



Automatisierte Verkehrszeichenkartierung aus mobil erfassten Stereobilddaten unter Verwendung der Tiefeninformation aus Dense-Stereo-Matching

STEFAN CAVEGN & STEPHAN NEBIKER, Muttenz, Schweiz

Keywords: mobile mapping, road signs, depth maps, dense stereo matching

Summary: *Automated 3D Road Sign Mapping with Stereovision-based Mobile Mapping exploiting Depth Information from Dense Stereo Matching.* This paper presents algorithms and investigations on the automated detection, classification and mapping of road signs which systematically exploit depth information from stereo images. This approach was chosen due to recent progress in the development of stereo matching algorithms enabling the generation of accurate and dense depth maps. In comparison to mono imagery-based approaches, depth maps also allow 3D mapping of the objects. This is essential for efficient inventory and for future change detection purposes. Test measurements with the mobile mapping system by the Institute of Geomatics Engineering of the University of Applied Sciences and Arts Northwestern Switzerland demonstrated that the developed algorithms for the automated 3D road sign mapping perform well, even under difficult to poor lighting conditions. Approximately 90 % of the relevant road signs with predominantly red, blue and yellow colours in the standard and small format in Switzerland can be detected, and 85 % can be classified correctly. Furthermore, fully automated mapping with a 3D accuracy of better than 10 cm is possible.

Zusammenfassung: In diesem Beitrag werden Algorithmen und Untersuchungen zur automatischen Detektion, Klassifizierung und Kartierung von Verkehrszeichen vorgestellt, die systematisch die Tiefeninformation aus Stereobildaufnahmen ausnutzen. Dieser Ansatz bietet sich an, da in den letzten Jahren neue Stereo-Algorithmen entwickelt wurden, welche die Generierung von dichten Tiefenkarten aus Stereobilddaten ermöglichen. Im Vergleich zu den auf Monobildern basierenden Verfahren erlauben die Tiefenkarten auch eine genaue 3D-Kartierung der Objekte, was für die effiziente Inventarisierung und für eine zukünftige Veränderungsanalyse wesentlich ist. Testmessungen mit dem mobilen Messfahrzeug des Instituts Vermessung und Geoinformation der Fachhochschule Nordwestschweiz haben gezeigt, dass die entwickelten Algorithmen für die automatisierte Verkehrszeichenkartierung auch bei schlechten Beleuchtungsverhältnissen sehr leistungsfähig sind. Es können etwa 90 % der relevanten Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen im Normal- und Kleinformat in der Schweiz detektiert und 85 % korrekt klassifiziert werden. Zudem ist eine vollautomatische Kartierung mit einer 3D-Genauigkeit von unter 10 cm möglich.

1 Einleitung

Entlang von Straßen in westeuropäischen Ländern sind sehr viele Verkehrszeichen angebracht. In der Schweiz sind es zurzeit etwa 5 Millionen. Vielfach fehlen den Eigentümern digitale Informationen über Position, Zustand sowie Zweckmäßigkeit dieser Verkehrszeichen. Nach Meinung verschiedener Experten sind beispielsweise viele Verkehrszeichen

überflüssig. So wird geschätzt, dass die unnötige Beschilderung in der Schweiz den Bund, die Kantone und die Gemeinden jährlich mehrere Millionen Franken kostet (SCHWEIZER FERNSEHEN 2008).

Zur Analyse und Lösung dieser Probleme und Fragestellungen bietet sich ein digitales Verkehrszeicheninventar an. Um ein solches Inventar anzulegen, wurden die Sachdaten und dokumentarische Bildinformation

der Verkehrszeichen in der Regel aufwändig vor Ort manuell erfasst und die Position mit einem GNSS-Empfänger im Meter- bis Dezimeterbereich bestimmt. In den letzten Jahren erfolgte die Kartierung und Inventarisierung aber auch zunehmend auf Grund von Daten, welche mit mobilen Messfahrzeugen erfasst wurden. Der Vorteil liegt darin, dass die 3D-Objekte sehr effizient kartiert werden können, ohne den Verkehrsfluss zu stören und das Vermessungspersonal zu gefährden. In Belgien konnten Verkehrszeichen über das ganze Landesgebiet anhand von Laserscanning-Daten detektiert werden; die Ableitung der Sachdaten erfolgte aber größtenteils manuell (TRIMBLE 2009). In den Niederlanden wurde

die Verkehrszeichenkartierung auf der Basis von Panoramabildern, welche im Abstand von 5 m vorliegen, manuell vorgenommen (DE WITH et al. 2010). Können Verkehrszeichen aus georeferenzierten Bilddaten größtenteils automatisch extrahiert werden, reduziert sich der manuelle Kartierungsaufwand massiv. Algorithmen für die Verkehrszeichenextraktion aus Monobildern sind in MALDONADO-BASCÓN et al. (2008) beschrieben. In diesem Beitrag werden Algorithmen präsentiert, die es erlauben, Verkehrszeichen aus mobil erfassten Stereobildsequenzen automatisch und benutzerunterstützt zu extrahieren, die 3D-Position im gewünschten geodätischen Referenzrahmen zu bestimmen und diverse Sachdaten aus den

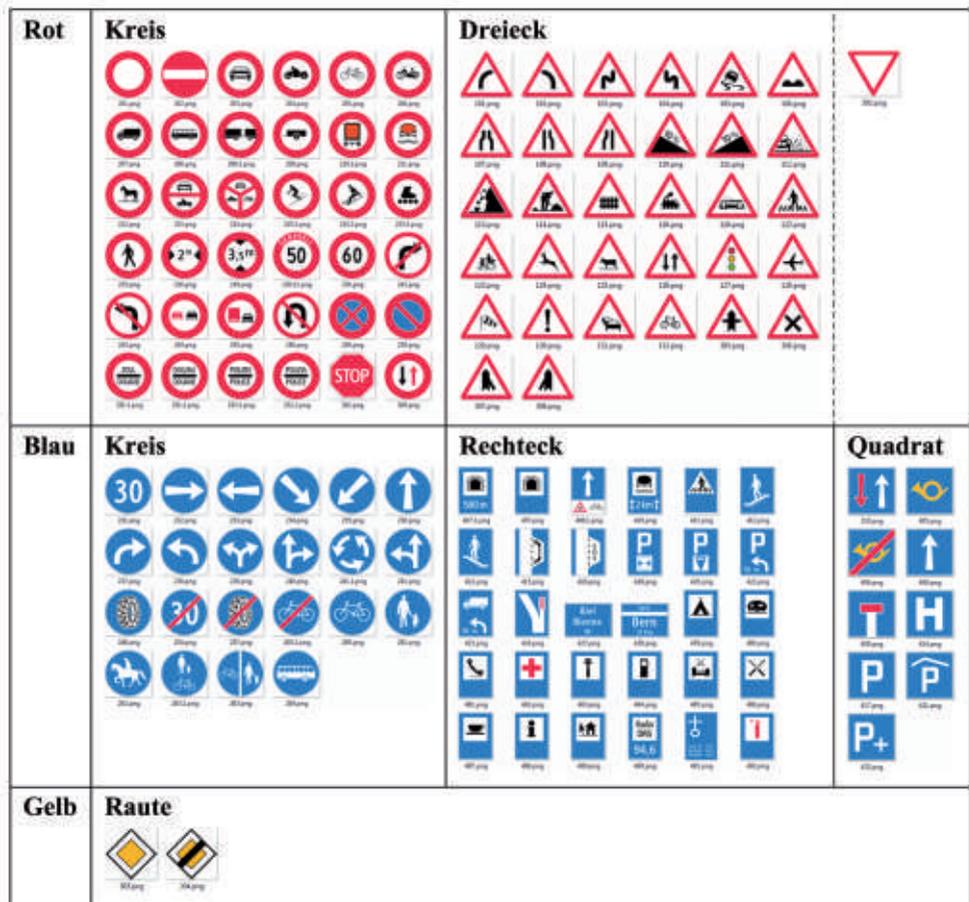


Abb. 1: Verkehrszeichen in der Schweiz, welche mit den entwickelten Algorithmen automatisch detektiert, klassifiziert und kartiert werden können. Gliederung nach dominanter Farbe (Zeilen) und nach Form (Spalten).

Bilddaten automatisch abzuleiten. Diese Algorithmen wurden vorerst auf die Verkehrszeichen in der Schweiz optimiert, welche unterschiedliche Farben und Formen aufweisen sowie je nach Straßentyp in vier verschiedenen Normgrößen vorkommen können (siehe Abb. 1). Sie lassen sich aber auch auf andere Länder übertragen. Da die Anwendung nicht für Fahrerassistenzsysteme oder intelligente autonome Fahrzeuge ausgelegt ist, sondern das Ziel explizit in der Kartierung, Inventarisierung und Bewirtschaftung von Verkehrszeichen liegt, hat die Echtzeitfähigkeit der Algorithmen nicht erste Priorität. Stattdessen liegt der Fokus auf Vollständigkeit, Korrektheit und geometrischer Genauigkeit.

Im folgenden Kapitel werden bestehende Ansätze für die Detektion, Klassifizierung und Kartierung von Verkehrszeichen erläutert. Nach der Vorstellung des Lösungsansatzes in Abschnitt 3 erfolgt die Präsentation der entwickelten Algorithmen und Softwaremodule in Abschnitt 4. Die Untersuchungen und Resultate sind in Abschnitt 5 beschrieben. Im letzten Abschnitt wird nach einem Fazit ein Ausblick auf laufende und zukünftige Entwicklungen gegeben.

2 Verwandte Arbeiten

Die meisten Ansätze zur Thematik der bildbasierten Verkehrszeichenextraktion lassen sich in zwei Hauptschritte unterteilen. Zuerst erfolgt die Detektion der Verkehrszeichen in der Bildszene mit dem Ziel, mögliche Kandidaten zu lokalisieren. Um für jeden Kandidaten die korrekte Verkehrszeichenart zu bestimmen, ist danach eine Klassifizierung erforderlich. Einige der maßgebenden existierenden Ansätze für die Detektion und Klassifizierung der Verkehrszeichen sind in diesem Abschnitt dokumentiert. Eine detaillierte Zusammenstellung von unterschiedlichen Methoden mit ihren Merkmalen ist in NGUWI & KOUZANI (2008) gegeben. Ist zusätzlich die absolute Position der detektierten Zeichen von Interesse, fällt als dritter Schritt noch die Kartierung der Verkehrszeichen an.

2.1 Detektion von Verkehrszeichen

Das Ziel der Detektion besteht darin, in einem Kamerabild Verkehrszeichenkandidaten zu finden. Ausgangsbasis ist dabei vielfach die Farbinformation. Wird eine Farbsegmentierung mit Schwellwerten vorgenommen, kann schnell auf die Suchregionen fokussiert werden. Da der RGB-Farbraum empfindlich auf Änderungen der Lichtverhältnisse durch Schatten, Beleuchtungs- und Betrachtungsgeometrie sowie auf starke Reflektionen reagiert, erfolgt die Segmentierung in der Regel im HSV-Farbraum mit den Farbton- und Sättigungs-Komponenten (FLEYEH 2006, MALDONADO-BASCÓN et al. 2008, FU & HUANG 2010). JAU et al. (2008) haben umfangreiche Untersuchungen zu den Unterschieden zwischen dem RGB- und HSV-Farbraum vorgenommen und festgestellt, dass für die Verkehrszeichendetektion die Farbsegmentierung im HSV-Farbraum besser als im RGB-Raum geeignet ist. MADEIRA et al. (2005) verwenden für die Farbsegmentierung die Farbton- und die chromatische RGB-Komponente, welche nach DE LA ESCALERA et al. (2004) berechnet wird. Dies weil die Sättigungs-Komponente im Gegensatz zur chromatischen RGB-Komponente für kleine Werte sehr empfindlich auf Rauschen reagiert (MADEIRA et al. 2005). Nach GARCÍA-GARRIDO et al. (2006) treten Segmentierungsprobleme bei vorwiegend weißen Verkehrszeichen und bei Beleuchtungsunterschieden auf. Formbasierte Methoden zeigen eine größere Robustheit bezüglich variabler Beleuchtungsverhältnisse. Obwohl diese Methoden schnell in der Ausführung und ziemlich einfach zu implementieren sind, produzieren sie viele Kandidatenobjekte, was für die folgenden Schritte mehr Aufwand bedeutet (GÓMEZ-MORENO et al. 2010).

2.2 Klassifizierung von Verkehrszeichen

Bei der Klassifizierung geht es darum, für den beim Detektionsprozess ermittelten Verkehrszeichenkandidaten den richtigen Typ zu bestimmen. Dabei kommen in vielen Fällen neuronale Netzwerke zum Einsatz (DE LA ESCALERA et al. 2003, NGUWI & KOUZANI 2008,

HOSSAIN et al. 2010). Diese erfordern zusätzlich zur Implementierung ein umfassendes Training der Algorithmen, das Eingangsbild muss aber nicht in eine andere Raumrepräsentation transformiert werden (FLEYEH & DOUGHERTY 2005). REN et al. (2009) zeigen, dass der Trainingsprozess im Fall einer umfassenden Verkehrszeichenklassifizierung sehr viele Trainingsbilder mit unterschiedlicher Skalierung, Orientierung und Beleuchtung benötigt. Neuronale Netzwerke werden daher vielfach nur für wenige bestimmte Verkehrszeichen wie Geschwindigkeitsbegrenzungen verwendet. GAO et al. (2006) schreiben, dass die neuronalen Netzwerke neu trainiert werden müssen, wenn ein weiteres Referenzmuster hinzugefügt wird, was viel Zeit in Anspruch nimmt. Beim Template Matching handelt es sich um ein intensitätsbasiertes Bildzuordnungsverfahren zwischen Referenz- und Suchbild. Für die Klassifizierung wird es unter anderem von PICCIOLI et al. (1996) und MALIK et al. (2007) eingesetzt. Ein neues Referenzmuster kann problemlos hinzugefügt werden. Das Template Matching in seiner Grundform weist aber keine Robustheit auf Skalierung, Rotation oder allgemein auf affine Transformationen auf und reagiert empfindlich auf Beleuchtungsunterschiede (REN et al. 2009).

2.3 Weitere Ansätze für die Detektion und Klassifizierung von Verkehrszeichen

Viele Ansätze lassen sich nicht nur für die Detektion oder Klassifizierung von Verkehrszeichen einsetzen, sondern sind für beide Aufgaben geeignet. Einige davon werden nachfolgend diskutiert.

Die Hough-Transformation reagiert gemäß NGUWI & KOUZANI (2008) tolerant auf Lücken und ist nicht sehr anfällig auf Rauschen. Durch die unterschiedlichen Normgrößen und Formen der Verkehrszeichen, sowie durch die teilweise ungleiche Ausdehnung in den Koordinatenrichtungen, müssen viele Maßstäbe geprüft werden, was sich negativ auf die Rechenzeit und den Speicherbedarf auswirkt. Somit können für Echtzeitanwendungen nur schnellere abgeänderte Methoden eingesetzt werden (NGUWI & KOUZANI 2008). Auch CHU-

TATAPE & GUO (1999) zählen die lange Rechenzeit und den hohen Speicherbedarf der standardisierten Hough-Transformation zu den Nachteilen. Um dies zu verbessern, wird eine modifizierte Variante der Hough-Transformation vorgeschlagen (CHUTATAPE & GUO 1999). Diese Variante verwenden KIM et al. (2006) für die Verkehrszeichendetektion, nachdem vorgängig Kanten mit dem Canny-Operator aus den Bilddaten extrahiert wurden. BARRILE et al. (2007) führen die Formdetektion mit der standardisierten Hough-Transformation durch; für die Klassifizierung verwenden sie die generalisierte Hough-Transformation. Diese wird auch von HABIB et al. (1999) auf Kanten angewandt, welche mit dem Canny-Filter extrahiert wurden.

Für die Detektion und Klassifizierung von Verkehrszeichen werden zunehmend die Ansätze von Support Vector Machines (SVM) und Scale Invariant Feature Transform (SIFT) verfolgt. Beim SIFT-Ansatz nach LOWE (2004) sind die extrahierten Merkmale invariant gegen Translationen, Rotationen und Skalierung und robust gegen Beleuchtungsänderungen, Bildrauschen und kleine geometrische Deformationen (REITERER et al. 2009, REN et al. 2009). FLEYEH & DOUGHERTY (2008) stellen eine Echtzeit-Anwendung für schwedische Verkehrszeichen vor. Mehr als 3400 Bilder kamen für das Training und diverse Tests zum Einsatz. Dabei wurden zwei Stufen von SVM trainiert. Die erste Stufe dient der Formdetektion und die zweite der Erkennung der Piktogramme (FLEYEH & DOUGHERTY 2008). MALDONADO-BASCÓN et al. (2007) verwenden für die Detektion und anschließende Formklassifizierung eine lineare SVM und für die Erkennung des Inhalts eine SVM mit Gauss-Kernel. Ein Verkehrszeichen wird nur dann korrekt detektiert, wenn es in mehr als einem Bild vorkommt. Somit ergeben sich wenige 'False Positives'. Die implementierten Algorithmen sind invariant bei Translation, Rotation, Skalierung und in vielen Fällen bei partiellen Verdeckungen (MALDONADO-BASCÓN et al. 2007).

2.4 Extraktion unterschiedlicher Verkehrszeichen

In der Regel wird entweder auf die Detektion oder die Klassifizierung von Verkehrszeichen fokussiert; praxistaugliche Gesamtlösungen mit zusätzlicher automatischer Kartierung sind kaum implementiert. In mehreren Fällen wird auch nur eine semantische Kategorie wie Geschwindigkeitszeichen behandelt. Im Gegensatz dazu haben REN et al. (2009) eine Echtzeit-Anwendung für fast alle Verkehrszeichen in Neuseeland implementiert. Die Detektion wird mit einer Farbsegmentierung im HSV-Farbraum vorgenommen. Es folgt eine Hough-Transformation und ein anschließendes Matching mit SIFT- oder SURF-Merkmalen. Mit den in MALDONADO-BASCÓN et al. (2008) vorgestellten Algorithmen lassen sich mehr als 300 verschiedene Verkehrszeichen in Spanien erkennen. Die Anwendung funktioniert sowohl für urbane als auch für ländliche Umgebungen. Anpassungen für Verkehrszeichen in anderen Ländern können ohne weiteres vorgenommen werden (MALDONADO-BASCÓN et al. 2008). Ein System, welches Verkehrszeichen verschiedener Länder detektieren und klassifizieren kann, wurde von KHAN et al. (2011) entwickelt. Sie haben diverse Farben und Formen berücksichtigt, und die Algorithmen sind invariant bei Translation, Rotation, Skalierung und partiellen Verdeckungen.

2.5 Kartierung von Verkehrszeichen

Um Verkehrszeichen zu inventarisieren, wurden bis vor kurzem die Koordinaten mit einem handgeführten GNSS-Empfänger gemessen, das Zeichen mit einer Digitalkamera fotografiert und die Sachdaten mit Hilfe eines GIS-Clients vor Ort erfasst. Anhand eines mobilen PDA-basierten Geoinformationssystems können all diese Arbeiten mit nur einem Gerät durchgeführt werden. Ein solches wurde von BENESOVA et al. (2007) für die Inventarisierung von Verkehrszeichen eingesetzt, wobei die Klassifizierung halbautomatisch erfolgte.

Die bisherige Forschung und Entwicklung im Bereich der Detektion und Klassifizierung von Verkehrszeichen war größtenteils getrieben durch die Bedürfnisse der Automobi-

bilindustrie (vgl. Abschnitte 2.1 bis 2.4). Dabei sind die Aspekte Fahrerassistenz und Sicherheit zentral, womit in der Regel auf rote Gefahren-, Vorschrifts- oder Vortrittssignale fokussiert wurde. Die automatisierte exakte Lokalisierung bzw. Kartierung von sämtlichen Verkehrszeichen spielte demgegenüber eine deutlich geringere Rolle, was sich in einer kleineren Anzahl entsprechender Forschungsarbeiten niederschlägt. Zu den besonders relevanten Arbeiten im Bereich der automatisierten Kartierung von Verkehrszeichen gehören sowohl MADEIRA et al. (2005), KIM et al. (2006) und BARÓ et al. (2009), welche die absolute 3D-Koordinatenbestimmung von Objektpunkten aufgrund von Stereobilddaten vornehmen, als auch SHI et al. (2008), die einen kombinierten Ansatz von Bild- und Laserscanning-Daten verwenden. Während SHI et al. (2008) eine Genauigkeit von circa 30 cm erreichen, können MADEIRA et al. (2005) Kartierungen nur im Meterbereich vornehmen. Die exakte 3D-Kartierungsmöglichkeit in einem globalen geodätischen Bezugssystem ist wesentlich und gewinnt zunehmend an Bedeutung, da dadurch digitale Inventare für eine Reihe allgemeiner GIS-basierter Aufgabenstellungen, z.B. Verkehrsplanung, Change Detection, Simulationen, Überlagerung und Nachführung von existierenden oder projizierten Geodaten in Mixed-Reality Umgebungen, angelegt werden können.

3 Lösungsansatz auf Basis der Stereodisparität

Beim entwickelten und nachfolgend vorgestellten Ansatz bildet die Nutzung der Tiefenkarten aus Stereobildaufnahmen ein zentrales Element. Obwohl die Tiefeninformation aus Dense-Stereo-Matching ein riesiges Potential hat, wurde sie für die Extraktion von Verkehrszeichen noch gar nicht oder wie von CYGANEK (2008) nur als optionaler Bestandteil für die Suchraumeinschränkung (Unterteilung in nahe Regionen und Hintergrund) verwendet. Anhand der Stereobilder können die Verkehrszeichen auch dreidimensional kartiert werden, was mit Monobildern nur sehr umständlich und mit beschränkter Genauigkeit möglich ist.

Für die Datenerfassung kann ein Stereo-bild-basiertes Mobile Mapping System wie dasjenige des Instituts Vermessung und Geoinformation (IVGI) der Fachhochschule Nordwestschweiz (FHNW) eingesetzt werden (siehe Abb. 2, links). Bei diesem Messsystem stehen für die Bildaufnahme zurzeit zwei Stereosysteme mit einer Stereobasis von je circa 90 cm und Industriekameras unterschiedlicher Auflösung (Full-HD und 11MPixel) zur Verfügung. Zur direkten Georeferenzierung der Stereobildpaare dient ein Navigationssystem POS LV 210 von Applanix, bestehend aus einem GNSS-Empfänger, einer Inertialmess-einheit und einem Distanzmessindikator. Diverse Testfahrten mit dieser multifunktionalen Forschungsplattform haben ergeben, dass für gut signalisierte Punkte bei einer ausreichenden GNSS-Konstellation 3D-Genauigkeiten von 2–3 cm möglich sind.

Um vereinfachte Verfahren für die Folgeoperationen zu ermöglichen, sind die mit dem mobilen Messfahrzeug erfassten und genähert parallelen Stereobilder mittels Rektifizierung in den Stereonormalfall umzurechnen, nachdem die Verzeichnungskorrektur angebracht wurde. Auf Grund der resultierenden Nor-

malbilder wird mit einem Stereo-Matching-Algorithmus für jedes Pixel die Disparität bzw. x-Parallaxe ermittelt. Für jede Disparität kann danach anhand der Stereogeometrie (Basis und Kamerakonstante) ein Tiefenwert berechnet werden, deren Gesamtheit eine Tiefenkarte bildet. Für die vorliegenden Untersuchungen erfolgt das Dense Matching mit dem in OpenCV implementierten Semi-Global-Block-Matching-Algorithmus (OPEN-CV 2012), der einige Unterschiede zu dem von HIRSCHMÜLLER (2008) entwickelten Semi-Global-Matching-Algorithmus aufweist (z.B. Berechnung der Matchingkosten). Die semi-globale Zuordnung basiert auf einer global zu optimierende Kostenfunktion, welche durch Aufsummierung der Kosten auf ausgewählten Pfaden approximiert wird. Dank dieser Approximation kann dieser semi-globale Zuordnungsalgorithmus für große Bilder, insbesondere auch für hochaufgelöste Luftbilder, eingesetzt werden. Ein Vergleich des SGM-Algorithmus mit einer Vielzahl alternativer Stereo-Algorithmen ist auf der Middlebury Stereo Webseite (SCHARSTEIN & SZELISKI 2012) gegeben.

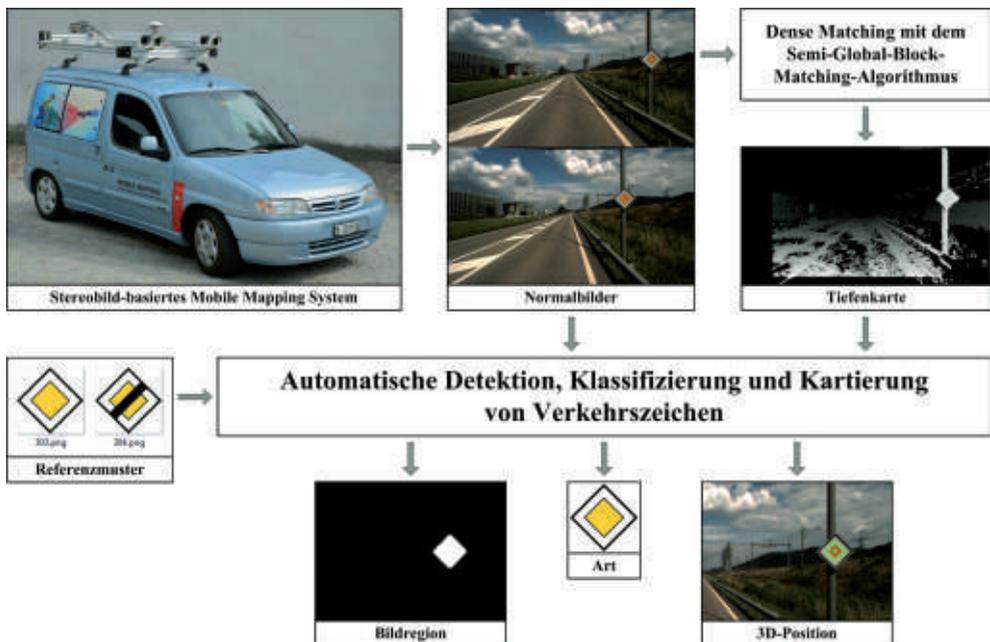


Abb. 2: Eingangs- und Ausgabedaten für die automatische Detektion, Klassifizierung und 3D-Kartierung von Verkehrszeichen.

Sowohl Normalbilder als auch Tiefenkarten sind für die anschließende automatische Detektion und Kartierung von Verkehrszeichen notwendig (siehe Abb. 2). Für die Klassifizierung werden zusätzlich Referenzmuster von allen möglichen Verkehrszeichen benötigt. Diese können im Falle der Schweiz beispielsweise von der Webseite WIKIPEDIA (2012) im PNG-Format heruntergeladen werden. Nach erfolgreicher Detektion, Klassifizierung und Kartierung sind die Bildregion, die Sachdaten sowie die 3D-Position des Verkehrszeichens bekannt.

Bei den entwickelten Algorithmen werden die Stereodisparitäten bzw. die daraus abgeleiteten Tiefenkarten für unterschiedliche Aufgaben verwendet:

- Einschränkung des Suchraums durch den vordefinierten Distanzbereich.
- Definition von distanzabhängigen Bedingungen für die Farbsegmente.
- Bestimmung von Regionen mit ähnlichen Tiefenwerten.
- Berechnung der 3D-Position.
- Einschränkung des Wertebereichs für die Maßstäbe bei der Hough-Transformation.

4 Entwickelte Algorithmen und Softwaremodule

Der vorgestellte Lösungsansatz auf der Basis von Stereobildern und Tiefenkarten wurde mit einer Reihe von Algorithmen und Softwaremodulen in Matlab umgesetzt. Diese decken den gesamten Ablauf von der automatischen Detektion über die Klassifizierung bis zur Kartierung von Verkehrszeichen ab (siehe Abb. 3) und werden nachfolgend erläutert.

4.1 Automatische Detektion von Verkehrszeichen

Als Eingangsdaten sind von jedem Stereobildpaar das linke Normalbild und die entsprechende Tiefenkarte erforderlich (siehe Abb. 4). Im unteren Drittel des Normalbildes sind keine permanenten Verkehrszeichen zu erwarten, weshalb dieser mit einem schwarzen Band ausmaskiert wird. Da die Farbton- und Sättigungs-Komponenten kaum auf unter-

schiedliche Beleuchtungsverhältnisse reagieren und mit diesen somit bessere Ergebnisse bei der Detektion von Verkehrszeichen erzielt werden können, wird das RGB-Normalbild in den HSV-Farbraum transformiert. Um den Suchraum im Normalbild einzuschränken, wird die Tiefenkarte auf einen vordefinierten Distanzbereich reduziert. Das gewählte Basis-Tiefenverhältnis liegt zwischen 0.06 und 0.25, womit auch Verkehrszeichen detektiert werden können, die an einer benachbarten Fahrbahn anliegen. Zudem kann mit einer hohen Bildaufnahmezeit das gleiche Zeichen redundant detektiert und klassifiziert werden. Die Segmentierung von roten, blauen und gelben Farbregionen erfolgt anhand von Schwellwerten für die Farbton- und Sättigungs-Komponenten, welche empirisch auf Grund von mehreren Bildern verschiedener Messfahrten ermittelt wurden. Für blaue Segmente ergab sich beispielsweise ein Farbtonbereich zwischen 0.52 und 0.72 sowie ein Sättigungsbereich zwischen 0.20 und 0.80. Liegt der Farbtonwert zwischen 0.04 und 0.19 und ist der Sättigungs-

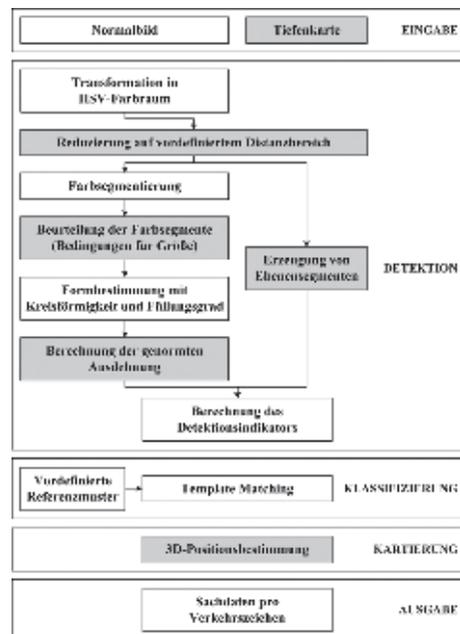


Abb. 3: Entwickelte Algorithmen für die automatische Detektion, Klassifizierung und 3D-Kartierung von chromatischen Verkehrszeichen (graue Felder: Operationen nutzen die Disparitäts- bzw. Tiefeninformation).

wert größer als 0.50 und kleiner als 0.98, handelt es sich um Pixel eines gelben Segments. Nach den morphologischen Bildoperationen Öffnen, Schließen und Füllen wird geprüft, ob die Farbsegmente einer bestimmten Flächengröße, welche von der jeweiligen Aufnahmedistanz des Segments abhängig ist, entsprechen. Ist dies der Fall, erfolgt die Bestimmung der Form mit den zwei Merkmalen Kreisförmigkeit und Füllungsgrad. Diese lassen sich wie folgt berechnen:

$$\text{Kreisförmigkeit} = \frac{4 \times \pi \times \text{Fläche des Segments}}{(\text{Umfang des Segments})^2} \quad (1)$$

$$\text{Füllungsgrad} = \frac{\text{Fläche des Segments}}{\text{Fläche des umschließenden Rechtecks}} \quad (2)$$

Die Differenzen zwischen der Ausdehnung der Segmente und den genormten Seitenlängen, welche in der Signalisationsverordnung (SCHWEIZER BUNDESRAT 1979) definiert sind, dürfen ein Toleranzintervall von beispielsweise 10 cm nicht überschreiten. Die Tiefen-

karten dienen dabei der Bestimmung der metrischen Breite und Höhe des Segments im Objektraum. Die gerundeten Seitenlängen sollten dem Klein- oder Normalformat in der Schweiz entsprechen. Nachdem auf der distanzreduzierten Tiefenkarte ein Schliessen und eine Dilatation ausgeführt wurden, erfolgt die Generierung von Ebenensegmenten. Dies sind Regionen, in denen die Bildpunkte sehr geringe Tiefenunterschiede aufweisen. Die Segmente werden wiederum mit morphologischen Bildoperationen (Öffnen, Schließen und Dilatation) generalisiert. Der Quotient aus Fläche des Ebenensegments innerhalb des Farbsegments (Verschnitt Abb. 4g und Abb. 4k) und Gesamtfläche des Farbsegments (Abb. 4g) bildet den Detektionsindikator. Dieser dient zur Beurteilung der Güte des Detektionsprozesses.

4.2 Automatische Klassifizierung und Kartierung von Verkehrszeichen

Für die Klassifizierung der Verkehrszeichen wird ein korrelationsbasiertes Template Matching mit vordefinierten Referenzmustern

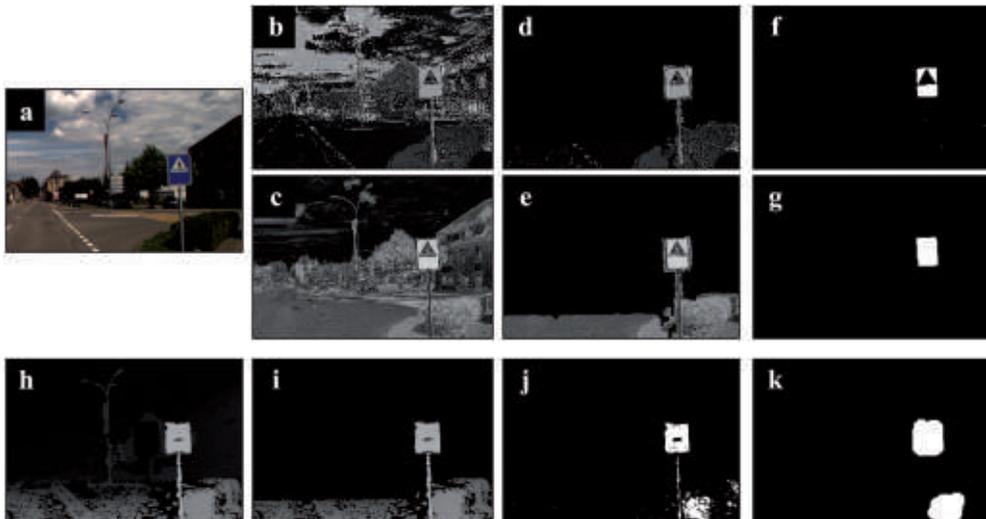


Abb. 4: Automatische Detektion eines rechteckigen Verkehrszeichens mit vorwiegend blauen Farbanteilen (a: linkes Normalbild, b: Farbton-Komponente, c: Sättigungs-Komponente, d: distanzreduzierte Farbton-Komponente, e: distanzreduzierte Sättigungs-Komponente, f: blaue Farbsegmente (binarisiert), g: blaue Farbsegmente nach morphologischen Operationen, h: Tiefenkarte, i: distanzreduzierte Tiefenkarte, j: Ebenensegmente, k: Ebenensegmente nach morphologischen Operationen).

durchgeführt (siehe Abb. 5). Dabei werden nicht alle Muster geprüft, denn durch einen hierarchischen Klassifizierungsansatz kann die Kandidatenmenge auf Grund von Farbe und Form wesentlich reduziert werden. Verkehrszeichenarten, welche auf den erfassten Straßenabschnitten nicht vorkommen, können vom Klassifizierungsprozess ausgeschlossen werden.

Es wird das Referenzmuster bestimmt, mit welchem über den Suchbereich der größte normierte Kreuzkorrelationskoeffizient ermittelt wird. Dieser Wert dient auch als Klassifizierungsindikator. Überschreitet er einen vordefinierten Schwellwert, gilt die Klassifizierung als erfolgreich.

Die Größe des Suchbildes ist durch den Verkehrszeichenausschnitt im Normalbild zu-

züglich eines definierbaren Bereichs (z.B. 10 Pixel) auf jeder Seite definiert. Das Referenzmuster wird auf die Ausdehnung des Farbsegments innerhalb des Suchbildes skaliert. Die Berechnung der Korrelation erfolgt auf Grund desjenigen Kanals, welcher bei empirischen Untersuchungen die größte Ähnlichkeit zwischen einem realen Verkehrszeichenbild und dem entsprechenden synthetischen Referenzmuster zeigte. Bei roten Verkehrszeichen ist dies der blaue Kanal, bei blauen der rote und bei gelben die Sättigungs-Komponente. Pixel des Suchbildes, welche im Referenzmuster weiß sind, weisen je nach Bildqualität einen zu tiefen Grauwert auf (siehe Abb. 5). Um die Matching-Ergebnisse zu verbessern, werden somit alle Pixelwerte des Suchbildes auf weiß (maximaler Grauwert) oder schwarz (Null)

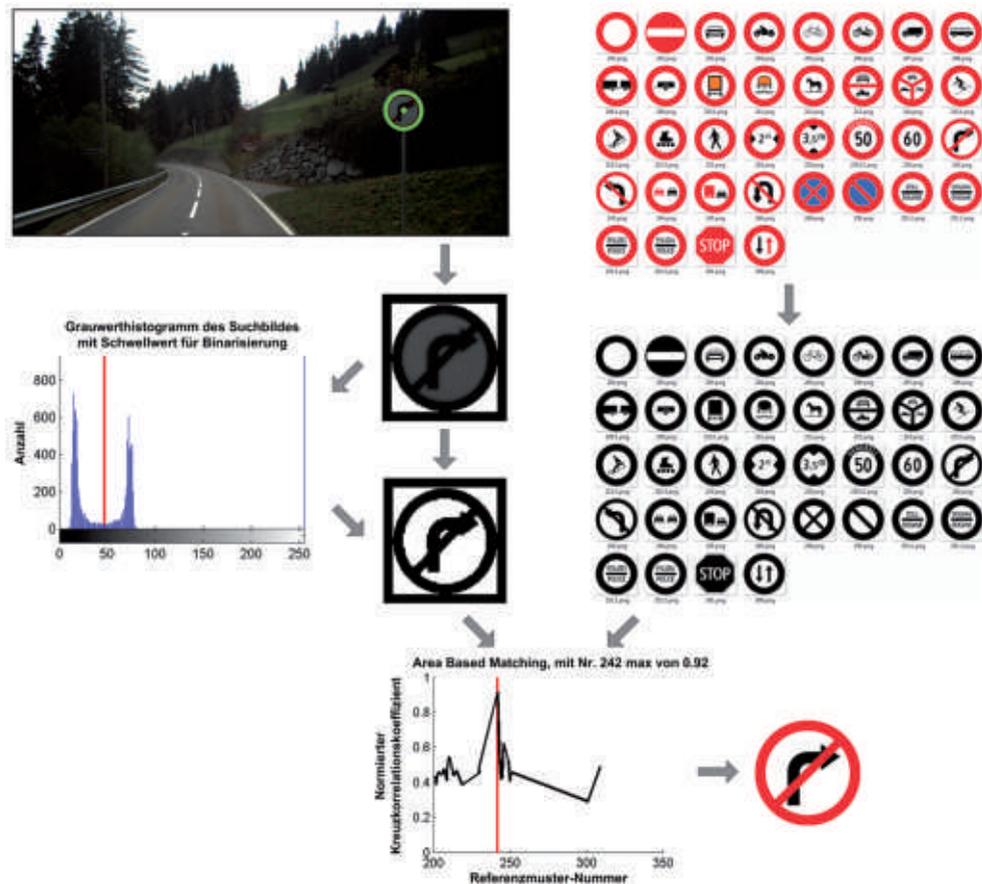


Abb. 5: Automatische Klassifizierung eines roten runden Verkehrszeichens mit kreuzkorrelations-basiertem Template Matching.

gesetzt. Die Ermittlung des dafür benötigten Schwellwertes erfolgt dynamisch auf Grund der Grauwertverteilung des Suchbildes. Dabei wird jener Grauwert bestimmt, bei welchem der Wert des Summenhistogramms größer als ein bestimmter prozentualer Anteil der betrachteten Pixel ist. Bei Dreiecken ist dieser Anteil 40 % und bei den restlichen Formen 75 %. Die Grauwerte vom Maximum des linken und rechten Bereichs im Grauerthistogramm werden ermittelt, der Mittelwert dient als Schwellwert.

Wenn ein Verkehrszeichen detektiert und klassifiziert werden konnte, erfolgt die Bestimmung der 3D-Objektkoordinaten. Für die Berechnung der Modellkoordinaten sind die Tiefeninformation und die Bildkoordinaten des Schwerpunktes sowie die Parameter der inneren Orientierung notwendig. Die anschließende Transformation in den gewünschten geodätischen Referenzrahmen setzt voraus, dass die äusseren Orientierungsparameter des linken Normalbildes bekannt sind.

Unterschiedliche Sachdaten wie die Landeskoordinaten des Schwerpunktes, die Referenzmustersnummer und die genormten Seitenlängen werden automatisch bestimmt und in eine Textdatei ausgegeben. Diese können für den Aufbau oder die Aktualisierung eines Geoinformationssystems genutzt werden.

4.3 Benutzerunterstützter Ansatz

Verkehrszeichenkandidaten, welche einen zu niedrigen Detektions- oder Klassifizierungsindikator aufweisen, müssen nachträglich durch den Benutzer verifiziert werden. Dafür steht eine grafische Benutzeroberfläche zur Verfügung. Fälschlicherweise als Verkehrszeichen segmentierte Regionen können verworfen und eine falsche Klassenzuweisung korrigiert werden. Wurde ein Verkehrszeichen nicht automatisch detektiert, kann es interaktiv aus den Bilddaten extrahiert werden, indem der Schwerpunkt digitalisiert wird. Die manuelle Digitalisierung übernimmt die Rolle der Suchraumeinschränkung aus Abschnitt 4.1, ansonsten kommen die gleichen Algorithmen wie beim automatischen Ansatz zum Zuge. Dies ist vor allem dann hilfreich, wenn schlechte Tiefenkarten vorliegen. Sind

die Schwellwerte für die Farbsegmentierung nicht gut gewählt, kann die ebenfalls implementierte Hough-Transformation zu besseren Ergebnissen führen. Dabei werden je nach Farbe die Formen Kreis, Rechteck, Quadrat, Dreieck und Raute getestet.

4.4 Detektion von achromatischen Verkehrszeichen

Für die Farbsegmentierung von weißen und grauen Verkehrszeichen erwies sich die Bestimmung von robusten Schwellwerten für die Farbton- und Sättigungs-Komponenten als untauglich, da für entsprechende Flächen in den Bilddaten keine klaren Grauwert-Intervalle ausgemacht werden konnten. Zudem weisen viele Bildpunkte in der Umgebung der Verkehrszeichen ähnliche Grauwerte auf. Der Ansatz von PAČLÍK et al. (2000) konnte auch nicht verfolgt werden, da die Bestimmung eines festen Intervalls für die Hellwert-Komponente nicht erfolgreich war.

Nach GÓMEZ-MORENO et al. (2010) kann die Segmentierung von achromatischen Farben durch die Verwendung der chromatischen/achromatischen Dekomposition verbessert werden. Dafür sind mehrere Ansätze denkbar, welche von ihnen vorgestellt und evaluiert werden. Der chromatische/achromatische Index, welcher von LIU et al. (2002) entwickelt wurde, ist einer davon und lässt sich wie folgt berechnen:

$$\text{CAD}(R,G,B) = \frac{|R-G| + |G-B| + |B-R|}{3D} \quad (3)$$

R, G, B repräsentieren die Farbkomponenten für ein gegebenes Pixel und D ist der Extraktionsgrad einer achromatischen Farbe. Ist der CAD-Wert kleiner oder gleich 1, handelt es sich um eine achromatische Farbe. Bei weißen Verkehrszeichen muss zusätzlich die Summe von R, G, B größer als ein definierter Schwellwert sein (GÓMEZ-MORENO et al. 2010). Bei LIU et al. (2002) und MALDONADO-BASCÓN et al. (2007) werden die besten Ergebnisse mit $D = 20$ erzielt, bei GÓMEZ-MORENO et al. (2010) mit $D = 30$ und einem Schwellwert für weiß von 180. Diese Resultate konnten mit den vorliegenden Bilddaten nicht reproduziert wer-

den. Somit wurde für die automatische Detektion von weißen und grauen Verkehrszeichen die generalisierte Hough-Transformation um die Tiefeninformation erweitert. Dabei werden zunächst Kanten mit dem Canny-Operator auf Grund der Hellwert-Komponente extrahiert. Alle Kantenpixel, welche sich in den Ebenensegmenten befinden, dienen als Eingabe für die Hough-Transformation. Geprüft werden die Formen Kreis und Rechteck. Die Einschränkung des Wertebereichs für die Maßstäbe erfolgt mit der Distanz zum Verkehrszeichen, wofür wiederum die Tiefenkarte zur Anwendung kommt. Für die anschließende Klassifizierung wird erneut das Template Matching aus Abschnitt 4.2 durchgeführt.

5 Untersuchungen und Resultate

Für die Güteuntersuchungen der entwickelten Algorithmen wurden mehr als 4000 georeferenzierte Stereobildpaare von einer Messfahrt mit dem mobilen Kartierungsfahrzeug des IVGI in Muttenz bei Basel verwendet. Die dabei eingesetzten Kameras, welche in einem Abstand von 92,7 cm montiert waren, weisen eine Kamerakonstante von 7,9 mm auf. Die Pixelgröße beträgt 7,4 µm bei einer Sensorgröße von 1926 x 1082 Pixel. Die Messfahrt erfolgte Mitte November 2010 bei bewölktem Himmel und entsprechend ungünstigen Beleuchtungsverhältnissen, was dunkle Bildaufnahmen zur Folge hatte. Pro Sekunde wurden 5 Full-HD-Stereobildpaare bei einer Fahrgeschwindigkeit von circa 40 km/h erfasst, was einem Stereobildabstand von etwa 2 m entspricht.

5.1 Detektions- und Klassifizierungsgüte

Für die Beurteilung der Detektions- und Klassifizierungsergebnisse erfolgte eine Unterteilung in Hauptstraßen und Quartierstraßen. Relevante Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen wurden bestimmt. In diesem Fall waren dies vor allem rechts an der Straße anliegende Verkehrszeichen, welche normalerweise senkrecht zur Fahrtrichtung ausgerichtet sind. Dies repräsentiert die Verkehrszeichen, welche in der Regel entlang von Hauptstraßen anzutreffen sind.

Bei der Prozessierung im Distanzbereich von 4 bis 14 m konnten mit den entwickelten Algorithmen 89 % der relevanten Verkehrszeichen automatisch detektiert werden (siehe Tab. 1). Korrekt klassifiziert wurden auf einer Hauptstraße 86 % und auf Quartierstraßen 82 %. Wie bei HÖFERLIN & ZIMMERMANN (2009) wurde auch der Indikator für die Präzision der Detektion berechnet:

$$\text{Präzision} = \frac{TP}{TP + FP} \quad (4)$$

Dabei ist die Anzahl der ‘True Positives’ TP durch die Summe der ‘True Positives’ und ‘False Positives’ FP zu dividieren. Die hohen Werte in Tab. 1 sagen aus, dass sehr wenige Segmente fälschlicherweise als Verkehrszeichen ermittelt wurden.

Dass einige Verkehrszeichen nicht detektiert oder klassifiziert werden konnten, beruht auf unterschiedliche Ursachen. Bei der Detektion im Bereich von Baustellen ist es möglich, dass aufgrund von Abschränkungen und Warnvorrichtungen viele rote Farbsegmen-

Tab. 1: Detektions- und Klassifizierungsgüte der entwickelten Algorithmen für eine Messfahrt in Muttenz bei Basel (Alle: Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen; Relevante: rechts an der Straße anliegende Verkehrszeichen, welche vielfach senkrecht zur Fahrtrichtung ausgerichtet sind).

| | | Anzahl | Detektionsrate | Klassifizierungsrate | Präzision |
|------------------------|------------------|-----------|----------------|----------------------|-----------|
| Hauptstraße | Alle | 90 | 72% | 70% | 94% |
| | Relevante | 63 | 89% | 86% | |
| Quartierstraßen | Alle | 152 | 55% | 47% | 98% |
| | Relevante | 65 | 89% | 82% | |

te ausgeschieden werden. Sind die Flächen nicht allzu groß, kann dies zu 'False Positives' führen. Obwohl die Verkehrszeichen in der Schweiz in der Regel einen guten Zustand aufweisen, sind gewisse Verkehrszeichen vergilbt. Dabei weist die Sättigungs-Komponente sehr niedrige Werte auf, was auch bei im Schatten liegenden Verkehrszeichen der Fall ist. Da der definierte Schwellwert für diese Komponente nicht überschritten werden kann, ist die Detektion nicht möglich. Verkehrszeichen, welche fast parallel zur Fahrtrichtung ausgerichtet sind, haben in den Normalbildern teilweise eine zu geringe Breite im Vergleich zur Höhe, womit diese Farbsegmente mit der aktuellen Lösung verworfen werden. Schwierigkeiten bereitet die automatische Detektion auch bei schlechten bzw. unvollständigen Tiefenkarten. Für mehrere Verkehrszeichen existierten keine vordefinierten Referenzmuster, was zu keiner oder zu einer falschen Klassifizierung führte. Ein nicht optimaler Schwellwert für die Binarisierung des Suchbildes kann die Ursache für einen zu niedrigen Korrelationskoeffizienten sein.

5.2 Kartierungsgenauigkeit

Für die Berechnung der Kartierungsgenauigkeit wurden die Referenzkoordinaten von 22 Verkehrszeichen mittels Tachymeter bestimmt. Die Differenzen zu den mit den entwickelten Algorithmen automatisch ermittelten Koordinaten wurden ausgewertet. Tab. 2 zeigt, dass für die Komponente quer zur Fahrtrichtung und für die Höhe eine empirische Standardabweichung von etwa 5 cm erreicht werden konnte und für die Komponente in Fahrtrichtung 6 cm. Erwartungsgemäß streut die Komponente in Fahrtrichtung bzw. senkrecht zur Stereobasis am meisten und ist somit am schlechtesten bestimmt. Die größ-

te Abweichung einer Komponente beträgt 16 cm, die meisten Differenzen pro Komponente liegen jedoch im Bereich bis 5 cm. Die empirische Standardabweichung einer Koordinatendifferenz für die 3D-Position beträgt 9,5 cm.

6 Fazit und Ausblick

Es konnte aufgezeigt werden, dass durch eine konsequente Nutzung der Tiefeninformation aus Dense-Stereo-Matching in der automatisierten Verkehrszeichenkartierung sehr hohe Detektionsraten erreicht werden können. Unter guten Bedingungen können etwa 90 % der Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen im Normal- und Kleinformat in der Schweiz automatisch detektiert und 85 % korrekt klassifiziert werden. Benutzerunterstützt lassen sich diese Werte um weitere 5 % erhöhen. Somit müssen nur noch 5 bis 10 % der Verkehrszeichen interaktiv in den Stereobildern oder im Feld digitalisiert werden. Zudem treten durch den benutzerunterstützten Ansatz und diverse Bedingungen nahezu keine 'False Positives' auf. Eine vollautomatische Kartierung mit einer 3D-Genauigkeit von unter 10 cm ist möglich.

Der entwickelte Ansatz reagiert robust auf Skalierung, Translation und kleine Rotationen. Obwohl mit nahe gelegenen Verkehrszeichen bessere Ergebnisse zu erwarten sind, kann das Signal im ganzen vordefinierten Distanzbereich detektiert werden. Das Verkehrszeichen kann beliebig im Bild positioniert sein. Kleine Rotationen werden toleriert. Zudem ist es möglich, im gleichen Bild mehrere Verkehrszeichen mit den Formen Kreis, Rechteck, Quadrat, Dreieck und Raute zu detektieren. Eine reduzierte Leistung des Verfahrens ist bei schief stehenden oder stark verdrehten Verkehrszeichen sowie auf sehr stei-

Tab. 2: Praktische Kartierungsgenauigkeit der entwickelten automatischen Verkehrszeichenkartierung für eine Messfahrt in Muttenz bei Basel (mm).

| | Δ quer | Δ längs | Δ Höhe | Δ Lage | Δ 3D |
|-------------------------|---------------|----------------|---------------|---------------|-------------|
| Mittelwert | -36 | 23 | -36 | 66 | 86 |
| Maximum | 152 | 146 | 157 | 154 | 159 |
| m_{Diff} | 46 | 64 | 53 | 79 | 95 |

len Straßen zu erwarten, was jedoch in der Praxis noch nicht beobachtet werden konnte.

Neben qualitativ guten Tiefenkarten ist auch eine gute Farbsegmentierung für den Detektionserfolg entscheidend. Dazu müssen passende Schwellwerte vorliegen. Bei den Untersuchungen wurde das Intervall für die jeweilige Komponente relativ groß gewählt. Dies ist aber nur möglich, weil der Suchraum durch die Tiefeninformation markant eingeschränkt wird und fälschlicherweise detektierte Verkehrszeichen durch gewisse Bedingungen verworfen werden.

Da Detektions- und Klassifizierungsgütern von 100 % kaum realistisch sind, besteht die Möglichkeit, die automatisch kartierten Verkehrszeichen in einem georeferenzierten 3D-Video zu überlagern. Die 3D-Videos können mit dem StereoVision-Client (BURKHARD et al. 2011) betrachtet, die Ergebnisse somit kontrolliert und die fehlenden Verkehrszeichen spektiv digitalisiert werden. Eine erste Implementierung der Algorithmen für weiße und graue Verkehrszeichen, welche die Tiefeninformation in Kombination mit der Hough-Transformation verwendet, wird in Zukunft weiter verbessert. Geplant sind auch die Detektion von weiteren komplexen Verkehrszeichen und die Erkennung von beliebigem Text (WU et al. 2005). Eine Erhöhung der geometrischen Qualität und Zuverlässigkeit lässt sich durch Matching in Stereobild-Sequenzen erreichen (HUBER et al. 2011). Die Verfolgung der Verkehrszeichen über mehrere Stereobildpaare würde insbesondere eine Steigerung der semantischen Qualität bewirken.

Laufende Arbeiten sollen die Auswirkung der Kameraauflösung auf die Detektions- und Klassifizierungsgenauigkeit aufzeigen. Erste Untersuchungen mit einem Stereosystem bestehend aus höher aufgelösten Industriemesskameras mit je 11 MPixeln deuten auf eine leichte Verbesserung der Ergebnisse hin. Für die Erkennung von Text ist die höhere geometrische Auflösung zwingend notwendig. Laufende Untersuchungen zeigen zudem, dass die Qualität der Tiefenkarten mit den höher aufgelösten Bildsensoren und mit geeigneten radiometrischen Anpassungen nochmals deutlich gesteigert werden kann, was sich wiederum positiv auf die automatisierte Verkehrszeichenkartierung auswirkt.

Referenzen

- BARÓ, X., ESCALERA, S., VITRIÀ, J., PUJOL, O. & RADEVA, P., 2009: Traffic Sign Recognition Using Evolutionary Adaboost Detection and Forest-ECOC Classification. – *IEEE Transactions on Intelligent Transportation Systems* **10** (1): 113–126.
- BARRILE, V., CACCIOLA, M., MEDURI, G.M. & MORABITO, F.C., 2007: Automatic Recognition of Road Signs by Hough Transform. – **5th International Symposium on Mobile Mapping Technology**, Padua, Italy.
- BENESOVA, W., LYPETSKYY, Y., ANDREU, J.-P., PALETTA, L., JEITLER, A. & HÖDL, E., 2007: A Mobile System for Vision Based Road Sign Inventory. – **5th International Symposium on Mobile Mapping Technology**, Padua, Italy.
- BURKHARD, J., NEBIKER, S. & EUGSTER, H., 2011: Stereobild-basiertes Mobile Mapping: Technologie und Anwendungen. – *Geomatik Schweiz* **109** (6): 295–298.
- CHUTATAPE, O. & GUO, L., 1999: A modified Hough transform for line detection and its performance. – *Pattern Recognition* **32** (2): 181–192.
- CYGANEK, B., 2008: Road-Signs Recognition System for Intelligent Vehicles. – **Second International Workshop, RobVis 2008**: 219–233, Auckland, New Zealand.
- DE LA ESCALERA, A., ARMINGOL, J.M. & MATA, M., 2003: Traffic sign recognition and analysis for intelligent vehicles. – *Image and Vision Computing* **21** (3): 247–258.
- DE LA ESCALERA, A., ARMINGOL, J.M., PASTOR, J.M. & RODRÍGUEZ, F.J., 2004: Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles. – *IEEE Transactions on Intelligent Transportation Systems* **5** (2): 57–68.
- DE WITH, P., HAZELHOFF, L., CREUSEN, I. & BRUINSMAN, H., 2010: Efficient Road Maintenance. Automatic Detection and Positioning of Traffic Signs. – *GEOInformatics* **13** (7): 10–12.
- FLEYEH, H. & DOUGHERTY, M., 2005: Road and Traffic Sign Detection and Recognition. – **10th EWGT Meeting and 16th Mini-EURO Conference**: 644–653, Poznań, Poland.
- FLEYEH, H., 2006: Shadow And Highlight Invariant Colour Segmentation Algorithm For Traffic Signs. – *IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand.
- FLEYEH, H. & DOUGHERTY, M., 2008: Traffic Sign Classification Using Invariant Features and Support Vector Machines. – *IEEE Intelligent Vehicles Symposium*: 530–535, Eindhoven, The Netherlands.

- FU, M.-Y. & HUANG, Y.-S., 2010: A Survey of Traffic Sign Recognition. – International Conference on Wavelet Analysis and Pattern Recognition: 119–124, Qingdao, China.
- GAO, X.W., PODLADCHIKOVA, L., SHAPOSHNIKOV, D., HONG, K. & SHEVTSOVA, N., 2006: Recognition of traffic signs based on their colour and shape features extracted using human vision models. – Journal of Visual Communication and Image Representation **17** (4): 675–685.
- GARCÍA-GARRIDO, M.A., SOTELO, M.A. & MARTÍN-GOROSTIZA, E., 2006: Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions. – 9th International IEEE Conference on Intelligent Transportation Systems: 811–816, Toronto, Canada.
- GÓMEZ-MORENO, H., MALDONADO-BASCÓN, S., GIL-JIMÉNEZ, P. & LAFUENTE-ARROYO, S., 2010: Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition. – IEEE Transactions on Intelligent Transportation Systems **11** (4): 917–930.
- HABIB, A.F., UEBBING, R. & NOVAK, K., 1999: Automatic Extraction of Road Signs from Terrestrial Color Imagery. – Photogrammetric Engineering & Remote Sensing **65** (5): 597–601.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.
- HÖFERLIN, B. & ZIMMERMANN, K., 2009: Towards Reliable Traffic Sign Recognition. – IEEE Intelligent Vehicles Symposium: 324–329, Xi'an, China.
- HOSSAIN, M.S., HASAN, M.M., ALI, M.A., KABIR, M.H. & ALI, S., 2010: Automatic Detection and Recognition of Traffic Signs. – IEEE Conference on Robotics Automation and Mechatronics: 286–291, Singapore.
- HUBER, F., NEBIKER, S. & EUGSTER, H., 2011: Image Sequence Processing in Stereovision Mobile Mapping – Steps towards Robust and Accurate Monoscopic 3D Measurements and Image-Based Georeferencing. – Photogrammetric Image Analysis 2011, Lecture Notes in Computer Science **6952**: 85–95, Springer, Berlin.
- JAU, U.L., TEH, C.S. & NG, G.W., 2008: A comparison of RGB and HSI color segmentation in real-time video images: A preliminary study on road sign detection. – International Symposium on Information Technology: 2576–2581, Kuala Lumpur, Malaysia.
- KHAN, J.F., BHUIYAN, S.M.A. & ADHAMI, R.R., 2011: Image Segmentation and Shape Analysis for Road-Sign Detection. – IEEE Transactions on Intelligent Transportation Systems **12** (1): 83–96.
- KIM, G.-H., SOHN, H.-G. & SONG, Y.-S., 2006: Road Infrastructure Data Acquisition Using a Vehicle-Based Mobile Mapping System. – Computer-Aided Civil and Infrastructure Engineering **21** (5): 346–356.
- LIU, H., LIU, D. & XIN, J., 2002: Real-time Recognition of Road Traffic Sign in Motion Image based on Genetic Algorithm. – First International Conference on Machine Learning and Cybernetics: 83–86, Beijing, China.
- LOWE, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.
- MADEIRA, S.R., BASTOS, L.C., SOUSA, A.M., SOBRAL, J.F. & SANTOS, L.P., 2005: Automatic Traffic Signs Inventory Using a Mobile Mapping System. – International Conference and Exhibition on Geographic Information GIS Planet, Estoril, Portugal.
- MALDONADO-BASCÓN, S., LAFUENTE-ARROYO, S., GIL-JIMÉNEZ, P., GÓMEZ-MORENO, H. & LÓPEZ-FERRERAS, F., 2007: Road-Sign Detection and Recognition Based on Support Vector Machines. – IEEE Transactions on Intelligent Transportation Systems **8** (2): 264–278.
- MALDONADO-BASCÓN, S., LAFUENTE-ARROYO, S., SIEGMANN, P., GÓMEZ-MORENO, H. & ACEVEDO-RODRIGUEZ, J., 2008: Traffic Sign Recognition System for Inventory Purposes. – IEEE Intelligent Vehicles Symposium: 590–595, Eindhoven, The Netherlands.
- MALIK, R., KHURSHID, J. & AHMAD, S.N., 2007: Road Sign Detection and Recognition using Colour Segmentation, Shape Analysis and Template Matching. – 6th International Conference on Machine Learning and Cybernetics: 3556–3560, Hong Kong.
- NGUWI, Y.-Y. & KOUZANI, A.Z., 2008: Detection and classification of road signs in natural environments. – Neural Computing and Applications **17** (3): 265–289.
- OPENCV, 2012: OpenCV v2.1 Documentation. Camera Calibration and 3d Reconstruction. StereoSGBM. – http://opencv.willowgarage.com/documentation/cpp/camera_calibration_and_3d_reconstruction.html#stereosgbm (31.1.2012).
- PAČLIK, P., NOVOTICOVÁ, J., PUDIL, P. & SOMOL, P., 2000: Road sign classification using Laplace kernel classifier. – Pattern Recognition Letters **21** (13/14): 1165–1173.
- PICCIOLI, G., DE MICHELI, E., PARODI, P. & CAMPANI, M., 1996: Robust method for road sign detection and recognition. – Image and Vision Computing **14** (3): 209–223.
- REITERER, A., HASSAN, T. & EL-SHEIMY, N., 2009: Robust Extraction of Traffic Signs from Georefer-

- erenced Mobile Mapping Images. – **6th** International Symposium on Mobile Mapping Technology, Presidente Prudente, São Paulo, Brazil.
- REN, F., HUANG, J., JIANG, R. & KLETTE, R., 2009: General Traffic Sign Recognition by Feature Matching. – **24th** International Conference on Image and Vision Computing New Zealand: 409–414, Wellington.
- SCHARSTEIN, D. & SZELISKI, R., 2012: Middlebury Stereo Evaluation – Version 2. – <http://vision.middlebury.edu/stereo/eval/> (31.1.2012).
- SCHWEIZER BUNDES RAT, 1979: Signalisationsverordnung. – <http://www.admin.ch/ch/d/sr/7/741.21.de.pdf> (31.1.2012).
- SCHWEIZER FERNSEHEN, 2008: Überflüssige Schilder: Teuer und gefährlich. – <http://www.tagesschau.sf.tv/Nachrichten/Archiv/2008/08/26/Schweiz/Ueberfluessige-Schilder-Teuer-und-gefaehrlich> (31.1.2012).
- SHI, Y., SHIBASAKI, R. & SHI, Z.C., 2008: Towards Automatic Road Mapping by Fusing Vehicle-Borne Multi-Sensor Data. – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **37** (B5): 867–872.
- TRIMBLE, 2009: Signs of Change in Belgium. – *Technology & more* **2009** (3): 12–13.
- WIKIPEDIA, 2012: Bildtafel der Verkehrszeichen in der Schweiz. – http://de.wikipedia.org/wiki/Bildtafel_der_Verkehrszeichen_in_der_Schweiz (31.1.2012).
- WU, W., CHEN, X. & YANG, J., 2005: Detection of Text on Road Signs From Video. – *IEEE Transactions on Intelligent Transportation Systems* **6** (4): 378–390.

Adresse der Autoren:

STEFAN CAVEGN & Prof. Dr. STEPHAN NEBIKER, FHNW Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, CH-4132 Muttenz, Tel.: +41-61-467-4336, Fax: +41-61-467-4460, e-mail: {stefan.cavegn}{stephan.nebiker}@fhnw.ch

Manuskript eingereicht: September 2011
Angenommen: März 2012