



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

# Ausarbeitung

Gregory Föll

Tracking bewegter Objekte in Videosequenzen

Gregory Föll  
Tracking bewegter Objekte in Videosequenzen

Ausarbeitung eingereicht im Rahmen der Veranstaltung Anwendungen 1  
im Studiengang Informatik (Master)  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr. rer. nat. Kai von Luck  
Zweitgutachter : Prof. Dr. rer. nat. Gunter Klemke

Abgegeben am 3. März 2010

# Inhaltsverzeichnis

<b>1 Einführung</b>	<b>4</b>
<b>2 Grundlagen</b>	<b>5</b>
2.1 Auswahl geeigneter Merkmale . . . . .	6
2.1.1 Histogramm . . . . .	6
2.1.2 Kanten . . . . .	7
2.1.3 Muster . . . . .	7
2.1.4 Optisches Flussfeld . . . . .	8
2.2 Objektrepräsentation . . . . .	8
2.2.1 Punkte . . . . .	9
2.2.2 Primitive Grundformen . . . . .	9
2.2.3 Konturen . . . . .	9
2.2.4 Modelle . . . . .	9
<b>3 Methoden der Objektverfolgung</b>	<b>11</b>
3.1 Merkmalsbasierte Methoden . . . . .	11
3.2 Regionenbasierte Methoden . . . . .	12
3.3 Konturbasierte Methoden . . . . .	12
3.4 Modellbasierte Methoden . . . . .	12
<b>4 Zusammenfassung und Ausblick</b>	<b>14</b>
<b>Literaturverzeichnis</b>	<b>15</b>

# 1 Einführung

Das Tracking oder das Verfolgen von bewegten Objekten durch die einzelnen Bilder einer Videosequenz, ist eine große Herausforderung im Bereich der digitalen Bildverarbeitung. Seitdem die Rechenleistung der Computer es ermöglicht, Bildfolgen in Echtzeit zu verarbeiten und auszuwerten, werden Methoden und Systeme entwickelt, die eine automatische Objektverfolgung in Video- oder Bildsequenzen ermöglichen.

Es gibt viele Gebiete in denen solche Systeme eingesetzt werden. Die automatische Objektverfolgung wird zum Beispiel oft bei der Sicherheits- und Verkehrsüberwachung, bei der Überwachung von Produktionsprozessen und bei der Verhaltensanalyse von Konsumenten oder Tieren eingesetzt. Einen breiten Einsatz finden solche Systeme auch beim Militär und bei der Entwicklung autonomer Roboter.

Im Rahmen des Masterstudiums soll ein robustes und echtzeitfähiges Verfahren für die Verfolgung von mehreren bewegten Objekten in Videosequenzen als Unterstützung für die Steuerung eines autonomen Modellfahrzeugs, welches am Carolo-Cup teilnimmt, entwickelt werden.

Der Carolo-Cup wird jährlich von der TU Braunschweig veranstaltet. Es ist ein Wettbewerb, bei dem Studententeams verschiedener Hochschulen mit selbst entwickelten, autonomen Modellfahrzeugen im Maßstab 1:10 gegeneinander antreten. Zum jetzigen Zeitpunkt müssen dabei folgende Disziplinen bewältigt werden: paralleles Einparken, Befahren einer, durch Fahrspuren gekennzeichneten, Rundstrecke mit Kreuzungen und mit stationären und dynamischen Hindernissen. Mit Hilfe des Trackingsystems sollen Hindernisse auf der Fahrbahn erkannt werden. Die Fahrspurerkennung und die Fahrbahnverfolgung wurden ebenfalls mit Methoden der digitalen Bildverarbeitung im Rahmen anderer Arbeiten bereits realisiert [Jenning \(2008\)](#); [Nikolov \(2009\)](#).

Bei autonomen Fahrzeugen werden zur Hindernisdetektion oft auch aktive Sensoren wie zum Beispiel Ultraschall- oder Infrarotsensoren eingesetzt. Aktive Sensoren sind auch mit passiven Videosensoren kombinierbar, sie können beim Navigieren bei Dunkelheit und zur schnellen und genauen Entfernungsmessung eingesetzt werden. Das zu entwickelnde Trackingsystem soll aber nur auf Methoden der Bildverarbeitung basieren.

In dieser Ausarbeitung werden die grundlegenden Methoden der digitalen Bildverarbeitung für die Objekterkennung und die Objektverfolgung in Bildsequenzen vorgestellt.

## 2 Grundlagen

Bevor ein Objekt durch eine Bildsequenz verfolgt werden kann, muss es erstmal in einem Bild oder einem Videoframe erkannt werden, dazu muss eine Kette verschiedener Bildverarbeitungsschritte durchlaufen werden.

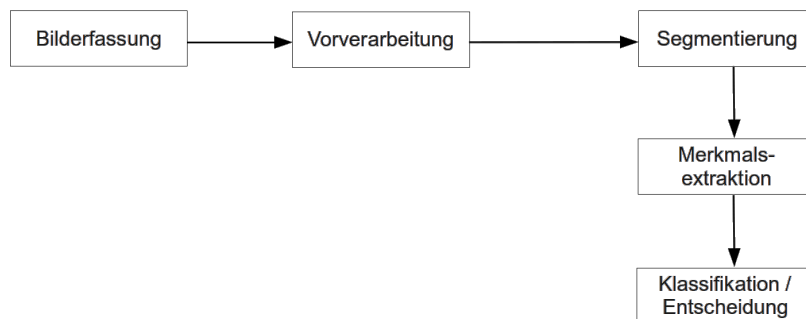


Abbildung 2.1: Kette einzelner Bildverarbeitungsschritte bei der Objekterkennung

Liegt das gewünschte Bild noch nicht in digitaler Form vor, so muss es zunächst, z.B. durch eine Videokamera, erfasst werden.

Oft besitzen die erfassten Bilder nicht die gewünschte Qualität, die Bildinformationen sind gestört, was die weitere Verarbeitung des Bildes erheblich beeinträchtigen kann. Solche Störungen sind zum Beispiel ungleichmäßige Bildausleuchtung oder Rauschen, die in einem Vorverarbeitungsschritt möglichst vollständig entfernt werden müssen. In manchen Fällen ist es sinnvoll, z.B. um die Echtzeitfähigkeit des Gesamtsystems zu garantieren, dass die Bildvorverarbeitung durch Hardwaremodule (z.B. durch FPGAs) übernommen wird.

Der nächste Schritt wird als Segmentierung bezeichnet. Dabei werden aus einem Bild alle zusammenhängenden Bereiche extrahiert.

Dann wird versucht aus jedem segmentierten Bereich bestimmte Merkmale, wie Farbe, Größe, Konturform, zu extrahieren.

Zum Schluss wird anhand extrahierter Merkmale entschieden, ob ein bestimmter Bereich das gesuchte Objekt beinhaltet.

Im Allgemeinen ist die Objekterkennung selbst nicht trivial und sollte in einer Extraausarbeitung näher untersucht werden. Im Folgenden werden die Merkmalsextraktion und die Klassifikation extrahierter Merkmale genauer beschrieben.

## 2.1 Auswahl geeigneter Merkmale

Merkmale charakterisieren unterschiedliche Eigenschaften von Objekten. Die Auswahl geeigneter Merkmale, mit deren Hilfe das gesuchte Objekt von anderen Bildinhalten unterschieden werden kann, ist eine wichtige und oft komplizierte Angelegenheit, die maßgebend für die Objekterkennung und für die spätere Objektverfolgung ist. Es ist in der Regel nicht möglich, eine allgemeine Auswahl an Merkmalen für ein beliebiges Szenario zu verwenden, die auszuwählenden Merkmale werden meistens für ein bestimmtes Anwendungsgebiet angepasst. Eine automatische Merkmalsextraktion ist schwierig, es ist ein weiteres eigenständiges Forschungsgebiet der Bildverarbeitung.

### 2.1.1 Histogramm

Bei Grauwertbildern, wird der Grauwert jedes Pixel im Bild mit einem Schwellwert (Threshold) verglichen und einer Klasse zugeordnet. Die Klassen werden gebildet, indem anhand der Extremwerte (Maxima) im Histogramm der Grauwertbereich aufgeteilt wird. Ein guter Punkt für die Bestimmung des Schwellwertes ist der zwischen zwei Extremwerten im Histogramm. Überschreitet der Grauwert eines Pixel den eingestellten Schwellwert, so wird dieses Pixel gelöscht (auf Weiß gesetzt), alle anderen Pixel werden auf Schwarz gesetzt.

Histogrammbasierte Segmentierungsverfahren werden sehr oft eingesetzt. Diese Verfahren sind einfach, schnell und ermöglichen in manchen Fällen eine vollständige Segmentierung. Leider sind diese Verfahren sehr anfällig gegen Rauschen und gegen Helligkeitsverläufe in den Bildern. In diesen Fällen kann man die Extremwerte im Histogramm nicht mehr genau unterscheiden und so keine genaue Aufteilung des Grauwertbereiches in Klassen vornehmen.

In Abbildung 2.2(c) sind ganz deutlich zwei Maxima erkennbar. Das linke Maximum zeigt die Häufigkeiten der dunkleren Bereiche und das rechte Maximum die Häufigkeiten der helleren Bereiche. Die Werte zwischen den beiden Extremwerten stellen das Rauschen dar und müssen entfernt werden.

Im Bild 2.2(d) sind dann nur noch zwei Häufigkeiten erkennbar, bei 0 (schwarz) und bei 255 (weiß).

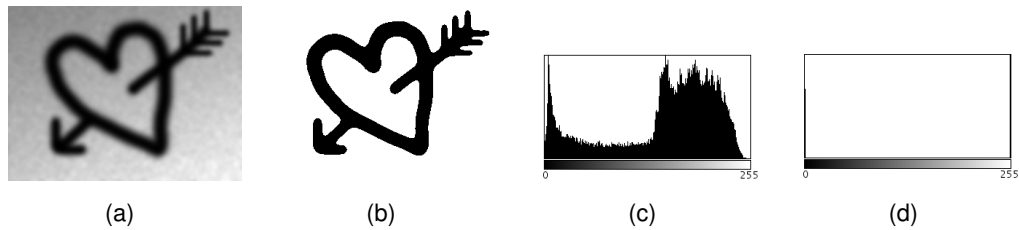


Abbildung 2.2: (a) Grauwertbild (b) Bild (a) nach Segmentierung mit Schwellwert 90 (c) Histogramm von Bild (a) (d) Histogramm von Bild (b)

Bei Farbbildern wird diese Technik für jedes der drei Merkmale (rot, grün, blau (RGB) oder Farbton, Intensität und Sättigung (HIS)), oder für die Summe der drei Merkmale angewandt.

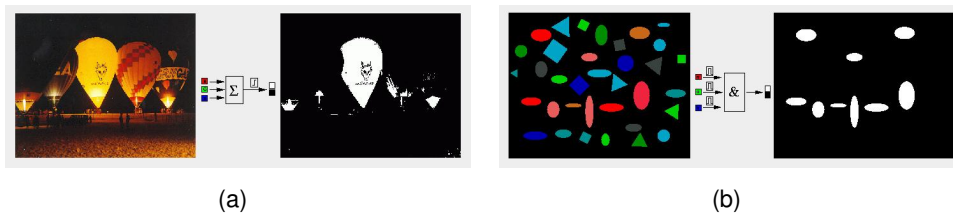


Abbildung 2.3: (a) Summierender Threshold (b) Threshold für jedes Merkmal einzeln

## 2.1.2 Kanten

Zwischen den Rändern eines physikalischen Objektes und dem Hintergrund, entstehen in Bildern normalerweise starke Helligkeitsunterschiede. Diese Unterschiede werden von verschiedenen Kantendetektionsalgorithmen, wie z.B. dem LaPlace- oder dem Sobel-Operator, ausgenutzt, um die Konturen eines Objektes zu bestimmen. Kanten haben den Vorteil, dass sie nicht so extrem auf Helligkeitsänderungen reagieren wie histogrammbasierte Merkmale.

Oft erhält man nach einer Kantensuche nicht geschlossenen Kantenzüge, diese müssen mit Kantenverfolgungs- und Füllalgorithmen komplettiert werden.

## 2.1.3 Muster

Muster oder Texturen sind wiederkehrende, ähnliche Strukturen oder schnelle Grauwertschwankungen im Bild. Mit Mitteln der mathematischen Statistik können bekannte lokale

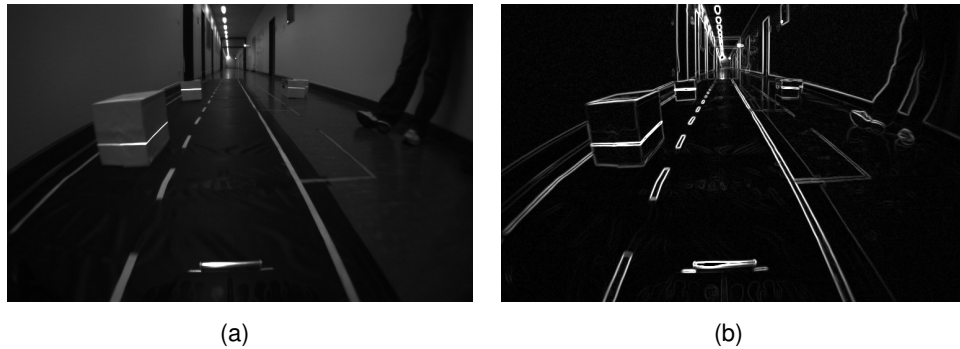


Abbildung 2.4: Kantendetektion

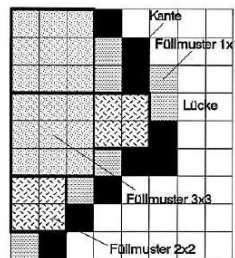


Abbildung 2.5: Beispiel für Füllmethode (Quelle: OpenCV)

Varianzen quantitativ erfasst werden. Ähnlich den Kanten, sind Muster weniger empfindlich gegen Helligkeitsänderungen als histogrammbasierte Merkmale.

### 2.1.4 Optisches Flussfeld

Optisches Flussfeld oder Optical Flow wird oft bei der Objektverfolgung eingesetzt. Es ist ein Vektorfeld, das die Bewegungsrichtung und -Geschwindigkeit für jeden Bildpunkt einer Bildsequenz angibt. Da der optische Fluss eine Größe ist, die nicht aus den Bilddaten berechnet werden kann, muss sie geschätzt werden, indem die Änderung des Helligkeitsmusters im Bild als Vektorfeld angegeben wird.

## 2.2 Objektrepräsentation

Bei der Objekterkennung ist es erforderlich zu wissen, wonach man im Bild suchen muss. Man muss sich die interessanten Bildsegmente herausgreifen, um diese dann weiter zu untersuchen. In der Regel muss für jedes Szenario oder Einsatzgebiet eine passende Reprä-



sentationsform gewählt werden. Diese Formen verwenden die im Abschnitt 2.1 vorgestellten Merkmale, wie zum Beispiel Kanten als Merkmal für eine konturbasierte Objektrepräsentation. Die Objektrepräsentation steht in einer sehr starken Beziehung zu dem verwendeten Trackingalgorithmus.

In diesem Abschnitt werden einige, für das Tracking wichtigen, Repräsentationsformen kurz beschrieben.

### 2.2.1 Punkte

Für das Erkennen sehr kleiner Objekte im Bild, ist die Punktrepräsentation gut geeignet. Die Objekte werden dabei durch einen Punkt oder durch eine Punktmenge repräsentiert. Diese Form der Repräsentation ist unempfindlich gegenüber Helligkeitsänderungen und der Kamerasicht auf das Objekt.

### 2.2.2 Primitive Grundformen

Primitive geometrische Formen sind z.B. Dreiecke, Rechtecke oder Ellipsen. Eine Objektrepräsentation mit diesen Formen ist geeignet für einfache, starre Objekte. Als Objektmerkmale kommen hier oft histogrammbasierte Merkmale wie zum Beispiel Farbe zum Einsatz.

### 2.2.3 Konturen

Bei dieser Form werden Objekte über ihre Umrisse durch umgebende Konturen repräsentiert. Konturen sind geeignet für die Beschreibung komplexer, dynamischer Objekte, wie beispielsweise Tiere oder Menschen.

### 2.2.4 Modelle

Anhand des Vorwissens über die Form und Grösse des gesuchten Objektes kann ein vereinfachtes Modell erstellt werden. Dieses Modell dient dann als Vorlage mit der Bildsegmente verglichen werden. Es werden meistens Drahtgittermodelle modelliert. Diese Modelle sind gut für die Verfolgung von starren Objekten, also Objekten, die ihre Form während der Verfolgung gar nicht oder nur geringfügig ändern, wie zum Beispiel Autos im Straßenverkehr.

Es können aber auch komplexe dynamische Modelle, die aus mehreren Teilen bestehen, erstellt werden. Die Bewegungsbeziehungen zwischen den einzelnen Modellelementen sollten

hier möglichst genau bekannt sein. Ein berühmtes Beispiel dafür sind Modelle menschlicher Körper, die aus Kopf, Rumpf, Armen, Beinen usw. bestehen. Im Idealfall kann das Modell aus primitiven geometrischen Formen zusammengesetzt werden.

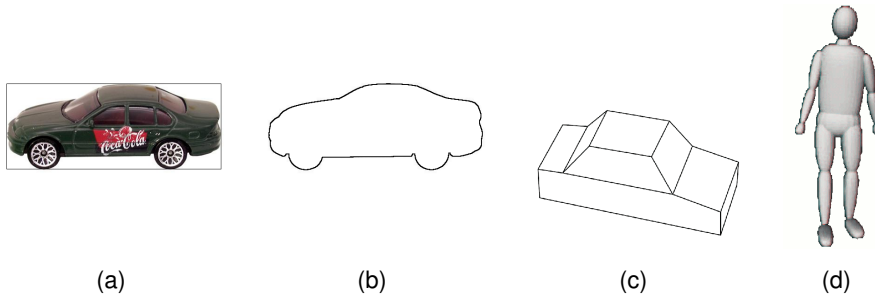


Abbildung 2.6: (a) primitive Form (Rechteck) (b) Kontur (c) Drahtgittermodell für einen PKW  
(d) Modell eines menschlichen Körpers

## 3 Methoden der Objektverfolgung

Die bekannten Methoden zur Objektverfolgung können grob in vier Klassen unterteilt werden. Jede Methode hat ihre Vor- und Nachteile und keine von ihnen ist allgemeingültig. Für ein bestimmtes Szenario müssen also nicht nur passende Objektmerkmale und die richtige Objektrepräsentation, sondern auch eine passende Verfolgungsmethode gewählt werden.

In diesem Kapitel wird die Klassifizierung der Methoden zur Objektverfolgung vorgestellt. Diese Ausarbeitung soll nur einen Überblick über das Themengebiet Objektverfolgung geben. Es soll die grundlegende Vorgehensweise bei der Entwicklung eines Trackingsystems mit Methoden der digitalen Bildverarbeitung verdeutlicht werden. In dieser Ausarbeitung werden daher keine bestimmten Algorithmen ausführlich erklärt oder hergeleitet, dies wäre ein Thema für eine weitere Ausarbeitung. Es sollen aber Anregungen für die Wahl eines passenden Verfolgungsalgorithmus gegeben werden.

### 3.1 Merkmalsbasierte Methoden

Bei merkmalsbasierten Methoden werden im ersten Schritt, durch eine geeignete Segmentierungsmethode, beliebige Bildelemente extrahiert. Im zweiten Schritt werden diese Bildelemente zu Merkmalen zusammengefasst. Merkmale können z.B. Farbe, Geraden oder Ecken sein. Durch verschiedene Ansätze wird versucht, die im aktuellen Bild gefundenen Merkmale denen im nächsten Bild zuzuordnen.

Oft ist die Bestimmung der Merkmale bei diesen Methoden sehr schwierig. Für eine genaue Zuordnung der Merkmale aus Bild 1 zu den Merkmalen aus Bild 2, wird eine sehr robuste Segmentierung der Merkmale vorausgesetzt. Des Weiteren wird eine gute Strategie für die Bild-zu-Bild-Merkmalzuordnung benötigt. Werden Merkmale im Folgebild verdeckt oder sind aus anderen Gründen, wie zum Beispiel plötzliche Änderung der Geschwindigkeit, nicht mehr zu erkennen, so kann die Rekonstruktion nicht mehr garantiert werden. Ein weiteres Problem ist eine große Anzahl ähnlicher Elemente im Bild, dies kann schnell zu Mehrdeutigkeit führen.

Die Vorteile dieser Methoden liegen in den robusten Ergebnissen auch bei stark verrauschtem Bildmaterial und in der guten Erkennung schneller Bewegungen. Bei ausreichender

Merkmalsstreuung wird Mehrdeutigkeit weitestgehend ausgeschlossen. Auch die Objektgrenzen können sehr genau bestimmt werden.

## 3.2 Regionenbasierte Methoden

Diese Methoden sind aufgrund ihrer Einfachheit und Geschwindigkeit in vielen Anwendungen weit verbreitet. Durch Auswahl geeigneter Merkmale werden zusammenhängende Bildelemente/Regionen durch primitive geometrische Formen, wie Rechtecke oder Ellipsen, dargestellt. Durch Untersuchen zeitlicher Veränderungen im Bild, wird versucht die neue Position einer bewegten Region zu berechnen.

Fast alle regionenbasierte Methoden sind auf feststehende Kameras beschränkt. Ein einfacher Algorithmus könnte also z.B. den Hintergrund, der sich nach der Zeit nur geringfügig ändert, von dem Gesamtbild abziehen und als Ergebnis bekommt man alle Bildregionen, die nicht zum Hintergrund gehören. Alle zusammenhängende Regionen können dann als bewegte Objekte interpretiert werden.

## 3.3 Konturbasierte Methoden

Sehr komplexe und dynamische Objekte können nur sehr ungenau von einfachen geometrischen Formen beschrieben werden. Bei konturbasierten Methoden werden Objekte durch ihre Umrisse repräsentiert. Bei Änderung der Konturform wird versucht, die neue Kontur mit Hilfe der Kontur aus dem vorgehenden Bild zu bestimmen. Durch die ständige Aktualisierung der gesuchten Kontur, ergeben sich von Bild zu Bild relativ kleine Formänderungen, was die Stabilität dieser Methoden unterstützt.

Diese Methoden sind relativ unempfindlich gegenüber Rauschen und können schnell gute Ergebnisse liefern. Es werden nur die gesuchten Objekte ohne Lücken segmentiert und die Extraktion ist meistens sehr detailliert.

## 3.4 Modellbasierte Methoden

Bei den modellbasierten Methoden wird im Bildmaterial nach bekannten, im Vorfeld modellierten, vereinfachten Modellen (Bsp. Abbildung 2.6(c)) des Objektes gesucht. Die Algorithmen können durch das Vorwissen über die Modelle stark vereinfacht werden. Zum Beispiel das komplexe Modell eines menschlichen Körpers (siehe Abbildung 2.6(d)), besteht aus

mehreren Teilen. Diese Einzelteile werden durch Verbindungen zusammengehalten. Für die Bewegungsbeziehungen zwischen den Teilen wird zusätzlich ein kinematisches Bewegungsmodell erstellt, so kann das Verhalten des gesamten Modells gut eingeschätzt werden.

Diese Methoden sind sowohl für die Verfolgung von einfachen, starren Objekten, als auch für die von sehr komplexen, dynamischen Objekten geeignet.

## 4 Zusammenfassung und Ausblick

In dieser Ausarbeitung wurde ein Überblick über die grundlegenden Ansätze zur Objektverfolgung in digitalen Bildern, über die wichtigsten Objektrepräsentationsformen und Objektmerkmale gegeben.

Man hat gesehen, dass die Objektdetektion und die Objektverfolgung sehr komplizierte Gebiete der digitalen Bildverarbeitung sind. Sowohl für die Objekterkennung als auch für die Objektverfolgung gibt es momentan keine allgemeinen Methoden, die für ein beliebiges Szenario eingesetzt werden könnten. Für jeden Einsatzgebiet müssen die richtigen Merkmale, eine passende Objektrepräsentation und eine passende Verfolgungsmethode mit dem richtigen Algorithmus ausgewählt werden.

Die Hinderniserkennung bei autonomen Fahrzeugen ist besonders schwierig, da hier die Echtzeitfähigkeit eine große Rolle spielt. Auf einem Modellfahrzeug kann nicht beliebig viel Hardware, für eine schnellere Verarbeitung der Bilddaten, verbaut werden. Des Weiteren kommt es oft zu starken Helligkeitsänderungen im Bild und es müssen gleichzeitig mehrere Objekte verfolgt werden. Weiterhin ist auch zu bedenken, dass nicht nur die zu verfolgenden Objekte, sondern auch die Kamera selbst ständig in Bewegung ist.

Im Rahmen des Masterstudiums soll ein robustes und echtzeitfähiges System für die Hinderniserkennung bei autonomen Modellfahrzeugen entwickelt werden.

# Literaturverzeichnis

- [Gyaourova u. a. 2003] GYAOUROVA, Aglika ; KAMATH, Chandrika ; CHEUNG, Sen ching: Block matching for object tracking / University of Nevada and Lawrence Livermore National Laboratory. URL <https://computation.llnl.gov/casc/sapphire/pubs/UCRL-TR-200271.pdf>, 2003. – Forschungsbericht
- [Handmann u. a. 1998] HANDMANN, Uwe ; KALINKE, Thomas ; TZOMAKAS, Christos: Videobasierte Hinderniserkennung / Ruhr-Universität Bochum. 1998. – Forschungsbericht
- [Homberg 2007] HOMBERG, Ulrike: *Multi-Objektverfolgung in Farbbildsequenzen auf der Basis von Partikelfiltern*, Otto-von-Guericke-Universität Magdeburg, Diplomarbeit, 2007. – URL [http://wwwisg.cs.uni-magdeburg.de/bv/theses/thesis\\_homberg.pdf](http://wwwisg.cs.uni-magdeburg.de/bv/theses/thesis_homberg.pdf)
- [Jenning 2008] JENNING, Eike: *Systemidentifikation eines autonomen Fahrzeugs mit einer robusten, kamerabasierten Fahrspurerkennung in Echtzeit*, Hochschule für Angewandte Wissenschaften Hamburg, Masterarbeit, 2008
- [Koller u. a. ] KOLLER, Dieter ; WEBER, Joseph ; MALIK, Jitendra: Robust Multiple Car Tracking with Occlusion Reasoning / University of California at Berkeley. URL <http://citeseer.ist.psu.edu/cache/papers/cs/206/http:zSzzSzwww.vision.caltech.edu/zSzzjweberzSzPaperszSzTRUCB-93-780.pdf/koller93robust.pdf>. – Forschungsbericht
- [Nikolov 2009] NIKOLOV, Ivo: *Verfahren zur Fahrbahnverfolgung eines autonomen Fahrzeugs mittels Pure Pursuit und Follow-the-carrot*, Hochschule für Angewandte Wissenschaften Hamburg, Bachelorarbeit, 2009
- [Schwarzenberg 2005] SCHWARZENBERG, Gregor: *Objektverfolgung mit Partikelfiltern*, Universität Karlsruhe (TH), Diplomarbeit, 2005. – URL <http://fsstud.uni-duisburg.de/is/publikationen/pdf/schwarzenberg05da.pdf>
- [Wang u. a. ] WANG, Yiwei ; DOHERTY, John F. ; DYCK, Robert E. V.: Moving Object Tracking in Video. . – URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.7856&rep=rep1&type=pdf>
- [Wedel 2006] WEDEL, Andreas: *Detektion stationärer Hindernisse in monokularen Bildsequenzen*, Rheinischen Friedrich-Wilhelms-Universität Bonn, Diplomarbeit, 2006

- [Yilmaz u. a. 2006] YILMAZ, Alper ; JAVED, Omar ; SHAH, Mubarak: Object tracking: A survey. In: *ACM Computing Surveys*, Vol. 38, No. 4 (2006), Nr. 13, S. 45. – URL <http://portal.acm.org/citation.cfm?doid=1177352.1177355>