



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

## **Ausarbeitung - Anwendungen 2**

Carsten Fries

Kameragesteuerter Knickarmroboter zur Lokalisierung von  
Gegenständen

Carsten Fries

Carsten.Fries@haw-hamburg.de

Kameragesteuerter Knickarmroboter zur Lokalisierung von  
Gegenständen

Ausarbeitung eingereicht im Rahmen der Veranstaltung Anwendungen 2

im Studiengang Informatik (Master)  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuer:

Prof. Dr.-Ing. Andreas Meisel

Prüfer:

Prof. Dr. rer. nat. Kai von Luck

Prof. Dr. rer. nat. Gunter Klemke

Abgegeben am 18. August 2010

**Carsten Fries**

**Thema der Ausarbeitung**

Kameragesteuerter Knickarmroboter zur Lokalisierung von Gegenständen

**Stichworte**

Objekterkennung, Lokalisierung, SURF, SIFT, Skalierungsinvariante Merkmale, Merkmalsextraktion, Greifarm, Knickarm-Roboter

**Kurzzusammenfassung**

In dieser Arbeit werden drei Lösungsansätze, zur autonomen Identifizierung und Lokalisierung von Gegenständen, vorgestellt. Als Grundlage dient ein kameragesteuerter Knickarmroboter. Dieser liefert Bildinformationen, welche sogenannte invariante Merkmale bezüglich Rotation, Translation, Skalierung, Änderung der Lichtverhältnisse und teilweise affiner Verzerrung beinhalten. Diese Robustheit gegenüber Veränderungen ermöglicht es, Merkmale zuverlässig in perspektivisch unterschiedlichen Kameraaufnahmen wiederzuerkennen. Damit kann die räumliche Pose eines Gegenstandes, im sichtbaren- und greifbaren Umfeld des Knickarmroboters, erfasst werden.

**Carsten Fries**

**Title of the paper**

Camera-articulated robot for object localization

**Keywords**

object detection, localization, SURF, SIFT, scale-invariant features, feature extraction, manipulator-arm

**Abstract**

This work focuses on three approaches that autonomously identify and locate objects. The approaches are based on a camera-articulated robot. The camera images include features which are invariant to rotation, translation, scale, illumination changes and partially invariant to affine distortion. This allows features to be recognized in different perspective images. This leads to the resolution that objects which are positioned in the visible and tangible environment of the robot can be gathered.

„Computer sind das bis heute genialste Werk menschlicher Faulheit.“

*(Slogan einer IBM Werbekampagne in den '70iger Jahren)*

# Inhaltsverzeichnis

<b>1 Problemstellung</b>	<b>1</b>
<b>2 Lösungsansätze</b>	<b>2</b>
2.1 Modellbasierte Posenbestimmung . . . . .	3
2.1.1 Initiale Bestimmung von Weltpunkten . . . . .	3
2.1.2 Weltpunkte aus weiteren Aufnahmen . . . . .	5
2.1.3 Berechnung der Pose . . . . .	6
2.2 Visual Servoing . . . . .	6
2.3 Kombiniertes Ansatz . . . . .	8
<b>3 Zusammenfassung und Ausblick</b>	<b>10</b>
<b>Abbildungsverzeichnis</b>	<b>11</b>
<b>Quellenverzeichnis</b>	<b>12</b>

# 1 Problemstellung

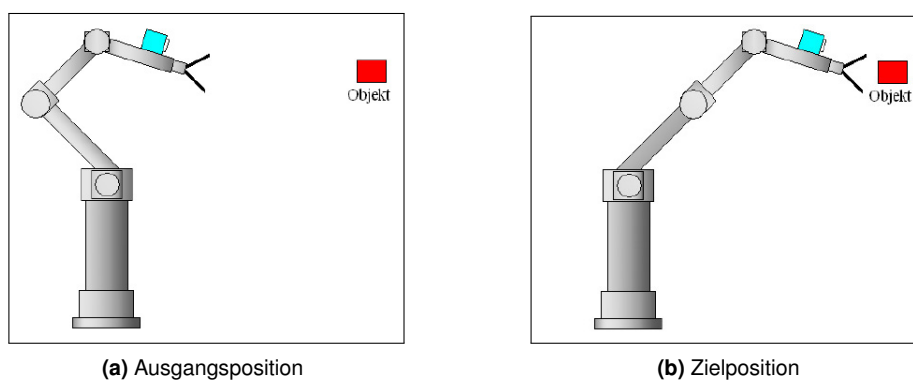
In dieser Arbeit werden drei Lösungsansätze vorgestellt, welche in Abhängigkeit eines kameragesteuerten Knickarmroboters Gegenstände autonom identifizieren und lokalisieren können.

Die Verfahren beruhen auf den nachfolgend aufgelisteten Knickarmabhängigkeiten.

- Bewegungsgeschwindigkeit von  $90^\circ$  pro Sekunde
- Ortsgebundener Roboterarm mit einem Bewegungsradius von 60cm
- Maximale Nutzlast von 500g
- Am Endeffektor montierte Kamera mit einem Weitwinkelobjektiv

Das Verfahren soll merkmalsbasiert arbeiten. Des Weiteren müssen sich Gegenstände im sichtbaren und greifbaren Umfeld des Knickarmroboters befinden.

Die nachfolgende Abbildung 1.1 veranschaulicht die in dieser Arbeit zu behandelnde Problemstellung. Die Grafik 1.1A stellt die Ausgangsposition des Knickarmroboters dar. Es soll der rot-dargestellte Gegenstand erkannt und lokalisiert werden. Die türkis-markierte Kamera liefert Bildinformationen. Anhand dieser Aufnahmen soll der Gegenstand identifiziert und lokalisiert werden, sodass sich der Knickarmroboter in die Zielposition 1.1B bewegen und den Gegenstand im anschließenden greifen kann.<sup>1</sup>



**Abbildung 1.1:** Identifizierung und Lokalisierung eines Gegenstandes.<sup>2</sup>

Das nachfolgende Kapitel 2 beschreibt Verfahren zur autonomen Erfassung und Lokalisierung von Gegenständen.

<sup>1</sup>Die (inverse) Kinematik, Trajektorienplanung sowie das Greifen ist nicht Bestandteil dieser Arbeit.

<sup>2</sup>Bildquelle: WAGNER (2009B)

## 2 Lösungsansätze

### Inhalt

---

2.1	Modellbasierte Posenbestimmung . . . . .	3
2.1.1	Initiale Bestimmung von Weltpunkten . . . . .	3
	Suche nach Merkmalskorrespondenzen . . . . .	3
	Berechnung der Fundamentalmatrix . . . . .	4
	Bestimmung der Projektionsmatrizen . . . . .	5
	Berechnung von 3D-Weltpunkten . . . . .	5
2.1.2	Weltpunkte aus weiteren Aufnahmen . . . . .	5
2.1.3	Berechnung der Pose . . . . .	6
2.2	Visual Servoing . . . . .	6
2.3	Kombinierter Ansatz . . . . .	8

---

In diesem Kapitel werden drei Verfahren, zur merkmalsbasierten Erfassung und Lokalisierung von Gegenständen, beschrieben.

Aufgrund des zu verwendenden Weitwinkelobjektivs sind die Aufnahmen bei der verwendeten Kamera stark verzerrt. Für alle drei Lösungsansätze ist es jedoch von Nöten, dass die tonnenförmige Verzeichnung für die Merkmalspunkte<sup>3</sup> rausgerechnet wird. Würde keine Entzerrung stattfinden, könnten die Merkmalspunkte nur schwer in anderen Aufnahmen wiedererkannt werden. Wenn ein Merkmalsvergleich stattfindet, wird deshalb in den jeweiligen Lösungsverfahren eine Entzerrung der Merkmale vorgenommen.

Die Lösungsansätze basieren auf den bekannten intrinsischen Kameraparametern. Oftmals sind diese vom Hersteller bekannt. Aufgrund der abweichenden Kameraparameter innerhalb der gleichen Baureihe ist eine Kalibrierung dennoch notwendig. Aus diesem Grund wird mit einer Kalibrierung sichergestellt, dass die intrinsischen Kameraparameter spätere Berechnungen nicht verfälschen.

Die intrinsischen Parameter beschreiben die interne Geometrie einer Kamera und somit Eigenschaften wie

- Brennweite
- Bildhauptpunktverschiebung
- Pixelskalierung
- Linsenverzeichnung

---

<sup>3</sup>Ein Merkmalspunkt wird u.a. durch einen Merkmalsvektor beschrieben und ist ein robuster Punkt im Kamerabild. Dieser besitzt invariante Eigenschaften bezüglich Rotation, Translation, Skalierung, Änderung der Lichtverhältnisse und teilweise affiner Verzerrung (Vgl. FRIES (2009)).

Dass in Matlab implementierte Kalibrierverfahren „Camera Calibration Toolbox“<sup>4</sup> kann u.a. für die Kalibrierung verwendet werden. Es basiert auf einem bekannten Kalibrierkörper, welches ein Schachbrettmuster dargestellt. Für den Kalibriervorgang sind mindestens zwei, jedoch möglichst viele Aufnahmen aus verschiedenen Perspektiven des Kalibrierkörpers aufzunehmen. Anhand dieser perspektivisch unterschiedlichen Aufnahmen werden die intrinsischen Kameraparameter berechnet.

Nachträgliche Änderungen der Kameraeigenschaften wie Auflösung oder Zoom, führen, aufgrund veränderter intrinsischer Kameraparameter, zu einer zwangsweisen Neukalibrierung.

## 2.1 Modellbasierte Posenbestimmung

Das erste Verfahren, zur Ermittlung der relativen räumlichen Orientierung eines Gegenstandes, basiert auf einem bekannten 3D-Merkmalmodell.

### 2.1.1 Initiale Bestimmung von Weltpunkten

Sind die intrinsischen Kameraparameter bekannt, können die ersten 3D-Oberflächenpunkte eines Gegenstandes ermittelt werden. 3D-Oberflächenpunkte werden im Folgenden als sogenannte „Weltpunkte“ bezeichnet. Für die Berechnung der ersten Weltpunkte sind zunächst zwei perspektivisch unterschiedliche Ansichten eines Gegenstandes aufzunehmen. Diese werden auf robuste Merkmale und Korrespondenzen zwischen den beiden Kameraaufnahmen untersucht.

#### Suche nach Merkmalskorrespondenzen

Die skaleninvariante Merkmalsextraktion (engl. Scale-invariant feature transform, kurz SIFT)<sup>5</sup> und SURF (engl. Speeded Up Robust Features)<sup>6</sup> sind zwei gebräuchliche Verfahren zur Suche nach robusten Merkmalen, welche eine Invarianz bezüglich

- Rotation
- Translation
- Skalierung
- Änderung der Lichtverhältnisse
- Teilweiser affiner Verzerrung

---

<sup>4</sup>Die „Camera Calibration Toolbox“ ist erhältlich unter BOUGUET (2008)

<sup>5</sup>Veröffentlichung des SIFT-Algorithmus: LOWE (1999)

Weiterentwicklungen sind in LOWE (2001), BROWN UND LOWE (2002) und LOWE (2004) beschrieben.

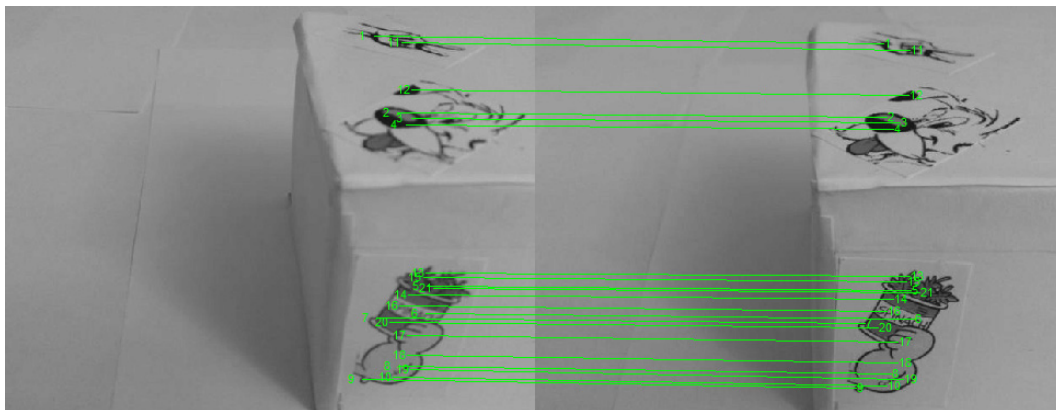
<sup>6</sup>Wurde erstmals in BAY U. A. (2006B) veröffentlicht. SURF besitzt im Vergleich zu SIFT einen ähnlichen Algorithmus. Er hat jedoch einen deutlichen Geschwindigkeitsvorteil. Dieser wird durch eine effizientere Suche nach Merkmalskandidaten erzielt. Detaillierte Informationen bezüglich SURF sind in BAY U. A. (2006A) und BAY U. A. (2008) zu finden.



aufweisen.

Ein Merkmal beschreibt einen markanten Punkt in einem Bild. Das Merkmal besteht aus einem Merkmalsvektor, einer Hauptorientierung und der Bildkoordinate des markanten Punktes. Ein Merkmalsvergleich zwischen zwei Ansichten kann anhand des kleinsten euklidischen Abstands bezüglich der Merkmalsvektoren durchgeführt werden (Vgl. FRIES (2009)<sup>7</sup>).

Die nachfolgende Abbildung 2.1 stellt ermittelte Merkmalskorrespondenzen dar.



**Abbildung 2.1:** Ermittelte Merkmalskorrespondenzen zwischen zwei perspektivisch unterschiedlichen Ansichten.

### Berechnung der Fundamentalmatrix

Nehmen zwei voneinander abweichend positionierte Kameras ein Objekt im dreidimensionalen Raum auf, entstehen zwei unterschiedliche zweidimensionale Bilder. Die Fundamentalmatrix  $F$  beschreibt die relative Orientierung zwischen diesen Bildern.

Mit Hilfe des 8-Punkte-Algorithmus<sup>8</sup> kann die Fundamentalmatrix berechnet werden (vgl. HARTLEY UND ZISSERMAN (2004) S. 281). Das RANSAC-Verfahren<sup>9</sup> wird gängigerweise herangezogen, um die Fundamentalmatrix mit Hilfe der verfügbaren Merkmalskorrespondenzen und dem 8-Punkte-Algorithmus optimal zu schätzen. Bei dieser Schätzung der Fundamentalmatrix werden, mit einer hohen Wahrscheinlichkeit, alle Fehlzuordnungen aufgedeckt. Diese falschen Merkmalskorrespondenzen werden als „Outlier“ bezeichnet. Dagegen besitzen sogenannte „Inlier“ eine richtige Zuordnung.

<sup>7</sup>In „Objekterkennung mit SIFT-Merkmalen“ wurde das SIFT-Verfahren für die Erfassung robuster Merkmale herangezogen und untersucht.

<sup>8</sup>Der 8-Punkte-Algorithmus wurde erstmals in LONGUET-HIGGINS (1981) vorgestellt. Dieser benötigt mindestens acht korrespondierende Punkte zwischen zwei Bildern. An dieser Stelle sei auf den 7-Punkte-Algorithmus verwiesen, welcher lediglich sieben Korrespondenzen benötigt. Informationen bezüglich des 7-Punkte-Algorithmus sind in HARTLEY UND ZISSERMAN (2004) auf S. 281 zu finden.

<sup>9</sup>RANSAC steht für „RANdom SAmple Consensus“. Detaillierte Informationen bezüglich des RANSAC-Verfahrens sind in RODEHORST (2004) ab S. 143 zu finden.

### Bestimmung der Projektionsmatrizen

Eine Projektionsmatrix bildet 3D-Weltpunkte auf 2D-Bildpunkte einer Kameraaufnahme ab. Die Projektionsmatrix kann anhand der ermittelten Fundamentalmatrix, den intrinsischen Kameraparametern und den Merkmalskorrespondenzen berechnet werden. Dafür wird die Fundamentalmatrix zusammen mit den intrinsischen Kameraparametern zu einer sogenannten „Essentialmatrix“ erweitert (Vgl. EBERT (2009), S. 35). Anschließend können die Projektionsmatrizen  $P_1$  und  $P_2$  der ersten beiden Kameraaufnahmen, auf Basis der Singulärwertzerlegung<sup>10</sup>, ermittelt werden.

### Berechnung von 3D-Weltpunkten

Sind die Projektionsmatrizen  $P_1$  und  $P_2$  vorhanden, dann kann der dreidimensionale Ursprung für alle Merkmalskorrespondenzen zurückgerechnet werden. Dafür sind mit Hilfe der Projektionsmatrizen die Merkmalskorrespondenzen zu triangulieren. Die Triangulation erfolgt durch eine Minimierung des Rückprojektionsfehlers eines 3D-Weltpunktes zu seinen zwei Bildpunkten (Vgl. EBERT (2009), S. 39). Das Ergebnis der Triangulation sind die gesuchten ersten 3D-Weltpunkte von einem Gegenstand.

Des Weiteren sind jedem 3D-Weltpunkt seine Merkmale zugeordnet. Dies ermöglicht in einem späteren Merkmalsvergleich, welcher zwischen zwei Bildern durchgeführt wird, dass der dreidimensionale Ursprung und damit die räumliche Lage eines Merkmales bekannt ist.

#### 2.1.2 Weltpunkte aus weiteren Aufnahmen

Für einen hohen Wiedererkennungsgrad eines Gegenstandes müssen Merkmale aus weiteren Aufnahmen in das bereits bestehende 3D-Merkmalmodell hinzugefügt werden. Dafür ist eine zu den ersten zwei Bildern ( $B_1, B_2$ ) abweichende perspektivische Aufnahme  $B_3$  aufzunehmen. Um neue Weltpunkte für das 3D-Merkmalmodell berechnen zu können, wird zuerst die Projektionsmatrix  $P_3$  des aktuellen Bildes  $B_3$  benötigt. Wird ein Merkmalsvergleich zwischen der Aufnahme  $B_3$  und den Merkmalen aller Weltpunkte durchgeführt, dann kann auf Basis dieser 3D-Welt- zu 2D-Bild-Korrespondenzen die Projektionsmatrix  $P_3$  berechnet werden. Die sogenannte „Direct Linear Transformation (DLT)“ ermöglicht die Berechnung von  $P_3$  mit Hilfe der 3D-2D-Korrespondenzen (Vgl. EBERT (2009), S. 47).

Im Anschließendenden sind die Merkmale zwischen der aktuellen Aufnahme  $B_3$  und der vorherigen Aufnahme  $B_2$  zu vergleichen. Alle Merkmalskorrespondenzen welche zu einem noch nicht vorhandenen Weltpunkt führen, sind mit dem Triangulationsverfahren (Siehe „Berechnung von 3D-Weltpunkten“ auf S. 5) zu berechnen.

Wenn viele Aufnahmen flächendeckend um einen Gegenstand herum aufgenommen wurden, dann ist ein komplettes 3D-Merkmalmodell von diesem Gegenstand möglich. Grundsätzlich gilt, je mehr Merkmale im 3D-Modell vorhanden sind, desto höher ist die Wahrscheinlichkeit, dass die räumliche Lage eines Gegenstandes auf Basis eines Bildes berechnet werden kann.

<sup>10</sup>Die Singulärwertzerlegung (engl. „Singular Value Decomposition“, kurz SVD) ist in RODEHORST (2004) auf S. 151 beschrieben.

### 2.1.3 Berechnung der Pose

Gesucht sei die Pose eines Gegenstandes in der räumlichen Umgebung des Knickarmroboters. Gegeben ist lediglich ein Kamerabild aus einer beliebigen Perspektive. Für die Ermittlung der Gegenstandsposition und -ausrichtung ist die einzige Bedingung, dass sich der Gegenstand im Kamerabild befinden muss. Ist diese Vorbedingung erfüllt, kann die Position eines Gegenstandes erkannt werden. Dafür sind zu Beginn die Merkmale der aktuellen Aufnahme mit dem 3D-Merkmalmodell zu vergleichen. Eine anschließende Berechnung der Projektionsmatrix mit den 3D-2D-Korrespondenzen erzielt die gesuchten Informationen bezüglich der Raumpose des Gegenstandes. Denn aus der Projektionsmatrix folgt die Rotation und Translation zwischen der Kamera (-aufnahme) und dem im Raum positionierten Gegenstand.

## 2.2 Visual Servoing

In WAGNER (2009A) wird das sogenannte „2D Visual Servoing“-Verfahren, zur Lösung der in dieser Arbeit gestellten Problemstellung, beschrieben. Dieses Verfahren arbeitet grundlegend in Abhängigkeit einer Regelung und wird im Folgenden erläutert.

Die am Endeffektor montierte Kamera nimmt in kontinuierlichen Abständen Aufnahmen auf. Eine Aufnahme enthält den zu lokalisierenden Gegenstand. Das Verfahren basiert auf einer vorab gespeicherten Zielansicht, aus der mindestens acht Gegenstandsmerkmale als Sollgröße  $s^*$  festzulegen sind. Damit setzt sich die Zielansicht durch die Orientierung von mindestens acht bekannten Gegenstandsmerkmalen (Siehe Merkmale, Detektion sowie Merkmalsvergleich ab S. 3) zusammen. Diesbezüglich besteht eine Merkmalsorientierung aus einer x- und y-Bildkoordinate. Aus Sicherheitsgründen wurden in WAGNER (2009A) 16 bekannte Merkmale detektiert und als Zielansicht definiert.

In Abbildung 2.2 ist der verwendete Regelkreis abgebildet.

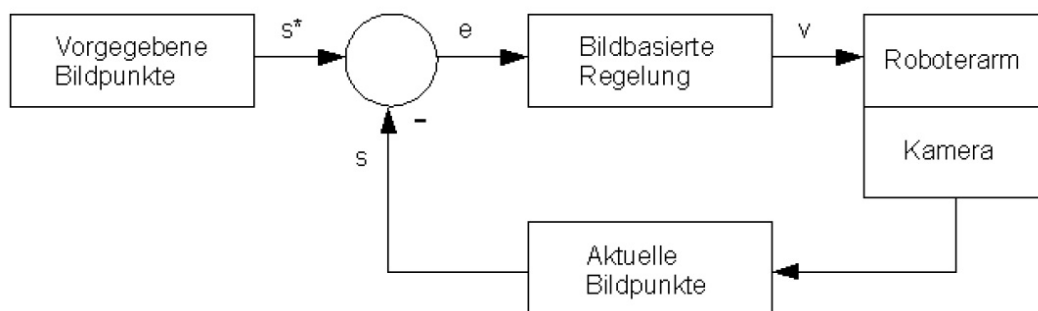


Abbildung 2.2: Regelkreisbasierte Posenbestimmung.<sup>11</sup>

In der Regelung ist die Sollgröße  $s^*$  die Orientierung der bekannten Merkmale aus der Zielansicht. Die Istgröße  $s$  ist die aktuelle Orientierung dieser Merkmale. Die Differenz der Ist- und

<sup>11</sup>Bildquelle: WAGNER (2009A)

Sollgröße gibt die momentane Regelabweichung  $e$  an. Aus der Regelabweichung wird die Bewegung des Knickarmroboters in Richtung zur Zielansicht abgeleitet<sup>12</sup>. Die Regelabweichung wird dabei im Verlauf minimiert. Der Knickarmroboter nähert sich dadurch der Zielansicht (Regelabweichung  $\approx 0$ ) an und lokalisiert letztendlich den gesuchten Gegenstand.

Befindet sich die Zielansicht jedoch zu weit von der Startansicht entfernt, kann das Verfahren instabil werden. Die Instabilität wirkt sich mit einer Entfernung zum gesuchten Gegenstand aus (Vgl. CHAUMETTE UND HUTCHINSON (2006)). Wenn die Verdrehung der Kamera um die optische Achse zwischen der Start- und Zielpose zu groß ist, dann führt dies zur gleichen Problematik.

Das Problem könnte durch eine Triangulation (Siehe Kapitel 2.1.1 auf S. 5) der Merkmalskorrespondenzen  $s$  und  $s^*$  behoben werden. Denn mit Hilfe einer Triangulation kann der zu zwei Merkmalskorrespondenzen zugehörige Raumpunkt berechnet werden. Sind die Raumpunkte der Merkmale bekannt, kann mit einer 3D-2D-Korrespondenz die Kameraentfernung zum Gegenstand berechnet werden. Für diese Berechnung ist die „Direkt Linear Transformation (DLT)“ zu verwenden, welche auf S. 5 beschrieben wurde. Die Kenntnis über die Entfernung ermöglicht eine Überprüfung ob sich die Kamera, im Gegensatz zur letzten Aufnahme, zum gesuchten Gegenstand entfernt oder angenähert hat.

Zusammenfassend ist der „2D Visual Servoing“-Algorithmus durch folgende sechs Schritte beschrieben.

1. Zielansicht (Orientierung von mindestens acht Merkmalen) festlegen
2. Aufnahme eines Kamerabildes
3. Detektion der aktuellen Merkmale  $s$ , welche mit den vorgegebenen Merkmalen  $s^*$  korrespondieren
4. Triangulation der Merkmalskorrespondenzen  $s$  und  $s^*$  zur Entfernungsmessung
5. Bewegung des Knickarmroboters, anhand der Regelabweichung  $e$  und der Abstandswerte, in Richtung zur Zielansicht ermitteln und durchführen
6. Wiederhole ab 2. bis Regelabweichung  $\approx 0$

Der Vorteil eines solchen Regelkreises ist die kontinuierliche Korrektur der aktuellen Roboterposition. Dies führt zu einer sehr genauen Lokalisation eines gesuchten Gegenstandes. Des Weiteren ist das Verfahren robust gegenüber Kalibrierungsungenauigkeiten.

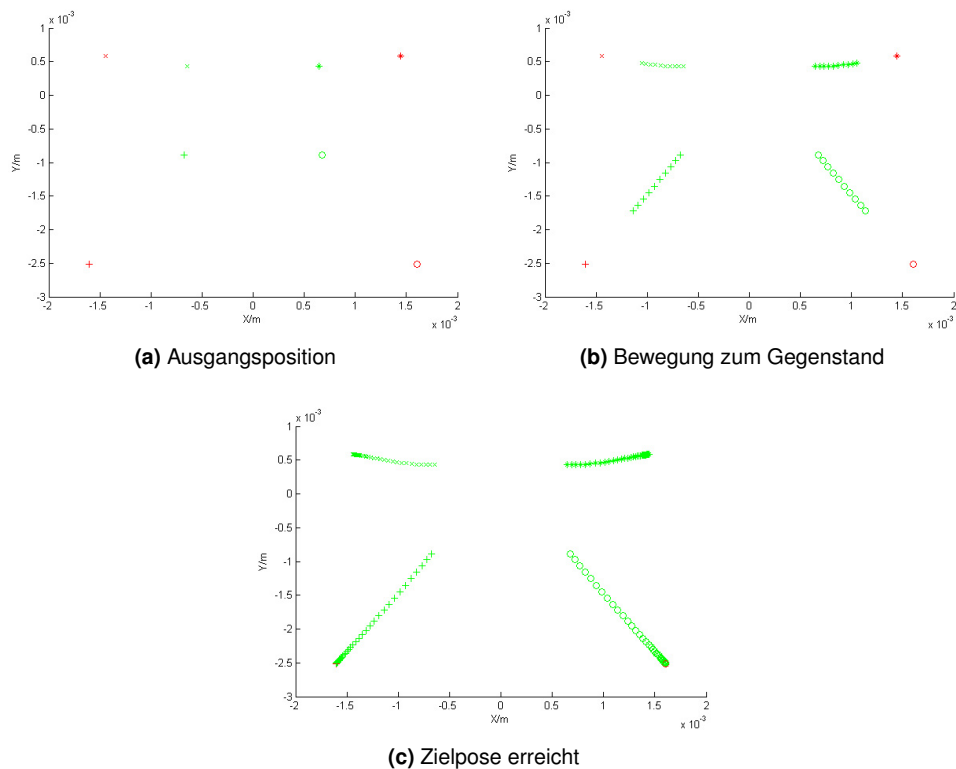
Die nachfolgende Abbildung 2.3 stellt das, mit der Matlab Robotics Toolbox<sup>13</sup> sowie der Epipolar Geometry Toolbox<sup>14</sup>, simulierte Verfahren dar. Aus Übersichtsgründen sind jeweils nur vier der 16 Orientierungsmerkmale dargestellt.

---

<sup>12</sup>Ein Regelgesetz (Vgl. CHAUMETTE UND HUTCHINSON (2006)) ermöglicht eine Abbildung der Regelabweichung  $e$  auf eine Roboterbewegung Richtung Zielansicht. Die Bewegung wird in Geschwindigkeiten  $v$  bezüglich des Knickarm-Basiskoordinatensystems angegeben. Das Basiskoordinatensystem befindet sich dabei im Fuß des Roboters.

<sup>13</sup>Die Robotics Toolbox for Matlab ist in CORKE (1996) zu finden.

<sup>14</sup>Beschrieben durch die Herausgeber in MARIOTTINI UND PRATTICIZZO (2005).



**Abbildung 2.3:** Simulierter Ablauf des „2D Visual Servoing“-Verfahrens.<sup>15</sup>

Die erste Grafik 2.3A stellt die Ausgangsposition des Knickarmroboters dar. Die grün-markierten Merkmale symbolisieren die detektierten Merkmale in der aktuellen Aufnahme. Die rot-dargestellten Merkmale kennzeichnen die Merkmale der Zielansicht. Im zweiten Bild 2.3B hat sich der Knickarmroboter auf Basis der Regelabweichung in Richtung zum gesuchten Gegenstand bewegt. Die bis dahin grün-markierten Merkmale symbolisieren den Bewegungsverlauf der Kamera sowie der Merkmalsorientierung. In der letzten Grafik 2.3C ist die Regelabweichung gegen 0 konvergiert. Der Knickarmroboter befindet sich in der Zielpose und hat damit den Gegenstand lokalisiert.

Weitere Informationen über das „2D Visual Servoing“-Verfahren sind aus WAGNER (2009A) und CHAUMETTE UND HUTCHINSON (2006) zu entnehmen.

## 2.3 Kombinerter Ansatz

Der modellbasierte Ansatz, zur Bestimmung der Pose eines Gegenstandes, besitzt eine hohe Geschwindigkeit. Die hohe Geschwindigkeit ist auf die einmalige Bildauswertung zurückzuführen. Wurde das aufgenommene Kamerabild analysiert, so kann die relative Pose des gesuchten Gegenstandes sofort ermittelt und angefahren werden. „Visual Servoing“ ist ein regelungsbasiertes Verfahren und benötigt dagegen mehrere Schritte um die Posenbestimmung durchzuführen. Dadurch ist das Verfahren deutlich langsamer.

<sup>15</sup>Bildquelle: WAGNER (2009A)

Wird die Genauigkeit betrachtet, dann weist der modellbasierte Ansatz gegenüber dem „2D Visual Servoing“-Verfahren eine geringere Genauigkeit auf. Dies ist auf die Regelung und die damit verbundene ständige Positionskorrektur zurückzuführen.

Ein kombinierter Ansatz der beiden Verfahren liegt daher nahe und wird im anschließenden Algorithmus beschrieben.

Zuerst wird für den modellbasierten Ansatz, in einem einmaligen Vorgang, ein 3D-Merkmalmodell erzeugt. Für „Visual Servoing“ dagegen wird eine Zielansicht (Orientierung von mindestens acht Merkmalen) definiert.

1. Vorbereitungen der Verfahren
  - a) Generierung des 3D-Merkmalmodells
  - b) Festlegen der Zielansicht

Im Anschluss der Vorbereitung kann ein Gegenstand initial mit dem modellbasierten Ansatz lokalisiert werden. Das Ergebnis der initialen Schätzung der Raumpose wird für die erste Bewegung zum gesuchten Gegenstand verwendet. In der letzten Phase wird auf Basis des regelbasierten Ansatzes die räumliche Position korrigiert.

1. Initiale modellbasierte Posenbestimmung
2. Bewegung zum Gegenstand
3. Korrektur der Pose mit „Visual Servoing“
  - a) Aufnahme eines Kamerabildes
  - b) Detektion der aktuellen Merkmale  $s$ , welche mit den vorgegebenen Merkmalen  $s^*$  korrespondieren
  - c) Triangulation der Merkmalskorrespondenzen  $s$  und  $s^*$  zur Entfernungsmessung
  - d) Bewegung des Knickarmroboters, anhand der Regelabweichung  $e$  und der Abstandswerte, in Richtung zur Zielpose ermitteln und durchführen
  - e) Wiederhole ab 3.a bis Regelabweichung  $\approx 0$

Der kombinierte Ansatz vereint den Geschwindigkeitsvorteil des modellbasierten Ansatzes sowie die hohe Genauigkeit des „Visual Servoing“-Verfahrens. Aus diesem Grund wird in fortlaufenden Arbeiten der kombinierte Ansatz zur Posenbestimmung weiter verfolgt.

### 3 Zusammenfassung und Ausblick

In dieser Arbeit wurden drei Ansätze, zur Erfassung und Lokalisierung eines Gegenstandes, beschrieben.

Der modellbasierte Ansatz zur Posenbestimmung wurde in Kapitel 2.1 als erstes Lösungsverfahren vorgestellt. Dieses benötigt für jeden Gegenstand ein bekanntes 3D-Merkmalmodell. Für eine möglichst hohe Wiedererkennungsratesollte das Merkmalsmodell aus vielen perspektivisch unterschiedlichen Aufnahmen berechnet werden.

Als zweites Verfahren wurde das „2D Visual Servoing“-Verfahren beschrieben (Vgl. Kapitel 2.2). Das Verfahren basiert auf einer Regelung und bewegt sich zum Zielgegenstand in mehreren kleinen Bewegungen. Die Regelung benötigt eine bekannte Zielansicht und vergleicht diese kontinuierlich mit der aktuellen Ansicht. Dabei beinhaltet die aktuelle Aufnahme mindestens acht bekannte Gegenstandsmerkmale, welche vorab aus der Zielansicht definiert wurden. Jedes dieser Merkmale besitzt eine Orientierung ( $x$  und  $y$  Bildkoordinate). Auf Basis der Differenz zwischen den aktuellen Merkmalsorientierungen und der Orientierungen der Merkmale aus der Zielansicht, wird die Bewegung zum gesuchten Gegenstand abgeleitet. Im Gegensatz zum modellbasierten Ansatz benötigt der „2D Visual Servoing“-Algorithmus, anstatt eines 3D-Merkmalmodells, eine Zielansicht. Diese beschreibt die Merkmalsorientierung bzw. wie der Endeffektor in der Zielsituation zu einem Gegenstand stehen soll.

Das „2D Visual Servoing“-Verfahren hat sich gegenüber dem modellbasierten Ansatz als langsameres Verfahren herausgestellt. Jedoch weist dieses eine höhere Genauigkeit auf. Ein dritter Lösungsansatz, welcher die Vorteile beider Verfahren vereint, wurde in Kapitel 2.3 vorgestellt.

Dieser lokalisiert in der ersten Phase, mit der modellbasierten Posenbestimmung, den gesuchten Gegenstand. Wurde der Gegenstand lokalisiert und angefahren, kann die zweite Phase zur Korrektur der Pose durchgeführt werden. Die Positionskorrektur erfolgt mit dem regelkreisbasierten „2D Visual Servoing“-Verfahren, welches aufgrund der Regelung eine sehr hohe Genauigkeit aufweist.

Für die in dieser Arbeit beschriebene Problemstellung, hat sich die Kombination der ersten beiden Verfahren als bester Lösungsansatz herausgestellt. Aus diesem Grund wird in weiterführenden Arbeiten der kombinierte Ansatz zur Posenbestimmung implementiert und evaluiert.

Auf Basis dieser Erkenntnisse kann ein autonomer Roboterarm zur Identifizierung, Lokalisierung und Manipulation von Gegenständen realisiert werden.

Für weitere Arbeiten ist ein mobiler Einsatz des Knickarmroboters in Betracht zu nehmen.

# Abbildungsverzeichnis

1.1	Identifizierung und Lokalisierung eines Gegenstandes . . . . .	1
2.1	Ermittelte Merkmalskorrespondenzen zwischen zwei perspektivisch unterschiedlichen Ansichten . . . . .	4
2.2	Regelkreisbasierte Posenbestimmung . . . . .	6
2.3	Simulierter Ablauf des 2D Visual Servoing Verfahrens . . . . .	8



# Quellenverzeichnis

## Literaturquellen

- [Bay u. a. 2008] BAY, Herbert ; ESS, Andreas ; TUYTELAARS, Tinne ; GOOL, Luc V.: *Speeded-Up Robust Features (SURF)*. Computer Vision and Image Understanding, 2008
- [Bay u. a. 2006a] BAY, Herbert ; FASEL, Beat ; GOOL, Luc V.: *Interactive Museum Guide: Fast and Robust Recognition of Museum Objects*. Computer Vision Laboratory (BIWI), 2006
- [Bay u. a. 2006b] BAY, Herbert ; TUYTELAARS, Tinne ; GOOL, Luc V.: *SURF: Speeded Up Robust Features*. Computer Vision (ECCV), 2006. – ISBN 978-3-540-33832-1
- [Beardsley u. a. 1996] BEARDSLEY, Paul ; TORR, Philip ; ZISSERMAN, Andrew: *3D Model Aquisition from Extended Image Sequences*. Springer-Verlag, 1996
- [Beis und Lowe 1997] BEIS, Jeffrey S. ; LOWE, David G.: *Shape indexing using approximate nearest-neighbour search in high-dimensional spaces*. Conference on Computer Vision and Pattern Recognition, 1997
- [Brown und Lowe 2002] BROWN, Matthew ; LOWE, David G.: *Invariant features from interest point groups*. British Machine Vision Conference, 2002
- [Chaumette und Hutchinson 2006] CHAUMETTE, Francois ; HUTCHINSON, Seth: *Visual servo control, Part I: Basic approaches*. IEEE Robotics and Automation Magazine 13, Nr. 4, S. 82-90, 2006
- [Corke 1996] CORKE, Peter I.: *A Robotics Toolbox for MATLAB*. IEEE Robotics and Automation Magazine 3, Nr. 1, S. 24-32,, 1996 ISSN 1070-9932
- [Cornelis und Gool 2008] CORNELIS, Nico ; GOOL, Luc V.: *Fast Scale Invariant Feature Detection and Matching on Programmable Graphics Hardware*. Computer Vision and Pattern Recognition, 2008
- [Ebert 2009] EBERT, Matthias: *Modellbasierte Posebestimmung aus 2-D/3-D SIFT-Korrespondenzen*. Diplomarbeit - Universität Koblenz, 2009
- [Fahrmeir u. a. 1996] FAHRMEIR, Ludwig ; HAMERLE, Alfred ; TUTZ, Gerhard: *Multivariate statistische Verfahren*. 1996. – ISBN 3-11-013806-9
- [Fries 2009] FRIES, Carsten: *Objekterkennung mit SIFT-Merkmalen*. Hochschule für Angewandte Wissenschaften Hamburg, 2009
- [Gordon und Lowe 2004] GORDON, Iryna ; LOWE, David G.: *Scene Modelling, Recognition and Tracking with Invariant Image Features*. International Symposium on Mixed and Augmented Reality (ISMAR) - University of British Columbia, 2004

- [Graf 2007] GRAF, Simone: *Kamerakalibrierung mit radialer Verzeichnung: Die radiale essentielle Matrix*. Dissertation - Universität Passau, 2007
- [Hartley und Zisserman 2004] HARTLEY, Richard I. ; ZISSERMAN, Andrew: *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press, 2004. – ISBN 0521540518
- [Hofmann 2009] HOFMANN, Robert: *Implementierung des SURF-Feature-Detektors auf der GPU mit Hilfe von CUDA*. Universität Koblenz, 2009
- [Krzanowski 2000] KRZANOWSKI, Wojtek J.: *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press, 2000. – ISBN 978-0198507086
- [Longuet-Higgins 1981] LONGUET-HIGGINS, Hugh C.: *A computer algorithm for reconstructing a scene from two projections*. University of Sussex, 1981
- [Lowe 1999] LOWE, David G.: *Object Recognition from Local Scale-Invariant Features*. International Conference on Computer Vision, 1999
- [Lowe 2001] LOWE, David G.: *Local Feature View Clustering for 3D Object Recognition*. International Conference on Computer Vision, 2001
- [Lowe 2004] LOWE, David G.: *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 2004
- [Mariottini und Prattichizzo 2005] MARIOTTINI, Gian L. ; PRATTICHIZZO, Domenico: *EGT: a Toolbox for Multiple View Geometry and Visual Servoing*. IEEE Robotics and Automation Magazine 3, Nr. 12, 2005
- [Rodehorst 2004] RODEHORST, Volker: *Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie*. wvb Wissenschaftlicher Verlag Berlin, 2004. – ISBN 978-3936846836
- [Torr 1995] TORR, Philip H. S.: *Outlier Detection and Motion Segmentation*. Dissertation - University of Oxford, 1995
- [Torr 2002] TORR, Philip H. S.: *Bayesian Model Estimation and Selection for Epipolar Geometry and Generic Manifold Fitting*. International Journal of Computer Vision, 2002
- [Torr u. a. 1999] TORR, Philip H. S. ; FITZGIBBON, Andrew W. ; ZISSERMAN, Andrew: *The Problem of Degeneracy in Structure and Motion Recovery from Uncalibrated Image Sequences*. International Journal of Computer Vision, 1999 ISSN 0920-5691
- [Torr und Murray 1993] TORR, Philip H. S. ; MURRAY, David W.: *Outlier Detection and Motion Segmentation*. SPIE volume 2059, 1993
- [Torr und Murray 1997] TORR, Philip H. S. ; MURRAY, David W.: *The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix*. International Journal of Computer Vision, 1997
- [Torr und Zisserman 2000] TORR, Philip H. S. ; ZISSERMAN, Andrew: *MLESAC: A New Robust Estimator with Application to Estimating Image Geometry*. Computer Vision and Image Understanding, 2000
- [Wagner 2009a] WAGNER, Benjamin: *Positionierung eines Roboterarms mittels bildbasierter Regelung*. Hochschule für Angewandte Wissenschaften Hamburg, 2009

- [Wagner 2009b] WAGNER, Benjamin: *Visuelle Regelung für einen Roboterarm*. Hochschule für Angewandte Wissenschaften Hamburg, 2009
- [Zhang 1999] ZHANG, Zhengyou: *Flexible Camera Calibration By Viewing a Plane From Unknown Orientations*. International Conference on Computer Vision, pages 666-673, 1999
- [Zhang 2000] ZHANG, Zhengyou: *A flexible new technique for camera calibration*. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 2000
- [Zimmermann 2007] ZIMMERMANN, Fabian: *Keypoint-Detektion und Deskriptoren-Berechnung auf der Grafikkarte*. Diplomarbeit - Technische Universität Kaiserslautern, 2007

## Internetquellen

- [Bouguet 2008] BOUGUET, Jean-Yves: *Camera Calibration Toolbox for Matlab*. Stand: 02.06.2008. 2008. –  
URL: [HTTP://WWW.VISION.CALTECH.EDU/BOUGUETJ/CALIB\\_DOC/](http://www.vision.caltech.edu/bouguetj/calib_doc/)
- [Canterakis 2007] CANTERAKIS, Nikos: *Computer Vision I*. Vorlesungsskript im Rahmen des Lehrstuhls für Mustererkennung der Universität Freiburg, Stand: 26.11.2007. 2007. –  
URL: [HTTP://LMB.INFORMATIK.UNI-FREIBURG.DE/LECTURES/COMPUTER\\_VISION/CVF\\_8.PDF](http://lmb.informatik.uni-freiburg.de/lectures/computer_vision/cvf_8.pdf)
- [Meisel 2008a] MEISEL, Prof. Dr.-Ing. A.: *3D Bildverarbeitung*. Vortragsfolien zu Grundlagen der 3D-Bildverarbeitung, Stand: 03.03.2008. 2008. –  
URL: [HTTP://WWW.INFORMATIK.HAW-HAMBURG.DE/1708.HTML](http://www.informatik.haw-hamburg.de/1708.html)
- [Meisel 2008b] MEISEL, Prof. Dr.-Ing. A.: *Robot Vision*. Vorlesungsskript des Faches Robot Vision, Stand: 07.03.2008. 2008. –  
URL: [HTTP://WWW.INFORMATIK.HAW-HAMBURG.DE/WP\\_ROBOT\\_VISION.HTML](http://www.informatik.haw-hamburg.de/wp_robot_vision.html)