20 Jahre medizinische Bildverarbeitung

Ränder, Regionen, Intelligenz und Wahrnehmung

Hans-Peter Meinzer

Deutsches Krebsforschungszentrum Heidelberg
Abteilung Medizinische und Biologische Informatik
Im Neuenheimer Feld 280, 69120 Heidelberg

Email: H.P.Meinzer@DKFZ-Heidelberg.de

Zusammenfassung. Diese Arbeit beschreibt die Erfahrungen vieler Jahre medizinischer Bildverarbeitung. Sie ist stark autobiographisch. Sie zeigt verschiedene Versuche medizinische Bilder zu verstehen, auszuwerten und die gewonnene Information dem Mediziner zur Verfügung zu stellen. Das Ziel ist die Verbesserung der Diagnose und der Therapieplanung. Trotz aller Bemühungen bleiben bis heute ungelöste Probleme bestehen.

Schlüsselwörter: Medizinische Bildverarbeitung

1 Einleitung

1979 besuchte der Gastwissenschaftler Gershom Zaijcek von der Hadassah Universität in Jerusalem unser Institut für Dokumentation, Information und Statistik, das von Gustav Wagner geleitet wurde. Wir vereinbarten, die Mitosen (sich teilende Zellen) in der Darmkrypte mit Hilfe digitaler Bildverarbeitung statistisch auszuwerten.

Zaijcek, als Pathologe (und Computerexperte), sollte die Dünnschnitte besorgen. Mir fiel die Hardwarebeschaffung und die Entwicklung der Bildverarbeitungssoftware und Statistik zu. Dieses Projekt ist an unüberwindlichen Schwierigkeiten gescheitert, die ich heute besser verstehe als damals. Auf der anderen Seite hat kein Projekt vorher oder nachher meine Arbeit so geprägt und fokussiert wie dieses. Seit dieser Zeit arbeiten wir an medizinischer Bildverarbeitung.

Zusammen mit Bengt Sandblad von der Universität Uppsala haben wir in der Folge durch Simulation der Zellkinetik entscheidende Einblicke in die Proliferation epithelialer Gewebe gewonnen. Berichten will ich hier aber über die zweite Entwicklung: die digitale Bildverarbeitung, zunächst an biologischen Daten und nachher in größerem Umfang an medizinischen CT und MR Daten.

2 Die Anfänge

Anfang der 80er Jahre gab es wenige Lehrbücher für die Bearbeitung von Rasterbildern [1-3], wie man das damals nannte, und untaugliche Geräte für Eingabe und Ausgabe. Genaugenommen waren auch die damaligen Rechner völlig überfordert.

Die Kamera konnte zeilenweise ausgelesen werden. So war das Bild sein eigener primärer Speicher und nahm keinen Platz im Hauptspeicher weg. Typischerweise las man nur immer die drei Zeilen für die nächste Faltung ein. Die 32-bit Adressierung beschränkte den Adressraum auf 4 Mbytes, davon standen real maximal 64 Kbytes Hauptspeicher zur Verfügung.

Die Ausgabegeräte waren eigentlich gar keine. In der Not, die Ergebnisse doch irgendwie ansehen zu müssen, wurden mehrere merkwürdige Krücken erfunden. So konnte man die Bilder auf Flutelektronenbildschirmen der Firma Tektronix ansehen. Eine ähnliche Technik wurde beim Übereinanderdrucken von mehreren Buchstaben auf dem Kettendrucker verwendet. Jedes Pixel wurde, je nach Grauwert, in einen Buchstabenhaufen abgebildet, die Löcher dazwischen haben arg gestört. Die beste Methode war eine Art Höhenlinienplot auf dem Plotter.

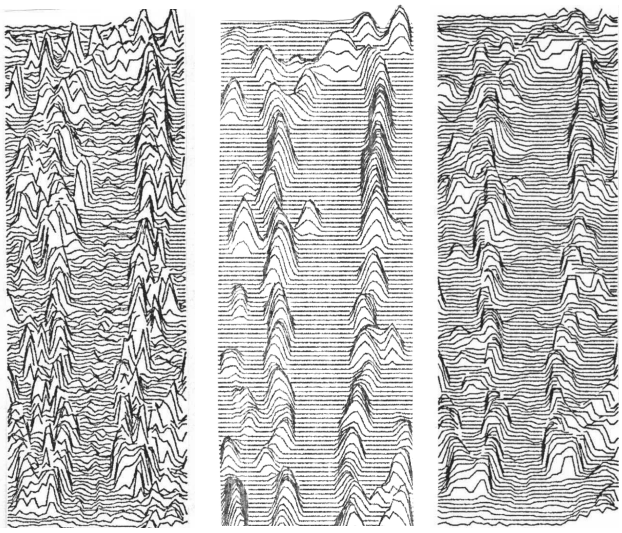

Abb. 1: Höhenlinienplot einer Darmkrypte, daneben nach Gaussglättung und nach einer Medianfilterung

Die Software war völliges Neuland, praktisch haben wir alles selbst erfunden. Um uns das Leben leichter zu machen und weil uns langsam die Schwierigkeit des Unterfangens geschwant hat, haben Uwe Engelmann und ich 1981 einen Bildverarbeitungsinterpreter gebaut, der immerhin schon über 100 Befehle verarbeitet hat: den Interpreter PIC.

Ich habe das Handbuch noch einmal zur Hand genommen und mich köstlich amüsiert. Ein paar der Abbildungen stammen daraus. Immerhin gab es schon ein Konzept zur Integration eigener Erweiterungen, man konnte eigene Programme über einen CALL Befehl aufrufen. Auch eine Beleuchtungskomponente gab es schon, allerdings zur Beleuchtung von Grauwertoberflächen, die Grauwerte waren als Höhen auf das 2D Bild aufgetragen.

Im Nachhinein kann man feststellen, daß unser Unwissen ebenso groß wie unsere Begeisterung war. Kein Chef hat uns gebremst und nur wenige haben verstanden, was wir da überhaupt machen. Der Interpreter PIC erwies sich bald als zu starr. So wurde in den Jahren 1985 bis 1988 auf der Basis des Interpreters APL ein viel größeres Paket entwickelt, das hatte dann schon 936 Funktionen. Da die Übersichtlichkeit gleichzeitig dahin ging, haben wir alles in 33 Hauptgruppen eingeordnet.

Zu den Operatorgruppen zählten solche zur Manipulation von Histogrammen, Pyramiden, Regions-of-Interest, Fourieroperatoren, Faltungen, Verschiebungsvektorfelder, Merkmalsextraktion, Morphologie sowie ein Visualisierungspaket für das Raytracing, das eigentlich schon ein Volumenvisualisierer war. Das gesamte japanische Softwarepaket SPIDER mit weiteren 350 Routinen war aufrufbar. Es gab einfach nichts mehr, was wir nicht eingebaut haben. Als wir bald gemerkt haben, daß die Rechnerei uns nicht weiterbringt, haben wir versucht, dem Gesamtwerk auch ein wenig Vernunft beizubringen.

Die AI Sprache PROLOG wurde komplett nachimplementiert, um eine bekannte AI Sprache direkt mit Bildverarbeitungsroutinen verknüpfen zu können. Bis dahin standen AI und BV in verschiedenen, nicht verknüpfbaren Sprachen quasi alleine da. Das gilt noch heute, allerdings erwartet man sich von den AI Sprachen heute weniger. Der beschriebene große Aufwand war weniger erfolgreich als wir uns das gewünscht und auch erwartet hatten.

Ich hatte damals die Idee, ein Buch über Bildverarbeitung zu schreiben. Bei der genaueren Planung wurde mir klar, wie unlogisch alles aufgebaut sein müßte. Das Buch wurde nie geschrieben. Mir tun die Studenten leid, die vor unbegründeten Haufen von Algorithmen sitzen und auch beim besten Willen nicht verstehen können, welche Methoden für welche Problemstellung womöglich geeignet sind. Vor allem können sie nicht wissen, was genaugenommen nur akademischen Reiz besitzt und keine praktische Relevanz.

3 Die 3D Visualisierung

1983 sah ich von dem Pionier Gabor Herman aus Philadelphia spektakuläre 3D Rekonstruktionen einer Wirbelsäule. Die Attraktion und Relevanz solcher Einsichten in den menschlichen Körper hat mich sehr beeindruckt und ich habe versucht, auch in diese Richtung tätig zu werden. Als erstes brauchten wir dazu eine 3D Visualisierung. Aus Unkenntnis der Verfahren, die auf Oberflächen beruhen, haben wir uns der Volumenvisualisierung zugewandt. Es war der Versuch, als Physiker, ein physikalisch begründetes Modell der Wechselwirkung von Licht und Materie zu bauen.

In der Tat ist der erste einer Reihe von 3D Visualisierern etwa 1986 entstanden. Damals haben wir diese fälschlicherweise Raytracer genannt, erst danach wurde dieses Wort semantisch eingegrenzt. Was wir hatten, heißt seit langem genauer Volumenvisualisierung. Die anderen Verfahren rechnen aus den Volumendaten auf diese oder jene Weise erst eine Oberfläche heraus und lassen dann diese geeignet reflektieren. Diese Verfahren waren im Prinzip von Lukasfilm und Spielberg entwickelt worden. Die medizinischen Anwendungen hatten typischerweise Probleme mit Objekten, die sich innerhalb von anderen Objekten befanden, ein Schatten fehlte. Diese Verfahren können halb transparente Objekte wie Wolken oder Texturen in Volumina gar nicht darstellen. Die Bestimmung einer Oberfläche innerhalb medizinischer Daten ist und bleibt sehr kritisch. Unser Verfahren, der Heidelberger Raytracer, konnte das alles, war aber sehr langsam. Wir brauchten auf dem damaligen Großrechner pro Bild etwa 1 Stunde Rechenzeit.

Immer wieder werden Anstrengungen unternommen, die Rechenzeit deutlich zu beschleunigen, Ziel ist der Bereich von 25 frames pro Sekunde. Dafür wurde und wird gerne spezielle Hardware gebaut. Nach unserer Erfahrung hat sich das nie gelohnt, weil man in der Regel mit der jeweils nächsten Hardwaregeneration die speziell dafür notwendige Software wegwerfen muß. Dafür werden in einem scheinbar unaufhaltsamen Prozeß die normalen Rechner immer schneller, mit dem Vorteil, alte Softwaremodule einfach viel schneller abzuarbeiten. Nach unserer Erfahrung ist es daher besser, in ordentlichen standardisierten Code zu investieren.

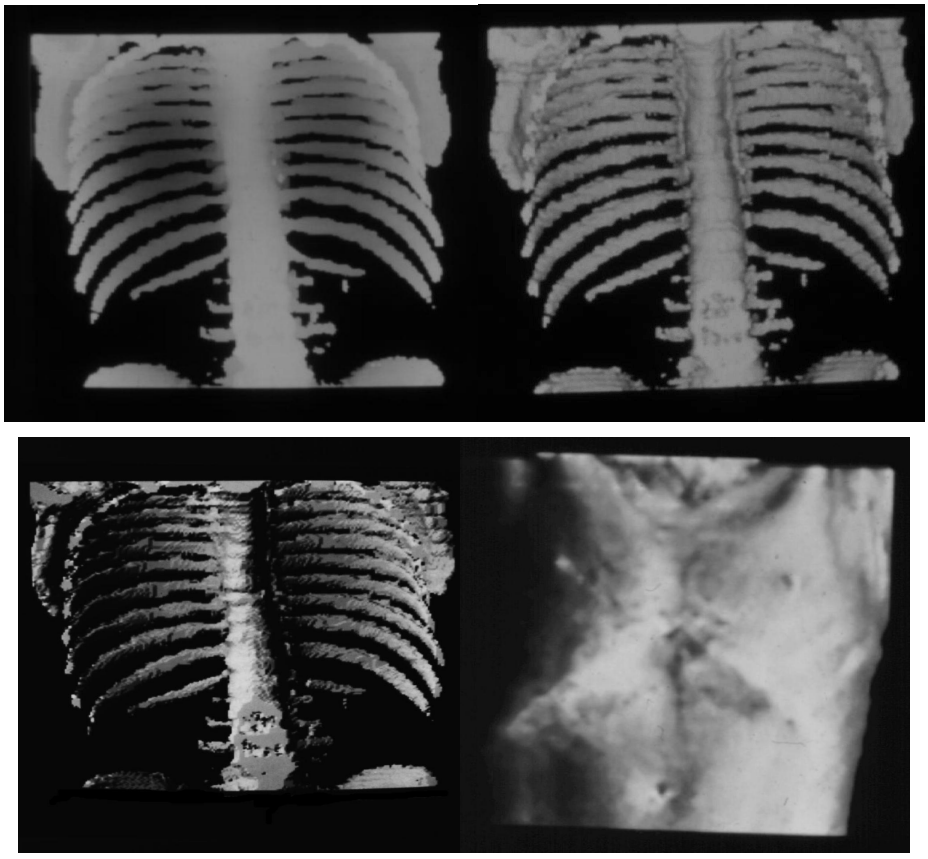

Abb.2 Frühe Visualisierung des Thorax mit distance shading, gradient shading und dem Heidelberger Raytracer 1986

4 Kurven und Flächen

Nachdem die Visualisierung heute gelöst ist, rückt als eigentliches Problem die Segmentierung in die Mitte der Aufmerksamkeit. So war es zwar schon immer, aber die technischen Randbedingungen und die Probleme mit der Visualisierung haben das zunächst überdeckt.

Schon der erste simple Ansatz vor 20 Jahren hatte Operatoren beinhaltet, die Ränder oder alternativ zusammenhängende Gebiete zu suchen. Heute findet man das alles bestenfalls rührend. Damals konnten noch Herren wie Sobel oder Roberts mit Trivialitäten berühmt werden. An den real existierenden medizinischen Bildern sind alle Operatoren aber gescheitert. Die Ränder waren eben dauernd, und sei es nur um ein Pixel, unterbrochen. Was der Mensch ohne nachzudenken als zusammenhängend erkennt, war für den Rechner zweierlei.

Wir haben in die Verbesserung der Gradientenfilter sehr viel gesteckt, erst mit Herrn Canny, dann durch look-ahead Strategien an Kurvenenden, dann durch dilatation-erosion. Der Höhepunkt war eine Arbeit zu sogenannten Illusionskanten, dem verwegenen Versuch, das menschliche Auge zu kopieren. Warum kann ein Mensch mit elegantem Schwung Grenzlinien ziehen, die objektiv im Datenraum nicht vorhanden sind? Traurigerweise haben alle Versuche erfolglos geendet, allein einige interaktive, d.h. durch den Mensch gesteuerte Verfahren werden heute noch benutzt.

Wenn es mit dem Rand nichts ist, müssen eben regionenbasierte Verfahren herhalten. Diese funktionieren immer nach dem selben Schema: man sieht nach, ob ein Nachbarpixel dem zu untersuchenden ähnlich ist. Das Dilemma steckt hier im Wort ähnlich. Man braucht erst mal ein Maß an sich, um etwas quantitativ erfassen zu können und dann eine Methode um den Abstand, oder besser die Nähe, beurteilen zu können.

Eine umfassende Studie über alle verfügbaren Nachbarschaftsmaße wurde angefertigt. Sage und schreibe 401 Texturmaße wurden untersucht, darunter Cooccurence Matrizen, Histogramm Differenzenstatistik, Runlength Matrizen, Powerspektren mit Ringsummen und Keilsummen, Gradientenstatistik sowie mehrere Dutzend Texturparameter des Robertsopearator.

Diese Arbeit erfüllt uns insofern mit Stolz, als daß man die sichtliche Intensität des Vorgehens nachvollziehen kann. Vom Ergebnis her war es eher traurig, die Maße korrelieren miteinander in erschreckendem Umfang. Ein für alle Fälle gültiges Maß konnten wir nicht finden, man muß für jede Modalität und jede spezielle Segmentierung das geeignete Maß ermitteln. Da die Maße mathematisch eher willkürlich sind, kann Erfolg oder Mißerfolg nicht interpretiert oder vorhergesagt werden.

Die Zusammenfassung der Maße und das Aufspannen eines Merkmalsraumes kann durch viele Verfahren versucht werden. Am primitivsten, am besten verstanden und am schnellsten wäre die klassische multivariate Statistik. Aber auch hier haben wir alles Mögliche versucht. Viel Energie haben wir in die topologischen Karten nach Kohonen gesteckt. Es gelang tatsächlich, die topologischen Beziehungen im Merkmalsraum auf zweidimensionalen Karten abzubilden. Es gelang nicht, diese Information wieder sinnvoll auszulesen, man scheitert zuletzt an der Abgrenzung der einzelnen Klassen untereinander. Eine böse Zunge ließ dazu die Bemerkung fallen, daß nichts besseres als bei der klassischen Statistik herauskäme, nur die Rechenzeit sich verhundertfacht.

5 Die Intelligenz

Wir haben erkannt, daß wir mit den blanken numerischen Waffen nicht viel weiter kommen. Die Zuordnung der Voxel zu Organen oder allgemein medizinischen Objekten erwies sich als diffiziler als vermutet. Weil es für den Menschen so einfach scheint, seine Umwelt intelligent wahrzunehmen, denkt man, daß es doch auch einer Maschine einfach beizubringen sein müßte. Dem ist nicht so.

Douglas Hofstadter hat in seinem Buch ‚Gödel, Escher, Bach' [4] eine für mich damals merkwürdige Definition eines Computers gegeben: ‚It's impossible for men to be unobservant'. Ich verstehe das so, daß er meint, es sei, im Gegensatz dazu, einer Maschine möglich, unaufmerksam zu sein.

Eine ähnliche Lehre will ich hier ziehen: ‚It's impossible for men to see pixels'. Wir sehen, wenn wir die Augen öffnen, die Welt in Echtzeit, farbig und dynamisch. Wir kommen gar nicht auf die Idee, diskrete Mikroeinheiten (pixel) wahrzunehmen. Wir sehen gestalt-orientiert Zusammenhänge, d.h. ganze Einheiten: ein Mensch, ein Bild, eine Leber. Exakt daran scheitert ein Rechner.

Es wurde daher folgerichtig versucht, dem Rechner eine Art von Intelligenz beizubringen. Das Gebiet der künstlichen Intelligenz war ja in den 80er Jahren sehr en vogue. Große Hoffnungen ruhten, nicht nur hier in Europa, auf diesen Entwicklungen. Die Japaner wollen in der Dekade 1980 bis 1990 die fünfte intelligente Generation von Computern entwickeln und auf den Markt bringen. Arg viel mehr als fuzzy logic im Staubsauger ist dabei nicht heraus gekommen.

Persönlich habe ich immer gehofft, daß diese Entwicklungen uns in der Bildanalyse wegweisend hilfreich sein könnten. Ich habe auch mit Interesse die Entwicklungen in der Sprachanalyse verfolgt, in der gleichen Hoffnung. Die Sprachforscher hatten es meiner Meinung nach leichter, weil es hier wenigstens einen Ansatz einer Theorie gab.

Naom Chomsky klassifizierte die Sprachen in vier Klassen, die regulären Ausdrücke, die kontextfreien Sprachen, die kontextsensitiven Sprachen und die sogenannten freien Sprachen. Zu den ersten beiden kann man Analysatoren bauen, die endlichen Automaten (finite automata) und die Kellerautomaten (pushdown automata) respektive. Für die beiden letzteren bräuchte man endliche Automaten mit endlichen oder unendlichen linearen Listen, das sind Turing Maschinen. Die letzen beiden können aus theoretischen Überlegungen heraus nicht gebaut werden. Das ist im Falle der Analyse von Computersprachen nicht weiter schlimm, weil diese doch nur den beiden ersten Klassen angehören. Problematisch sind echte Sprachen, wie z.B. Deutsch, das vermutlich der kontextsensitiven Klasse angehört. Die heutigen Spracheingabeanalysatoren sind daher nicht auf einer soliden Theorie aufgebaut, sondern auf einer simplen, trotzdem wirkungsvollen Triplett-Wortfolgenstatistik, die sich einer Phonemanalyse anschließt.

Leider hilft uns das für die Fragestellung der Bildanalyse, dem Bildverstehen, nicht weiter. Vielleicht mußte der Versuch, Rechnern Intelligenz beizubringen, sogar ganz systematisch und logischerweise scheitern. Man kann versuchen, die Struktur des Gehirns im Rechner abzubilden, das läuft auf die Modellierung von Nervenzellen und Synapsen hinaus. Das ist denkbar, setzt aber voraus, daß vergleichbare Mengen an Instanziierungen von Nerven und Vernetzungen zur Verfügung stehen und noch wichtiger, daß dieses elektrische Modell überhaupt so richtig ist.

Die meiner Meinung nach letztlich gescheiterten Modelle der letzten 20 Jahre waren noch primitiver: man beschied sich mit dem vorhandenen Werkzeug, das Rechner zur Verfügung stellen. Ein Rechner kann Bitmuster speichern, umladen, vergleichen, negieren und addieren. Alles basiert auf dem Neumann'schen Konzept von seriellen Operationen auf Speicherinhalten. Ein solches Werkzeug eignet sich gut zum Rechnen, daher der Name, aber schon viel schlechter für Text- und graphische Aufgaben. Alles, was sich in irgend einer Form als flacher Graph im Sinne der Graphentheorie darstellen läßt, ist mit guter Chance auch auf eine heutige Rechnerarchitektur abbildbar.

Dazu gehören eben Computersprachen und Rechnungen, deshalb geht das so gut. Ich vermute, daß, selbst wenn wir wüßten, wie die Wahrnehmung und Verarbeitung optischer Information vor sich geht, dann unsere heutigen Rechner zu dumm wären, diese Modelle in sich abbilden

zu lassen. Wenn wir uns in unserer Ratlosigkeit an die medizinischen Experten wenden, so wird uns hier leider auch nicht geholfen. Die rein topologische und strukturelle Beschreibung von Organen und deren nervöser Vernetzung nutzt solange nichts, wie die zugrunde liegende Funktionalität nicht klar ist.

Wenn ich keinen Plan und kein Modell habe, muß meine Modellierung der visuellen Wahrnehmung auf einem Rechner ganz logischerweise scheitern. Ein bisschen Statistik kann trösten, ein Verständnis wäre aber hilfreicher. Das Nichtwissen und Nichtverstehen hinter gewaltiger Mathematik zu verbrämen, halte ich für übel und eine Gemeinheit am einfachen Studenten.

6 Rechnen statt Verstehen

Man könnte an dieser Stelle über die existierenden Studien- oder Lehrbücher über Bildverarbeitung klagen. Wir wissen wie schwer es ist, ein didaktisch vernünftiges Werk zu verfassen. Alle Wissenschaften haben für lange Zeit ihres Bestehens den Mangel, daß die dazugehörigen Bücher ein historisierendes Vorgehen pflegen. Der berühmte Gerthsen, bis heute die Bibel der Physiker, fängt mit der Mechanik an und arbeitet sich dann Kapitel für Kapitel durch die Jahrhunderte, um mit der Atom- und Teilchenphysik zu enden.

In den Bildverarbeitungsbüchern stehen nacheinander, ohne erkennbare Logik im Aufbau, Bildverbesserung, Kantenoperatoren, Texturmaße usw. Vielleicht muß man das eben akzeptieren, weil noch niemand die geniale Idee der inhärenten Logik gehabt hat. Die Lehrbücher der Chemie waren genauso ungeordnet, damals nannte man das Alchemie (Libavius), bis endlich das Periodensystem der Elemente ein nachvollziehbares und vorhersagbares Verhalten in chemischen Verhalten der Elemente ermöglicht hat. Das ist ein schönes Beispiel für die Reifung einer Wissenschaft. Der medizinische Bildverarbeitung fehlt diese Art der Reife.

Über die angebliche Analogie der faltungsorientierten Operatoren und der Fourieroperationen will ich mich nicht auslassen. Beide haben viel gemeinsam und unterscheiden sich doch nicht nur im Vorgehen. Wie beim Dualismus von Welle und Teilchen in der Physik, so stehen auch diese beiden Methoden nebeneinander und praktischerweise setzt man einfach immer die gerade nachvollziehbar nützlichere ein.

Die klassische Faltung, also eine arithmetische Operation in kleinen Fenstern (3*3, 5*5...), die auf alle Bildpunkte nacheinander angewandt wird, zerfällt in mehrere Gruppen. Bei den einen ist irgendwo in der Arithmetik zwischen den Pixeln ein Pluszeichen. Diese Addition führt unabhängig von allen weiteren Machenschaften zu einer Glättung des Ergebniswertes. Ebenso rabiat wirkt ein einziges Minus zwischen zwei Pixeln wie eine hochfrequente Ableitung, meistens mit der Wirkung einer Rauschverstärkung. Beide Effekte sind meist ebenso unsichtbar, unerwartet wie ungewollt, systematisch aber nicht verhinderbar.

Eine weitere schwerwiegende Einsicht ist die, daß bei jedweder Operation auf einem Bild man immer das Ergebnis erhalten wird, nach dem man gefragt hat. Die Wahl einer Beobachtungsgröße, meistens simpel die Kantenlänge einer Faltung, legt den untersuchten Bandpass fest. Das entspricht im Fourierraum der Wahl eines bestimmten Frequenzbandes. Man sucht mit einer bestimmten Filtergröße eben gezielt nach Objekten genau dieser Größe. Kleinere (hochfrequentere) Objekte oder Bildinhalte fallen ebenso durch den Beobachtungsraum wie größere (niederfrequentere).

Statt die Größe der Operatoren, die auf ein Bild angewendet werden, zu verändern, um andere Frequenzen anzusprechen, kann man auch die Filter gleich lassen und dafür die Bilder verkleinern. Dieses populäre Verfahren führt zu der Familie der Pyramiden, ein Bild wird nach und

nach mehrfach verkleinert, in der Regel halbiert. Man kann nun den gleichen Operator über eine Reihe von Frequenzbändern schicken. So ein ähnlicher Prozeß scheint sich im Auge abzuspielen, wenn man die Augen zusammenkneift und im unscharfen Bild andere Details besser sieht, konkret die niederfrequenteren.

In einer großen Arbeit haben wir uns den kognitiven Texturparametern zugewandt. Es war der Versuch, über die hölzerne Mathematik hinaus (den genannten 401 Texturmaßen), vermeintlich humane Wahrnehmungsprozeduren abzubilden. Die Hauptwaffe dabei waren Fourieroperationen auf größeren Teilbildern. Der Mensch ist in der Lage, in stark verrauschten Bildern doch noch Regelmäßigkeiten zu erkennen. Wir konnten das nicht gut genug nachvollziehen. Die Hauptschwierigkeit dabei war, die rechte Fenstergröße auf dem rechten Bildausschnitt anzusetzen. Warum kann der Mensch uninteressante Gebiete so schnell als eben solche aussortieren?

Wir rechnen mehr, als dass wir verstehen. In einem wütenden Leserbrief hat einmal ein Kollege im renomierten Journal ‚Pattern Analysis and Machine Intelligence' die dort geschätzte Mathematisierung und Formalisierung in Frage gestellt. Er meinte, wir sollten doch zugeben, daß wir optische Information nicht verstehen könnten, mit einer einzige Ausnahme: dem Strichcode im Supermarkt. Den könnten wir aber nur deshalb einscannen, weil wir ihn selbst erschaffen haben. Ich gebe zu, ich habe gelacht, der Mann sprach mir aus dem Herz.

7 Wohin denn dann?

Unser Traum, mit einer AUTOMATISCHEN Bildverarbeitung den Medizinern bei Diagnose und Therapieplanung nützlich sein zu können, hat sich als ungemein viel schwieriger herausgestellt, als alle Pioniere vor 20 und noch vor 10 Jahren glauben wollten. Wir befinden uns tröstlicherweise in guter Gesellschaft. Bei den Anderen (KI, Spracherkennung) klappt es auch nicht, oder nur schlecht und recht.

Das ist aber nur der eine Teil der Problematik, den wir in unserer Eigenbetrachtung zuerst sehen. Selbst wenn wir einige der obengenannten Mängel gelöst hätten, blieben uns noch große Hürden zu überwinden.

Nach unserem Selbstverständnis soll unsere Arbeit den Ärzten und vor allem den Patienten nützlich sein. Das bedeutet aber ganz einfach, daß unsere Arbeit dort Wirkung zeigen muß, wo therapeutische Prozesse stattfinden und das ist nicht in unseren Softwarelabors. Unsere Arbeit muß vor Ort in der Praxis und in der klinischen Routine etwas nutzen. Das bedeutet mit anderen Worten, daß wir hinaus müssen ins feindliche Leben (Schiller).

Wenn ich die mir bekannten Arbeiten in Deutschland und auch in den USA Revue passieren lasse, ist davon selten die Rede. Es wird zu gern in den geschützten Biotopen der Labors herumgebastelt. Schon die Besorgung von mehr als ein paar Testbildern gestaltet sich zum großen Hindernis. Das Abgreifen und Lesen der primären Bilddaten ist selten gut gelöst.

Die Integration in die medizinische Routine liegt ebenfalls sehr im Argen. Wir medizinischen Bildverarbeiter haben uns bisher um die Arbeitsabläufe in der Klinik oder Praxis definitiv zu wenig gekümmert, so als ginge uns das gar nichts an. Es kommt dazu, daß eben diese Abläufe (workflows) oft unregelmäßig, unlogisch und Ritualen unterworfen sind. Sie bilden oft Historie und Hierarchie ab, statt Vernunft und praktisches Denken. Das ist aber nicht unsere Sache, da einzugreifen und verbessern zu wollen. Wir müssen naht- und fugenlos unsere Werkzeuge integrieren in die nun mal vorhandenen ritualisierten Abläufe.

Gerne vergessen wird auch, daß eine erfolgreiche Integration nicht mit der Installation abgeschlossen ist, wie glatt auch immer im Arbeitsablauf. Der nächste oft vergessene Baustein ist

die klinische Evaluation. Unsere tollen Algorithmen müssen einen Härtetest der Wahrheit und Nützlichkeit über sich ergehen lassen. Alles muß an hunderten Patienten ausgetestet und mit konkurrierenden Verfahren verglichen werden.

Hier handelt es sich um deutlich personenbezogene Daten, also gelten alle Datenschutzbestimmungen wie sonst kaum. Es handelt sich auch um extrem relevante Prozesse, wo es schnell um Leben und Tod gehen kann. Es ist also nicht unbillig, wenn hier Gesetze wie z.B. das Medizingerätegesetz gilt.

Die Gruppen, die die letztgenannten neuralgischen Punkte berücksichtigen, soll man loben. Es geht hier sehr schnell aus der reinen medizinischen Bildverarbeitung hinaus. Diese Arbeit wird manchmal als unakademisches Engineering abgetan. Das ist ein sehr kurzsichtiger und unfairer Schluß.

8 Schluß

Erst die Integration von vernünftiger und wenn möglich eleganter medizinischer Bildverarbeitung mit den klinischen Handlungsabläufen, mit einer pingeligen Qualitätskontrolle, wird die von uns allen gewünschte Verbesserung der Diagnose und Therapie für uns alle, Ärzte und Patienten, bringen. Wir sind nicht da, wir sind auf dem Weg.

9 Literatur

1. Rosenfeld A: Digital Picture Processing, Academic Press, New York (1976)
2. Pratt WK: Digital Image Processing, Wiley, New York (1978)
3. Serra, J: Image Analysis and Mathematical Morphology, Academic Press, London (1982)
4. Hofstadter DR: Gödel, Escher, Bach ein Endloses Geflochtenes Band, Klett-Cotta, Stuttgart (1994)
5. Gerthsen C, Kneser HO, Vogel H: Physik, Springer Verlag, Berlin, Heidelberg, New York (1982)