



# Visual Servoing

Jürgen Beyerer, Michael Heizmann, Helge-Björn Kuntze, Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, Karlsruhe

Mit bildgebenden Sensoren lassen sich weitaus mehr und umfassendere Informationen für die Überwachung und Regelung von technischen Prozessen gewinnen, als mit manch anderen Sensoren möglich ist. Dies eröffnet völlig neuartige Möglichkeiten der Realisierung von Prozess-Monitoring und Regelungssystemen für ein weites Spektrum von Anwendungsfeldern. Die notwendigen Anforderungen an die Echtzeitfähigkeit und Praxisreife beim Einsatz bildgebender Sensoren setzen jedoch eine hohe Verarbeitungsgeschwindigkeit und Robustheit der Hardware- und Softwarekomponenten voraus, so dass erst vor ca. 20 Jahren Anwendungen zur Regelung von Robotern realisierbar wurden.

Die 1996 am Stanford Research Institute (SRI) geprägte Definition von „Visual Servoing“<sup>1</sup> beschränkte sich zunächst auf die Positions- und Bewegungsregelung von Robotern und anderen mobilen Mechatroniksystemen (z. B. Fahrzeuge, Krane etc.) mit Hilfe von bildgebenden Sensoren. Mit der stark wachsenden Leistungsfähigkeit und Funktionalität der Hardware- und Softwarekomponenten und der zunehmenden Durchdringung anderer Anwendungsbereiche mit bildgebenden Sensoren ging auch ein Wandel der Begrifflichkeit einher. Immer häufiger wird unter „Visual Servoing“ die allgemeingültigere Definition der „Prozessregelung mit bildgebenden Sensoren“ verstanden.

Auch den Herausgebern dieses Themenheftes scheint diese weniger eingeschränkte Sichtweise sinnvoller zu sein. Sehr anschaulich wird dies am Beispiel der Regelung von Schweißprozessen mit bildgebenden Sensoren sichtbar. So werden gegenwärtig am Markt ausgereifte Systemlösungen und -komponenten mit bildgebenden Sensoren angeboten (z. B. beim Laser- oder WIG-Schutzgasschweißen), die nicht nur die Messung und Regelung der 3D-Position des Brenners entlang der Schweißnaht ermöglichen, sondern auch die thermi-

sche Regelung des Schmelzbades und mithin die Güte der erzeugten Schweißraupe. Durch den generischen Charakter einer Vielzahl von extrahierbaren Bildinformationen sind Regelungskonzepte mit bildgebenden Sensoren offenbar nicht an spezielle technologische oder physikalische Prozesse gebunden, sondern lassen sich im Sinne der Systemtheorie universell einsetzen.

Die spezifischen Herausforderungen bildgestützter Regelungen liegen vor allem in den folgenden Bereichen:

Die Verarbeitung des erfassten Bildes, in dem die Information über die Regelgröße enthalten ist, ist meist komplexer als bei sonst üblichen Messgliedern und beinhaltet eine Vielzahl von Verarbeitungsschritten (z. B. Vorverarbeitung, Segmentierung, Merkmalsextraktion, ggf. Klassifikation), bis das gewünschte Nutzsignal zur Bestimmung der Regelabweichung zur Verfügung steht. Diese Verarbeitungskette muss konzipiert und parametrisiert werden, was beim Entwurf der bildgestützten Regelung besondere Sorgfalt erfordert.

In diesem Zusammenhang stellt sich die Frage, welche Größe als Führungsgröße modelliert wird: Wird dazu eine im Bild direkt erfassbare Größe herangezogen (z. B. eine Sollposition für ein Zielobjekt im Bild), so wird die Regelung als „bildbezogenes Visual Servoing“ (Image Based Visual Servoing) bezeichnet. Findet eine Transformation der im Bild enthaltenen Information statt, so dass z. B. die Position eines Objekts im Raum bestimmt und auf diese geregelt wird, bezeichnet man diesen Ansatz üblicherweise als „modell- oder positionsbezogenes Visual Servoing“ (Position Based Visual Servoing). In beiden Fällen erfordert der Einsatz der Bildsensorik als Messglied eine problemangepasste Herangehensweise zur Konzeption des Regelkreises.

Die Nutzinformation über die Regelgröße in den erfassten Bildern ist durch Störungen überlagert, die spezifisch für Systeme der Bildverarbeitung sind. So treten z. B. Abschattungen, Empfindlichkeiten gegen die Beleuchtungssituation, Fremdlicht, Variationen in der optischen Erscheinung der zu regelnden Szene etc. auf,

<sup>1</sup> Hutchinson, S.; Hager, G. D.; Corke, P. I.: A Tutorial on Visual Servo Control. In: IEEE Transactions on Robotics and Automation 12 (1996), No. 5, pp. 651–670



die in Regelkreisen mit anderen Messgliedern nicht zu befürchten sind.

Wie jedes System der Bildverarbeitung sind bildgestützte Regelungen den Gesetzen der Optik unterworfen. Dies bedeutet z. B., dass sich beim dimensionellen Messen mit der üblicherweise verwendeten entozentrischen Optik der Abbildungsmaßstab in Abhängigkeit vom Abstand ändert, was einer zusätzlichen Störgröße entspricht. Eine weitere Auswirkung ist, dass die Orts- und Wertequantisierung der bildgestützt erfassten Größe oft stärker ins Gewicht fällt als bei sonst üblichen Messgliedern.

Schließlich benötigen Systeme der Bildverarbeitung für die Abarbeitung der Kette der Verarbeitungsschritte eine gewisse Verarbeitungszeit, die als Totzeit in die Regelung eingeht und für die Dynamik und Stabilität der Regelung unerwünscht ist. Häufig wird für die Bildverarbeitung selbst und das darunterliegende Betriebssystem Software verwendet, die keine garantierten Antwortzeiten aufweist, so dass die unvermeidliche Totzeit auch noch stochastisch variiert.

Andererseits bieten bildgestützte Regelungen auch Chancen, die sich auf anderem Weg kaum realisieren lassen:

Mittels bildgestützter Messglieder können Messgrößen erfasst werden, die anders kaum zugänglich wären. Dies sind z. B. mehrdimensionale positionsbezogene Informationen im Fall des positionsbezogenen Visual Servoing. Ein anderes Beispiel sind objektbezogene Größen (etwa in der Service-Robotik), für deren Bestimmung zunächst das richtige Objekt klassifiziert werden muss, bevor die eigentliche Messgröße bestimmt werden kann.

Für die vorausschauende optimale kollisionsfreie Bahnplanung, Navigation und Bewegungsregelung von mobilen Robotern mit modellprädiktiven Regelungskonzepten bieten bildgebende Sensoren entscheidende Vorteile, die andere vergleichbare Sensoren nicht aufweisen können. Sie erfassen in jedem Bild sowohl den Nahbereich (Vordergrund) als auch den Weitbereich (Hintergrund) des Fahrzeuges und ermöglichen dadurch – zusammen mit darauf aufbauenden heuristischen oder modellprädiktiven Algorithmen – die Imitation des charakteristischen taktischen und strategischen Navigationsverhaltens des Menschen.

Die Erfassung der Messgröße erfolgt stets berührungslos. Auf diese Weise sind prinzipbedingt keine unerwünschten Rückwirkungen auf die Regelgröße zu befürchten. Sensoren müssen auch nicht invasiv in das zu regelnde System eingebracht werden, was den Aufwand bei der Umsetzung einer Regelung reduziert.

Schließlich werden Kameras als „Messaufnehmer“ des bildgestützten Messglieds immer besser und gleichzeitig günstiger. Dies macht sie für viele Anwendungen zunehmend attraktiv und öffnet neue Anwendungsgebiete. Gleichzeitig profitieren Anwendungen in der professionellen Automatisierungstechnik von Fortschritten, die in der Informationstechnik – v. a. auch getrieben von der Unterhaltungselektronik – sehr schnell erfolgen. Ein pro-

minentes Beispiel aus der jüngsten Vergangenheit ist die 3D-Kamera *Microsoft Kinect*, die in kürzester Zeit den Sprung aus der Spieleindustrie in die Robotik geschafft hat.

Dieses Themenheft greift einige der angesprochenen Chancen und Herausforderungen auf und bietet einen Überblick über aktuelle Themen und Anwendungen bildgestützter Regelung.

Der Beitrag von Malzahn et al. verwendet Visual Servoing, um das Führungs- und das Störungsverhalten eines gliedelastischen Roboters zu verbessern. Dabei kommt als Sensorik eine RGB-D-Kamera auf dem Endeffektor des Roboters zum Einsatz, die zusätzlich zu Helligkeitswerten für jeden Bildpunkt einen Entfernungswert liefert. Aus den erfassten Bild- und Entfernungsdaten wird die Eigenbewegung des Roboters rekonstruiert und zur Bewegungskompensation verwendet.

Damit ein Roboter z. B. als Mitglied eines Teams von Menschen und Robotern auch komplexere mehrphasige Handhabungsaufgaben arbeitsteilig erfolgreich bewältigen kann, ist es nicht ausreichend, dass er unter Verwendung von bildgebenden Sensoren nur visuell mit seiner Umwelt kommuniziert. Er muss zusätzlich über weitere Sinne verfügen, damit er z. B. auch akustisch oder bei physikalischem Kontakt mit umgebenden Gegenständen oder Menschen auch haptisch mit Kraft-Momenten-Sensoren seine Umwelt wahrnehmen kann. Zur multisensoriellen Regelung solcher mehrphasigen Aufgabenstellung in einem komplexen mehrfach beschränkten Arbeitsumfeld wird von De Laet et al. erstmals eine generische umfassende Entwurfsmethodologie vorgestellt.

Ein wichtiges Anwendungsgebiet bildgestützter Regelungen sind autonome Roboter. Grundlage für die Handlungsfähigkeit solcher Roboter sind zuverlässige und genaue Karten der Umgebung, die während der Roboterbewegung erstellt werden (Simultaneous Localization and Mapping, SLAM). Der Beitrag von Endres et al. zeigt, welches Potenzial RGB-D-Sensoren für die autonome Robotik haben.

Im Beitrag von Milighetti und De Luca geht es um eine bildgestützte Blickregelung für einen humanoiden Roboterkopf. Die Bilddaten eines Stereokamerasystems werden verwendet, um die Ausrichtung eines mit kinematischer Redundanz ausgestatteten Kopfes zu regeln. Damit wird das Ziel verfolgt, die dynamischen Zielobjekte immer im Blickfeld der Stereokamera zu halten, wobei der Kopf des humanoiden Roboters ein möglichst menschenähnliches Bewegungsmuster aufweisen soll.

Im Beitrag von Böge et al. wird ein Visual-Servoing-Konzept für humanoide Roboter vorgestellt, das sowohl einarmiges als auch zweiarmiges visuell geregeltes Greifen und Manipulieren von Objekten ermöglicht. Hierzu werden sowohl die Objekte wie auch die Roboterhände während des Greifvorgangs mithilfe von Blickrichtungswechseln visuell verfolgt. Durch Fusion der Kamera- und Gelenkwinkeldaten sowie Kraftsensordaten werden

Bewegungsvorgaben für die Arme und den Roboterkopf generiert. Die Validierung der Funktionalität und Leistungsfähigkeit wurden an dem humanoiden Roboter ARMAR-III anhand verschiedener Greif- und Manipulationsaufgaben durchgeführt.

Der Beitrag von Gräser et al. berichtet von dem mobilen manipulationsfähigen Assistenzroboter FRIEND, der körperbehinderten, z. B. querschnittsgelähmten Menschen sowohl Mobilität im Weitbereich als auch Manipulationsfähigkeit bei täglichen Handhabungsaufgaben im Nahbereich vermittelt. Mit Hilfe von bildgebenden Sensoren und adaptiven Visual-Servoing-Konzepten kann der Nutzer über unterschiedliche Modalitäten (z. B. Augen, Sprache) anspruchsvolle Handhabungs- und Greifaufgaben aktivieren, die von einem mehrachsigen visuell geregelten Roboterarm auch in einem komplexen häuslichen Umfeld (z. B. Kühlschrank) autonom umgesetzt werden.

Im Beitrag von Strom et al. wird schließlich mit Zellulären Neuronalen Netzwerken eine Technologie vorgestellt, mit der die Totzeiten bildgestützter Messglieder deutlich reduziert werden können. Durch eine in Hardware realisierte Bildauswertung bereits auf dem Kamerachip werden sehr hohe Verarbeitungsgeschwindigkeiten erzielt, so dass auch hochdynamische Prozesse wie z. B. Schweißprozesse geregelt werden können.



**Prof. Dr.-Ing. Jürgen Beyerer** ist Leiter des Fraunhofer-Instituts für Optronik, Systemtechnik und Bildauswertung IOSB in Karlsruhe und Inhaber des Lehrstuhls für Interaktive Echtzeitsysteme am Institut für Anthropomatik am Karlsruher Institut für Technologie KIT.

Adresse: Fraunhofer IOSB, Fraunhoferstraße 1, 76131 Karlsruhe,  
E-Mail: juergen.beyerer@iosb.fraunhofer.de



**Dr.-Ing. Michael Heizmann** ist Leiter der Abteilung Mess-, Regelungs- und Diagnosesysteme des Fraunhofer-Instituts für Optronik, Systemtechnik und Bildauswertung IOSB in Karlsruhe.

Adresse: Fraunhofer IOSB, Fraunhoferstraße 1, 76131 Karlsruhe,  
E-Mail: michael.heizmann@iosb.fraunhofer.de



**Dr.-Ing. Helge-Björn Kuntze** ist wissenschaftlicher Berater des Fraunhofer-Instituts für Optronik, Systemtechnik und Bildauswertung IOSB in Karlsruhe.

Adresse: Fraunhofer IOSB, Fraunhoferstraße 1, 76131 Karlsruhe,  
E-Mail: helge-bjoern.kuntze@iosb.fraunhofer.de