbt werden. Damit entspricht also jedem digitalen Bild sofort ein kontinuierliches, das jeweils innerhalb der Pixel (wir werden zur Vereinfachung die Unterscheidung zwischen Pixel und Voxel nicht aufrechterhalten) konstant ist. Umgekehrt kann man aus jedem kontinuierlichen Bild sofort ein digitales, indem man die Mittelwerte von u über alle Pixel berechnet.

Auf den ersten Blick erscheint die Definition eines kontinuierlichen Bildes als beliebige mathematische Idealisierung. In der Praxis zeigt sich jedoch, dass diese Definition grundlegend für viele verschiedene Aspekte der Bildverarbeitung ist, etwa die Folgenden:

- Digitale (diskrete) Bilder können in verschiedener Auflösung vorliegen, d.h. mit verschiedener Pixelanzahl. Um den Zusammenhang zwischen verschiedenen Auflösungen zu verstehen ist es sehr nützlich alle möglichen digitalen Bilder als diskrete Versionen (durch Mittelung in Pixeln) eines zu Grunde liegenden kontinuierlichen Bildes zu sehen. Durch die Konsistenz mit dem Übergang zwischen kontinuierlichen und digitalen Bildern ist auch der Übergang zwischen verschiedenen digitalen Bildern klar definiert.
- Viele intuitiv klar erfassbare Eigenschaften von Bildern sind in digitalen Bildern schwer zu definieren. Ein wichtiges Beispiel dafür sind Kanten, die wir als Übergang zwischen sehr verschiedenen Grauwerten verstehen. Während die Definition von Kanten im kontinuierlichen Bild als die Menge der Unstetigkeiten

der Funktion u naheliegend und eindeutig ist, ist die Definition im digitalen Fall schwierig und nicht eindeutig, da ja meist von jedem Pixel zum nächsten der Wert wechselt. Eine Kante kann im digitalen Bild aber sinnvoller als die Mittelung der Kante im kontinuierlichen Bild über Pixel (in der jeweiligen Auflösung) erhalten werden.

- Die mathematische Modellierung von Bildverarbeitungsaufgaben kann im kontinuierlichen Setup sehr einheitlich durchgeführt werden, hier hat man alle grundlegenden Konzepte der Integral- und Differentialrechnung zur Verfügung. Wie wir sehen werden sind diese fast in allen Techniken von enormer Bedeutung. Hat man ein Variationsmodell für das kontinuierliche Bild konstruiert, kann man dies durch eine Diskretisierung mit Standardmethoden der Numerik sofort konsistent auf alle diskreten Bilder, unabhängig von der Auflösung, übertragen. Insbesondere ist auch die Kondition des Problems interessant, die für hohe Auflösung bei sinnvoller Diskretisierung gegen die Kondition des kontinuierlichen Problems konvergiert. Folglich sollte man versuchen die Kondition des kontinuierlichen Problems zu analysieren, was insbesondere durch geeignete Wahl von Regularisierungsfunktionalen zu realisieren ist und auf interessante mathematische Fragen führt, die sich in den meisten praktischen Anwendungen aber zufriedenstellend beantworten lassen.

- Das Design von Algorithmen wird ebenfalls konsistenter wenn man ein zu Grunde liegendes kontinuierliches Bild annimmt. Zum Beispiel möchte man vermeiden, dass ein iterativer Algorithmus bei Erhöhung der Auflösung eine deutlich höhere Anzahl an Iterationen benötigt um eine gewisse Auflösung zu erreichen. Dies ist gewährleistet wenn man den Algorithmus als Diskretisierung eines Algorithmus für das kontinuierliche Bild erhält. Letzterer muss dann als Algorithmus für Funktionen, d.h. in Funktionenräumen konstruiert werden, analog zu Verfahren für partielle Differentialgleichungen.

1.2 Variationsmethoden

Variationsmethoden sind eine beliebte und recht allgemein einsetzbare Technik zur Verarbeitung von a-priori Information in der Bildgebung. Dabei konstruiert man ein Optimierungsproblem mit einem Funktional aus zwei Termen. Der erste Teil dient zur Approximation der Daten, die wir im kontinuierlichen Fall als f und im digitalen Fall als F bezeichnen, der zweite ist das Regularisierungsfunktional, das die a-priori Information einfliessen lässt. Damit erhalten wir ein Problem der Form

$$J(u) = D(u, f) + \alpha R(u) \to \min_u \tag{1.1}$$

wobei weitere a-priori Information in der Klasse der Bilder u über die wir minimieren enthalten sein kann (etwa Nichtnegativität bei Dichtebildern). Der Datenterm D sollte so konstruiert sein, dass D sein Minimum (meist normiert als Null) erreicht, wenn die Daten exakt reproduziert werden. Das Regularisierungsfunktional sollte so konstruiert werden, dass R einen kleinen Wert annimmt, wenn das Bild sehr gut mit der a-priori Information übereinstimmt, und einen sehr grossen Wert, wenn das Bild nicht gut zur vorhandenen a-priori Information passt. Damit werden solche Bilder wohl kaum ein Minimum des Funktionals J werden. Der Parameter α gewichtet die beiden Anteile, für sehr kleines α werden vor allem die Daten reproduziert, mit wachsendem α erhält die Regularisierung stärkeren Anteil, dies ist besonders wichtig wenn Datenfehler oder Rauschen zu erwarten ist.

Wir wollen das Prinzip zunächst an einem einfachen Beispiel verdeutlichen. Wir suchen eine kleine Lösung der quadratischen Gleichung

$$u^2 - 2f_1u + f_2, \tag{1.2}$$

mit den gegebenen Daten $f_1 = 1000.001$ and $f_2 = 1$. Mit der Lösungsformel für quadratische Gleichungen erhalten wir zwei Lösungen

$$u_1 = 0.001 \qquad \text{und} \qquad u_2 = 1000. \tag{1.3}$$

Mit unserer a-priori Information einer kleinen Lösung wählen wir offensichtlich die Lösung u_1 aus. Dies ist allerdings in dem speziellen Fall nur möglich, da wir tatsächlich alle Lösungen berechnen können. Bei komplizierteren Problemen ist dies nicht möglich, dazu müssen wir mit Datenfehlern rechnen, also sollten wir besser robust eine approximative Lösung ausrechnen als mehrere exakte Lösungen. Ein Variationsansatz kann in der Form

$$\frac{1}{2}(u^2 - 2f_1u + f_2)^2 + \alpha|u|^p \to \min_{u\in\mathbb{R}} \tag{1.4}$$

realisiert werden. Für $p > 1$ und α klein hat dieses Funktional ein globales Minimum in der Nähe von $u = 0.0001$, approximiert also automatisch die kleinere Lösung.

1.3 Mathematische Aufgaben in der Bildgebung

Biomedizinische Bildgebung ist ein interdisziplinäres Gebiet, mit Fragestellungen in verschiedenen Gebieten. Wir werden im Folgenden einige mathematische Aufgaben kurz vorstellen, auch um einige Grundbegriffe zu klären.

- **Rekonstruktion (Reconstruction):** In den seltensten Fällen entspricht das aufgenommene Bild tatsächlich den Grau- oder Farbwerten, an denen man interessiert ist. So nimmt man etwa bei einem Computertomographen Mittelwerte der Dichte über Linien durch das Objekt auf, interessiert sich aber für die Dichte selbst, die im Bild dargestellt werden soll. Analoge Probleme treten auch bei allen anderen medizinischen bildgebenden Verfahren auf, sowie bei vielen anderen der oben genannten Anwendungen. Damit benötigt man als Grundlage jeder Weiterverarbeitung zunächst die Bildrekonstruktion, üblicherweise ein schlecht gestelltes Problem bei dessen (approximativer) Lösung einige Information verloren geht.

- **Entzerren (Deblurring):** Ein besonders häufig auftretendes Beispiel der Bildrekonstruktion ist die Entzerrung, die mathematisch meist auf die Lösung einer Faltungsgleichung (oder allgemeiner Integralgleichung) führt. Die Verzerrung entsteht, weil man (z.B. bei Fotografietechniken) nicht direkt den Farbwert jedes Punktes, sondern eigentlich eine lokale (oder manchmal auch globale) Mittelung aufnimmt. Passiert die lokale Mittelung auf einer hinreichend kleinen Skala, so wird keine Entzerrung durchgeführt, sondern direkt das verzerrte Bild betrachtet. Den typischen Radius bei der lokalen Mittelung bezeichnet man dann als Auflösungsgrenze des bildgebenden Verfahrens (bzw. der Kamera, des Mikroskops, ..). Die Auflösung liesse sich dann durch eine Entzerrung eventuell noch weiter steigern, allerdings ist in vielen Fällen die Art der lokalen Mittelung (die rückgängig gemacht werden soll) nicht genau bekannt. Mathematisch führt dies auf die Lösung einer Faltungsgleichung mit unbekanntem Kern, und Lösungsansätze für solche Probleme werden als blinde Entzerrung (Blind Deconvolution) bezeichnet.

- **Entrauschen (Denoising):** Aufgrund von Messfehlern, Übertragungsfehlern oder ähnlichem hat man oft mit verrauschten Bildern zu tun, bei denen einzelne Pixelwerte komplett falsch oder teilweise gestört sind. Entrauschen bezeichnet dann die Aufgabe das Rauschen zu eliminieren, bzw. Bilder ähnlich zum verrauschten Bild zu erzeugen, die aber für das menschliche Auge "vernünftiger" aussehen.

- **Kontrastkorrektur (Contrast Correction):** Aus verschiedenen Gründen (zu wenig Licht, falsche Kameraeinstellungen, zu ähnliche Objekte) erhält man manchmal Bilder mit wenig Kontrast, d.h. alle auftretenden Grau- oder Farbwerte sind in einem sehr ähnlichen Bereich. Um die Unterschiede besser sichtbar zu machen, führt man dann eine Kontrastkorrektur durch, um die Spanne der Farbwerte zu vergrössern. Dafür hat man mehrere Möglichkeiten (lineare oder nichtlineare Transformationen etc.), und die Hauptfragestellung ist dabei geeignete Transformationen zu finden, die genau die interessanten Teile sichtbar machen.

- **Segmentierung (Segmentation):** Als Segmentierung bezeichnet man das automatische Erkennen von Objekten ("objektbasierte Segmentierung") oder von Kanten ("kantenbasierte Segmentierung") in einem Bild. In den meisten Fällen hängen Kanten und Objekte auch direkt zusammen, den die Objektgrenzen sind ja meist als Kanten in einem Bild sichtbar. Schwierig wird die Segmentierung dadurch, dass man eigentlich nur wichtige Kanten (bzw. solche die auch das menschliche Auge als Kanten klassifizieren würde) erkennen sollte. Die Segmentierung ist deshalb ein Aufgabengebiet der Computer Vision.

  Mathematisch definiert man eine Kante meist als eine abrupte Änderung des Grau- und Farbwerts im digitalen Modell, bzw. eine Unstetigkeit der Intensitätsfunktion im Kontinuumsmodell.

- **Bildzerlegung (Image Decomposition):** Die Bildzerlegung ist eigentlich die Synthese aus Techniken wie Entrauschen und Segmentierung. Bei der Zerlegung versucht man die verschiedenen Anteile des Bildes, wie etwa Kanten, Texturen oder den Cartoon, zu trennen, um daraus weitere Rückschlüsse ziehen zu können.

- **Registrierung (Registration):** Vor allem in medizinischen Anwendungen tritt oft das Problem auf, dass man Information aus zwei Bildern zusammensetzen oder vergleichen will, die aus verschiedenen Modalitäten (z.B. Computertomograph und Ultraschall) gewonnen worden. Dabei tritt neben den Eigenheiten jeder Modalität auch das Problem auf, dass Teile des Bildes bewegt wurden (der Patient und seine Körperteile sind eventuell in verschiedener Lage, seine Organe ändern Grösse und Form wie z.B. das schlagende Herz, ...). Registrierung bezeichnet dann die Aufgabe diese Bewegungen und Verformungen zu korrigieren, sodass man zwei Bilder erhält, in denen zwei Punkte in den Bildern mit jeweils gleichen Koordinaten auch dem selben Punkt in der Realität entsprechen. Ein verwandtes Problem ist auch das *Warping*, bei dem man an einem ganzen, möglichst fliessenden, Übergang zwischen zwei Bildern. Solche Effekte werden z.B. im Fernsehen gerne verwendet um einen Übergang zwischen zwei Gesichtern (verschiedene Personen oder gleiche Person zu verschiedenen Zeiten) darzustellen.

- **Restauration (Inpainting):** Die Restauration von Bildern ohne Computer ist ein Problem, mit dem viele Museen konfrontiert sind. Dabei werden z.B. beschädigte Bereiche eines Bildes wieder mit Farbe gefüllt, sodass die Restauration ein plausibles Bild ergibt. Ein analoges Problem stellt sich in der digitalen Welt, wenn man z.B. den Schmutz in einem gescannten Bild ausbessern möchte. Man muss das gegebene Bild in einer gewissen (meist nicht allzu grossen) Bildregion fortsetzen (weitermalen, daher der engl. Name Inpainting). Dies automatisch zu tun, damit es für das menschliche Auge vernünftig aussieht, ist eine nicht zu unterschätzende (mathematische) Aufgabe.

- **Bildfusion (Matching):** Bildfusion bezeichnet das Zusammenbringen von Teilen verschiedener Bilder zu einem einzigen Bild. Beispiele dafür sind Einzelbilder in verschiedene Richtungen, die zu einem Panoramabild zusammengesetzt werden sollen, oder mehrere Gruppenfotos, bei denen von jeder Person das beste Bild benutzt werden soll. Schwierigkeiten entstehen dabei durch möglicherweise verschiedenen Kontrast der Bilder sowie beim Verschmelzen der Ränder von einzelnen Teilbildern.

Neben den klassischen Aufgaben der Bildverarbeitung betrachtet man heute auch viele Aufgaben für Videos und Grafiken mit ähnlichen Techniken:

- **Bewegungserkennung (Motion Detection):** Bei der Bewegungserkennung versucht man aus einer Bildsequenz die sich bewegenden Objekte und deren Geschwindigkeiten zu erkennen. Dies ist kein einfaches Problem und auch kein eindeutig lösbares, wenn man keine Informationen über Vorder- und Hintergrund hat. Ein Beispiel ist eine weiße Scheibe, die sich vor schwarzem Hintergrund bewegt. Man kann dabei nicht unterscheiden, ob sich der weiße oder schwarze Anteil bewegt, denn es könnte sich ja auch eine schwarze Schablone vor ruhigem weißen Hintergrund bewegen.
- **Oberflächenglättung (Surface Smoothing):** Zur Visualisierung oder Restaurierung von Oberflächen ist Oberflächenglättung nötig, die kleine Oszillationen entfernt oder die Oberfläche glatt interpoliert. Die mathematischen Aufgaben dabei sind analog zum Entrauschen oder zur Restauration von Bildern.