

# DAViS – Dortmunder Active Vision System

## Projektleitung

Prof. Dr.-Ing.  
Ferdinand Kallmeyer †

## Kompetenzplattform

Kommunikationstechnik  
und Angewandte  
Signalverarbeitung

## Forschungsschwerpunkt

Process Improvement  
& CAQ

## Zeitraum

2006 – 2008

## Förderung

Fachhochschule  
Dortmund  
Forschungsbudget

## Weiterführung der Forschungsarbeiten

Prof. Dr.-Ing.  
Thomas Feldehoff  
Fachbereich  
Informations- und  
Elektrotechnik  
Institut für  
Informationstechnik  
Fachhochschule  
Dortmund  
Sonnenstraße 96  
44139 Dortmund  
Tel.: (0231) 9112-386  
E-Mail: feldehoff  
@fh-dortmund.de

† Prof. Dr.-Ing.  
Ferdinand Kallmeyer  
verstarb am 28.02.2008  
bei einem tragischen  
Unfall.

## Abstract

Die fortschreitende Entwicklung der technischen Disziplinen, allen voran der Computertechnologie, ermöglicht zunehmend die Übernahme komplizierter Aufgaben durch Automaten. Insbesondere periodische Prozesse mit immer gleich bleibenden Abläufen können heute fast vollständig automatisch betrieben werden, was beispielsweise in der Serienproduktion der Automobilindustrie eindrucksvoll demonstriert wird. Schwierigkeiten bei der automatischen Verarbeitung ergeben sich in der Regel dann, wenn Prozesse nicht vorhersehbaren Situationen oder Änderungen unterliegen. Diese nicht vorhersehbaren Situationen müssen sensorisch erfasst werden, um geeignete Reaktionen zu initiieren. Insbesondere in der Robotik sind solche nicht geplanten oder vorhersehbaren Situationen oft räumlicher Natur. Sind zum Beispiel in einem Montageprozess Bauteile durch einen Roboter zu greifen, deren genaue räumliche Position nicht a priori bekannt ist, so muss eine sensorische Erfassung erfolgen. Entsprechende Aufgabenstellungen finden sich bei autarken Navigationsanwendungen für den Land- und Lufttransport oder zukünftigen Fahrerassistenzsystemen.

## I. Einleitung

Als Lösungsansätze der angeführten Aufgabenstellung werden in der Praxis oft aktive Verfahren genutzt. Bei aktiven Verfahren [1] werden zur räumlichen Vermessung spezielle Strahlungsquellen wie Ultraschall oder Laserlicht eingesetzt. Die Position der bestrahlten Objekte wird dann durch Reflexion der Strahlung an den Objekten bestimmt. Diese Verfahren bieten sich insbesondere für Aufgaben der Kollisionsvermeidung an. Sollen allerdings Umgebungsobjekte selektiv manipuliert werden, so müssen diese automatisch erkannt werden.

Neben aktiven Verfahren werden seit längerem passive Systeme intensiv erforscht. Passive Systeme [1] zeichnen sich durch den Verzicht auf eine aktive Strahlungsquelle aus, d.h. sie nutzen lediglich die diffuse Umgebungsbeleuchtung. Der Reiz der passiven Systeme liegt darin, dass die Natur im Zuge der Evolution hocheffiziente passive Strategien zur Identifikation und Lagebestimmung von Objekten hervorgebracht hat [2], die heutigen technischen Lösungen weit überlegen sind. Als ein Beispiel ist hier neben vielen anderen höheren Lebewesen der Mensch mit seinem hoch entwickelten Sehsystem zu nennen. Zur Objekterkennung werden bei passiven Systemen in der Regel monokulare Kamerasysteme

eingesetzt. Soll neben der reinen Objekterkennung zusätzlich die genaue Objektposition im Raum bestimmt werden, so wird der Einsatz von binokularen oder polyokularen Kamerasystemen notwendig, da bei der Bildaufnahme durch die Projektion des dreidimensionalen Raumes auf die zweidimensionale Bildebene die Tiefeninformation der Szene verloren geht. Zentrale Aufgabe bei diesen Ansätzen ist die Bestimmung der Objektposition bzw. der Tiefeninformation aus zwei oder mehreren Bildpaaren. Beim Stereosehen besteht aufgrund der aus unterschiedlichen Perspektiven aufgenommenen Szene die Möglichkeit, diese Aufgabe durch passive Triangulation zu lösen. Die Idee ist, aus den unterschiedlichen Positionen, auf die ein Raumpunkt projiziert wird, auf dessen Position im dreidimensionalen Raum zu schließen.

## II. Konzeption

DAViS dient als Basis für die Ausweitung des Bereichs Angewandte Signalverarbeitung in der Kompetenzplattform Kommunikationstechnik und Angewandte Signalverarbeitung (KOPF-CAS) auf Themenstellungen der digitalen Bildsignalverarbeitung und ist am Institut für Informationstechnik bearbeitet worden. Anwendungen im Bereich der industriellen Bildverarbeitung können synergetisch in den Forschungsschwerpunkt Process Improvement & CAQ integriert werden. Die Entwicklungsumgebung von DAViS besteht im Wesentlichen aus drei Komponenten: Einem aktiven stereoskopischen Kamerasystem, einer Softwareentwicklungsumgebung sowie einer Zielhardwareplattform. Die Komponenten erfüllen im Einzelnen diese Anforderungen:

- Das stereoskopische Kamerasystem kann entsprechend dem menschlichen Sehsystem mit hoher Genauigkeit aktiv auf einen beliebigen Raumpunkt ausgerichtet und fokussiert werden, um Stereobildpaarsequenzen für die Aufgaben der Objekterkennung und Objektlageschätzung zu generieren.
- Die Softwareentwicklungsumgebung ermöglicht den modellbasierten Entwurf von Anwendungen der digitalen Bildverarbeitung. Die Modelle sind als Simulation lauffähig und können vor der Portierung auf die Zielhardware zur Evaluierung der verwendeten Algorithmen genutzt werden. Die Evaluierung ist aufgrund des gewählten modellbasierten Ansatzes unabhängig von der Zielhardware. Die Portierung der Modelle auf die Zielhardware erfolgt weitgehend automatisch durch target-spezifische Codegenerierung.

- Entsprechend der vielfältigen Anwendungsfelder stereoskopischer Sehsysteme und den daraus resultierenden spezifischen Anforderungen ist die Zielhardwareplattform flexibel im Hinblick auf die jeweilige Aufgabenstellung gewählt. In einem ersten Schritt sind als Zielplattformen ein PC sowie das Digitale Signalprozessor Entwicklungsboard DM642 von Texas Instruments verfügbar. Bei der gewählten Zielplattform wurde auf die Fähigkeit zur Verarbeitung von Stereo-Videodaten geachtet.

### III. Objekterkennung und Objektlageschätzung

Bei der Umsetzung der Konzeption wird dem modellbasierten Entwurf Rechnung getragen. Hierdurch ist es möglich, die vielfältigen und in Teilen extrem komplexen Aufgaben zur Objekterkennung und Objektlageschätzung in einzelne Teilbereiche zu zerlegen und diese allgemeingültig und wieder verwertbar für die übergeordnete Aufgabe zur Verfügung zu stellen. Nur die konsequente Anwendung des modellbasierten Entwurfs ermöglicht eine kontinuierliche wissenschaftliche Arbeit auf diesem Themengebiet trotz ständig wechselnder Studierender [3-7,9,11].

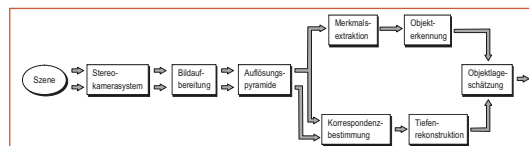


Abb. 1: Arbeitsaspekte zur Objekterkennung und Objektlageschätzung

In Abbildung 1 sind einzelne Aspekte zur Objekterkennung und Objektlageschätzung veranschaulicht. Ausgehend von einer dynamisch veränderlichen Szene erstellt das Stereokamerasystem zwei Videosequenzen. Aus den Videosequenzen wird eine Sequenz von Stereobilddaten generiert, die für die weiteren Schritte aufbereitet werden. Zur Verkürzung der Reaktionszeit können anschließend mit Hilfe einer Auflösungs-pyramide Stereobildpaare mit unterschiedlicher Bildgröße und unterschiedlichem Detaillierungsgrad erzeugt werden. Dieser Verarbeitungsschritt dient dazu, Bildbereich im optischen Zentrum hochauflösend zu analysieren, während periphere Bildbereiche nur mit niedrigerer Ortsauflösung berücksichtigt werden. Die resultierenden Bilder bilden die Basis für die zwei parallel ausführbaren Schritte zur Objekterkennung und Tiefenrekonstruktion. Da die Objekterkennung und Tiefenrekonstruktion auf Basis der gleichen vorverarbeiteten Bilder erfolgt,

kann bei der abschließenden Objekt-lageschätzung der erkannten Objekte auf eine aufwendige Sensorfusion verzichtet werden.

### A. Stereokamerasystem

Die Architektur des Stereokamerasystems ist in Abb. 2 dargestellt. Der Stereokamerakopf besitzt insgesamt zehn aktive Freiheitsgrade, die über RS-232 Schnittstellen ansprechbar sind. Er ist modular aus zwei Vergenzmodulen [3], einem Neigemodul und einem Schwenkmodul [4] aufgebaut. Jedes Vergenzmodul beinhaltet eine Kamera, einen Schrittmotor und einen Encoder zur Bestimmung der Vergenzposition. Als Kameras werden preisgünstige gehäuselose Serien-Blockkameras der Firma Sony eingesetzt. Jede Kamera verfügt über drei optische Freiheitsgrade (Zoom, Blende, Fokus). Da zur Vereinfachung der Rekonstruktion der 3D-Weltkoordinaten aus dem Korrespondenz- bzw. Tiefenbild das Bildzentrum der Kamera exakt im Schnittpunkt von Neige- und Vergenzachse liegen soll, sind bei der Konstruktion der Kamera-aufhängung in jedem Vergenzmodul sechs zusätzliche passive Freiheitsgrade konstruktiv berücksichtigt, die eine exakte Justage [6] der Kamera im Vergenzmodul ermöglichen.

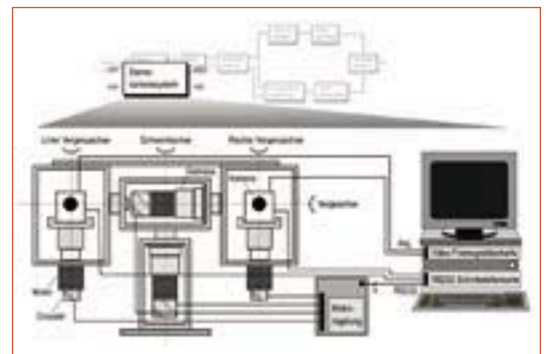


Abb. 2: Systemarchitektur des Stereokamerasystems

Um die hohen Anforderungen im Hinblick auf Positionieraufgaben auch ohne Direktantrieb erfüllen zu können, ist als Getriebtyp ein hochgenaues Harmonic Drive Getriebe gewählt [4]. Hier wird ein Wavegenerator direkt auf die Motorwelle montiert. Der gewählte Motor wird über einen Adapter des Getriebeherstellers an das Getriebe montiert. Der gewählte Schrittmotor kann im Mikroschrittbetrieb betrieben werden. Um die Antriebstechnik einheitlich zu gestalten, wird der gleiche Motortyp für alle vier Antriebseinheiten eingesetzt. Auf diese Weise kann eine einheitliche Steuerungselektronik für alle Antriebe verwendet werden [7].



$$\dot{x}_k = \lambda_k \cdot x_k - (b + c) \sum_{v \neq k} x_v^2 \cdot x_k - c \cdot x_k^3$$

In der Hakenschen Mustererkennungsgleichung sind  $b$  und  $c$  positive Konstanten,  $x_k$  ist ein Ordnungsparameter und  $\lambda_k$  ein individueller Gewichtsparameter des exponentiellen Wachstums für den  $k$ -ten Ordnungsparameter. Für  $\lambda_k > 0$  findet hierbei ein Wettbewerb zwischen den einzelnen Ordnungsparametern  $x_k$  statt. Die Gleichung wurde von Trapp, [8], um die implizite Erkennung von Okklusionen erweitert. Bei der Umsetzung im Hinblick auf praktische Aufgabenstellungen wird besonders auf eine optimale Verarbeitungsgeschwindigkeit bei hoher Robustheit geachtet [11].

Zur Initialisierung des Selbstorganisationsprozesses nach obiger Gleichung wurde vom Projektleiter unter MATLAB ein komplexwertiges Ähnlichkeitsmaß für Stereobildpaare implementiert. Ein hellerer Bildpunkt kennzeichnet einen geringeren Objektstand zur Bildebene als ein dunklerer Bildpunkt. Der Realteil des Ähnlichkeitsmaßes dient zur Initialisierung des Selbstorganisationsprozesses, dessen Ergebnis als Disparitätskarte, Abb. 6, dargestellt ist. Hierbei kennzeichnen schwarze Bildpunkte Okklusionen, d.h. Bildbereiche der aufgenommenen Szene, die nur von einer Kameraperspektive aus zu sehen sind. Der Imaginärteil des komplexwertigen Ähnlichkeitsmaßes kann nach Erstellung der Disparitätskarte genutzt werden. Die Disparitätskarte kann aufgrund der Pixelzuordnung von rechtem und linkem Stereobild nur pixelgenau sein und führt durch das diskrete Abtastraster zu Quantisierungsfehlern.

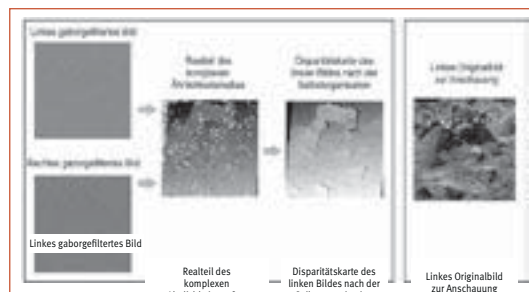


Abb. 6: Disparitätsbestimmung

Rechtes gaborfiltertes Bild

#### IV. Zusammenfassung und Ausblick

Ein vollständig steuer- und beobachtbares stereoskopisches Sehsystem steht am Institut für Informationstechnik zur Verfügung. Aufgaben zur Bildverarbeitung sind modellbasiert in MATLAB/Simulink umgesetzt. Hierzu gehört auch ein

entsprechendes Kameramodell, so dass neben der Auswertung realer Messdaten Vergleiche mit simulierten Daten möglich macht. Eine Evaluierung des Kameramodells bzw. der verwendeten Projektionsvorschriften dient der ständigen Verbesserung der Modelle und des Verständnisses für auftretende Artefakte. Auf diese Weise optimierte Modelle führen im Zusammenspiel sowohl zu einer Objekterkennung als auch zu einer Objektlagenschätzung.

Aufgrund der optischen Freiheitsgrade (Zoom, Fokus, Blende) sind vielschichtige Untersuchungen zur Optimierung des Kameramodells erforderlich. Somit ist gerade unter Beachtung dieser Freiheitsgrade bei der Kalibrierung des stereoskopischen Sehsystems noch Verbesserungspotential vorhanden. In weiterführenden Arbeiten wird auch die Abhängigkeit der Positionsgenauigkeit von Objekten im Raum von der Kalibrierung und damit von der Komplexität und Genauigkeit des Kameramodells untersucht.

Die Erkennung von Objekten und auch die Verfolgung eines Objektes im Raum sind mit der verwendeten Systemarchitektur, Abbildung 2, grundsätzlich demonstrierbar. Der Schwerpunkt liegt auf der algorithmisch korrekten Umsetzung. Da der PC und insbesondere die Entwicklungsumgebung MATLAB/Simulink nur bedingt echtzeitfähige Anforderungen erfüllen, ist in einem ersten Schritt auf dem DM642 Entwicklungsboard von Texas Instruments die Auswertung von Stereobilddaten implementiert [9]. In weiteren Arbeiten wird ein Embedded System entstehen, das die Anforderungen an eine kontinuierliche Objektverfolgung gewährleistet. Hierdurch wird die eng mit dem modellbasierten Entwurf verknüpfte Idee des Rapid Prototyping demonstrierbar.

#### V. Danksagung

DAViS wurde durch die hochschulinterne Forschungsförderung der Fachhochschule Dortmund finanziert. Durch die Beschaffung der einzelnen Komponenten war ein zeitnahe Aufbau des stereoskopischen Sehsystems gegeben. Dieses Sehsystem gestattet vielfältige Untersuchungen, die im Rahmen von Projekt- und Diplomarbeiten von Studierenden erfolgreich durchgeführt wurden. DAViS stellt einen Kernkompetenzbereich der digitalen Bildverarbeitung am Institut für Informationstechnik und auch in der KOPF-CAS dar.

Bei einem tragischen Unfall verstarb im Februar 2008 der liebeswürdige und hoch geschätzte Kol-

lege Prof. Dr.-Ing. Ferdinand Kallmeyer. Wie DAVIS verdeutlicht gelang es Prof. Dr. Kallmeyer immer, seine überwältigende Fachkompetenz mit einem engen Praxisbezug der ihn interessierenden Fragestellungen zu verbinden. Seine Zufriedenheit spiegelte sich in Prof. Dr. Kallmeyers menschlich sehr angenehmer und herzlicher Art wider; die Zusammenarbeit mit ihm war immer eine Bereicherung. Die von Prof. Dr. Kallmeyer begonnenen Themengebiete werden im Institut für Informationstechnik beheimatet bleiben.

## VI. Literatur

- [1] Gonzalez, R.C., Woods, R.E. und Eddins, S.L., Digital Image Processing Using Matlab, Prentice Hall, 2004
- [2] Haken, H., Erfolgsgeheimnisse der Natur, Deutsche Verlagsanstalt, 1981
- [3] Jotzo, D. und Wischeloh, M., Aufbau und Inbetriebnahme einer Entwicklungsumgebung zur kameragestützten Objekterkennung, Ingenieurmäßige Arbeit, FH Dortmund, 2006
- [4] Kiwitt, D., Entwurf eines Schwenkmoduls für ein stereoskopisches Sehsystem, Ingenieurmäßige Arbeit, FH Dortmund, 2006
- [5] Kräher, B., Verfahren zur Parametrisierung von Verhaltensmodellen elektrischer Verstellantriebe in Kraftfahrzeugen, Diplomarbeit, FH Dortmund, 2004
- [6] Macecek, M., Konzeption eines Vergenzmoduls für einen stereoskopischen Kamerakopf, Ingenieurmäßige Arbeit, FH Dortmund, 2006
- [7] Macecek, M., Entwurf und Inbetriebnahme eines stereoskopischen Sehsystems, Diplomarbeit, FH Dortmund, 2006
- [8] Trapp, R., Stereoskopische Korrespondenzbestimmung mit impliziter Detektion von Okklusionen, Dissertation, Universität Paderborn, 1998
- [9] Triesch, S., Automatische Codegenerierung von Videoquellen auf DSP's auf Basis von MATLAB/Simulink, Diplomarbeit, FH Dortmund, 2007
- [10] Tsai, R.Y., A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, IEEE Journal of Robotics and Automation, 1987
- [11] Vrieler, S., Kalibrierung eines stereoskopischen Kamerasystems unter MATLAB, Projektarbeit, FH Dortmund, 2008