Zeitschrift: Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association suisse des électriciens, de l'Association des entreprises électriques suisses

Herausgeber: Schweizerischer Elektrotechnischer Verein ; Verband Schweizerischer Elektrizitätsunternehmen

Band: 77 (1986)

Heft: 7

Artikel: Automatische Analyse industrieller Bildszenen

Autor: Ott, R.

DOI: https://doi.org/10.5169/seals-904188

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 09.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Automatische Analyse industrieller Bildszenen

R. Ott

Für den Fortschritt bei der Automatisierung ist die Handhabung von Werkstücken durch Roboter von grösster Bedeutung. Dazu muss diesen die Fähigkeit zur Bilderkennung beigebracht werden. Der Aufsatz beschreibt den heutigen Stand dieser Technik und insbesondere auch den Beitrag, welche die Forschung auf dem Gebiet der künstlichen Intelligenz dabei leistet.

Pour la poursuite de l'automatisation, la manipulation de pièces par des robots a une très grande importance. Ceux-ci doivent apprendre à reconnaître les images. L'auteur décrit l'état actuel de cette technique et notamment la contribution que fournit la recherche dans la domaine de l'intelligence artificielle.

Der Aufsatz entspricht dem Fachreferat, das der Autor am 20. Technischen Presse-Colloquium der AEG am 23./24. Oktober 1985 in Berlin gehalten hat.

Adresse des Autors
Dr.-Ing. *Rainer Ott*, Abteilung Zeichen- und Signalerkennung im AEG-Forschungsinstitut, D-7900 Ulm.

Für den Menschen ist das Auge das weitaus wichtigste Organ zur Wahrnehmung seiner Umwelt. Zusammen mit dem Gehirn bildet es ein nahezu unvorstellbar leistungsfähiges und komplexes Sinnessystem, dem wir unser Bild von der Welt, in der wir leben, verdanken –, das uns aber auch bei den scheinbar einfachsten, alltäglichen Aktionen eine unerlässliche Hilfe ist. Da das menschliche Sehen nahezu unbewusst abläuft, wird die Komplexität der Vorgänge, die notwendig sind, um auch die einfachsten Szenen zu analysieren, häufig unterschätzt. Wie sehr, erkennt man spätestens dann, wenn man versucht, solche Leistungen in technische Systeme zu übertragen. Das Fehlen dieses Sehvermögens führt bei modernen Handhabungs- und Robotersystemen im industriellen Fertigungsbereich dazu, dass die Prozesse starr ablaufen müssen. Das System, das keine Gegenstände erkennen kann, greift stets blind auf vorgegebene Positionen, unabhängig davon, ob das richtige Teil an der erwarteten Stelle liegt oder nicht. Dies bedeutet ein entscheidendes Hemmnis beim Einsatz von Robotern für Montage-, Palettierungs- und Greifaufgaben sowie andere anspruchsvolle Tätigkeiten.

Bildanalysesysteme

Eine in der Fertigung immer wiederkehrende Aufgabe ist die Versorgung automatisierter Fertigungseinrichtungen mit Werkstücken und Teilen, die ungeordnet in Behältern angeliefert werden (Fig. 1). Um solche Gegenstände zu greifen und geordnet an vorgeschriebener Stelle abzulegen, muss man zunächst einmal die einzelnen Teile erkennen und feststellen, welches Teil wo liegt. Man muss die Drehlage kennen und wissen, ob der Gegenstand frei liegt oder etwa durch andere Teile eingeklemmt ist. Das erfordert eine Vielzahl von Erkennungsvorgängen; jedesmal sind grosse Mengen an bildhafter Information unter ganz bestimmten Gesichtspunkten auszuwerten und zu interpretieren. Die Aufgabe erfordert aber gleichzeitig allgemeines Wissen über Werkstücke und ihre Greifpositionen, über die Ablageeinrichtungen und manches mehr.

Technische Systeme, die derartige Seh- und Wahrnehmungsaufgaben erfüllen, waren bis vor kurzem gar nicht denkbar. Erst die Fortschritte auf dem Gebiet der Mikroelektronik haben ihre Verwirklichung in den Bereich der Möglichkeit gerückt. Zwar gibt es bereits Bildanalysesysteme, die einfachere Bildszenen automatisch bearbeiten. Immer noch Gegenstand von Forschungs- und Entwicklungsarbeiten sind Systeme, die anspruchsvollere Bildanalyseaufgaben zu lösen vermögen, wobei das Einbringen und das Auswerten von Wissen (Hintergrundinformation) über die Bildszene eine immer grössere Rolle spielt. Zu diesem Wissen gehören die Aufgabenstellung und Regeln, die angeben, wie die Ergebnisse zu verstehen und zu interpretieren sind. Dieses Wissen muss in den Computer eingebracht (Wissensakquisition) und in rechnerinterner Form dargestellt werden (Wissensrepräsentation).

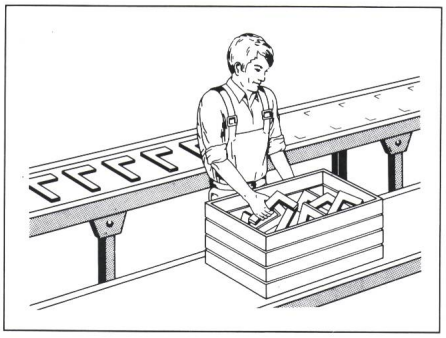


Fig. 1 Versorgung automatisierter Fertigungseinrichtungen mit Werkstücken
Ziel: Automatisierung

Begriffsdefinitionen

Abarbeitungsgraph: Der Vorgang, die beste unter den möglichen Zuordnungskonfigurationen zu suchen, lässt sich durch einen Graphen darstellen. Dabei entspricht jeder Pfad von der Wurzel bis zu einem terminalen Knoten einer der möglichen Zuordnungskonfigurationen. Der Graph enthält so viele Pfade, wie es Zuordnungskonfigurationen gibt. Gesucht ist der günstigste Pfad von der Wurzel des Graphen zu einem seiner terminalen Knoten.

Abbildungsfunktion: Siehe Operatoren.

Attributierter Graph: Graph, dessen Knoten und Kanten Attribute zur näheren Beschreibung zugeordnet sind.

Bildanalysesystem: Technisches System, das bildhafte Information aufnimmt und automatisch auswertet. Die Auswertung erfolgt auf elektronischem Wege, die Wandlung der optischen Information in elektrische Signale mit einem Bildaufnahmesystem und die Bildverarbeitung und Bildauswertung bis zur Deutung, Zählung, Vermessung und Interpretation der Bildinhalte durch Computer und geeignete Programme.

Bilddatenaufnahme: Aufnahme des Bildes in das Bildanalysesystem. Jedes Bild wird als regelmässige Anordnung von Bildelementen gespeichert. Für jede Bildelementenposition wird die Bildhelligkeit - gegebenenfalls auch der Farbwert - gemessen und als digitaler Wert abgespeichert.

Bildgrapherzeugung: Auf symbolischer Ebene stellt sich der Bildinhalt als eine Aufzählung von Bildobjekten dar, die untereinander bestimmte Relationen aufweisen. Es ist zweckmässig, sich diese Beziehungen als Bildgraphen vorzustellen. Den Knoten des Graphen sind die im Bild gefundenen Objekte zugeordnet sowie zur genaueren Beschreibung ein Satz von Parametern, die Objektattribute. Die Kanten des Graphen spiegeln die Beziehungen zwischen den Objekten wider - so etwa die einfache geometrische Beziehung «liegt innerhalb». Die Bildgrapherzeugung ist der Teil im Bildanalyseablauf, der die bildhafte Information in die symbolische Form überträgt.

Bildobjektvermessung: An einem Bildobjekt, das bereits die Segmentierung durchlaufen hat und von dem sowohl der genaue Verlauf der Berandung als auch sämtliche Helligkeitswerte im gesamten Objektinnern bekannt sind, werden Messungen durchgeführt. Einfache Messungen sind beispielsweise die Messung von Objektfläche, Objektumfang, Schwerpunktposition, Richtung der Hauptachse.

Bildsegmentierung: Zerlegung des Bildes in sinnvolle Bestandteile.

Bildverarbeitung (ikonisch, symbolisch)

Automatische Auswertung von Bildern und Bildinhalten. Wenn dabei aus Bildern wieder Bilder entstehen, spricht man von ikonischer Bildverarbeitung, werden dagegen die Objekte mit Symbolen gekennzeichnet, Objektmerkmale vermessen und den Objektnamen zugeordnet, Lagebeziehungen unter den Objekten ermittelt und Schlüsse über den Bildinhalt gezogen, so spricht man von symbolischer Bildverarbeitung.

Erkennungsergebnis: Ergebnis der Klassifikation oder Erkennung. Die Erkennungsergebnisse umfassen im allgemeinen sowohl die Klassenzugehörigkeit als auch ein Glaubwürdigkeitsmass, aus dem die Sicherheit der Erkennung abgelesen werden kann.

Grauwert: Bildhelligkeit in einem Bildelement eines monochromen Bildes. Bei Farbbildern gibt es drei Farbauszugbilder für die Grundfarben Rot, Grün und Blau. Jeder Farbauszug wird entsprechend durch einen Grauwert dargestellt.

Interpretation, Deutung einer Szene: Beschreibung einer Bildszene in einer ähnlichen Weise, wie es ein Mensch verbal tun würde. Einfaches Beispiel: Im Bild sind die Gegenstände Schraube, Schraubenzieher und Zange erkennbar. Die Gegenstände liegen auf einem Tisch.

Kandidat: Siehe Zuordnungskonfiguration.

Kanten eines Graphen: Verbindungslinie zwischen zwei Knoten des Graphen.

Klassifikation von Objekten: Erkennung von Objekten auf der Basis von Merkmalen, die unmittelbar am Objekt gemessen oder aus solchen Messgrössen abgeleitet wurden. Die Klassifikation ist immer die Zuordnung des gegebenen Objekts zu einer bestimmten Objektklasse. Um ein Objekt zu klassifizieren, muss man schon vorher wissen, was es überhaupt sein könnte. Die Klassifizierung führt im allgemeinen über Schätzwerte für die Zugehörigkeit zu den möglichen Objektklassen. Einfaches Beispiel: Ein Bildobjekt wird erkannt als «Schraube».

Konturcode: Der Verlauf der Berandung eines Objektes wird mit einem Code beschrieben. Die Elemente eines Konturcodes sind Richtungs- und Längeninformation. Es ist möglich, Konturcodes aus Geradenabschnitten konstanter oder wohldefinierter Länge aufzubauen, dann besteht der Konturcode lediglich aus einer Folge von Richtungsangaben.

Künstliche Intelligenz: Relativ junges wissenschaftliches Arbeitsgebiet, das aus der Berührung der Arbeitsgebiete Psychologie, Linguistik, Logik und Computerwissenschaften entstanden ist und das sich mit Themen wie automatisches Bild- und Sprachverstehen, mit Expertensystemen, mit Systemen, die natürlich-sprachliche Dialoge verstehen und führen können, und mit den Techniken des automatischen Beweisens beschäftigt.

Modellgraph: Graph, der das Wissen über ein Realweltobjekt enthält. Für jedes Realweltobjekt gibt es ein Modell. Wenn sich das Realweltobjekt im aufbereiteten Bild nicht als Ganzes, sondern nur in mehr als einem Bildobjekt wiederfinden lässt, dann wird ein Modell angelegt, das beschreibt, aus welchen Bestandteilen und in welcher Anordnung das Objekt aus Teilobjekten zusammengesetzt ist. Diese Beziehungen lassen sich in einem Graphen darstellen, der als Modellgraph bezeichnet wird. Die Knoten des Modellgraphen entsprechen den Modellteilobjekten, die Kanten geben die Beziehung «besteht aus» wieder.

Modellteilobjekt: Siehe Zuordnungskonfiguration.

Modellwissen: Die Kenntnis über Sachverhalte wird mit dafür geeigneten Begriffen beschrieben und mit geeigneten Massen erfasst. Die Gesamtheit dieser Daten wird als Modell bezeichnet. Beispiel: Gestalt, Eigenschaften und Abmessungen eines zu erkennenden Werkstückes sind im Modell des Werkstückes erfasst und werden zur automatischen Erkennung herangezogen.

Objektattribute: Aussagen über Objekteigenschaften, aus Messwerten abgeleitet und in einer Datenbasis abgelegt. Siehe auch Bildobjektvermessung.

Operatoren: In der automatischen Bildverarbeitung angewandte Prozeduren, die aus einem gegebenen Rasterbild ein neues erzeugen, indem sie aus den Werten einer ganzen Gruppe benachbarter Bildelemente des Eingabebildes nach einer vorgegebenen Vorschrift (Abbildungsfunktion) einen neuen Zahlenwert bilden und dem Ergebnisbildelement im Ausgabeteil zuordnen. Um das vollständige Bild zu erzeugen, muss dieser Schritt für alle Bildelemente des Ausgabebildes wiederholt werden.

Plausibilitätskriterium: Gewissermassen eine Prüfvorschrift, die die Nachprüfung erlaubt, ob eine Gruppierung von Bestandteilen insgesamt eine sinnvolle Konstellation ergibt oder nicht.

Symbole, Symbolketten: Siehe Bildverarbeitung.

Terminaler Knoten: Knoten, der keine Nachfolgeknoten besitzt.

Wissen: Unter Wissen versteht man in diesem Zusammenhang Hintergrundinformationen über die Gegebenheiten der Aufgabenstellung, wie sie auch der Mensch bei der Lösung der Bildanalyseaufgabe heranziehen würde. Dazu gehören die Bedingungen und Erfordernisse der Aufgabenstellung und Regeln darüber, in welcher Art und Weise Ergebnisse zu verstehen und zu interpretieren sind.

Zuordnungskonfiguration, Modellteilobjekt, Kandidat: Nur in simplen Fällen entspricht genau eines der durch die automatische Bildverarbeitung gebildeten Bildobjekte wiederum genau einem Realweltobjekt. Meist bildet sich das Realweltobjekt in einer ganzen Gruppierung von Bildobjekten ab, die dann nicht mehr dem vollständigen Realweltobjekt entsprechen, sondern nur Teilobjekten davon. Das Modellwissen beschreibt, aus welchen Teilobjekten und in welcher Anordnung das gesamte Objekt bestehen darf. Aufgabe der modellgesteuerten Bildanalyse ist es, die Bruchstücke, die man im Bild vorfindet, den Bestandteilen des Modells zuzuordnen. Normalerweise wird das aufbereitete Bild mehr Bildteilobjekte enthalten, als das Modell Modellteilobjekte besitzt, woraus sich sehr schnell eine unübersehbare Anzahl an möglichen Zuordnungskonfigurationen ergibt. Die Zahl der zu überprüfenden Zuordnungskonfigurationen kann drastisch eingeschränkt werden, wenn man zunächst einmal für jedes Modellteilobjekt Kandidaten einführt. Das sind solche Bildteilobjekte, die aufgrund ihrer Eigenschaften überhaupt als Kandidat für das betreffende Modellteilobjekt in Frage kommen. Zuordnungskonfigurationen zwischen Bild- und Modellteilobjekten werden nur für die jeweiligen Kandidaten aufgebaut. Eine Zuordnungskonfiguration ist nur ein Schema, das jedem Kandidaten genau ein Modellteilobjekt zuordnet. Für jede Zuordnungskonfiguration gibt es ein Bewertungsmass, das anzeigt, wie gut alles zusammenpasst.

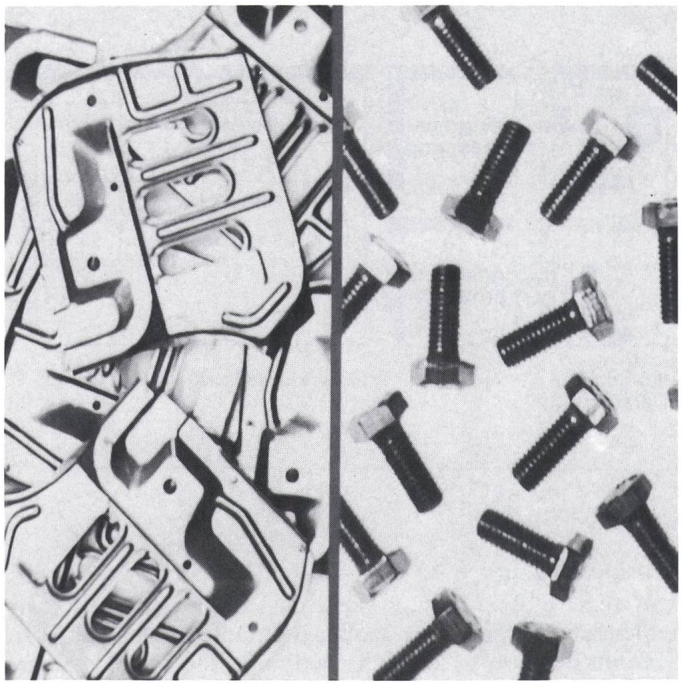


Fig. 2
Beispiele von Industrieszenen

Die folgenden Ausführungen betreffen den Entwurf von Methoden, Algorithmen und Systemkonzepten zur Realisierung automatischer Bildanalysesysteme, die in einem Produktionsprozess zur Deutung komplexer Szenen eingesetzt werden sollen.

Die Figur 2 vermittelt einen Eindruck von typischen Szenen aus der industriellen Fertigung. Es geht darum, Teile zu erkennen, Position und Drehlage zu ermitteln und Bereiche zu bestimmen, an denen die Teile etwa von einem Handhabungsroboter gegriffen werden können. Die Analyse der Szene ist einfach, wenn die zu erkennenden Teile freiliegen wie z. B. im rechten Bildteil; sie kann aber auch weitaus schwieriger sein, wie im linken Bildteil gezeigt wird.

Die *Bildverarbeitung* und die *Bildsegmentierung*, durch die das Bild aufbereitet und in sinnvolle Bestandteile (Bildobjekte) zerlegt wird, werten unmittelbar bildhafte Information aus und erzeugen selbst wieder Bilder. Diese Bild-zu-Bild-Transformation wird als *ikonische* Bildverarbeitung bezeichnet. Da zur Darstellung eines Bildes ausserordentlich grosse Datenmengen erforderlich sind, ist die Verarbeitung sehr aufwendig und wird insbesondere bei hoher Arbeitsgeschwindigkeit von speziellen Prozessoren durchgeführt.

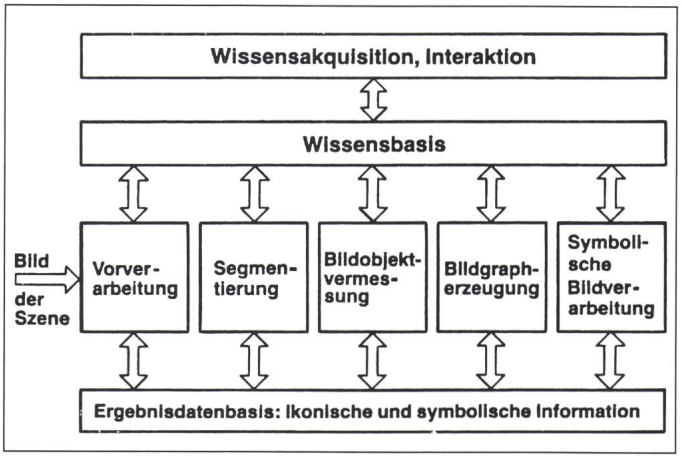
Die Teilsysteme *Bildobjektvermessung* und *Bildgrapherzeugung* vollziehen den Übergang von der ikonischen zur *symbolischen* Bildauswertung. Ergebnisse dieser Prozeduren sind Messdaten der durch die Bildsegmentierung gewonnenen Bildobjekte sowie von diesen Messwerten abgeleitete Grössen. Diese Aussagen über die Objekteigenschaften werden allgemein als *Attribut* bezeichnet. Bei der Klassifikation, dem Erkennen von Objekten, dienen sie als Merkmale. Den Objekten selbst werden Symbole zugeordnet, mit denen sie identifiziert werden können. Die Berandung der Objekte wird mit Symbolketten codiert (Randliniencodierung). Schliesslich wird die gegenseitige geometrische Lagebeziehung der Objekte untereinander erfasst. Die Gesamtheit dieser Ergebnisse kann in der Form eines *Bildgraphen* dargestellt werden. Die Knoten des Graphen sind die im Bild gefundenen Objekte (z. B. Tisch, Stuhl, Tasse), die Kanten (Verbindungslinien) des Graphen entsprechen den Beziehungen zwischen den Objekten (z. B. Tasse *ist auf* Tisch).

Die symbolische Bildverarbeitung zieht Schlüsse über den Bildinhalt aus den Objektmerkmalen und den Lagebeziehungen der Objekte untereinander; sie wertet also den Bildgraphen aus und nimmt auf diese Weise eine Interpretation und Deutung der Szene vor. Dieser Bereich der Szenendeutung ist nach dem Muster der menschlichen Vorgehensweise organisiert. Entscheidend sind hier die Kenntnisse grösserer Zusammenhänge, das Wissen über vergleichbare Szenen, das Erkennen von Ähnlichkeiten unter Bildobjekten, das Prüfen von Erkennungsergebnissen mit Hilfe anschaulicher Plausibilitätskriterien und das Ableiten logischer Schlussfolgerungen. Erst die Einbeziehung einer derart organisierten Verarbeitung macht es möglich, komplexere Szenen mit den Mitteln der automatischen Bildanalyse zu deuten.

Ablauf einer automatischen Bildanalyse

Die Figur 3 zeigt das Blockdiagramm eines Bildanalysesystems mit den wesentlichen Teilsystemen. Die Analyse beginnt mit der *Bilddatenaufnahme*, bei der die optische Information erfasst und abgespeichert wird. Ähnlich der Rasterung von Bildern beim Zeitungsdruck wird das Bild gleichmässig in eine grosse Anzahl von Bildelementen zerlegt, denen je ein Grauwert zugeordnet wird. Dieser wird in digitale Form umgesetzt und abgespeichert. Es entsteht so ein Rasterbild, das den Bildinhalt des Originals hinreichend getreu wiedergibt.

Fig. 3
Blockdiagramm eines Bildanalysesystems



Techniken der ikonischen Bildanalyse

Zweck der *Vorverarbeitung* ist es, nicht erwünschte Bildbestandteile nach Möglichkeit zu unterdrücken und solche Strukturen hervorzuheben, die für die weitere Verarbeitung von Interesse sind. Aus einem gegebenen Rasterbild wird dabei ein neues erzeugt. Hierzu wird aus den Werten einer ganzen Gruppe benachbarter Bildelemente des Eingabebildes nach einer vorgegebenen Vorschrift - der Abbildungsfunktion - ein neuer Zahlenwert berechnet und dem Ergebnisbildelement zugeordnet. Um das vollständige Ergebnisbild zu erzeugen, muss dieser Schritt für alle Bildelemente wiederholt werden. Je nach Wahl der Abbildungsfunktionen werden lineare oder nichtlineare Filter verwendet. Beispiele hierfür sind Hochpass- und Tiefpassfilterung zur Betonung von Bildfein- und Bildgrobstrukturen, Gradientenfilterung zur Hervorhebung von Bildobjektkanten, Erosion und Dilatation zur Beseitigung einzelner dunkler bzw. heller Flecken, Medianfilter zur Beseitigung von Störungen bei Erhaltung der Bildschärfe.

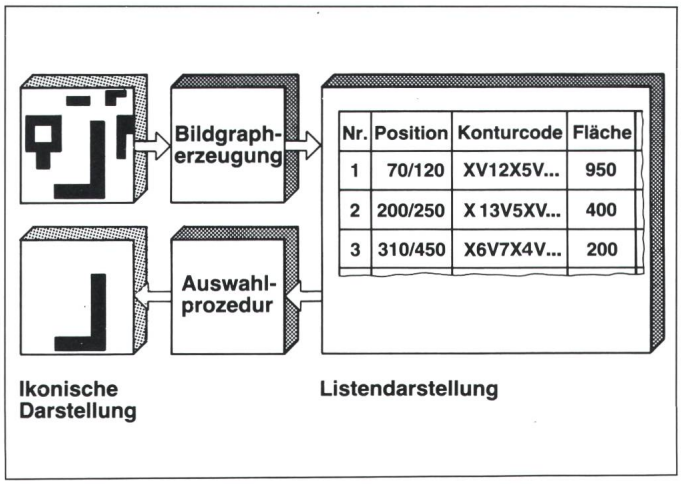
Die Figur 4 verdeutlicht die Wirkung solcher Bildfilterungen an zwei Beispielen. Das Original ist ein von einer Fernsehkamera aufgenommenes Grauwertbild. Es zeigt ungeordnet abgelegte Stanzteile. In das Bild eingeblendet sind die Ergebnisse von Filterungen, die die Kanten im Originalbild verstärken (eingeblendetes Fenster links oben) und Kanten detektieren, d.h. lokalisieren (Fenster rechts unten). Dabei werden die Strukturen hervorgehoben, die tatsächlich für die weitere Bildauswertung von Interesse sind.

Die *Segmentierung* zielt darauf ab, im aufgenommenen Kamerabild solche Gebiete, die sich mit Objekten oder Teilobjekten der realen Welt identifizieren lassen, aufzusuchen und deren Berandung zu ermitteln. Das Aussehen der Realweltobjekte ist dabei keinerlei Einschränkungen unterworfen. Sie können flach sein oder räumlich ausgedehnt. Auch ihre Lage ist völlig beliebig. Das Bild wird auf diese Weise in seine Bestandteile zerlegt. Dabei bestimmt jeweils der konkrete Anwendungsfall, welche Kriterien zur Definition solcher Zusammenhanggebiete herangezogen werden. Von Bedeutung sind besonders Verfahren, die die Gebietsberandung aufgrund von Grauwert- oder von Texturänderungen bestimmen. Diese Operation führt schliesslich für jedes Bildelement des Rasterbildes zu der binären Aussage, ob es einem bestimmten Gebiet angehört oder nicht. Häufig ergibt sich die Segmentierung nach Bildfilterung durch einfache binäre Quantisierung des gefilterten Bildes. Bei vielen Anwendungen muss beim Segmentierprozess auch ein weitergehendes Modellwissen berücksichtigt werden.

Das Ergebnis der Segmentierung ist der vollständig (mathematisch) beschriebene Verlauf aller Gebietsberandungen. Das Gebietsbild kann unmittelbar als Binärbild dargestellt werden. Die Schwarzweiss-Übergänge sind dabei die Gebietsgrenzen. Die auf den Segmentierprozess folgenden Verarbeitungsschritte *Bildobjektvermessung* und *Bildgrapherzeugung* deuten alle schwarzen oder alle weissen zusammenhängenden Gebiete als interessierende Bildobjekte. Je nach Anwendung und Bildinhalt entsprechen diese Objekte unmittelbar Realweltobjekten oder Teilen davon.

Die Figur 5 zeigt diesen Auswertungsvorgang in schematisierter Form. Es entsteht eine Tabelle, die für jedes ermittelte Bildobjekt eine Zeile enthält, in der Objektattribute eingetragen sind. Wichtige Tabelleneinträge sind etwa Position, Konturcode, einfache Merkmale wie Fläche oder Umfang und kompliziertere Merkmale

Fig. 5
Umsetzung der bildhaften Information in eine symbolische Form



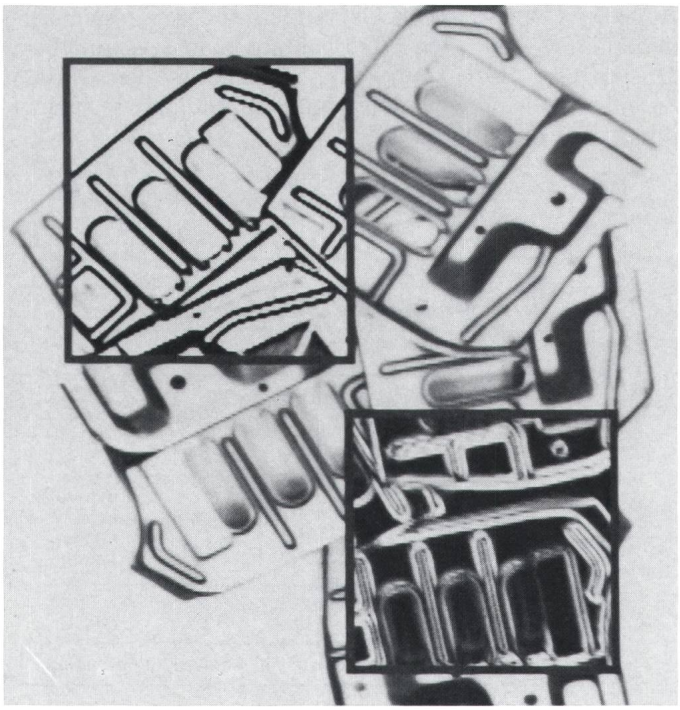


Fig. 4
Bildfilterung zur Kantenverstärkung

wie Fourierkoeffizienten[1] der Objektkontur. Ergänzt werden diese unmittelbar auf das einzelne Objekt bezogenen Daten durch geometrische Angaben über die Lagebeziehungen der Objekte untereinander und durch Daten, die die Hierarchie der Gebietsverschachtelung beschreiben. Insgesamt entsteht eine Datenbasis, die den gezielten Zugriff auf Objekte mit ganz bestimmten vorgebbaren Eigenschaften ermöglicht. Die in der Datenbasis abgelegten Parameter beschreiben die Objekte ausreichend, dass sich fehlerfreie Bilder der Objekte erzeugen lassen.

Die erfasste Information über die auszuwertende Bildszene kann als Bildgraph dargestellt werden, wobei dessen Knoten die gefundenen Bildobjekte repräsentieren. Jedem Knoten ist der Satz der ermittelten Attribute zugeordnet. Die Kanten des Graphen beschreiben die Relationen zwischen den Objekten und weisen damit darauf hin, ob z. B. ein Objekt innerhalb eines andern liegt. In den allermeisten Fällen ist die Anzahl der aus der Segmentierung hervorgehenden Bildobjekte so gross, dass ein Graph von verwirrender Komplexität entsteht.

Techniken der symbolischen Bildanalyse

Im industriellen Anwendungsbereich sind Erkennen und Vermessen von Objekten sowie besonders die Bestimmung von Drehlage und Position wichtige Aufgaben für die automatische Bildanalyse. Soweit es sich um einfach zu analysierende Objekte (mit nur einer Kontur) handelt, die tatsächlich Realweltobjekte repräsentieren, kann die Klassifikation, also das Erkennen, mit Hilfe der in der Datenbasis abgespeicherten Merkmale und Merkmalkombinationen durchgeführt werden. Die in Figur 5 dargestellte *gezielte* Auswahl von Objekten mit ganz bestimmten Eigenschaften stellt in diesem Fall den Lösungsweg der Objekterkennung dar. In einer Art Lernphase werden an einem typischen Vertreter der zu erkennenden Objektklasse charakteristische Merkmale erfasst. Anhand eines solchen Merkmalsatzes können – unter allen in der Szene gefundenen Bildobjekten – gezielt die erwünschten ausgewählt werden.

[1] Dabei wird die Kontur mit Hilfe einer komplexen Funktion beschrieben und transformiert.

Für das Erkennen von Realweltobjekten, die sich im Bild nur in einer Vielzahl von Teilbildobjekten (viele Konturen) wiederfinden lassen, wird die modellgesteuerte Bildanalyse angewendet. Als rein symbolische Bildanalyse arbeitet sie mit *Graphsuchverfahren* und somit mit Methoden, wie sie im Bereich der *künstlichen Intelligenz* verwendet werden. Ausgangspunkt dieser Art von Bildanalyse ist ein Modell des zu erkennenden Bildobjektes, etwa des Werkstückes. Unter Modell wird in diesem Zusammenhang die Gesamtheit der vom Modellentwerfer für wichtig erachteten Daten über das Realweltobjekt verstanden. Diese Daten werden in einer Art vorbereitender Lernphase durch den menschlichen Experten ausgewählt, wobei dieser Vorgang durch den Rechner unterstützt wird. Das aufgenommene und bis zur Bildgrapherzeugung aufbereitete Bild eines typischen Vertreters der zu erkennenden Objektklasse wird auf einem Monitor dargestellt. Mit dem Cursor können beliebige Bildobjekte als Teilobjekte des Modells ausgewählt und markiert werden. Ausgehend von den in der Datenbasis der Bildobjekte festgehaltenen Parametern werden automatisch die gewünschten Attribute des Modells berechnet. Wesentlich dabei ist, dass die Erstellung des Modells für ein neues Objekt auf diese Weise ausserordentlich einfach und schnell möglich ist und keinerlei Programmierarbeiten erfordert. Das Modell enthält also im wesentlichen die Liste der Teilobjekte, aus denen das Realweltobjekt besteht, sowie deren Attribute und die geometrischen Beziehungen der Teilobjekte untereinander. Es kann selbst als attributierter Graph (Modellgraph) interpretiert werden (Fig. 6 links), bei dem der Wurzelknoten das Objekt ist und die Knoten die Teilobjekte. Jedes der Teilobjekte kann wiederum aus Teilobjekten zusammengesetzt sein.

Aufgabe der modellgesteuerten Bildanalyse ist es, die Bruchstücke (Teilobjekte), die man im Bild vorfindet, den Bestandteilen des Modells zuzuordnen, also die richtige Zuordnungskonfiguration zu finden. Normalerweise wird das aufbereitete Bild mehr Bildteilobjekte enthalten, als das Modell Modellteilobjekte besitzt. Hieraus ergibt sich sehr schnell eine unübersehbare Anzahl an möglichen Zuordnungskonfigurationen. Es wird bald unmöglich, sie alle zu überprüfen, um die beste zu finden. Die Zahl der zu überprüfenden Zuordnungskonfigurationen kann einmal drastisch eingeschränkt werden, wenn man, wie in Figur 6 gezeigt wird, zunächst für jedes Modellteilobjekt (1, 2, 3) Kandidaten aus der Menge der Bildteilobjekte (A, B, C, . . . H, I) einführt. Das sind solche Bildteilobjekte, die aufgrund ihrer eigenen Eigenschaften wie Grösse, Gestalt usw. eine genügende Ähnlichkeit aufweisen, um überhaupt als Kandidaten für eine Zuordnung zum betreffenden Modellteilobjekt in Frage zu kommen. Zuordnungskonfigurationen zwischen Bild- und Modellteilobjekten werden nur für die jeweiligen Kandidaten aufgebaut.

Eine weitere Einschränkung der Zahl der zu überprüfenden Zuordnungskonfigurationen wird durch die spezielle Bearbeitungstechnik erreicht. Ihr wesentlicher Kern liegt darin, das Schema der Zuordnungen von Kandidat zu Modellteilobjekt schrittweise gezielt aufzubauen. Der schrittweise Aufbau einer Zuordnungsliste wird sofort unterbrochen, wenn das für diese Konfiguration berechnete *Bewertungsmass* schlechter ist als die aktuelle Bewertung der besten konkurrierenden Zuordnung. Den Zuschlag für die weitere Bearbeitung und damit die Fortsetzung des Listenaufbaus erhält in jedem Schritt die am besten bewertete Konfiguration. Bei grösserer Anzahl von Teilobjekten lässt sich mit einer solchen Probierstrategie das Ziel der Analyse, die Suche nach der optimalen Übereinstimmung zwischen den gefundenen Teilobjekten und den Teilobjekten eines vorgegebenen Modells, überhaupt erst erreichen; der Suchvorgang scheitert nicht an der kombinatorischen Explosion der Anzahl der zu untersuchenden Konfigurationen. Das für die Steuerung des Analyseablaufes herangezogene Bewertungsmass ist dabei natürlich von grosser Bedeutung. Es orientiert sich am Mass der Übereinstimmung zwischen den Merkmalen der Teilobjekte des Modells und denen der Kandidaten. Ganz entscheidend ist die Einbeziehung der Lagerelation der Teilobjekte untereinander. Das Ergebnis dieser Analyse ist eine sinnvolle Interpretation der Szene. Dabei werden weitestgehend offensichtliche Erkennungsfehler und Fehldeutungen ausgeschlossen.

Die Figur 6 zeigt anschaulich die Vorgehensweise der modellgesteuerten Bildanalyse mit Hilfe von Graphen. Aus dem Modellgraph und den jeweils aus den Bildteilobjekten ermittelten Kandidaten für die Modellteilobjekte

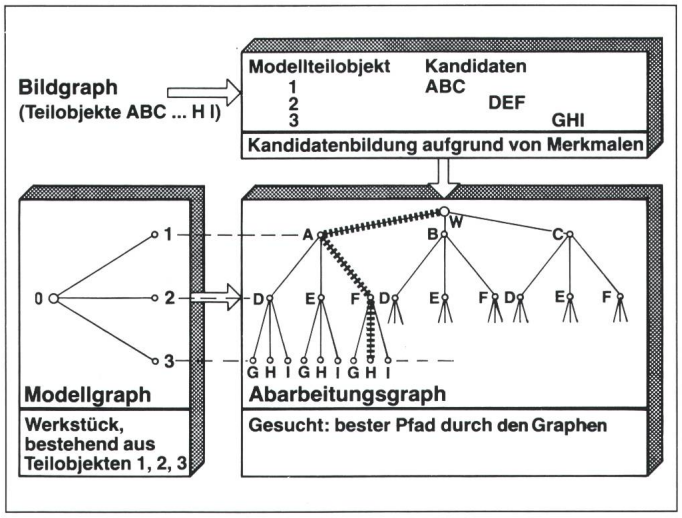


Fig. 6
Modellgesteuerte Bildanalyse
O: Wurzelknoten des Modellgraphs ≙ Modellobjekt
W: Wurzelknoten des Abarbeitungsgraphs ≙ Bildobjekt

wird ein Abarbeitungsgraph konstruiert. Hierzu sind die möglichen Alternativen für jedes Modellteilobjekt (1, 2, 3) als je ein Knoten entlang einer horizontalen Geraden (gestrichelt) einzuzeichnen. Werden die Knoten wie in Figur 6 mit Geradenstücken verbunden, so entsteht der Abarbeitungsgraph. Jeder Pfad von der Wurzel des Graphen (W) bis zu einem Knoten entspricht genau einer Zuordnungskonfiguration. Diese ist noch unvollständig, wenn der betrachtete Knoten nicht terminal ist, das heisst, wenn von diesen Knoten noch weitere Kanten weggehen. Sie ist erst vollständig, wenn der Knoten ein Terminalknoten (G, H, F) ist und damit ein Pfad durch den ganzen Graphen beschrieben wird. Jede auch unvollständige Konfiguration wird durch das gewählte Bewertungsmass bewertet. Das Ergebnis wird dem Knoten zugeordnet, an dem der Pfad der betrachteten Konfiguration endet. Das Aufsuchen der besten Zuordnung entspricht der schrittweisen Abarbeitung des Graphen. Der Algorithmus endet, wenn der am besten bewertete Pfad durch den ganzen Graphen gefunden wurde. Derartige Graphsuchverfahren spielen im Arbeitsgebiet der künstlichen Intelligenz eine wichtige Rolle.

Die Wirksamkeit des Verfahrens soll hier an einem Beispiel illustriert werden. Es geht dabei um die Erkennung von Stanzteilen, die in ungeordneter Schüttung teilweise übereinanderliegen, einschliesslich der Vermessung von Drehlage und Position. Die Figur 7 zeigt eines dieser Teile, so wie es für den Aufbau des Modells verwendet wird. Bildvorverarbeitung und Segmentierung haben sämtliche Zusammenhanggebiete bestimmt, von denen der Modellentwerfer einige, ihm charakteristisch erscheinende als Teilobjekte des Modells auswählt. In der Figur sind beispielsweise drei ausgesuchte Modellteilobjekte hervorgehoben.

Die Figur 8 stellt die zu analysierende Szene dar: Mehrere Stanzteile liegen ungeordnet, in verschiedenen Drehlagen teilweise übereinander. In dieser Figur sind die Ergebnisse der modellgesteuerten Bildanalyse eingetragen. Das zuoberst liegende Stanzteil wurde gefunden. Eingezeichnet sind die erkannten Modellteilobjekte und die Berandung des gesamten Werkstücks, so wie sie aus den Modelldaten rekonstruiert wurde. Drehlage und Koordinaten des Greifpunktes sind verfügbar. In einer solchen Situation liefert das Verfahren auch dann noch richtige Ergebnisse, wenn die Teile mehr oder weniger schräg liegen. Die gesamten Ergebnisdaten können an eine Robotersteuereinheit weitergegeben werden. Der Industrieroboter ist dann in der Lage, das Werkstück sichtgesteuert zu greifen.

Im industriellen Einsatz wäre die Aufgabe damit gelöst; das Werkstück wird gegriffen und von dem ungeordneten Haufen entfernt. Danach kann dieselbe Aufgabe erneut gelöst werden, bis der Stapel abgetragen ist. Die modellgesteuerte Bildanalyse ist darüber hinaus auch imstande, aus einem Bild wie Figur 8 auch die übrigen teilweise verdeckten Werkstücke so lange richtig zu erkennen, als auch ein

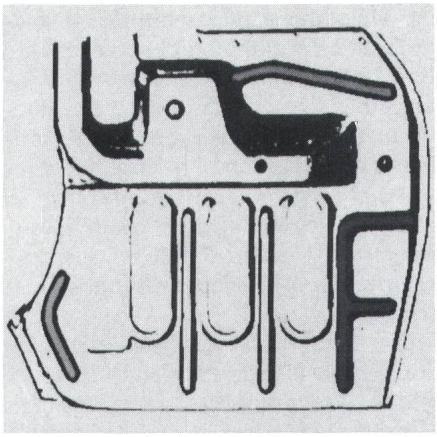
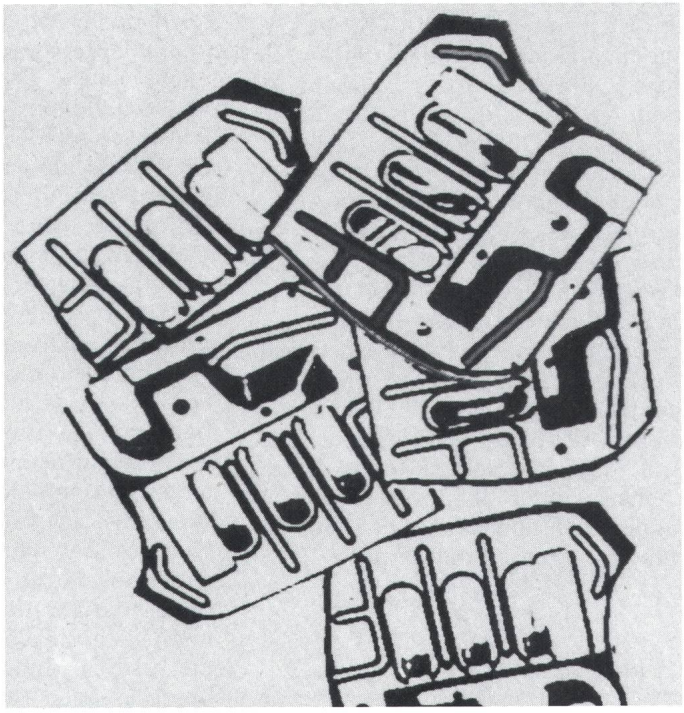


Fig. 7 Modellerstellung
Auswahl typischer Zusammenhangsgebiete

Fig. 8
Erkennung eines Stanzteils durch die modellgesteuerte Bildanalyse



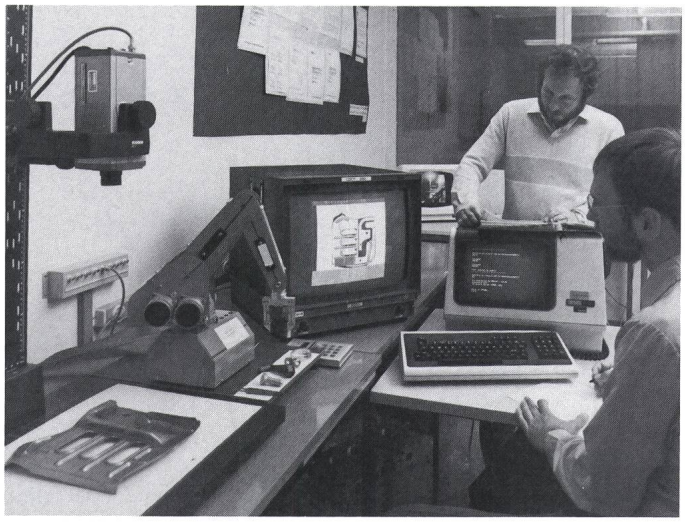


Fig. 9
Rechnergestütztes Experimentiersystem
AEG Forschungsinstitut Ulm

Bruchteil der im Modell festgelegten Teilobjekte ausreicht, um das Teil eindeutig zu identifizieren.

Experimentiersystem am Forschungsinstitut der AEG

Die beschriebenen Verfahren wurden am Forschungsinstitut der AEG in Ulm erarbeitet und finden fortlaufend Eingang in die Bildanalysesysteme, die das Unternehmen für den industriellen Einsatz entwickelt und auf den Markt bringt. Durch den Ausbau des Bereichs der symbolischen Bildverarbeitung und die Einführung von Strategien und Techniken aus dem Bereich der künstlichen Intelligenz ergeben sich deutliche Leistungssteigerungen gegenüber Bildverarbeitungssystemen des Typs, wie er gegenwärtig Stand der Technik ist – vor zwei Jahren wurde an dieser Stelle über ein derartiges System im Zusammenhang mit der Entpalettierung von Motorblökken berichtet. Für die Methodenentwicklung wird am Forschungsinstitut ein Experimentiersystem verwendet, das aus einem Rechner und einem speziell für Forschungszwecke eingerichteten Bildverarbeitungssystem mit Aufnahme- und Wiedergabeeinrichtungen besteht (Fig. 9). Sämtliche hier beschriebenen Methoden der ikonischen und symbolischen Bildanalyse sind auf diesem System implementiert und ablauffähig. Für Demonstrationszwecke wurde ein vom Rechner ansteuerbarer Tischroboter in das System integriert. Der Roboter ist in der Lage, vor ihm liegende Gegenstände zu sehen und zu greifen. Um die Vielfalt heute möglicher Interaktionen mit dem Robotersystem zu demonstrieren, wurde die Konfiguration um je ein im Hause entwickeltes Sprachein- und -ausgabesystem erweitert. Dadurch wird ein akustischer Dialog mit dem Experimentiersystem möglich. In vollsynthetisch generierter Sprache teilt das System mit, was es in der Szene sieht. Durch Spracheingabe können Anweisungen gegeben werden, welche der in der Szene erkannten Teile gegriffen werden sollen. Mit allen Einschränkungen, denen ein solches Experimentiersystem unterworfen ist, wurde damit doch gewissermassen ein sehender, sprechender und hörender Roboter aufgebaut.