
Zusammenfassung

Das Ziel der vorliegenden Arbeit ist es, einen Prototypen für ein Computerprogramm zu entwickeln, welches Fahrzeugtrajektorien aus vorab aufgezeichneten Videos bestimmen kann. Dabei wird der Verkehr auf einem Straßenabschnitt oder an einem Knotenpunkt mit einer Videokamera aufgezeichnet, welche fest auf einem Stativ o.ä. in der Nähe der Straße aufgestellt ist. Die Auswertung des Videomaterials findet im Anschluss mit einem Computer statt.

Geschwindigkeitstrajektorien von Fahrzeugen können als Grundlage für verschiedene Anwendungen dienen:

- Die Analyse von Unfallstrecken oder punktuellen Unfallschwerpunkten geschieht häufig nur basierend auf der Auswertung einzelner Unfälle. Um den möglichen Zusammenhang zwischen der Trassierung oder anderen Eigenschaften der Straße und den Unfällen zu untersuchen, kann eine große Menge von Geschwindigkeitstrajektorien hilfreich sein, um statistisch abgesicherte Aussagen über das Fahrerverhalten in diesem Bereich zu treffen, wie z.B. spätes oder sehr starkes Abbremsen.
- Eine große Menge von Trajektorien kann für weitere wissenschaftliche Anwendungen hilfreich sein, wie z.B. die Untersuchung, wie sich Geschwindigkeitsbeschränkungen auf das tatsächliche Fahrverhalten auswirken, bspw. in Verbindung mit einer fahrerpsychologischen Untersuchung.
- Verkehrszählungen und Fahrzeug-Klassifizierungen werden noch immer manuell durchgeführt, falls keine fest installierten Geräte zur Verkehrsüberwachung vorhanden sind. Mit einer videogestützten Auswertung kann der Arbeitsaufwand signifikant reduziert werden, da die Aufzeichnung mit einer einzelnen Videokamera genügt, um auf mehreren Fahrstreifen gleichzeitig die Fahrzeuge zu zählen und grob zu klassifizieren.

Das gesamte Projekt kann in drei wichtige Module unterteilt werden:

1. Bei der *Kamera-Kalibrierung* werden die geometrischen Eigenschaften der zentralperspektivischen Abbildung zwischen dem zweidimensionalen Bild-Koordinatensystem und dem dreidimensionalen Welt-Koordinatensystem der realen Umgebung bestimmt (der Name Welt-Koordinatensystem ist gängig, obwohl eine Georeferenzierung nicht stattfinden muss). Sind die Eigenschaften dieser Abbildung bekannt, so können Bildpunkte des Bild-Koordinatensystems in dreidimensionale Welt-Koordinaten zurück transformiert werden, so dass im Bild Positionen und damit auch Geschwindigkeiten und Beschleunigungen gemessen werden können. Diese Rückprojektion kann nur unter zusätzlichen Annahmen *eindeutig* stattfinden, falls kein stereoskopisches Verfahren zum Einsatz kommt.
2. Die Methoden der *digitalen Bildverarbeitung* machen den eigentlichen Kern des Projekts aus. Die wichtigsten Funktionen sind die *Fahrzeugdetektion* und das *Fahrzeugtracking*.
3. Die im Anschluss erfolgende *Offline-Auswertung* kombiniert die beim Fahrzeugtracking gefundenen Bild-Koordinaten und die mit der Kamera-Kalibrierung ermittelten Eigenschaften der Koordinatentransformation. Sind die Welt-Koordinaten und die zugehörigen Zeitpunkte der getrackten Fahrzeuge bekannt, so können die Geschwindigkeiten und Beschleunigungen unmittelbar durch numerische Differentiation berechnet werden. In Kombination mit den bekannten Zeitpunkten können weitere, abhängige Größen abgeleitet werden, wie z.B. Zeitlückenverteilungen oder Verkehrsstärken.

Die Methoden der Kamera-Kalibrierung können in drei Kategorien eingeteilt werden:

1. Die *parametrische Koordinatentransformation* kann verwendet werden, wenn sich alle zu projizierenden Punkte (annähernd) in einer Ebene befinden, oder die Höhen der Punkte über einer definierten Bezugsebene (z.B. der Fahrbahnebene) bekannt sind. Diese Annahme trifft für ebene Straßenabschnitte und Knotenpunkte zu. Die Parameter der perspektivischen Abbildung, wie z.B. Position und Drehwinkel der Kamera im Raum, können dabei z.B. mit Hilfe von Passpunkten oder Fluchtpunkten, welche vom Benutzer anzugeben sind, berechnet werden.

-
2. Die *stereoskopische Koordinatentransformation* verwendet die redundante Information von mindestens zwei Videokameras, die denselben Abschnitt der Straße filmen. Mit Hilfe der mindestens vier Bild-Koordinaten können unmittelbar dreidimensionale Welt-Koordinaten berechnet werden, falls die relative Position und Ausrichtung der beiden Kameras bekannt sind. Dieser Ansatz wird in der vorliegenden Arbeit nur knapp untersucht und nicht implementiert.
 3. Es wird im Rahmen der Arbeit eine einfache Methode zur *nicht-parametrischen Koordinatentransformation* vorgestellt, welche kontinuierliche Richtungen und Maßstäbe verwendet. Diese Informationen können z.B. mit Hilfe einer Messlatte bestimmt werden, welche bei laufender Kamera entlang des Straßenabschnitts bewegt wird. Diese sehr einfache Technik stellte sich als sehr effektiv heraus, da alle Parameter der geometrischen Abbildung, also Kameraposition und -ausrichtung im Raum, Brennweite, Bildhauptpunkt, Verzeichnung und die dreidimensionale Form der Trasse, *implizit* in der gemessenen Information enthalten sind. Die Forderung eines in einer Ebene liegenden Straßenabschnitts besteht hier nicht. Zwischen den Punkten, an denen der Maßstab und die Richtung bestimmt wurden, können die Werte interpoliert werden.

Im Rahmen der Arbeit wurden Methoden der ersten und der dritten Kategorie implementiert und getestet. Die Qualität der Methoden der Kamera-Kalibrierung konnte dabei teilweise mit Hilfe einer einfachen 3D-Simulation untersucht werden.

Durch den Einfluss von Wind kann es passieren, dass die Videokamera in Schwingungen versetzt wird oder sogar nachhaltig ihre ursprünglichen Position verliert. Für die Projektion kann dies insbesondere bei Punkten, die sich nah am Horizont befinden, erhebliche Auswirkungen haben. Es wird ein Algorithmus zur subpixelgenauen Bildregistrierung vorgestellt, der imstande ist, diese Wackelbewegungen zu erkennen und heraus zu rechnen. Während eine bleibende Fehlstellung der Kamera zuverlässig und genau erkannt werden kann, können geometrische Verzerrungen, die durch das zeilenweise Auslesen des Bildsensors während einer Wackelbewegung auftreten, kaum reduziert werden. Die ermittelten Trajektorien sind dann deutlich verrauschter.

Die Fahrzeugdetektion ist ein zentraler Teil des Projektes. Ein Hintergrundmodell wird automatisch gelernt und fortlaufend aktualisiert, so dass sich bewegende Objekte auf sehr einfache Weise im Bild gefunden werden können. Hierzu wurde eine sehr effiziente Approximierung eines Median-Filters verwendet, welche die Fahrzeugerkennung im Video in Echtzeit zulässt. Wurde ein sich bewegendes Objekt gefunden, so muss dieses verschiedene Eigenschaften erfüllen, um als Fahrzeug erkannt zu werden. Hierzu gehören vom Benutzer definierbare Eigenschaften wie Mindestgröße, Mindestgeschwindigkeit und die Bewegungsrichtung. Eine Prüfung der äußeren Form des Fahrzeugs findet nicht statt, da die Fahrzeuge unter Umständen nur wenige Pixel groß im Bild erscheinen. Die im Rahmen dieser Arbeit implementierte Fahrzeugerkennung stellte sich als zuverlässig heraus: in den ausgewerteten Testaufnahmen mit über 1500 Fahrzeugen wurden weniger als 1 Prozent der Fahrzeuge schlecht erkannt (z.B. wurde nur ein Teil des Fahrzeugs erkannt, oder das Fahrzeug war mit einem anderen Objekt verschmolzen und dadurch deutlich zu groß). Kein Fahrzeug wurde vom Algorithmus gar nicht erkannt, außer sehr kleine Fahrzeuge wie z.B. Fahrräder, die eine der Mindestanforderungen nicht erfüllten.

Der Tracking-Algorithmus stellt den Kern des gesamten Auswerteprozesses dar. Um Trajektorien mit guten Genauigkeiten zu erhalten, wird ein subpixelgenauer Tracking-Algorithmus vorgestellt. Die Idee hinter diesem Algorithmus ist es, die exakte Auswertung der zweidimensionalen, normalisierten Kreuzkorrelationsfunktion in der Nähe des gesuchten lokalen Maximums mit Hilfe einer Approximierung mit einer Flächenfunktion zu beschleunigen. Es konnte gezeigt werden, dass der Fehler dieser Approximierung im Mittel unter 0,01 Pixeln liegt. Um eine Genauigkeit von 0,15 Pixeln zu erzielen, genügt es, die Kreuzkorrelationsfunktion bis auf 0,5 Pixel genau auszuwerten. Da die Rechenintensität der Kreuzkorrelationsfunktion mit der dritten Potenz (in der naiven Implementierung sogar mit der vierten Potenz) der geforderten Genauigkeit ansteigt, betrug die Zeitersparnis durch die Approximierung *über 90 Prozent*. Der mittlere Tracking-Fehler

von etwa 0,15 Pixeln wurde in verschiedenen Experimenten mit guten Umgebungsbedingungen (wenig Wind, kein Niederschlag) ermittelt. Er ist unabhängig vom Bild-Maßstab oder der realen Fahrzeuggröße und wird vor allem durch Beleuchtungseffekte wie Reflexionen, durch die graduelle Veränderung des Aussehens und der Größe der Fahrzeuge im Bild, sowie durch Artefakte der Videokompression verursacht.

Für das Tracking ist eine Strategie wichtig, wie *Verdeckungen* durch statische Objekte oder auch durch andere Fahrzeuge erkannt werden können. Wurde eine Verdeckung erkannt, so muss das Template, das das zuletzt bekannte Erscheinungsbild des Fahrzeugs repräsentiert, beibehalten werden, damit das Fahrzeug später wiedergefunden werden kann. Es wird ein Verfahren vorgestellt, welches den gemessenen Grad der Ähnlichkeit bei der Korrelation für das Erkennen einer Verdeckung verwendet: Wird das zu trackende Fahrzeug plötzlich durch ein anderes Objekt verdeckt, so sinkt der Grad der Ähnlichkeit normalerweise erheblich. Die vorgestellte Methode funktioniert zufriedenstellend bei kurzzeitigen Verdeckungen (z.B. durch Fahrzeuge incl. LKW auf perspektivisch vorgelagerten Fahrstreifen). Problematisch ist es allerdings, Fahrzeuge nach lang andauernden Verdeckungen wieder im Bild zu erkennen, da diese bei Wiedererscheinen häufig eine deutlich andere Gestalt, Helligkeit und Größe im Bild besitzen.

Eine weitere, wichtige Fragestellung ist der Umgang mit Perspektiven, bei denen die Fahrzeuge im Bild deutlich ihre Größe verändern (z.B. wenn eine Straße von einer Brücke aus gefilmt wird, d.h. die Fahrzeuge sich auf die Kamera zu oder von der Kamera weg bewegen). Der Suchraum bei der Suche nach dem lokalen Maximum der Kreuzkorrelationsfunktion wird dann drei- oder mehrdimensional, d.h. die Komplexität wird deutlich erhöht. Auch wenn die Größe des Fahrzeuges selbst nicht von Interesse ist, so ist die Kenntnis derselben von großer Wichtigkeit für das Tracking: Besteht das Referenz-Bild (Template) nur aus einem kleinen Teil des Fahrzeugs, oder beinhaltet das Template sehr viel Umgebung, so wird das Tracking ungenau und instabil, d.h. Fahrzeuge könnten beim Tracking „verloren“ gehen. Abhilfe schafft hier eine Strategie, die sich die Beobachtung zu Nutze macht, dass sich die relativen Größen aller Fahrzeuge im Bild *gleichartig* beim Passieren des Straßenabschnitts verändern. Diese Methode misst die Größen einiger Fahrzeuge im Bild fortlaufend mit Hilfe der Fahrzeugdetektion, die ansonsten nur auf einen kleinen Bild-Bereich zur erstmaligen Erkennung von Fahrzeugen angewendet werden muss. Diese gemessenen, relativen Größenveränderungen werden sodann mit einem Median-Filter und einem kubischen Spline geglättet. Der Suchraum bei der Suche nach dem lokalen Maximum bleibt hierdurch zweidimensional.

Die in Bild-Koordinaten gemessenen Fahrzeugpositionen werden anschließend in dreidimensionale Koordinaten zurück transformiert. Hierfür kommt entweder eine parametrische Koordinatentransformation, oder die Verwendung kontinuierlicher Richtungen und Maßstäbe zum Einsatz, welche im Rahmen der Kamera-Kalibrierung ermittelt wurden. Die Geschwindigkeiten und Beschleunigungen werden durch einfaches bzw. zweifaches Bilden des Differenzenquotienten berechnet. Die Geschwindigkeiten schwanken üblicherweise um wenige km/h um den realen Wert (je nach Maßstab). Die Beschleunigungen hingegen sind deutlich verrauschter, da sich der Fehler durch Bilden des Differenzenquotienten überdimensional fortpflanzt. Um die Verläufe von Geschwindigkeit und Beschleunigung in eine für den Benutzer verständliche Form zu bringen, werden diese mit Hilfe von kubischen Splines geglättet. Der Ruck als dritte Ableitung der gemessenen Positionen liefert in ungefilterter Form kaum nützliche Information. Mit Hilfe einer einfachen Glättung kann er aber durchaus dazu dienen, ungefähre Brems- und Beschleunigungspunkte anzugeben. Diese können für verkehrswissenschaftliche Auswertungen von Interesse sein.

Um die gemessenen Geschwindigkeiten zu überprüfen, wurden einige Testfahrten mit einem mit GPS ausgestatteten Fahrzeug bei laufender Videokamera durchgeführt. Durch die begrenzte zeitliche Auflösung der GPS-Messungen von einer Sekunde konnten die geglätteten Verläufe eher dynamischer Trajektorien nicht sinnvoll überprüft werden. Testfahrten mit annähernd konstanter Geschwindigkeit ergaben Abweichungen zur mit GPS gemessenen Geschwindigkeit von weniger als 1 km/h.

Insgesamt wurden in vier verschiedenen Beispiel-Szenarien über 800 Trajektorien gemessen und ausgewertet.