

Differenzbasierte Repräsentation räumlicher Relationen zur probabilistischen Szenenerkennung mittels hierarchischen Constellation Models

Bachelorarbeit
von

Joshua Enrico Link

An der Fakultät für Informatik
Institut für Anthropomatik und Robotik
Lehrstuhl Prof. Dr.-Ing. R. Dillmann

Erstgutachter:	Prof. Dr.-Ing. R. Dillmann
Zweitgutachter:	???
Betreuender Mitarbeiter:	Dipl.-Inform. Pascal Meißner

Bearbeitungszeit: 11. Juni 2017 – 10. September 2017

Hiermit erkläre ich an Eides statt, dass ich die von mir vorgelegte Arbeit selbstständig verfasst habe, dass ich die verwendeten Quellen, Internet-Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen – die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Karlsruhe, den **(Datum)**

ToDo

Joshua Enrico Link

Inhaltsverzeichnis

Abbildungsverzeichnis	iv
1 Einführung	1
2 Motivation und Problemstellung	2
2.1 Motivation	2
2.2 Fokus der Arbeit	2
3 Grundlagen	4
4 Konzept	5
5 Implementierung	6
6 Evaluation	7
7 Zusammenfassung und Ausblick	8
Literaturverzeichnis	9

Abbildungsverzeichnis

2.1	Beispiel: Relative Position eines Objekts zu einem anderen	3
3.1	eine Grafik ohne Sinn und Verstand	4

1. Einführung

In der Robotik ist die Servicerobotik wohl der Forschungsbereich, welcher den größten Alltagsbezug für den Menschen hat, da er sich mit der Entwicklung und Weiterentwicklung von autonomen Robotern beschäftigt, welche dem Menschen im Alltag assistieren. Man findet mittlerweile Roboter im Privaten, die das Putzen, Staubsaugen oder Rasenmähen übernehmen, in der Industrie, bei Montage und Fertigung, sowie auch in der Medizin, als Pflegehilfe, Botengänger oder Assistent.

Allerdings müssen die Roboter ihre Umwelt für komplexere Aufgaben so präzise wie möglich wahrnehmen und verstehen. Sie könne Aufgaben übernehmen bei denen sie gezielt Objekte umfahren, suchen und auch aufnehmen und benutzen. Dieser Funktionsumfang kann mit dem Prinzip Programmieren durch Vormachen (PdV) ermöglicht werden, bei dem die Roboter Objekte und Tätigkeiten ihrer Umgebung kennen lernen, wieder erkennen und nachahmen können. So lässt sich die hohe Komplexität umgehen, die die manuelle Programmierung vieler Aufgaben mit sich bringen würde.

Um tatsächlich selbstständige Serviceroboter zu schaffen muss man aber noch zu einer Objekterkennung ein Kontextverständnis hinzufügen. Die Roboter müssen erkannte Objekte in einen Zusammenhang bringen, um die dadurch resultierenden Aufgaben zu verstehen. Zum Beispiel hat ein Teelöffel, welcher neben einer Tasse Tee liegt eine andere Aufgabe zu verrichten, als wenn er neben einem Becher Joghurt platziert ist. Nur am Kontext lässt sich dort entscheiden warum im einen Fall umgerührt und im anderen gelöffelt wird. Ebenso wäre ein Stück Butter verschieden zu verwenden, wenn es auf einem Frühstückstisch steht als wenn es mit anderen Zutaten neben einer Rührschüssel vorkommt.

Somit braucht man eine Szenenerkennung, welche zuverlässig die Objekte erkennen und ihren jeweiligen Kontext verstehen und einschätzen kann. Diese Erkennung ist nicht immer eindeutig, da der eben genannte Löffel ebenso zwischen einem Becher Joghurt und einer Tasse Tee liegen könnte, deshalb bietet es sich an mit Wahrscheinlichkeitsabschätzungen des vorliegenden Kontexts zu arbeiten.

2. Motivation und Problemstellung

2.1 Motivation

Wie schon in der Einführung erwähnt ist es elementar wichtig, dass man eine zuverlässige Szenenerkennung und ein gutes Kontextverständnis schafft um den Robotern die Möglichkeit zu bieten, sinnvoll mit ihrer Umwelt zu interagieren. Mit einer präzisen Wahrscheinlichkeitseinschätzung wie die momentane Umgebung beschaffen ist, lässt sich abschätzen welche Aufgaben es zu bewältigen und welche Probleme zu lösen gilt. Um dies zu gewährleisten muss man in das System möglichst viele Referenzdaten einspeisen, damit es jeden vorhandenen Kontext erkennen kann. Szenen werden zu diesem Zweck aufgebaut um der Erkennung als neue Szene vorgestellt. Jede Szene die die Szenenerkennung auf diese Weise lernt, hilft das Chaos der sie umgebenen Objekte mehr und mehr zu interpretieren und einzuordnen. Natürlich könnte man das System auch dynamisch mit jedem auftreten eines Objekts neue Kontexte beziehungsweise Szenen lernen lassen, das würde allerdings die Erkennung an Präzision einbüßen lassen.

Die Szenenerkennung nutzt also die Daten die sie zur Verfügung gestellt bekommt, bereits gelernte Szenen wiederzuerkennen. In dem bereits vorhandenen PSM(Probabilistic Scene Model)-System werden die Daten pro Szene zu einem Modell zusammengefasst, bei dem Auffällige Zusammenhänge berücksichtigt werden und scheinbar nicht miteinander in Verbindung stehende Objekte voneinander gelöst betrachtet werden. Dadurch gibt es Vorteile in der Laufzeit und möglicherweise auch eine signifikantere Erkennung, allerdings findet natürlich auch ein Informationsverlust statt, der zu Fehlern führen kann. In dieser Arbeit wird ein Ansatz getestet, der dichter an den erhaltenen Daten arbeitet.

2.2 Fokus der Arbeit

1. datengetriebene Entwicklung
2. Problem formulieren
3. Einschränkungen / Annahmen

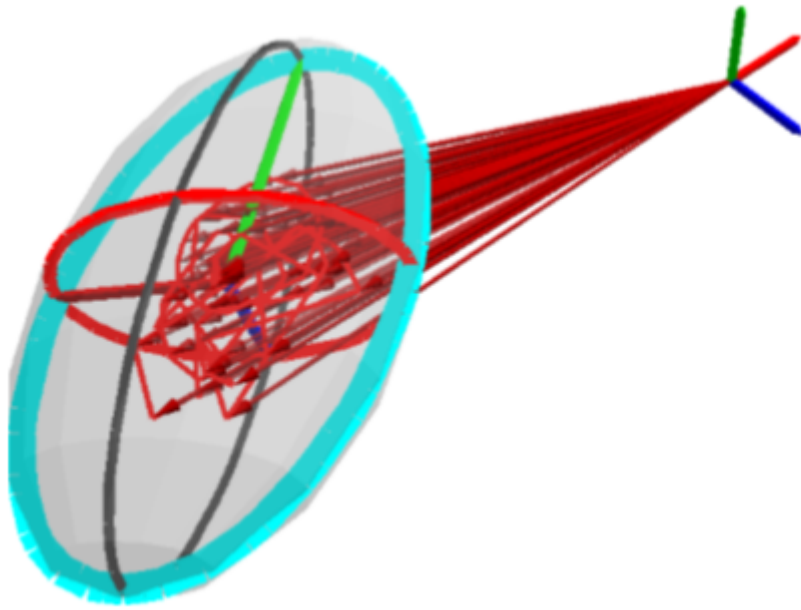


Abbildung 2.1: Beispiel: Relative Position eines Objekts zu einem anderen

4. parametrisches Modell ist ungenau - deshalb dichter an Daten

[DSS93]

3. Grundlagen

Länge max. halb so lang wie Konzept + Implementierung

simpel beschreiben

konkret:

bestehendes System : PSM

Relevanz erklären?

Datengetriebene Entwicklung erklären

Viele Bilder benutzen, auch aus Joachims Arbeit

Auch aus Joachims Arbeit

[Geh14]



Abbildung 3.1: eine Grafik ohne Sinn und Verstand

4. Konzept

komplexer mathematischer formulieren
Vergleichsbasierte Erkennung erklären
Stochastische Richtigkeit beweisen

[DSS93]

5. Implementierung

Umbau PSM
alle Klassen die umgebaut wurden
neuer differencebased modus

[DSS93]

6. Evaluation

Viele Bilder, beschreiben Daten
Text interpretiert
Fazit am Ende

[DSS93]

7. Zusammenfassung und Ausblick

Zwei Sätze zu jedem größeren Kapitel

[DSS93]

Literaturverzeichnis

- [DSS93] Randall Davis, Howard Shrobe und Peter Szolovits: *What is a Knowledge Representation?* AI Magazine, 14(1):17–33, 1993. <http://www.aaai.org/ojs/index.php/aimagazine/article/view/1029>.
- [Geh14] Joachim Gehring: *Probabilistische Szenenerkennung durch hierarchische Constellation Models über räumliche Relationen aus demonstrierten Objekttrajektorien*. Seiten 5–63, 2014. <http://www.aaai.org/ojs/index.php/aimagazine/article/view/1029>.