

Visual Servoing für ein- und zweiarmige Manipulationsaufgaben bei humanoiden Robotern

Visual Servoing for Single and Dual Arm Manipulation Tasks in Humanoid Robots

Christian Böge, Nikolaus Vahrenkamp, Tamim Asfour, Rüdiger Dillmann

Unter dem Begriff Visual Servoing werden Ansätze zur visuell geregelten Ausführung z.B. von Roboterbewegungen zusammengefasst. Hierbei werden durch kontinuierliche Analyse von Kamerabildern Stellgrößen erzeugt, um beispielsweise einen Roboterarm anzusteuern. In diesem Artikel stellen wir das Visual-Servoing-System des humanoiden Roboters ARMAR-III vor, welches eine zuverlässige Ausführung von ein- und zweiarmigen Greif- und Manipulationsaufgaben ermöglicht. Durch den Closed-loop-Ansatz können Ungenauigkeiten in Aktorik und Sensorik ausgeglichen werden.

The term Visual Servoing subsumes all control approaches, that use feedback information extracted from vision sensors. Therefore corrective actions are being generated on the continuous analysis of camera images, e.g. for controlling a robot arm. In this article we present the Visual-Servoing system for the humanoid robot ARMAR-III which enables safe execution of single and dual-handed grasping and manipulation actions. Inaccuracy of actuators and sensor systems can be compensated by the closed-loop approach.

Schlagwörter: Visual Servoing, Greifen, Manipulation, Bildverarbeitung, Regelung, Kraftsensorik, Hand-Auge Kalibrierung, Inverse Kinematik.

Keywords: Visual Servoing, Grasping, Manipulation, Vision, Control, Force sensors, Hand-Eye calibration, Inverse Kinematics.

1 Einleitung

Humanoide Roboter werden unter anderem entwickelt, um Menschen bei alltäglichen Aufgaben zu unterstützen. Die rasante Entwicklung der letzten Jahre zeigt, dass mobile Manipulatoren, Serviceroboter und humanoide Robotersysteme unter realen Bedingungen in einem menschenzentrierten Umfeld eingesetzt werden können. Als langfristiges Ziel werden Roboter angestrebt, die autonom handeln und das hierzu benötigte Wissen selbstständig akquirieren und verwalten. Weiterhin müssen solche Roboter über eine Vielzahl von Fähigkeiten zur Sensordatenverarbeitung, Planung und zur Aktionsausführung verfügen. Eine grundlegende Fähigkeit ist dabei das Greifen und Manipulieren von Gegenständen, da hierdurch eine physische Interaktion mit der Umwelt ermöglicht wird.

In diesem Beitrag wird ein Rahmenwerk vorgestellt, das einarmiges sowie zweiarmiges Greifen und Manipulieren von Objekten basierend auf der Rückkopplung visueller Wahrnehmung ermöglicht. Hierzu werden sowohl die Objekte wie auch die Roboterhände während des Greifvorgangs mithilfe von Blickrichtungswechseln visuell verfolgt. Aus den Daten der Kameras, Gelenkwinkelencoder und Kraftsensoren werden Bewegungsvorgaben für die Arme und den Roboterkopf erzeugt. Die Evaluation wurde an dem humanoiden Roboter ARMAR-III anhand verschiedener Greif- und Manipulationsaufgaben durchgeführt.

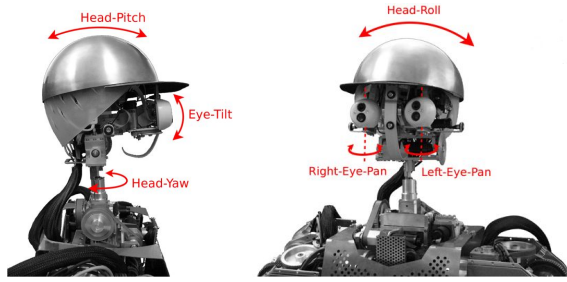


Bild 1: Bewegungsfreiheitsgrade des Kopfes von ARMAR-III.

2 Der humanoide Roboter ARMAR-III

Der in unseren Experimenten eingesetzte humanoide Roboter ARMAR-III [1] besitzt insgesamt 43 Bewegungsfreiheitsgrade. Mit Ausnahme der Hände sind alle Bewegungsfreiheitsgrade mit Gleichstrommotor-Getriebe-Einheiten aktuiert. Die beiden Arme des Roboters verfügen über jeweils 7 Bewegungsfreiheitsgrade. Die Motor-Getriebe-Einheiten des Ellenbogen und der Unterarmdrehung sind aus Gewichtsgründen im Torso des Roboters platziert. Die Kraftübertragung zu den Gelenken der Arme findet über Drahtseilzüge statt. An den Armen befinden sich zwei anthropomorphe Hände mit jeweils 8 pneumatisch aktuierten Bewegungsfreiheitsgraden [2]. Die Druckluft wird zentral von einem Kompressor zur Verfügung gestellt, der sich in der Plattform des Roboters befindet. Die 6D-Kraft-/Momentenmessdosen in den Handgelenken ermöglichen die Messung von einwirkenden Kräften und Momenten am Endeffektor. Die Hüfte des Roboters sowie dessen omnidirektional verfahrbare Plattform besitzen jeweils 3 Bewegungsfreiheitsgrade. Der Kopf des Roboters ist mit 4 Kameras ausgestattet. Pro *Auge* ist jeweils eine Kamera mit kurzer Brennweite und eine Kamera mit längerer Brennweite verbaut. Dieser Aufbau dient dazu, die Charakteristik der menschlichen Netzhaut nachzubilden, bei der die Auflösung im Sehzentrum (fovealer Bereich) größer ist als am Rande des Bildes (peripherer Bereich). Der Hals besitzt 4 Bewegungsfreiheitsgrade und weitere 3 Bewegungsfreiheitsgrade dienen zur Bewegung der Kameras. Diese lassen sich gemeinsam um eine Neige-Achse (engl. tilt) und unabhängig voneinander um jeweils eine Schwenk-Achse (engl. pan) bewegen (siehe Bild 1). Durch die Beweglichkeit des Kopfes und der Kameras lässt sich ein großer Arbeitsbereich visuell überwachen. Dabei lassen sich schnelle Blickrichtungswechsel mithilfe von Augenbewegungen realisieren, da die Augen eine geringe Trägheit besitzen. Alle Bewegungsfreiheitsgrade des Roboters sind mit Positionscodern ausgestattet, über die sich die aktuellen Gelenkwinkel bestimmen lassen.

3 Visual Servoing

Um einen Roboter in die Lage zu versetzen, auf neue Umweltsituationen, fehlerhafte Positionierungen oder Kollisionen reagieren zu können, muss der Roboter die Möglichkeit haben, Aktionsausführungen zu überwachen. Dem Menschen stehen zur Überwachung eigener Bewegungen Augen, Haut und Muskeln zur Verfügung. Bei einem Roboter können Kameras, die Sensorik der Gelenke und des Antriebs sowie Kraft-/Momenteninformationen zur Regelung eingesetzt werden.

Verfahren zur visuell geregelten Ausführung von Roboterbewegungen werden unter dem Begriff *Visual Servoing* zusammengefasst. In der Literatur wird meist zwischen drei Standardverfahren unterschieden: Bildbasiert (*Image-Based Visual Servoing*, *IBVS*), Positions basiert (*Position-Based Visual Servoing*, *PBVS*) und Hybrid (auch *2,5D Visual Servoing*). Die Stellgröße wird bei bildbasierten Verfahren aus dem Fehler zwischen aktueller und gewünschter Merkmalposition in der Bildebene gebildet, ohne dass dabei eine Lageschätzung des Objektes durchgeführt werden muss (siehe z.B. [3]). Bei PBVS-Ansätzen wird die Stellgröße aus der Differenz einer gewünschten Lage im kartesischen Raum und der aktuell visuell berechneten Lage bestimmt [4]. Hybride Ansätze verwenden eine Kombination beider Verfahren [5].

3.1 Visual Servoing für ARMAR-III

Das hier vorgestellte Verfahren basiert auf PBVS-Ansätzen, welche in [6] für einarmige Greif- und Manipulationsaufgaben vorgestellt wurden. In [7] wurde das System auf zweiarmige Problemstellungen erweitert. Eine umfassende Beschreibung des Rahmenwerks findet sich in [8]. In Bild 2 ist der Regelkreis für den bildbasierten Ansatz auf ARMAR-III dargestellt. Über die *Visual-Servoing-Schnittstelle* können Zielobjekte sowie Griffe für die linke und/oder rechte Hand vorgegeben werden. Das Visual-Servoing-Rahmenwerk nutzt die Roboter-Schnittstelle, um die aktuellen Gelenkwinkel c , einwirkende Kräfte und Momente f sowie Bilddaten zu ermitteln. Hieraus werden in jedem Regelzyklus Vorgaben für die Gelenkwinkelstellungen des Halses und der Augen c_{Kopf} sowie Bewegungsvorgaben für die Gelenke der Arme Δc_{links} und Δc_{rechts} bestimmt und dem Roboter vorgegeben.

Für jeden Arm steht eine Steuerungs- und Regelungskomponente zur Verfügung, welche aus einer visuell erfassten Objektpose $x_{Objekt} \in SE(3)$ und einem anzuwendenden Griff (gegeben durch eine Objekt-zu-Hand Transformation $x_{Griff} \in SE(3)$) eine kartesische Ziellage $x_{Ziel} \in SE(3)$ bestimmt (Gleichung 1).

$$x_{Ziel} = x_{Objekt} \cdot x_{Griff} \quad (1)$$

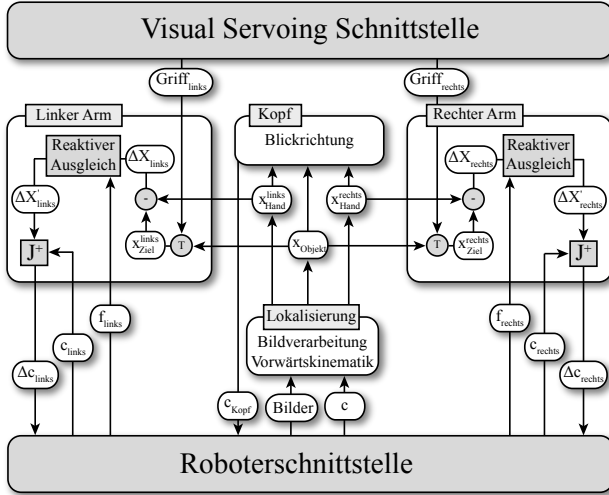


Bild 2: Visual Servoing für ein- und zweiarmige Bewegungen.

Aus x_{Ziel} und der visuell ermittelten Lage der Hand x_{Hand} wird die kartesische Regelungsdifferenz Δx bestimmt. Durch die reaktive Behandlung auftretender Kräfte und Momente wird $\Delta x'$ ermittelt und durch Methoden der differentiellen Kinematik in eine Geschwindigkeitsvorgabe Δc überführt.

Die Komponente zur Bestimmung der Blickrichtung passt die Ausrichtung des Kopfes und somit der Kameras an die aktuelle Situation an. Hier sind verschiedene Strategien zur Lenkung der Blickrichtung implementiert, um einhändige sowie eng-gekoppelte und lose gekoppelte zweiarmige Bewegungen zu realisieren. In Bild 3 sind die relevanten Koordinatensysteme für die Ausführung einer Greifbewegung mit einem Arm eingezeichnet. Aus der Objektlage x_{Objekt} resultiert die Greifpose x_{Ziel} , welche die Ziellage der visuell geregelten Bewegung der Hand beschreibt. Wichtig hierbei ist, dass sowohl die Hand des Roboters wie auch das Zielobjekt in einem Bildpaar sichtbar sind, damit die räumliche Relation zwischen Hand und Objekt genau bestimmt werden kann.

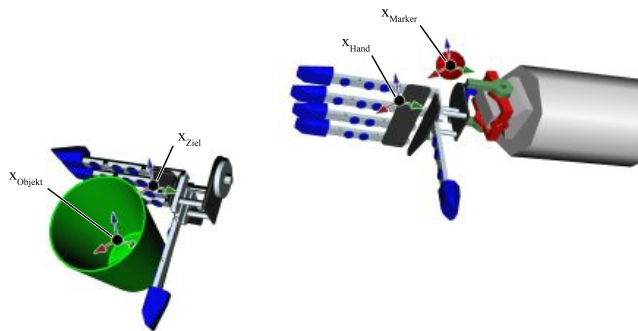


Bild 3: Die Koordinatensysteme für den Visual Servoing Controller.

3.2 Bestimmung der Regelungsvorgaben

Die Regeldifferenz Δx zwischen den visuell bestimmten Soll- und Ist-Posen x_{Ziel} und x_{Hand} ist im kartesischen Raum definiert. Um eine Beschädigung des Roboters oder der Umwelt zu verhindern, wird aus Δx eine Regeldifferenz $\Delta x'$ berechnet, die auftretende Kräfte und Momente berücksichtigt (siehe Bild 2). Hierzu wird aus den einwirkenden Kräften und Momenten f , welche mittels der 6D-Kraft-/Momentenmessdose im Handgelenk bestimmt werden, eine Korrekturvorgabe $\Delta x_f \in SE(3)$ ermittelt. Die kartesische Korrekturvorgabe Δx_f wird hierbei in Anlehnung an Methoden zur Nullkraftregelung bestimmt, so dass Kräfte oder Momente, welche durch unerwünschte Kontakte entstehen, ausgeglichen werden können (siehe Gleichung 2).

Die Regeldifferenz $\Delta x'$ wird nun über die differentielle inverse Kinematik des Roboters in Geschwindigkeitsvorgaben Δc für die einzelnen Gelenke umgesetzt. Der differentiellen Kinematik liegt die Pseudoinverse J^+ der Jacobi-Matrix J zugrunde. Die Jacobi-Matrix sowie deren Pseudoinverse sind nur für die aktuelle Gelenkwinkelkonfiguration c gültig und müssen somit in jedem Regelzyklus neu berechnet werden. Gleichung 3 beschreibt den Zusammenhang zwischen kartesischer Bewegungsvorgabe und resultierender Bewegung im Gelenkwinkelraum. Der auf diese Weise realisierte Regler entspricht einer P-Geschwindigkeitsregelung.

$$\Delta x' = \Delta x \cdot \Delta x_f \quad (2)$$

$$\Delta c = J^+(c) \cdot \Delta x' \quad (3)$$

4 Visuelle Wahrnehmung

Eine exakte und effiziente visuelle Wahrnehmung ist entscheidend für eine erfolgreiche Realisierung von *Visual-Servoing*-Ansätzen. Im Folgenden werden die eingesetzten Verfahren sowie deren Anwendung kurz umrissen.

4.1 Objekterkennung

Das visuelle Sensorsystem von ARMAR-III ermöglicht die kontinuierliche Erkennung und 6D-Lokalisierung von Objekten, die a priori eingelernt werden [9, 10]. Hierbei können sowohl einfarbige als auch texturierte Objekte verarbeitet werden. Weiterhin können Türgriffe erkannt und lokalisiert werden [11].

4.2 Visuelle Lokalisierung der Hand

Die visuelle Erkennung der Roboterhand, die sich auch in verschiedenen Konfigurationszuständen befinden kann, stellt eine große Herausforderung dar. Zur Bestimmung der Hand-Position (*Tool Center Point, TCP*) wird daher ein künstlicher Marker in Form einer roten

Kugel verwendet. Die Lokalisierung des Markers erfolgt ähnlich der Lokalisierung einfarbiger Objekte [6], wobei die erwartete Form und Größe im Kamerabild genutzt wird, um falsch erkannte Markerpositionen herauszufiltern und somit eine zuverlässige Erkennung zu garantieren. Wie in Bild 3 zu sehen ist, wird aus der Position des Markers x_{Marker} die Lage der Hand x_{Hand} bestimmt. Hierbei kommt ein hybrides Verfahren zum Einsatz, welches aus der visuell ermittelten Position des Markers, einem vordefinierten Offset zwischen Marker und TCP sowie der durch die Vorwärtskinematik des Roboters ermittelten Orientierung der Hand die Lage der Hand ermittelt.

Sichtverlust kann durch Verdeckung, Blickrichtungswechsel oder schlechte Bedingungen auftreten. Um die Regelung der Arme trotz nicht durchführbarer visueller Lokalisierung der Hand aufrecht zu erhalten, werden die Handposen während des Sichtverlusts geschätzt, indem die zu erwartende visuell ermittelte Lage $\hat{x}_{visuell}^{t+1}$ der Hand bestimmt wird. Hierzu wird der Unterschied Δx^t zwischen der zuletzt visuell ermittelten Lage $x_{visuell}^t$ und der durch die Vorwärtskinematik bestimmten Lage x_{kin}^t für die Zeit des Sichtverlustes als konstant angenommen.

$$\Delta x^t = (x_{kin}^t)^{-1} \cdot x_{visuell}^t, \quad (4)$$

$$\hat{x}_{visuell}^{t+1} = x_{kin}^{t+1} \cdot \Delta x^t. \quad (5)$$

4.3 Blickrichtung

Zur Steuerung der Blickrichtung stehen auf ARMAR-III unterschiedliche Ansätze zur Erzeugung von Bewegungen zum Blickrichtungswechsel zur Verfügung. Je nach Einsatz kann man sich auf die Bewegungsfreiheitsgrade des Halses beschränken, oder auch die kompletten Gelenke des Kopfes einschließlich der Augen einsetzen. Letzteres ermöglicht eine Vergrößerung des Bereiches, welcher visuell überwacht werden kann. Sollen die Gelenke der Augen eingesetzt werden, wird zusätzlich zu der Stereokalibrierung der Kameras, eine kinematische Kalibrierung der Kopfgelenke benötigt [12].

5 Visual Servoing für zweiarmige Aufgabenstellungen

Die bereits beschriebenen Ansätze zum positionsbasierten Visual Servoing werden in diesem Abschnitt erweitert, so dass zweiarmige Aufgabenstellungen realisiert werden können. Das Ziel ist es, gekoppelte zweiarmige Manipulationen zu ermöglichen, um beispielsweise große Objekte mit beiden Händen zu greifen. Weiterhin sollen nicht gekoppelte zweiarmige Aktionen durchgeführt werden können, um z. B. das gleichzeitige Greifen zweier Gegenstände zu ermöglichen.

5.1 Perzeption für zweiarmige Aufgaben

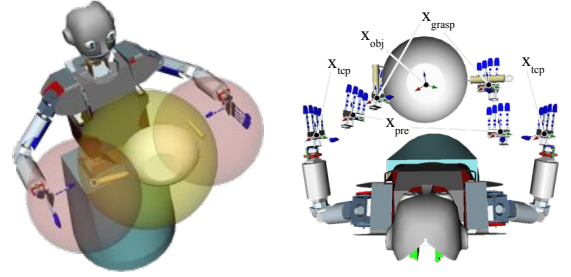


Bild 4: Drei Manipulationsbereiche von ARMAR-III (links). Ein Objekt mit Annäherungs- und Greifposen für die linke und die rechte Hand (rechts).

Das für die visuelle Perzeption genutzte Stereokamerasystem verfügt über ein begrenztes Blickfeld, weshalb die Ansätze zur Lokalisierung angepasst werden müssen, um einen größeren Aktionsbereich visuell überwachen zu können. Die Bildwinkel der Kameras sind hierbei auf einen kleinen Ausschnitt des möglichen Aktionsradius beschränkt, so dass Situationen auftreten, bei denen beide Hände nicht gleichzeitig lokalisiert werden können. Um trotz des eingeschränkten Sichtfeldes robuste, visuell überwachte Bewegungen zu ermöglichen, wird der aktive Kopf [13] des humanoiden Roboters ARMAR-III genutzt. In Bild 4 sind drei Aktionsbereiche für eine zweiarmige Manipulationen zu sehen. In Abhängigkeit von der auszuführenden Aktion sowie der Position der Hände werden diese Bereiche abwechselnd fokussiert, um die Lage der Hände bzw. Zielobjekte zu aktualisieren.

Bei schnellen Blickrichtungswechseln kommt es zu verschwommenen Kamerabildern (*motion blurring*). Um trotz dieses Effektes zuverlässige Lokalisierungsergebnisse zu erzielen, werden die Kamerabilder nur dann verarbeitet, wenn die Winkelgeschwindigkeit der Augen bei einem Blickrichtungswechsel unter einem bestimmten Grenzwert liegt.

5.2 Greifposen für zweihändige Manipulationen

Bild 4 zeigt rechts ein Objekt mit jeweils einer Annäherungs- und einer Greifpose (x_{pre} bzw. x_{grasp}) für die linke und die rechte Hand des Roboters. Die Posen x_{pre} und x_{grasp} sind zusammen mit der Objektrepräsentation in einer Datenbank abgelegt, und können beispielsweise mit Hilfe kinästhetischen Führens (engl. *kinesthetic teaching*) bei einer Nutzung der Nullkraftregelung der Arme eingelesen werden (siehe Bild 5). Bei einem Greifvorgang werden zuerst die Annäherungsposen x_{pre} als Zielvorgabe der Regelung verwendet. Fallen die translatorischen und rotatorischen Abstände zwischen dem TCP und x_{pre} unter einen Schwellwert, wird die Greifpose x_{grasp} als Ziel gewählt. Somit wird eine definierte Annäherungsrichtung beim Greifen realisiert.

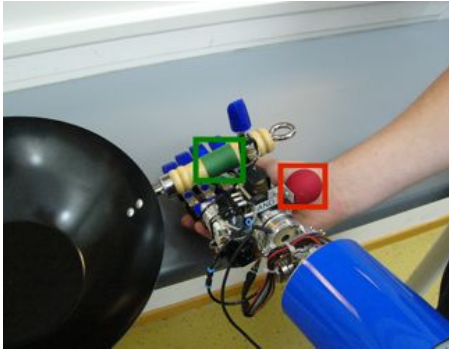


Bild 5: Einlernen von Griffen mithilfe einer Nullkraftregelung. Die Roboterhand wird hierbei durch den Menschen in die zu lernende Griffpose geführt. Anschließend wird die Relation der Roboterhand zum Objekt visuell bestimmt und als Griff in einer Datenbank abgelegt.

6 Anwendung auf ARMAR-III

Die folgenden Beispiele zeigen den Einsatz der hier beschriebenen Methoden auf dem humanoiden Roboter ARMAR-III. Basierend auf den vorgestellten Methoden können unterschiedliche Objekte, wie beispielsweise ein Becher oder eine Mülschachtel, zuverlässig gegriffen werden.

6.1 Einhändige Greifbewegungen

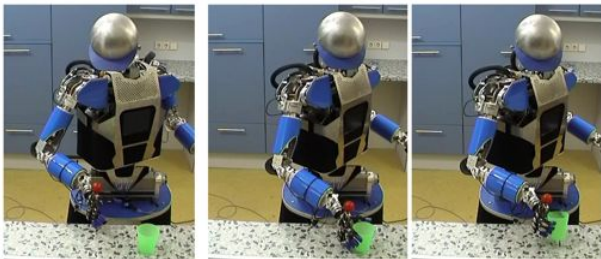


Bild 6: ARMAR-III greift einen Becher.



Bild 7: Die interne Sicht des Roboters ARMAR-III während des Greifens einer Schachtel.

Bild 6 zeigt die Ausführung auf dem Roboter. In Bild 8 sind die kartesischen Distanzen der Hand zur Zielposition aufgetragen. Da der Handmarker erst ab Zeitschritt 185 in den Kamerabildern lokalisiert werden kann (zuvor sind Objekt und Hand zu weit voneinander entfernt), ergibt sich zu diesem Zeitpunkt eine Korrektur der Handposition, da in den vorigen Zeitpunkten die Lage der Hand nur aus den kinematischen Berechnungen erfolgt ist.

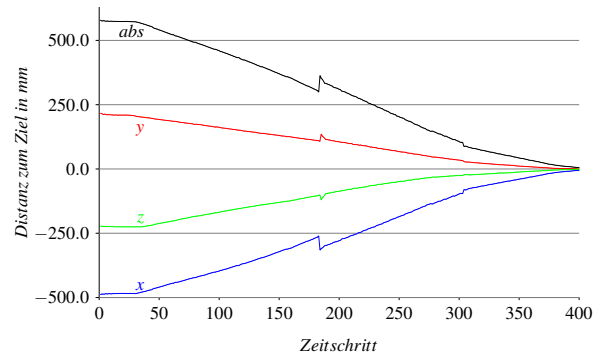


Bild 8: Die kartesischen Distanzen des TCP zur Zielposition, in blau (x), rot (y) und grün (z). Die Gesamtdistanz ist in schwarz aufgetragen. Bis zu Zeitschritt 185 sind Hand und Zielobjekt zu weit voneinander entfernt, um beide visuell zu erfassen, so dass die Position der Hand mit Hilfe der Vorwärtskinematik bestimmt wird. Durch die Erkennung der Hand und eine genauere Positionsbestimmung kommt es zu einem Sprung in der Regeldifferenz.

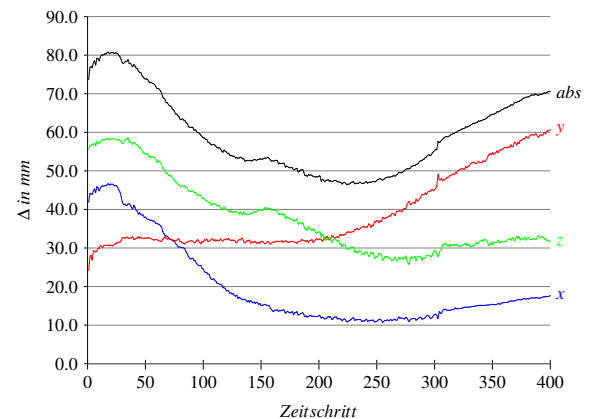


Bild 9: Unterschied zwischen kinematisch und visuell ermittelter TCP Position während der Ausführung (x in blau, y in rot, z in grün und Absolutwert in schwarz).

Bild 7 zeigt die interne Sicht des Roboters während eines Greifvorgangs. In Bild 9 ist der Unterschied zwischen visuell ermittelter Position des TCP und den durch die Vorwärtskinematik bestimmten Handpositionen dargestellt. Die hohe Differenz von bis zu 8 cm ist der langen kinematischen Kette zwischen den Koordinatensystemen der Kameras und der Hand geschuldet. Es ist weiterhin zu sehen, dass dieser Wert nicht konstant ist, und somit erst die kontinuierliche visuelle Überwachung eine erfolgreiche Ausführung von Greifaktionen ermöglicht.

6.2 Zweihändiges Greifen und Manipulieren

In dem hier vorgestellten Experiment wird ein großes Objekt (ein Wok) mit beiden Händen gegriffen. Der Wok steht hierbei vor dem Roboter und beide Hände sowie die Griffe des Woks werden visuell erfasst. In Bild 11 ist der Ablauf einer Greifaktion dargestellt. Zu sehen ist die Lokalisierung der Marker und Griffe während der Ausführung (links) sowie eine Abbildung des Roboters

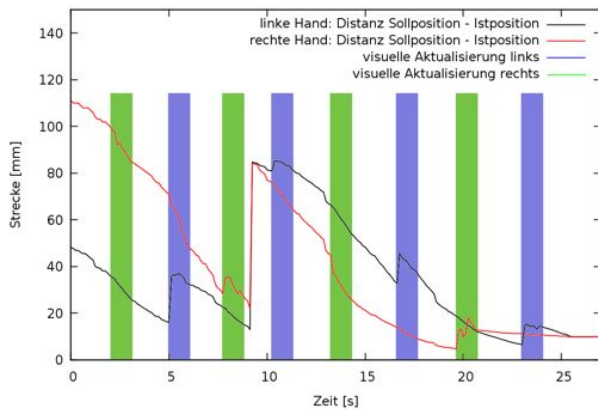


Bild 10: Distanzen der Soll- und Ist-Positionen beider Hände während der visuellen Positionsregelung. Bei Zeitindex 9s findet der Übergang von Pre- zur Greifpose statt. Erkennbar sind die zeitlich versetzten Aktualisierungen der linken und rechten Distanzen bedingt durch die Blickwechsel.

nachdem das Objekt gegriffen wurde (rechts).

Die während des zweiarmigen Greifens auftretenden Distanzen zwischen Soll- und Ist-Positionen sind in Bild 10 im Zusammenhang mit den Blickwechseln dargestellt. Hierbei können während den farbig hinterlegten Zeitabschnitten die Lage der linken Hand (grün) bzw. der rechten Hand (blau) visuell ermittelt werden. Zwischen diesen Phasen findet ein Blickwechsel statt und es kann keine zuverlässige Lokalisierung durchgeführt werden. Es ist erkennbar, dass bei Wiederherstellung des Sichtkontakts durch die Blickwechsel die geschätzte Handposition fehlerhaft ist und durch die visuelle Bestimmung korrigiert wird. Der dadurch entstandene Positionierungsfehler wird ausgegeregelt sobald wieder Sichtkontakt besteht.

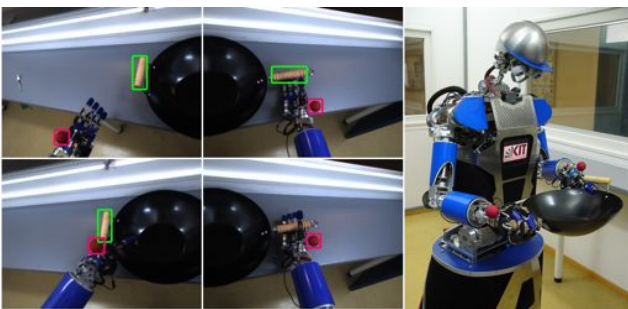


Bild 11: Der Wok wird simultan mit beiden Händen gegriffen, wobei die Hand über den Marker lokalisiert wird (rot) und über die Lage der Griffe (grün) die Greifposen ermittelt werden. Die Zielposen sind relativ zum Objekt definiert und wurden über Nullkraftregelung eingeplant.

In einem weiteren Experiment werden beide Arme genutzt, um einen Becher und einen Getränkekarton mit der rechten und linken Hand zu greifen. Anschließend wird eine vorgegebene Trajektorie ausgeführt. Durch den Einsatz von Visual Servoing kann die Trajektorie mit hoher Genauigkeit bezüglich der relativen Posen



Bild 13: Evaluierung des Rahmenwerks bei der Ausführung einer komplexen Tätigkeit: Bringen eines Objektes aus dem Kühlschrank

der beiden Objekte zueinander wiedergegeben werden und somit wird die Realisierung eines Einschenkvorgangs ermöglicht (siehe Bild 12).

Bild 13 zeigt den Einsatz des Rahmenwerks bei der Ausführung einer komplexen Tätigkeit. Auf den Sprachbefehl, einen Getränkekarton zu bringen, fährt ARMAR-III zum Kühlschrank und öffnet diesen. Das Lokalisieren und Greifen des Türgriffes findet hierbei über das vorgestellte Verfahren statt. Nachdem die Kühlschranktür kraftgeregelt geöffnet wurde, wird das gewünschte Objekt im Kühlschrank lokalisiert, gegriffen und entnommen. Ein anschließender Umgreifvorgang, bei dem das Objekt von der rechten in die linke Hand übergeben wird, erlaubt ein anschließendes Schließen der Kühlschranktür mit der freigewordenen rechten Hand. Für den Vorgang des Umgreifens wird das vorgestellte zweiarmige Visual Servoing eingesetzt.

7 Fazit

Der Einsatz von Visual Servoing ermöglicht eine zuverlässige Ausführung von ein- und zweiarmigen Greifbewegungen mit einem humanoiden Roboter. Durch die kontinuierliche Bestimmung der aktuellen Handpose und der gewünschten Zielpose können Positionierungsfehler ausgeglichen werden, ohne dass hierzu eine explizite Hand-Auge Kalibrierung notwendig ist. Voraussetzung für den erfolgreichen Einsatz ist ein effizientes und zuverlässiges Bildverarbeitungssystem, welches auf ARMAR-III mit dem *Integrating Vision Toolkit* (IVT) zur Verfügung steht [14].

Eine mögliche Verbesserung des hier vorgestellten Ansatzes lässt sich durch eine erweiterte Umsetzung der Handlokalisierung erreichen, um eine vollständige Ermittlung der aktuellen Lage des TCPs aus visuellen Daten zu erreichen. Hierdurch ließe sich die Hand-Objekt-Relation noch genauer bestimmen und somit könnte eine genauere Positionierung der Hand, beispielsweise für Präzisionsgriffe, realisiert werden. Weitere Anwendungsfelder ergeben sich in einer Mensch-Roboter-Kooperation, bei der beispielsweise bewegte Objekte bei einer Übergabe von Mensch zu Roboter mit dem hier vorgestellten Ansatz gegriffen werden.

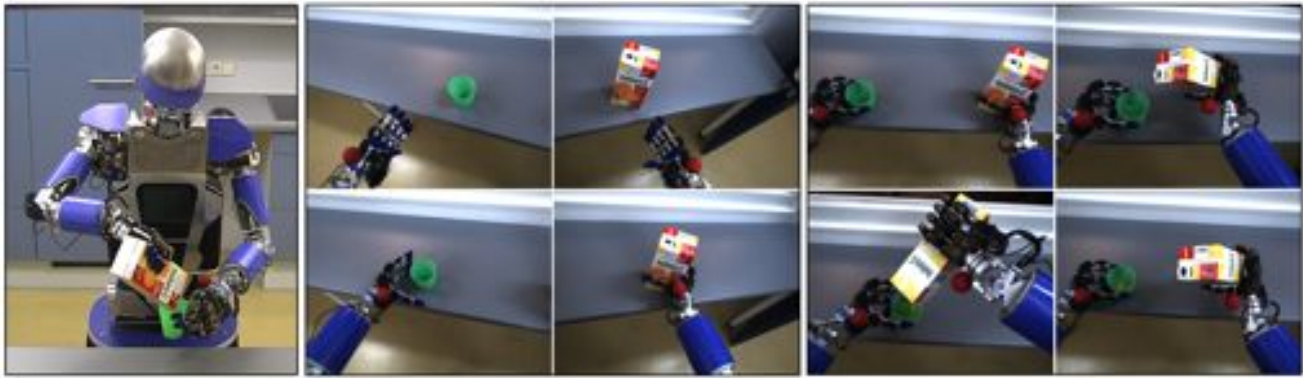


Bild 12: Der humanoide Roboter ARMAR-III nutzt Visual Servoing, um ein Getränk einzuschenken.

Literatur

- [1] T. Asfour, K. Regenstein, P. Azad, J. Schröder, A. Bierbaum, N. Vahrenkamp, and R. Dillmann, "Armar-III: An integrated humanoid platform for sensory-motor control." in *IEEE-RAS International Conference on Humanoid Robots (Humanoids 2006)*, December 2006, pp. 169–175.
- [2] I. Gaiser, S. Schulz, A. Kargov, H. Klosek, A. Bierbaum, C. Pylatiuk, R. Oberle, T. Werner, T. Asfour, G. Bretthauer, and R. Dillmann, "A new anthropomorphic robotic hand," in *IEEE-RAS International Conference on Humanoid Robots (Humanoids 2008)*, Daejeon, Korea, 2008.
- [3] L. Weiss, A. C. Sanderson, and C. P. Neuman, "Dynamic sensor-based control of robots with visual feedback," *IEEE Journal on Robotics and Automation*, vol. RA-3, no. 5, October 1987.
- [4] W. Wilson, C. W. Hulls, and G. Bell, "Relative end-effector control using cartesian position based visual servoing," *IEEE Transactions on Robotics and Automation*, vol. 12, pp. 684–696, Oct. 1996.
- [5] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 d visual servoing," *IEEE Transaction on Robotics and Automation*, vol. 15, no. 2, pp. 234–246, April 1999.
- [6] N. Vahrenkamp, S. Wieland, P. Azad, D. Gonzalez, T. Asfour, and R. Dillmann, "Visual servoing for humanoid grasping and manipulation tasks," in *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*, Dec. 2008, pp. 406–412.
- [7] N. Vahrenkamp, C. Böge, K. Welke, T. Asfour, J. Walter, and R. Dillmann, "Visual servoing for dual arm motions on a humanoid robot," in *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*, Dec. 2009, pp. 208–214.
- [8] N. Vahrenkamp, "Bewegungsplanung und sensorgestützte ausführung für das greifen auf humanoiden robotern." Ph.D. dissertation, Karlsruhe Institute of Technology (KIT), 2011.
- [9] P. Azad, T. Asfour, and R. Dillmann, "Stereo-based 6d object localization for grasping with humanoid robot systems," in *International Conference on Intelligent Robots and Systems (IROS)*, San Diego, USA, 2007.
- [10] P. Azad, T. Asfour, and R. Dillmann, "Accurate shape-based 6-dof pose estimation of single-colored objects," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2009, pp. 2690–2695.
- [11] D. Gonzalez-Aguirre, S. Wieland, T. Asfour, and R. Dillmann, "On environmental model-based visual perception for humanoids," in *14th Iberoamerican Congress on Pattern Recognition (CIARP)*, 2009, pp. 901–909.
- [12] K. Welke, M. Przyblyski, T. Asfour, and R. Dillmann, "Kinematic calibration for saccadic eye movements," Universität Karlsruhe (TH), Fak. f. Informatik, Institute for Anthropomatics, Tech. Rep., 2008.
- [13] T. Asfour, K. Welke, P. Azad, A. Ude, and R. Dillmann, "The karlsruhe humanoid head," in *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*, Dec. 2008, pp. 447–453.
- [14] P. Azad, "Integrating Vision Toolkit (IVT)," <http://ivt.sourceforge.net>, [Online; accessed 01-Jan-2012].

Manuskripteingang: 16. Januar 2012.