

Tagung
Bildverarbeitung in der industriellen Praxis
Steyr, Österreich
5.-6. September 1995

MASCHINELLES SEHEN:
WAS WIR FÜR DIE ZUKUNFT ERWARTEN

Franz Leberl, Axel Pinz
Institut für Computerunterstützte Geometrie und Graphik
Technische Universität Graz
A-8010 Graz, Austria

Zusammenfassung

Die Automatisierung des menschlichen Sehvorganges beruht auf den Prinzipien der digitalen Bildverarbeitung, wie sie im US-NASA Raumfahrtprogramm ab Mitte der 60er-Jahre entwickelt und von da aus in verschiedene Anwendungen gebracht wurde. Anwendungen in der Industrie, vor allem in der Robotik, werden oft mit dem Begriff des maschinellen Sehens (Computer Vision) bezeichnet. Die 30-jährige Geschichte dieser gänzlich neuen Technologie wurde sehr von der Entwicklung des Computers bestimmt. Erst die derzeitigen Prozessoren haben die wesentlichen Neuerungen von der Hardware in die Software und Methoden verlagert. Es werden die methodischen Entwicklungen daher die Zukunft des maschinellen Sehens prägen.

1 MASCHINELLES SEHEN

Es bestehen etwa drei historische Entwicklungen, welche zum heutigen maschinellen Sehen geführt haben. Das ist zunächst das Raumfahrtprogramm der NASA, welches im Jahre 1964 erstmals Videobilder des Planeten Mars in digitalisierter Form zur Erde übertrug (Nathan, 1966, Rindfleisch, 1971). Daraus entwickelte sich ein intensives Arbeitsgebiet der klassischen Bildvorverarbeitung zur nachfolgenden visuellen Betrachtung. Etwa gleichzeitig wurde der Gedanke geboren, die landwirtschaftlichen Nutzflächen in den USA aus Flugzeugen mittels Multispektralscannern zu beobachten, um daraus Erntevorhersagen zu entwickeln (Hoffer, 1967). Etwa ab 1966 entspann sich daraus ein intensives Arbeitsgebiet zur automatisierten Klassifizierung von Bildpixeln in die diversen landwirtschaftlichen Anbauflächen. "Bildverarbeitung" und "Bildanalyse" waren und sind die Begriffe, welche diese Arbeit beschreiben.

Etwas abgesondert von diesen Entwicklungen befaßte sich das entstehende Fachgebiet der künstlichen Intelligenz besonders mit dem menschlichen Sehsinn. Gerade am Massachusetts Institute of Technology wurden grundlegende Verfahren der rechnergestützten Bildanalyse entwickelt (Marr, 1982). Diese Überlegungen waren im Gegensatz zu den genannten planetaren und landwirtschaftlichen Motiven durch den Wunsch beflügelt, einen Roboter mit einem Sehsinn auszustatten. Dies wiederum stellt daher eine Entwicklung zur industriellen Nutzung digitaler Bilder dar und prägte die Begriffe des "maschinellen Sehens", auch "Computer Vision", in enger Verbindung mit dem Wunsch, Fertigungsmethoden weiterführend zu automatisieren.

Die Medizin erhielt mit der Erfindung der Computertomographie (CT) ein inhärent digitales, computergestütztes, bildgebendes System, welches ebenfalls als eine der verfahrensbegründenden Entwicklungen gesehen werden muß (Hounsfield, 1973). Obwohl die Dreidimensionalität der CT-Daten sehr wohl eine Eigenheit der medizinischen Anwendung darstellt, sind die Bildanalysemethoden der Medizin oft weitgehend von Entwicklungen in anderen Gebieten befruchtet worden, nicht aber umgekehrt. Die vielen heutigen digitalen Bildgeber in der Medizin nutzen Bildverarbeitungsmethoden, welche oft aus anderen Fachgebieten kommen.

Natürlich beinhaltet eine Aufzählung der Einflüsse auf die Innovationen digitaler Bildanalyse auch Neuerungen, welche aus militärischen Gründen entstanden sind. Hier gilt es, zwischen den Erkundungsverfahren aus Satelliten oder Flugzeugen und den Steuerungen von Lenkwaffen zu unterscheiden. Ersteres entspricht eher dem Methodenschatz der Planeten- und Landwirtschaftsforschung, letzteres eher dem maschinellen Sehen eines Roboters.

Damit haben wir aus einer Betrachtung der geschichtlichen Wurzeln zwischen den Begriffen der digitalen Bildverarbeitung, der Bildanalyse und dem maschinellen Sehen unterschieden. Es ist deutlich, daß es der letzte Begriff ist, welcher auf die industrielle Anwendungsumgebung anzuwenden ist. Der heutige Wissensstand ist in neueren Lehrbüchern gut dokumentiert, etwa in dem 2-bändigen Text von Haralick und Shapiro (1992,1993). Im folgenden wollen wir die heutige Situation der Bearbeitung digitaler Bilder charakterisieren. Daraus wollen wir begründen, daß digitale Bilder in viele Lebensbereiche eindringen werden, daß die derzeit noch scharfe Trennung zwischen natürlichen Bildern von Objekten und der Umwelt einerseits, und den künstlichen Bildern der Computergrafik andererseits, verwischt wird, und daß schließlich ebenso wie im Computerwesen allgemein auch in diesem Bereich gilt, daß die persönliche Unterhaltung und Freizeitgestaltung genauso oder gar mehr als andere, etwa technische Anwendungen, die Zukunft bestimmen werden.

2 DER STATUS QUO

2.1 Anwendungen

Pixelverarbeitung mag ein Begriff sein, der vielleicht noch allgemeiner als das maschinelle Sehen umschreibt, worüber wir hier berichten, wobei wir uns zunächst auf Pixel beschränken, welche sogenannte natürliche Bilder betreffen, also mit einem Sensor erfasste Daten. Wesentliche Anwendungsgebiete digitaler Bilder wurden schon genannt: (a) bildgebende Verfahren der Medizin, (b) die automatische Fertigung, (c) die Lenkwaffensteuerung und militärische Erkundung, (d) die Planetenforschung und Erdbeobachtung. Dazu kommt nun vermehrt (e) das automatisierte Büro mit seinem "Workflow-Management" von digitalen Dokumenten, ganz besonders im Bankwesen; (f) die hypermedialen Informationssysteme mit 2- und 3-dimensionalen Objekten; und schließlich (g) das digitale Fernsehen und die Unterhaltungselektronik, wie es durch eine Kombination von Computergrafik und natürlichen Digitalbildern verkörpert wird.

Die wirtschaftliche Nutzung der Technologie der Pixelverarbeitung ist zwar weitläufig, aber in hohem Maße fragmentiert. Es besteht keine ausgeprägte bildverarbeitende Industrie, obgleich es viele kleine Softwarefirmen gibt, welche bildorientierte System- und Softwarelösungen für gewisse Märkte anbieten. Im allgemeinen findet man aber das technische Wissen über das maschinelle Sehen innerhalb größerer EDV-Projekte, oder die Bildverarbeitung ist eine technische Komponente (ein Werkzeug unter vielen) eines Betriebes, dessen Erfolgsgeheimnis aber oft in anderen Stärken begründet ist.

2.2 Bildgebung und Datenraten

Maschinelles Sehen beruht auf digitalen Pixel-Matrizen. Diese entstammen aus elektro-optischen, Schall-, Mikrowellensensoren, aus der Analog-Digital-Wandlung von Papier- oder Filmvorlagen durch

Scannen, oder entstehen aus der Interpretation 2-dimensionaler Meßdaten als "Bild", wie etwa in der Seismik. Kaum wird heute die Erfahrung gemacht, daß die Sensorik für die Lösung einer Sehaufgabe der entscheidende Engpunkt ist. Wir können im allgemeinen erwarten, daß die 512 x 512, die 1024 x 1024 oder auch größeren Flächen-CCD-Kameras, oder die Zeilen-CCD-Kameras mit 2000 bis 8000 Elementen ausreichen, um die notwendigen Bilder zu erzeugen. Nur in gewissen Grenzfällen besteht ein Sensorik-Engpaß. Eher kann die Datenrate, mit welcher die Pixel aus dem Sensor fließen, Anlaß zu Sorgen werden, aber auch hier bestehen Lösungen im Bereich von Multi-MBytes/Sekunde.

2.3 Rechnerleistungen

Die Zeit der speziellen Bildverarbeitungsrechner ist seit Mitte der 80er Jahre vorbei. Sie wurde zunächst durch die Nutzung von speziellen Platinen als Einschübe in Arbeitsplatz-Rechner und auch Personal Computer ersetzt. Derzeit sind auch diese Sonderlösungen nur mehr beschränkt notwendig und die 100 bis 200 MHz Hochleistungschips mit 400 MIPs und 100 MFLOPs schaffen die meisten der in der Bildverarbeitung anfallenden Aufgaben ohne weitere Hilfsplatinen. Signalverarbeitungs-chips, etwa von der Art des Texas Instruments C80, sind in der Lage, für einen Beschaffungspreis von unter öS 5.000,- den Output einer Videokamera in Echtzeit sowohl geometrisch als auch radiometrisch zu verarbeiten. Die typischen zeitgemäßen Konfigurationen solcher Chips mit mehreren zusammenarbeitenden Elementen dienen heute in den höchsten Anforderungen bei Verwendung komplexer Algorithmen.

3 BEISPIELE

Wir unterscheiden in unserer eigenen Grazer Institutsarbeit zwischen 2- und 3-dimensionalen Anwendungen der Bildanalyse. Im folgenden präsentieren wir einige eigene Arbeiten, um den Zustand der derzeitigen Anwendung des maschinellen Sehens zu charakterisieren.

3.1 Stahlschliffe

Die Qualitätskontrolle von Schnellarbeitsstählen wird von Metallographen mittels visueller Klassifizierung speziell präparierter Stahlschliffproben unter Verwendung von Lichtmikroskopen durchgeführt. Im Falle der Karbidverteilung erfolgt eine Zuordnung der Proben zu einer Richtreihe mit 28 Referenzbildern. Dieser Vergleich ist mittels Bildanalyse automatisierbar. Es wird hiezu zusätzlich zu den Referenzbildern ein klassifizierter Bilddatensatz angelegt, dessen wesentliche Elemente bestimmt werden (Abbildung 2). Ein neuer Stahlschliff wird nun ebenfalls abgebildet und analysiert. Das ihm ähnlichste Referenzbild wird gefunden und zur Beurteilung der Stahlgüte verwendet (Abbildung 1). Dieses Verfahren wurde von Wiltschi u.a. (1995) in Zusammenarbeit mit Böhler-Edelstahl entwickelt. Es befindet sich derzeit in der Prüfung für die operationelle Anwendung.

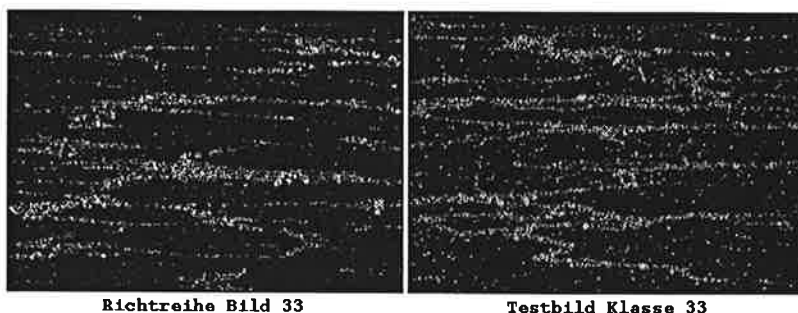


Abbildung 1 Gegenüberstellung: Referenzbild und klassifiziertes Testbild einer präparierten Stahlschliffprobe. Gezeigt wird eine Metalloberfläche von 0.89mm x 0.64mm

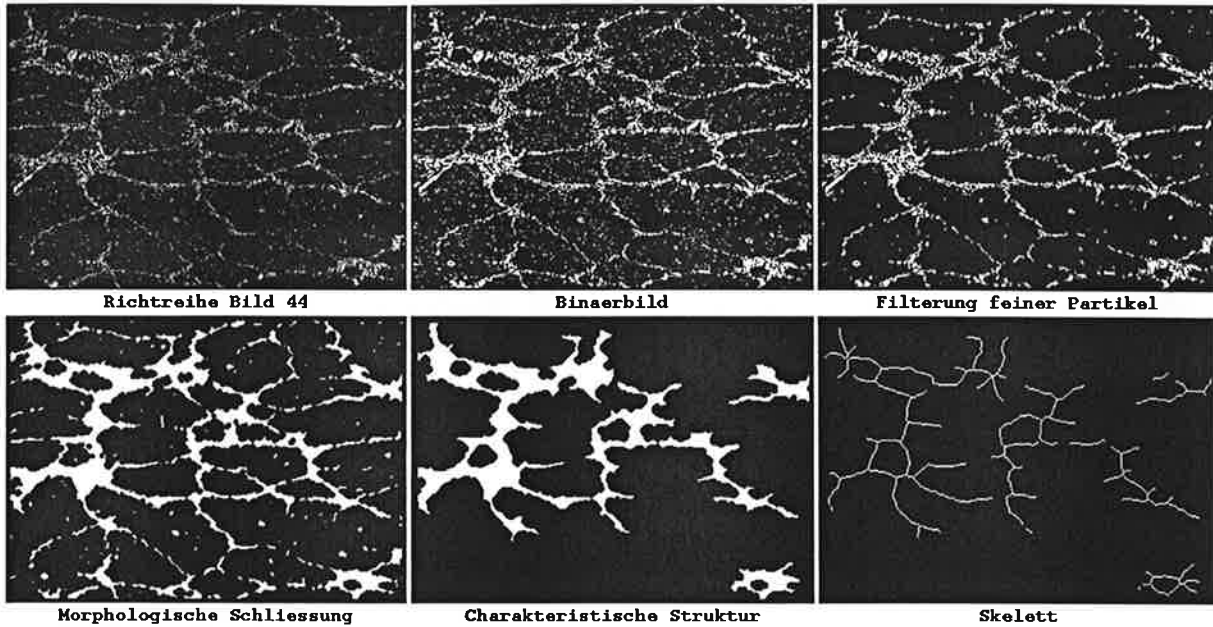


Abbildung 2 Extraktion der wesentlichen Struktur einer Karbidverteilung in Schnellarbeitsstählen

3.2 Melanome in der Dermatologie

In der Hautmedizin ist wegen des wachsenden Auftretens von Hautkrebs die Sorge um eine Dokumentation und verbreiterte Anwendung von Expertenwissen groß. Es gilt, die Betrachtung und manuelle Prüfung von Hautläsionen mittels maschineller Prüfung zu unterstützen. Hierzu wurde in einem gemeinsamen Projekt des Allgemeinen Krankenhauses Wien, dem Forschungszentrum Seibersdorf und unserem Grazer Institut ein Verfahren entwickelt, welches sich nun einer genauen Prüfung zu stellen hat (Ganster u.a., 1995). Bilder von Hautläsionen werden maschinell geprüft und einer von drei Klassen zugeordnet: gutartig; bösartig; nicht eindeutig (Abbildung 3 und Abbildung 4 zeigen einige typische automatisch extrahierte Merkmale). Erste Versuche waren ermutigend. Es wird nun eine große Anzahl von Patienten mit dem Verfahren zu untersuchen sein, um in einem Vergleich mit herkömmlichen manuellen Prüfmethode die Verlässlichkeit der maschinellen Methoden besser zu verstehen.

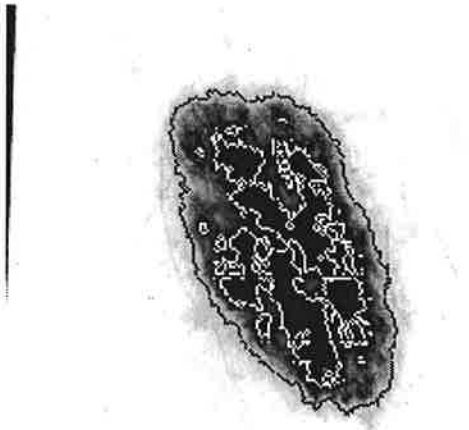


Abbildung 3 Hautläsion mit ihrer äußeren Begrenzung (schwarz) und Flecken (weiß). Diese Läsion ist einige mm lang.

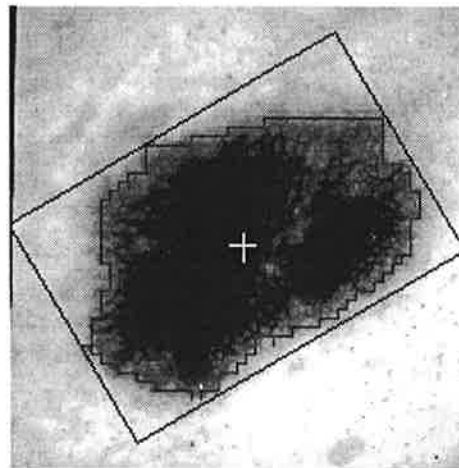


Abbildung 4 Einige Parameter einer Hautläsion: Umgebendes Rechteck (schwarz), Begrenzung (grau), Schwerpunkt (weiß)

3.3 Ophthalmologie

In diesem Projekt, das gemeinsam mit der Augenklinik am Wiener AKH durchgeführt wird, sollen neue Diagnose- und Therapieverfahren für die altersbedingte Makuladegeneration, die Hauptursache für bleibende Sehverschlechterungen bis hin zur völligen Erblindung, entwickelt werden. Zu diesem Zweck bedient man sich eines Scanning Laser Ophthalmoskopes (SLO). Das Gerät liefert in mehreren verschiedenen Untersuchungen Bilder, deren wesentliche Inhalte zu einer "Landkarte" der Retina verschmolzen werden. Abbildung 5 zeigt ein SLO-Bild, Abbildung 6 die daraus abgeleitete Landkarte mit Papille, Blutgefäßen, Fovea, und krankheitsbedingten Veränderungen. Die Verfahrensentwicklung läßt eine baldige vollautomatische Lösung dieser Aufgabe erwarten (Pinz u.a., 1995).

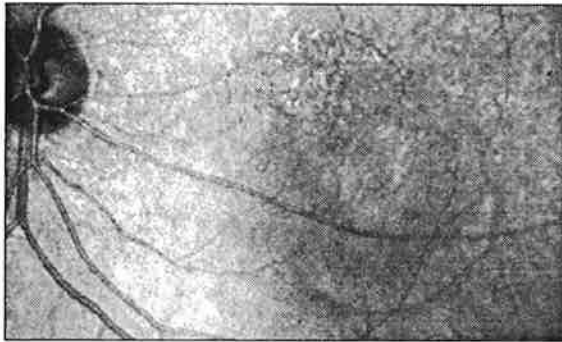


Abbildung 5 Ein SLO-Bild (Infrarot-Laser) der menschlichen Retina (ca. 6mm x 4mm)

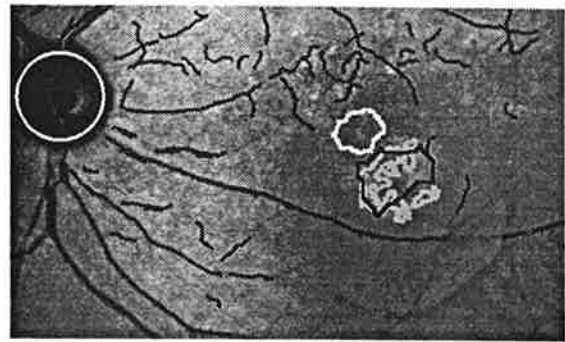


Abbildung 6 Die zugehörige Landkarte mit anatomischen und pathologischen Merkmalen

3.4 Aluminiumgußteile

Die Massenfertigung von mehreren hunderttausend Teilen eines Motors erfordert wegen des harten internationalen Wettbewerbs die Nutzung jeden kleinen Vorteils beim Vermeiden von Kosten und beim Erhöhen der Zuverlässigkeit und Qualität des Produktes. Die Entdeckung von Fabrikationsfehlern schon sehr früh in der Bearbeitungskette kann die Kosten senken; die automatische Endkontrolle der gefertigten Teile kann die dem Kunden sichtbare Leistung eines Betriebes steigern. Am Beispiel gefertigter Aluminiumgußteile (Abbildung 7) war zu zeigen, ob ein Fertigungsroboter instruiert werden kann, ein Teil als fehlerhaft zu erkennen und die Fertigung zu unterbrechen. Abbildung 8 illustriert die Kamera- und Beleuchtungseinrichtung, welche in der Interaktion mit dem Manipulatorarm des Roboters die visuelle Prüfung ermöglicht. Abbildung 9 zeigt einen typischen Fabrikationsfehler, Abbildung 10, daß dieser in der Tat maschinell erkennbar und ausscheidbar ist (Nischelwitzer u.a., in Vorbereitung).

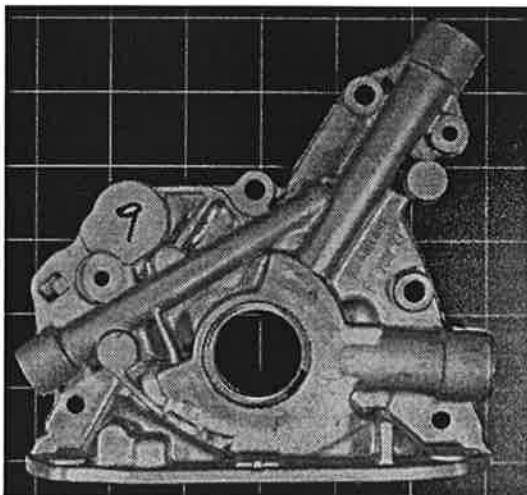


Abbildung 7 Aluminium-Druckgußteil aus dem Motorenbau, zirka 15cm groß



Abbildung 8 Kamera und Beleuchtung sind innerhalb des Schutzgitters angebracht, um das vom Roboter gehaltene Werkstück abzubilden

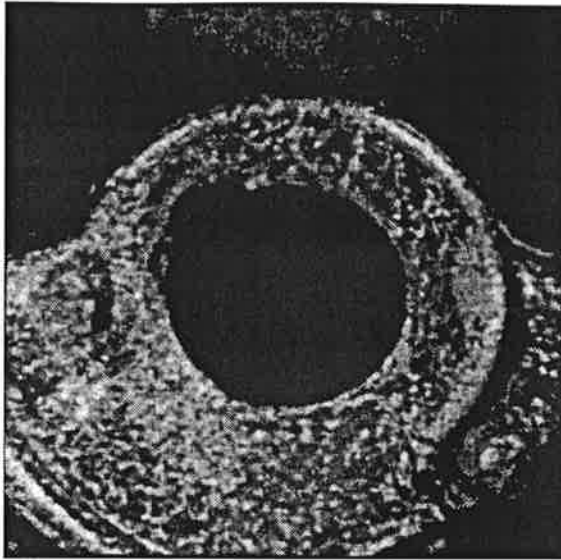


Abbildung 9 Bohrung mit Ausbruch im Aluminiumteil aus Abb.7

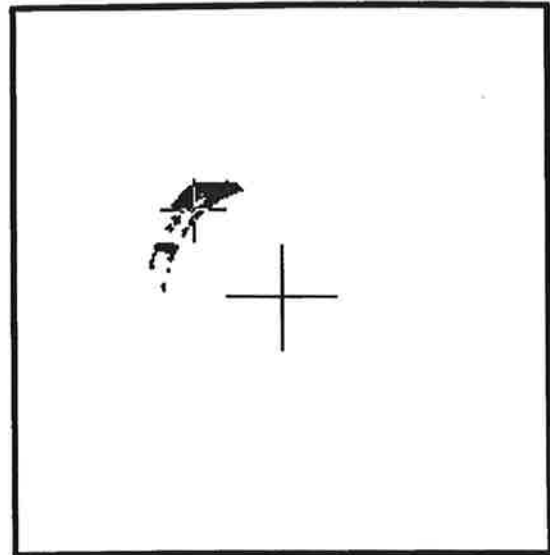


Abbildung 10 Detektion des Fehlers in Abb.9

3.5 Endoskopische Analyse von Verbrennungsabläufen in Motoren

Als letztes Beispiel zwei-dimensionaler Fragestellungen sei die derzeit laufende endoskopische Studie von Verbrennungsabläufen in Motoren genannt. Abbildung 11 ist ein durch ein Endoskop erfaßtes Bild einer Explosion im Zylinder eines Motors. Aus einer Serie solcher Bilder ist die Ausbreitung des Explosionsvorganges feststellbar. Die einzelnen Bilder einer Bildfolge dienen dem Zweck, den Vorgang zu rekonstruieren, wobei die Flamme von Bild zu Bild zu verfolgen ist. Die Berechnung der Flammentemperatur (Abbildung 12) und der Rußkonzentration aus den aufgenommenen Farbbildern liefern weitere Informationen über den Verbrennungsvorgang.

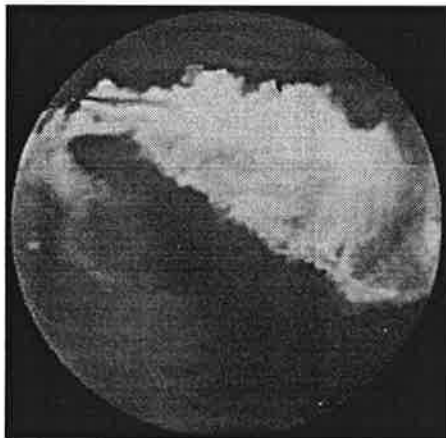


Abbildung 11 Endoskopisches Verbrennungsbild aus dem Inneren eines Motors

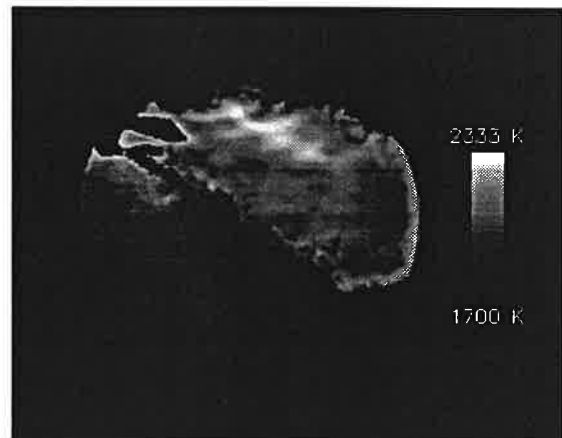


Abbildung 12 Berechnung der Flammentemperatur

3.6 Analyse von Bruchoberflächen aus Stahlproben

Zur Bestimmung von Materialparametern ist es oft notwendig, die Topographie der Oberfläche nach Bruchversuchen zu bestimmen. Abbildung 13 zeigt ein mittels Rasterelektronenmikroskop erzeugtes Stereobildpaar einer solchen Stahlbruchstelle, Abbildung 14 stellt die 3-D Oberflächenrekonstruktion dar. Aus einer Formanalyse des 3-D "Bruchgebirges" lassen sich Aussagen über die Festigkeit und das Bruchverhalten des Werkstoffes herleiten (Scherer et al. 1995).

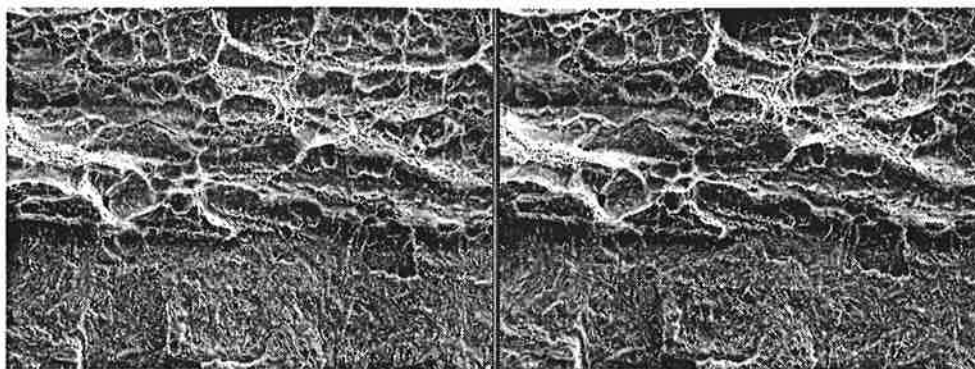


Abbildung 13 Stereobildpaar einer Stahlbruchstelle, (Elektronenmikroskopaufnahmen $120\mu\text{m} \times 90\mu\text{m}$)

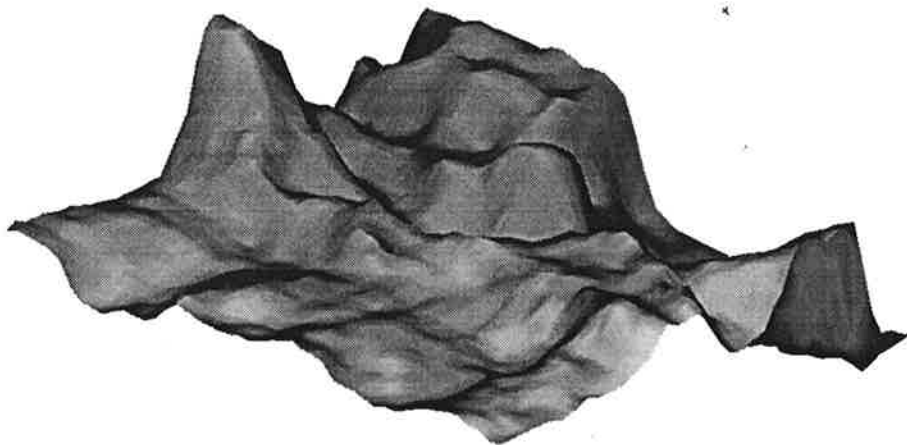


Abbildung 14 3-D Oberflächenrekonstruktion aus Abb.13, Höhenunterschiede betragen etwa $25\mu\text{m}$.

3.7 Motorblock und Dieseleinlaßteil

Ein genaues 3-D Modell eines Motorblocks wird benötigt, wenn Strömungsuntersuchungen für einen Motor durchzuführen sind. Dann ist die tatsächliche Form mit Fehlern von weniger als $\pm 0,5\text{ mm}$, nicht die Form aus dem Entwurfsdokument zu verwenden. Abbildung 15 ist eine Photographie und Abbildung 16 das 3-D Modell des Wasserkernes eines Motorblocks, mit dem gezeigt wurde, daß bildgebende Verfahren durchaus geeignet sind, die Form mit ausreichender Genauigkeit zu bestimmen.

Ähnlich ist die Aufgabe bei der Vermessung von gefertigten Dieseleinspritzkanälen. Abbildung 17 zeigt ein Ergebnis, mit dem illustriert werden konnte, daß auch Genauigkeiten im Bereich von $\pm 0,1\text{ mm}$ eingehalten werden können.

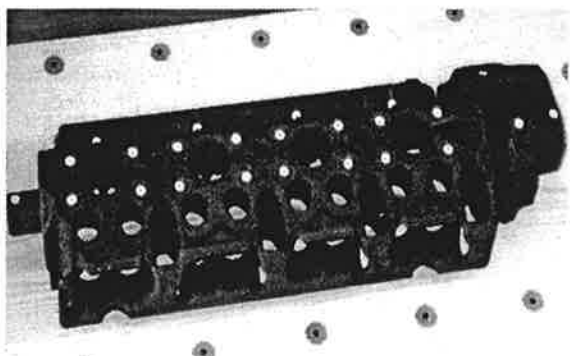


Abbildung 15 Wasserkern eines Motorblocks; ein etwa $0,5\text{m}$ langes Objekt

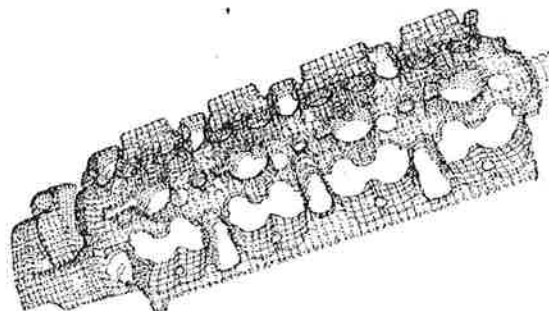


Abbildung 16 3-D Modell des Wasserkerns

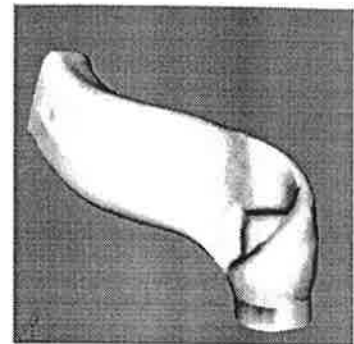
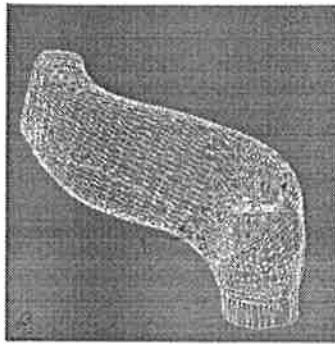
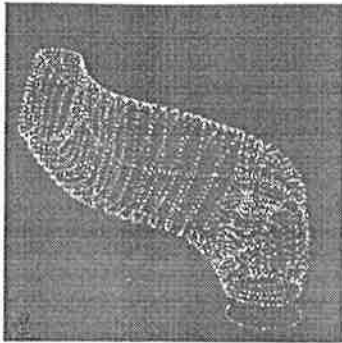


Abbildung 17 Vermessung, 3-D Rekonstruktion und Visualisierung von Dieseleinspritzkanälen

3.8 *Drei-dimensionales Modell einer Statue*

Stellvertretend für eine an unserem Institut bearbeitete Projektserie zur Modellierung von Gebäuden und Gebäudeinnenräumen, als auch von Rauminhalten wie etwa Möbel und Statuen, sei hier nur auf die Herstellung eines 3-D Modells einer Statue des Habsburger Kaisers Karl VI verwiesen. Abbildung 18 ist eine Photographie der Statue, Abbildung 19 die Darstellung einer Ansicht des 3-D Modells im Rechner. Die Arbeit von Leberl, Gruber u.a. (1994) und Uray und Pinz (1995) befaßt sich mit der Qualität und den Kosten solcher Rekonstruktionen und der Verbesserung der Kostensituation durch Entwicklung automatischer Bildanalysemethoden.



Abbildung 18 Marmorstatue Kaiser Karl VI



Abbildung 19 3-D Modell der Statue mit aufgebracht Phototextur

3.9 Immersive Benutzerschnittstellen

Die Interaktion mit 3-D Objekten am Bildschirm wird erst dann wirklich lebensnah möglich, wenn der Benutzer ein stereoskopisches Bild wahrnehmen kann. Dann ist aber auch ein Cursor anzubieten, der es gestattet, die Objekte 3-dimensional zu markieren, zu verändern, zu erweitern. Abbildung 20 zeigt ein experimentelles und auf 2 Kameras gestütztes System, welches die Bewegungen des Benutzers verfolgt. Die Stereobrillen des Anwenders und sein 3-D Griffel werden in ihrer Lage und Orientierung im 3-D Raum vermessen und übermitteln dem Rechner die notwendigen Informationen, um auf die Bewegungen des Kopfes und der Hand mit dem Griffel zu reagieren. Damit entsteht ein geschlossener Regelkreis mit dem Betrachter, der die 3-D Informationen mit einbeziehen kann. Man kann daher von "immersiver", also einer eintauchenden Benutzerschnittstelle sprechen, da der Benutzer sich inmitten der Daten befindet und mit ihnen realitätsnahe interagieren kann.

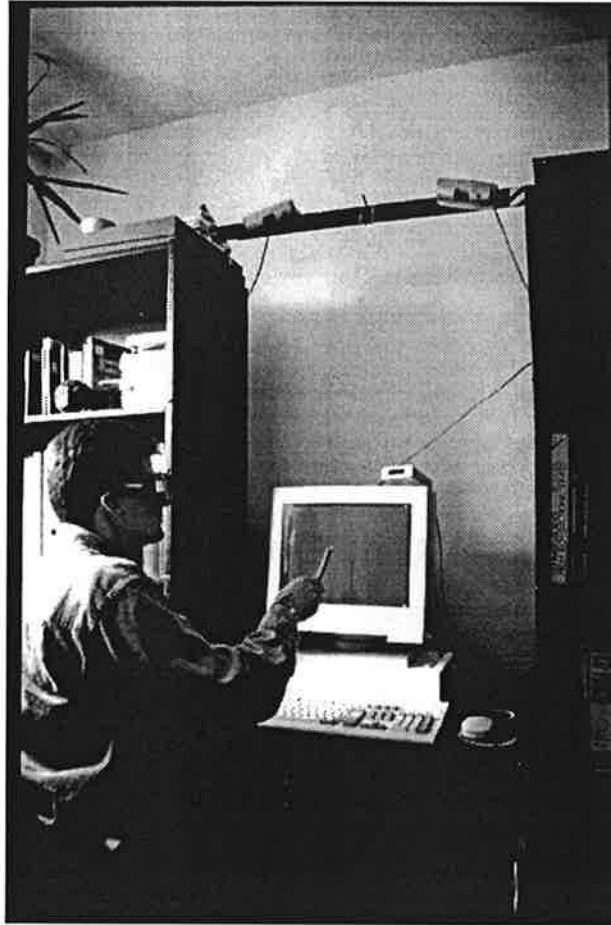


Abbildung 20 Experimentelles System für eine immersive Benutzerschnittstelle mit 2 Kameras, welche die Stereobrille und den 3-D Griffel verfolgen.

3.10 Planetenoberflächen

Stellvertretend für den weiten Anwendungsbereich automatischer Bildanalyseverfahren in der Fernerkundung und Erdbeobachtung zeigt Abbildung 21 ein Radarbild aus NASAs Magellan-Mission zum Planeten Venus (1990 bis 1994). Dieses Gebiet der Venusoberfläche wurde durch Nutzung von Verfahren der Stereomessung und des Shape-from-Shading 3-dimensional rekonstruiert (Abbildung 22, Leberl u.a., 1994). Diese Formmodelle sind wiederum notwendig, um Bilder unterschiedlicher Provenienz, etwa aus verschiedenen Phasen der Mission, miteinander zu verbinden.



Abbildung 21 Magellan-Radarbild der Venusoberfläche, der Ausschnitt ist ca. 180km breit.



Abbildung 22 3-dimensionale Rekonstruktion durch Kombination von Stereo und Shape-from-Shading, hergestellt aus Abb.21.

3.11 Schluß

Die gezeigten Anwendungen der Bildanalyse und des maschinellen Sehens sind nur ein ganz kleiner Ausschnitt dessen, was in Medizin, industrieller Prüfung, Robotersteuerung, Erdbeobachtung usw. heute bearbeitet wird. Dies ist so vielfältig, daß nicht einmal der Erfahrungsschatz unseres doch noch recht neuen (seit Ende 1992 bestehenden) Institutes hier vollständig zusammengefaßt werden kann. Dies stimmt für die Zukunft recht optimistisch.

4 ENTWICKLUNGSRICHTUNGEN

Um nun den Blick in die Zukunft zu werfen, wollen wir drei gesonderte Aspekte des maschinellen Sehens betrachten: das Systemdenken; die Sensorik und Rechnerinfrastruktur; die Anwendung.

4.1 Systemdenken versus Algorithmen

Es wird immer deutlicher, daß die bisherige Trennung in Bildgebung, Algorithmik und Rechnerarchitekturen nicht haltbar ist. Erfolgreiche Verfahren des maschinellen Sehens optimieren die Lösung mit allen ihren Komponenten.

Pinz (1994) propagiert eine Denkweise, wie sie in Abbildung 23 dargestellt ist. Hier wird auch angeregt, daß Raum für Mehrfachbilder, für aktives Sehen im Sinne von Bildgebung als Ergebnis einer vorläufigen Bildanalyse, und schließlich auch für die Verknüpfung verschiedener Methoden der Bild- und Wissensverwendung gegeben sein sollte. Dies erweitert die eher traditionellere Denkweise, welche sich aus dem Konzeptdenken nach Gonzalez und Woods (1992) ergibt (siehe Abbildung 24).

Klar ist jedenfalls, daß die Sensorik, die Beleuchtung, das Expertenwissen, die Rechnerarchitektur, die Verfahrensbausteine usw. alle als Parameter bei der Suche nach einer Lösung veränderbar sein müssen. Diese Vielfalt an Möglichkeiten verwirrt und behindert die Verwendung des maschinellen Sehens. Daher wird zu erwarten sein, daß für diese Technologie höherwertige Werkzeuge entstehen werden, welche nicht als C- oder C++ Code angeboten werden, sondern in höherwertiger Sprache in Verbindung mit Bildgebung und Rechnerkonfigurationen bereitstehen müssen.

Die Wissensvermittlung zum Benutzer wird stark zu verbessern sein. Derzeit ist es notwendig, in der internationalen Literatur mühevoll nach Rollenmodellen für gewisse Anwendungen zu suchen, um Ideen und Algorithmen für eine gegebene Problemstellung zu finden. Die heutige Alternative ist die "Neuerfindung des Rades". In Zukunft wäre zu hoffen, daß sich ganze Kataloge erstellen lassen, wie dies in anderen Ingenieurdisziplinen Tradition hat, in denen die Problemstellungen standardisiert werden. Daraus wäre dann für den Anwender ein System aus geeigneten Komponenten zu konstruieren.

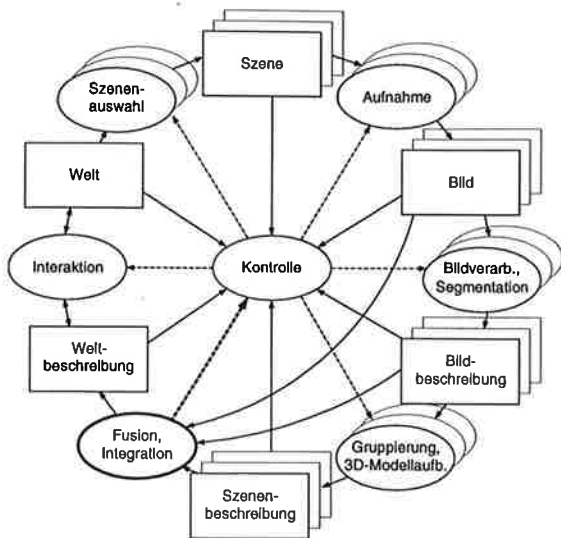


Abbildung 23 Systemmodell für Bildverstehen mit einer aktiven Fusion-Komponente nach Pinz (1994)

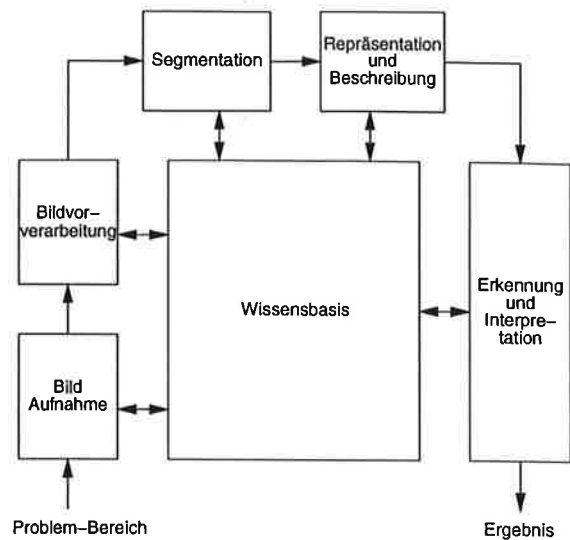


Abbildung 24 Grundlegende Verarbeitungsschritte in einem Bildanalyse-System nach Gonzalez und Woods (1992)

4.2 Sensorik, Rechner und Telekommunikations-Infrastruktur

Die Bildgebung ist heute schon rudimentär durch die Multi-Media-Anwendung in die Arbeitsplatz-rechner integriert. Die Leistung dieser Bildgebung sollte sich dramatisch verbessern, wenn die Beschränkungen im Speicher, in der Rechnerleistung und in der Datenübertragung verringert werden. Die Massenfertigung von höher auflösenden elektro-optischen Kameras steht daher zu erwarten. Ob es nicht doch zu einem Standard der Flächenkameras mit 1024 x 1024 Pixel kommt? Die Akzeptanz dieser Norm ist heute vor allem wegen der im Fernsehen bestehenden Formate behindert. Mit HDTV und den schnellen Datenübertragungsverfahren wie ATM würde die Infrastruktur gegeben sein, die nach Hochleistungsanwendungen sucht, welche aus der digitalen Pixelbearbeitung entstehen.

Wir erwarten daher, daß nicht die Anwendung selbst die treibende Kraft für die Weiterentwicklung der Technologie des maschinellen Sehens ist, sondern daß die Innovationen in der Rechnerleistung und in der Datenübertragung die Kosten von Komponenten beeinflussen und daß daraus die Akzeptanz der Technologie gesteigert wird.

In der Rechnerleistung wird ja noch immer erwartet, daß die bekannten Leistungssteigerungen fortgesetzt werden und in den kommenden 15 Jahren noch einmal ein Preis-Leistungsschub um einen Faktor von 10.000 erreicht wird. Gerade die Pixelverarbeitung wird von diesen Steigerungen profitieren, die ja für die Massenverwendung des PC als Schreibmaschine und Tabellenkalkulator nicht relevant wären, wohl aber im Zusammenhang mit Multimedien .

Bei der Datenübertragung stehen wir noch am Anfang. Die einfachen Telefon- und Modemverbindungen mit Datenraten im Bereich von 9,6 Kbaud gehören bald der Vergangenheit an und werden durch die ISDN-Übertragung mit 128 Kbaud ersetzt. Dazu kommen aber die sehr hohen Übertragungsraten aus Glasfaserverbindungen und ATM-Protokoll im Bereich von zunächst 155 Mbaud und schließlich mit vielen Gbaud. Wenn dann zusätzlich durch den Konkurrenzdruck auch die Übertragungspreise purzeln, steht der verstärkten Nutzung der digitalen Bildtechnologie kaum noch ein Hindernis im Wege.

4.3 Anwendungen

4.3.1 Produkt versus Ingenieurleistung

Es wurde schon angedeutet, daß die digitale Bearbeitung von Pixelmatrizen ein Stück Technologie mit dem Charakter einer Systemkomponente darstellt. Anwendungen entstehen zwar derzeit in rascher Folge, aber ein wirtschaftlich relevanter Industriezweig ist damit nicht begründet. Es bleibt überhaupt die Frage, ob das "maschinelle Sehen" nicht Inhalt von Leistungen beratender Ingenieure (Zivil-Ingenieure) bleiben wird, statt sich als "Produkt" auszuprägen. Auf jeden Fall gilt es aber, die Möglichkeiten und Komponenten des automatischen Sehens in der traditionellen Methode der technischen Handbücher zu dokumentieren.

4.3.2 Dominierende Industrieanwendung

Die breite Palette der Anwendungen wurde schon im Abschnitt 3 charakterisiert. Innerhalb der heutigen Industrie-Anwendungen dominieren die folgenden vier Bereiche: (a) Platinenprüfung in der Elektronik; (b) pharmazeutische Prüfung von Pillen; (c) Robotersichtsysteme in der Fertigung; und (d) Flaschenprüfung in der Getränkeindustrie.

An sich ist die Fokussierung auf diese Industriezweige vor allem durch die offensichtlich einfach erkennliche Verbesserung des Warenflusses begründet. Aber andere Anwendungen werden folgen, wo immer ein menschlicher Operateur eine Maschine oder gefertigte Teile mit seinem Sehsinn zu prüfen hat. Es bestehen keine besonderen Hindernisse, die sich einer totalen Durchdringung aller Sichtungsaufgaben mit maschinellen Methoden in die Quere setzen.

4.4 Bildverarbeitung und Multi-Medien

Wollte man die breitere Palette von maschinellen Pixelaufgaben prüfen und die Beschränkung auf industrielle Fertigung fallen lassen, so wird rasch ersichtlich, daß die bisherige Trennung in Bildbearbeitung und Computergrafik zunehmend unsinnig wird. Wir tragen dieser Erwartung an der Technischen Universität Graz Rechnung, indem diese beiden Themen in ein einziges Lehrgebiet verschmolzen wurden. Bilder sind die Quelle jener Informationen, welche oft in der Computergrafik zur Bilderzeugung verwendet werden: die Computergrafik verwendet ja Objektmodelle, die irgendwie in den Computer eingespeist werden müssen. Bei bestehenden Objekten sind dies gerade aus natürlichen Bildern gewonnene Inhalte, und zwar sowohl geometrischer als auch radiometrischer Art (Texturen).

Die Unterhaltungselektronik wird zunehmend eine Triebfeder für Innovationen in der Bildverarbeitung und Computergrafik. Multimediale Systeme benötigen Inhalte, welche aus Bildern gewonnen und mittels Computergrafik vermittelt werden. Dabei betrachten wir die Multimedia-Systeme als Anwendung des maschinellen Sehens, allerdings in Verbindung mit der Computergrafik.

5 EINE ÖSTERREICHISCHE PERSPEKTIVE

5.1 Institutionen

Zum maschinellen Sehen werden in Österreich im wesentlichen an 4 Institutionen in größerem Umfang Innovationen entwickelt. Dies sind an der Technischen Universität Wien die Abteilung für Mustererkennung (Prof. Kropatsch) im Rahmen der Informatik; an der Technischen Universität Graz das Institut der Autoren dieses Beitrages, ebenfalls in der Informatik; am Forschungszentrum Seibersdorf die Gruppen im Engineering (Ing. Wallisch) und in der Informationsverarbeitung (Dipl.-Ing.

Schoitsch); und am Joanneum Research das Institut für Digitale Bildverarbeitung (Dr. Pölzleitner). Einzelprojekte bestehen auch an anderen akademischen Einrichtungen, etwa am Research Institute for Symbolic Computing RISC der Universität Linz, oder im Rahmen der geodätischen Forschung an den Technischen Universitäten Wien und Graz.

Der Kreis ist etwas größer, wenn Innovationen aus Anwendungen in der Fernerkundung und Erdbeobachtung hinzugezählt werden. Dann werden weitere Gruppen betroffen, welche in einem Projektteam "Fernerkundung" des österreichischen Bundesministeriums für Wissenschaft, Forschung und Kunst und unter dem Vorsitz des ersten Autors dieses Beitrages organisiert sind. Auch aus medizinischen Anwendungen ergeben sich österreichische Innovationen des maschinellen Sehens, wobei aber die schon genannten Gruppen betroffen sein dürften.

5.2 Industrie/Gewerbe

In der Industrie besteht in Österreich derzeit noch keine Orientierung auf diese Technologie, wie es etwa in Deutschland der Fall zu sein scheint. Dort besteht ein "Vision Club" von etwa 20 Firmen, welche sich als die wesentlichen Unternehmen mit Bildverarbeitungswissen verstehen. Weitere 150 Firmen sollen bestehen, welche ebenfalls mit Bildverarbeitung ihr Fortkommen suchen, allerdings meist als Betriebe mit weniger als 5 Mitarbeitern.

Damit ist deutlich, daß es in Österreich keine einschlägige Industrie gibt und sich aber auch im nahen Ausland kaum eine Industrie gebildet hat. Wohl besteht dort ein reges Kommen und Gehen im Bereich der kleinen Software-Häuser im Sinne von Ingenieurbüros, ein Phänomen, welches sich in Österreich noch nicht manifestiert hat.

5.3 Forschung und Entwicklung

Die Forschung in Österreich hat im wesentlichen 3 Finanzierungsquellen:

den Fonds zur Förderung der wissenschaftlichen Forschung (FWF), um die akademische Forschung zu befruchten;

den Forschungsförderungsfonds (FFF) zur Finanzierung von industriellen und gewerblichen Projekten, welche auf Produktentwicklung abzielen;

Auftragsforschung der Ministerien, vor allem des BM für Wissenschaft, Forschung und Kunst.

Derzeit läuft beim FWF ein Forschungsschwerpunkt "Digitale Bildverarbeitung" mit insgesamt 12 Teilaufgaben, von denen einige die industrielle Anwendung berühren. Dieses Projekt begann im Jahre 1990 mit einer kleineren Vorstudie und ist nun seit 1994 in einem 5-jährig konzipierten Unterfangen tätig (bis 1999).

Das Thema "Erdbeobachtung" ist als derzeitiger Leitschwerpunkt der ministeriellen Auftragsforschung definiert und sollte daher in der laufenden Legislaturperiode zu weiteren Forschungsarbeiten im Bereich der digitalen Bildanalyse führen.

Des weiteren liefert die österreichische Initiative zu einem Multimedia-System Hyper-G eine Gelegenheit und die Verpflichtung, die digitale Bildbearbeitung in diesen Tätigkeitskreis einzuführen. Erste Ansätze wurden mit dem schon genannten Projekt Kaiser Karl und ähnlichen Aufgaben geliefert. Gerade dort besteht aus der österreichischen kulturellen Tradition die Gelegenheit, einen internationalen Beitrag zu liefern und eine Nische zu besetzen.

6 SCHLUSSFOLGERUNGEN

Wir haben die derzeitige Lage der digitalen Bildverarbeitung und des maschinellen Sehens dargestellt und zwar sowohl die internationalen Trends, als auch die spezielle österreichische Entwicklung.

Daraus haben wir unsere Ansicht begründet, daß diese Technologie in Zukunft eine weite Verbreitung erfahren wird und zwar in den schon heute absehbaren Anwendungen der Industrie, der Medizin, der Erdbeobachtung, aber auch in neueren Themen, etwa in multimedialen Systemen und in der Unterhaltungselektronik.

Durch die heute bereits verfügbaren Hochleistungs-Prozessoren haben sich die wesentlichen Neuerungen von der Hardware in die Software und in die Methoden verlagert. Dieser Trend wird sich in Zukunft verstärkt fortsetzen, etwa bei der Integration verschiedener Informationsquellen (Sensoren) oder bei der Entwicklung aktiver Inspektionssysteme.

DANKSAGUNG

Die hier vorgestellten eigenen Arbeiten unseres Institutes wurden durch nachfolgend genannte Projekte unterstützt:

Fonds zur Förderung der wissenschaftlichen Forschung FWF, Projekte S7001-MAT und S7003-MAT,

Jubiläumsfonds der Österreichischen Nationalbank, Projekt Nr. 4871,

Bundesministerium für Wissenschaft, Forschung und Kunst, GZ 601.560/2-II/6/94 und GZ 76.008/2-II/5/94,

Österreichische Akademie der Wissenschaften, GZ 6070/3-II/5/94.

LITERATUR

- Ganster H., Gelautz M., Pinz A., Binder M., Pehamberger H., Bammer M., Krocza J. (1965)
Initial Results of Automated Melanoma Recognition, in: Proc. 9th SCIA, Uppsala, pp.209-218
- Gonzalez R.C., und Woods R.E. (1992)
Digital Image Processing, Addison-Wesley
- Haralick R.M., Shapiro L.G. (1992)
Computer and Robot Vision, Volume I, Addison-Wesley
- Haralick R.M., Shapiro L.G. (1993)
Computer and Robot Vision, Volume II, Addison-Wesley
- Hoffer R. (1967)
Interpretation of Remotely Sensed Imagery of Agricultural Crops. Research Bulletin Nr. 1, Laboratory for Agricultural Remote Sensing, Purdue Univ., West Lafayette, Indiana
- Houndsfield G.N. (1973)
Computerised Transverse Axial Scanning (Tomography), Part 1: Description of Systems, Brit.J.Radiol. 46, pp.1016-1022

- Leberl F.W., Maurice K., Thomas J.K., Millot M. (1994)
Automated radar image matching experiment, ISPRS Journal of Photogrammetry and Remote Sensing 49(3), Elsevier
- Leberl F., Gruber M., Uray P., Madritsch F. (1994)
Trade-offs in the reconstruction and rendering of 3-D objects, in: Mustererkennung 1994, Informatik Xpress, pp.58-73
- Marr D. (1992)
Vision, Freeman
- Nathan R. (1966)
Digital Video Handling Report, Technical Report Nr. 32877, Jet Propulsion Laboratory, Pasadena, California
- Nischelwitzer A., Pinz A., Gruber M. (in Vorbereitung)
CERES - Casting Error Recognition and Evaluation System
- Pinz A. (1994)
Bildverstehen, Springers Lehrbücher der Informatik, Springer
- Pinz A., Prantl M., Datlinger P. (1995)
Mapping the Human Retina, in: Proc. 9th SCIA, Uppsala, pp.189-198
- Rindfleisch T.C., et al. (1971)
Digital Processing of Mariner 6 and 7 Pictures, J.Geophys.Res., Vol.76, pp.394-417
- Scherer S., Stampfl J., u.a. (1995)
Automatische Oberflächenrekonstruktion mikroskopischer Objekte, Angenommen zur MicroMat 1995, Berlin
- Uray P., Pinz A. (1995)
Semiautomatic triangulation of irregular 3D point data, in: Proc. 19th ÖAGM and 1st SDRV Workshop, Maribor, in print
- Wiltschi K., Pinz A. (1995)
Feature extraction for carbide classification of high speed steel, in: Proc. 19th ÖAGM and 1st SDRV Workshop, Maribor, in print