

Hinderniserkennung für mobile Arbeitsmaschinen durch Fusion von Radar- und Videodaten

Obstacle detection for mobile machinery by fusion of radar and vision data

Gregory Föll
Rainer Bruns

Lehrstuhl für Maschinenelemente und Technische Logistik
Fakultät für Maschinenbau, Institut für Konstruktions- und Fertigungstechnik
Helmut-Schmidt-Universität, Universität der Bundeswehr, Hamburg

In diesem Artikel wird ein Verfahren zur automatischen Hinderniserkennung für mobile Arbeitsmaschinen vorgestellt. Das Verfahren basiert auf der Fusion von Daten eines Radarsensors und einer Monokamera. Die Hinderniserkennung erfolgt dabei, indem im Kamerabild die für das Fahrzeug befahrbare Fläche segmentiert wird. Die Segmentierung basiert auf der sogenannten Wasserscheidentransformation. Die Koordinaten der vom Radarsensor erkannten Hindernisse, werden in das Kamerabild transformiert und dienen der Wasserscheidentransformation als Hintergrundmarker.

[Schlüsselwörter: Hinderniserkennung, Datenfusion, Wasserscheidentransformation, Radar, Kamera]

This article introduces an automatic obstacle detection method based on radar and mono-camera fusion. The obstacle detection is carried out by segmenting the traversable ground plane in the camera image. The segmentation is based on the so-called watershed transformation. The coordinates of obstacles detected by the radar sensor, are transformed to image coordinates and used by the watershed transformation as background markers.

[Keywords: obstacle detection, data fusion, watershed transformation, radar, camera]

1 EINLEITUNG

Fahrerlose Transportsysteme (FTS) und Fahrerassistenzsysteme (FAS) werden im industriellen Umfeld zunehmend wichtiger. In vielen Bereichen sind diese Systeme mittlerweile nicht mehr wegzudenken. Sie tragen unter anderem zur Kostenoptimierung und Sicherheit bei. Die Einsatzgebiete von FTS und FAS sind sehr vielfältig. Zu ihren Aufgaben gehören unter anderem: Einparkhilfe, Mindestabstand halten, Spurverfolgung, Routenplanung und Navigation. Vor allem spielt die Reduzierung der

Unfallfolgen und damit die passive Sicherheit für Arbeiter eine bedeutende Rolle.

Um die oben genannten Aufgaben zu meistern und zur Steigerung der Wettbewerbsfähigkeit von Fahrerlosen Transportsystemen, müssen zukünftige Fahrerlose Transportfahrzeuge (FTF) vorausschauend und situationsabhängig auf Hindernisse in den Verkehrswegen reagieren können. Dazu bedarf es einer automatischen Hinderniserkennung, die die Objekte in der Umgebung des Fahrzeuges möglichst vollständig und genau erfasst.

Zahlreiche Untersuchungen zeigen, dass der Einsatz eines einzelnen Sensortyps für die Hinderniserkennung in vielen Fällen nicht ausreichend ist [GER2001, COU2002, DEC2009, RAP2010]. Die Anforderungen an eine Hinderniserkennung im industriellen Umfeld unterscheiden sich deutlich von denen im Straßenverkehr. Landstraßen und Autobahnen haben einen relativ monotonen Fahrbahnverlauf. Auch die gesamte Umgebung des Fahrzeuges ändert sich nicht so rasant wie in dynamischen und unstrukturierten Umgebungen. Durch die kleinere Anzahl potenzieller Hindernisse, entstehen, bei aktiven Sensoren, weniger Fehlreflektionen, die die Hinderniserkennung erschweren. Des Weiteren erleichtern Fahrbahnmarkierungen und Leitpfosten die Orientierung auf der Fahrbahn und man hat es überwiegend mit bekannten Hindernissen wie Lkw, Pkw, Motorrädern, Verkehrsschildern und Schutzplanken zu tun. In unstrukturierten industriellen Umgebungen, entstehen oft hochkomplexe Situationen, die sich mit nur einem Sensortyp nicht bewältigen lassen.

Um die Einsatzgebiete von FTS auszuweiten und die Zuverlässigkeit der Systeme zu verbessern, können die Daten mehrerer unterschiedlicher Sensortypen fusioniert werden. In dieser Arbeit werden die Daten eines Radarsensors und einer Monokamera fusioniert. Dabei wird mithilfe des entwickelten Fusionsmodells, die für das Fahrzeug befahrbare Fläche im Kamerabild segmentiert. Die Hinderniserkennung erfolgt durch die Unterteilung der Bildbereiche in Hindernisse und Fahrweg. Alles was nicht zum Fahrweg gehört wird dabei als Hindernis betrachtet. Das entwickelte Verfahren soll sowohl im Indoor- als auch im Outdoor-Bereich zuverlässige Ergebnisse liefern.

2 SENSORIK

Für die Hinderniserkennung haben sich zahlreiche Sensoren etabliert. Diese Sensoren lassen sich grob in zwei Gruppen unterteilen: passive und aktive Sensoren. Zu den aktiven Sensoren gehören zum Beispiel Kameras, die mit strukturiertem Licht arbeiten, Time of Flight Kameras, Laserscanner, Ultraschall- und Radarsensoren. Zu den passiven Sensoren gehören Stereokamera- und Monokamerasysteme. Aktive Sensoren senden selbständig bestimmte Signale aus, deren Reflektionen sie wieder einfangen und auswerten. Die passiven Kamerasensoren nutzen nur das einfallende Licht, ähnlich dem menschlichen Auge, um eine Abbildung ihrer Umgebung zu erhalten. Die Bilddaten werden anschließend von einem Rechner verarbeitet.

Jeder Sensor hat seine Vor- und Nachteile. Viele heutige Lösungen für eine Hinderniserkennung basieren auf einem bestimmten Sensortyp. In aktuellen industriellen Anwendungen kommen zum Beispiel oft Laserscanner zum Einsatz. Zum einen sind diese Systeme noch sehr teuer, zum anderen bereiten, in industriellen Anwendungen oft auftretende, Verschmutzungen der Optik und eingeschränkte Sichtverhältnisse (durch Regen, Schnee, Nebel und Sonnenlicht) Schwierigkeiten. Ähnliche Probleme haben auch aktive Kamerasensoren. Die mechanischen Wellen eines Ultraschallsensors werden ebenfalls von den Umgebungsbedingungen beeinflusst [MOE2012]. Ultraschallsensoren sind auch wegen der verhältnismäßig geringen Signalausbreitungsgeschwindigkeit und Reichweite in ihren Einsatzmöglichkeiten begrenzt.

Radar-Systeme haben sich im Einsatz von Abstandsregelautomat-Anwendungen (ACC) bewährt [TOK1996, OHS1999]. Der große Vorteil von Radar-Sensoren ist ihre Unempfindlichkeit gegenüber ungünstigen Umweltbedingungen wie starkes Sonnenlicht, Regen, Nebel, Dunkelheit und Verschmutzungen des Sensors. Des Weiteren lassen sich mit einem Radarsignal innerhalb eines Messzyklus die Entfernung, die relative Geschwindigkeit und der Lateralwinkel mehrerer Objekte bestimmen. Der Nachteil dieses Sensors ist die geringe laterale Auflösung.

Als passive Sensoren werden Kamerasysteme von Umgebungsbedingungen wie Lichtverhältnissen und schlechtem Wetter beeinflusst. Monokamerasystemen fehlen Tiefeninformationen die die Objekterkennung vereinfachen. Das Lösen des Korrespondenzproblems bei Stereokamerasystemen ist sehr rechenintensiv [SHU2013] und die resultierende Entfernungsmessung ist nicht so akkurat wie bei Radarsensoren. Dennoch ist eine Monokamera ein kompakter, kostengünstiger und leistungsfähiger Sensor. Der Sensor bietet einen großen Öffnungswinkel und eine hohe Objektauflösung, bei der für den FTS-Einsatz relevanten Entfernung. Daraus resultiert auch ein hohes laterales Auflösungsvermögen des Sensors. Durch Mustererkennung lassen sich mithilfe einer Kamera die

betrachteten Objekte klassifizieren. Zum Beispiel können Fahrbahnmarkierungen und auch schlecht reflektierende Objekte, wie Personen, die von einem Radar nicht immer erfasst werden können, erkannt werden.

Aufgrund ihrer komplementären Eigenschaften lassen sich Radar- und Kamerasensoren sinnvoll miteinander kombinieren [KAT2002, SOL2004, PRO2008, BAE2009, MIL2011].

3 WASSERSCHIEDENTRANSFORMATION

3.1 BILDSEGMENTIERUNG

Die Bildsegmentierung spielt eine sehr bedeutende Rolle für die Interpretation von Bilddaten. Als Segmentierung bezeichnet man den Prozess, bei dem das Bild in verschiedene, zusammenhängende Bildbereiche unterteilt wird. Ein Bildbereich oder eine Bildregion ist eine Menge von Pixeln mit ähnlichen Eigenschaften. Nach der Segmentierung werden die einzelnen Bildbereiche nach bestimmten Merkmalen, wie zum Beispiel Form, Grauwertverteilung, Textur, durchsucht. Anschließend können auf den relevanten Bildbereichen Messungen durchgeführt werden. Die Segmentierung eines Bildes ist eine sehr herausfordernde Aufgabe. Es existiert kein allgemeingültiges Segmentierverfahren, das sich bei jedem beliebigen Szenario einsetzen lässt. Aus diesem Grund werden in der Praxis Speziallösungen eingesetzt, die ein gewisses Vorwissen über das zu untersuchende Bildmaterial beinhalten.

In der digitalen Bildverarbeitung wird zwischen folgenden Segmentierverfahren unterschieden:

- Pixelorientiert
- Kantenbasiert
- Regionenorientiert
- Modellbasiert
- Morphologische Segmentierung (Mischverfahren)

Bei den pixelorientierten Verfahren werden die Eigenschaften jedes einzelnen Pixels im Bild untersucht. Anhand dieser Eigenschaften wird entschieden, zu welchem Bildbereich oder Segment diese Pixel gehört. Kantenbasierte Verfahren durchsuchen die Bilddaten zunächst nach Kanten. Man geht davon aus, dass die Kanten zwischen unterschiedlichen Bildregionen verlaufen. Durch Kantenverfolgungsalgorithmen können die gefundenen Kanten anschließend vervollständigt werden, sie beschreiben dann die Grenzen der einzelnen Bildbereiche. Bei regionenorientierten Verfahren werden nicht wie bei pixelorientierten Verfahren einzelne Pixel, sondern Mengen benachbarter Pixel untersucht, die dann zu zusammenhängenden Objekten vereinigt werden. Modellbasier-

te Verfahren durchsuchen das Bildmaterial nach Modellen bekannten Objekte. Die Wasserscheidentransformation gehört zur Gruppe der morphologischen Segmentierverfahren. Es wird versucht, die Vorteile unterschiedlicher Verfahren zu vereinen, ohne deren Nachteile zu übernehmen.

3.2 EINFÜHRUNG

Der Begriff „Wasserscheide“ kommt aus den Geowissenschaften. Als Wasserscheide wird die Trennungslinie der Einzugsgebiete von Bächen oder Flüssen bezeichnet. Gebirgskämme stellen natürliche Wasserscheiden dar. Trifft ein Wassertropfen auf die Spitze eines Berges, so fließt es so lange entlang des steilsten Abhangs ab, bis es ein Minimum erreicht hat. Mit der Zeit bildet sich an dieser tiefsten Stelle ein See oder ein Fluss, der von dem Gebirge umschlossen wird.

Für die Segmentierung wird das Bild zunächst in ein Graustufenbild konvertiert. Im Weiteren wird die topographische Darstellung des Bildes betrachtet. Als topographische Darstellung wird hier das Höhenprofil des Graustufenbildes bezeichnet.



Abbildung 1. Graustufenbild

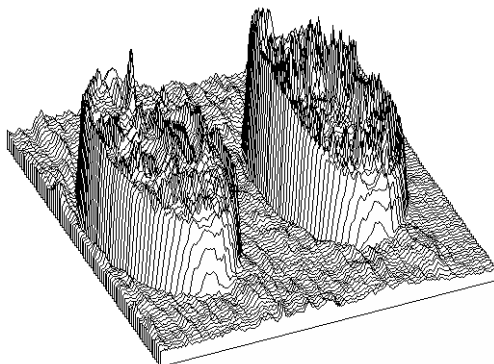


Abbildung 2. Topographische Darstellung

Abbildung 2 zeigt die topographische Darstellung des in Abbildung 1 dargestellten Graustufenbildes. Die maximale Höhe des Grauwertgebirges, in der topographischen Darstellung, hängt dabei von der Codierung der Grauwerte des Graustufenbildes ab. Bei einer Codierung mit 8 Bit ($2^8 = 256$ Werte) beträgt die minimale Höhe 0 und die maximale Höhe 255.

Im Folgenden wird die Wasserscheidentransformation anschaulich anhand der Simulation von Überflutungen

erklärt. Weitere Informationen zum Algorithmus und die genaue Formalisierung sind u.a. in [SOI1998] und [GON2008] zu finden.

3.3 SIMULATION VON ÜBERFLUTUNGEN

Unter Betrachtung des Graustufenbildes als eine topographische Oberfläche, nimmt man an, dass in die regionalen Minima dieser Oberfläche Löcher gebohrt werden. Danach wird die gesamte Oberfläche langsam in Wasser eingetaucht. Dabei wird das Gebirge, beginnend mit den niedrigsten Minima, durch die Löcher von unten geflutet. Dieser Vorgang wird in den folgenden Abbildungen dargestellt. In Abbildung 3 sind die lokalen Tiefpunkte durch kleine blaue Kreise gekennzeichnet. Die Nummern geben die Reihenfolge der Flutung an.

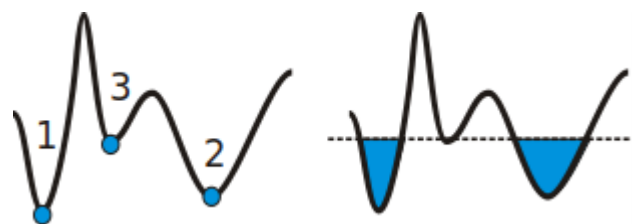


Abbildung 3. Überflutung des Grauwertgebirges von unten durch die Löcher in den lokalen Tiefpunkten

Durch das steigende Wasser bilden sich an Stellen der regionalen Minima Seen. An den Stellen, an denen sich zwei Seen, bei weiterem Wasseranstieg, miteinander vereinigen würden, werden Dämme oder Wasserscheiden errichtet. Die Dämme werden dabei immer bis zur Höhe des höchsten Berges gebaut. Nach der kompletten Überflutung sind nur noch die Wasserscheidelinien zu sehen, welche die Grenzen der segmentierten Objekte darstellen.

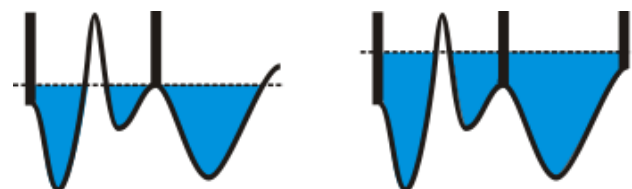


Abbildung 4. Bau von Dämmen an den Stellen, an denen sich zwei benachbarte Seen mit steigendem Wasser vereinigen würden

3.4 MARKERKONTROLLIERTE SEGMENTIERUNG

Die Wasserscheidentransformation ist ein sehr leistungsfähiges Segmentierverfahren. Bei entsprechender Qualität des Eingangsbildes lässt sich das Bild mithilfe dieses Werkzeuges in aussagekräftige Bildbereiche zerlegen. Ein großes Problem dieses Verfahrens ist die Übersegmentierung. Durch Bildrauschen oder durch lokale Unregelmäßigkeiten im Bild, kommt es in der Praxis oft zu stark übersegmentierten Ergebnissen. Abbildung 5 zeigt ein verrauschtes Graustufenbild und Abbildung 6 das Ergebnis der Segmentierung mithilfe der Wasserscheidentransformation. Die schwarzen Partikel stellen im

Bild die gesuchten Objekte dar. Im Ergebnisbild erkennt man deutlich, dass die Grenzen der gesuchten Objekte sehr ungenau gekennzeichnet werden und es existieren viele kleine Segmente, die kein gesuchtes Objekt beinhalten.

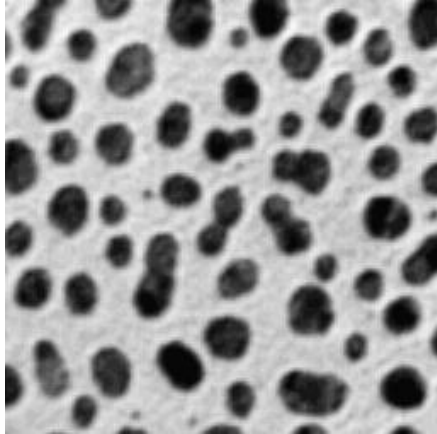


Abbildung 5. Verrauschtes Graustufenbild

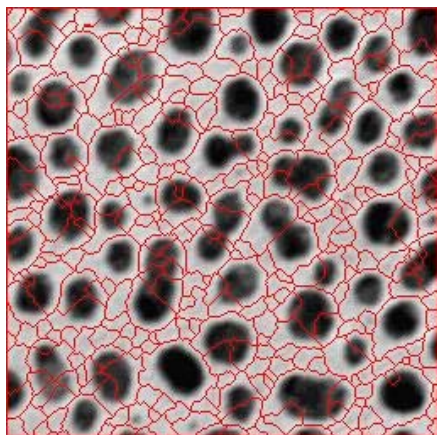


Abbildung 6. Ergebnis der Wasserscheidentransformation

Zu der Übersegmentierung kommt es durch die Existenz von vielen nicht relevanten Minima. Um ein besseres Ergebnis zu erreichen ist es also das Ziel, die irrelevanten Minima vor der Segmentierung im Ausgangsbild zu entfernen. Oft ist es jedoch schwierig oder sogar unmöglich einen passenden Filter für diese Aufgabe zu implementieren.

Die Lösung ist das manuelle oder automatische Markieren des Bildes. Man unterscheidet zwischen Vordergrund- und Hintergrundmarkern. Ein Marker kann ein einzelnes Pixel oder eine beliebig große Pixelmenge darstellen, wobei mit größeren Markern bessere Ergebnisse erzielt werden. Vordergrundmarker kennzeichnen die relevanten Objekte im Bild und Hintergrundmarker die Stellen, die mit Sicherheit keine gesuchten Objekte beinhalten. Die markerkontrollierte Wasserscheidentransformation ignoriert irrelevante Minima und generiert dadurch deutlich bessere Ergebnisse.

4 MODELLBESCHREIBUNG

Die Hauptaufgabe des Fusionsmodells in dieser Arbeit, ist die Segmentierung des Fahrweges im Kamerabild. Der Algorithmus soll die befahrbare Fläche von allen anderen Gegenständen im Bild unterscheiden. Als Segmentierverfahren wird die Wasserscheidentransformation eingesetzt. Das automatische Markieren durch Merkmalsdetektoren erfordert eine Vorverarbeitung des Bildes und viel Vorwissen über die gesuchten Objekte. Da sich im industriellen Umfeld viele Gegenstände mit sehr unterschiedlichen Eigenschaften wie Geometrie, Farbe, Form usw. befinden, ist das Markieren allein durch Merkmalsdetektoren kaum möglich. Aus diesem Grund ist die Idee, zusätzlich zum Kamerasensor, einen aktiven Sensor, in diesem Fall Radar, zum Generieren von Markern einzusetzen.

4.1 DATENFUSION

Als Datenfusion bezeichnet man die sinnvolle Kombination von Daten mehrerer unterschiedlicher Sensoren, mit dem Ziel, die Schwächen der einzelnen Sensoren zu reduzieren und/oder die Kosten des Gesamtsystems im Vergleich zu einem Einzelsensor zu senken.

Die in der Literatur beschriebenen Modelle für die Fusion von Radar- und Kameradaten unterscheiden sich grundsätzlich in den Fusionsebenen Rohdatenebene (engl.: *low level*), Merkmalsebene (engl.: *middle level*) und Entscheidungsebene (engl.: *high level*) [KLE1999]. Die Fusionsebene gibt Aussage darüber, welche Daten wie und zu welchem Zeitpunkt in der Verarbeitung fusioniert werden. Einige Verfahren vereinen beide Fusionsebenen und können als Mischverfahren eingestuft werden.

Bei dem Low-Level-Ansatz versucht man die Rohdaten beider Sensoren sinnvoll miteinander zu kombinieren, um durch Vervollständigung und Redundanz ein genaueres Ergebnis zu erzielen. Die Fusion erfolgt hier in der Regel unmittelbar nach der Gewinnung der Daten. Bei der Fusion auf Merkmalsebene werden die Rohdaten vor der Fusion zu Merkmalen zusammengefasst. In vielen Arbeiten wird die Merkmalsebene auch als Rohdatenebene bezeichnet. Die Rohdaten des Radarsensors sind die Entfernung, die relative Geschwindigkeit und der Lateralwinkel eines Objektes. Als Rohdaten des Kamerasensors können nicht nur einzelne Pixel, sondern auch primitive Bildmerkmale wie Ecken, Kanten und Texturen angesehen werden.

Beim High-Level-Ansatz werden die Daten jedes Sensors separat ausgewertet. Die ausgewerteten Daten des einen Sensors validieren dann die Ergebnisse des anderen Sensors oder man interpretiert die ausgewerteten Daten beider Sensoren als Teilergebnisse um daraus eine Entscheidung zu generieren.

Des Weiteren unterscheidet man bei Datenfusionsmodellen zwischen synchronen und nicht synchronen Systemen. Bei synchronen Systemen ist die Datengewinnung aller Sensoren zeitlich aufeinander abgestimmt. Bei nicht synchronen Systemen ist der Takt der Datengewinnung der beteiligten Sensoren individuell. Ein Vorteil synchroner Systeme ist die Kenntnis über deren zeitliches Verhalten. Auf der anderen Seite stellen diese Systeme höhere Ansprüche an die verwendete Hard- und Software und können in manchen Fällen zu längeren Verarbeitungszeiten führen.

Das vorgestellte Verfahren führt die Datenfusion auf Entscheidungsebene durch. Der Radarsensor wertet seine Rohdaten selbständig aus und generiert daraus Objektkoordinaten, die anschließend zu Hintergrundmarkern umgewandelt werden. Im Kamerabild werden Vordergrundmarker gesetzt. Die ausgewerteten Ergebnisse beider Sensoren werden dann von der Wasserscheidentransformation zur Segmentierung verwendet. Um eine höhere Verarbeitungsgeschwindigkeit zu erreichen, werden die beiden Sensoren nicht synchronisiert.

4.2 AUFBAU

Abbildung 7 zeigt den Aufbau des Systems anhand eines Schubmaststaplers.

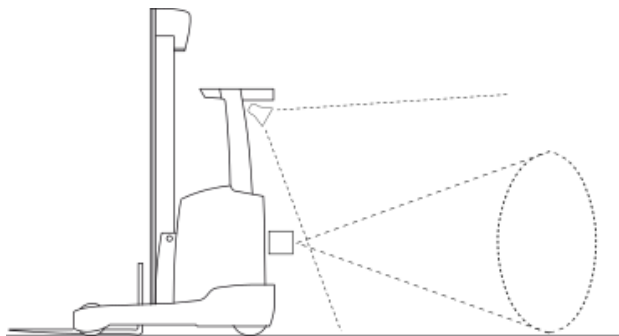


Abbildung 7. Systemaufbau

Der Radarsensor wird horizontal zum Fahrweg angebracht. Die Kamera befindet sich über dem Radar und ist unter einem bestimmten Winkel auf den Fahrweg gerichtet. Abbildung 8 zeigt die Szene von oben aus der Sicht der Kamera. Werden durch den Radarsensor keine Hindernisse unmittelbar vor dem Fahrzeug festgestellt, dann wird davon ausgegangen, dass dieser Bereich zu befahrbarer Fläche gehört. Im Kamerabild wird dann an dieser Stelle ein statischer Vordergrundmarker gesetzt (grün dargestellt). Zusätzlich zum statischen Marker wird versucht einen dynamischen Vordergrundmarker zu generieren, der sich automatisch an den Fahrweg unmittelbar vor dem Fahrzeug anpasst. Alle durch den Radarsensor ermittelten Hindernisse werden als Hintergrundmarker interpretiert. Dieser Vorgang wird kontinuierlich fortgesetzt.

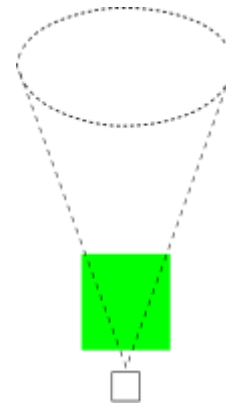


Abbildung 8. Statischer Vordergrundmarker im Kamerabild

4.3 KOORDINATENTRANSFORMATION

Das Gesamtsystem beinhaltet mehrere Koordinatensysteme. Die Transformation zwischen zwei Koordinatensystemen im Raum wird durch eine Rotationsmatrix und einen Translationsvektor beschrieben. Die Beziehung zwischen einem Raumpunkt und einer Kamera wird in Abbildung 9 verdeutlicht.

Man unterscheidet zwischen dem sogenannten Weltkoordinatensystem, dem Kamerakoordinatensystem, dem Radarkoordinatensystem, dem Fahrzeugkoordinatensystem, dem Objektkoordinatensystem, dem Polarkoordinatensystem des Radarsensors, dem Bildkoordinatensystem und Pixelkoordinatensystem.

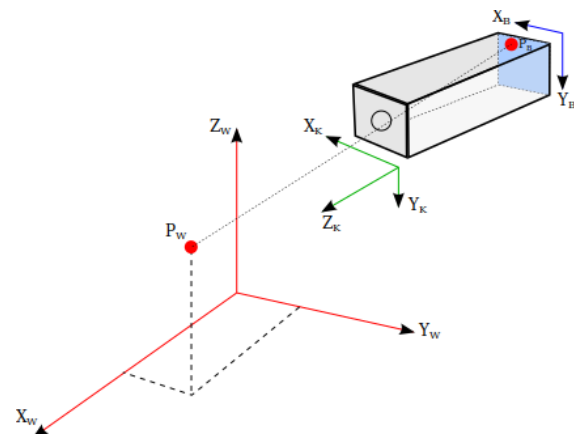


Abbildung 9. Koordinatensysteme: Weltkoordinatensystem (rot), Kamerakoordinatensystem (grün), Bildkoordinatensystem (blau)

Das 3D-Weltkoordinatensystem ist das Referenzsystem auf das man sich bezieht, wenn Objektpositionen im Raum beschrieben werden. Es bezieht sich auf die betrachtete Szene, also auf im Raum liegende Objekte. Die Einheiten dieses Systems werden in m angegeben.

Das Kamerakoordinatensystem ist mit der aufnehmenden Kamera verknüpft. Ihre z -Achse entspricht dabei der optischen Achse des Kamerasystems. Ihr Ursprung liegt im Projektionszentrum der Kamera.

Entsprechend definiert das Radarkoordinatensystem die Position des Radarsensors im Raum (bezogen auf das Weltkoordinatensystem). Dasselbe gilt für das Objektkoordinatensystem und das Fahrzeugkoordinatensystem. Der Ursprung des Fahrzeugkoordinatensystems liegt üblicherweise in der Mitte der Vorderachse des Fahrzeugs.

Das Bildkoordinatensystem ist zweidimensional. Seine x - und y -Achse verlaufen parallel zu den entsprechenden Achsen des Kamerakoordinatensystems. Das Bildkoordinatensystem beschreibt die Position eines abgebildeten Raumpunktes im Bild. Sein Ursprung liegt in der Mitte des Bildes, die Einheiten werden in m oder in μm angegeben.

Das für die Abbildungsgeometrie eigentlich irrelevante System, ist das Pixelkoordinatensystem. Es ist das Koordinatensystem eines digitalen Bildes, bei dem die Zeilen- und Spaltennummer eines Pixels als dessen Koordinaten verwendet werden. Der Ursprung dieses Koordinatensystems liegt in der linken oberen Ecke des Bildes. Die Einheit des Systems ist Pixel.

Das Radar gibt die Position von erkannten Hindernissen in Polarkoordinaten wieder. Diese müssen dann in Raumkoordinaten umgerechnet werden. Zuerst werden die Polarkoordinaten folgendermaßen in kartesische Koordinaten umgerechnet:

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

Ein großes Problem ist die Herstellung der Transformation oder der Beziehung zwischen den beiden Sensoren. In vielen Arbeiten werden der Translationsvektor und die Neigungswinkel der Sensoren per Hand ausgemessen, die Rotation der Sensoren im Raum wird in der Regel vernachlässigt. Diese ungenaue Methode kann zu falschen Ergebnissen und dadurch zu einem höheren Verarbeitungsaufwand führen. In der Literatur werden zwar einige weitere Methoden beschrieben, doch diese sind aufwendig und die Ergebnisse der Kalibrierung sind dabei auch nicht sehr genau. Aus diesem Grund wurde für dieses Zuordnungsproblem eine einfache und hochgenaue Methode entwickelt. Dabei wird in Höhe des Radarsensors, parallel zur Fahrbahn, unmittelbar vor dem Radarsensor, ein bestimmtes Muster platziert. Das Muster wird so angebracht, dass es vollständig von der Kamera erfasst wird. Dabei wird die minimale Rotation des Sensors (lässt sich in der Regel durch Montageungenauigkeiten nicht vermeiden oder ist gewollt) berücksichtigt. Die linke obere Ecke des Musters ist nur auf einer Achse (y -Achse) in der Ebene vom Mittelpunkt des Radarsensors verschoben. Diese Verschiebung lässt sich leicht und sehr genau mit einem Lineal ausmessen. Der Aufbau dieser Kalibrierung wird in Abbildung 10 dargestellt.

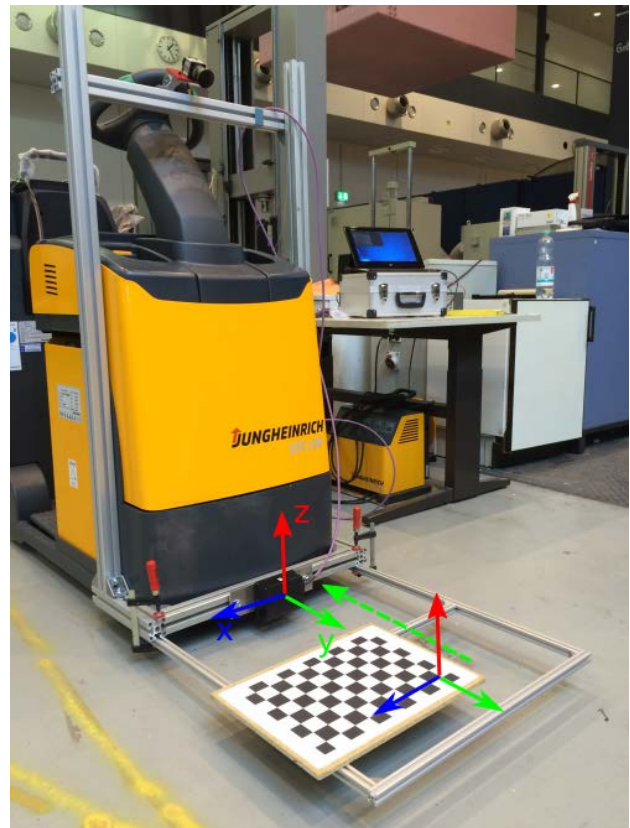


Abbildung 10. Kalibrierung der Sensoren zueinander

Mithilfe der im Vorfeld kalibrierter Kamera wird nun das angebrachte Muster im Bild erkannt und die Beziehung zwischen Muster und Kamera im Raum berechnet. Die Berechnung erfolgt durch das Lösen des sogenannten Perspective-n-Point-Problems [XIA2003]. Danach ist die Pose der Kamera im Raum, bezogen auf die linke obere Ecke des Musters bekannt. Die Pose der Kamera setzt sich aus der Rotationsmatrix R und dem Translationsvektor t_M zusammen. Die y -Komponente des Vektors t_M wird nun um die gemessene Entfernung, zwischen dem Mittelpunkt des Radarsensors und der linken oberen Ecke des Schachbrettmusters, minimiert. Daraus erhält man den Vektor t_R der die Translation zwischen dem Radarsensor und der Kamera beschreibt. Die Rotationsmatrix R bleibt erhalten. Lässt man nun die Ursprünge des Radarkoordinatensystems und des Fahrzeugkoordinatensystems zusammenfallen, so erhält man automatisch die Beziehung Fahrzeug-Radar-Kamera.

Wird ein Hindernis durch den Radarsensor festgestellt, dann werden die x - und y -Komponenten des Vektors t_R um die entsprechenden kartesischen Koordinaten des Hindernisses erweitert und man erhält den, mit dem Hindernis oder dem Objekt verknüpften, Translationsvektor t_O . Da der eingesetzte Radarsensor keine Höheninformationen zu den erkannten Hindernissen liefert, wird die z -Komponente des Vektors t_O , unter Berücksichtigung der Anbringungshöhe des Radarsensors und dessen vertikalen Öffnungswinkels, geschätzt. Die Rotationsmatrix bleibt

hier wieder erhalten. Danach können die Weltkoordinaten des Hindernisses mithilfe der projektiven Transformation in Bildkoordinaten umgerechnet werden.

$$s p_B = K [R|t_O] p_W$$

Dabei steht $p_W = [x \ y \ z \ 1]^T$ für die homogenen Koordinaten eines Punktes im Weltkoordinatensystem (Hinderniskoordinaten). Die korrespondierenden homogenen Bildpunktkoordinaten werden mit $p_B = [u \ v \ 1]^T$ bezeichnet. K ist die Kalibrierungsmatrix, die die intrinsischen Kameraparameter beinhaltet und s bezeichnet den Skalierungsfaktor eines Bildpunktes. Die Kalibrierungsmatrix wird im Vorfeld bei der Kamerakalibrierung ermittelt.

5 ERGEBNISSE

Abbildungen 11 und 12 zeigen die ersten Ergebnisse des vorgestellten Verfahrens im In- und Outdoor-Bereich.

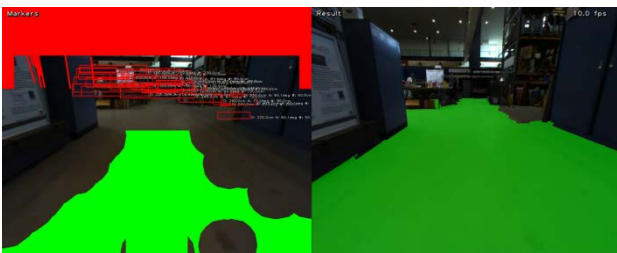


Abbildung 11. Ergebnis im Indoor-Bereich. Links ist das markierte Eingangsbild mit roten Hintergrund- und grünen Vordergrundmarkern. Rechts ist das Ergebnis der Segmentierung. Die befahrbare Fläche im rechten Bild wird grün dargestellt.



Abbildung 12. Ergebnis im Outdoor-Bereich.

Im linken Bild werden die Hintergrundmarker rot und die Vordergrundmarker grün dargestellt. Oben sieht man einen breiten statischen Hintergrundmarker, dieser kann anhand der Kameraausrichtung bestimmt werden. Dieser Marker wird gesetzt, da in diesem Bildbereich keine Abbildung der befahrbaren Fläche zu erwarten ist. Die roten Rechtecke in der Mitte des Bildes, sind die vom Radar erkannten Hindernisse, die in das Kamerabild transformiert wurden. Durch Beobachtung wurde festgestellt, dass es hilfreich sein kann, die im oberen Drittel des Bildes senkrecht verlaufende Linien als Hintergrundmarker zu betrachten. Die Suche nach senkrechten Linien wird mithilfe der Hough-Transformation durchgeführt, nähere Informationen dazu findet man unter anderem in [HAR2009]. Un-

ten befinden sich ein statischer, rechteckiger, grüner Vordergrundmarker sowie ein dynamischer grüner Marker. Im rechten Bild sieht man das ausgewertete Ergebnis. Die befahrbare Fläche wird grün dargestellt.

6 ZUSAMMENFASSUNG UND AUSBLICK

In diesem Beitrag wird ein Verfahren zur automatischen Hinderniserkennung für mobile Arbeitsmaschinen, wie Reinigungsmaschinen, Schlepper, Stapler usw. vorgestellt. Das Verfahren basiert auf der Fusion von Daten eines Radarsensors und eines Monokamerasystems. Die Hinderniserkennung erfolgt dabei durch die Segmentierung des Fahrweges im Kamerabild. Als Segmentierverfahren wird die sogenannte Wasserscheidentransformation verwendet.

Erste Messungen mit dem System zeigen bereits erfolgsversprechende Ergebnisse. Sowohl in der Halle, im Schatten und im direkten Sonnenlicht kann die Fahrbahn vor dem Fahrzeug segmentiert werden.

Der verwendete Radarsensor hat einen relativ kleinen horizontalen Öffnungswinkel. Dadurch kann nicht der gesamte Frontbereich des Fahrzeuges abgedeckt werden. Außerdem ist in dem Sensor eine sehr hohe Empfindlichkeit voreingestellt, dadurch kommt es zu vielen Fehlreflektionen, die nicht gefiltert werden können. An dieser Stelle sind weitere Untersuchungen nötig. In Zukunft wird durch den Einsatz eines anderen Radarsensors versucht, die Ergebnisse des Verfahrens weiter zu verbessern.

LITERATUR

- [GER2001] A. Gern, U. Franke und P. Levi. "Robust vehicle tracking fusing radar and vision". In: *Conference Documentation International Conference on Multisensor Fusion and Integration for Intelligent Systems. MFI 2001 (Cat. No.01TH8590)*. VDI/VDE Soc. Meas. & Autom. Control, Feb. 2001, S. 323–328. ISBN: 3-00-008260-3. DOI: 10.1109/MFI.2001.1013555.
- [COU2002] C. Cou u. a. "Multi-sensor data fusion using Bayesian programming : an automotive application". In: *IEEE/RSJ International Conference on Intelligent Robots and System*. Bd. 1. IEEE, 2002, S. 141–146. ISBN: 0-7803-7398-7. DOI: 10.1109/IRDS.2002.1041379.

- [DEC2009] S. Decker, T. Camus und J. Eledath. "Collision Sensing by Stereo Vision and Radar Sensor Fusion". In: *IEEE Transactions on Intelligent Transportation Systems* 10.4 (Dez. 2009), S. 606–614. ISSN: 1524-9050. DOI: 10.1109/TITS.2009.2032769.
- [RAP2010] Raphaël Labayrade u. a. "Sensor Data Fusion for Road Obstacle Detection: A Validation Framework." 2010.
- [MOE2012] Michael Möser. *Technische Akustik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, S. 537. ISBN: 9783642309328. DOI: 10.1007/978-3-642-30933-5.
- [TOK1996] S. Tokoro. "Automotive application systems of a millimeter-wave radar". In: *Proceedings of Conference on Intelligent Vehicles*. IEEE, 1996, S. 260–265. ISBN: 0-7803-3652-6. DOI: 10.1109/IVS.1996.566388.
- [OHS1999] S. Ohshima u. a. "Phase-comparison monopulse radar with switched transmit beams for automotive application". In: *1999 IEEE MTT-S International Microwave Symposium Digest (Cat. No.99CH36282)*. Bd. 4. IEEE, 1999, S. 1493–1496. ISBN: 0-7803-5135-5. DOI: 10.1109/MWSYM.1999.780235.
- [SHU2013] Shuvra S. Bhattacharyya u. a. *Handbook of Signal Processing Systems*. Springer Science & Business Media, 2013, S. 1420. ISBN: 1461468590.
- [KAT2002] T. Kato, Y. Ninomiya und I. Masaki. "An obstacle detection method by fusion of radar and motion stereo". In: *IEEE Transactions on Intelligent Transportation Systems* 3.3 (Sep. 2002), S. 182–188. ISSN: 1524-9050. DOI: 10.1109/TITS.2002.802932.
- [SOL2004] A. Sole u. a. "Solid or not solid: vision for radar target validation". In: *IEEE Intelligent Vehicles Symposium, 2004*. IEEE, 2004, S. 819–824. ISBN: 0-7803-8310-9. DOI: 10.1109/IVS.2004.1336490.
- [PRO2008] Zhengping Ji und D. Prokhorov. "Radar-vision fusion for object classification." 2008.
- [BAE2009] Bernard Bäker. *Moderne Elektronik im Kraftfahrzeug IV*. expert verlag, 2009, S. 213. ISBN: 3816926681.
- [MIL2011] A. Milella u. a. "Combining radar and vision for self-supervised ground segmentation in outdoor environments". In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Sep. 2011, S. 255–260. ISBN: 978-1-61284-456-5. DOI: 10.1109/IROS.2011.6094548.
- [SOI1998] Pierre Soille. *Morphologische Bildverarbeitung - Grundlagen, Methoden, Anwendung*. 1998. DOI: 10.1007/978-3-642-72190-8.
- [GON2008] Rafael C. Gonzalez und Richard E. Woods. *Digital Image Processing, 3rd edition*. 2008. DOI: 10.1049/ep.1978.0474.
- [KLE1999] L.A. Klein. *Sensor and data fusion concepts and applications*, 2. Aufl. Bd. 35. SPIE, Bellingham, Wash (1999)
- [XIA2003] Gao Xiao-Shan, Hou Xiao-Rong; Tang Jianliang, Cheng Hang-Fei (2003). "Complete Solution Classification for the Perspective-Three-Point Problem." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 25 (8): 930–943. DOI: 10.1109/tpami.2003.1217599
- [HAR2009] Peter E. Hart. "How the Hough transform was invented." In: *IEEE Signal Processing Magazine* November 2009 S. 18–22

Gregory Föll, M.Sc. studierte Informatik an der Hochschule für angewandte Wissenschaften in Hamburg und ist seit 2013 am Lehrstuhl für Maschinenelemente und Technische Logistik als wissenschaftlicher Mitarbeiter tätig.

Prof. Dr.-Ing. Rainer Bruns hat Luft- und Raumfahrt-technik in Braunschweig studiert und ist seit 1992 Leiter des Lehrstuhls für Maschinenelemente und Technische Logistik an der Helmut-Schmidt-Universität in Hamburg.

Adresse: Helmut-Schmidt-Universität, Universität der Bundeswehr Hamburg, Fakultät für Maschinenbau, Institut für Konstruktions- und Fertigungstechnik, Holstenhofweg 85, 22043 Hamburg

Telefon (R. Bruns): +49 40 6541-2855, E-Mail: gregory.foell@hsu-hh.de, rainer.bruns@hsu-hh.de