

Universität des Saarlandes



Fachrichtung 6.1 – Mathematik

Preprint Nr. 310

**Mathematische Bildverarbeitung mit Ideen
aus der Natur**

Joachim Weickert

Saarbrücken 2012

Mathematische Bildverarbeitung mit Ideen aus der Natur

Joachim Weickert

Universität des Saarlandes
Fakultät für Mathematik und Informatik
Campus E1.7
66041 Saarbrücken
Deutschland
`weickert@mia.uni-saarland.de`

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-Mail: preprint@math.uni-sb.de
WWW: <http://www.math.uni-sb.de/>

Dieser Beitrag ist eine allgemeinverständliche Darstellung einiger Grundideen zur mathematischen Bildanalyse. Er wird als eingeladenen Artikel in den Mitteilungen der DMV erscheinen.

Einleitung

Die Faszination der angewandten Mathematik beruht nicht zuletzt auf der Tatsache, dass sie sich als ein äußerst erfolgreiches Modellierungswerkzeug etabliert hat. Viele Vorgänge in Natur und Technik lassen sich durch mathematische Gleichungen sehr kompakt und oft auch mit hoher Genauigkeit darstellen. Differentialgleichungen, die Zusammenhänge zwischen einer unbekannten Funktion und ihren Ableitungen beschreiben, haben sich dabei als besonders wichtige Modelle erwiesen. Da es häufig angebracht ist, Funktionen mit mehreren Variablen zu verwenden, können partielle Ableitungen auftreten. Entsprechende Gleichungen nennt man partielle Differentialgleichungen (englisch: *partial differential equations*, kurz *PDEs*). Seit vielen Jahren weiß man, dass zahlreiche physikalische Vorgänge, wie z.B. Wärmeleitung, Diffusion, strömungsmechanische Prozesse und elektrodynamische Phänomene, partiellen Differentialgleichungen genügen. Auf mathematischer Seite hat man solche Gleichungen intensiv untersucht und hocheffiziente numerische Techniken entwickelt, um sie schnell und mit hoher Genauigkeit auf Computern zu approximieren. Damit hat sich die numerische Simulation zu einem dritten Standbein neben der Theorie und dem Experiment etabliert und das Gebiet des wissenschaftlichen Rechnens begründet, das zahlreichen Mathematikern ein spannendes Betätigungsfeld bietet.

Weniger bekannt ist jedoch die Tatsache, dass sich auch wichtige Teilgebiete in der digitalen Bildverarbeitung und der Computergrafik mit großem Erfolg mathematischer Modelle bedienen, die auf partiellen Differentialgleichungen beruhen. Dabei wurden zahlreiche Anstrengungen unternommen, praktisch das gesamte Spektrum dieser Prozesse nutzbar zu machen, angefangen von elliptischen über parabolische bis hin zu hyperbolischen Gleichungen. Ziel dieses Beitrags ist es zu illustrieren, wie einige dieser grundlegenden Modelle aus Natur und Technik Eingang in die moderne Bildverarbeitung und Computergrafik gefunden haben. Nicht zuletzt diesen Erfolgen ist es zu verdanken, dass sich die sogenannte mathematische Bildverarbeitung als eigenständiges Forschungsgebiet etabliert hat, bei dem Deutschland international sehr gut positioniert ist. Wir werden Prozesse betrachten, die Inspirationen aus fünf Bereichen schöpfen: Elektrostatik, Diffusion, Wellenausbreitung, Wärmeleitung und Osmose.

1 Bilder als Funktionen

Um überhaupt mathematische Bildverarbeitung betreiben zu können, müssen wir zunächst definieren, was man sich als Mathematiker unter einem Bild vorstellen sollte. Hierzu betrachten wir einen rechteckigen Bildbereich $\Omega \subset \mathbb{R}^2$.

Unter einem kontinuierlichen Grauwertbild verstehen wir eine Abbildung $f : \Omega \mapsto \mathbb{R}$. Es handelt sich also um eine skalarwertige Funktion mehrerer Veränderlicher, die jeder Ortskoordinate (x, y) einen Grauwert $f(x, y)$ zuordnet. Typischerweise interpretiert man große Grauwerte als helle Bildstrukturen und ordnet dunklen Bildstrukturen niedrige Grauwerte zu. Mit diesem Funktionsbegriff kann man Bilder mit Werkzeugen der mehrdimensionalen Analysis untersuchen. Beispielsweise identifiziert man Bildkanten mit Orten, an denen der Grauwertgradient $\nabla f = (\partial_x f, \partial_y f)^\top$ einen besonders großen Betrag aufweist (Abb. 1).

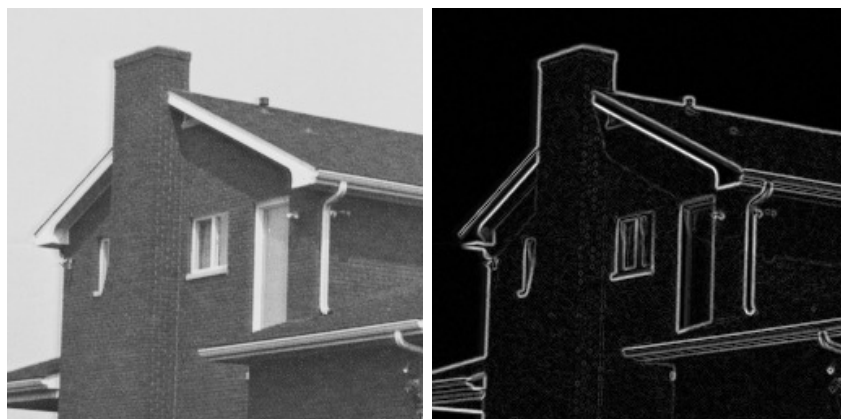


Abbildung 1: **Links:** Originalbild, 256×256 Pixel.. **Rechts:** Indem man den Betrag des Bildgradienten berechnet, kann man Kanten identifizieren.

Im Fall von Farbbildern repräsentiert man eine Farbe durch ihren Rot-, Grün- und Blauanteil (RGB-Modell) und betrachtet Abbildungen $f : \Omega \mapsto \mathbb{R}^3$. Farbbilder sind für uns also vektorwertige Funktionen. In unserem Artikel werden wir mathematische Modelle insbesondere für Grauwertbilder studieren, jedoch diverse Beispiele zeigen, bei denen Farbbilder zum Einsatz kommen. Im einfachsten Fall verwendet man dabei ein Konzept für Grauwertbilder für jeden der drei Farbkanäle getrennt, aber auch kompliziertere, gekoppelte Modelle können sinnvoll sein. Eine genauere Diskussion würde jedoch den Rahmen dieses einführenden Artikels sprengen.

Um Bilder einer Verarbeitung auf dem Rechner zugänglich zu machen, diskretisiert man den Definitionsbereich Ω , indem man das Bild auf einem

äquidistanten Gitter abtastet. Man gelangt so von einem kontinuierlichen zu einem diskreten Bild. Die einzelnen Bildpunkte werden als Pixel – eine Abkürzung für den englischen Begriff *picture elements* – bezeichnet. Ableitungen kann man nun durch ihre finiten Differenzenapproximationen ersetzen. Möchte man partielle Differentialgleichungen auf Bildern betrachten, so kann man sich auf natürliche Weise der Numerik der finiten Differenzenverfahren bedienen [7]. Häufig gelangt man durch solche Differenzenapproximationen zu großen linearen oder nichtlinearen Gleichungssystemen. Jede Unbekannte repräsentiert den Grauwert des gefilterten Bildes an einem einzigen Pixel. Man kann sich leicht vorstellen, welche numerischen Herausforderungen sich aus der hohen Auflösung moderner Digitalbilder ergeben. Da wir uns jedoch auf Modellierungsaspekte konzentrieren möchten, soll dies hier nicht weiter vertieft werden.

2 Elektrostatik

Wir beginnen unsere Exkursion in die mathematische Modellierung von Bildverarbeitungsprozessen mit einem Ansatz, der auch ohne Kenntnisse über Differentialgleichungen intuitiv erfasst werden kann: Er verwendet Ideen aus der Elektrostatik, die wir seit unserer Schulzeit kennen.

Das Bildverarbeitungsproblem, das wir betrachten möchten, wird als Halftoningproblem bezeichnet: Wir möchten ein Grauwertbild drucken, haben aber lediglich schwarzen Toner oder schwarze Tinte zur Verfügung. In der Druckbranche werden hierzu Verfahren eingesetzt, die Grauwerte durch geeignet verteilte schwarze Kreisscheiben approximieren. Das Problem der optimalen zweidimensionalen Approximation ist mathematisch sehr anspruchsvoll, und es gibt eine ganze Reihe von Algorithmen, die in Publikationen und Patenten beschrieben wurden und die diese Aufgabe mehr oder weniger gut lösen. Viele dieser Methoden sind recht heuristisch und auch teilweise relativ kompliziert. Ein neuartiges Verfahren, das auf einfachen elektrostatischen Prinzipien beruht und Ergebnisse liefert, die qualitativ zum Stand der Kunst gehören, wird in [10] beschrieben. Nehmen wir zunächst an, dass wir lediglich schwarze Kreisscheiben von einer einheitlichen, vorgegebenen Größe zur Verfügung haben. Jeder dieser “Partikel” soll die gleiche negative Ladung tragen und frei beweglich sein. Aufgrund ihrer Abstoßungskräfte werden diese Teilchen nun versuchen, sich möglichst gleichmäßig im Bildbereich zu verteilen, um ihren Abstand zu maximieren. Wenn wir nun eine zusätzliche positive Ladungsverteilung einführen, die ortsfest ist und proportional zu den dunklen Grauwerten im Bild wirkt, stellt sich ein neues Kräftegleichgewicht ein, bei dem sich die schwarzen Kreisscheiben in dunklen Bildbereichen stärker konzen-

trieren. Dabei modellieren wir die Kräfte zwischen zwei geladenen Teilchen als invers proportional zu ihrem Abstand und als proportional zum Produkt ihrer Ladungen. Ein entsprechendes Halftoningresultat sehen wir in Abbildung 2(b) und (c) für zwei verschiedene Teilchengrößen. Interessanterweise genügen schon wenige Kreisscheiben, um die Strukturen des Grauwertbildes gut zu approximieren.

Selbstverständlich kann man diese Modelle auf die verschiedensten Arten weiterentwickeln, beispielsweise durch unterschiedliche Teilchengrößen und Ladungen. Darüber hinaus ist es möglich, mehrere Teilchen durch eine starre Achse miteinander zu koppeln, um damit anisotrope Strukturen durch Geradenstücke zu approximieren [11]. Dadurch erhält man künstlerische Repräsentationen, die dem Charakter einer Bleistiftzeichnung besonders nahe kommen (Abbildung 2(d)).

Zusätzliche Möglichkeiten ergeben sich bei Farbbildern, wo wir drei verschiedenen Teilchenarten einführen können (rot, grün und blau) [11]. Einerseits sollten wir in einem solchen Fall ein Kräftegleichgewicht innerhalb der selben Teilchenart anstreben. Andererseits ist es sinnvoll, auch Interaktionen zwischen unterschiedlichen Teilchenarten zu modellieren, um insgesamt zu einer ausgewogenen Verteilung zu gelangen. Eine unerwartete Anwendung in der Computergrafik sehen wir in Abbildung 3. Hier verwenden wir anstelle von roten, grünen und blauen Kreisscheiben Pizzazutaten in Form von Schinkenstücken, Champignons und Salamisscheiben unterschiedlicher Größe. Geben wir nun verschiedene kontinuierliche Verteilungen vor, so erzeugt unser elektrostatisches Pizzahalftoning einen Belag nach Wunsch. Während ein anonymer Gutachter diese Resultate als “yummy” beschrieben hat, hat uns ein italienischer Student darauf aufmerksam gemacht, dass diese Pizzen noch zu “deutsch” wirken, da der Belag insgesamt zu wohlgeordnet erscheint. Wir sind zuversichtlich, dass sich mit einer stärkeren Randomisierung der Algorithmen der italienische Charakter unserer Pizzen deutlicher herausarbeiten lässt. Hat man das Grundprinzip des elektrostatischen Halftonings verstanden und kann es numerisch simulieren, liegt es auf der Hand, es auch auf andere Anwendungen in der Computergrafik zu übertragen, beispielsweise für die Computersimulation von Wäldern und Streuobstwiesen.

Aus mathematischer Sicht führt der Ansatz des elektrostatischen Halftonings auf interessante Optimierungsprobleme im Bereich der konvexen Analysis. Darüber hinaus ist zu beachten, dass naive numerische Implementationen der Interaktion von N Teilchen auf iterative Verfahren führen, die eine Komplexität von $\mathcal{O}(N^2)$ pro Iterationsschritt besitzen. Da N recht groß sein kann, ist dies schon relativ ungünstig. Mit mathematisch anspruchsvolleren Methoden, beispielsweise mit schnellen Summationstechniken aus dem Bereich der nichtäquidistanten Fouriertransformation, lässt sich die Komplexität auf

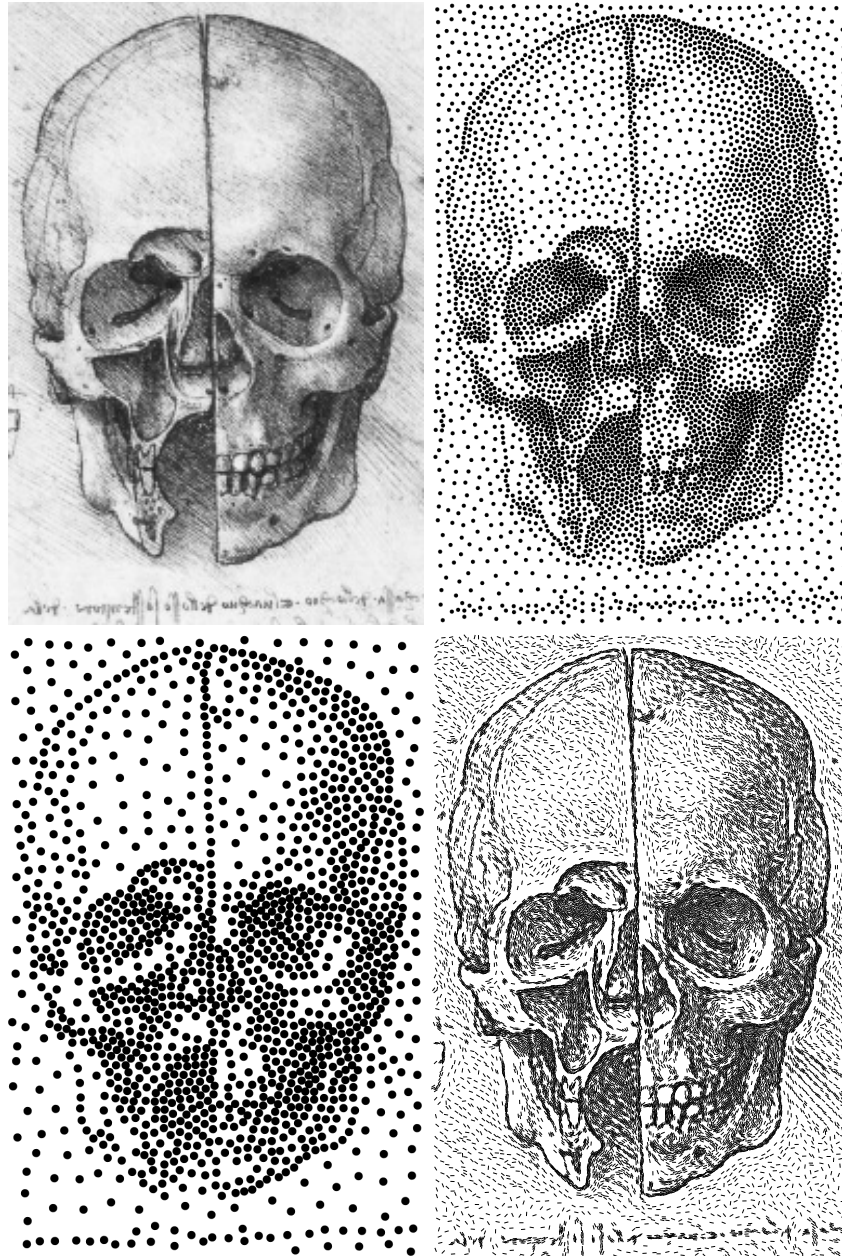


Abbildung 2: (a) **Oben links:** Originalzeichnung von Leonardo da Vinci. (b) **Oben rechts:** Elektrostatisches Halftoning mit zahlreichen kleinen Punkten [10]. (c) **Unten links:** Mit wenigen großen Punkten [10]. (d) **Unten rechts:** Elektrostatische Repräsentation mit Linien [11].

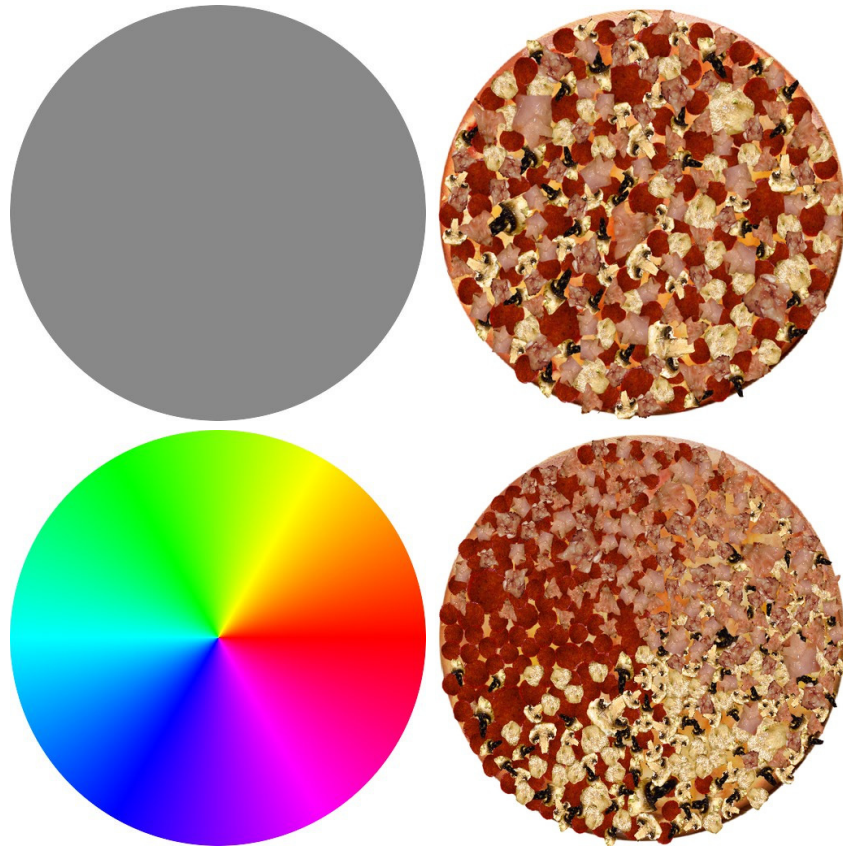


Abbildung 3: **Links:** Originaldaten. **Rechts:** Elektrostatische Repräsentation der RGB-Kanäle mit Hilfe von Schinken, Champignons und Salamischeiben in drei verschiedenen Größen [11]. Quelle der Zutaten: (User:fugutabetai_shyashin, flickr.com, CC-BY-NC-SA), (User:cyclonebill, flickr.com, CC-BY-SA), (User:vagabondvince310, flickr.com, CC-BY).

$\mathcal{O}(N \log N)$ reduzieren. Für weitere Details sei der Leser auf [13] verwiesen.

3 Diffusion

Als nächstes wenden wir uns einem grundlegenden Bildverarbeitungsproblem zu, das bereits seit langem intensiv erforscht wird: dem Glätten und Entrauschen von Bildern.

Es ist unmittelbar einsichtig, dass sich Rauschen typischerweise in Grauwertschwankungen zwischen Nachbarpixeln niederschlägt. Um solche Fluktuationen zu entfernen, betrachten wir einen wohlbekannten Glättungsprozess aus der Physik: Diffusion ist ein massenerhaltender Mechanismus, der Konzen-

trationsunterschiede ausgleicht. In einer zweidimensionalen Welt mit Konzentrationen $u(x, y, t)$ im Ort (x, y) zur Zeit t genügt er im einfachsten Fall der Differentialgleichung

$$\partial_t u = \Delta u \quad (1)$$

wobei $\Delta u = \partial_{xx} u + \partial_{yy} u$ den örtlichen Laplaceoperator darstellt. Diese Gleichung bezeichnet man als homogene (lineare) Diffusionsgleichung. Typischerweise kennt man die Konzentrationsverteilung zur Zeit $t = 0$ (Anfangsbedingung),

$$u(x, y, 0) = f(x, y), \quad (2)$$

und ist an der zeitlichen Evolution für $t > 0$ interessiert.

Um diesen Ausgleichsvorgang auf Grauwertbilder zu übertragen, müssen wir lediglich die Anfangskonzentrationsverteilung mit unserem Grauwertbild $f(x, y)$ identifizieren. Üblicherweise nimmt man darüber hinaus an, dass der Fluss über den Rand unseres Bildbereichs verschwindet (homogene Neumann-Randbedingungen). Die Lösung $u(x, y, t)$ stellt dann ein geglättetes Bild dar, wobei die Zeit t als Glättungsparameter fungiert: Größere Werte liefern glattere Lösungen.

Homogene Diffusion wird in der Bildverarbeitung schon seit über 50 Jahren verwendet: Bereits im Jahr 1959 hat Taizo Iijima diesen Prozess axiomatisch fundiert, indem er bewiesen hat, dass dies der einzige Glättungsvorgang ist, der fünf Axiomen genügt: Linearität, Translationsinvarianz, Skaleninvarianz, der Halbgruppeneigenschaft und der Erhaltung der Nichtnegativität [5]. Diese auf Japanisch publizierten Ergebnisse blieben in der westlichen Welt lange Zeit unbekannt. In Japan selbst wurden homogene Diffusionsfilter im Bereich der optischen Zeichenerkennung (OCR) mit großem Erfolg eingesetzt.

Abbildung 4 zeigt die Wirkung der homogenen Diffusionsfilterung anhand eines Beispiels aus der computergestützten Qualitätskontrolle von Spinnvlies. Hierbei möchte man die Wolkigkeit auf unterschiedlichen Skalen darstellen und analysieren. Die Resultate zu unterschiedlichen Diffusionszeiten liefern diese gewünschte Skalenraumrepräsentation. Indem man für die jeweiligen Bilder die Varianz berechnet, hat man ein Maß für die Wolkigkeit auf verschiedenen Skalen.

Während homogene Diffusion für die obige Anwendung ideal erscheint, ist sie bei anderen Problemen zu unflexibel: Deutliche Grauwertunterschiede zwischen Nachbarpixeln müssen nicht notwendigerweise auf Rauschen zurückgehen, sondern können beispielsweise auch durch Bildkanten verursacht werden. Solche Kanten möchte man häufig nicht wegglätten, da sie essentielle Bildinformation beinhalten. Es wäre daher schön, wenn man unterschiedliche Bildstrukturen unterschiedlich behandeln könnte. Dies ist im Rahmen der homogenen Diffusionsfilterung leider nicht möglich und erfordert

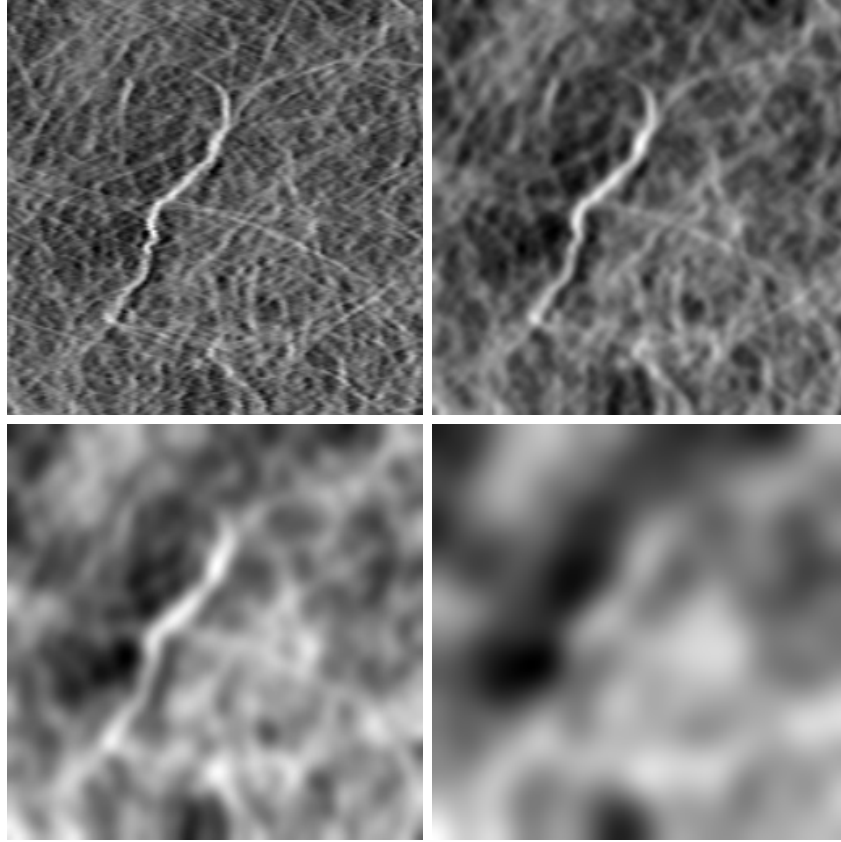


Abbildung 4: Homogene Diffusionfilterung zur Visualisierung der Wolkigkeit in Spinnvlies auf verschiedenen Skalen. **Von links oben nach rechts unten:** Wachsende Diffusionszeit.

ortsvariante Prozesse. Im Jahr 1990 haben Perona und Malik hierzu folgende nichtlineare Diffusionsgleichung vorgeschlagen [9]:

$$\partial_t u = \operatorname{div} \left(g(|\nabla u|^2) \nabla u \right), \quad (3)$$

wobei div den örtlichen Divergenzoperator beschreibt. Hierbei wird die Diffusion durch eine nichtlineare Diffusivität

$$g(|\nabla u|^2) = \frac{1}{1 + |\nabla u|^2 / \lambda^2} \quad (4)$$

gesteuert, die eine monoton fallende Funktion im Betrag des Bildgradienten ist. Wichtige Kanten, bei denen $|\nabla u| \gg \lambda$ gilt, möchte man dabei möglichst wenig glätten, während Rauschen, das sich in geringeren Fluktuationen niederschlägt ($|\nabla u| \ll \lambda$), wirksam entfernt werden soll.

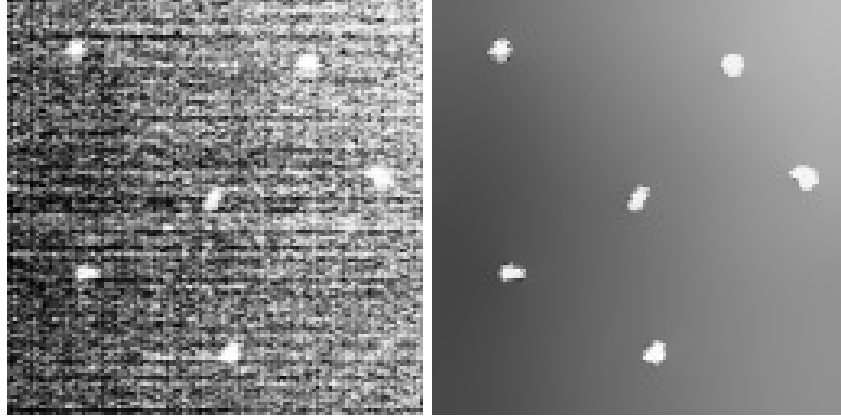


Abbildung 5: Nichtlineare Diffusion in der Mammographie. **Links:** Ausschnitt aus einem Mammogramm, 128×128 Pixel. Die hellen Strukturen sind Mikrokalzifikationen, die bei der Krebsfrüherkennung eine Rolle spielen. **Rechts:** Nach einer nichtlinearen Diffusionsfilterung werden Mikrokalzifikationen besser sichtbar.

Während das ursprüngliche Perona-Malik-Modell in der kontinuierlichen Formulierung mit einigen Gutgestelltheitsproblemen zu kämpfen hat, gibt es eine Reihe von Modifikationen, die sich als sehr stabil und äußerst nützlich in der Praxis erwiesen haben. So kann man beispielsweise den “Kantendetektor” $|\nabla u|$ durch eine robustere Variante $|\nabla u_\sigma|$ ersetzen, wobei u_σ durch eine Faltung von u mit einem Gaußkern G_σ mit Standardabweichung σ entstanden ist [3]. Abbildung 5 zeigt eine Anwendung, bei der durch eine solche nichtlineare Diffusionsfilterung winzige Verkalkungen in der Mammographie herausgearbeitet werden. Für weiterführende Diskussionen zu nichtlinearen Diffusionsfiltern wird auf die Monografie [14] verwiesen, die im Internet frei erhältlich ist.

4 Wellenausbreitung

Die bisher betrachteten Diffusionsgleichungen sind Prototypen für sogenannte parabolische Differentialgleichungen. Sie üben einen stark glättenden Effekt in positiver Zeitrichtung aus. Wir wollen uns nun einem anderen Typ partieller Differentialgleichungen widmen, der als hyperbolisch bezeichnet wird und dessen Verhalten sich durch die Ausbreitung von Wellen illustrieren lässt. Hier sind Vorwärts- und Rückwärtsrichtung in der Zeit grundsätzlich gleichberechtigt, und es kommt nicht notwendigerweise zu Glättungseffekten. Ganz im Gegenteil: Von hyperbolischen Differentialgleichungen weiß man,



Abbildung 6: Illustration der morphologischen Grundprozesse Dilatation und Erosion. **Linke Spalte:** Dilatation eines Binärbilds von Island mit wachsenden Evolutionszeiten. **Rechte Spalte:** Entsprechende Erosion.

dass sie selbst bei glatten Anfangsdaten nicht differenzierbare Lösungen erzeugen können. Solche “Schocks” können in der Bildverarbeitung durchaus wünschenswert sein, wenn sie als Berandungen von Bildobjekten dienen.

Als Prototypen für hyperbolische Differentialgleichungen betrachten wir folgende zwei Evolutionsgleichungen, die mit dem Originalbild f als Anfangswert zur Zeit $t = 0$ starten:

$$\partial_t u = +|\nabla u|, \quad (5)$$

$$\partial_t u = -|\nabla u|. \quad (6)$$

Die erste Gleichung beschreibt eine sogenannte Dilatation. Ihre Lösung $u(x, y, t)$ zur Zeit $t > 0$ ersetzt den Grauwert $f(x, y)$ durch das Maximum innerhalb einer Kreisscheibe vom Radius t . In analoger Weise ersetzt die zweite Gleichung, die eine Erosion simuliert, den Grauwert durch das Mi-

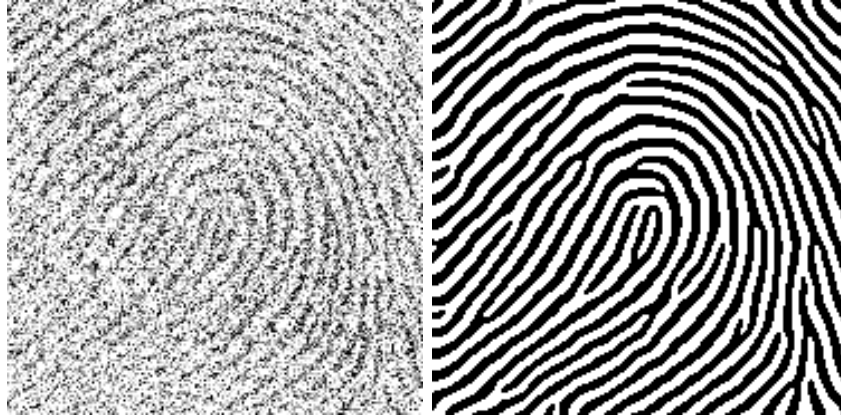


Abbildung 7: **Links:** Fingerabdruck (186×186 Pixel) mit additivem Gauß-Rauschen. Die Rauschvarianz beträgt das Vierfache der Bildvarianz. **Rechts:** Nach kohärenzverstärkender Schockfilterung.



Abbildung 8: Wendet man kohärenzverstärkende Schockfilter auf Bilder von Personen an, entstehen karikaturartige Effekte. **Links:** Originalbild von Professor Albrecht (Universität des Saarlandes), 480×640 Pixel. **Rechts:** Nach kohärenzverstärkender Schockfilterung.

nimum innerhalb einer Kreisscheibe vom Radius t . Beide Prozesse bilden die Grundbausteine der sogenannten mathematischen Morphologie. Dieses Gebiet wurde um 1965 von Serra und Matheron an der ENS des Mines in Fontainebleau begründet und zunächst mengentheoretisch formalisiert [12]. Seit rund zwanzig Jahren verwendet man auch entsprechende PDE-basierte Formulierungen; siehe z.B. [1].

Wie der Leser wohl bereits vermutet hat, wurde der Begriff der Erosion durch reale Erosionsvorgänge in der Natur motiviert. Dilatation beschreibt einen Ausbreitungsprozess, wie man ihn in der Physik vom Huygens-Prinzip her kennt: Eine Wellenfront breitet sich in äußerer Normalenrichtung aus. Dilatation und Erosion agieren komplementär: Einer Dilatation des Vordergrunds entspricht eine Erosion des Hintergrunds und umgekehrt. Abbildung 6 illustriert die Wirkung von Dilatation und Erosion auf ein Binärbild von Island. Während Dilatation und Erosion für sich genommen nur von begrenztem Nutzen sind, entstehen durch ihre geschickte Kombination mächtige Bildverarbeitungswerkzeuge. Beispielsweise verwendet man bei den sogenannten Schockfiltern die Dilatationsgleichung im Einzugsbereich der Maxima, während die Erosionsgleichung im Einzugsbereich der Minima agiert [8]. Somit breiten sich sowohl Maxima als auch Minima aus, und an der Grenze beider Einzugsbereiche entsteht eine Diskontinuität (Schock). Die einzelnen Schockfilter unterscheiden sich insbesondere durch verschiedene Definitionen der Einzugsbereiche. Beispielsweise kann man diese Einzugsbereiche so wählen, dass kohärente, flussartige Strukturen herausgearbeitet werden [15]. Derartige kohärenzverstärkende Schockfilter erlauben es unter anderem, Fingerabdruckbilder in hoher Qualität aufzubereiten (Abbildung 7). Wie so oft im Leben und in der Bildverarbeitung gehen die erwünschten Wirkungen auch mit Nebenwirkungen einher, die manchmal sogar interessanter sind: Wendet man kohärenzverstärkende Schockfilter auf Porträtaufnahmen von Personen an, so entstehen Karikaturen, bei denen Frisur, Bart und Brille mit besonderer Hingabe verstärkt werden (Abbildung 8).

5 Wärmeleitung

Neben parabolischen und hyperbolischen Differentialgleichungen sind nicht zuletzt auch elliptische Differentialgleichungen von großer Bedeutung. Sie entstehen beispielsweise als stationäre Zustände ($t \rightarrow \infty$) parabolischer Prozesse.

Wir möchten in diesem Zusammenhang insbesondere stationäre Wärmeleitungsprozesse betrachten. Bekanntermaßen folgen Wärmeleitungsvorgänge denselben Differentialgleichungen wie Diffusionsprozesse, so dass wir auf un-



Abbildung 9: Interpolation mithilfe von Wärmeleitungsmodellen. **Links:** Farbbild (438×297 Pixel) mit überlagertem Text. **Rechts:** Eine anisotrope Wärmeleitungsgleichung rekonstruiert die fehlende Information.

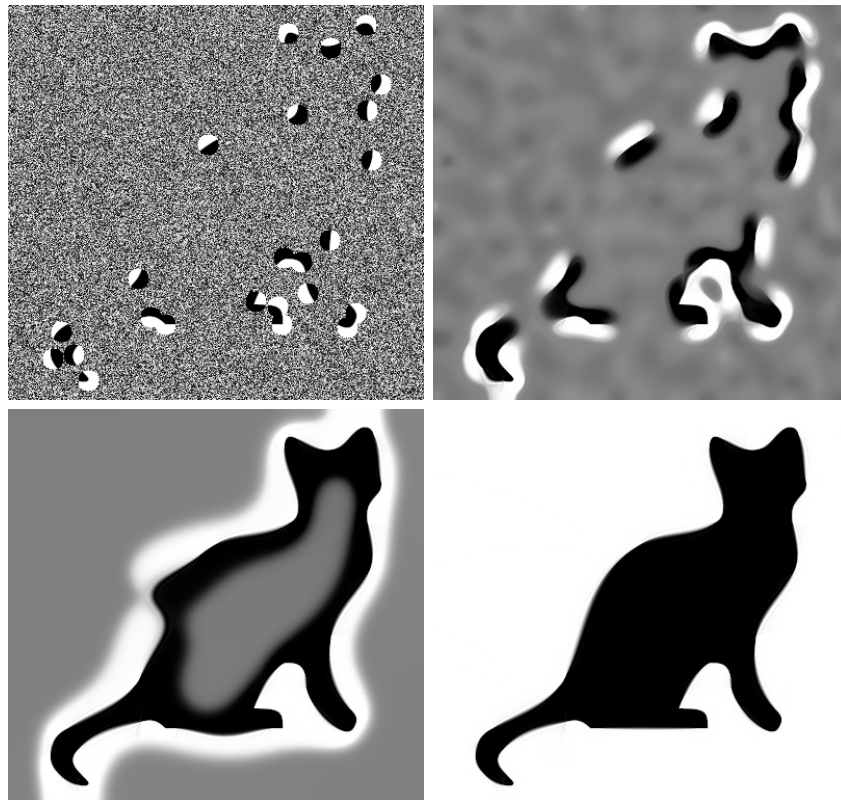


Abbildung 10: **Links oben:** Die Form eines zu rekonstruierenden binären Objekts wurde auf 24 kleinen Kreisscheiben vorgegeben. Diese dienen als Daten für die Interpolation. An allen anderen Stellen wurde mit gleichverteiltem Rauschen initialisiert. **Rechts oben bis rechts unten:** Zeitliche Evolution eines Interpolationsprozesses, der auf einer anisotropen Wärmeausbreitung beruht.

seren Kenntnissen aus Abschnitt 4 aufbauen können. Der Unterschied liegt lediglich in der physikalischen Interpretation, die im Fall des stationären Zustands für die Wärmeleitung etwas intuitiver erscheint als für Diffusion.

Wir möchten Ideen aus dem Bereich der Wärmeleitung verwenden, um fehlende Strukturen in Bildern zu rekonstruieren. Dies ist eine typische Interpolationsaufgabe. In der Bildverarbeitung spricht man auch von Inpainting. Um Inpaintingprobleme mit Wärmeleitungsmodellen zu lösen, stellen wir uns vor, dass wir alle bekannten Pixel als kleine Heizkörper auffassen, die mit einem Thermostat zu jeder Zeit auf einer festen Temperatur gehalten werden, die ihrem Grauwert entspricht. In den unbekannten Pixeln lösen wir nun eine Wärmeleitungsgleichung für $t \rightarrow \infty$. Im einfachsten Fall einer homogenen Wärmeleitung bedeutet dies, dass wir die Laplacegleichung

$$\Delta u = 0 \tag{7}$$

verwenden. Die stationäre Temperaturverteilung interpretieren wir als rekonstruiertes Bild.

Neben der homogenen Wärmeleitung haben sich anisotrope Wärmeleitungsprozesse als besonders nützlich erwiesen, insbesondere wenn fehlende Kanten in hoher Qualität rekonstruiert werden sollen. Hierzu betrachten wir Gleichungen der Form

$$\operatorname{div}(\mathbf{D}(\nabla u_\sigma) \nabla u) = 0. \tag{8}$$

Der Wärmeleitungstensor $\mathbf{D}$ ist eine positiv definite 2×2 -Matrix mit Eigenvektoren ∇u_σ und $\nabla u_\sigma^\perp$, wobei u_σ wiederum eine mit einem Gaußkern geglättete Version von u beschreibt. Die Richtung der Eigenvektoren verläuft somit senkrecht bzw. parallel zu Bildkanten. Die entsprechenden Eigenwerte seien durch

$$g(|\nabla u_\sigma|^2) = \frac{1}{\sqrt{1 + |\nabla u_\sigma|^2/\lambda^2}} \tag{9}$$

bzw. 1 gegeben. Wir möchten also einen ausgeprägten Wärmeausgleich längs der Bildkanten, während wir den Ausgleich über Kanten hinweg deutlich einschränken (je steiler die Kante, desto stärker). Derartige Modelle wurden zuerst im Zusammenhang mit Diffusionsfiltern zum Entrauschen von Bildern vorgeschlagen [14], zeigen jedoch für Interpolationsaufgaben ihre Qualitäten noch deutlicher. Abbildung 9 präsentiert eine Inpaintinganwendung unseres anisotropen Wärmeleitungsmodells. Wir erkennen eine realistische Dateninterpolation.

Wer in dieser Richtung einmal Blut geleckt hat, fragt sich schnell, wie weit man dabei gehen kann. Abbildung 10 zeigt ein Experiment, bei dem nur ein kleiner Teil der gesamten Bilddaten vorgegeben ist. Nichtsdestotrotz gelingt

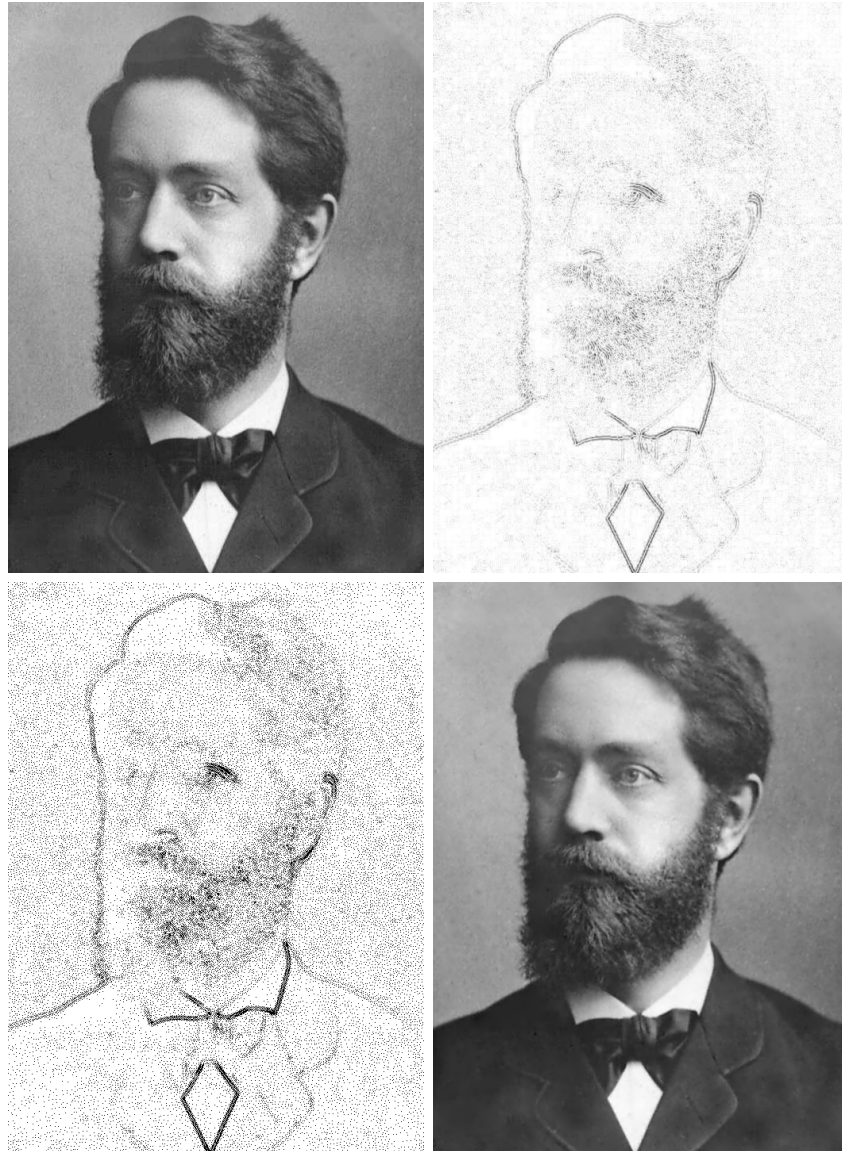


Abbildung 11: **Oben links:** Originalbild von Felix Klein, 512×720 Pixel. **Oben rechts:** Betrag des Laplaceoperators. Dunkle Farben repräsentieren hier hohe Werte. **Unten links:** Ausgewählte Interpolationspunkte bei einer gewünschten Dichte von 10 %. Hierzu wurde elektrostatisches Halftoning verwendet. **Unten rechts:** Nach Interpolation mit der Laplacegleichung.



Abbildung 12: Vergleich verschiedener Kompressionsverfahren bei einer Kompressionsrate (Reduktion der Dateigröße) von 104:1. **Oben links:** Originalbild, 380×431 Pixel. **Oben rechts:** Komprimiert mit JPEG. Man erkennt deutliche Blockartefakte. **Unten links:** Mit dem moderneren, waveletbasierten JPEG2000-Verfahren ist die Qualität wesentlich besser. **Unten rechts:** Eine Kompression mit der homogenen Wärmeleitungsgleichung liefert die besten Ergebnisse [6].

es der anisotropen Wärmeleitung, im stationären Fall die Kontur der Katze sehr gut zu rekonstruieren.

Ein nächster Schritt wäre es nun, diese Konzepte für Kompressionsanwendungen nutzbar zu machen. Die Grundidee erscheint simpel: Man speichert die Bilddaten nur an wenigen Pixeln und rekonstruiert die fehlende Information durch wärmeleitungsbasiertes Inpainting. Erwartungsgemäß hängt die Qualität der Rekonstruktion von der Wahl der Pixel ab. Während die anisotrope Wärmeleitungsgleichung aufgrund ihrer ausgefeilten Modellannahmen hier recht robust ist, profitiert insbesondere die einfachere homogene Wärmeleitung sehr von einer guten Pixelwahl.

Was sind also gute Interpolationsdaten? Im Fall der homogenen Wärmeleitung lässt sich diese Frage sogar analytisch beantworten. In [2] wurde mit Methoden aus dem Bereich der Shapeoptimierung gezeigt, dass man die Dichte der zu speichernden Pixel so wählen sollte, dass sie eine monoton wachsende Funktion von $|\Delta f|$ ist. Wie gut dies in der Praxis funktioniert, illustriert Abbildung 11. Um von der kontinuierlichen Pixeldichte $|\Delta f|$ zu einer konkreten Pixelauswahl zu gelangen, wurde das bereits vorgestellte elektrostatische Haftoning verwendet. Wir erkennen, dass sich das Bild von Felix Klein schon aus lediglich 10 % aller Pixel mit hoher Genauigkeit rekonstruieren lässt.

Aktuelle Kompressionsstandards wie JPEG und JPEG 2000 sind völlig andersartig aufgebaut und beruhen auf orthogonalen Transformationen (diskrete Kosinustransformation bei JPEG, diskrete Wavelettransformation bei JPEG 2000). In der Transformationsdomäne werden visuell wenig relevante Koeffizienten ungenauer gespeichert oder sogar komplett weggelassen. Diese Kompressionsansätze sind hochgradig optimiert. Um mit ihnen mithalten zu können, muss man ebenfalls zahlreiche Optimierungsschritte durchführen, bei denen insbesondere auch den Ort der zu speichernden Pixel effizient kodiert wird [6].

Glücklicherweise scheint sich dieser Aufwand auszuzahlen, wie man in Abbildung 12 erkennt. Bei nicht allzu stark texturierten Bildern und hohen Kompressionsraten ist unser Wärmeleitungsansatz sowohl dem JPEG-Standard als auch seinem Nachfolger JPEG 2000 qualitativ überlegen. Zur Zeit forscht unsere Arbeitsgruppe sehr intensiv an weiteren Verbesserungen dieser PDE-basierten Methoden zur Datenkompression.

6 Osmose

Nachdem wir Diffusionsgleichungen erfolgreich für eine Reihe von Bildverarbeitungsproblemen eingesetzt haben, wollen wir nun untersuchen, ob sich

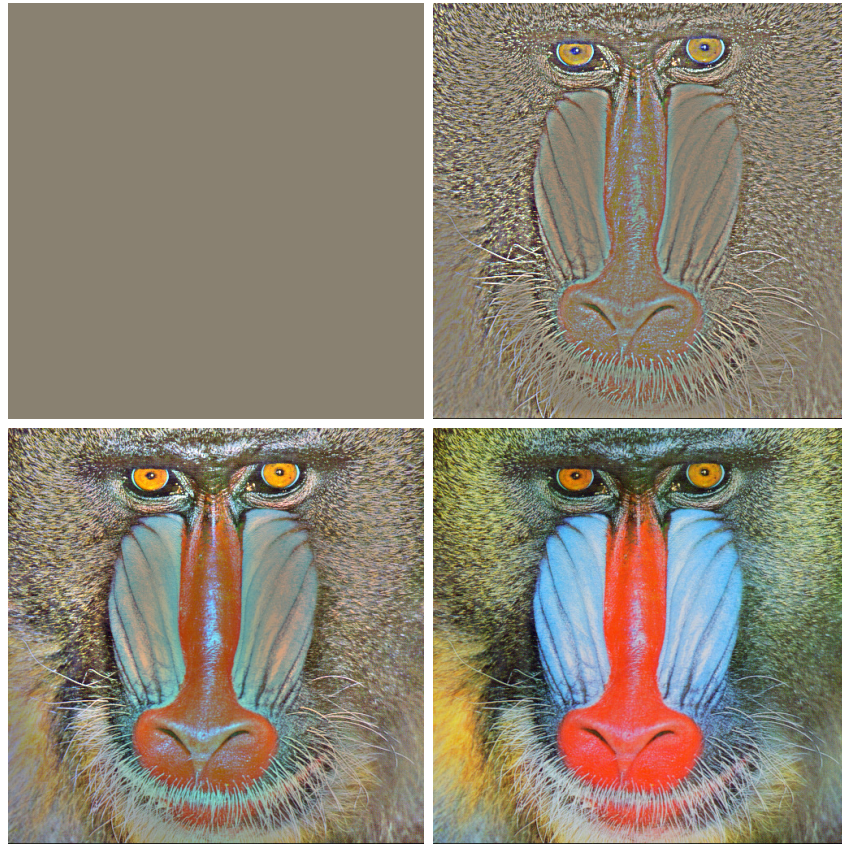


Abbildung 13: Osmose erlaubt die Spezifikation eines beliebigen stationären Zustands. **Von links oben bis rechts unten:** Zeitliche Evolution eines Osmoseprozesses.

durch eine Modifikation dieser Prozesse weitere Anwendungsfelder erschließen lassen [4]. Diskretisiert man Diffusionsgleichungen mithilfe expliziter Finite-Differenzen-Verfahren, so gelangt man zu symmetrischen Markowketten. Symmetrie bedeutet, dass die Übergangswahrscheinlichkeit von Grauwerten zwischen einem Pixel A und seinem Nachbarpixel B identisch ist zur Übergangswahrscheinlichkeit von B zu A. Diese Symmetrieeigenschaft wird benötigt, um ein Maximum-Minimum-Prinzip sowie Konvergenz zu einem konstanten stationären Zustand zu beweisen (falls man nicht wie im vorherigen Abschnitt zusätzliche Daten spezifiziert) [14].

Es wäre spannend zu untersuchen, welche zusätzlichen Modellierungsmöglichkeiten sich ergeben, wenn man auf die Symmetrieeigenschaft verzichtet. Stellt man sich einen Grauwertübergang zwischen benachbarten Pixeln als einen Teilchentransport durch eine virtuelle Membran vor, die sich zwischen den Pixeln befindet, so erinnert eine unsymmetrische



Abbildung 14: Osmose zum Editieren von Bildern. **Oben links:** Gemälde von Euler. **Oben rechts:** Zeichnung von Lagrange. **Unten links:** Direktes Einsetzen des Kopfes von Lagrange in das Gemälde von Euler ist sofort als plumpe Bildmanipulation erkennbar. **Unten rechts:** Fusionierung über die Driftvektoren eines Osmoseprozesses erlaubt einen nahtlosen Übergang [4].

Durchlässigkeit in beiden Richtungen an Osmoseprozesse. Bei diesen können sich bekanntermaßen unterschiedliche stationäre Konzentrationen auf beiden Seiten einstellen. Verwandte osmoseartige Modelle für die Bildverarbeitung lassen sich durch Drift-Diffusions-Prozesse beschreiben. Im einfachsten (linearen) Fall haben diese folgende Struktur:

$$\partial_t u = \Delta u - \operatorname{div}(\mathbf{d}u). \quad (10)$$

Der Driftvektor $\mathbf{d}$ generiert die erwünschte Unsymmetrie und ist für nicht konstante stationäre Zustände verantwortlich. Wählt man beispielsweise $\mathbf{d}(x, y) = \nabla(\ln v(x, y))$ mit einem nichtnegativen Bild v , so konvergiert der Prozess für $t \rightarrow \infty$ gegen ein Vielfaches von v . Dieser Zusammenhang erlaubt es uns also, durch Vorgabe von Ableitungsdaten eines Zielbildes $\ln v$ einen Evolutionsprozess zu konstruieren, dessen stationärer Zustand w die Gleichung $\nabla(\ln w) = \nabla(\ln v)$ erfüllt. In diesem Sinne gibt es zu jedem differenzierbaren positiven Bild $v(x, y)$ einen induzierten Driftvektor $\mathbf{d}(x, y) = \nabla(\ln v(x, y))$. Abbildung 13 illustriert die Konvergenz eines Osmoseprozesses gegen ein vorgegebenes Bild.

Selbstverständlich erscheint es zunächst wenig sinnvoll, Prozesse zu modellieren, die im Sinne dieser logarithmischen Ableitungsdaten gegen einen bereits bekannten stationären Zustand konvergieren. Spannender wird die Sache jedoch, wenn wir “Ableitungsdaten” vorschreiben, die überhaupt nicht exakt integrierbar sind. Dann erzeugt die Osmose stationäre Zustände, die versuchen, alle ihnen auferlegten widersprüchlichen Bedingungen möglichst gut zu approximieren. Dies kann man beispielsweise dazu verwenden, Bilder nahtlos zu fusionieren, indem man in einem Bereich die Ableitungsinformationen (induzierte Driftvektoren) des einen Bildes und im anderen Bereich Ableitungsinformationen des anderen Bildes vorgibt und einen Osmoseprozess startet. Das Anfangsbild ist hierbei kaum relevant: Lediglich sein mittlerer Grauwert wird erhalten. Ein entsprechendes Beispiel findet sich in Abbildung 14, wo es gelungen ist, mithilfe eines Osmoseprozesses den Kopf von Lagrange nahtlos in ein Gemälde von Euler zu integrieren [4].

7 Zusammenfassung

Mathematische Bildverarbeitung ist ein vergleichsweise junges Gebiet der angewandten Mathematik, dessen Faszination auch auf der Tatsache beruht, dass hier die unterschiedlichsten Ideen aus Natur und Technik Einzug gehalten haben. Es ist eine spannende Frage, warum es gerade diese Ideen sind, die sich als besonders fruchtbar erwiesen haben. Viele Prozesse aus der Natur folgen gewissen Optimalitätsprinzipien und weisen zahlreiche Invarianzen

auf. Beides ist auch für die Bildverarbeitung von enormer Bedeutung. Zudem ergibt sich hieraus auch die Möglichkeit, von dem reichhaltigen Wissen zu profitieren, das man in anderen Bereichen über diese Prozesse gewonnen hat.

Mathematische Bildverarbeiter genießen sämtliche Freiheiten, Modelle nach den jeweiligen Zielsetzungen zu modifizieren, und sie können die Auswirkungen ihrer mathematischen Modellierungsbestrebungen sofort visuell studieren. Auch wenn wir uns in diesem Beitrag hauptsächlich auf Prozesse konzentriert haben, die durch Differentialgleichungen beschrieben werden können, soll nicht unerwähnt bleiben, dass nahezu alle Gebiete der Mathematik ihren Stempel hinterlassen haben. Dadurch ergeben sich täglich neue und unerwartete mathematische Herausforderungen, und jeder Mathematiker ist herzlich eingeladen, zu ihrer Lösung beizutragen.

Danksagung

Es ist mir eine Freude, mich bei allen Kollegen, Mitarbeitern und Studierenden zu bedanken, die durch ihr Engagement und ihre Ideen zum Erfolg unserer gemeinsamen Forschungsanstrengungen beigetragen haben. Nicht zuletzt gebührt mein besonderer Dank der Deutschen Forschungsgemeinschaft. Sie hat durch die Förderung zahlreicher Einzelprojekte sowie durch den Gottfried-Wilhelm-Leibniz-Preis die finanziellen Voraussetzungen geschaffen, ohne die die jeweiligen Forschungsvorhaben nicht realisierbar gewesen wären.

Literatur

- [1] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel. Axioms and fundamental equations in image processing. *Archive for Rational Mechanics and Analysis*, 123:199–257, 1993.
- [2] Z. Belhachmi, D. Bucur, B. Burgeth, and J. Weickert. How to choose interpolation data in images. *SIAM Journal on Applied Mathematics*, 70(1):333–352, 2009.
- [3] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 32:1895–1909, 1992.
- [4] K. Hagenburg, M. Breuß, J. Weickert, and O. Vogel. Novel schemes for hyperbolic pdes using osmosis filters from visual computing. In A. M. Bruckstein, B. ter Haar Romeny, A. M. Bronstein, and M. M. Bronstein,

- editors, *Scale Space and Variational Methods in Computer Vision*, volume 6667 of *Lecture Notes in Computer Science*, pages 532–543. Springer, Berlin, 2012.
- [5] T. Iijima. Basic theory of pattern observation. In *Papers of Technical Group on Automata and Automatic Control*. IECE, Japan, December 1959. In Japanese.
 - [6] M. Mainberger, A. Bruhn, J. Weickert, and S. Forchhammer. Edge-based compression of cartoon-like images with homogeneous diffusion. *Pattern Recognition*, 44(9):1859–1873, 2011.
 - [7] K. W. Morton and L. M. Mayers. *Numerical Solution of Partial Differential Equations*. Cambridge University Press, Cambridge, UK, second edition, 2005.
 - [8] S. Osher and L. I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 27:919–940, 1990.
 - [9] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.
 - [10] C. Schmaltz, P. Gwosdek, and J. Weickert. Electrostatic halftoning. *Computer Graphics Forum*, 29(8):2313–2327, December 2010.
 - [11] C. Schmaltz, P. Gwosdek, and J. Weickert. Multi-class electrostatic halftoning. Technical Report 301, Department of Mathematics, Saarland University, Saarbrücken, Germany, October 2011.
 - [12] P. Soille. *Morphologische Bildverarbeitung*. Springer, Berlin, 1998.
 - [13] T. Teuber, G. Steidl, P. Gwosdek, C. Schmaltz, and J. Weickert. Dithering by differences of convex functions. *SIAM Journal on Imaging Sciences*, 4(1):79–108, January 2011.
 - [14] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart, 1998. (<http://www.mia.uni-saarland.de/weickert/book.html>).
 - [15] J. Weickert. Coherence-enhancing shock filters. In B. Michaelis and G. Krell, editors, *Pattern Recognition*, volume 2781 of *Lecture Notes in Computer Science*, pages 1–8. Springer, Berlin, 2003.