

Robuste Erkennung von Verkehrsteilnehmern zur Vermeidung von Unfällen beim Türöffnen für ein Fahrerassistenzsystem

Christian Scharfenberger und Georg Färber
Lehrstuhl für Realzeit Computersysteme
Technische Universität München, Germany
{christian.scharfenberger, georg.farber}@rcs.ei.tum.de

Zusammenfassung—Seit einigen Jahren gewinnen Fahrerassistenzsysteme zunehmend an Bedeutung. Diese sollen den Fahrer während des Fahrens unterstützen und das Risiko eines Unfalls reduzieren. Viele Unfälle passieren jedoch, wenn das Fahrzeug steht und eine der Seitentüren plötzlich geöffnet wird. Besonders Fahrradfahrer sind hier gefährdet, da sie relativ nahe an Fahrzeugen entlang fahren und leicht übersehen werden. Dieser Beitrag präsentiert ein Assistenzsystem integriert in eine intelligente, aktivierte Fahrzeugtür zur Vermeidung von gefährlichen Unfällen beim Türöffnen. Das System basiert auf einer omnidirektionalen Kamera, die in die Seitenspiegel integriert den Außenraum um die Türen vollständig auf annähernde Verkehrsteilnehmer überwacht, und ein Assistenzsystem, das die Passagiere vor einem gefährlichen Türöffnen warnt oder aktiv Türbewegungen blockiert.

I. EINLEITUNG

In modernen Fahrzeugen gewinnen Fahrerassistenzsysteme immer mehr an Bedeutung. Als Beispiele seien die Systeme Lane Departure Warning (LDW), Adaptive Cruise Control (ACC), Forward Collision Warning (FCW) und Blind-spot Detection (BSD) genannt [1]. Diese Sicherheitssysteme sind primär darauf ausgelegt, den Fahrer während des Fahrens zu unterstützen und das Risiko eines Unfalls aktiv zu vermeiden. Viele Unfälle passieren aber, wenn das Fahrzeug steht und eine von seinen Seitentüren plötzlich geöffnet wird: Gängige Praxis ist das Beachten von Hindernissen mit einem Blick in den Seitenspiegel oder neben die Tür, bevor die Tür geöffnet wird. Doch genau dies vergessen viele Passagiere beim Türöffnen. Eine plötzlich geöffnete Tür kann dann schwere Unfälle hervorrufen. Besonders gefährdet sind hierbei Fahrradfahrer, da sie relativ nahe am Fahrzeug entlang fahren und leicht übersehen werden.

In diesem Beitrag wird ein Sicherheitssystem vorgestellt, welches aus einer intelligenten Fahrzeugtür und aus einem omnidirektionalen Kamerasystem besteht. Jeweils eine Kamera ist auf jeder Fahrzeugseite in die Außenspiegel integriert und kann aufgrund ihres großen Sichtfeldes die Umgebung um die Türen vollständig auf annähernde Hindernisse überwachen. Bei annähernden Objekten werden Passagiere und Türnutzer vor Hindernissen wie Fahrzeugen, Fahrradfahrer oder Fußgänger gewarnt. Ein Kollisionserkennungssystem nutzt die generierten Hindernisinformationen, um ein mögliches Türöffnen zu blockieren oder um eine bereits

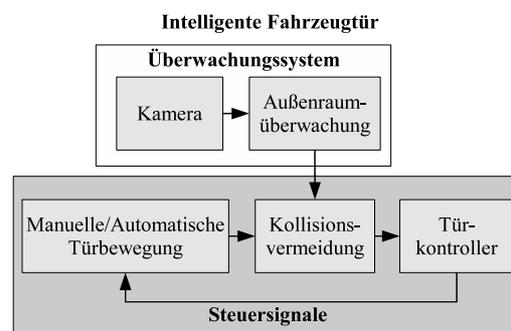


Abbildung 1. Intelligentes Türsystem: Die Informationen der Kamera und des Überwachungssystems dienen als Input für die Türregelung, z.B. zum Blockieren der Tür zur Vermeidung von Unfällen.

geöffnete Tür anzuhalten. Abbildung 1 gibt eine High-Level Übersicht von der verwendeten intelligenten Tür. In dem Beitrag von Strolz *et al.* [2] wird ein generisches Steuersystem für aktivierte, geregelte Fahrzeugtüren mit mehr als einem Freiheitsgrad vorgestellt. Der Fokus des genannten Beitrages liegt im mechanischen Design und der Regelung dieser Fahrzeugtüren. Eine wichtige Komponente einer solchen Tür ist jedoch das Bilderfassungs- und -verarbeitungssystem, dessen Hindernisdaten als Input für die nachfolgende Türregelung dient.

Der Fokus dieses Beitrages liegt daher in der Beschreibung des Kamerasystems, der Präsentation robuster Algorithmen zur Detektion annähernder Hindernisse sowie der Bestimmung deren Gefährdungsgrades. Studien zeigen, dass statische Objekte neben der Fahrzeugtüre in den meisten Fällen von den Passagieren beachtet werden. Diese können aber mit anderen Sensoren oder bildbasiert mit Hilfe von Motion-Stereo Algorithmen [3], [4] mit nur einem Kamerasystem beim Türöffnen oder beim Einparken erfasst werden.

Das in diesem Prototyp verwendete Kamerasystem besteht aus einer omnidirektionalen Kamera (siehe Abbildung 2), die aufgrund ihrer physikalischen Eigenschaften in den Außenspiegel eines Fahrzeuges integriert den Außenraum um die Tür vollständig überwachen kann. Abbildung 3 zeigt die prototypische Integration einer solchen Kamera. Abbildung 4 zeigt das Panoramabild einer typischen Parksituation, welches von der

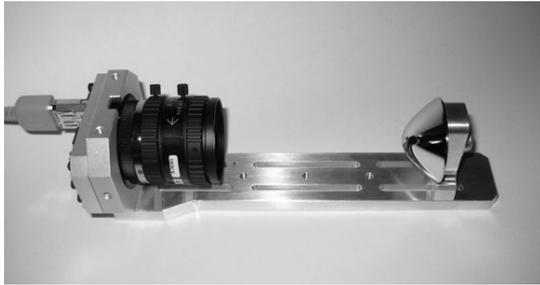


Abbildung 2. Omnidirektionales Kamerasystem mit einem Sichtfeld von 360°.



Abbildung 3. Prototypisch in den Außenspiegel integriertes Kamerasystem.

omnidirektionalen Kamera aufgenommen wurde. Aufgrund des sehr großen Sichtbereiches der Kamera ist es möglich, annähernde Hindernisse frühzeitig zu detektieren und deren Gefährdungsgrad einzuschätzen.

Für eine sichere manuelle oder automatische Türbetätigung müssen Hindernisse detektiert sein, bevor die Tür geöffnet wird. Annähernde Objekte müssen daher erfasst und deren Gefährdungsgrad bestimmt sein, auch wenn diese noch relativ weit vom Fahrzeug entfernt und auf sehr wenige Pixel im Bild abgebildet werden. Des weiteren gibt es zwischen einem Einparkvorgang und einem Türöffnen ein kleines Zeitintervall, welches zur Extraktion von Hindernissen genutzt werden kann. Dies erlaubt das Benennen von Bedingungen, unter denen der vorgestellte Algorithmus arbeitet: (i) Das Kamerasystem kann für ein kleines Zeitintervall als statisch angenommen werden. Dieses Zeitintervall kann beispielsweise zum Lernen der Türumgebung verwendet werden. (ii) Objekte, die weit vom Fahrzeug entfernt sind, werden auf nur wenig Pixel abgebildet und sind damit schwer vom Hintergrund unterscheidbar. (iii) Aufgrund des großen Sichtfeldes der Kamera gibt es verschiedene Ansichten von einem sich bewegenden Objekt (vorne, Seite, hinten) (siehe Abbildung 5). (iv) Der ausgewählte Algorithmus muss in Realzeit arbeiten, da sehr schnell fahrende Fahrzeuge innerhalb weniger Bilder extrahiert sein müssen. (max. 2 - 3).

In der Literatur gibt es viele Techniken zur Extraktion von Objekten. Als Beispiele seien Detektionsmethoden basierend auf Hidden Markov Modellen [5], Template Matching [6], ba-



Abbildung 4. Panoramabild der Türumgebung aufgenommen mit einem omnidirektionalen Kamerasystem.

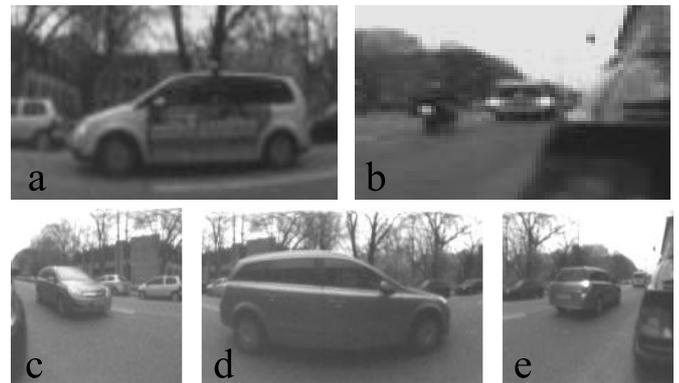


Abbildung 5. a) Große und b) sehr kleine Objekte. c) - e) Verschiedene Ansichten auf ein fahrendes Fahrzeug.

sierend auf Featurepunkten, Optischen Fluss [7],[8] und basierend auf Hintergrundmodellen genannt. Tabelle I gibt eine kurze Übersicht auf gängige Extraktionsmethoden und vergleicht diese in Bezug auf Rechenzeit, Parallelisierbarkeit und Robustheit gegenüber perspektivischen Änderungen. Hintergrundmodelle sowie optischer Fluss sind unempfindlich gegenüber perspektivischen Änderungen und benötigen im Vergleich zu anderen Methoden kein Vorwissen über Objekteigenschaften wie Farbe oder Objektgröße. Hintergrundmodelle haben gegenüber optischen Fluss auch den Vorteil, dass bei einem einmal gelernten Hintergrundbild auch statische Objekte detektiert werden können. Allerdings benötigen Hintergrundmodelle ein kleines Zeitintervall zum Lernen eines Hintergrundbildes: Dieses Zeitintervall existiert aber in gängigen Parksituationen und kann zur Bestimmung eines Hintergrundbildes verwendet werden. Objekte werden in dieser Anwendung mithilfe eines Hintergrundmodells extrahiert, welches zur Extraktion dieser in Realzeit geeignet ist. Abbildung 6 zeigt das Blockdiagramm des vorgeschlagenen Extraktionsalgorithmus sowie die abschließende Bestimmung des Gefährdungsgrades. In diesem Beitrag wird das verwendete Hintergrundmodell sowie Erweiterungen wie Belechtungskompensation und Schattenelementation beschrieben. Diese Methoden wurden speziell für die Anwendung in der intelligenten Fahrertüre angepasst. Im folgenden Kapitel werden die Details des gewählten Extraktionsalgorithmus ausführlicher beschrieben.

	Initialisierung	Vorwissen	Statische Objekte	Änderungen Objektgröße	Perspektivische Änderungen	Parallelisierbarkeit	Rechenzeit
Template Matching	nein	benötigt	-	schlecht	schlecht	mittel	langsam
HMM	nein	benötigt	-	schlecht	schlecht	mittel	langsam
Feature Based	nein	benötigt	-	schlecht	schlecht	gut	schnell
Optical Flow	nein	nein	schlecht	-	-	gut	mittel
Hintergrundmodelle	ja	nein	-	-	-	gut	schnell

Tabelle I

VERGLEICH VERSCHIEDENER OBJEKTEXTRAKTIONSTECHNIKEN IN BEZUG AUF EXTRAKTION VON VERKEHRSTEILNEHMERN. *Hintergrundmodelle* SCHEINEN FÜR UNSERE APPLIKATION AM BESTEN GEEIGNET ZU SEIN (-: *unabhängig*).

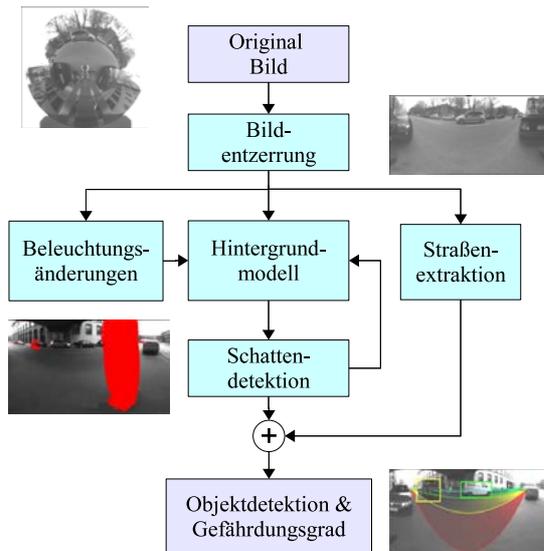


Abbildung 6. Blockdiagramm zur Objektdetektion mit Hilfe von Hintergrundbestimmung, Schatteneeliminierung und Kompensation von Beleuchtungsänderungen.

A. Stand der Technik und Forschungsansatz

In der Literatur sind viele Methoden und Anwendungen zur Objektextraktion vorgestellt worden, wie zum Beispiel Umgebungsüberwachung [9], Verkehrsbeobachtungen [10] und Fahrzeugführung. Für die meisten Anwendungen ist die Separation von Vordergrund und Hintergrund der erste Schritt zur Objektdetektion. Extraktionstechniken basierend auf Hintergrundmodellen sind mächtige Methoden, deren Vorteil in der objekt- und featureunabhängige Detektion liegt. Einige bekannte Techniken für Hintergrundmodelle basieren auf Kalman-Filterung [11], Kernel-Density-Funktion [12], Hidden-Markov-Modellen [13], Mixture-of-Gaussians Modellen [14], [15] sowie die Verwendung von farbusabhängigen Features [16]. Viele von diesen Algorithmen repräsentieren ein Hintergrundpixel basierend auf einer Probabilistischen Dichte Funktion (PDF) und klassifizieren Pixel als Vordergrund oder Hintergrund abhängig dieser Dichtefunktionen. Bhaskar *et al.* [17] entwickelten eine Methode basierend auf Gaussian-Mixture-Modellen zur Modellierung von dynamischen Hintergrund und zur Detektion und Kompensation von Beleuchtungsänderungen. Zhong *et al.* [18] verwenden Kalman-Filter zur iterativen Schätzung dynamischer Hintergrundtexturen sowie zur Extraktion von Regionen, die Vordergrund enthal-

ten. Auch Karman *et al.* verwenden zur Modellierung von dynamischen Hintergrund Kalman-Filter. Dafür schlagen sie zwei Verstärkungsfaktoren vor, die zum einen eine schnelle Adaption von Hintergrundveränderungen und zum anderen eine langsame Adaption von Vordergrund in den Modellen ermöglichen. Ridder *et al.* [19] verbessern diesen Ansatz und präsentieren einen Algorithmus zur Kompensation von Schattenpixeln. Dabei nehmen sie an, dass Schatten nur kleine Änderungen zum Hintergrund erzeugt.

Obwohl viele Hintergrundmodelle entwickelt und vorgestellt wurden, basieren die Mehrzahl dieser Modelle auf Farbinformationen zur Detektion von Schatten und Beleuchtungsänderungen [12], [18]). Für viele Anwendungsfälle, in denen nur monochromatische Kameras verwendet werden, sind Methoden – z.B. zur Schatteneeliminierung – basierend auf Farbinformationen nicht mehr anwendbar. In monochromatischen Bildern ist es weiterhin schwierig, zwischen Schatten oder kleinen, dunklen Vordergrundobjekten zu unterscheiden, die nur eine kleine Differenz zum Hintergrundmodell aufweisen. Diese Probleme, zusammen mit der Genauigkeit einer Extraktion von Vordergrundobjekten sowie die Kompensation von Beleuchtungsänderungen ist die Motivation dieses Beitrages.

Inspiziert von dem Hintergrundmodell von Ridder *et al.* und von dem Schattendetektor präsentiert von Jacques Jr. *et al.* [20], wird in diesem Beitrag eine Methode zur robusten Detektion von Verkehrsteilnehmern sowie eine Methode zur Einschätzung deren Gefährdungsgrads präsentiert. Schatten wird in der Arbeit von Ridder [19] detektiert, aber wenn Schlagschatten einmal als Vordergrund detektiert worden ist, ist es unmöglich, mit deren Methode zwischen Schatten und Vordergrund zu unterscheiden. Jacques Jr. *et al.* [20] präsentieren einen guten Schattendetektor basierend auf der normierten Kreuzkorrelation (NCC). Dieser Detektor nimmt Schatten als verdunkelte Hintergrundpixel an, so dass die Ähnlichkeit zwischen Schatten und Hintergrundbild, also der Betrag der NCC in einer definierten Nachbarschaft 1 wird. Allerdings werden kleinere Objekte wie Motorräder, die sich kaum vom Hintergrund unterscheiden, fälschlicherweise als Schattenpixel klassifiziert. Um dies zu vermeiden, wurde das Hintergrundmodell von Ridder *et al.* um den Schattendetektor von Jaques Jr. *et al.* ergänzt. Dieser Schattendetektor wurde weiterhin um die mittelwertbefreiten normierten Kreuzkorrelation (ZNCC) erweitert, um zwischen Schatten- und Vordergrundpixeln zu unterscheiden. Beleuchtungsänderungen

können bei Hintergrundmodellen zu Fehldetektionen führen und müssen daher kompensiert werden. Dies wird mittels lokaler Suchfenster realisiert, und das Hintergrundmodell entsprechend aktualisiert. Damit können sowohl schnelle und als auch langsame Beleuchtungsunterschiede kompensiert werden.

Der Rest dieses Beitrages ist folgendermaßen aufgebaut: Kapitel II-A beschreibt die angewandten Methoden zur Initialisierung des Hintergrundmodells sowie die Straßenextraktion zur Einschätzung des Gefährdungsgrades von annähernden Objekten. Das dynamische Hintergrundmodell, die Schatten-detektion und die Kompensation von Beleuchtungsunterschieden werden ab Kapitel II-C vorgestellt. Messergebnisse werden in Kapitel III präsentiert und diskutiert. Dieser Beitrag endet mit einer kurzen Zusammenfassung und einen Ausblick.

II. HINTERGRUNDMODELL

Wie in Kapitel I beschrieben, werden Hintergrundmodelle zur Extraktion von annähernden Objekten wie Fahrzeugen oder Motorräder verwendet. Das hierbei verwendete Hintergrundmodell basiert auf den Methoden von Karman *et al.* [11] und Ridder *et al.* [19] und wurde erweitert, um eine bessere Erkennung von Schatten sowie eine Steigerung der Robustheit gegenüber Beleuchtungsunterschieden zu gewährleisten. In Kapitel II-A und Kapitel II-C werden die mathematischen Zusammenhänge des Hintergrundmodells basieren auf Kalman-Filterung beschrieben. Dazu werden in einem ersten Schritt Pixel als Hintergrund oder als möglicher Vordergrund klassifiziert. Danach werden Pixel initial mittels NCC (siehe Kapitel II-D) als Schatten identifiziert und das Ergebnis mittels ZNCC (siehe Kapitel II-E) verbessert. Dies hat den Vorteil, dass kleine Objekte nicht als Schatten klassifiziert werden. Beleuchtungsänderungen werden mit den vorgestellten Methoden in Kapitel II-F detektiert und das Hintergrundmodell dementsprechend angepasst. Kapitel II-H beschreibt die abschließende Gefährdungsgradbestimmung detektierter Objekte.

A. Initialisierung des Hintergrundmodells

Für viele Hintergrundmodelle wird ein Hintergrundbild benötigt, das möglichst frei von Vordergrundobjekten sein sollte. In typischen Parksituation, z.B. an belebten Straßen, ist es nicht möglich, ein leeres Hintergrundbild aufzunehmen. In der vorgestellten Methode von Ridder *et al.* wird ein fester Initialwert für das Hintergrundmodell verwendet, der während einer Trainingsphase an die Umgebungssituation angepasst wird. Jacques *et al.* verwenden eine Median-basierte Hintergrundinitialisierung. Diese Initialisierung erlauben die Extraktion des Hintergrundbildes in belebten Szenarien, jedoch wird angenommen, dass der leere Hintergrund für mehr als die Hälfte der Trainingsdaten vorhanden ist. Reale Szenarien zeigen aber, dass diese Voraussetzung gerade an belebten Straßen nicht immer gegeben ist. Farin *et al.* schlagen in [21] eine interessante Methode vor, um dieses Problem zu lösen: Die Idee von ihrem Algorithmus besteht in der groben Selektion von Vordergrund und Hintergrund. Dazu wird in der Nachbarschaft eines jeden Pixels und für jedes Eingangsbild ein Block gebildet. Hintergrundpixel sind Pixel, in denen der

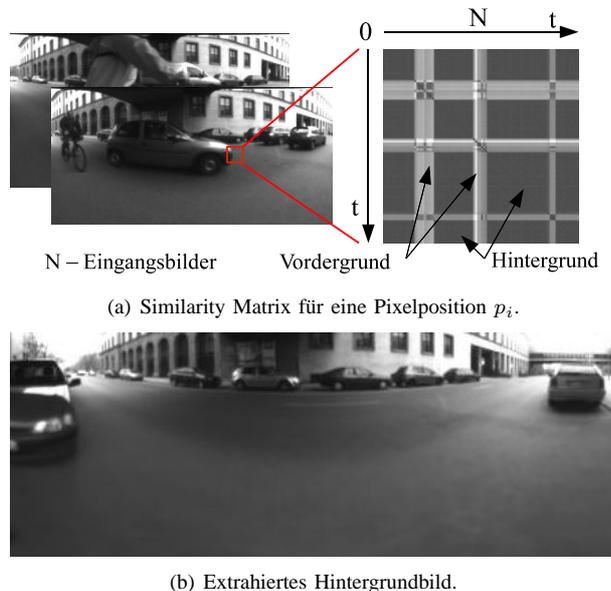


Abbildung 7. Abbildung (a) zeigt eine Similarity-Matrix für eine Pixelposition p_i . Kleine Differenzen zwischen den Blöcken werden dunkel, große Differenzen zwischen den Blöcken werden hell angezeigt. (b): Extrahiertes Hintergrundbild zur Initialisierung des Hintergrundmodells.

Kontext in den dazugehörigen Blöcken stabil ist. Um diese Blöcke mit Hintergrundpixeln zu finden, wird die Ähnlichkeit der Blöcke für jedes Pixel während einer festen Trainingsperiode innerhalb einer *Similarity Matrix* gespeichert. In diese Matrix werden an die einzelnen Positionen die Differenzen der einzelnen Blocks eingetragen, wobei die Differenz durch die *Summe der Absoluten Differenzen* gebildet werden. Kleine Ergebnisse an den einzelnen Matrixeinträgen bedeuten Hintergrundpixel, während große Werte an den einzelnen Stellen Veränderungen anzeigen, die durch Vordergrund hervorgerufen wurden. Die Einträge der Matrizen für jedes Pixel werden dann in zwei Teile aufgeteilt, um Vordergrund und Hintergrund zu separieren. Das entsprechende Hintergrundbild wird abschließend aus den Hintergrundpixeln berechnet, wobei Ausreißer durch Median-Filterung [22] beseitigt werden. Das so generierte Hintergrundbild wird zur Initialisierung des Kalman-basierten Hintergrundmodells mit dem Systemzustand $\hat{I}_{x,y}(t_0)$ verwendet. Abbildung 7 zeigt eine *Similarity Matrix* für eine Pixelposition und das dabei extrahierte Hintergrundbild.

B. Straßenextraktion und Zoneneinteilung

Aus dem gewonnenen Hintergrundbild wird in einem nächsten Schritt die Straße extrahiert und diese in Zonen zur späteren Bestimmung des Gefährdungsgrad eingeteilt. Als erstes wird mit Hilfe von Straßenextraktionsalgorithmen [23] die Straße extrahiert und als eigenständiger Bereich für weitere Verarbeitungsschritte gespeichert. Die dabei verwendete Methode von Alvarez *et al.* basiert auf einem Region-Growing-Algorithmus [24], der in einem beleuchtungsunabhängigen Farbraum Straßen extrahiert. In diesem Beitrag wird nach einem Helligkeits- und Beleuchtungsausgleich die Fahrbahn mittels Region-Growing extrahiert, mit deren Hilfe

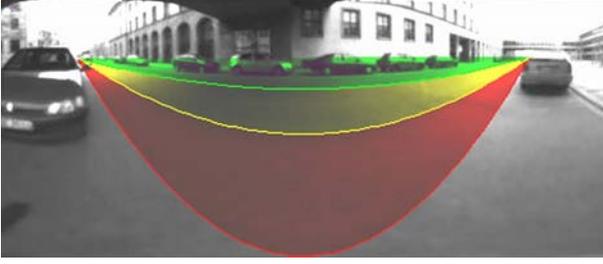


Abbildung 8. Typisches Straßenszenario mit eingeteilten Zonen zur späteren Bestimmung des Gefährdungsgrades.

im nächsten Schritt die Einteilung in die gefährdeten Bereiche erfolgt. Zur Bestimmung der gefährdeten Bereiche helfen Straßenmarkierungen, sofern diese in der jeweiligen Parksituation vorhanden sind. Fahrbahnmarkierungen werden in den bestimmten Straßenbereichen mittels Bandpassfilterung nach Kantenverstärkung extrahiert und angezeigt. Typische Verfahren wie Extraktion von Straßenmarkierungen mittels Hough-Transformation können in dieser Anwendung aufgrund von ungeraden Linienverläufen in Panoramabildern nicht angewandt werden. Mit Hilfe des Straßenverlaufes und basierend auf Straßenmarkierungen kann nun der Bereich um die Tür in drei Zonen eingeteilt werden. Objekte im roten Bereich können mit einer geöffneten Tür kollidieren, während Objekte im grünen Bereich ungefährlich sind. Objekte im gelben Bereich erzeugen keine unmittelbare Kollision, können sich aber relativ schnell in den roten Bereich hineinbewegen und sind daher als gefährdet anzusehen. Abbildung 8 zeigt die eingeteilten Zonen zur Schätzung des Gefährdungsgrades.

C. Kalman basiertes Hintergrundmodell

In diesem Kapitel wird das Hintergrundmodell beschrieben, auf dem die Objekterfassung basiert. Die Intensität eines Pixels an der Position (x, y) zum Zeitpunkt t wird als $I_{x,y}(t)$ definiert. Der geschätzte Systemzustand des Hintergrundmodells wird als $\hat{I}_{x,y}(t)$ und dessen Ableitung als $\dot{\hat{I}}_{x,y}(t)$ bezeichnet. Der geschätzte Hintergrundverlauf für ein Pixel ergibt sich somit als:

$$\begin{bmatrix} \hat{I}_{x,y}(t) \\ \dot{\hat{I}}_{x,y}(t) \end{bmatrix} = \begin{bmatrix} \tilde{I}_{x,y}(t) \\ \tilde{\dot{I}}_{x,y}(t) \end{bmatrix} + K_{x,y}(t) \cdot \left(I_{x,y}(t) - H \cdot \begin{bmatrix} \tilde{I}_{x,y}(t) \\ \tilde{\dot{I}}_{x,y}(t) \end{bmatrix} \right) \quad (1)$$

Nach Gleichung 2 ist die Prediktion $\tilde{I}_{x,y}(t)$ des Systemzustandes $\hat{I}_{x,y}(t)$ und dessen Ableitung $\tilde{\dot{I}}_{x,y}(t)$ zum Zeitpunkt t wie folgt definiert:

$$\begin{bmatrix} \tilde{I}_{x,y}(t) \\ \tilde{\dot{I}}_{x,y}(t) \end{bmatrix} = S \cdot \begin{bmatrix} \hat{I}_{x,y}(t-1) \\ \dot{\hat{I}}_{x,y}(t-1) \end{bmatrix} \quad (2)$$

Mit der Systemmatrix S , der Messmatrix H der Kalman-Verstärkung K ergibt sich:

$$S = \begin{bmatrix} 1 & s_{1,2} \\ 0 & s_{2,2} \end{bmatrix}, H = [1 \quad 0] \quad \text{and} \quad K_{x,y}(t) = \begin{bmatrix} k_{1,x,y}(t) \\ k_{2,x,y}(t) \end{bmatrix} \quad (3)$$

Gemäß [11] werden $s_{1,2} = s_{2,2} = 0.7$ als Werte zur Modellierung der Dynamik der Hintergrundmodells verwendet. Da die Kamera nur Grauwerte $I_{x,y}(t)$ generiert, wird die Messmatrix H als konstant betrachtet. Die Kalman-Verstärkung K wird anhand detektiertem Vordergrund oder Hintergrund bestimmt. Dazu wird eine Pre-estimation vom nächsten Systemzustand vorgenommen (siehe Gleichung 4 und Gleichung 5).

$$m_{x,y}(t) = \begin{cases} 1, & \text{if } \left[\begin{array}{l} d'_{x,y}(t) \geq th_{bg} \\ (d'_{x,y}(t) < th_{bg}) \wedge \\ (d''_{x,y}(t) \geq th_{bg}) \end{array} \right] \vee \\ 0, & \text{if } \left[\begin{array}{l} (d'_{x,y}(t) < th_{bg}) \wedge \\ (d''_{x,y}(t) < th_{bg}) \end{array} \right] \end{cases} \quad (4)$$

$$\begin{aligned} d'_{x,y}(t) &= |I_{x,y}(t) - \tilde{I}_{x,y}(t)| \\ d''_{x,y}(t) &= |I_{x,y}(t) - \dot{\tilde{I}}_{x,y}(t)| \end{aligned} \quad (5)$$

with $\dot{\tilde{I}}_{x,y}(t) = \tilde{\dot{I}}_{x,y}(t) + \beta \cdot [I_{x,y}(t) - \tilde{I}_{x,y}(t)]$

Pixel, deren Differenz zwischen Hintergrundmodell (Systemzustand) kleiner als ein bestimmter Schwellwert ($d' < th_{bg}$) sind, müssen nicht zwangsläufig zum Hintergrund gehören. Um eine falsche Zuordnung zu vermeiden, wird eine Pre-Estimation $\dot{\tilde{I}}_{x,y}(t)$ vom nächsten Systemzustand vorgenommen. Dies geschieht unter der Annahme, dass solche Pixel zum Hintergrund gehören. Ist der geschätzte Messwert d'' größer als der Schwellwert th_{bg} , handelt es sich bei diesem Pixel um ein Vordergrundpixel. Anderenfalls gehört dieser Pixel zum Hintergrund. Die Segmentierung Hintergrund/Vordergrund wird in einer binären Karte $m_{x,y}(t)$ gespeichert, bei der 1 Vordergrund und 0 Hintergrund repräsentiert. Davon abhängig wird die Kalmanverstärkung $k_{1,2,x,y}(t) = \alpha$ oder $k_{1,2,x,y}(t) = \beta$ gewählt (siehe Gleichung 6).

$$k_{1,2,x,y}(t) = \begin{cases} \alpha, & \text{if } m_{x,y}(t) = 1 \\ \beta, & \text{if } m_{x,y}(t) = 0 \end{cases} \quad (6)$$

D. Schattendetektion

Soll die Fahrzeugtüre sicher geöffnet oder bei einer möglichen Kollision rechtzeitig abgebremst werden, ist die Einschätzung des Gefährdungsgrades eines Objektes von essentieller Bedeutung. Diese Einschätzung basiert auf der aktuellen Position des Objektes und seiner geschätzten Bewegungstrajektorie. Schatten, der durch Objekte erzeugt wird, führt zu einer Verfälschung der Objektposition und damit auch zu einer Verfälschung der geschätzten Bewegungstrajektorie. Von daher muss der Schatten von Objekten unterdrückt werden. Zur initialen Bestimmung von Schattenpixeln wird die normierte Kreuzkorrelation verwendet (NCC - [20]), wobei das Ergebnis anschließend mit der mittelwertbefreiten normierten Kreuzkorrelation (ZNCC) verfeinert wird. Somit

können Schattenpixel von Vordergrundpixel mit kleiner Differenz zum Hintergrund unterschieden werden.

Sei $\tilde{I}_{x,y}(t)$ das geschätzte Hintergrundbild und $I_{x,y}(t)$ ein aktuell von der Kamera aufgenommenes Bild. Zusätzlich wird um jedes Vordergrundpixel eine Template $T_{xy}(n,m) = I_{x+n,y+m}(t)$ mit der Größe $-N \leq (n,m) < N$ generiert, wobei dem es sich bei \bar{t} um den Mittelwert des Templates $T_{xy}(n,m)$ handelt. Analog dazu wird ein Template $B_{xy}(n,m)$ von jeden Hintergrundpixel mit $B_{xy}(n,m) = \hat{I}_{x+n,y+m}(t)$ gebildet, wobei auch hier \bar{b} der Mittelwert des Templates $B_{xy}(n,m)$ sei. Die Ähnlichkeit beider Templates kann nun mittels der mittelwertbefreiten normierten Kreuzkorrelation ZNCC und der normierten Kreuzkorrelation NCC mit ($\bar{t} = 0, \bar{b} = 0$) für jede Pixelposition (x,y) nach Gleichung 11 berechnet werden:

$$ZNCC_{x,y} = \frac{EZR_{x,y}}{EZB_{x,y} \cdot EZT_{x,y}} \quad (7)$$

mit

$$\begin{aligned} EZR_{x,y} &= \sum_{n=-N}^N \sum_{m=-N}^N |(B_{xy}(n,m) - \bar{b})|(T_{xy}(n,m) - \bar{t})| \\ EZB_{x,y} &= \sqrt{\sum_{n=-N}^N \sum_{m=-N}^N (B_{xy}(n,m) - \bar{b})^2} \\ EZT_{x,y} &= \sqrt{\sum_{n=-N}^N \sum_{m=-N}^N (T_{xy}(n,m) - \bar{t})^2} \end{aligned} \quad (8)$$

Dabei beschreibt $EZT_{x,y}$ die Energie des Bildtemplates T und $EZB_{x,y}$ die Energie des geschätzten Hintergrundbildes (Systemzustand) B . Ein Pixel wird als mögliches Schattenpixel identifiziert, wenn der NCC zwischen Template und Hintergrund nahezu eins ist und die Energie $ET_{x,y}$ des Templates T kleiner als die Energie $EB_{x,y}$ des Hintergrundtemplates B ist (siehe Gleichung 9).

$$NCC_{x,y} \geq th_{NCC} \text{ and } EB_{x,y} > ET_{x,y} \quad (9)$$

Sowohl $EB_{x,y}$ als auch $ET_{x,y}$ kann mit Hilfe von $EZB_{x,y}(\bar{b} = 0)$ und $EZT_{x,y}(\bar{t} = 0)$ berechnet werden.

E. Schattenverbesserung

Abhängig des gewählten Schwellwert th_{NCC} mit ($th_{NCC} < 1.0$) können viele Vordergrundpixel mit nur geringer Differenz zu Hintergrundpixel als Schattenpixel klassifiziert werden. Um dieses Problem zu umgehen, werden die Ergebnisse der initialen Schattendetektion mit Hilfe der ZNCC verbessert, um Schatten- von Vordergrundpixeln zu unterscheiden. Der Vorteil der ZNCC gegenüber der NCC ist, dass Beleuchtungsänderungen keinen Einfluss auf die ZNCC haben. Von daher erzeugen nur Änderungen in der Textur signifikante Änderungen der ZNCC. Der hier vorgeschlagene Verbesserungsprozess mittel ZNCC überprüft, ob Änderungen der genannten Schattenpixel durch Schatten (=Beleuchtung) oder durch Textur hervorgerufen wurde. Obwohl die ZNCC lichtinvariant ist, erzeugt Bildrauschen (=Texturveränderungen) einen Offset. Dieser Offset kann während der Lernphase bestimmt und bei der Wahl des Schwellwertes th_{ZNCC} berücksichtigt werden. Ähnlich der

Klassifikation mittels der NCC ist ein Vordergrundpixel ein Pixel bei dem die ZNCC beider Templates nahe dem gelernten Initialwert ist und die Energie $ET_{x,y}$ des Bildtemplates T kleiner als die Energie des Hintergrundtemplates $EB_{x,y}$ ist. Im Gegensatz zur NCC repräsentieren die Energien $EZT_{x,y}$ und $EZB_{x,y}$ nur die Energien der Texturen von Hintergrund und Bild. Von daher kann die Energie eines Template um ein Vordergrundpixel auch kleiner als die Energie des Hintergrundtemplates sein (z.B. Hervorgerufen durch große homogene Objekte wie LKWs). Ein Pixel ist also nur dann ein Schattenpixel, wenn die Energie $EZT_{x,y}$ des Template T ungefähr so groß ist wie die Energie $EZB_{x,y}$ des Hintergrundes B . Gleichung 11 zeigt diesen Zusammenhang:

$$\begin{aligned} |ZNCC_{x,y} - (1.0 - \theta)| &\leq th_{ZNCC} \text{ and} \\ |EZB_{x,y} - EZT_{x,y}| &\leq th_{comp} \text{ and} \\ ET_{x,y} &< EB_{x,y} \end{aligned} \quad (10)$$

F. Aktive Kompensation von Beleuchtungsunterschieden

Hintergrundmodelle basierend auf Kalman-Filterung können langsame Beleuchtungsänderungen in das Modell aufnehmen. Wenn jedoch Vordergrundobjekte den Hintergrund verdecken, können Beleuchtungsunterschiede nicht vollständig detektiert werden. Weiterhin können sehr schnelle Beleuchtungsunterschiede nicht kompensiert werden, da diese abhängig vom gewählten Schwellwerte als Vordergrund betrachtet werden. Von daher muss das Hintergrundmodell derart modifiziert werden, dass dieses sowohl langsame als auch schnelle Veränderungen detektiert und kompensiert. Jedes eingehende Bild wird daher in m Subimages zu den jeweiligen Positionen $p_x(m)$ und $p_y(m)$ unterteilt. Weiterhin wird zu jedem Subimage der Mittelwert berechnet, mit dessen Hilfe globale Beleuchtungsunterschiede detektiert und kompensiert werden (siehe Gleichung 11).

$$\mu(m,t) = \frac{1}{J \cdot I} \sum_{j=-J/2}^{J/2} \sum_{i=-I/2}^{I/2} I(p_x(m)+j, p_y(m)+i, t) \quad (11)$$

Dabei sei J und I die Größe der einzelnen Subimages. Die globale Beleuchtungsänderung $\Delta(t)$ kann nun aus den einzelnen, lokalen Beleuchtungsänderungen $\delta(m,t)$ mittels Median berechnet werden:

$$\Delta(t) = \text{median}_m \delta(m,t) \quad (12)$$

mit

$$\delta(m,t) = \mu(m,t) - \mu(m,t-1) \quad (13)$$

Da kleine Beleuchtungsänderungen vom Hintergrundmodell adaptiert werden, wird ein Schwellwert verwendet, bei dessen Überschreitung erst das Hintergrundmodell angepasst wird. Dies vermeidet ein zu häufiges Update des Hintergrundmodells.

$$\Delta(t) = \begin{cases} 0, & \text{if } \Delta(t) < th_{\Delta} \\ \Delta(t), & \text{if } \Delta(t) \geq th_{\Delta} \end{cases} \quad (14)$$

Gleichung 2 zur Bestimmung des Hintergrundes (Systemzustand) wird nun so modifiziert, um Beleuchtungsänderungen

zu kompensieren. Mit dieser Modifikation können sowohl schnelle als auch langsame Beleuchtungsunterschiede detektiert und kompensiert werden.

$$\begin{bmatrix} \tilde{I}_{x,y}(t) \\ \tilde{I}_{x,y}(t) \end{bmatrix} = S \cdot \begin{bmatrix} \hat{I}_{x,y}(t-1) \\ \hat{I}_{x,y}(t-1) \end{bmatrix} + \begin{bmatrix} \Delta(t) \\ 0 \end{bmatrix} \quad (15)$$

G. Parallelisierung der Objektextraktion

Alle oben beschriebenen Techniken zur Extraktion von Vordergrundobjekten und zur Schattenelimination arbeiten pixelunabhängig und können daher parallel, z.B. auf einem Multi-Core System bearbeitet werden. Dies dient der Verkürzung von Rechenzeiten für zeitkritische Task wie Detektion von Verkehrsteilnehmern. Dies ist wichtig, da der Algorithmus in Realzeit arbeiten muss, um eine Gefährdung rechtzeitig zu erkennen, damit das Türsteuersystem Türbewegungen rechtzeitig anhalten kann. Dazu wird das Originalbild des Kamerasystems in n -Subimages unterteilt und die gleiche Anzahl an Threads generiert, die unabhängig auf einer Multi-Core Plattform arbeiten können. Das Ergebnis der Objekterkennung wird anschließend für ein weitere Klassifikationsalgorithmen oder zur nachfolgenden Bestimmung des Gefährdungsgrades zu einem Bild zusammengeführt und gegebenenfalls interpoliert. Abbildung 9 zeigt das verwendete Parallelisierungskonzept.

H. Bestimmung des Gefährdungsgrades

Zur Bestimmung des Gefährdungsgrades wird sowohl die aktuelle Position (x_n, y_n) des Fahrzeuges in einem der drei Bildbereiche als auch dessen geschätzte Bewegungstrajektorie bestimmt. Ähnlich wie die in Kapitel II-B dargestellte Einteilung der Bereiche handelt es sich aufgrund des Panoramabildes nicht um Geraden, sondern um kurvenförmige Trajektorien. Diese werden mithilfe eines Polynoms 2. Ordnung aus den erfassten Positionsdaten eines Objektes über der Zeit berechnet und die Lage des Scheitelpunktes bestimmt. Aufgrund der Lage des Scheitelpunktes wird der Gefährdungsgrad bestimmt: Befindet sich der Scheitelpunkt der Trajektorie innerhalb einer Zone, wird sich das Objekt mit hoher Wahrscheinlichkeit auch durch diese Zone bewegen. Sowohl die Analyse des Gefährdungsgrades anhand aktueller Positionsdaten als auch anhand von Bewegungstrajektorien ist wichtig, da eine Einschätzung des Gefährdungsgrades allein durch Trajektorien nicht ausreicht. Als Beispiel sei ein detektiertes Fahrzeug genannt, welches sich aus einer gefährdeten Zone herausbewegt und aufgrund der Bewegungsrichtung als ungefährlich eingeschätzt wird. Während es sich aber aus der gefährdeten Zone bewegt, kann es immer noch zu einer Kollision mit der Türe kommen, was z.B. aufgrund der aktuellen Position des Objektes erkannt werden kann.

III. ERGEBNISSE

Die vorgestellte Methode zur Extraktion von Objekten und Einschätzung deren Gefährdungsgrades wurde anhand komplexer Szenarien mit schwachen und starken Schatten, unter veränderlichen Lichtbedingungen, mit kleinen und mit großen Objekten sowie mit kleinen Objekten, die nur einen geringen

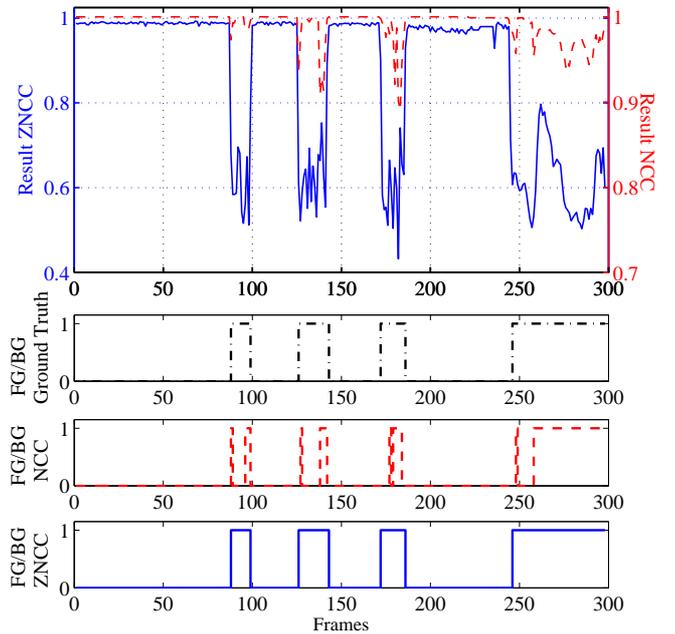


Abbildung 10. Oben: Charakteristik des NCC- und ZNCC-Verlaufs für ein Bildpixel. Mitte und Unten: Klassifikation von Schattenpixeln mittels NCC ($FG/BG = 1/0$) und Verbesserung des Detektionsergebnisses. Mit Hilfe der ZNCC kann besser zwischen Schatten und Vordergrund unterschieden werden.

Unterschied zum Hintergrundmodell aufweisen, getestet, verifiziert und evaluiert. Dazu wurde das in Kapitel I vorgestellte Konzept der in den Außenspiegel integrierten omnidirektionalen Kamera verwendet.

A. Detektion von Schattenpixeln

Um Objekte mit geringem Unterschied zwischen Vordergrund und Hintergrund zu detektieren, muss der Schwellwert th_{bg} (siehe Gleichung 4) sehr klein sein. Ein guter, aus Experimenten ermittelter Wert ist $th_{bg} \geq 7$, der gerade die Detektion von diesen kleinen Objekten ermöglicht. Dieser kleine Wert führt aber auch dazu, dass viele Schattenpixel oder Bildrauschen als Vordergrund detektiert werden. Abbildung 10 zeigt die Ergebnisse der Vordergrunddetektion für ein Pixel über der Zeit, wobei Schatten unter Verwendung der NCC und ZNCC eliminiert werden. Die normierte Kreuzkorrelation ist nützlich zur initialen Detektion von Schattenpixeln, doch gültige Vordergrundpixel werden fälschlicherweise als Schattenpixel erkannt (siehe Abbildung 10, BG/FG NCC). Werden zusätzlich Texturen mittels der ZNCC berücksichtigt, können fälschlicherweise als Schattenpixel klassifizierte Pixel als Vordergrund erkannt werden (Abbildung 10, BG/FG ZNCC).

B. Beleuchtungsänderungen und Hintergrundadaptation

Abbildung 11 zeigt die Ergebnisse von Experimenten mit verschiedenen Beleuchtungsänderungen. Der Ground-Truth Kurvenverlauf des Hintergrundes ist in Abbildung 11(a) dargestellt, während verschiedene Kurvenverläufe des Hintergrundes

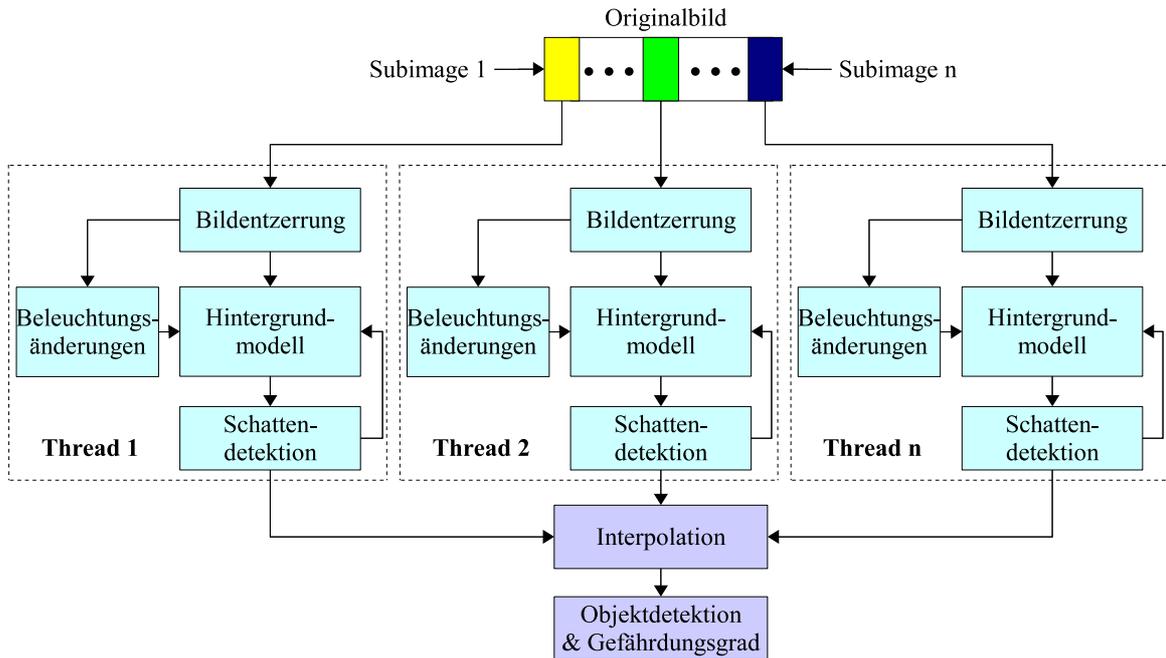


Abbildung 9. Parallelisierung des Objekterkennungsalgorithmus.

des unter Einfluss verschiedener Beleuchtungsänderungen in Abbildung 11(b) und in Abbildung 11(c) dargestellt sind. Das Kalman-basierte Hintergrundmodell kann langsame Beleuchtungsänderungen adaptieren, auch wenn der Hintergrund kurzzeitig durch Vordergrund verdeckt wird und die Beleuchtungsänderungen nicht zu schnell sind (siehe Abbildung 11(c), Bilder 0 – 100). Schnelle Beleuchtungsänderungen, die größer als th_{bg} sind, erzeugen allerdings falsche Vordergrundinformationen (siehe Bilder (280 - 310) und (380 - 400)). Abbildung 11(d) zeigt, dass Beleuchtungsänderungen erkannt und erfolgreich im Modell kompensiert werden können (siehe Gleichung (15)). Zur Bestimmung der dafür benötigten optimalen Anzahl an Suchfenstern zur Detektion von globalen Beleuchtungsunterschieden wurden ebenfalls Experimente durchgeführt. Die Anzahl an Suchfenstern muss so gewählt werden, dass der Einfluss von Beleuchtungsunterschieden hervorgerufen durch große Objekte (LKWs) reduziert wird (siehe Gleichung (12)). Dazu wurde ein künstlich veränderliches Beleuchtungsprofil generiert, das mit Hilfe verschiedener Anzahl an Suchfenstern zu tracken versucht wurde. Experimente zeigten, dass mindestens 60 Suchfenster notwendig sind, um das Beleuchtungsprofil erfolgreich zu tracken. Eine geringere Anzahl an Suchfenstern erhöht den Einfluss von größeren Objekten auf die Detektion von Beleuchtungsänderungen. Der Einfluss von Vordergrundobjekten wird bei einer Anzahl von 90 Suchfenstern (NoW) nahezu vollständig kompensiert. Die durchgeführten Experimente durch Tracking des künstlich generierten Lichtprofils mit unterschiedlicher Anzahl an Suchfenstern zeigt Abbildung 12. Weitere Messung zeigen, dass ein Suchfenster nicht weniger als (15×15) Pixel besitzen sollte. Bei kleineren Suchfenstern steigt der Einfluss von

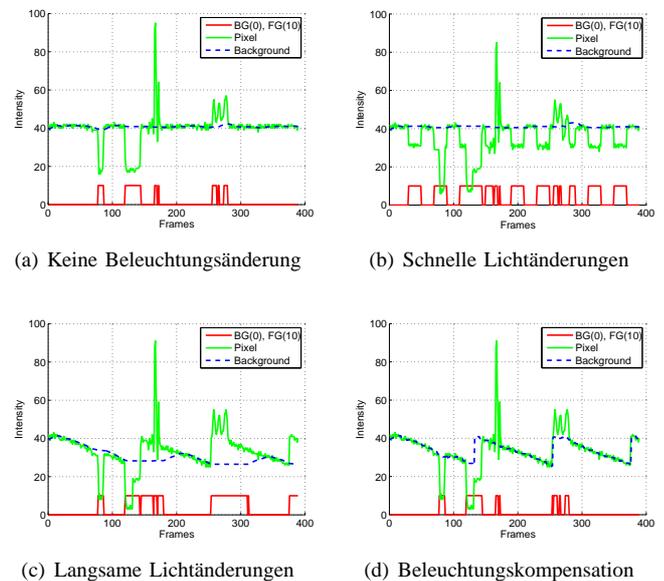
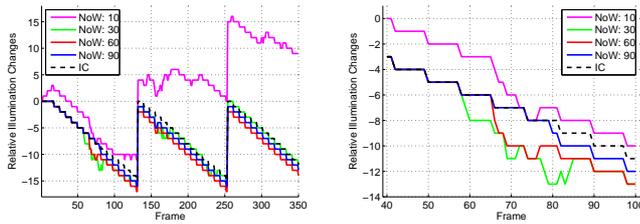


Abbildung 11. (a) Pixelverlauf und detektiert Vordergrund eines Szenarios ohne Beleuchtungsänderungen (Referenz). Als Vordergrund detektierte Pixel hervorgerufen durch schnelle (b) und langsame (c) Beleuchtungsänderungen. (d) Adaption von schnellen und langsamen Beleuchtungsänderungen.

Bildrauschen signifikant. Ein Offset zwischen künstlichem Lichtprofil und geschätztem Lichtprofil von 2 ist tolerierbar und wird automatisch von Hintergrundmodell adaptiert.

C. Validierung detektierter Vordergrundpixel

Nicht alle detektierten Vordergrundpixel müssen gültige Vordergrund (true positives = t.p.), sondern auch Schattenpixel



(a) Beleuchtungsprofil-Änderungen und (b) Bilder 40 - 100

Abbildung 12. Je mehr Suchfenster (NoW) zur Detektion von Beleuchtungsunterschieden verwendet werden, desto besser werden Beleuchtungsänderungen erfasst. Ein gutes Detektionsergebnis erzielt eine Anzahl zwischen 60 - 90 Suchfenster.

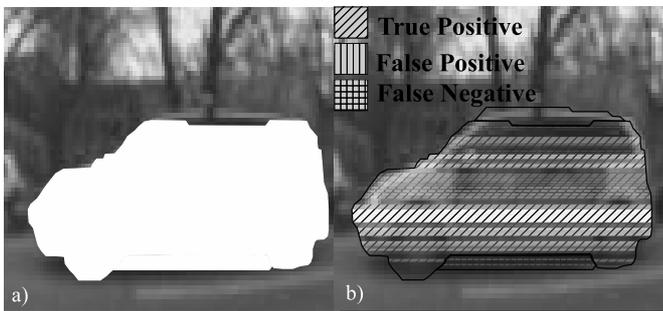


Abbildung 13. a) Detektierte Pixel eines Vordergrundobjektes. b) Einige Pixel werden fälschlicherweise als Vordergrund und Vordergrundpixel als Hintergrund erkannt. Weiterhin gibt gültige Vordergrundpixel, die nicht als Vordergrund angezeigt werden (false negatives).

können fälschlicherweise als Vordergrund (false positives = f.p.) erkannt worden sein. Auf der anderen Seite können auch Vordergrundpixel, die nur eine geringe Differenz zum entsprechenden Hintergrundpixel haben, als Hintergrund klassifiziert werden. (false negatives = f.n.). Abbildung 13 zeigt ein typisches Straßenszenario sowohl mit true- und false positives als auch false negatives. Die Methode zur Detektion von Objekten wurden unter verschiedenen Bedingungen wie diffuse Beleuchtung, direktes Sonnenlicht und Innenraumbeleuchtung evaluiert. Dafür wurden die gemessenen Daten mit den Ergebnissen verglichen, die einer perfekten Detektion (Ground Truth) entsprechen würden. Die dabei ermittelten Ergebnisse zeigt Tabelle II, wobei die angegebenen Ergebnisse auf 300 Testbildern basieren. Gute Detektionsraten konnten bei großen

Szenario	Obj. Größe	t.p.	f.n.	f.p.
Diffuses Licht	klein	85%	15%	2%
	groß	95%	5%	3%
Sonnenlicht	klein	76%	24%	7%
	groß	93%	7%	10%
Innenraum	klein	90%	10%	1%
	groß	97%	3%	4%

Tabelle II
ÜBERBLICK DETEKTIONSERGEBNISSE.

Vordergrundobjekten in allen getesteten Szenarien erreicht

werden. Allerdings führen falsch erkannte Vordergrundpixel bei kleineren Objekten zu einer schlechteren Detektionsrate als die gleiche Anzahl an Pixel in großen Objekten. Weiterhin werden Schattenpixel in Szenarien mit direkter Sonneneinstrahlung leichter als Vordergrund klassifiziert.

D. Rechenzeit und Parallelisierung

Um die benötigten Rechenzeiten zu bestimmen, wurde Experimente anhand eines komplexen Innenraumszenarios durchgeführt. Dieses Szenario zeichnete sich durch drei laufende Personen, schwache und starke Schatten sowie durch veränderliche Lichtverhältnisse aus. Dafür wurden verschiedene Datensets von jeweils 400 Bilder aufgenommen und die Ausführungszeiten berechnet. Tabelle III zeigt die gemittelten Ausführungszeiten sowie deren Standardabweichung für Bildentzerrung, Hintergrundbestimmung, Schattendetektion und Beleuchtungsänderungen unter Verwendung einer CPU einer 2.3 GHz AMD Phenom 9650 Quad-Core CPU. Wie in Ka-

	mittlere Zeit	Std. Abw.	Bildgröße
Entzerrung	≈ 3.6 ms	0.7 ms	640 × 480
Hintergrund	≈ 30.1 ms	1.2 ms	480 × 204
Schatten	≈ 24.1 ms	2.3 ms	''
Beleuchtung	≈ 10.4 ms	1.6 ms	''
Total Time	68.2 ms		

Tabelle III
BERECHNUNGSZEITEN FÜR DIE OBJEKTEXTRAKTION AUF EINEM 2.3 GHz AMD PHENOM 9650 QUAD-CORE FÜR DEN NICHT PARALLELISIERTEN ALGORITHMUS.

pitel II-G beschrieben, wurde der vorgestellte Algorithmus parallelisiert und kann parallel auf einer Quad-Core CPU ausgeführt werden. Dabei wird das Bild in n -Subimages unterteilt und die gleiche Anzahl an unabhängigen Threads erzeugt. Die dabei gemessenen Ausführungszeiten sind in Tabelle IV dargestellt. Während die Gesamtausführungszeiten sinkt, steigt hingegen die Ausführungszeiten für Interpolation mit der Anzahl erzeugter Threads an. Zur Einschätzung des

	1 CPU	2 CPU	4 CPU
Entzerrung	≈ 3.6 ms	≈ 1.9 ms	≈ 1.0 ms
Hintergrund	≈ 30.1 ms	≈ 15.1 ms	≈ 7.5 ms
Schatten	≈ 24.1 ms	≈ 12.2 ms	≈ 6.3 ms
Beleuchtung	≈ 10.4 ms	≈ 5.6 ms	≈ 2.8 ms
Interpolation	0.0 ms	≈ 2.0 ms	≈ 2.2 ms
Total Time	68.2 ms	36.8 ms	19.8 ms

Tabelle IV
DIESE TABELLE ZEIGT DIE BERECHNUNGSZEITEN FÜR DEN OBJEKTEXTRAKTION AUF EINE 2.3 GHz AMD PHENOM 9650 QUAD-CORE FÜR DEN NICHT PARALLELISIERTEN ALGORITHMUS.

Gefährdungsgrades mittels Trajektorien und Objektpositionen werden für eine Anzahl von bis zu 20 Objekten lediglich eine Rechenzeit von ≈ 4.3ms benötigt. Aus diesem Grund ist eine weitere Threadgenerierung nicht erforderlich. Abbildung 14 zeigt die einzelnen Ausführungszeiten für jeden Extraktionsschritt sowie die insgesamt benötigte Rechenzeit für mehr als

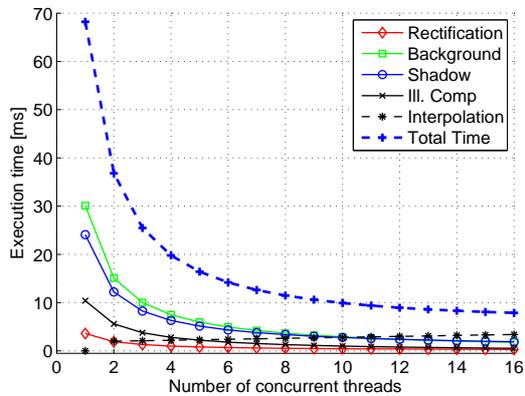


Abbildung 14. Schätzung der Ausführungszeit für mehr als 4 CPUs.

4 CPUs. Diese Abschätzung basiert auf den Messergebnissen zur Laufzeitbestimmung für 1, 2, 3 und 4 CPUs und wurde basierend darauf für bis zu 16-CPUs extrapoliert. Abbildung 14 zeigt auch, dass sich keine Verbesserung für mehr als 16-CPUs erreichen lässt. Weiter wurden die vorgestellten Algorithmen evaluiert und mit anderen Methoden zur Objektextraktion verglichen (siehe Kapitel Stand der Technik I-A). Dazu wurden verschiedene Szenarien mit bis zu 500 Bildern evaluiert. Diese Szenarien waren unterschiedlich texturiert, hatten sowohl schwache als auch starke Schatten und unterschiedliche Lichtverhältnisse. So konnte gezeigt werden, dass die von Ridder *et al.* vorgeschlagene Methode den Hintergrund gut modelliert, Schatten aber in vielen Fällen nicht detektiert werden. Dahingegen werden Schatten bei der von Jaques Jr. *et al.* vorgeschlagenen Methode relativ gut kompensiert. Allerdings kann das in deren Beitrag präsentierte Hintergrundmodell nicht aktualisiert werden, so dass Veränderungen nicht adaptiert werden können. Die Kombination beider Methoden sowie die Erweiterung zur verbesserten Schattendetektion und eine aktive Beleuchtungskompensation erlaubt eine gute Detektion von Verkehrsteilnehmern für unsere Anwendung. Die dabei detektierten Verkehrsteilnehmer sind in Abbildung 15 dargestellt. Nach erfolgreicher und präziser Extraktion von Verkehrsteilnehmern wäre der nächste Schritt die Identifikation von Verkehrsteilnehmern, um deren Gefährdungsgrad noch präziser bestimmen zu können.

IV. ZUSAMMENFASSUNG

In diesem Beitrag wurde ein robuster Algorithmus zur Detektion von Verkehrsteilnehmern in Videodaten vorgestellt. Der Gefährdungsgrad der dabei detektierten Verkehrsteilnehmer wurde anhand der Objektposition und deren geschätzter Bewegungstrajektorie bestimmt. Diese Information dient als Input für eine Intelligente Fahrzeugtüre, welche bei einer möglichen Kollision beim Türöffnen die Türe blockiert und/oder den Benutzer vor einer Kollision warnt. In einem ersten Schritt wurde dazu ein Hintergrundbild von einer Parksituation aufgenommen, welches die Grundlage einer

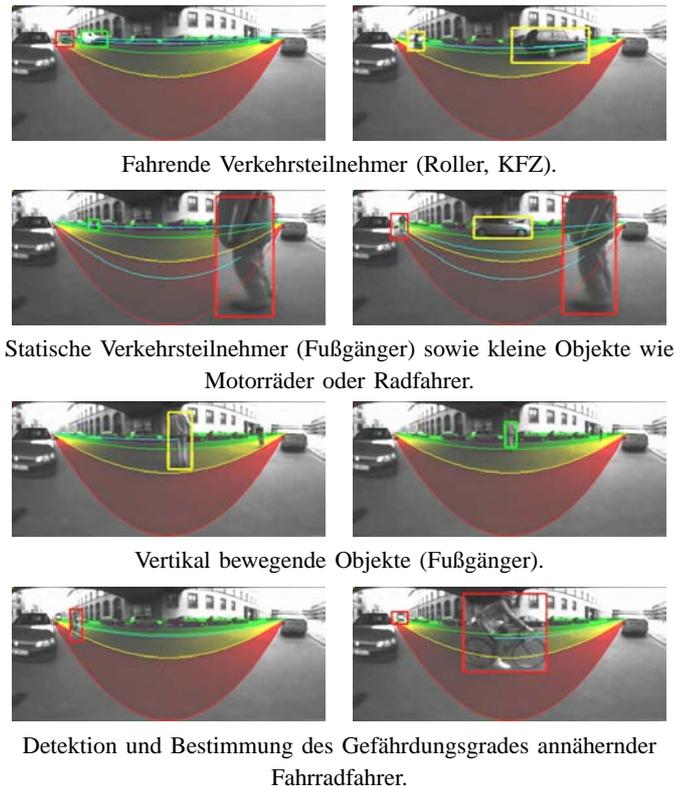


Abbildung 15. Evaluierung vorgestellter Methode mit verschiedenen Objekten in einem Referenzszenario.

Objektdetektion ist. Während der Lernphase können sich auch bewegende Vordergrundobjekte im Bild befinden. Abschließend werden mittels des aufgenommenen Hintergrundbildes und Kalman-Filterung Vordergrundobjekte extrahiert. Dabei als Vordergrund detektierte Schattenpixel werden mit Hilfe der normierten Kreuzkorrelation (NCC) initial detektiert und mit Hilfe der mittelwertbefreiten normierten Kreuzkorrelation (ZNCC) verbessert.

Um den Einfluss von Beleuchtungsänderungen zu kompensieren, werden Eingangsbilder in kleine Subimages unterteilt und deren Beleuchtungsänderung berechnet. Basierend auf diesen Daten wird abschließend das Hintergrundmodell angepasst, um Objekte auch bei schnellen Änderungen zu detektieren. Der Gefährdungsgrad derart detektierter Objekte wird zuletzt anhand deren aktuellen Position sowie deren Objektbewegungstrajektorie geschätzt.

Der vorgestellte Algorithmus wurde erfolgreich implementiert und kann selbst kleine und schnelle Objekte wie Motorräder oder Fahrradfahrer detektieren. Um den Algorithmus für ein Vielzahl an Parksituationen robust zu gestalten, sind weitere Studien für eine optimale Parameterwahl der vielfältigen Parameter notwendig. Eine weitere Forschungsrichtung wäre die Portierung des Algorithmus auf ein Embedded System (z.B. FPGA), so dass die Ausführungszeiten weiter verkürzt und damit die Reaktionszeit der intelligenten Türe weiter verbessert werden kann.

LITERATUR

- [1] European Commission, "Europeans information society, driver-assistance systems," *t*, 2009. [Online]. Available: http://ec.europa.eu/information_society/activities/intelligentcar/technologies/das/index_en.htm
- [2] M. Strolz, Q. Mühlbauer, C. Scharfenberger, G. Färber and M. Buss, "Towards a generic control system for actuated car doors with arbitrary degrees of freedom," *In Proceedings of IEEE Intelligent Vehicles Symposium 2008, Eindhoven, The Netherlands*, pp. 391–397, 2008.
- [3] M. Okutomi and T. Kanade, "A multiple-baseline stereo," 1993.
- [4] J. Suhr, H. Jung, K. Bae, and J. Kim, "Vorrichtung und Verfahren zum monokularen Motion-Stereo-basierten Detektieren von freien Parkplätzen," 2010, german Patent DE102009012435A1.
- [5] R. R. Lawrence, "A tutorial on hidden markov models and selected applications in speech recognition." *Proceedings of the IEEE*, vol. 77, no. 2, 1989.
- [6] R. Brunelli, "Template matching techniques in computer vision: Theory and practice," *Wiley, ISBN 978-0-470-51706-2 ([1] TM book)*, 2009.
- [7] O. Achler, M.M. Trivedi, "Real-time traffic flow analysis using omnidirectional video network and flatplane transformation," *Computer Vision and Robotics Research Laboratory - University of California, San Diego*, 2002.
- [8] T. Ghandi, M.M. Trivedi, "Motion based vehicle surround analysis using an omni-directional camera," *In IEEE Intelligent Vehicle Symposium*, 2004.
- [9] I. Haritaoglu, D. Harwood, L.S. Davis, "W4: Who? when? where? what? a real-time system for detecting and tracking people." *In Proceedings of the third IEEE International Conference on Automatic Face and Gesture Recognition.*, pp. 222–227, 1998.
- [10] N. Friedman, S. Russel, "Image segmentation in video sequences: A probabilistic approach." *In Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, 1997.
- [11] K.-P. Karman, A. Brandt von, "Moving object recognition using an adaptive background memory," *In V. Cappellini (ed.), Time-varying Image Processing and Moving Object Recognition, 2, Elsevier Publishers B.V., Amsterdam, The Netherlands*, pp. 297–307, 1990.
- [12] A. Elgammal, R. Duraiswami, D. Harwood, L. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, 2002.
- [13] D. Stenger, V. Ramesh, N. Paragios, F. Coetzer, J.M. Buhmann, "Topology free hidden markov models: application to background modeling." *In Proceedings of the International Conference on Computer Vision*, 2001.
- [14] Z. Zivkovic, F.V. Heijden, "Efficient adaptive density estimation per image pixel for task of background subtraction." *In Pattern Recognition Letters.*, pp. 773–780, 2006.
- [15] Z. Tang, Z. Miao, "Fast background subtraction and shadow elimination using improved gaussian mixture model," *Proceedings of the IEEE International Workshop on Haptic, Audio and Visual Environments and Games, HAVE 2007*, pp. 38–41, 2007.
- [16] H. Ardoe, R. Berthilsson, "Adaptive background estimation using intensity independent features," *In Proceedings of the International Conference on the 17th British Machine Vision Conference*, 2006.
- [17] H. Bhaskar, L. Mihaylova, S. Maskell, "Background modeling using adaptive cluster density estimation for automatic human detection," *In Lecture Notes in Informatics, GI Jahrestagung (2)*, pp. 130–134, 2007.
- [18] J. Zhong, S. Sclaroff, "Segmenting foreground objects from a dynamic textured background via a robust kalman filter," *In Proceedings of the International Conference on Computer Vision, October*, 2003.
- [19] C. Ridder, O. Munkelt, H. Kirchner, "Adaptive background estimation and foreground detection using kalman-filtering," *In Proceedings on International Conference on ICRAM*, pp. 193–199, 1995.
- [20] J. Jacques, C. Jung, S. Musse, "Background subtraction and shadow detection in grayscale video sequences," *In Proceedings of the 18th Brazilian Symposium on Computer Graphics and Image*, pp. 189–196, 2005.
- [21] D. Farin, P. de With, and W. Effelsberg, "Robust background estimation for complex video sequences," 2003.
- [22] M. Massey and W. Bender, "Salients stills: Process and practice," 1993.
- [23] A. B. R. Alvarez, J.M. Lopez, "Illumination-invariant model-based road segmentation." 2008.
- [24] R. Adams and L. Bischof, "Seeded region growing: an extensive and comparative study," vol. 26, no. 8, 2005.