

# PiRo - Ein autonomes Kommissioniersystem für inhomogene, chaotische Lager

PiRo - An autonomous picking system for inhomogeneous, chaotic storages

*Dominik Colling\*, Jonathan Dziedzitz\*, Patric Hopfgarten\*, Kai Markert\*, Kira Neubehler\*, Florian Eberle, Maximilian Gilles, Matthias Jung, Alp Kocabas, Kai Furmans*

*Institut für Fördertechnik und Logistiksysteme (IFL)  
Karlsruher Institut für Technologie (KIT)*

**D**as Team IFL PiRo des KIT nahm 2017 an der Amazon Robotics Challenge in Nagoya, Japan teil. Bei diesem Wettbewerb müssen Kommissionieraufträge in inhomogenen und chaotischen Lagern durchgeführt werden. Diese Veröffentlichung behandelt die getroffenen Entscheidungen bei der Entwicklung des Robotersystems PiRo und dessen Steuerung. Das Abschneiden bei dem Wettbewerb und die gewonnenen Erfahrungen werden dargestellt.

*[Amazon Robotics Challenge, Griff in die Kiste, Automatisierung, Kommissionierung, Objekterkennung]*

**T**eam IFL PiRo of KIT participated in the Amazon Robotics Challenge in Nagoya, Japan. In this competition, picking orders must be carried out in inhomogeneous and chaotic storages. This paper describes the decisions taken during the development of the robot system PiRo and its control. The performance in the competition and the gained experiences are presented.

*[Amazon Robotics Challenge, Bin Picking, Automation, Order Picking, Object Detection]*

## 1 EINLEITUNG

Die Produktion von Gütern hat mittlerweile einen hohen Automatisierungsgrad erreicht. Auch in der Logistik werden immer neue Bereiche automatisiert. Weit fortgeschritten sind mittlerweile die Handhabung und der Transport von Standardladungsträgern. Behälter, Boxen und Paletten können problemlos automatisch mithilfe von Regalbediengeräten oder Shuttles entnommen und eingelagert werden und dann mittels Stetig- oder Unstetigförderern transportiert werden.

Was noch nicht funktioniert, ist die automatisierte Entnahme und Einlagerung von Einzelprodukten mit unter-

schiedlichen Formen und Produkteigenschaften. Beispielsweise wird bei der Ware-zu-Mann-Kommissionierung die Box oder sogar das ganze Regal autonom zum Kommissionierer gebracht, die Entnahme findet aber weiterhin manuell statt.



Abbildung 1: PiRo bei der Amazon Robotics Challenge in Nagoya, Japan

Durch Weiterentwicklungen in der bildgebenden Sensorik (2D und 3D) und Steigerungen der Rechenleistung scheint auch dieser Prozess in den nächsten Jahren automatisierbar. Um diese Entwicklung weiter zu beschleunigen, gibt es seit 2015 die Amazon Picking Challenge, die mittlerweile Amazon Robotics Challenge (ARC) heißt und jährlich stattfindet. Durchgeführt wird dieser Wettbewerb von der Firma Amazon Robotics.

Das Institut für Fördertechnik und Logistiksysteme (IFL) des Karlsruher Instituts für Technologie (KIT) hat am Wettbewerb, der vom 27.07. bis zum 31.07.2017 in Nagoya, Japan, stattfand, das erste Mal teilgenommen. Für den Wettbewerb wurde ein eigenes Robotersystem entwickelt.

\* Erstautoren/lead authors

Die bei der Konzeption des Systems getroffenen Designentscheidungen und deren Konsequenzen werden im Folgenden erläutert. Die Regeln, die für den Wettbewerb gelten und die Designentscheidungen maßgeblich beeinflussen, werden in Kapitel 2 vorgestellt. In Kapitel 3 wird die verwendete Hardware erläutert, in Kapitel 4 die entwickelte Software. Die aus der Roboterentwicklung und der Teilnahme an der ARC gewonnenen Erkenntnisse werden in Kapitel 5 dargestellt.

## 2 REGELN DER AMAZON ROBOTICS CHALLENGE 2017

Das hier beschriebene Regelwerk für den Wettbewerb hatte maßgeblichen Einfluss auf die Entwicklung des Designs und die Steuerungslogik des Robotersystems. In den Regeln [1] werden der Bewerbungsprozess, der Ablauf der einzelnen Wettbewerbsdurchläufe und ihre Wertung, als auch die technischen Randbedingungen für Roboter und Lagersystem beschrieben.

### 2.1 WETTBEWERB

In einer Bewerbungsphase werden von Amazon Robotics 16 internationale Teams ausgewählt und zur Teilnahme an der Amazon Robotics Challenge eingeladen. Die selbst gebauten Systeme müssen jeweils zwei Aufgaben lösen. Es gibt einen Wettbewerb für das Einlagern von Artikeln in ein Lagersystem, den sogenannten Stow Task, und einen Wettbewerb für Auslagern und Kommissionieren, den sogenannten Pick Task.

Die besten Teams aus diesen Wettbewerben dürfen sich am dritten Tag noch in einem sogenannten Final Task messen, bei dem die Artikel in ein teilweise gefülltes Lagersystem eingelagert werden. Anschließend müssen daraus Artikel für drei Kommissionieraufträge entnommen und in Kartons abgelegt werden.

Die von Amazon Robotics zur Verfügung gestellten Artikel [2] unterscheiden sich in Größe, Form, Gewicht und Verformbarkeit. Die Artikelabmessungen und das Artikelgewicht (max. 2 kg) sind limitiert. Einige Wochen vor dem Wettbewerb bekommt jedes Team ein Set von 40 Artikeln, das sogenannte Training Set. In den einzelnen Wettbewerben sind die Artikel hälftig aus diesen „alten“ Artikeln und aus „neuen“ Artikeln zusammengesetzt. Die neuen Artikel samt Artikelbilder und -daten bekommt jedes Team 45 Minuten vor dem jeweiligen Wettbewerbsdurchlauf für 30 Minuten zur Verfügung gestellt. Diese Zeit muss dem Team reichen, um die neuen Artikel in das System einzulernen.

#### Stow Task

Beim Stow Task erhält jedes Team eine Kiste (Tote) mit 20 Artikeln. Diese sind ungeordnet und liegen teilweise übereinander. Der Roboter hat 15 Minuten Zeit diese Artikel in das eigene Lagersystem einzuräumen. Dabei müssen

die Artikel vom Endeffektor gehandhabt werden – die Kiste darf nicht einfach über dem Lagersystem ausgekippt werden. Am Ende muss eine Datei mit der jeweiligen Artikelposition ausgegeben werden. Nicht erkannte Artikel dürfen ohne Dokumentation in die sogenannte Amnesty Tote gelegt werden, ohne dass dafür Strafpunkte vergeben werden.

Punkte gibt es für jeden ins Lagersystem umgelagerten Artikel, für neue Artikel die doppelte Punktzahl. Abzüge gibt es für falsche Dokumentation, kleinere oder größere Beschädigungen der Artikel, das Fallenlassen von Artikeln aus mehr als 15 cm Höhe und für Artikel die mehr als 2 cm über das Lagersystem herausragen. Wurde die Aufgabe vollständig gelöst, gibt es Zusatzpunkte für die verbleibende Zeit.

#### Pick Task

Beim Pick Task müssen vom Team insgesamt 32 Artikel per Hand in das Lagersystem eingelagert werden. Dabei bekommt das Team die Artikel einzeln angereicht und muss jeden Artikel sofort einräumen. Danach dürfen die Artikel vom Team nicht mehr umgeräumt werden, sehr wohl aber von den Juroren. Konnten einzelne Artikel aus Platzgründen nicht ins Lagersystem geräumt werden, führt das zu Punktabzug. Das Robotersystem muss dann innerhalb von 15 Minuten drei Aufträge mit insgesamt zehn Artikeln kommissionieren. Die Aufträge bestehen aus fünf, drei und zwei Artikeln, die in je eine von fünf verschiedenen Verpackungskartons eingelagert werden müssen. Eine Dokumentation ist beim Pick Task nicht nötig. Die Punktevergabe ist ähnlich der des Stow Tasks. Zusatzpunkte gibt es, wenn sich vollständig eingeräumte Kartons ohne Umräumen der Artikel schließen lassen. Beschädigungen an den Kartons führen zu Punktabzug.

#### Final Task

Der Final Task ist eine Kombination aus Stow und Pick Task, für den 30 Minuten zur Verfügung stehen. Es müssen 16 Artikel vom Team von Hand eingeräumt werden, danach muss der Roboter weitere 16 Artikel ins Lagersystem einlagern. Erklärt das Team den Vorgang für beendet, verbleibt die Restzeit für den nachfolgenden Pick Task. Hierbei müssen wieder zehn Artikel in drei Aufträgen kommissioniert werden.

### 2.2 TECHNISCHE RANDBEDINGUNGEN FÜR DAS LAGERSYSTEM

Das Lagersystem muss aus ein bis drei Teilen bestehen, die nicht fest mit dem Robotersystem verbunden sind und die prinzipiell von einem mobilen Amazon Robotersystem transportiert werden können (siehe Abbildung 2). Für den Wettbewerb ist es erlaubt, das Lagersystem zu fixieren. Es darf zwischen zwei und zehn Fächer enthalten. Das umschließende Volumen, die benötigte Stellfläche und

die maximale Ausdehnung sind dabei limitiert. Das Lagersystem darf keine eigene Aktorik haben. Sensoren dürfen bis zu einem Gesamtwert von 50\$ verbaut werden. Für Pick und Stow Task muss dasselbe Lagersystem verwendet werden. Schubladen und ähnliche Erweiterungen sind erlaubt, müssen jedoch am Anfang und am Ende eines Tasks geschlossen sein.

### 3 HARDWARE

Für die verwendete Hardware gibt es von Amazon Robotics nur wenige Vorgaben. Das Roboter- und Lagersystem soll auf einer Fläche von 2,5m x 2,5m untergebracht werden. Die Anforderungen an das Lagersystem wurden in Kapitel 2.2 beschrieben.



Abbildung 2: Ein mobiles Robotersystem von Amazon Robotics (ehem. Kiva Systems) beim Transport eines Regals [3]

#### 3.1 KONZEPTION DES GESAMTSYSTEMS/ROBOTERSYSTEMS

Der zur Verfügung stehende UR5 von Universal Robots ist ein Roboter mit 6 Freiheitsgraden und einer Nutzlast von 5 kg. Er besitzt einen begrenzten Aktionsradius von 85 cm. Eine Erweiterung des Roboters durch eine mobile Fahrplattform oder ein kompaktes, integriertes und flächenoptimiertes Gesamtsystem wurden als Lösung in Erwägung gezogen. Für die Fahrplattform spricht eine hohe Flexibilität, dagegen stehen langsame Fahrbewegungen sowie mögliche Unsicherheiten bezüglich der Position verbunden mit höherem Steuerungsaufwand. Beim integrierten System sind die Koordinatensysteme von Roboter und Lagersystem fest verbunden. Daher muss das System nur einmal kalibriert werden und kann schneller mit weniger Unsicherheiten betrieben werden.

##### 3.1.1 LAGERSYSTEM

Dass für die ARC 2017 nun auch das Lagersystem selbst zu entwerfen war, bot neue Freiheiten. Naheliegender ist ein System, ähnlich einem üblichen Fachbodenregallager, mit Zugriff auf die Artikel von vorne. Eine weitere Möglichkeit ist der Einsatz von Behältern, die von oben bedient werden. Da das Lagersystem bezüglich Fläche und Volumen begrenzt ist, wurde der Flächen- und Raumnutzungsgrad in die Entscheidung mit einbezogen.

Ein horizontales, also von oben zu bedienendes Lagersystem hat viele Vorteile: Ablegen ist einfacher als Abstellen. Ein Artikel kann in der gleichen Lage abgelegt werden, in der er auch aus einer Kiste aufgenommen wurde. Das Ermitteln oder Schätzen von Schwerpunkten und damit korrespondierenden Flächen, die einen sicheren Stand ermöglichen, entfällt. Die Gefahr, dass ein Artikel während oder nach dem Manipulieren aus dem Lagersystem fällt, ist geringer. Auch das Aufnehmen von Artikeln von oben ist meist einfacher. Hinzu kommt, dass das Verfahren/Bewegen der Artikel an den gewünschten Ort sicherer ist, wenn dabei die Lage des Artikels nicht geändert werden muss. Es stellte sich heraus, dass der Aktionsradius des Roboterarms für eine integrierte, horizontale Lagerung ausreicht.

Artikel, die sich ganz oder teilweise überlappen, erschweren die Objekterkennung und das Greifen deutlich. Eine möglichst große Fläche im Lagersystem kann helfen, Überlappungen zu minimieren oder sogar zu verhindern. Die Fläche ist jedoch auf 5.000 cm<sup>2</sup> beschränkt. Daher wurde ein Teil des Lagersystems mit einer oben aufliegenden Schublade (B) versehen (siehe Abbildung 3). Zwei weitere Behälter komplettieren das System. Ein Behälter (C) ist schmal und hoch, um große aber flache Objekte stehend aufzunehmen, die sonst viel Fläche beanspruchen oder andere Artikel verdecken würden. Der vierte Behälter (D) des Lagersystems hat eine mittlere Höhe und kann große, würfelförmige Artikel aufnehmen.

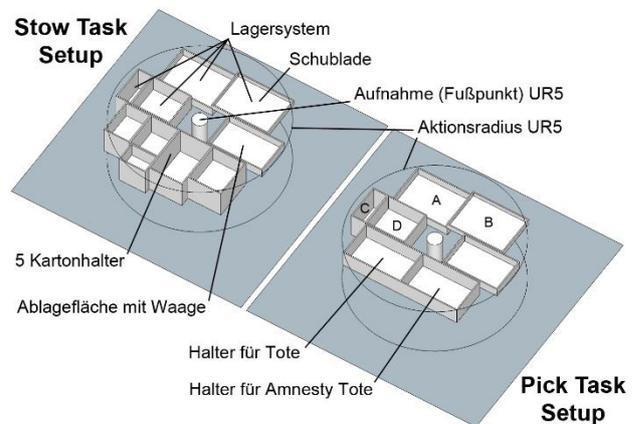


Abbildung 3: Anordnung der Behälter und Komponenten beim Stow Task und beim Pick Task

##### 3.1.2 ANORDNUNG DER KOMPONENTEN

Nach der Entscheidung für das horizontale Lagersystem und das integrierte Gesamtsystem war das Systemlayout zu entwerfen. Es wurde von einem zentral montierten Roboter ausgegangen, um den herum die verschiedenen Behälter (Lagersystem, Karton-Halter für Verpackungskartons, Totes, u.a.) platziert sind. Aufgrund der geringen Reichweite des Roboters ist dies auch die einzig mögliche Lösung. Für den Stow Task wird Platz für zwei Totes mit Artikeln (Stow Tote und Amnesty Tote) benötigt, für den

Pick Task Platz für drei von fünf möglichen Kartons verschiedener Abmessungen. Da die Totes und Kartons nicht zeitgleich benötigt werden, teilen sich die Halterungen hierfür jeweils einen gemeinsamen Bereich. Da beim Pick Task drei Kartons mit jeweils unterschiedlichen Abmessungen zum Einsatz kommen, wurde für jeden der fünf möglichen Kartons eine eigene Position im System definiert.

Der durch die Roboterreichweite begrenzte Raum in der Ebene wurde auch bei der Dimensionierung des Lagersystems und aller weiteren Behälter und Komponenten berücksichtigt. Zusätzlich zu den obligatorischen Elementen (Roboter, Lagersystem, Totes und Kartons) konnte noch eine Ablagefläche integriert werden, die mit einer Waage versehen wurde, um bisher nicht erkannte Objekte darauf abzulegen und zuzuordnen.

Das Gesamtsystem wurde überwiegend mit Alu-Systemprofilen aufgebaut. Schon bei der Konstruktion wurde auf die Möglichkeit eines einfachen, partiellen Zerlegens für den Versand nach Japan geachtet.

### 3.1.3 KAMERA

Das System ist mit einem einzigen bildgebenden Element, einer DepthSense DS 325/525 ausgestattet. Diese besitzt einen 2D-RGB-Sensor und ein 3D-Modul, welches nach dem Time-of-Flight-Prinzip (ToF) arbeitet.

Zwischen dem letzten Robotergelenk und dem Greifer ist eine Plattform montiert, an der die Kamera befestigt ist. Diese Plattform und ein Schutzbügel schützen die Kamera vor Beschädigungen, vor allem durch Kabel und Saugschlauch bei Drehungen des Endeffektors.

### 3.1.4 GREIFER

Unterhalb der Kamerahalterung ist ein elektrisch angetriebener Greifer montiert (Abbildung 4). Der 2-Finger-Greifer von Robotiq kann Gegenstände bis 140 mm greifen. Er wird als Parallelgreifer ohne Nutzung der adaptiven Funktion betrieben. Die Finger wurden im Auslieferungszustand belassen und sind innen mit Moosgummi belegt.



Abbildung 4: Greifen von verformbaren Gegenständen

Die Fingerspreizung (Position), die Greifgeschwindigkeit als auch die maximale Kraft lassen sich für jeden Greifvorgang einstellen. Die Ansteuerung erfolgt über die USB-Schnittstelle.



Abbildung 5: Saugen von Gegenständen

### 3.1.5 SAUGGREIFER

Der überwiegende Teil der Artikel kann mit einem Sauggreifer manipuliert werden. Daher ist für ein erfolgreiches Abschneiden bei der ARC die Möglichkeit zum Saugen wichtig. Allerdings können einige Artikel nicht oder nur schwer gesaugt werden, wie bspw. ein Becher mit Lochgitter.

Sollen mechanischer Greifer und Sauggreifer gemeinsam betrieben werden, so ist die Anordnung der beiden Greifer nicht trivial. Nach Möglichkeit soll die Greifachse identisch sein. Daher wurde ein System entwickelt, bei dem der Sauggreifer mit dem mechanischen Greifer gegriffen wird (Abbildung 5). Somit ist die Greifachse identisch, lediglich die Länge des Manipulators wird etwas erhöht, was aber per Software leicht zu berücksichtigen ist.

Befindet sich der Sauggreifer nicht am elektrischen Greifer, so ist er am Roboter, am „Oberarm“ arretiert (Abbildung 6). Hierzu besitzt der Sauggreifer ein metallisches Element und die Halterung einen Elektromagneten. Damit die Position nach jedem Ablegen des Sauggreifers identisch ist, befinden sich am Halter Zentrierkegel. Das Holen oder Ablegen des Sauggreifers dauert unter 10 Sekunden.



Abbildung 6: Holen und Ablegen des Greifers

### 3.1.6 VAKUUMERZEUGUNG UND ÜBERWACHUNG

Zur Vakuumerzeugung für Handhabungsaufgaben im industriellen Umfeld werden Ejektoren, Vakuumpumpen

und –gebläse eingesetzt. Für eine hohe Leckage-Kompensation und um Artikel zu greifen, die nicht saugdicht sind oder nicht saugdicht angefahren werden können, sind Vakuumgebläse besonders geeignet. Das erzeugte Vakuum ist zwar nicht so hoch wie bei den Ejektoren und Pumpen, reicht aber für die zu erwartenden Artikel der Amazon Robotics Challenge aus. Die Auswahl fiel auf eine verbreitete Art von Vakuumgebläsen, einen gewöhnlichen Haushaltsstaubsauger mit ca. 2kW Leistung. Ungünstig wirkt sich bei den Gebläsen aber der hierfür nötige größere Schlauchdurchmesser aus, der die Bewegungsfreiheit des Robotersystems unter Umständen einschränken kann.

Wird ein Sauggebläse ausgeschaltet, so dauert es einige Zeit bis der Motor zum Stillstand kommt und das Vakuum abgebaut ist. Gerade leichte Gegenstände können dann noch mehrere Sekunden am Sauger hängen, da sie sich erst lösen, wenn am Sauger praktisch Umgebungsdruck herrscht. Um schneller wieder die nächste Zielposition anfahren zu können, wurde eine Entlüftung konstruiert. Mittels eines Servomotors wird eine flexible Gummimatte von einer größeren Entlüftungsöffnung abgerollt. Hierdurch erfolgt die Entlüftung des Vakuumsystems innerhalb einer Sekunde und man kann sicher gehen, dass kein Artikel mehr vom Sauger gehalten wird.

Zur Kontrolle des anliegenden Vakuums ist in den Sauggreifer ein Vakuumsensor integriert.

### 3.1.7 LASERSCANNER

Die horizontale Lagerstruktur erlaubt den Einsatz eines 2D-Laserscanners direkt am Roboterarm. Die Montage wurde am Schultergelenk realisiert, sodass der Arbeitsraum des Endeffektors stets abgedeckt ist (Abbildung 7).

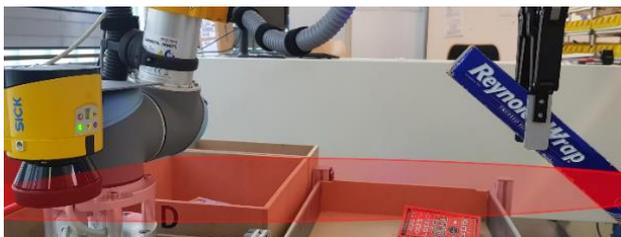


Abbildung 7: Laserfeld überwacht Greifvorgang und sichere Verfahrenhöhe

Während des Betriebs von PiRo übernimmt der Scanner vier verschiedene Aufgaben:

- Erfolgskontrolle beim Greif- oder Saugvorgang
- Höhenkorrektur vor dem Transport
- Falldetektion während des Transports
- Erfolgskontrolle beim Ablegevorgang

Für die Erfolgskontrolle wird beim Durchfahren des Laserfeldes aus den Rohdaten ein 3D-Modell des Artikels erstellt. Hieraus kann mit Hilfe von Oriented-Bounding-Boxes (OBB) [4] sowie des Iterative-Closest-Point (ICP) Algorithmus [5] eine Aussage über den Greif- oder Saugerfolg getroffen werden. Die Höhenkorrektur sowie Falldetektion wird durch eine Überwachung des Laserfeldes realisiert.

### 3.1.8 HAUPTRECHNER

Für die gesamte Objekterkennung inkl. Bildverarbeitung, Ablaufsteuerung, das Auslesen der Sensorik und die Ansteuerung von Roboter und weiterer Aktorik ist ein Standard-PC ausreichend. Er ist ausgestattet mit einem Intel i7 Prozessor, einer NVIDIA GeForce 1070 GTX Grafikkarte und 16 GB Arbeitsspeicher.

## 4 SOFTWARE

Die verwendete Software-Struktur ist in unterschiedliche Module aufgeteilt. Eine Übersicht ist in Abbildung 8 dargestellt:

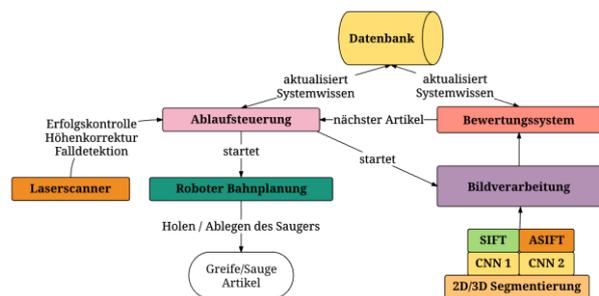


Abbildung 8: Software-Struktur PiRo

### 4.1 BILDVERARBEITUNG

Die Objekterkennung verwendet unterschiedliche Verfahren:

- (1) Scale-Invariant-Feature-Transform (SIFT) [6]
- (2) Affine SIFT (ASIFT) [7]
- (3) Conv. Neural Network (CNN): yolo v2 [8]
- (4) CNN: Inception Resnet v2 [9]
- (5) 2D/3D-Segmentierung

Gesteuert werden diese vom Software-Modul „Bildverarbeitung“ (siehe Abbildung 8).

Die beiden SIFT-Verfahren (1) und (2) können als merkmalsbasierte Erkennungen beschrieben werden und sind besonders für die Erkennung von neuen Objekten geeignet (Abbildung 9).

Zum Erlernen neuer Objekte werden mehrere Bilder (etwa 2-8) von dem zu erkennenden Objekt von unterschiedlichen Seiten und Orientierungen aufgenommen. Zusätzlich können auf den Bildern wichtige Punkte, wie beispielsweise gute Saug- oder Greifpunkte markiert werden. Die SIFT-Merkmale werden extrahiert und gespeichert. Für die Erkennung des Objektes wird ein Bild der Szene aufgenommen und die SIFT-Merkmale extrahiert. Diese werden mit den bereits vorhandenen Merkmalen abgeglichen. Existiert eine Übereinstimmung oberhalb eines definierten Grenzwertes, wird die bestmögliche Transformation in Bezug zu den bereits bekannten Bildern berechnet und eine Bounding-Box sowie die definierten Saug- und Greifpunkte ausgegeben.

Ein großer Vorteil bei der Verwendung von merkmalsbasierter Objekterkennung ist die geringe Anzahl an Bildern, die für eine korrekte Erkennung benötigt werden. Die Hinzunahme von weiteren Informationen, wie Saug- und Greifpunkte, ist möglich. Deren Positionen werden relativ zu den gefundenen Merkmalen abgelegt. Der Hauptunterschied zwischen (1) und (2) ist die Berechnung von affinen Transformationen bei (2) bezüglich der Objektorientierung in Bezug zur Kamera und ermöglicht somit eine höhere Erkennungsrate bei unbekanntem Objektorientierungen.

Die beiden Neuronalen Netze (3) und (4) bieten eine robuste Erkennung für verformbare und merkmalsarme Artikel (Abbildung 9). Für das überwachte Training wurden Bilder mit einer variierenden Anzahl unterschiedlicher Objekte und unterschiedlicher Orientierung, Belichtung und Entfernung aufgenommen. Anschließend wurden die Objekte durch manuelles Labeln markiert. Hierbei entstand ein Datensatz von über 6.000 Bildern mit ca. 150 Bildern pro Objekt. Eine bestmögliche Verallgemeinerung der Erkennung wurde mit einer Aufteilung in 70 % Trainingsbilder und 30 % Validierungsbilder erreicht. Das Training wurde lokal auf zwei GPUs für ca. 120 Stunden durchgeführt. Ein kurzfristiges Einlernen von neuen Objekten ist somit für die beiden Verfahren ungeeignet.

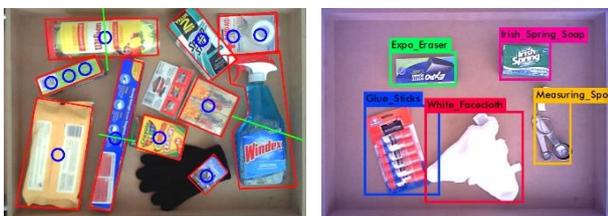


Abbildung 9: Unterschiedliche Erkennung von ASIFT (2) mit Bounding-Boxen, Saug- und Greifpunkten und yolo v2 (3) mit Bounding-Boxen

Die 2D/3D-Segmentierung (5) besteht aus zwei unterschiedlichen Segmentierungen: 3D-Segmentierung basierend auf einer Punktwolke und eine 2D/3D-Segmentierung basierend auf einem fusionierten Farb- und Tiefenbild [10]. Auf die gefundenen Segmente wird eine Hauptkomponentenanalyse angewendet und entweder eine 3D-Bounding-

Box ausgegeben (3D-Segmentierung) oder eine 2D-Bounding-Box (2D/3D-Segmentierung). Je nach Objektstruktur und Abstand liefern beide Verfahren sehr unterschiedliche Ergebnisse. Aus diesem Grund wurde für jeden Behälter des Lagersystems eine Kalibrierung durchgeführt und der Hintergrund entsprechend entfernt.

## 4.2 BAHNPLANUNG DES ROBOTERS

Für den Griff in die Kiste, das Öffnen der Schublade, das Holen und das Ablegen des Saugers mit Hilfe des Greifers werden genaue und kontrollierbare Bewegungen benötigt. Aus diesem Grund wurde eine eigene, vereinfachte Bahnplanung und Bewegungssteuerung entwickelt.

Da mit einem liegenden Lagersystem gearbeitet wird und Artikel nur von oben gegriffen bzw. abgelegt werden, müssen nicht alle möglichen Positionen vom Roboter angefahren werden. In diesem Fall wurde die Ausrichtung des Endeffektors eingeschränkt und nur Positionen mit nach unten zeigendem Endeffektor als gültige Positionen deklariert.

Als Eingabe erwartet die Bahnplanung die Position des Endeffektors mit Ausrichtung. Die Bahnplanung berechnet mithilfe eines linearen Bewegungsprofils die Strecke zwischen aktueller Position und dem Ziel in kartesischen Koordinaten. Durch die inverse Kinematik werden diese Positionen auf Gelenkwinkel projiziert. Um eine flüssige Bewegung zu gewährleisten, werden aus dieser Folge von Gelenkwinkeln, mithilfe der kubischen Spline-Interpolation, Geschwindigkeits- und Beschleunigungsprofil berechnet. Die Kombination aus Position, Geschwindigkeit und Beschleunigung wird anschließend an den Robotercontroller gesendet.

Mit dieser Bahnplanung wäre eine Bewegung auf die gegenüberliegende Seite des Roboters nicht möglich. Um dies zu ermöglichen, wird vor der Bewegung der aktuelle Fußgelenkwinkel mit dem Ziel-Fußgelenkwinkel verglichen. Wird ein Schwellwert überschritten, wird zuerst das Fußgelenk entsprechend gedreht, bevor die lineare Bahnplanung startet. Hierdurch werden auch das Überdrehen des Roboters und daraus entstehende Kabelschäden durch mechanische Belastung verhindert.

Das Öffnen und Schließen der Schublade ist eine Folge von linearen Bewegungen. Die dafür benötigten Positionen werden aus der Schubladenposition berechnet. Da der Greifer über eine bessere Kraftregelung als der UR5 verfügt, wird der letzte Zentimeter durch den Greiferhub geöffnet oder geschlossen. Dies schützt Schublade und Roboter.

Für das Holen und Ablegen des Saugers mit Hilfe des Greifers wird keine Bahnplanung in kartesischen Koordinaten, sondern eine Bahnplanung in Gelenkwinkelkoordinaten durchgeführt. Für die Kalibrierung von Aufnahme- und Ablegeposition werden einmalig vier Gelenkwinkel

bestimmt. Die Werte für Fuß- und Schultergelenk sind für die Berechnung nicht von Interesse und erlauben einen Wechsel bei beliebigen Roboterfußwinkel.

### 4.3 ABLAUFSTEUERUNG

In der Ablaufsteuerung werden die verschiedenen Softwaremodule für Hardwaresteuerung und Sensordatenverarbeitung aufgerufen. Das von Ablaufsteuerung und Bewertungssystem benötigte Systemwissen ist in einer Datenbank hinterlegt. Da die Aufträge und Lagerinhalte im Rahmen der ARC im json-Format vorliegen, wurde die Datenbank mit MongoDB umgesetzt, welches Datenbankzugriffe im json-Format ermöglicht.

Der Ablauf eines Kommissionier- und eines Einlagervorgangs sind einander sehr ähnlich, weshalb beide Fälle durch eine gemeinsame Softwareimplementierung abgedeckt werden können. Zunächst muss das nächste Ziel bestimmt und angefahren werden, anschließend wird die Bildverarbeitung gestartet. Die Ergebnisse der Bildverarbeitung werden mit Hilfe eines Bewertungssystems gewichtet und der nächste zu greifende Artikel sowie dessen Ablageort ermittelt. Daraufhin wird der Artikel mit dem geeigneten Werkzeug (Greifer oder Sauger) gegriffen, an den Zielort transportiert und abgelegt. Der Ablauf eines Pick oder Stow Tasks ist in Abbildung 10 dargestellt.

Das System kommt mit einer einzigen, am Endeffektor montierten, Kamera aus. Durch diese Designentscheidung ist es zum einen durch eine geeignete Bewegung des Manipulators möglich, beliebige Orte des Lagersystems optisch zu erfassen, zum anderen ist eine Identifizierung des Behälterinhalts auch nur möglich, wenn der Endeffektor eine bestimmte Position und Ausrichtung einnimmt.

Eine Reduktion der Rechenzeit bei der Bildverarbeitung kann durch die Verwendung von Weltwissen und eine geeignete Kombination der Algorithmen erzielt werden. Das Wissen darüber, welche Artikel sich im Lagersystem befinden, dient in zweierlei Hinsicht dazu, die Performance der Bildverarbeitung zu erhöhen: Die durch den notwendigen Abgleich unterschiedlicher Ansichten jedes Artikels zeitaufwändigen merkmalsbasierten Verfahren (SIFT und ASIFT) können deutlich beschleunigt werden, wenn lediglich nach vorhandenen bzw. im Auftrag befindlichen Artikeln gesucht wird. Die Wahrscheinlichkeit für Fehldetektionen der beiden Neuronalen Netze kann somit ebenfalls eingeschränkt werden. Die Ergebnisse der einzelnen Erkennungsalgorithmen werden kombiniert, wobei Artikel, die von mehreren Komponenten erkannt wurden, besser bewertet werden.

Die Greif- bzw. Saugpunkte für einen Artikel werden aus den Ergebnissen der Bildverarbeitung bestimmt. Bei den SIFT-basierten Verfahren werden festgelegte Greifpunkte mit ausgegeben (siehe Kapitel 4.1). Bei den Neuronalen Netzen werden die Saug- und Greifpunkte anhand der Bounding-Boxen berechnet.

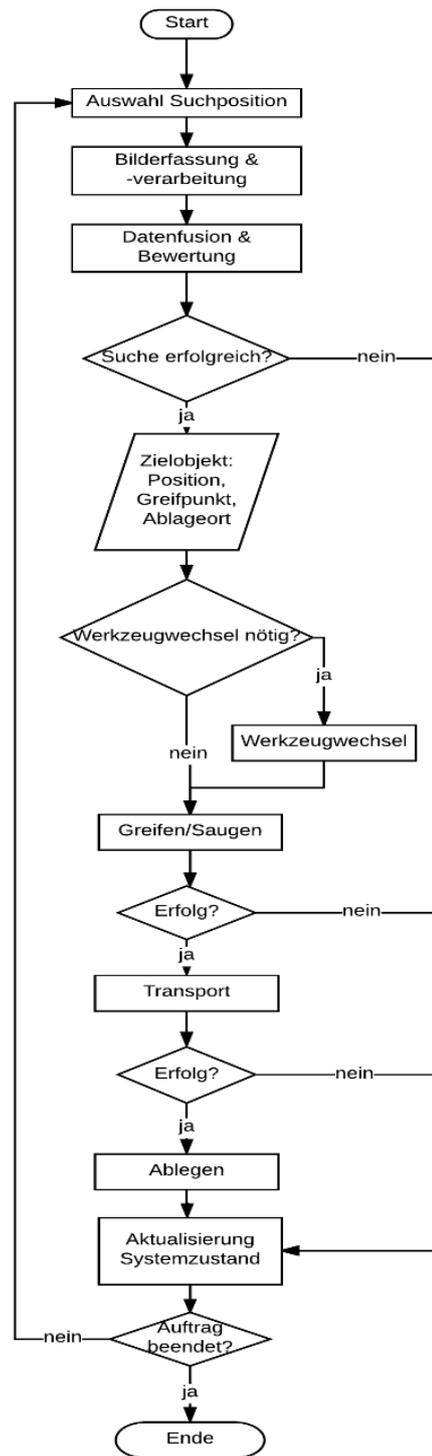


Abbildung 10: Ablauf von Pick oder Stow Task

Bei der Bewertung und Auswahl der Greifpunkte und des Werkzeugs werden der maximale Fingerabstand des Greifers, die Entfernung der Greifpunkte zur Behälterwand sowie im Vorfeld definiertes Wissen über Objekteigenschaften berücksichtigt. So können beispielsweise ver-

formbare Artikel mit einer engeren Fingerstellung gegriffen werden, während für nicht verformbare Artikel der Abstand der Finger den Ausmaßen der Bounding-Box angepasst wird.

Die Ablaufsteuerung muss in der Lage sein, mit verschiedenen Arten von Unsicherheit umzugehen. Dazu zählen zum einen die Möglichkeit von Fehlerkennungen sowie die Möglichkeit, dass ein Artikel mittels der genutzten Algorithmen nicht erkannt werden kann (dies ist z.B. bei zuvor unbekanntem Artikeln ohne eindeutige Merkmale der Fall). In der Ablaufsteuerung wird daher stets das Zielobjekt mit der höchsten Erfolgswahrscheinlichkeit zuerst ausgewählt. Die Erfolgswahrscheinlichkeit wird über eine Heuristik unter Berücksichtigung der Arteikeigenschaften, der Nähe zu anderen Artikeln, der ermittelten Greif- und Saugpunkte, des Zuverlässigkeitswertes aus der Bildverarbeitung sowie der Anzahl an vorangegangenen Fehlversuchen ermittelt. Nach einem Fehlversuch für einen Artikel wird die Greiferausrichtung bzw. Sauger-Position verändert, oder – sofern die Arteikeigenschaften es zulassen – das jeweils andere Werkzeug gewählt. So kann durch Variation des Greifvorgangs ein Teil der Unzulänglichkeiten der Bildverarbeitung ausgeglichen werden. Da die Erkennung von Artikeln auf 2D-Daten stattfindet, ist die exakte Position der Artikel im Behälter nicht bekannt. Daher wird der Endeffektor in das Lagersystem abgesenkt, bis durch einen Anstieg am Kraftsensor des Roboterarms Widerstand detektiert wird. Der Erfolg des Greifvorgangs, des Transports und des Ablegevorgangs wird mit Hilfe des in Abschnitt 3.1.7 beschriebenen Laserscanners sowie durch Auslesen der Greifersensorik bzw. des Vakuumsensors beurteilt. Wenn das System einen fehlgeschlagenen Greifversuch oder den Verlust eines Artikels beim Transport detektiert hat, wird der aktuelle Transport abgebrochen. Hierdurch können unnötige Leerfahrten verhindert und der Lagerzustand des Systems konsistent gehalten werden.

Wird kein Artikel mehr erkannt, wird die Bildverarbeitung zunächst an zwei alternativen Blickpositionen durchgeführt. Sollte auch hier kein gesuchter Artikel mehr gefunden werden, kommt die Segmentierung (siehe Kapitel 4.1) zum Einsatz. Die erhaltenen Segmente werden mit den Ergebnissen der Objekterkennungsalgorithmen abgeglichen, um Segmente bekannter Objekte auszuschließen. Ein unbekanntes Segment wird anschließend ausgewählt und auf die Ablage gelegt. Dort wird das Objekt optisch vermessen und gewogen und so einem der möglichen, noch nicht erkannten Artikel zugewiesen. Nach der Zuweisung kann der Zielpunkt für den Artikel bestimmt und der Transport des Artikels abgeschlossen werden.

## 5 ERFahrungen und Ergebnisse

Das Team IFL PiRo konnte mit dem hier vorgestellten System ins Finale der Amazon Robotics Challenge einziehen und erzielte in der Gesamtwertung den 7. Platz von ins-

gesamt 16 teilnehmenden Teams. Bei der Challenge aufgefallene Stärken und Schwächen des Systems werden im Folgenden dargestellt:

**Team:** Das Team ist interdisziplinär aus den Fachrichtungen Informatik, Elektrotechnik, Mechatronik, Maschinenbau und Wirtschaftsingenieurwesen zusammengestellt. Mögliche Lösungsansätze wurden gemeinsam diskutiert und erarbeitet. Durch unterschiedliche Kernkompetenzen entstand ein produktives und effektives Arbeitsumfeld. Insgesamt arbeiteten ca. 15 Personen an PiRo, während der Challenge vor Ort wurde die kleine Teamgröße von sieben Mitgliedern als vorteilhaft empfunden.

**Roboter:** Das Lagersystem befindet sich zwar vollständig im Arbeitsraum des UR5, allerdings kommt es bei der Anfahrt von Positionen am Randbereich teilweise zu Extremstellungen des Roboters. Zusätzlich ist das Anfahren in eine sichere Verfahrenhöhe bei großen Objekten, die sich in Ecken befinden, ohne weiteres nicht möglich. Das gegriffene oder gesaugte Objekt musste erst Richtung Roboterfuß bewegt werden und anschließend konnte die Höhenkorrektur durchgeführt werden. Diese Problematik könnte durch einen Roboter mit größerem Aktionsradius vermieden werden.

**Hardware:** Das entwickelte System zeichnet sich durch einen klaren und kompakten Aufbau aus. Kabel und Schläuche wurden durch Führungen vor Beschädigungen geschützt, ihre Bewegungsmöglichkeiten im Voraus definiert. Hierdurch wurde das Risiko, durch sich lösende Kabelverbindungen Funktionalität während der Tasks zu verlieren, auf ein Minimum reduziert. Bei der Konstruktion der Hardware wurden auf hohe Zuverlässigkeit geachtet. Ein ordentlich aufgebautes System führt nicht zwangsläufig zu guten Ergebnissen, es reduziert jedoch die Wahrscheinlichkeit zu scheitern.

**Robustheit:** Da es sich um die erste Teilnahme des Teams an einer Challenge handelte, konnte nicht auf ein bestehendes System aufgebaut werden, sondern die Hardware und Software musste vollständig neu entwickelt werden. Dies führte dazu, dass Funktionalitäten oder Abläufe (wie bspw. der Final Task) vor Ort in Nagoya angepasst werden mussten und die Zeit zum ausführlichen Testen fehlte. Während der Challenge sind neue und bisher unvorhergesehene Kombinationen aufgetreten, sodass der Stow Task auf Grund eines Treiberproblems leider abgebrochen werden musste. Ausführliche Komponenten- und Systemtests wären nötig gewesen, um solche Probleme im Vorhinein aufzudecken. Das System darf sich weder durch Schwachstellen in externen Softwarepaketen noch durch Sicherheitsfunktionen des Roboters zum Stoppen oder im schlimmsten Fall zum Absturz bringen lassen. Der Einsatz von möglichst wenig Sensorik und Aktorik vereinfachte bei PiRo nicht nur die Ansteuerung, sondern erhöhte auch die Ausfallsicherheit. Beim Aufbau der Ablaufsteuerung wurde zudem das Prinzip der Modularität verfolgt, sodass

Ausfälle einzelner Komponenten nicht zum Systemabsturz führen. So ist das System beispielsweise auch bei Ausfall einzelner Objekterkennungskomponenten oder einer der beiden Greifmethoden (Vakuumsystem bzw. Fingergreifer) noch einsatzfähig.

**Lagersystem:** Die Verwendung des horizontalen Lagersystems in Verbindung mit einer Schublade erlaubte es, beim Pick Task alle Objekte nebeneinander mit geringer Überlappung zu platzieren. Hierdurch konnte die Kamera von einer Position oberhalb eines Behälters alle darin befindlichen Artikel erkennen. Das Greifen von oben erleichtert die Berechnung von Saug- und Greifpunkten sowie die kollisionsfreie Anfahrt des Endeffektors. Da die Schublade durch den Roboter zu Beginn des Tasks geöffnet und am Ende geschlossen werden muss, ist ein vom Team definiertes, vorzeitiges Beenden eines Tasks ohne Punktabzug kaum möglich. Dies kann bei fehlerhaften Systemverhalten ein Nachteil darstellen.

**Greifen und Saugen:** Während viele Artikel ausschließlich durch Saugen umgelagert werden können, entstehen bei schweren, verformbaren oder nicht saugdichten Artikeln Probleme. Der schnelle Wechsel von Greifer zu Sauger und umgekehrt erlaubt eine artikelspezifische Handhabung und somit größte Erfolgchancen beim Umlagern.

**Bildverarbeitung:** Redundanz in der Bildverarbeitung durch unterschiedliche Bilderkennungsalgorithmen und deren Kombination führt zu einer höheren Erkennungs- und Ausfallsicherheit. Die Fehlklassifikation von Neuronalen Netzen bei für sie unbekanntem Artikeln ist hoch. Hohe Konfidenzwerte sind hier kein haltbares Gütekriterium. Aus diesem Grund wurde eine Kombination aus SIFT, ASIFT und zwei unterschiedlichen Neuronalen Netzen verwendet und führte in vielen Situationen zu einer robusten Erkennung. Durch den modularen Softwareaufbau ist das System leicht um weitere Klassifizierungsverfahren, erweiterbar. Denkbar wäre z.B. die Ergänzung einer Farbklassifizierung und/oder einer 3D-Klassifizierung.

**Kamera:** Die Objekterkennung von PiRo basiert ausschließlich auf der Auswertung von farbigen 2D-Bilddaten. Die verwendete DepthSense DS 325/525 liefert sowohl 2D- als auch 3D-Daten. Während das Tiefenbild wichtige Informationen über den Füllstand der Fächer liefern kann, wäre eine bessere 2D-Hardware, wie beispielsweise eine HDR-Kamera, bei schwierigen Lichtverhältnissen von Vorteil. Die Verwendung eines Polfilters könnte Fehlklassifizierungen durch auftretende Reflexionen vorbeugen. Durch Einsatz zusätzlicher stationärer Kameras, die unabhängig von der aktuellen Roboterposition das Lagersystem überwachen, könnte eine höhere Systemgeschwindigkeit erreicht werden.

## 6 FAZIT

Trotz der beschriebenen Schwierigkeiten während der Challenge konnte sich das Team IFL PiRo im internationalen Vergleich behaupten. Die maßgeblichen Erfolgsfaktoren waren die hervorragende Leistung des gesamten Teams, seine interdisziplinäre Zusammensetzung und die gute Unterstützung unseres Instituts.

## LITERATUR

- [1] Amazon Robotics. 2017. "2017 Amazon Robotics Challenge Official Rules." <http://www.amazonrobotics.com/site/binaries/content/assets/amazonrobotics/arc/2017-amazon-robotics-challenge-rules-v3.pdf> (aufgerufen am 01.08.2017)
- [2] Amazon Robotics. 2017. "Product Grid." [https://www.amazonrobotics.com/site/binaries/content/assets/amazonrobotics/arc/arc-2017\\_product-grid\\_2a.pdf](https://www.amazonrobotics.com/site/binaries/content/assets/amazonrobotics/arc/arc-2017_product-grid_2a.pdf) (aufgerufen am 01.08.2017)
- [3] Clements, A. 2016. "Guest comment: How UK distribution centres can replicate Amazon's robotics success." <http://edelivery.net/2016/09/guest-comment-uk-distribution-centres-can-replicate-amazons-robotics-success/> (aufgerufen am 01.08.2017)
- [4] O'Rourke, J. 1985. "Finding minimal enclosing boxes." *International Journal of Parallel Programming* 14.3: 183-199.
- [5] Besl, P., McKay, N. 1992. "A Method for Registration of 3-D Shams." *IEEE Trans. Parrem Analysis and Machine Infelligene*.
- [6] Lowe, D. 1999. "Object Recognition from Local Scale-Invariant Features." *ICCV '99 Proceedings of the International Conference on Computer Vision. Band 2: 1150-1157*
- [7] Morel, J.-M. Yu, G. 2009. "ASIFT: A New Framework for Fully Affine Invariant Image Comparison." *SIAM Journal on Imaging Sciences, Band 2, Ausgabe 2: 438-469*

- [8] Redmon, J. et al. 2016. “You Only Look Once: Unified, Real-Time Object Detection.” *University of Washington, Allen Institute for AI, Facebook AI Research*. <https://arxiv.org/abs/1506.02640>
- [9] Szegedy, C. et al. 2016. “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.” *Google Inc. 1600 Amphitheatre Pkwy, Mountain View, CA*. <http://arxiv.org/abs/1602.07261>
- [10] Toscana, G. Rosa, S. 2016. “Fast Graph-Based Object Segmentation for RGB-D Images.” *Politecnico di Torino Corso D. Degli Abruzzi 24, 10100, Turin, Italy*. <http://arxiv.org/abs/1605.03746>

---

**M.Sc. Dominik Colling\***, Wissenschaftlicher Mitarbeiter in der Abteilung Steuerungstechnik

**M.Sc. Jonathan Dziedzitz\***, Wissenschaftlicher Mitarbeiter in der Abteilung Robotik und Assistenzsysteme

**M.Sc. Patric Hopfgarten\***, Wissenschaftlicher Mitarbeiter in der Abteilung Robotik und Assistenzsysteme

**Dipl.-Ing. (FH), M.Sc. Kai Markert\***, Wissenschaftlicher Mitarbeiter in der Abteilung Steuerungstechnik

**M.Sc. Kira Neubehler\***, Wissenschaftliche Mitarbeiterin in der Abteilung Steuerungstechnik

**M.Sc. Florian Eberle**, Wissenschaftliche Hilfskraft

**Maximilian Gilles**, Student

**B.Sc. Matthias Jung**, Wissenschaftliche Hilfskraft

**B.Sc. Alp Kocabas**, Wissenschaftliche Hilfskraft

**Prof. Dr.-Ing. Kai Furmans**, Institutsleiter

[Alle Autoren sind am Institut für Fördertechnik und Logistiksysteme (IFL) des Karlsruher Instituts für Technologie (KIT) tätig]

Adresse: Institut für Fördertechnik und Logistiksysteme (IFL), Karlsruher Institut für Technologie (KIT), Gotthard-Franz-Str. 8, 76131 Karlsruhe, Germany

E-Mail: [vorname.nachname]@kit.edu

---

\* Erstautoren/lead authors