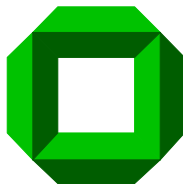


**Einführung in das CaRo-Projekt:
Geometrie- und Texturerfassung von 3D-Objekten mit
robotergeführter Videokamera**

Interner Bericht 11/98

A. Schmitt, M. Fautz, P. Oel

15. 5. 1998



Universität Karlsruhe
Fakultät für Informatik
Institut für Betriebs- und Dialogsysteme

Kurzzusammenfassung

Mit dem CaRo-Projekt (CaRo = Camera Roboter) wird an der Universität Karlsruhe ein neuartiger Ansatz verfolgt, um das Problem der Erfassung von Form- und Oberflächendaten für 3D-Objekte zu lösen. Bei CaRo führt ein Roboterarm eine Kamera und richtet sie gemäß der jeweiligen Erfassungsstrategie aus verschiedenen Richtungen auf das zu digitalisierende Objekt, z. B. ein Werkstück des Maschinenbaus. Mit Bildanalyseverfahren können nun die Koordinaten von Oberflächenpunkten des Werkstücks bestimmt werden. Diese Vorgehensweise hat mehrere Vorteile. Die Kameraführung erfolgt adaptiv und wird von der Analysesoftware vorgegeben. Zwischen Globalsichten und Detailvergrößerungen kann ständig gewechselt werden. Die Oberfläche des Werkstücks kann in hochauflösenden Farbbildern repräsentiert werden, so daß Textur und Farbe zusammen mit den Geometriedaten geliefert werden können.

Die Flexibilität des Ansatzes wird dadurch deutlich, daß sich mit ihm z.B. unterschiedlichste Anforderungen erfüllen lassen. So besteht die Möglichkeit, auch Bücher mit nicht planliegenden Seiten vollautomatisch zu digitalisieren, wobei nicht die 3D-Geometrie, sondern die Textrepräsentation im Vordergrund steht. Weiterhin können mit dem CaRo-Ansatz dreidimensionale Daten erzeugt werden, wie sie z. B. beim Reverse Engineering und in der Werbe- und Filmindustrie (virtuelle Welten, Computeranimationen in der Werbung, Trickfilm) benötigt werden.

Die bisher bekanntgewordenen Digitalisiergeräte können die geforderte Breite und Flexibilität der Objektdigitalisierung nicht leisten. Es besteht die begründete Hoffnung, daß der CaRo-Ansatz, also das bewegte Kameraauge, zu einer Standard-Eingabetechnik für die graphische Datenverarbeitung ausgebaut werden kann.

Abstract

The CaRo project (CaRo = Camera Robot) at the University of Karlsruhe uses a new approach, to solve the problem of the acquisition of geometry and surface data of three-dimensional objects. With CaRo a robot arm moves a camera according to the acquisition strategy and points it from various directions towards the object which should be digitized, e.g. a workpiece of mechanical engineering. Using image analysis algorithms the coordinates of surface points of the workpiece can be determined. This method has several advantages. The camera guidance given by the analysis software is adaptive. Between global views and detail enlargements can easily be changed. Texture and color as well as the geometry data can be supplied since the surface of the workpiece can be represented using high-resolution color images.

The flexibility of the approach becomes clear by looking at the different requests that can be fulfilled. It is even possible to digitize books with cambered pages, which means not getting a three-dimensional representation but getting a textual representation. Furthermore three-dimensional data can be produced. This is needed e.g. in reverse engineering and in advertising and film industry (virtual worlds, computer animations and commercial trickfilm).

The digitization hardware known so far does not provide the required width and flexibility for object digitization. There is the justified hope that the CaRo approach, the moveable camera eye, will become a standard input technique in computer graphics.

Inhalt

EINLEITUNG	5
1. PROBLEMSTELLUNG	5
2. LÖSUNGSANSATZ.....	6
3. HARDWAREARCHITEKTUR UND INTERFACING	8
4. DATENERFASSUNGSSTRATEGIEN UND FORSCHUNGSTHEMEN.....	11
4.1 WERKSTÜCKGROBUMRIß BESTIMMEN	11
4.2 NAH-FERN-REKONSTRUKTIONSTECHNIK	13
4.3 REKONSTRUKTION MIT LASER-HILFSSTRAHL.....	14
4.4 REKONSTRUKTION MIT OPTISCH PROJIZIERTEN HILFSMUSTERN.....	15
4.5 REKONSTRUKTION MIT REFERENZGITTER	16
4.6 KOMBINATION VON TEXTUR UND GEOMETRIE: DAS CARO-DATENFORMAT.....	17
4.7 SEMIPLANARE OBJEKTE.....	18
4.8 SICHTSYSTEM-ANWENDUNGEN IM KLEINEN	19
5. AUSBLICK.....	19
LITERATUR.....	21

Einleitung

Wir geben hier einen Überblick über das CaRo-Projekt und erläutern einzelne geplante Verfahrensweisen.

In den nächsten beiden Abschnitten werden zuerst die Problemstellung und dann der Lösungsansatz genauer erläutert. Abschnitt 3 gibt genauere Informationen zur CaRo-Hardware. Schließlich werden in Abschnitt 4 Datenerfassungsstrategien und Forschungsthemen dargelegt. Der Bericht schließt mit einem Ausblick ab.

1. Problemstellung

Wer heute speziell im 3D-Bereich graphische Datenverarbeitung betreibt, möchte 3D-Objekte in flexibler Weise bearbeiten. Dazu gehören nicht nur Entwurf und Manipulation von Objekten auf dem graphischen Bildschirm sowie die Speicherung evtl. in Datenbanken, sondern auch die von [Schmitt 83, S. 189] in Zusammenhang mit der allgemeinen Struktur von Entwurfssystemen so genannten Einschleuse- und Ausschleusefunktionen, die der graphischen Eingabe- und Ausgabefunktionalität entsprechen. Unter Einschleusen versteht man die datenmäßige Erfassung von in der realen Welt existierenden Objekten, um sie im Graphik-System weiterbearbeiten zu können [Faugeras 93, Jain et al. 95]. So sollen existierende Werkstücke des Maschinenbaus, Gußrohlinge, Skelett-Knochensteile, Zähne, Pflanzen und Blattformen in der Biologie, aber auch z. B. bereits existierende Produkte wie Teller, Werkzeuge usw. speziell für die Computeranimation und Visualisierung in eine 3D-Datenrepräsentation umgesetzt werden.

Die Ansprüche und Bedürfnisse der Anwender an die eingeschleusten 3D-Daten sind sehr unterschiedlich. Der Ingenieur möchte das Werkstück verändern, muß also die Daten in das Format umsetzen können, das von seinem CAD-System benötigt wird. Der Paläontologe möchte die Knochenstücke genau ausmessen, um daraus Schlußfolgerungen zu ziehen. Beim Zusammenspiel von Designern und Ingenieuren werden oft erst Gipsmodelle erstellt, deren Geometrie dann in CAD-Systeme übernommen werden soll. Der Computeranimateur soll ein Werbevideo mit real vor ihm liegenden kommerziellen Produkten herstellen, er benötigt also auch genaue Oberflächentexturen, um mit Raytracing hochrealistische Visualisierungen erzielen zu können. Während also im einem Fall nur die nackte Geometrie, diese aber genau gewünscht ist, werden im anderen Fall auch die genaue Oberflächenbeschaffenheit inklusive Farbe, Texturen, Beschriftungen etc. angestrebt.

Invers zum Einschleusen ist das Ausschleusen, die Umsetzung von 3D-Daten in echte Modelle, dazu gehört z.B. die CNC-Technik und das Formfräsen, aber auch sehr aufwendige Verfahren. Beide Richtungen, sowohl das Ein- als auch das Ausschleusen sind in der graphischen Datenverarbeitung stets problematisch

und werden normalerweise durch Spezialverfahren mit beschränktem Einsatzbereich gelöst.

Wir haben uns daher die Frage gestellt, wie die Einschleusefunktion etwas genereller angegangen werden kann. Probleme entstehen bereits, wenn man Bücher vollautomatisch, also ohne Handarbeit, digitalisieren möchte. Der Flachbett-Scanner üblicher Bauart kann nicht plan liegende Buchseiten schlecht abtasten. Beim Nachdenken über eine sinnvolle Lösung wurde die Idee geboren, eine robotergeführte Kamera als Scanner zu benutzen. Eine solche im Raum bewegliche Kamera kann mehrerlei: Eine gewölbt liegende Buchseite, wie bei einem normal aufgeschlagenen Buch, kann auch schräg angepeilt werden, um stets eine senkrechte Sicht auf das Papier zu gewährleisten, was natürlich voraussetzt, daß das Gesamt-Pixelbild aus vielen sich überlappenden Einzelaufnahmen zusammengesetzt werden muß. Die Kamera kann aber auch abhängig vom Detailreichtum, also vom Bildinhalt, den Vergrößerungsmaßstab beliebig variieren, sich also an den jeweils erforderlichen Detaillierungsgrad der Vorlage anpassen. Wenn die OCR-Analyse gleichzeitig mitläuft, können problematische Zeichen genauer und in höherer Auflösung angeschaut werden. So kommt man zu einer weitgehenden Analogie zum lesenden Auge, wir fixieren schwer erkennbares ebenfalls durch näheres Herangehen. Wenn wir Details eines Werkstücks überprüfen wollen, z.B. die Sauberkeit einer Kante, müssen wir ebenfalls genauer hinschauen.

2. Lösungsansatz

Dem CaRo-Projekt liegt folgender technischer Systemaufbau zugrunde: Das zu digitalisierende 3D-Objekt liegt im Arbeitsraum eines Roboterarmes. Der Roboter ist aber nicht mit einer Greifhand ausgestattet, sondern führt statt der Hand eine CCD-Videokamera, die ein gut aufgelöstes Farb-Videosignal liefert. Der Roboter kann die Kamera um das 3D-Objekt herumführen und von allen Seiten betrachten, auch von oben, aber natürlich nicht von unten. Die so entstehenden Video-Standbilder werden bei Bedarf von einer Framegrabber-Karte in voller Farbauflösung digitalisiert und stehen dem Analysesoftware zur Verfügung.

Nun sind photographische Techniken für die Objektrekonstruktion nichts Neues, sie konnten sich aber wegen einer Reihe technischer Probleme noch nicht recht durchsetzen. Zu erwähnen sind die Bildanalyse und vor allem die Kalibrierprobleme der Kamera und speziell die genaue Bestimmung der Kameraposition nur auf der Basis des Bildmaterials. Schon in den Jahren 1988 bis 1991 haben wir am Institut das Problem untersucht, wie aus mehreren von verschiedenen Standorten aufgenommenen Fotos eine 3D-Rekonstruktion realisiert werden kann [Leister 91]. Es ergaben sich deutliche Probleme mit der Genauigkeit, insbesondere gelang es nur sehr grob, die Kamerastandorte aus

dem Bildmaterial hinreichend genau zu erschließen, obwohl dies theoretisch möglich ist.

Das Kamera-Kalibrierproblem wird durch unseren Ansatz entscheidend entschärft. Die äußeren Kameraparameter sind bekannt, denn der Roboter arbeitet in seinem eigenen festen Koordinatensystem, so daß man bei jedem von der Kamera aufgenommenen Bild sowohl Kameraposition als auch Blickrichtung genau kennt. Zur Bestimmung innerer Kameraparameter kann man die Kamera gezielt auf Kalibriermuster mit bekannter Geometrie richten und kann aus den Bilddaten mit relativ einfachen Verfahren die Objektivverzerrungen und die rechnerische Brennweite bestimmen [Tsai 86]. Dadurch wird auch die Kamerakalibrierung weitgehend automatisierbar. Wie eine solche Kalibrierung auch bei den in der Anwendung wesentlich flexibleren Zoom-Objektiven vorgenommen werden kann, bleibt noch zu untersuchen, vgl. z. B. [Tarabanis et. al. 92, Watanabe 96, Li, Lavest 95]. Objektive mit fester Brennweite und bekannter Entfernungseinstellung sind demgegenüber problemlos zu kalibrieren.

Ein weiterer wesentlicher Vorteil des CaRo-Konzeptes ergibt sich durch die Möglichkeit, die Kamera auch sehr nahe an Objektdetails heranfahren zu können und so stark vergrößerte Detailansichten zu gewinnen. Bei geeigneter Objektivausstattung kann man fast in den mikroskopischen Bereich vordringen. Daher können auch ziemlich kleine Objekte abgetastet werden und feine Texturen können bis zu einer Auflösung von ca. 100 Pixel/mm erfaßt werden. Allgemein kann man feststellen, daß die Auflösung des Rekonstruktionsverfahrens in weiten Bereichen vorgebar ist und daß Oberflächen farblich genau digitalisiert werden können, was den Anwendungsbereich deutlich ausweitet. Z.B. können so auch Qualitätsprüfungen von Oberflächen automatisiert durchgeführt werden.

Bei der Rekonstruktion größerer Objekte wirkt sich der beschränkte Arbeitsraum des Roboters aus, denn die Kamera kann das Objekt nicht mehr von allen Seiten betrachten. In solchen Fällen wird die gut bekannte Drehtellertechnik eingesetzt. In geeigneter Entfernung vom Roboter wird der Drehteller aufgestellt, auf ihm wird das zu digitalisierende Objekt fixiert. Der Teller kann nun von der Steuersoftware in einem 360°-Winkelbereich gedreht werden, so daß auch seitliche und rückwärtige Ansichten zugänglich sind. Insbesondere können so ganz einfach Konturdaten des Objektes erfaßt werden.

Erwähnenswert als weiteren wichtigen Pluspunkt ist auch die Möglichkeit, daß die Kamera abhängig von der jeweiligen Digitalisierstrategie bewegt werden kann. Wir müssen also bei der Digitalisierung nicht nach festen Schemata vorgehen, sondern können abhängig von der Datenauswertung in der Digitalisierungssoftware variable Strategien einschlagen [Jarvis 83]. So kann z. B. ein charakteristischer Punkt auf dem zu digitalisierenden Objekt mehrfach und

von verschiedenen Seiten angepeilt werden, um mittels Ausgleichsrechnung besonders hohe Koordinatengenauigkeit zu erzielen.

In der Summe ergeben sich bei dem vorgestellten Lösungsansatz eine Reihe von gewichtigen Vorteilen, die uns hoffen lassen, daß wir auf diesem Wege in der Tat eine flexible Lösung für das Einschleuseproblem in der graphischen Datenverarbeitung gefunden haben. Das Forschungsprojekt ist jedoch von einer beachtlichen Komplexität und erfordert umfangreiche experimentelle Arbeiten sowie die Entwicklung neuartiger Softwarearchitekturen, um zuverlässige Aussagen über die Praxistauglichkeit zu gewinnen.

3. Hardwarearchitektur und Interfacing

In der ersten Hälfte 1997 wurde ein Roboter des Typs Typ VS-6354 der Fa. Schuster in Denklingen/Bayern mit 6 Freiheitsgraden erworben und die umgebende Arbeitszelle aufgebaut, um die für das CaRo-Projekt erforderliche Laborausrüstung bereitzustellen, siehe Abbildungen 1 und 2. Der angeschaffte Roboter-Arm hat folgende Leistungsdaten:

- 6 Achsen,
- Nutzlast 3 kg,
- Wiederholgenauigkeit in x-y-z-Koordinaten 0,02 mm,
- Motoren/Getriebe: bürstenlose AC Servo-Motoren mit absoluter Winkelmessung, Harmonic-Drive-Technik.

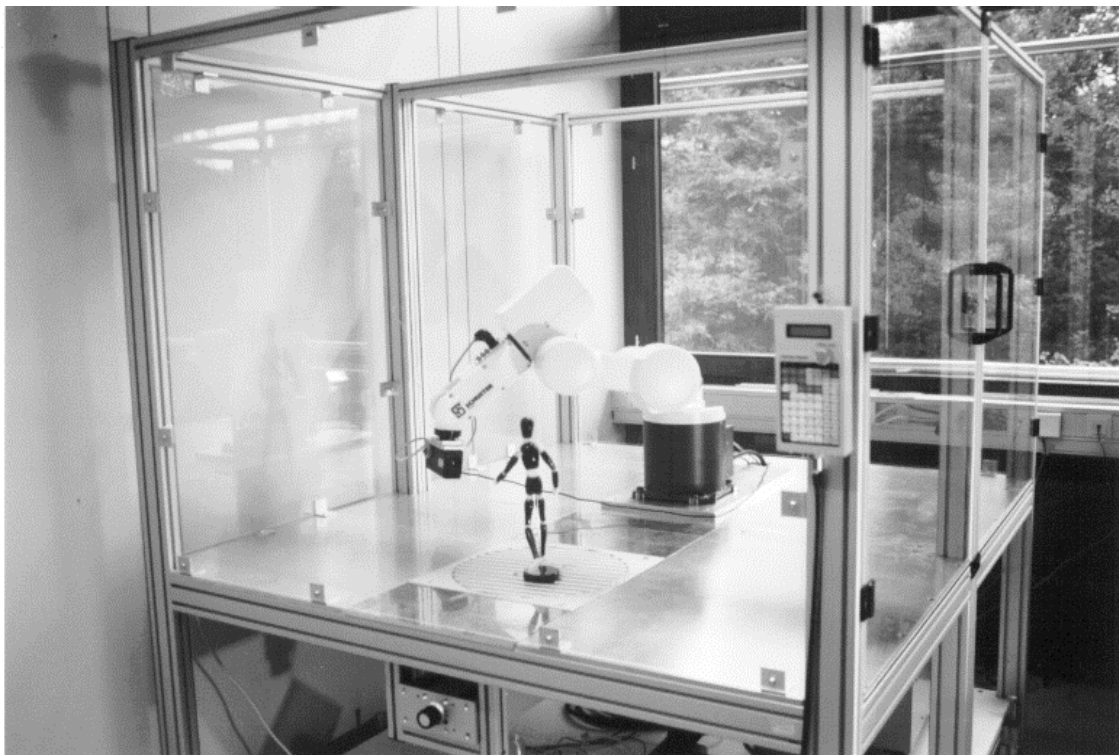


Abbildung 1: Arbeitszelle mit Roboter, Kamera an der Roboterhand fest montiert. Auf dem im Arbeitstisch eingelassenen Drehteller befindet sich das zu digitalisierende Modell, eine stilisierte menschliche Gliederpuppe.

An der Roboterhand wurde eine 1/3"-CCD-Farb-Videokamera Modell CD-08 von Chugai Boyeki Co. befestigt. Die Kamera ist derzeit mit einem Zoom-Objektiv 6.5-52 mm, 1:1.8 bestückt und erlaubt über eine serielle Schnittstelle die Einstellung von Focus, Zoom-Brennweite, Verschußzeit, Weißbalance, Blende und einige weitere Parameter. Derzeit verwenden wir noch die Fokuseinstellautomatik der Kamera. Ob dies auf Dauer so bleiben kann, werden die weiteren Erfahrungen zeigen. Insbesondere wird angestrebt, die Kamera bei Bedarf auch sehr nahe an zu digitalisierende Objekte heranzufahren, um sehr hohe Detailauflösungen zu erzielen. Erfahrungsgemäß ist dabei mit einer geringen Tiefenschärfe der Einzelaufnahmen zu rechnen, so daß man die Strategie anwenden muß, die Kamera stets möglichst senkrecht auf die Objektoberflächen zu richten.

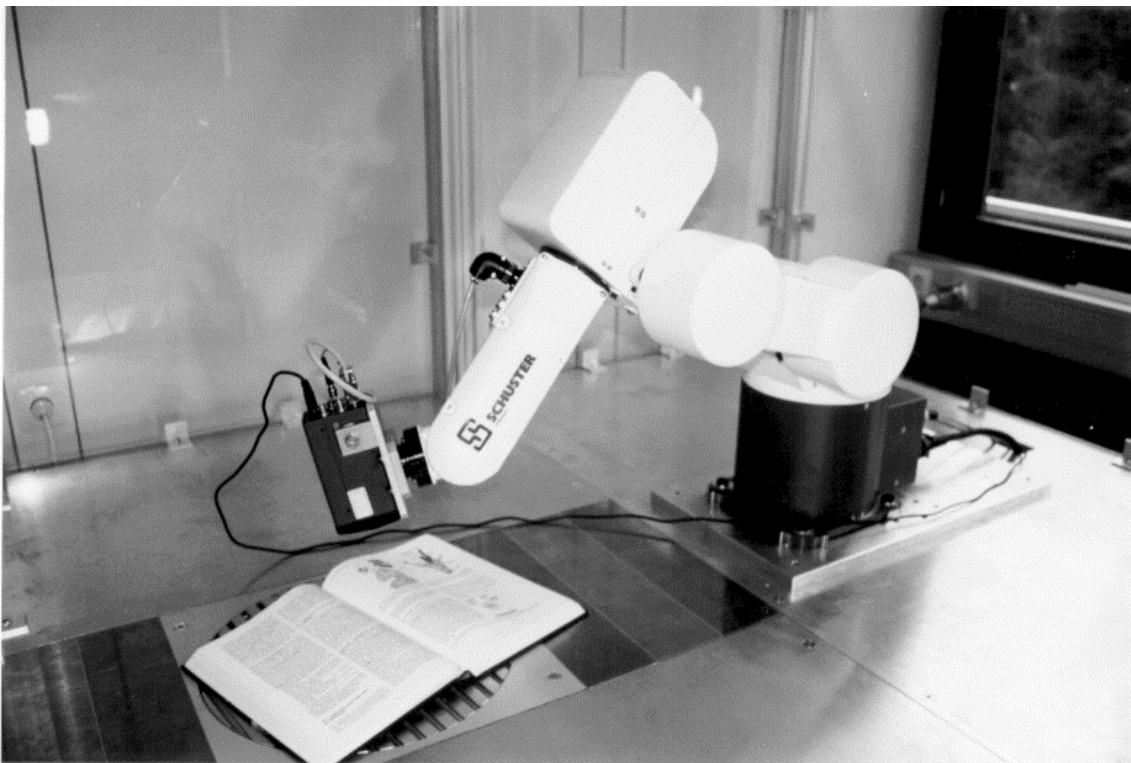


Abbildung 2: Roboter und Kamera in Nahaufnahme. Auf dem Drehteller befindet sich ein aufgeschlagenes Buch, dessen gekrümmt liegende Seiten von der Kamera digitalisiert werden sollen, vergleiche hierzu auch Anmerkungen zur Buchlesemaschine in Abschnitt 4.7.

Die wichtigsten Hardware-Module und Datenflüsse gehen aus Abbildung 3 hervor. Aus der Sicht des Anwendungsrechners, auf dem also die eigentliche Anwendungssoftware läuft, wird die gesamte Roboterzelle über einen transparenten Schnittstellentreiber gesteuert.

Die Funktionalität dieser zentralen Schnittstelle kann durch Auflistung ihrer wichtigsten Funktionen beschrieben werden:

- a.) Roboter in Initialposition versetzen.

- b.) Roboter von der aktuellen Position direkt oder interpolierend über Zwischenpositionen in eine neue Position bringen. (Eine Positionsangabe ist ein x-y-z-Koordinatentripel, die Einstellung der Kameraorientierung erfordert zwei weitere Koordinatentripel, vgl. Abbildung 4)
- c.) Einlesen des aktuellen Videobildes von der Roboterkamera.
- d.) Drehteller-Winkelstellung vorgeben.
- e.) Fehlermeldungen abfragen.

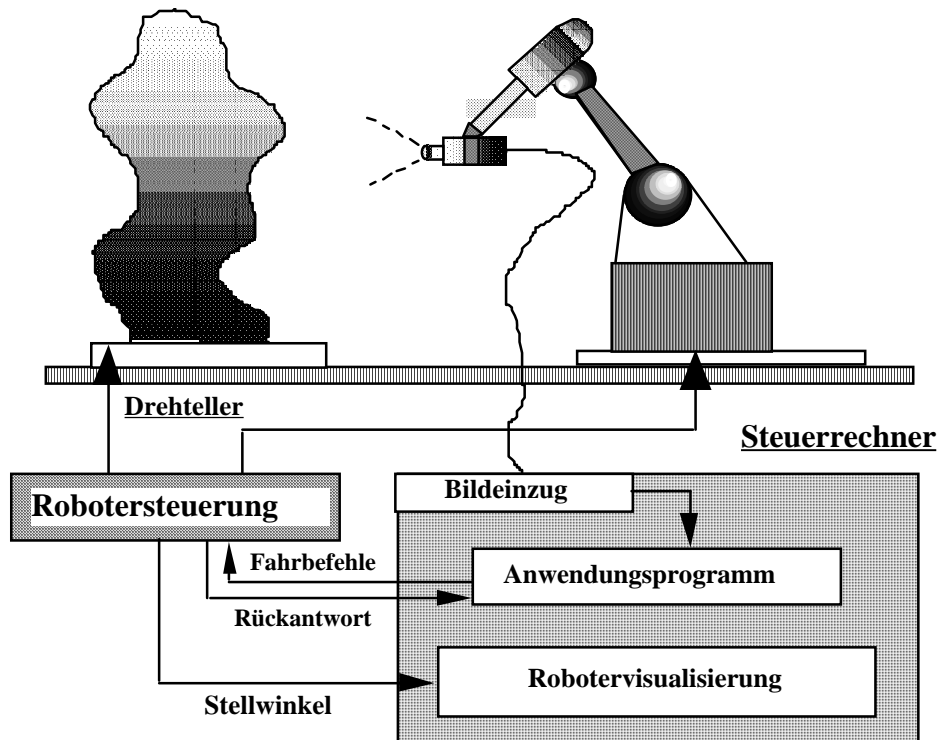


Abbildung 3: Blockstrukturbild bei Anwendung auf dem Steuerrechner

Auch wenn zukünftig noch Ausleuchtungsbefehle (Licht und Laser-Hilfsstrahl) und Befehle für die Kamera-Objektiveinstellung hinzukommen werden, ist dies ein relativ einfaches und übersichtliches Interface. Anwendungsprogramme können so alle Einstellungen vornehmen, ihre zu verarbeitenden Daten erhalten sie ausschließlich über den Einlesebefehl für Videobilder.

Derzeit haben wir noch keine Kollisionsvermeidungssoftware im Einsatz. Längerfristig muß ein solches Überwachungssystem realisiert werden, um Beschädigungen an der Kamera und am Roboter selbst zu vermeiden. Die Anwendungsprogramme sollten von derartigen Zusatzprüfungen entlastet sein. Beim normalen fabrikmäßigen Einsatz von Robotern tritt das Kollisionsproblem kaum auf, weil mit vorgefertigten eintrainierten Fahrprogrammen gearbeitet wird. Wenn solche Programme fehlerfrei laufen, sind auch bei weiteren Wiederholungen keine Kollisionsfehler zu erwarten.

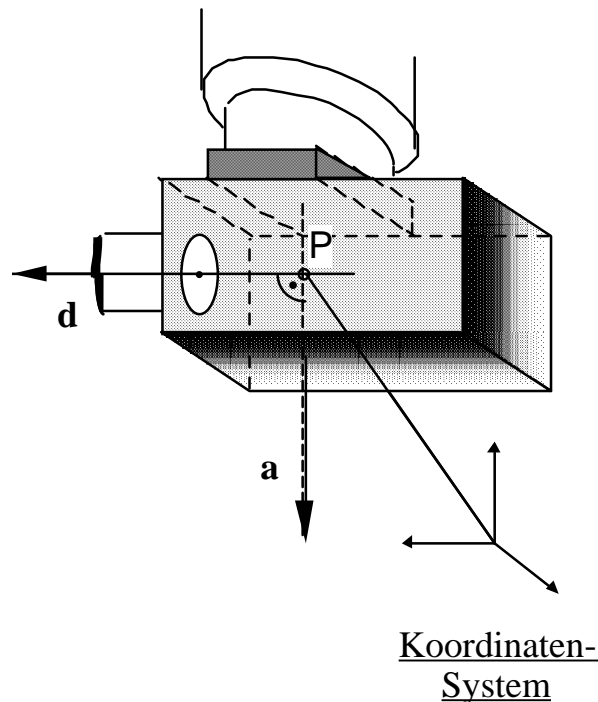


Abbildung 4: Eine Position ist ein x-y-z-Koordinatentripel (Punkt P), die Einstellung der Kameraorientierung erfordert 2 weitere Vektoren (direction d, approach a)

Um von verschiedenen Arbeitsplatzrechnern aus Anwendungsprogramme entwickeln zu können, denken wir beim Interfacing längerfristig an eine Client-Server-Architektur. Die CaRo-Maschinerie kann dann auch von entfernten, am Netz hängenden Rechnern aus betrieben werden. Um die Roboterzelle von der Ferne überwachen zu können, sind dann noch 2 bis 3 fest installierte Übersichtskameras vorzusehen, deren aktuelle Bilder auf den Workstation-Bildschirmen kontrolliert werden können. Insbesondere sind auch interaktive Betriebsweisen von CaRo nur von einem Bildschirm-Arbeitsplatz aus denkbar, vor allem für Entwicklung und Test neuer Verfahren.

4. Datenerfassungsstrategien und Forschungsthemen

Die folgenden etwas genauer diskutierten Einzelverfahren haben vorläufigen Charakter und sollen etwas mehr Klarheit in die strategische Ausrichtung des CaRo-Projektes liefern. Es werden Ideen dargestellt, wie die anstehenden Probleme einer Lösung zugeführt werden können. An der einen oder anderen Stelle lauern sicher noch unvorhersehbare Detailprobleme. Das soll uns aber nicht davon abhalten, unsere Vorstellungen bereits hier am Anfang unserer Arbeit auszubreiten.

4.1 Werkstückgrobumriß bestimmen

Ein automatisierter Digitalisierprozeß muß ganz am Anfang, wenn noch nichts über Größe und Ausdehnung des Werkstücks bekannt ist, herausfinden, wo die Kamera nicht hinfahren darf. Es muß also so etwas wie eine "bounding Box"

um das Werkstück bzw. das zu digitalisierende Objekt gelegt werden. Das geschieht am besten durch entferntere Kameransichten, von oben, von vorne und von seitlich auf das Werkstück. Wegen der bereits genau kalibrierten Kamera und wegen der genau bekannten Kameraposition ist es wohl nicht schwierig, jeweils die Außenkontur des Objektes auf den Sichtbildern zu bestimmen. Insbesondere ist der Innenraum der Roboterzelle entweder völlig weiß oder völlig schwarz oder in einer selten vorkommenden Farbe ausgekleidet, was die Konturerkennung in den digitalisierten Bildern erheblich erleichtern sollte. Geometrisch ergibt sich dann aus jedem Bild ein Innenvolumen, das einer Pyramide mit dem Grundriß der Konturkurve entspricht, mit der Pyramidenspitze im Kamerazentralpunkt, siehe Abbildung 5. Die durch die verschiedenen Sichten so gewonnenen Innenvolumen werden mengentheoretisch miteinander geschnitten, das Schnittgebilde, von dem auch der Halbraum unter der Tischebene abzutrennen ist, ist eine Grobapproximation des zu digitalisierenden Objekts [Martin, Aggarwal 83].

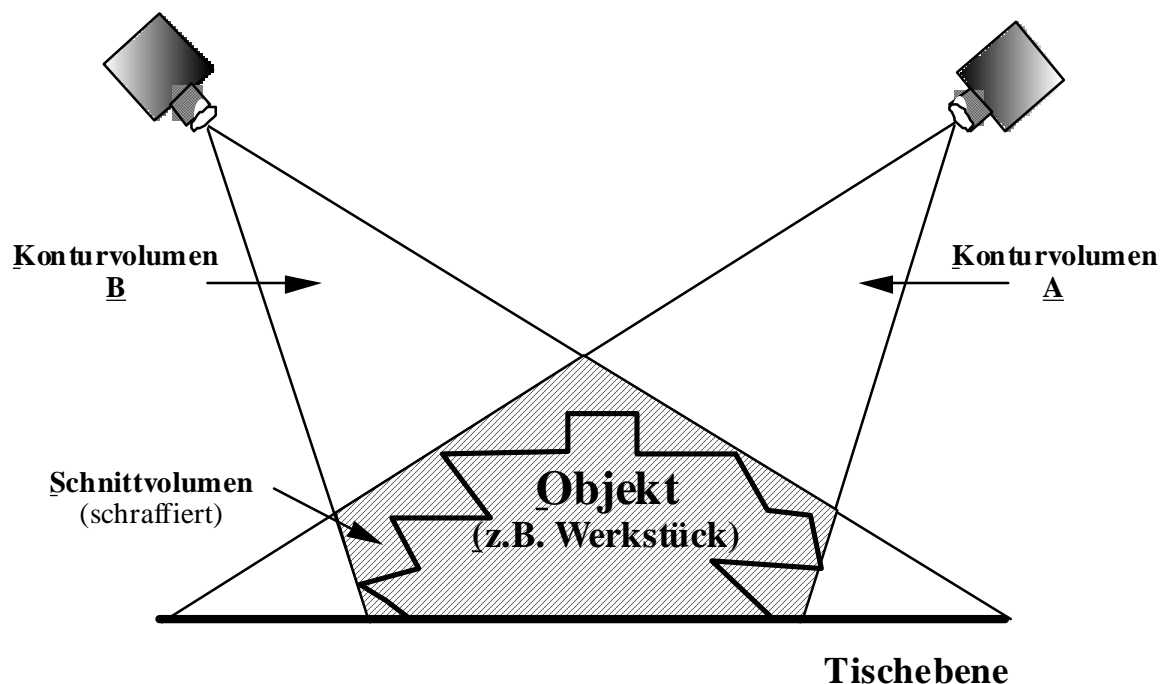


Abbildung 5: Volumenapproximation

Treibt man mehr Aufwand mit vielen verschiedenen Kamera-Positionen, so wird das Objekt bis nahe an die konvexe Hülle approximiert.

Das Berechnen des Schnitts der Innenvolumen geht wohl am einfachsten durch die Voxel-Methode, also durch eine hinreichend feine reguläre Zellstruktur innerhalb des Zellraums. Am Schluß sind die Voxel dieses Raumes markiert, die in allen Konturvolumina liegen, sie repräsentieren dann ein Hüllvolumen für das Werkstück.

4.2 Nah-Fern-Rekonstruktionstechnik

Dies ist eine vom Ansatz her außerordentlich durchsichtige Technik, um die Geometrie eines texturiert/strukturierten Körpers zumindest approximativ zu bestimmen. Es werden nur zwei Bilder ausgewertet, das eine näher am Objekt, das andere ferner. Die Kameras sind genau auf den gleichen Hauptsehstrahl ausgerichtet, siehe Abbildung 6.

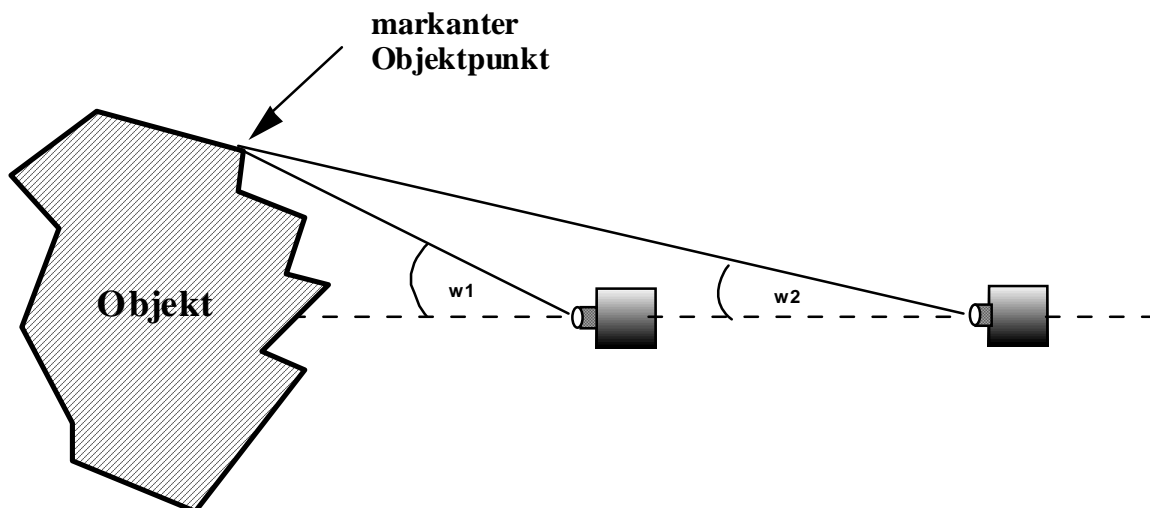


Abbildung 6: Nah-Fern-Rekonstruktionstechnik

Die genau ausgerichteten Kameras haben den markanten Objektpunkt jeweils auf ihren Aufnahme-Bildern. Wir müssen voraussetzen, daß der markante Objektpunkt durch ein geeignetes Bildanalyseverfahren auf beiden Bildern identisch zugeordnet werden kann. Der Rest ist einfach: Die geometrische Lage der Strahlen unter dem Winkel w_1 bzw. w_2 sind vollständig bekannt. Ihr Schnittpunkt kann berechnet werden und dies ist der Ort, an dem sich der markante Objektpunkt befindet, dessen Koordinaten nun zumindest näherungsweise bekannt sind. Für jeden in beiden Bildern durch Bildanalyse identifizierbaren markanten Punkt können so Näherungskordinaten bestimmt werden. Die Bildanalyse sollte deshalb nicht so schwierig sein, weil das hintere Bild im wesentlichen eine Verkleinerung des vorderen Bildes ist, es gibt nur Änderungen durch die Perspektive.

Problematisch bei diesem Ansatz ist die erreichbare Genauigkeit [Delherm et al. 96]. In der Umgebung der optischen Achse wird der Triangulationswinkel sehr klein, insbesondere können Punkte, die auf der opt. Achse liegen, nicht rekonstruiert werden. Da die Kamera frei und in jede Position bewegt werden kann, ist es stets möglich, den interessierenden markanten Punkt an den Bildrand zu bringen, was dann genaueres Anpeilen ermöglicht.

Die Rekonstruktionsgenauigkeit dieses Verfahrens kann sofort erhöht werden, indem mehr Kamerasichten auch aus der Nähe auf den markanten Punkt gerichtet werden. Offensichtlich ist es nicht notwendig, die Kamera auf einem Hauptsehstrahl auszurichten. Es muß nur gewährleistet sein, daß der markante Punkt durch die Bildanalyseverfahren überall eindeutig zu identifizieren ist.

Fehlerhafte Lokalisation des Punktes in einem Bild fällt auf, weil sich dann bei verschiedenen Anpeilungen voneinander abweichende x - y - z -Koordinaten für den Oberflächenpunkt ergeben.

Bei Oberflächen ohne markante Punkte ist dieses Verfahren nicht anwendbar. In der Praxis hilft man sich dann oft, indem man die Oberfläche mit Farbtröpfchen besprüht, die dann wieder markante Punkte liefern. Aber auch ohne derartige Veränderungen bzw. Verschmutzungen der Oberfläche ist die Oberflächenrekonstruktion möglich.

4.3 Rekonstruktion mit Laser-Hilfsstrahl

Hat man keine markanten Punkte, keine Kanten und Ecken zum Anpeilen zur Verfügung, so gibt es den Ausweg mit dem Laser-Hilfsstrahl. Wenn z.B. ein Hühnerei, exakt wie es ist, digitalisiert werden soll, so haben wir auf der Eierschale keine markanten Punkte, mit denen die Bildanalyse etwas anfangen könnte. In solchen Fällen können wir einen dünnen Laserstrahl erzeugen, der parallel zum Hauptsehstrahl verläuft, siehe Abbildung 7.

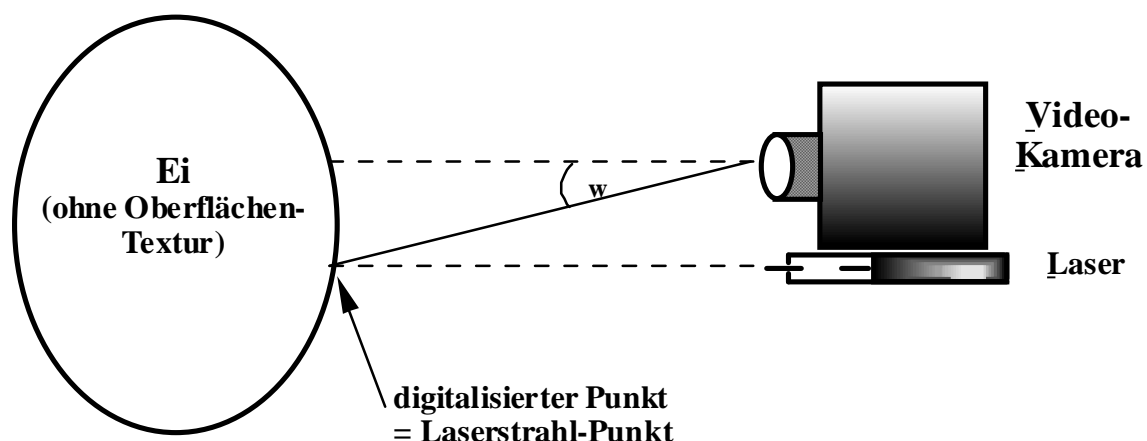


Abbildung 7: Rekonstruktion ohne markante Punkte

Da auch der Laserstrahl kalibriert ist, das heißt, aus der genauen Kameraposition und Blickrichtung der Kamera können wir auch den genauen Verlauf des Laserstrahls bestimmen, ist es sofort möglich, den Laserpunkt im Bild zu suchen; die unter dem Winkel w von der Kamera ausgehende, den Laserpunkt treffende Gerade zu bestimmen und mit der Laser-Geraden zu schneiden [Curless, Levoy 96]. Der so gefundene Punkt liegt auf der Oberfläche des Eies.

Durch systematisches Anpeilen der Nachbarschaft von bereits digitalisierten Punkten ist es möglich, ein Punktraster über die Oberfläche zu legen. Wenn z.B. gefordert wird, daß die Digitalisierpunkte ca. 3 mm voneinander entfernt liegen sollen, so kann das problemlos in unsere Abtaststrategie eingebaut werden, denn wir haben vollständige geometrische Information über die bereits digitalisierten Punkte und können sehr genau auf deren Nachbarschaft zielen. Da der Laserpunkt im Videobild sehr schnell gefunden werden kann, der Roboter

wenig Zeit zum Anfahren benachbarter Punkte (wenige mm Weglänge) benötigt, kann ein einzelner Punkt nach derzeitigen Schätzungen in deutlich unter 1 Sekunde digitalisiert werden. Vorteilhaft an diesem Verfahren ist der Umstand, daß nur ein sehr einfaches Bildanalyseverfahren benötigt wird, während z. B. die automatische Identifikation markanter Punkte erheblich aufwendiger ist.

4.4 Rekonstruktion mit optisch projizierten Hilfsmustern

Bei den hier kurz beschriebenen Methoden werden fehlende markante Punkte durch statische Lichtmusterprojektion auf das zu digitalisierende Objekt erzeugt [Will, Pennington 72]. In der Praxis bewährt haben sich hier Linien oder Gittermuster. Auf der Oberfläche gekrümmter Objekte zeichnet sich ihre Projektion als Höhenlinien ab, siehe Abbildung 8.

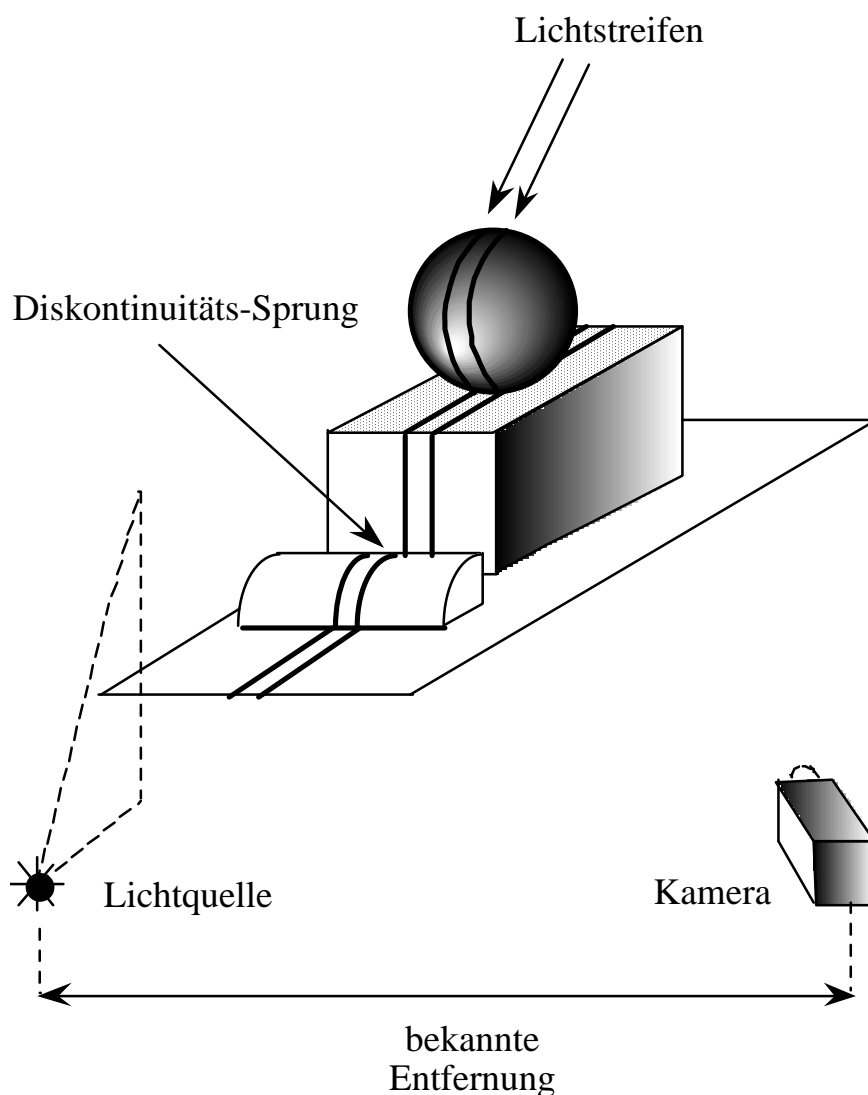


Abbildung 8: Aus dem Verlauf von projizierten Linien lassen sich Krümmungs- und Tiefeninformation gewinnen, vgl. z. B. [Jarvis 83] und [Jaine et al. 95, S. 301 ff].

Aus dem Verlauf der projizierten Linien lassen sich Aussagen über das Krümmungsverhalten der zu rekonstruierenden Oberfläche gewinnen. Außerdem können die nun zahlreich vorhandenen markanten Punkte zur Gewinnung der Tiefeninformation verwendet werden.

Weiterentwicklungen des Ansatzes der Projektion von Liniemustern sind unter dem Namen "codierter strukturierter Lichtansatz" bekannt. Hierbei werden Liniemuster unterschiedlicher Dichte auf das zu erfassende Objekt projiziert. Als Ergebnis erhält man ein $2\frac{1}{2}$ -dimensionales Tiefenbild. Dieses Verfahren wird heute bereits in der Praxis eingesetzt [Breuckmann 93]. Z.B. wird es beim Prototypentwurf in der Automobilindustrie verwendet. Derzeit kann noch nicht entschieden werden, ob im CaRo-Projekt dieser Ansatz weiter verfolgt werden sollte.

4.5 Rekonstruktion mit Referenzgitter

Bisher wurden Verfahren vorgestellt, die von einer kalibrierten Kamera ausgehen. Treten in den Bildern starke Verzerrungen auf, wie sie z.B. durch Aufnahmen im Nahbereich hervorgerufen werden können, so ist es jedoch oft schwierig, genaue Kameraparameter zu bestimmen. Dies liegt daran, daß das mathematische Kameramodell nur näherungsweise die tatsächlichen physikalischen Gegebenheiten widerspiegelt [Delherm et al. 96].

Bei Verfahren, die mit einer unkalibrierten Kamera arbeiten, werden in der Regel Bilder von einem oder mehreren Referenzgittern aufgezeichnet. In diesen Referenzbildern sind nun die Abbildungsparameter des Kamera-Linsensystems enthalten. Um die 3D-Koordinaten eines Objektpunktes zu bestimmen, geht man nun in zwei Schritten vor: Zunächst wird für einen Objektbildpunkt $P(u,v)$ der Sektor $(G_{1,1} G_{1,2} G_{2,1} G_{2,2})$ innerhalb der Aufnahme des Referenzgitters bestimmt, in dem der Punkt liegt (vgl. Abbildung 9). Aus dem Wissen über die Geometrie des Referenzgitters läßt sich dann seine Lage im Ausgangsreferenzgitter ermitteln.

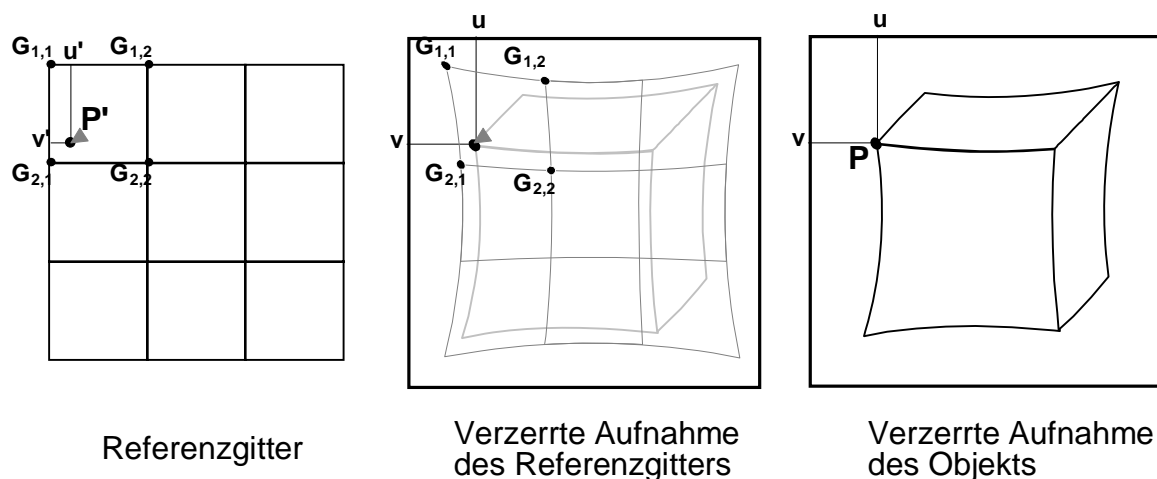


Abbildung 9: Bestimmung von Objektkoordinaten bezüglich eines Referenzgitters.

Ist das optische Zentrum bekannt, dann kennt man bereits den Sehstrahl, auf dem der Objektpunkt liegt. Andernfalls muß der erste Schritt bei gleicher Kamera- und Objektposition mit einem anderen Gitter wiederholt werden. Da der Abstand D zwischen den Referenzgittern genau bekannt ist, läßt sich nun ein Sehstrahl für diesen Kamerastandpunkt und den beobachteten Objektpunkt bestimmen [Martins et al. 81]. Aus mehreren solcher Sehstrahlen kann nun die Lage des Objektpunkts im 3D berechnet werden (vgl. Abbildung 10).

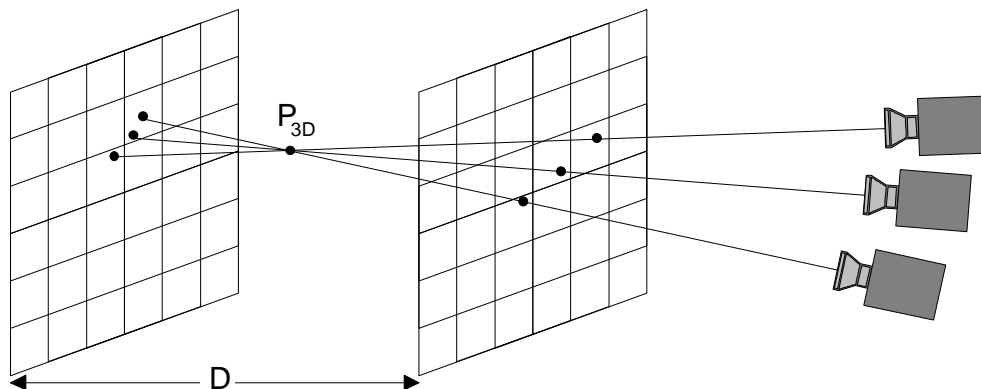


Abbildung 10: Berechnung von 3D-Punktkoordinaten aus mehreren Sehstrahlen

Voraussetzung für eine hohe Rekonstruktionsgenauigkeit ist hierbei eine hohe Wiederholgenauigkeit der Positionierung der Kamera bei der Aufnahme der Referenz- und Objektbilder. Diese Anforderung ist mit dem CaRo-Ansatz leicht zu erfüllen, da der verwendete Roboterarm über eine sehr hohe Wiederholgenauigkeit verfügt.

4.6 Kombination von Textur und Geometrie: Das CaRo-Datenformat

Nach den oben geschilderten Verfahren sollte es möglich sein, beliebig feine Triangulationen der Form von Objekten zu digitalisieren. Wenn dies geleistet ist, so kann nun quasi für jedes einzelne Dreieck der Triangulation auch Farbe und Textur aufgenommen werden [Niem, Wingbermhühle 97]. Da die Geometrie des zu digitalisierenden Objektes nach der Triangulation hinreichend genau bekannt ist, kann die Kamera approximativ orthogonal zur einzelnen Dreiecksfläche an diese heranfahren und in geeignetem Abstand Farbaufnahmen erzeugen, aus denen direkt und genau das Flächenstück ausgeschnitten werden kann, das dem Zieldreieck auf der Objektoberfläche entspricht. Das CaRo-3D-Datenformat, das die Digitalisiersoftware liefern soll, kann also so aufgebaut werden: Triangulationsnetz vorgegebener Feinheit über die gesamte Objektoberfläche, falls gewünscht, zusätzlich schwächer oder höher aufgelöste Oberflächenbilder im Dreiecksformat, die jeweils den Dreiecken der Triangulation entsprechen.

Da Kanten von Objekten sehr genau repräsentiert werden sollten, muß man bei Kantenverdacht versuchen, die Triangulationspunkte genau auf die Kanten zu

legen. Kanten müssen nicht optisch, sondern können durch Glattheitsanalyse der Triangulation erkannt werden, was durchsichtiger zu implementieren ist als wenn man Schattierungseigenschaften oder ähnliches zu Hilfe zieht. In unserem Forschungsprogramm werden wir aber sicher auch Schattierungsanalysen durchführen, um auch abgerundete Kanten genauer erfassen zu können.

Zur optischen Prüfung dessen, was die Digitalisiersoftware geliefert hat, können alle Verfahren der Bilderzeugung von der Vektorgraphik über die Tiefenpuffer-Verfahren bis zum Raytracing oder Radiosity-Verfahren angewandt werden. Es müssen als Geometrie-Primitive nur Dreiecke verarbeitbar sein. Wie man die Texturbilder bzw. die aus ihnen leicht berechenbare Mittelfarbe der einzelnen Dreiecke abbilden will, bleibt dem Anwender überlassen. Mit gewöhnlichem Texturmapping können jedenfalls schon perfekte und realistische Abbildungen erzeugt werden, auch wenn die digitalisierten Objekte z.B. kleine Beschriftungen und Bildaufdrucke tragen.

Aber man kann auch noch einen Schritt weitergehen und zusätzlich die Reflexivität, also das Lichtreflexionsverhalten von Oberflächendreiecken, bestimmen lassen. Die Kamera müßte zu diesem sicherlich sehr speziellen Zweck mit einer schwenkbaren Lichtquelle etwa wie in Abbildung 11 angedeutet ausgerüstet werden.

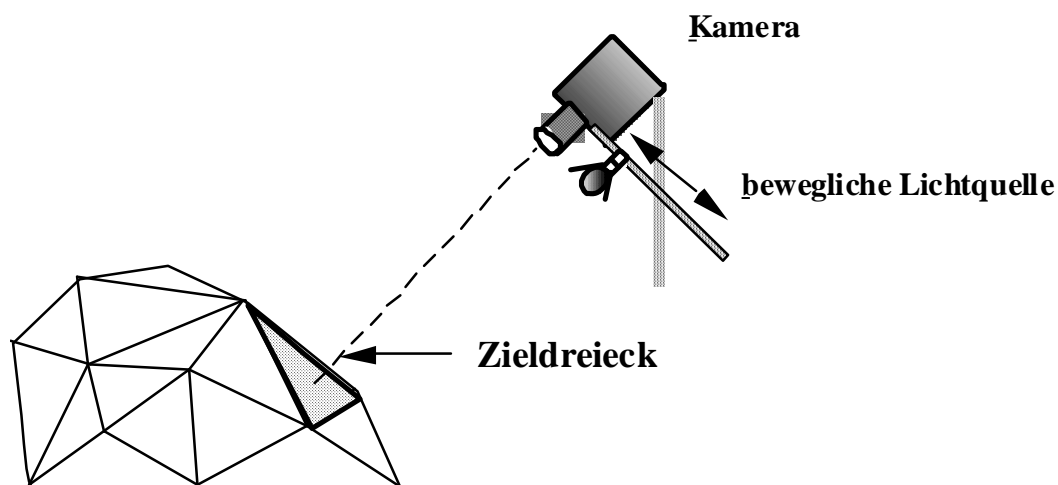


Abbildung 11: Selektive Lichtreflexions-Messung mit beweglicher Lichtquelle an Kamera

Durch mit unterschiedlichen Winkeln auf die Oberfläche fallendes Licht und den sich dabei ergebenden Änderungen der Bildhelligkeiten kann vollautomatisch festgestellt werden, ob sich die Oberfläche ganz matt verhält oder sehr stark spiegelt. Der entsprechende Reflexionskoeffizient wäre so mit für die fotorealistische Bilderzeugung ausreichender Genauigkeit meßbar.

4.7 Semiplanare Objekte

Das bereits oben angesprochene Problem, leicht gekrümmte Buchseiten sehr genau abzutasten, um ein präzises, letztendlich planares Pixelbild zu erstellen,

kann mit den oben erläuterten Methoden angegangen werden. Allerdings steht jetzt nicht mehr die Geometrie-Erfassung im Vordergrund, es soll vielmehr die ihrer Natur nach planare Textur rekonstruiert werden. Daher kommt es eher darauf an, mehrere sich überlappende Einzelbilder der Buchseite nahtlos, übergangsfrei und entzerrt aneinanderzufügen. Das gilt auch für das Scannen großformatiger Vorlagen, die schon ziemlich plan liegen.

Eine vollautomatisch arbeitende Buchdigitalisiermaschine wäre sehr hilfreich, um das in älteren Büchern schlummernde Wissen in die zukünftige volldigitalisierte Informationswelt hinüberzuretten. Das CaRo-Projekt kann dazu einen Beitrag leisten. Es fehlt allerdings noch eine zuverlässig arbeitende Seitenumblättermaschine, um auch noch diese letzte Handarbeit überflüssig zu machen. Insbesondere die Möglichkeit, eine kritische Textstelle in Zusammenarbeit mit der OCR-Software mehrmals und mit wachsender Pixelauflösung zu inspizieren, verspricht deutlich erhöhte OCR-Erkennungsleistung. Im Zweifelsfall, wenn also Mustererkennung mit Präzisionsscanning nicht weiterhilft, kann ein kleiner Bereich als Pixelmap in das endgültige Text-Dokument aufgenommen werden. So muß ja auch bei allen Abbildungen verfahren werden.

4.8 Sichtsystem-Anwendungen im Kleinen

Die CaRo-Hardware kann nicht nur für die Objektrekonstruktion, sondern auch als Sehhilfe genutzt werden. In einer Studienarbeit [Holfelder 97] haben wir bereits mit bewegter Kamerafahrt Stereo-Videosequenzen von Stilleben-Szenen aufgenommen. Auf einem Stereobildschirm (z.B. einer SiliconGraphics Workstation) kann man dann mit einer Shutterbrille ausgerüstet die Videos mit hervorragendem räumlichem 3D-Eindruck anschauen. Die CaRo-Kamera erzeugt dabei ein Stereobildpaar so, daß die Kamera von zwei verschiedenen Positionen, also z. B. im Augenabstand, jeweils ein Stereobildpaar aufnimmt.

Weitere Anwendungen in dieser Richtung bieten sich an, wenn Wissenschaftler über Kontinente hinweg ein interessantes Objekt konferenzartig diskutieren wollen. Über das Internet könnte die Kamera über große Entfernungen gezielt positioniert werden, um z. B. auffällige oder interessierende Detail bei Bedarf auch stark zu vergrößern.

5. Ausblick

Mit dem CaRo-Projekt wird eine universelle und flexible Technik für das Digitalisieren von 3D-Objekten angestrebt. Die Technik, eine Kamera robotergesteuert zu bewegen, scheint wegen der flexiblen Verfahrensweisen sehr erfolgversprechend zu sein. Erstmals werden dabei typische Eigenschaften des menschlichen Sehens kombiniert, nämlich das foveale, scharf auf einen Punkt gerichtete sowie das Übersichtssehen. Es handelt sich aber nicht um eine Low-Cost-Lösung, denn sowohl die Hardware, als auch die Software sind von

erheblicher Komplexität. Die Hardware kann jedoch leicht aus kommerziell erhältlichen Komponenten aufgebaut werden und erfordert kaum Spezialentwicklungen.

Welche Rekonstruktionsgenauigkeit mit einem gewöhnlichen Industrieroboter erreicht werden kann, ist derzeit noch nicht vorhersagbar. Allerdings gibt es eine Reihe von Möglichkeiten, die Ergebnisse durch Ausgleichsrechnungen erheblich zu verbessern, z.B. durch Methoden, die in der Photogrammetrie entwickelt wurden. Mit sehr genau positionierenden Robotern kann die Digitalisiergenauigkeit entsprechend verbessert werden, eine ebenfalls sehr wichtige Option.

Bemerkenswert ist die Vielfalt der Rekonstruktionsverfahren, die im einzelnen realisiert und untersucht werden können. Praktisch jedes bisher vorgeschlagene Kamera-basierte 3D-Rekonstruktionsverfahren ist im Prinzip simulierbar, wobei die bereits genau bekannten Kamerapositionen die Rekonstruktionsverfahren erheblich vereinfachen.

Das System ist unabhängig von der Software größenskalierbar. Mit sehr kleinen Robotern kann in den Bereich der Feinmechanik vorgedrungen werden, mit großen Robotern in Portalbauweise können auch größere Werkstücke digitalisiert werden.

Wir haben damit die Hoffnung, ein allgemein einsetzbares Verfahren von großer Praxisrelevanz gefunden zu haben, welches die Einschleuseprobleme der graphischen Datenverarbeitung in einem umfassenden Sinne anzugehen gestattet.

Literatur

[Breuckmann 93]

B. Breuckmann: Bildverarbeitung und optische Meßtechnik. Franzis-Verlag GmbH, München, 1993.

[Curless, Levoy 96]

B. Curless, M. Levoy: A Volumetric Method for Building Complex Models from Range Images. SIGGRAPH 96 Conference Proceedings, Annual Conference Series, Addison Wesley, August 1996, pp. 303-312.

[Delherm et al. 96]

C. Delherm, J.M. Lavest, M. Dhome, J.T Lapresté: Dense Reconstruction by Zooming. In B. Buxton, R. Cipolla (Hrsg.), Computer Vision - ECCV '96, Proceedings, Lecture Notes in Computer Science 1065, Springer Verlag, 1996, pp. 427-438.

[Faugeras 93]

O. Faugeras: Three-Dimensional Computer Vision, a Geometric Viewpoint. MIT Press, Cambridge, Massachusetts (u.a.), 1993.

[Holfelder 97]

H. Holfelder: Monokulares Stereo-Sehen. Studienarbeit, Fakultät für Informatik, Universität Karlsruhe 1997.

[Jain et al. 95]

R. Jain, R. Kasturi, B. Schunck: Machine vision. McGraw-Hill, New York (u.a.), 1995.

[Jarvis 83]

R. A. Jarvis: A Perspective on Range Finding Techniques for Computer Vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, 5(2), 1983, pp. 122-139.

[Leister 91]

W. Leister: Geometrisches Modellieren durch interaktive Rekonstruktionsmethoden. Dissertation, Fakultät für Informatik, Universität Karlsruhe 1991.

[Li, Lavest 95]

M. Li, J.M. Lavest: Some aspects of zoom-lens camera calibration. Technical Report, Computational Vision and Active Perception Laboratory Royal Institute of Technology, Stockholm, Sweden, Number ISRN KTH/NA/P-95/03--SE, February 1995.

[Martin, Aggarwal 83]

W. N. Martin, J. K. Aggarwal: Volumetric Descriptions of Objects From Multiple Views. IEEE Transactions on Pattern Analysis and Machine Intelligence 5(2), 1983, pp. 150-158.

[Martins et al. 81]

H.A. Martins, J.R. Birk, R.B. Kelly: Camera Models Based on Data from Two Calibration Planes. *Computer Graphics and Image Processing* 17, 1981, pp. 173-180.

[Niem, Wingbermühle 97]

W. Niem, J. Wingbermühle:

Automatic Reconstruction of 3D Objects Using a Mobile Monoscopic Camera. *Proceedings of the International Conference on Recent Advances in 3D Imaging and Modelling*, Ottawa, Canada, 12-15 May 1997.

[Schmitt 83]

A. A. Schmitt: Dialogsysteme - Kommunikative Schnittstellen, Software-Ergonomie und Systemgestaltung. BI Wissenschaftsverlag, Mannheim 1983.

[Tarabanis et al. 92]

K. Tarabanis, R.Y. Tsai, D.S. Goodman: Modelling of a Computer-Controlled Zoom Lens. *IEEE International Conference on Robotics and Automation*, 1992, pp. 1545-1551.

[Tsai 86]

R.Y. Tsai: An Efficient and Accurate Camera Calibration Technique for 3-D Machine Vision. *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, 22-26 Juni 1986, pp. 364-374.

[Watanabe 96]

M. Watanabe, S.K. Nayar: Telecentric Optics for Computational Vision. In B. Buxton, R. Cipolla (Hrsg.), *Computer Vision - ECCV '96, Proceedings, Lecture Notes in Computer Science 1065*, Springer Verlag, 1996, pp. 439-454.

[Will, Pennington 72]

P. Will, K. Pennington: Grid Coding, A Novel Technique for Image Processing. *Proceedings of the IEEE* 60(6), 1972, pp. 669-680.