Automatische Erkennung von dreidimensionalen Objekten

Autor(en): **Bunke, H. / Gmür, E.**

Objekttyp: **Article**

Zeitschrift: **Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association Suisse des Electriciens, de l'Association des Entreprises électriques suisses**

Band (Jahr): **79 (1988)**

Heft 15

PDF erstellt am: **24.05.2024**

Persistenter Link: https://doi.org/10.5169/seals-904057

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

Automatische Erkennung von dreidimensionalen Objekten

H. Bunke und E. Gmür

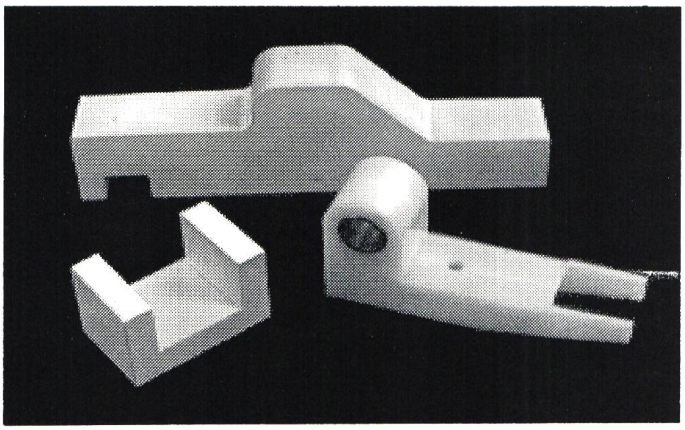
Dieser Artikel skizziert die grundsätzliche Problematik, mögliche Lösungswege sowie die wichtigsten Anwendungen der Disziplin der automatischen Bildanalyse mit Hilfe von Computern. Als konkretes Beispiel für ein Bildanalysesystem wird das am Institut für Informatik und angewandte Mathematik der Universität Bern entwickelte Sichtsystem Phi-1 für Roboter vorgestellt.

Le présent article esquisse la problématique fondamentale qui est d'étudier des voies de solutions possibles ainsi que les applications les plus importantes de la discipline de l'analyse automatique des images, à l'aide d'ordinateurs. Par un exemple concret pour un système d'analyse des images, on présente le système de vision pour robots développé à l'Institut d'informatique et de mathématiques appliquées de l'Université de Berne.

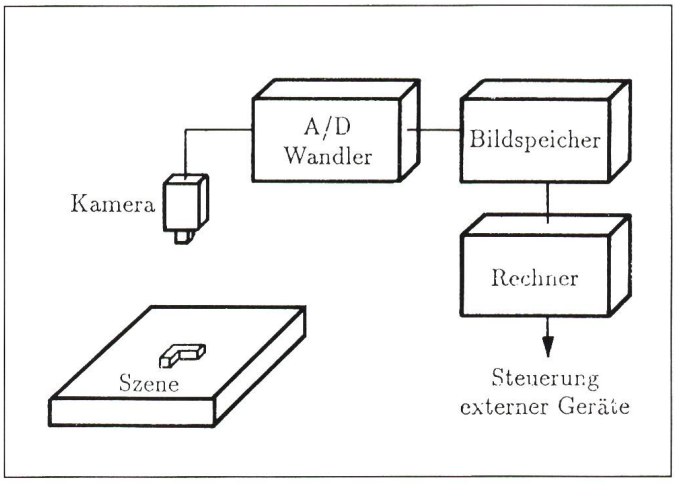
Adresse der Autoren
Prof. Dr.-Ing. *Horst Bunke* und dipl. Phys. *Edgar Gmür*, Institut für Informatik und angewandte Mathematik der Universität Bern, Länggass-Strasse 51, 3012 Bern.

Bildanalyse, auch als Bildinterpretation oder Bildverstehen bezeichnet, befasst sich mit der Aufgabe, mit Hilfe von Computerprogrammen aus einem vorgegebenen Bild automatisch seinen Inhalt bzw. seine Bedeutung zu erschliessen. Eine einheitliche Theorie der Bildanalyse ist heute (noch) nicht verfügbar, jedoch ist charakteristisch, dass die angewendeten Verfahren auf der Manipulation nichtnumerischer symbolischer Datenstrukturen beruhen. In diesem Artikel wird nach einer allgemeinen Einführung und einem Überblick über die Bildanalyse ein am Institut für Informatik und angewandte Mathematik der Universität Bern entwickeltes System zur Erkennung von Objekten in Bildern auf der Basis von CAD-Modellen (CAD, Computer-Aided Design) vorgestellt.

Die enorme Leistungsfähigkeit des menschlichen Gesichtssinns und die Tatsache, dass visuelle Wahrnehmung fast ausschliesslich im Unterbewusstsein abläuft, täuschen häufig über die grossen Schwierigkeiten hinweg, die grundsätzlich bei der automatischen Bildanalyse mit Hilfe des Computers existieren. Die Probleme werden aber offenbar, wenn wir uns vor Augen führen, wie Bilder im Computer dargestellt, d.h. abgespeichert, werden. In Figur 1 ist eine Szene mit verschiedenen Objekten gezeigt. Diese Szene wurde mit Hilfe einer Fernsehkamera aufgenommen, die über ein spezielles Interface (inklusive A/D-Wandler) mit einem Computer verbunden ist. Das Aufnahmeprinzip entspricht dem Blockbild in Figur 2. Die Szene von Figur 1 wird im Computer als zweidimensionales Feld von 512×512 Zahlenwerten, auch *Bildpunkte* oder *Pixel* genannt, abgespeichert. Jeder Wert ist ganzzahlig und liegt zwischen 0 und 255, was einer Grauwertdarstellung durch 8 Bit entspricht. Helle Punkte der Szene werden durch grosse Zahlenwerte dargestellt, während Werte nahe bei Null durch dunkle Punkte repräsentiert sind. Ein Vergleich von Figur 1 mit Figur 3 zeigt unmittelbar die Problematik der Bildanalyse mit Hilfe des Computers. Während ein menschlicher Betrachter spontan die einzelnen Objekte in Figur 1 erkennen und ihnen eine Bedeutung zuordnen kann, «sieht» der Computer bei «Betrachtung» der Szene nur eine zweidimen-



Figur 1 Eine Beispiel-Szene mit verschiedenen dreidimensionalen Objekten

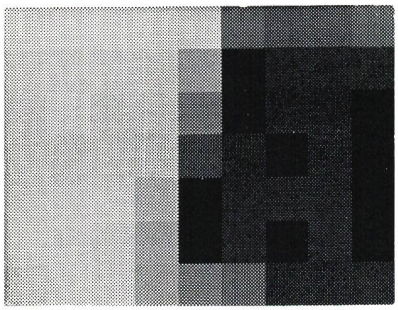


Figur 2 Aufnahme und Digitalisierung von Bildern

sionale Anordnung von Zahlen entsprechend Figur 3b. Es ist unmittelbar klar, dass es keine einfache Aufgabe sein kann, Verfahren, also Computerprogramme, anzugeben, die aus einer Darstellung, wie in Figur 3 gegeben, auf die Bedeutung des Bildinhalts schliessen können. Ein durchschnittlicher menschlicher Betrachter wäre jedenfalls völlig überfordert.

Das Gebiet der Bildanalyse mit dem Computer ist nun schon seit über 30 Jahren Gegenstand zahlreicher Forschungs- und Entwicklungsarbeiten, und so verwundert es nicht, dass trotz der offensichtlich schwierigen Aufgabenstellung eine Reihe von Fortschritten erzielt wurde. Die wichtigsten Anwendungsgebiete mit kommerziell angebotenen Geräten sind heute die Erkennung von Schriftzeichen, die optische Qualitätskontrolle, Sichtsysteme für Industrieroboter, die Erdfernerkundung sowie medizinische Bildanalyse.

a



b

144	148	144	148	135	33	30	31	34
140	141	141	142	103	6	10	9	11
141	144	139	136	56	6	10	8	5
141	141	139	130	19	14	7	8	7
142	139	138	112	2	11	8	8	7
139	142	134	96	7	13	7	9	7
134	138	130	89	52	38	13	13	9

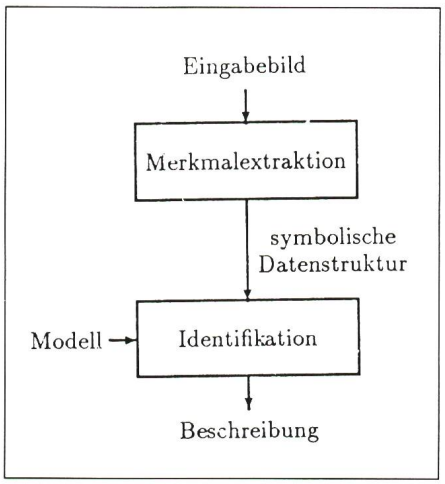
Figur 3 Ausschnitt aus Szene der Figur 1
a Grauwertdarstellung
b rechnerinterne Darstellung

Grundsätzliche Prinzipien und Lösungswege

Obwohl die in der obigen Aufzählung erwähnten Anwendungsgebiete sehr unterschiedlich erscheinen, existieren dennoch eine Reihe grundlegender Prinzipien und Verarbeitungsverfahren, die nahezu unverändert in den verschiedenen Anwendungen verwendet werden können. Typischerweise ist ein Bildanalysesystem nach dem Schema von Figur 4 aufgebaut.

Als *Eingabebilder* finden, je nach Anwendung, neben Grauwertbildern (Fig. 3) eine Vielzahl anderer Sensordaten Verwendung. So wird z.B. im industriellen Bereich auch mit Röntgenbildern gearbeitet, wenn es um die zerstörungsfreie Prüfung von Gussteilen auf das Vorhandensein von Einschlüssen oder anderen Unregelmässigkeiten im Innern des Materials geht. Ein aktuelles Gebiet ist die Entwicklung und Anwendung von Sensoren, welche die Distanz zwischen Beobachter, d.h. Sensor, und Punkten auf der Oberfläche der zu analysierenden Objekte messen und aufzeichnen. In der Erdfernerkundung liegen häufig sog. Multispektralaufnahmen vor; eine solche besteht aus 3 bis 12 Einzelbildern desselben Gebiets, wobei ein Einzelbild die Strahlungsintensität in einem bestimmten Spektralbereich widerspiegelt. Auch in der Medizin finden heute eine grosse Zahl bildgebender Verfahren Verwendung, und es existieren somit eine Fülle verschiedener Eingabebilder für Bildanalysesysteme, z.B. Röntgenbilder, Szintigramme, Ultraschallbilder, Röntgentomogramme, NMR-Aufnahmen usw.

In der Phase der *Merkmalextraktion* geht es darum, signifikante Merkmale der zu erkennenden Objekte in den Eingabebildern mit Hilfe geeigneter Methoden, d.h. Computerprogrammen, zu lokalisieren und zu identifizieren. Die Merkmale müssen hierbei so



Figur 4 Typischer Aufbau eines Bildanalysesystems

gewählt werden, dass es in der anschliessenden Phase der Identifikation möglich ist, ein Objekt eindeutig zu erkennen und von möglichen anderen Objekten zu unterscheiden. Sehr häufig verwendete Merkmale sind die Konturlinien der zu erkennenden Objekte, d.h. die Grenzen zwischen dem betrachteten Objekt und dem Bildhintergrund oder eventuell angrenzenden oder teilweise verdeckten anderen Objekten. Andere häufig verwendete Merkmale sind charakteristische Eigenschaften von Oberflächen von Objekten wie etwa Grösse, Umfang oder mittlerer Grauwert einer Fläche.

Die während der Merkmalextraktion erhaltenen Zwischenergebnisse werden in einer symbolischen Datenstruktur gespeichert, auf welche in der anschliessenden Phase der Identifikation der Objekte zugegriffen wird. Die in dieser Datenstruktur enthaltene In-

formation charakterisiert z.B. die im Bild gefundenen Konturlinien mit ihren gegenseitigen Relationen wie z.B. relative Winkel zwischen Geraden oder eventuelle Schnitt- oder Berührungspunkte zweier Linien. Ein konkretes Beispiel für eine symbolische Datenstruktur folgt bei der Beschreibung des Systems zur Erkennung von Objekten auf der Basis von CAD-Modellen. Um Objekte in Bildern mit Hilfe eines Computers erkennen zu können, ist es unumgänglich, Wissen über die charakteristischen Merkmale der zu identifizierenden Objekte zu speichern. Die hierzu verwendeten Datenstrukturen werden meistens als *Modell* bezeichnet. Abhängig von der konkreten Aufgabenstellung kann das als Modell verwendete Wissen sehr unterschiedlicher Natur sein und von einfachen Merkmallisten über geometrische Modelle bis hin zu symbolischen Strukturen reichen, die auch in der Künstlichen Intelligenz Verwendung finden. Die Aufgabe während der Phase der *Identifikation* besteht nun darin, eine Zuordnung zu treffen zwischen den in der symbolischen Datenstruktur enthaltenen Merkmalen, die aus dem aktuellen Eingabebild abgeleitet wurden, und den im Modell gespeicherten Objekteigenschaften. Ein Objekt gilt dann als erkannt, wenn seine im Modell angegebenen charakteristischen Eigenschaften hinreichend gut mit den in der symbolischen Datenstruktur enthaltenen Merkmalen übereinstimmen.

Als Ergebnis der Bildanalyse liegt eine Beschreibung vor, die typischerweise aus einer Liste der im Bild enthaltenen Objekte besteht. Ein einzelnes Objekt ist z.B. charakterisiert durch Name, Lage, Vergrösserungs- bzw. Verkleinerungsfaktor, Rotationswinkel sowie eine Liste von Abweichungen bezüglich eines Objektprototyps.

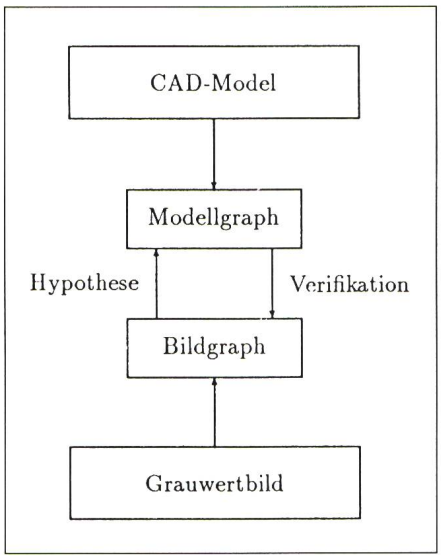
Das Robotersichtsystem Phi-1

Zur Konkretisierung der bisherigen allgemeinen Ausführungen soll in diesem und den folgenden Abschnitten das am Institut für Informatik und angewandte Mathematik der Universität Bern entwickelte Robotersichtsystem Phi-1 genauer dargestellt werden. Das Ziel des Systems liegt in der Identifikation und Bestimmung der Lage und Orientierung dreidimensionaler Werkstücke in einem Grauwertbild. Im jetzigen System Phi-1 können hellere, nicht zu stark glänzende, massive Objekte erkannt werden, deren Oberfläche durch ebene und zylindrische Flächen beschrieben werden kann. Das System ist ein Forschungs-Prototyp, mit dem neue Verfahren zur 3D-Objekterkennung und Merkmalmodellierung untersucht werden sollen. In einer industriellen Anwendung würde Phi-1 mit einem Roboter kombiniert, welcher in Figur 2 in den externen Geräten enthalten ist. Phi-1 übernimmt die Szenenanalyse und gibt deren Ergebnis an den Roboter weiter. Mit Hilfe der interpretierten Szene werden die weiteren Handhabungsschritte geplant.

Die Verwendung von CAD-Modellen als Modelle für die Identifikation entsprechend der Figur 4 wäre ideal, da CAD-Modelle der zu erkennenden Objekte im industriellen Fertigungsprozess meist ohnehin verfügbar sind und somit eine erneute Erfassung der Geometrie entfallen könnte. Leider sind die internen Darstellungen der geometrischen Modellierer in CAD-Systemen für die Objekterkennung nicht geeignet, da sie vor allem den interaktiven Modellierungsprozess und die graphische Darstellung an Ausgabegeräten wie Bildschirm und Plotter, nicht aber die Bildanalyse unterstützen. Gesucht ist eine Modellrepräsentation von Werkstücken, die vor allem diejenigen Merkmale enthält, welche in der Phase der Merkmalextraktion (Fig. 4) in einem Eingabebild lokalisiert werden können. Da ein 3D-Objekt im Raum drei Translations- und drei Rotationsfreiheitsgrade besitzt, sollte die Modelldarstellung in erster Linie lage- und orientierungsunabhängig sein. Beim Phi-1 wird das CAD-Modell in einem Off-line-Schritt in ein Merkmalmodell umgewandelt.

Zur Repräsentation des Modells wird eine attributierte Graphenstruktur verwendet. Wir sprechen hier von *Modellgraphen*. Die Ecken und Kanten des Graphen entsprechen den Ecken und Kanten der modellierten Objekte. Damit ein Vergleich der Modell- und Bildstruktur durchgeführt werden kann, muss für beide der gleiche Repräsentationsformalismus verwendet werden. Deshalb werden die extrahierten Merkmale des Bildes ebenfalls in einen attributierten Graphen, den sogenannten *Bildgraphen*, umgewandelt. Daraus ergibt sich die in Figur 5 dargestellte Architektur des Systems Phi-1.

Phi-1 wurde auf einer Prime-750 in Pascal implementiert. Die Quellenprogramme umfassen heute etwa 8000 Zeilen. Obwohl Phi-1 heute bereits einfachere und mittelkomplexe Teile identifizieren und lokalisieren kann, ist die Forschungsarbeit bei weitem nicht abgeschlossen. Zur Erhöhung der Flexibilität und der Fehlertoleranz werden gegenwärtig die einzelnen Verfahren noch verbessert. Für einen industriellen Einsatz in Echtzeit müssten die Algorithmen etwa um einen Faktor hundert beschleunigt werden. Die Lösung liegt hier in parallelen Rechnerarchitekturen, die zurzeit einen grossen Aufschwung erleben. Die folgenden Abschnitte stellen die einzelnen Teile des Systems genauer vor.



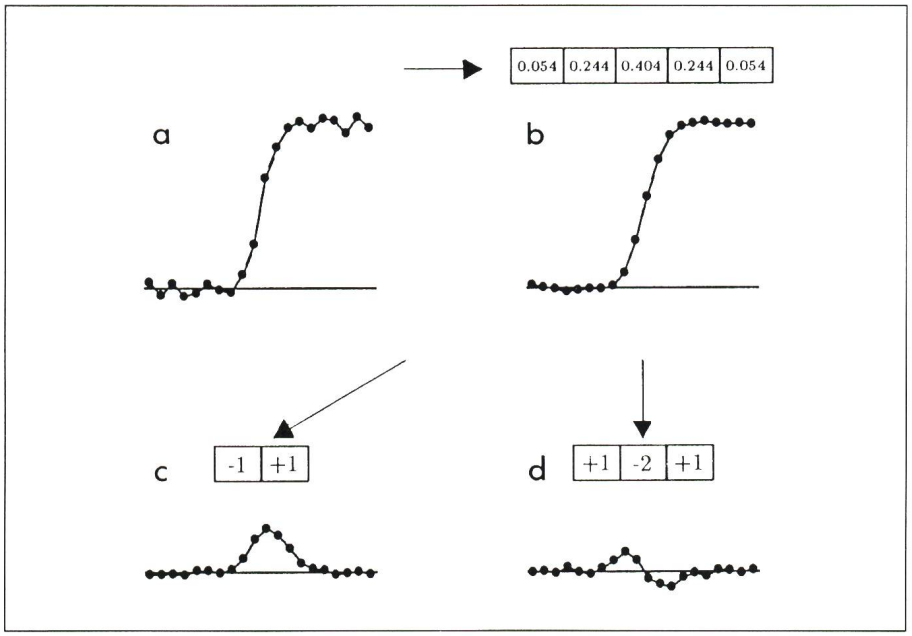
Figur 5 Systemarchitektur des PHI-1
Die Objekterkennung (Identifikation) ist in zwei Phasen aufgeteilt. In der ersten werden Hypothesen generiert, in der zweiten Phase werden diese verifiziert.

Extraktion der Bildmerkmale beim Phi-1

Die beim Phi-1 verwendeten Merkmale sind die Kanten und Ecken der zu erkennenden Objekte. Die Gewinnung der Merkmale aus dem Grauwertbild lässt sich in drei Hauptschritte aufteilen:

1. Bildverbesserung und Kantendetektion,
2. Segmentierung der Kantenzüge,
3. Vervollständigung der Linienzeichnung durch Verbinden der einzelnen Kantenstücke.

Der erste Schritt lässt sich am besten durch eine Betrachtung des eindimensionalen Falles erklären. In Figur 6a ist das Profil einer Kante dargestellt, wobei Punkte mit niedrigen bzw. ho-



Figur 6 Kantendetektion
am Beispiel einer eindimensionalen Grauwertfunktion

a verrauschtes Intensitätsprofil

b mit Gaussoperator gemitteltes Profil

$f[i] = 0{,}054\ g[i-2] + 0{,}244\ g[i-1] + 0{,}404\ g[i] + 0{,}244\ g[i+1] + 0{,}054\ g[i+2]$ wobei $f[i]$ bzw. $g[i]$ den Wert der Intensitätsfunktion nach bzw. vor der Glattung an der Stelle i darstellen.

c erste Ableitung, Approximation durch Differenz

$\triangle_1[i] = g[i+1] - g[i]$

d zweite Ableitung, Approximation durch Differenz

$\triangle_2[i] = g[i+1] - 2\,g[i] + g[i-1]$

hen Werten hellen bzw. dunklen Bildpunkten entsprechen. Kantenpunkte werden durch abrupte Sprünge der Grauwerte definiert. Überlagert wird das Profil durch Rauschen des elektronischen Aufnahmesystems. Das Rauschen täuscht falsche Kantenpunkte vor. Deshalb müssen diese Störungen durch eine Glättung eliminiert werden.

Glättung und Kantendetektion in Grauwertbildern geschehen beim Phi-1-Projekt mit Hilfe von linearen Operatoren. Dabei werden die Pixel in einer Umgebung des aktuellen Punktes mit den Gewichten des Operators multipliziert und die Werte summiert. In Figur 6b ist ein Operator zur Gauss-Mittelung und daneben die geglättete Funktion dargestellt (s. Legende Fig. 6). Die Bestimmung der Kantenpunkte geschieht entweder durch ein- oder zweimaliges Ableiten der Intensitätsfunktion. Der Ort der Kante ist im ersten Fall durch das Maximum der ersten Ableitung und im zweiten Fall durch den Nulldurchgang der zweiten Ableitung bestimmt. In den Figuren 6c und 6d ist die Anwendung zweier Operatoren dargestellt, welche die erste bzw. zweite Ableitung durch Differenzenbildung nähern. Der Vorteil der zweiten Ableitung liegt in der einfacheren Bestimmung der Kante, da sich Nulldurchgänge mit Hilfe eines Programms leichter als Maxima lokalisieren lassen.

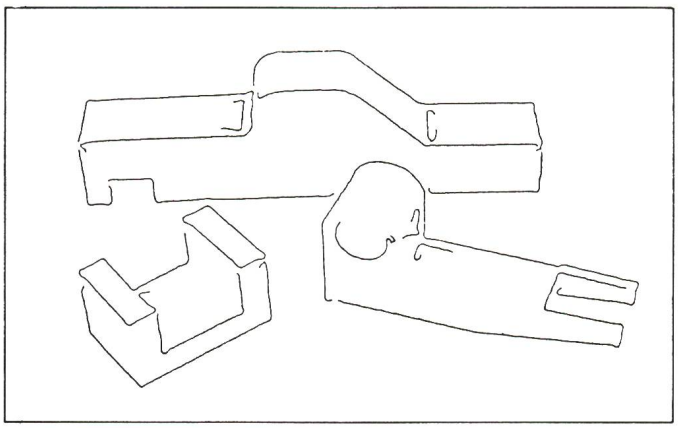
Die Kantenpunkte werden im ersten Schritt zu Kantenzügen in einer Listenstruktur (Folge von Linienkoordinaten) verbunden. Diese Folge von Punkten wird im zweiten Schritt durch gerade Strecken, Bogenstücke und Ellipsen approximiert. Im dritten Schritt werden die geometrischen Stücke zu einer möglichst vollständigen Linienzeichnung verbunden. Hier wird Wissen über die Kanten und Ecken in einer Polyederszene eingebracht. Aus der Linienzeichnung wird der sogenannte Bildgraph für die nachfolgende Objekterkennung generiert.

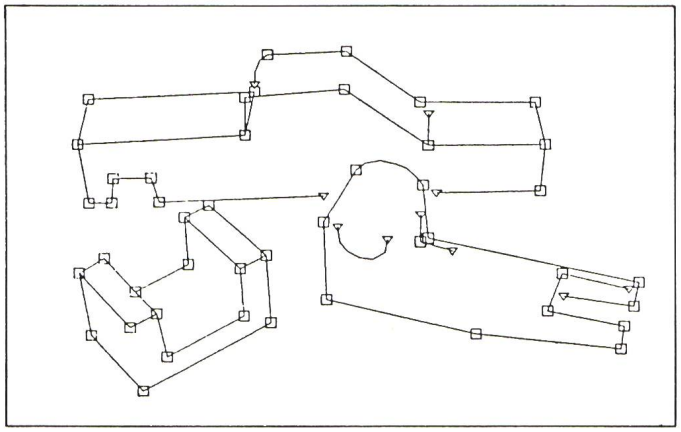
Als Beispiel betrachte man die Figur 1. Die Berechnung der 2. Ableitung mit Hilfe einer zweidimensionalen Verallgemeinerung (Matrix) des in Figur 6d gezeigten Operators sowie die anschliessende Bestimmung der Nulldurchgänge liefern die Kanten in Figur 7. Nach dem dritten Schritt liegt als Ergebnis der Merkmalextraktion der in Figur 8 gezeigte Bildgraph vor. Durch spezielle Operationen, die hier nicht näher beschrieben werden, wurde dieser Bildgraph in verschiedene Teile, die sog. *Objektgraphen*, welche den einzelnen Objekten entsprechen, zerlegt.

Modelldarstellung in Phi-1

Ein zentraler Aspekt von Phi-1 ist die Bereitstellung des Merkmalmodells aus einem CAD-Modell. Wir verwenden das kommerziell erhältliche CAD-System Prime-Medusa. Der 3D-Modellierer beruht auf einer Polyederranddarstellung, wobei ge-

Figur 7 Ergebnis der Kantendetektion





Figur 8 Der aus Figur 7 gewonnene Bildgraph

Die Knoten des Bildgraphen sind durch Rechtecke bzw. Dreiecke dargestellt. Besitzt ein Knoten nur einen Nachbarn, so ist er als Dreieck, andernfalls als Rechteck repräsentiert. Knoten des Bildgraphen treten überall dort auf, wo sich mehr als zwei Linienendpunkte treffen oder wo ein Übergang stattfindet zwischen verschiedenen Geraden, Bögen oder Ellipsen, die durch Approximation gewonnen werden. Die Geraden, Bögen und Ellipsen selbst sind durch die Kanten des Bildgraphen dargestellt.

krümmte Flächen durch ebene Facetten angenähert werden. Die Markierung der gekrümmten Kanten bleibt jedoch erhalten. Die interne Datenstruktur wird mit einem Medusa-Hilfsprogramm auf eine Datei geschrieben. Dieser File bildet die Eingabe eines Programmes zur Analyse der Geometrie und Topologie sowie zur Generierung des Modellgraphen. Dieser Algorithmus kann in vier Schritte unterteilt werden.

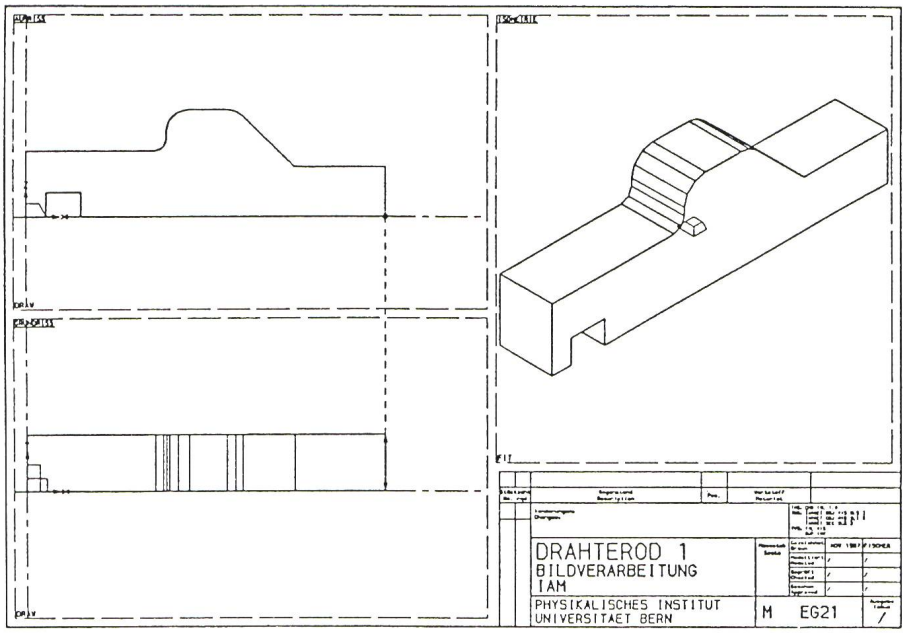
1. Bestimmung der Oberflächennormalen zur Unterstützung der folgenden Klassifizierungsschritte,
2. Klassifizierung und Markierung der Kanten in konvexe und konkave Kanten,
3. Klassifizierung und Markierung der Ecken mit Hilfe der Kantenmarkierungen,
4. Darstellung der gekrümmten Kanten als Polygonzüge.

Ein Beispiel für ein mit Prime-Medusa modelliertes Objekt ist in Figur 9 gezeigt. Die Darstellung in Figur 10 entspricht dem Zwischenergebnis nach Schritt 2. Hierbei kennzeichnen + und - konvexe bzw. konkave Kanten des Objekts. Die Oberflächennormalen sind durch Pfeile dargestellt. Nach Durchführung aller vier Schritte zur Modellanalyse wird der attributierte Graph auf eine Datei geschrieben und steht für den späteren Erkennungsprozess zur Verfügung.

Objekterkennung in Phi-1

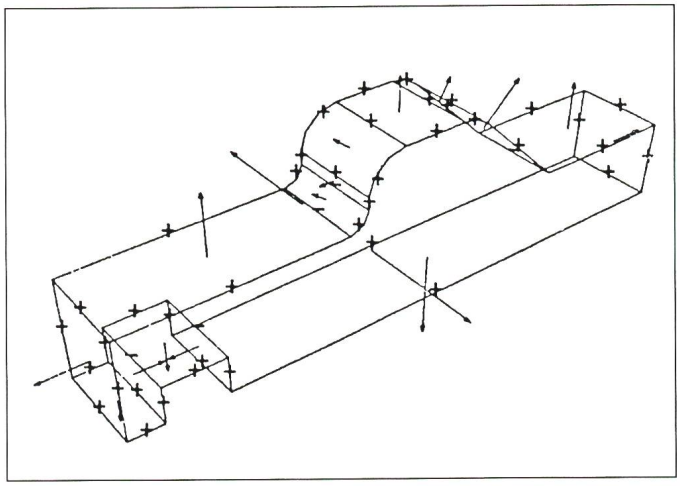
Die Aufgabe bei der Objekterkennung besteht darin, die im Bild vorhandenen Objekte, welche nach der Merkmalextraktion durch ihre Objektgraphen repräsentiert sind, zu identifizieren, indem die Objektgraphen mit den Modellgraphen aller möglichen Objekte verglichen werden. Dieser Vergleich wird in zwei Phasen aufgeteilt. In der ersten Phase werden *Hypothesen* generiert und für jeweils ein festes Objekt bestimmte mögliche Modellgraphen als Kandidaten vorselektiert. Der hierzu verwendete Algorithmus beruht auf einem strukturell-topologischen Vergleich eines Objektgraphen mit allen vorhandenen Modellgraphen. Ein Modellgraph kommt dann als Kandidat in Frage, wenn er den Bildgraphen als Teilgraphen enthält. Diese Vorgehensweise beruht darauf, dass ein Modellgraph ein Objekt immer als dreidimensionale Einheit unabhängig von einem Beobachtungspunkt darstellt, während der aus dem Bild abgeleitete Objektgraph ein Objekt nur als zweidimensionale Projektion unter einem bestimmten Blickwinkel zeigt. Das zugrunde liegende mathematische Konzept ist unter dem Begriff *Teilgraphisomorphismus* in der Informatik seit langem bekannt (siehe auch die im letzten Abschnitt dieses Artikels angegebene Literatur). Ein Beispiel ist in Fig. 11 angegeben. Man erkennt hier sofort, dass der Objektgraph aus f als Teil im Modellgraphen b, nicht jedoch im Modellgraphen d enthalten ist. Dies erlaubt, die Pyramide als möglichen Objekttyp für die in e gezeigte Projektion auszuschliessen.

Der in der ersten Phase durchge-



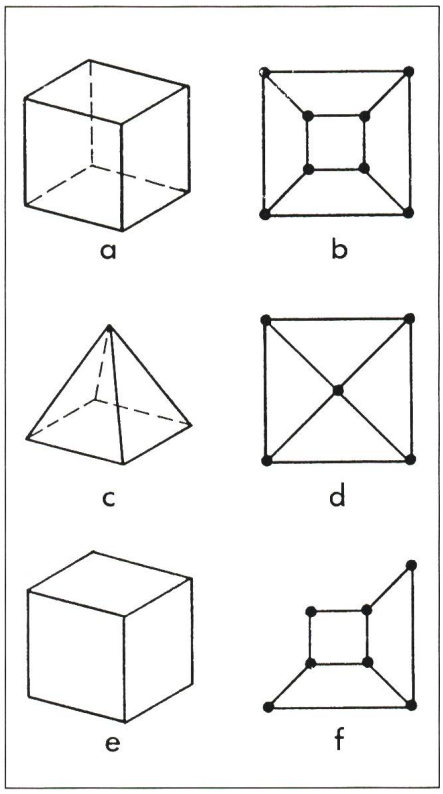
Figur 9 Ein mit Prime-Medusa modelliertes Objekt

Die modellierten Objekte werden bei der Ausgabe auf einen Plotter standardmässig in den drei gezeigten Projektionen dargestellt.



Figur 10
Ein Zwischenschritt bei der Erzeugung des Modellgraphen

→ Normalenvektor
+ - Krümmungsvorzeichen (Konvex, Konkav)



Figur 11 Modell- und Objektgraphen

Es ist nur der rein topologische Informationsgehalt dargestellt. Angaben über Längen, Krümmungen usw. sind der Übersichtlichkeit halber weggelassen worden.

a Quader als Modellobjekt
b Graphendarstellung des Quaders a (über die Länge der Kanten wird nichts ausgesagt)
c Pyramide als Modellobjekt
d Graphendarstellung der Pyramide c
e Abbild eines Quaders als zu erkennendes Objekt
f Graphendarstellung von e. Man beachte, dass nur die sichtbaren Kanten im Objektgraphen erscheinen

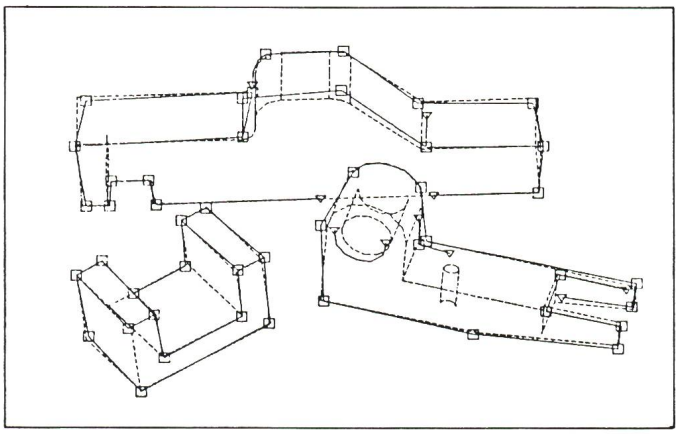
Ein Vergleich von f (Objektteilgraph) mit b und d (Modellgraphen) zeigt, dass der zu erkennende Körper keine Pyramide, wohl aber ein Quader sein kann.

führte Vergleich verwendet im wesentlichen nur strukturell-topologische Information des Modells und macht keinen Gebrauch von Attributen wie Länge oder Typ einer Kante (konvex oder konkav), Winkeln zwischen Kanten usw. Deshalb erlaubt dieser Vergleich noch keine eindeutige Objektidentifikation, sondern gestattet lediglich, eine Reihe möglicher Hypothesen aufzustellen. Die endgültige Objekterkennung erfolgt in einer zweiten Phase, der sog. *Verifikation*, bei welcher die vorselektierten Objektkandidaten der Reihe nach auf die Bildebene projiziert werden und die so entstandenen Objektkonturen mit den während der Merkmalextraktion bestimmten Kanten verglichen werden. Erweist sich eine in Phase 1 erzeugte Hypothese als nicht verifizierbar, so versucht das System, weitere Hypothesen zu generieren. Auf diese Weise entsteht der in Figur 5 gezeigte Zyklus, bestehend aus Hypothesenerzeugung und Verifikation. Die in der ersten Phase ermittelten Zuordnungen zwischen Kanten des Objekt- und Modellgraphen erlauben, die a priori sehr grosse Anzahl von Projektionen stark einzuschränken. Aufgrund von Verzerrungen bei der Bildaufnahme und anderen Störungen, z.B. Verdeckungen, wird die Projektion eines Modells in die Bildebene i.a. nicht völlig deckungsgleich mit dem entsprechenden Objekt sein. Dies macht es nötig, eine Toleranzschwelle einzuführen. Ein Objekt gilt dann als erkannt, wenn sich sein zugehöriger Objektgraph im Bild vom aktuell projizierten Modellgraphen um weniger als die vorgegebene Toleranzschwelle unterscheidet. Eine genauere Behandlung des Begriffs der Unterschiedlichkeit von Graphen findet sich ebenfalls in der angegebenen Literatur.

In Figur 12 ist das Endergebnis des Erkennungsprozesses gezeigt. Die in die Bildebene projizierten Kanten der Modelle sind gestrichelt dargestellt. Alle im Bild vorhandenen Objekte wurden korrekt identifiziert. Phi-1 wurde mit verschiedenen Objektmodellen an einer Reihe von Szenen getestet. Einfachere und mittelkomplexe Teile, wie z.B. die in Figur 1 gezeigten, werden mit grosser Zuverlässigkeit erkannt.

Ausblick

Trotz der Probleme, welche die Aufgabe der Bildanalyse mit Hilfe von



Figur 12 Ergebnis der Erkennung

——— Kanten des Bildgraphen, wie bei Figur 8 beschrieben
- - - - - - Projektion der dreidimensionalen CAD-Modelle auf die Bildebene

Computern in sich birgt, hat das Gebiet enorme Fortschritte während der letzten Jahre zu verzeichnen. Diese Fortschritte beruhen einerseits auf verbesserten und anspruchsvolleren Algorithmen und anderseits auf fortschrittlicherer Hardware. Der gegenwärtig noch ungebremste Trend zu schnellerer und billigerer Hardware sowie neuartige Computerarchitekturen bieten noch einen breiten Spielraum für weitere Anwendungen und die Verbesserung herkömmlicher Verfahren der Bildanalyse.

Bis zum Ende der 80er Jahre und darüber hinaus werden weltweit im Rahmen spezieller Forschungs- und Entwicklungsprogramme aus öffentlicher und privater Hand erhebliche Fördersummen aufgewendet, um das Gebiet der Künstlichen Intelligenz (KI) voranzutreiben. Als Teilgebiet der KI kommt auch die automatische Bildanalyse in den Genuss beträchtlicher Zuwendungen. Auch deshalb kann mit weiteren Fortschritten in der Bildanalyse gerechnet werden. Bereits im Jahr 1984 gab es eine Prognose des amerikanischen Bureau of Standards, dass innerhalb der kommenden Dekade «90% aller visuellen Inspektionsaufgaben im industriellen Bereich mit Hilfe von Bildanalysesystemen auf Computern automatisiert sein werden».

Bei der Abschätzung ihrer zukünftigen Bedeutung darf die automatische Bildanalyse nicht isoliert betrachtet werden. Vielmehr kann davon ausgegangen werden, dass zukünftige Computersysteme bei bestimmten Anwendungen Module für automatische Bildanalyse als Teilkomponenten enthalten. Konkrete Ansätze hierzu finden sich bei modernen Industrierobotern für die flexible Montage sowie beim japanischen Konzept für die 5. Generation von Computern. Erste Rechner dieses Typs können in der Mitte der 90er Jahre am Markt erwartet werden. Ein weiteres grosses Potential für die automatische Bildanalyse liegt in ihrer Kombination mit Expertensystemen, bei welchen neben der herkömmlichen Benützereingabe über die Tastatur auch graphisch-bildliche Eingabedaten direkt verarbeitet werden können.

Literatur

Einen Überblick über die Disziplin der modellgesteuerten Bildanalyse gibt [1]. Dieses Buch enthält auch die detaillierte Beschreibung eines unter der Leitung des Autors entwickelten wissensbasierten Bildanalysesystems aus dem Bereich der Medizin. Eine Behandlung der Bildanalyse unter spezieller Betonung von Methoden der Künstlichen Intelligenz findet sich in [2]. Weitere Standardwerke der Bildverarbeitung und Bildanalyse sind [3,...,6]. Einen Überblick über kommerziell verfügbare und eingesetzte Systeme enthält [7]. Das am Institut für Informatik und angewandte Mathematik der Universität Bern entwickelte Roboter-Sichtsystem Phi-1 ist detailliert in [8] beschrieben.

Literatur

[1] *H. Bunke:* Modellgesteuerte Bildanalyse. Dargestellt anhand eines Systems zur automatischen Auswertung von Sequenzszintigrammen des menschlichen Herzes. Stuttgart, Teuber-Verlag, 1985.

[2] *H. Niemann* und *H. Bunke:* Künstliche Intelligenz in Bild- und Sprachanalyse. Stuttgart, Teuber-Verlag, 1987.

[3] *A. Rosenfeld* and *A.C. Kak:* Digital picture processing. 2 volumes. Second edition. New York a.o., Academic Press, 1982.

[4] *D.H. Ballard* and *C.H. Brown:* Computer vision. Englewood Cliffs/N.J., Prentice-Hall, 1982.

[5] *M. Levine:* Vision in man and machine. New York, McGraw-Hill, 1985.

[6] *Y. Shirai:* Three-dimensional computer vision. – Symbolic Computation/Computer Graphics – Berlin/Heidelberg, Springer-Verlag, 1986.

[7] *D. Braggins* and *J. Hollingum:* The machine vision sourcebook. Berlin/Heidelberg, Springer-Verlag, 1987.

[8] *E. Gmür:* Ein Roboter Sichtsystem basierend auf CAD-Modellen. Dissertation der Universität Bern, 1988.