

Integration von Sensor und Aktor in einem aktiven binokularen Sehsystem

Michael Hansen¹, Gerd Diesner, Henrik Schmidt, Gerald Sommer
Lehrstuhl für Kognitive Systeme, Christian-Albrechts-Universität zu Kiel

Zusammenfassung

Auf dem Gebiet des maschinellen Sehens gewinnt das aktive Sehen als Methodik zur Lösung visueller Aufgaben zunehmend an Bedeutung, bietet doch die Variation der Kameraparameter wie Position, Vergenz, Zoom oder Fokus die Möglichkeit, auf die visuellen Stimuli aktiv und zielgerichtet zu reagieren. In der Arbeitsgruppe wird am verhaltensbasierten Entwurf eines aufmerksamkeitsgesteuerten visuellen Systems gearbeitet. Neben Arbeiten an einem Attentionsmechanismus werden neuronale und Reinforcement Lernverfahren entwickelt.

Im Bereich Okulomotorik wird eine binokulare Kameraeinheit eingesetzt. Ein binokulares regelbares Kamerasystem als aktiver Sensor zur zielgerichteten Informationsgewinnung muß die Fähigkeit besitzen, verschiedene Signalarten und Systemkomponenten in einem Regelkreis zu integrieren.

Am Beispiel einer monokularen Kameranachführung als visuelles Verhaltensmodul werden das Funktionieren, aber auch die Schwachstellen dieses Regelkreises aufgezeigt, die weniger in der Kommunikation der Systemkomponenten untereinander als vielmehr in der Transformation Bilddaten - Stellgrößen liegen.

In einem Ausblick werden Ansätze zu weiteren visuellen Verhaltensmodulen beschrieben, die das okulomotorische Verhalten des aufmerksamkeitsgesteuerten visuellen Systems bilden werden.

1 Aktives Sehen

In den letzten Jahren hat sich auf dem Gebiet des maschinellen Sehens infolge steigender Rechnerleistung und der Entwicklung von kleinen CCD-Kameras die Möglichkeit ergeben,

neue Wege bei der Bildgewinnung durch aktive Steuerung der Kameraparameter zu gehen. Die klassischen Ansätze im maschinellen Sehen basierten auf der Kamera als passiven Sensor, der in den Prozeß des Sehens nicht eingreift. Die Bildauswertung lieferte dann die Beantwortung der Frage „Was ist wo?“, auf die *Marr* die Aufgabe des maschinellen Sehens reduzierte und die eine umfangreiche Modellierung der Welt voraussetzte (aus [6]). Von *Bajcsy* [3] wurde die Kamera als aktiver Sensor, der in den Bildgewinnungsprozeß eingreift und dadurch das Sehen und Erkennen aus den Bilddaten beeinflusst, eingeführt. Die Modellierung und Untersuchung geeigneter Steuerungsstrategien zur visuellen Wahrnehmung stehen bei ihr im Vordergrund des aktiven Sehens. *Aloimonos* [1] zeigt, daß unterbestimmte, nichtlineare Aufgaben wie die Formerkennung aus Schattierung durch aktives Sehen regularisiert und linearisiert werden können. Von *Ballard* und *Brown* [2] wird der Begriff des „Animate Vision“ eingeführt, der eine Verbindung des aktiven Sehens zum menschlichen Sehen knüpft. Das Lernen gehört neben der Blicksteuerung zu einem der drei Prinzipien des „Animate Vision“. Die Notwendigkeit zum Lernen ergibt sich in ihrem Sinne aus der Unvorhersagbarkeit der Welt, in der der Prozeß des Sehens stattfindet. Eine Zusammenstellung der verschiedenen Aspekte des aktiven Sehens findet sich in [11].

Welche Beispiele für aktives Sehen zur Lösung von visuellen Aufgaben gibt es? Für ein aktives Sehsystem, das über ähnliche visuelle Fähigkeiten wie der Mensch verfügen soll, ist die Steuerung der Blickrichtung zur Erweiterung des Sehfeldes unerlässlich. Erst dadurch ist es in der Lage 'herumzusehen', um die Aufgabe der Exploration in einer unbekanntem Umgebung lösen zu können. Zusätzlich ist durch eine Steuerung der Kamerablende eine Anpassung an die Beleuchtungsbedingungen möglich.

¹Institut für Informatik, Preußerstr.1-9, 24105 Kiel, Tel: 0431/56 04 72.

Ein Vorteil kann sich auch aus der Steuerbarkeit der Brennweite ergeben. Wird zum Beispiel ein sich bewegendes Objekt durch die Steuerung der Kameraposition im Bildzentrum gehalten, so kann durch Steuerung der Brennweite die Größe der Abbildung konstant bleiben und sich die Wiedererkennung des Objektes vereinfachen, auch wenn sich das Objekt von der Kamera entfernt.

Wir beschäftigen uns mit der Entwicklung eines künstlichen Systems zur aktiven visuellen Wahrnehmung. Dabei steht in einem Teilbereich die Bereitstellung der visuellen Fähigkeiten, wie Sakkaden, Augenfolgebewegung, Akkomodation und Vergenzbewegung, sowie die Untersuchung deren Vorteile bez. visueller Wahrnehmung im Vordergrund. Ein anderer Bereich umfaßt die Erlangung der Lernfähigkeit für ein solches System, um es in die Lage zu versetzen, sich adaptiv an die Realwelt anzupassen und optimal der Aufgabe entsprechend zu verhalten. Dazu werden neuronale und Reinforcement Lernverfahren untersucht. Die Entwicklung eines Attentionsmechanismus, der entsprechende Verhalten des Systems auslösen kann oder selbst ein Verhaltensmodul *Attention* des Systems darstellt, ist ein weiterer Bereich der Forschung. In diesem Manuskript wird jedoch die Integration von einem Sensor und einem Aktor zu einem aktiven Sehsystem beschrieben.

2 Das aktive binokulare Kamerasystem

Unser verfügbares aktives Sehsystem besteht aus folgenden Komponenten, die über Bus- und Signalverbindungen miteinander agieren müssen, um dem System visuelle Fähigkeiten zu ermöglichen:

- Die binokulare Kameraeinheit BiSight der Firma TRC.
- Die Motorensteuerkarte PMAC der Firma DELTA TAU.
- Der Pipeline-Rechner MaxVideo 200 der Firma DATACUBE.

- Eine Workstation Sparc Station 10 der Firma SUN.

Die Kommunikation der Komponenten untereinander erfolgt über den VME-Bus. Lediglich die Motorensteuerkarte und die Kameraeinheit kommunizieren nicht über den Bus sondern über analoge und digitale Signalwege.

2.1 Die binokulare Kameraeinheit BiSight

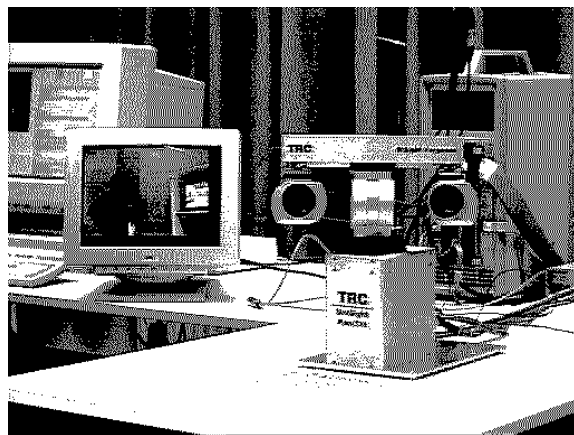


Abb.1: Das aktive binokulare Kamerasystem.

Die Kameraeinheit verfügt über einen anthropomorphen Aufbau, d.h. sie besitzt dem Menschen nachempfundene Freiheitsgrade. Dieser Aufbau eignet sich für die Nachahmung des okulomotorischen System des Menschen, da es die vier mechanischen Freiheitsgrade Schwenk (pan), Neigung (tilt) und Vergenz (rechts/links) ermöglichen, den Augenbewegungen ähnliche Kamerabewegungen durchzuführen. Lösungen visueller Aufgaben lassen sich am menschlichen Vorbild orientieren. Eine wesentliche Vereinfachung von Koordinatentransformationen ergibt sich dadurch, daß sich die Rotationsachsen von Vergenz und Neigung im optischen Zentrum der Kamera schneiden.

| | Schwenk | Neigung | Vergenz |
|--------------|---------------------|---------------------|----------------------|
| max.Winkel | $\pm 160^\circ$ | $\pm 90^\circ$ | $\pm 45^\circ$ |
| Auflösung | $0,0225^\circ$ | $0,0225^\circ$ | $0,006^\circ$ |
| Max. Geschw. | $650^\circ s^{-1}$ | $500^\circ s^{-1}$ | $1000^\circ s^{-1}$ |
| Max. Beschl. | $1350^\circ s^{-2}$ | $1350^\circ s^{-2}$ | $12000^\circ s^{-2}$ |

Tab.1: Technische Daten der Kameraeinheit.

Die maximalen Winkel (siehe Tab.1) entsprechen denen eines Menschen bei Hals- und Augenbewegungen. Der Abstand der beiden CCD-Kameras beträgt 25 cm. Der Mensch hat im Vergleich einen kleineren Augenabstand (6-8 cm). Er erreicht beim sprunghaften Blickwechsel, den Sakkaden, eine Winkelgeschwindigkeit von bis zu $800^{\circ}s^{-1}$ [13].

Die optischen Freiheitsgrade der Kameras bilden das Zoom (Brennweite 11,5mm-69mm), der Fokus (1m- ∞) unter Closed-Loop-Regelung und die Iris in Open-Loop-Steuerung. Je nach Zoomstellung ergibt sich ein Bildwinkel von 7° bis 51° (diagonal).

Die Beschreibung der weiteren Systemkomponenten findet sich im technischen Anhang.

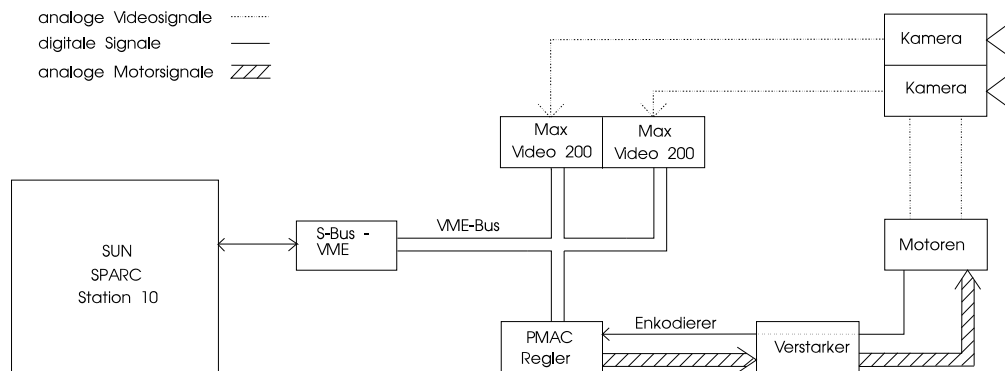


Abb.2: Das aktive binokulare Sehsystem im Übersichtsschema der Kommunikationsverbindungen.

Die Kommunikation der einzelnen Komponenten erfolgt gemäß Abb. 2. Im einzelnen sind die Schnittstellen wie folgt realisiert:

- Die Digitalisierung des Videosignals auf dem MaxVideo 200 System in Echtzeit liefert ein 512x480 Pixel Bild. Die volle Auflösung der Kamera wird horizontal noch nicht erreicht.
- Die Abtastung der Encodersignale auf der Steuerkarte (PMAC-Regler) ist in verschiedenen Modi möglich. Auf die Encoderwerte kann von der Workstation aus zugegriffen werden.
- Die Generierung der Steuersignale erfolgt auf der Motorensteuerungskarte durch einen 16-Bit D/A-Konverter.

3 Eine Anwendung: Monokulare Kameranachführung

Ein Anwendungsbeispiel, das die Funktionsweise des Regelkreises Kameranachführung—Bildverarbeitung—Kamerasystem verdeutli-

chen soll, ist das Verhalten der Kameranachführung zur Verfolgung einer Lichtquelle. Das menschliche Analogon zur Kameranachführung ist die Augenfolgebewegung, die meist mit einer Sakkade auf den zu verfolgenden Punkt initialisiert wird. Die Augenfolgebewegung stabilisiert das Bild auf der Netzhaut und hält den attendierten Punkt in der Fovea des Auges, dem Bereich der höchsten retinalen Auflösung. Wird die Objektgeschwindigkeit für eine gleichmäßige Augenbewegung zu groß ($> 100 - 200^{\circ}s^{-1}$), wird die Lichtquelle mit einer folgenden Sakkade wieder 'eingeholt'[13].

Das Modul *monokulare Kameranachführung* gliedert sich in die Bildverarbeitung und die Kamerasteuerung.

3.1 Bildverarbeitung

Die Bildverarbeitung besteht bei der Lichtquellenverfolgung aus der Binarisierung des 256 Graustufen umfassenden Kamerabildes $f(i, j)$. Sei h ein geeignet gewählter Schwellwert, um die Lichtquelle im Bild zu detektieren. Dann

gilt für das Binärbild $b(i, j)$:

$$b(i, j) = \begin{cases} 0 & f(i, j) \leq h \\ 255 & f(i, j) > h \end{cases}$$

Die Binarisierung und die Schwerpunktberechnung erfolgt in Echtzeit ($t < 33$ ms) auf dem MaxVideo 200 System. Das Schwerpunktkoordinatenpaar wird der Workstation übergeben. In einem gemeinsamen Speicherbereich (shared memory) steht es dem Programm zur Kamerasteuerung zur Verfügung.

3.2 Kamerasteuerung

In diesem Beispiel der monokularen Kameranachführung werden nur die Neigungsachse und eine Vergenzachse zur Regelung verwendet. Das Kugelkoordinatensystem (χ, ϕ, R) habe seinen Ursprung im optischen Zentrum der Kamera. Drehungen um die Neigungsachse werden durch den Winkel ϕ und Drehungen um die Vergenzachse durch den Winkel χ beschrieben. Der Vektor $(0, 0, R)$ sei parallel zur optischen Achse der Kamera gewählt. Der Abstand R kann bei der Steuerung der Blickrichtung vernachlässigt werden.

Der Algorithmus zur Regelung der Kameranachführung gliedert sich in folgende Teilschritte:

1. Digitalisierung des Kamerabildes.
2. Binarisierung des Bildes und Detektion der Lichtquelle.
3. Berechnung der Schwerpunktkoordinaten der Lichtquelle und Bestimmung deren Abweichung vom Bildmittelpunkt (x_t, y_t) .
4. Berechnung der Kugelkoordinaten (χ_t, ϕ_t) des Schwerpunktes aus den Bildkoordinaten (x_t, y_t) zum Zeitpunkt t .
5. Abschätzung der Position des Schwerpunktes (χ_{t+1}, ϕ_{t+1}) zum Zeitpunkt $t + 1$ unter der Annahme gleichbleibender Beschleunigung.
6. Anfahren der Position (χ_{t+1}, ϕ_{t+1}) als neue Kameraposition.

Die Digitalisierung und Binarisierung erfolgt in Echtzeit auf dem Pipelinerechnersystem. Die Schwerpunktberechnung wird mit dem Hardwaremodul *Houghtransformation* realisiert.

Nach der Umrechnung des Schwerpunktes in Kugelkoordinaten (χ_t, ϕ_t) erfolgt die Abschätzung der möglichen Schwerpunktposition zum Zeitpunkt $(t + 1)$ gemäß folgendem Ansatz:

Die Bahnfunktion des Schwerpunktes im Raum wird mit Hilfe einer Taylorentwicklung zweiter Ordnung extrapoliert. D.h. es wird angenommen, daß sich der Schwerpunkt gleichförmig oder gleichförmig beschleunigt bewegt. Es ergibt sich für die abgeschätzte Schwerpunktposition:

$$\begin{aligned} \chi_{t+1} &= \chi_t + (\chi_t - \chi_{t-1}) + \frac{1}{2}(\chi_t - 2\chi_{t-1} + \chi_{t-2}) \\ \phi_{t+1} &= \phi_t + (\phi_t - \phi_{t-1}) + \frac{1}{2}(\phi_t - 2\phi_{t-1} + \phi_{t-2}). \end{aligned}$$

Für die Initialisierung bis zum Zeitpunkt $t = 2$ gilt:

1. $(\chi_1, \phi_1) = (\chi_0, \phi_0)$. Dies entspricht einer Sakkade auf den detektierten Schwerpunkt.
2. $\chi_2 = \chi_1 + (\chi_1 - \chi_0)$ (analog für ϕ_2).

3.3 Auftretende Probleme, experimentelle Resultate

Die bisherigen Experimente zeigen, daß die präzise Ermittlung der Achsenstellwinkel (χ, ϕ) aus den Bildkoordinaten (x_s, y_s) des Objektschwerpunktes für die Genauigkeit der Kameraregelung entscheidend ist. Eine ungenaue Schwerpunktberechnung oder eine Ungenauigkeit der Kameraposition spielen eine geringe Rolle.

Der Schwerpunkt der Lichtquelle wird mit einer Genauigkeit von unter einem Pixel detektiert und berechnet. Das Anfahren einer bestimmten Position entsprechend der Achsenstellwinkel ist durch die Closed-Loop-Regelung exakt, führt aber je nach geforderter Genauigkeit zu einer langen Bewegungsphase der Kamera. So dauert eine Bewegung mit einer Genauigkeit

von 50 Encoderwerten etwa 300 Millisekunden, während sich die Dauer bei einer Genauigkeit von 1 Encoderwert etwa verdoppelt.

Quantitative Untersuchungen zur Genauigkeit der Kameranachführung sind bisher nicht durchgeführt worden, da noch keine reproduzierbaren Lichtquellenbewegungen möglich sind. Die Untersuchung der Genauigkeit einer Sakkade auf eine ruhende Lichtquelle ergab abhängig von der Distanz und der Sakkadenrichtung eine Ungenauigkeit in der Größenordnung von 2–10 Pixel

Der Achsenstellwinkel χ berechnet sich für Sakkaden entlang der x -Achse des Bildkoordinatensystem zu: $\chi = \arctan(s_x(x_s - x_m)/f)$.

Dabei ist x_m die x -Koordinate des Kamerabildzentrums, f die Brennweite und s_x der Skalierungsfaktor der Kamera. Ein Vergleich des theoretischen Fehlers $\Delta\chi$ und des experimentellen Fehlerverlaufes läßt den Schluß auf eine ungenaue Bestimmung der internen Kameraparameter zu.

Um diese bestimmen zu können, wird eine Kalibrierung mit einem Würfel bekannter Abmessungen vorgenommen. Dieses Verfahren verwendet die Kenntnis der 3-D Koordinaten bez. eines festen Koordinatensystems und der Bildkoordinaten der Punkte, die sich auf dem Würfel befinden. Über ein Optimierungsverfahren läßt sich eine Matrix berechnen, die die Transformation Welt-Bildkoordinaten beschreibt und die internen Kameraparameter enthält.

Eine ungenaue Orientierung des CCD-Chip in der Kamera ist eine weitere Fehlerquelle. Ist der Chip nicht parallel zur Neigungsachse ausgerichtet, erhält man bei horizontalen Sakkaden eine Abweichung entlang der Vertikalen. Als Lösung kommt zur Zeit nur eine manuelle Ausrichtung der Orientierung des CCD-Chip in Frage. Diese kann nach Reduzierung der Abweichung durch ungenaue interne Parameter mit Hilfe von Sakkaden überprüft werden. Bessere Kalibrierungsverfahren würden sich durch eine gesteuerte Torsionsbewegung der Kamera ergeben, wie sie beim KTH-Kamerakopf von Eklundh möglich ist [5].

Darüberhinaus müssen auch die Annahmen

über die wirkliche Lage der Rotationsachsen und des optischen Zentrums überprüft werden, um ein optimiertes kinematisches Kameramodell zu erhalten und die visuelle Rückkopplung des Regelkreises zur Steuerung effektiv ausnutzen zu können. Gleichzeitig stellt sich jedoch die Frage nach Algorithmen, die keiner umfangreichen Kalibrierung bedürfen, sondern sich aus einem Lernprozeß heraus optimieren.

Trotz der bisher nicht optimalen Kalibrierung ist das Kamerasystem in der Lage, die Kamera auf eine Lichtquelle nachzuführen. Die bisher erreichte Verarbeitungsgeschwindigkeit liegt, bedingt durch den im Moment genutzten Bewegungsmodus der Steuerkarte, bei 3 Kamerabewegungen pro Sekunde. Entsprechend kritisch reagiert das Modul auf schnelle Richtungsänderungen der Lichtquelle. Bei extremen Änderungen kann die Lichtquelle auch ganz aus dem Blickfeld der Kamera verschwinden.

Gegenstand der laufenden Arbeit ist, neben der Optimierung und einer Untersuchung der zeitlichen Verzögerungen des bestehenden Moduls, die Erweiterung zu einem binokularen Modul Kameranachführung. Dabei ist nicht an die parallele Implementierung des gleichen monokularen Algorithmus gedacht, sondern an eine Vorgehensweise, bei der die Regelung der zweiten Kamera, sich der Geometrie der Kameraanordnung und der Informationen der ersten Kamera bedient. Dies ist der menschlichen Augenfolgebewegung ähnlich, bei der ein Auge dominant ist und das andere Auge motorisch nachgeführt wird [13].

4 Weitere visuelle Verhaltensmodule

Ein Sehsystem muß über visuelle Verhaltensmodule verfügen, die in geeigneter Weise miteinander agieren und aufeinander reagieren können, um autonomes Verhalten zeigen zu können. Neben dem Modul Kameranachführung ist ein Modul zur Sakkadenregelung notwendig, da das Verhalten der Kameranachführung, das dem der Augenfolgebewegung entspricht, durch eine Initialsakkade eingeleitet wird [10].

Dabei ist auch eine neuronale Sakkadenregelung mittels der in der Gruppe von Bruske entwickelten *Dynamischen Zellstrukturen* [4] vorgesehen.

Die Bestimmung der Entfernung von Punkten und Objekten ist eine visuelle Fähigkeit, die bei einem mobilen autonomen Sehsystem zur Hindernisvermeidung wichtig ist. Ohne eine genaue Entfernungsbestimmung ist die visuelle Navigation eines mobilen Systems unmöglich. Eine hohe Genauigkeit ist auch zur visuell basierten Steuerung eines Manipulators erforderlich. Beim Greifen muß die Pose des zu greifenden Objektes bekannt sein, da sich sonst nur schwer geeignete Greifpunkte finden lassen. Mit einem binokularem Kamerasystem kann aufgrund der Horizontaldisparität von korrespondierenden Punkten des Bildpaares die Entfernung mittels Triangulation bestimmt werden. Die Entfernungsbestimmung aus Stereobildpaaren ist im Nahbereich (Bis 3 m, abhängig von der Basislinie) aufgrund der großen auftretenden Disparitäten ein geeignetes Verfahren. Dabei stellt sich in der Praxis das Auffinden der korrespondierenden Punkte als nicht allgemein lösbares Problem dar. In der Literatur wird zwischen merkmalsbasierten und regionenbasierten Verfahren zur Lösung des *Korrespondenzproblems* unterschieden. Werden bei merkmalsbasierten Verfahren Merkmale, meist Grauwertkanten und Linien, aus den Kamerabildern extrahiert und zugeordnet, so werden in regionenbasierten Verfahren Korrelationen über die Grauwerte berechnet. Über 200 verschiedene Verfahren zur Bestimmung der Entfernung aus Stereoverfahren sind in der Literatur bekannt [7].

Neben Verfahren, die die Korrelationen auf den sog. Epipolargeraden der Kameraanordnung berechnen, um den korrespondierenden Punkt zu finden, bietet sich das Verfahren von Mallot und Theimer an [9], das über Phasenantworten von Gaborfiltern die Entfernung berechnet und keine explizite Lösung des Korrespondenzproblem erfordert.

Ferner wird untersucht, ob mit einem Verfahren, das auf einen Punkt vergiert, verbesserte Ergebnisse erzielt werden können, da das Kamerasystem im Verhältnis zu anderen Sy-

stemem über eine gute Vergenzwinkelauflösung verfügt. Ein Punkt, auf den die Kameras vergieren, besitzt keine Querdisparität. Über die Vergenzwinkel erhält man die Entfernung dieses Punktes.

Ein weiteres Verfahren der Entfernungsbestimmung ergibt sich durch Akkomodation. Dieses Teilmodul wird über die Scharfstellung eines Bildausschnittes geregelt. Es gibt eine Vielzahl von Funktionen, die die Schärfe eines Bildausschnittes bewerten [12]. Meist nutzen diese Funktionen den hochfrequenten Anteil eines Bildes aus, der Grauwertkanten repräsentiert. Der Fokus wird solange variiert, bis das Maximum der Schärfefunktion gefunden ist. Da die Fokusbewegung zeitaufwendig ist (6 Sekunden für einen vollen Durchlauf), müssen schnelle Algorithmen zur Bestimmung des Maximums verwendet werden. Über die Bildweite kann gemäß der Formel für dünne Linsen die Entfernung zu einem Punkt bei bekannter Brennweite f berechnet werden.

Die Kombination der Berechnung aus Stereobildpaaren und Akkomodation lassen ein stabiles Verhalten des Moduls Entfernungsbestimmung erwarten [8].

5 Technischer Anhang

5.1 Die Motorensteuerkarte PMAC

Die Motorensteuerkarte PMAC ist eine sehr umfangreich einzusetzende Karte, die in Verbindung mit unterschiedlichsten Sensoren, Signalverstärkern und Motoren verwendet werden kann. Sie ist mit einer eigenen CPU bestückt und kann über den VME-Bus, den Standard-Bus und die serielle Schnittstelle (RS232, RS422) mit anderen Rechnern verbunden werden. Die Karte ermöglicht die Steuerung von 8 Motoren unter Closed-Loop-Regelung und zwei Motoren unter Open-Loop-Regelung, sowie die Verwendung von bis zu 3 Encoderkanälen pro Motor. Der Analogkreis zur Leistungsteuerung über die Signalverstärker ist vom digitalen Kreis zur Motorensteuerung optisch entkoppelt, um eine Rauschreduzierung zu erreichen.

Für die Programmierung der Motorensteuerung stehen eine Vielzahl von Registervariablen zur Verfügung, mit denen die Position, Geschwindigkeit und Beschleunigung der Motoren geregelt werden können. Die Regelung wird in sog. Bewegungsprogrammen festgelegt, in denen auch S-Kurven für den Geschwindigkeitsverlauf eines Motors definiert werden können. Sowohl über externe als auch über den internen Timer ist die Echtzeitfähigkeit möglich. Dabei ist vom Programmierer selbst auf die Ausführungszeit der Bewegungsbefehle zu achten.

5.2 Der Pipelinerechner MaxVideo 200

Die MaxVideo 200 Karte ist ein leistungsfähiges und echtzeitfähiges Bildverarbeitungssystem. Es enthält einen 8-Bit A/D-Wandler (26 MHz Pixelrate), eine lineare ALU, 24 MB Speicher und verschiedene Hardwaremodule für Faltungen, Verzerrungen (warper) und morphologische Operationen. Über den sog. Max-Bus (8-Bit) können weitere Hardwaremodule aufgerüstet werden. Die Taktfrequenz beträgt 20 MHz für alle Hardwaremodule und 40 Mhz für den Speicher. Die Programmierung erfolgt durch die Konfiguration von sog. Pipes. Eine Pipe beschreibt den Verarbeitungsfluß durch die gewünschten Hardwaremodule. Nach der Konfiguration wird die Pipe gestartet, d.h. die Bilddaten werden entsprechend verarbeitet. Eine Pipe kann kontinuierlich oder auch nur einmal von den Bilddaten durchlaufen werden. Zur Darstellung des Kamerabildes auf einem Monitor ist das Abarbeiten einer kontinuierlichen Pipe erforderlich.

5.3 Die Sparc 10 Workstation

Eine Sparc 10 Workstation (64 MB RAM, 40 MHz) dient zur Kommunikation des Benutzers mit der Motorensteuerkarte PMAC. Die Programme zur Initialisierung und Ansteuerung der Karte, sowie die Algorithmen, die das Verhalten des visuellen Systems beschreiben, laufen auf der Workstation ab. Die Workstation ist über ihren S-Bus und einen VME-Adapter am VME-Bus angeschlossen. Sie übernimmt auch

die Kommunikation mit dem Pipelinerechner MaxVideo 200. Mit dem Betriebssystem Solaris 2.x ist Echtzeitverhalten implementierbar.

Literatur

- [1] Y. Aloimonos (Editor). *Active Perception*. Lea Lawrence Erlbaum Associate, Willestdal, New Jersey, 1993.
- [2] D.H. Ballard, C. Brown. *Principles of Animate Vision Image Understanding*, Vol.56,1992, 3-21.
- [3] R. Bajcsy. *Active Perception* Proceedings of the IEEE, Vol.76, No.8,1988, 996-1005.
- [4] J. Bruske. *Dynamic Cell Structures*. Technischer Bericht 9403. Christian-Albrechts-Universität zu Kiel, 1993.
- [5] K. Pahlavan, J.A. Eklundh. *Heads, Eyes and Head-Eye Systems* International Journal of Pattern Recognition and Artificial Intelligence, Vol.7, No.1, 1993, 3-49.
- [6] J.-M. Jolion. *Computer Vision Methodologies* Image Understanding, Vol.59, No.1,1994, 53-71.
- [7] A. Koschan. *A Survey on Current Stereo Papers*. Technischer Bericht 92-22. Technische Universität Berlin, 1993.
- [8] E.P. Krotkov. *Active Computer Vision by Cooperative Focus and Stereo*. Springer Verlag, New York, 1989, 82ff.
- [9] W.M.Thiemer, H.A. Mallot. *Binocular Vergence Control and Depth Reconstruction Using a Phase Method* Artificial Neural Networks, 2, 1992, 517 -520.
- [10] D.A. Robinson. *Control of Eye Movements*. In Handbook of Physiology, The Nervous System II, American Physiological Society, 1981, 1275-1320.
- [11] M.J.Swain, M.A. Stricker *Promising Directions in Active Vision*. International Journal of Computer Vision, 11:2, 1993.
- [12] T. Yeo, S. Ong, Jayasooriah, R. Sinniah. *Autofocusing for tissue Microscopy* Image and Vision Computing, Vol.11, 1993.
- [13] A. L.Yarbus. *Eye Movements and Vision*. Verlag Plenum Press, New York, 1967.

(Zusätzlich wurden die Manuals und Anleitungen der Firmen DATACUBE, DELTA TAU, SUN und TRC verwendet.)