



Erkennung und Klassifizierung von Verkehrszeichen

Fabian Ruff & David Latotzky

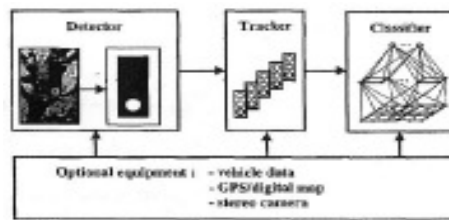


Motivation

- Fahrassistenzsysteme mit Verkehrszeichenerkennung könnten:
 - Vor gefährlichen Situationen **warnen** (z.B. Geschwindigkeitsbegrenzung überschritten).
 - Auf wichtige Zeichen **hinweisen** (z.B. Straßenschäden, Baustelle).
 - Über aktuelle Beschränkungen der Strasse **informieren** (z.B. Überholverbot)

Systemkomponenten

- Häufig: 3 Module
 - Detektion: Finden der Kandidaten (ROI) in jedem Bild.
 - Tracking: Verfolgen eines Kandidaten über mehrere Bilder.
 - Klassifizieren: Ordnet jeden Kandidaten einem Verkehrszeichen zu oder verwirft ihn.



- Alternativ: Detektion + Klassifizieren in einer Phase.



Rahmen des Referates

- Übersicht über Verfahren zur Detektion und Klassifizierung
- Echtzeit-Aspekte berücksichtigen
- Tracking nur im Ansatz
- Mit nur einer Kamera
- Keine weiteren Hilfsmittel (GPS, etc...)

Modul: Detektion

- Ziel: Im Kamerabild Kandidaten für VZ finden (region of interest, ROI)
- größte Herausforderung im System
 - VZ verpassen vs. false positives
 - rechenintensiv

false positives: Irrtümlich für VZ gehaltene Regionen.

Der Detektor sollte alle VZ wahrnehmen.

Regionen in denen fälschlicherweise ein VZ vermutet wird sollten zwar nicht im Übermaß auftreten, werden aber in Kauf genommen, um kein VZ übersehen wird.

Es ist dann die Aufgabe des Klassifizierers diese „false positives“ zu verwerfen.

Die Erkennungs- und die „false positive“- Rate stehen in direktem Gegensatz zueinander.

Farbbasierte Verfahren

- Segmentierung eines Bildes nach relevanten Farben (z.B. rot, gelb, blau, schwarz, weiß)
- Zusammenhängende Regionen bestimmen
- Seitenverhältnis und Größe einer Region als weiteres Kriterium für eine ROI

Die Segmentierung erfolgt pixelweise.

Die Farbklassifizierung kann erfolgen durch

- Festkodierte Schwellwerte

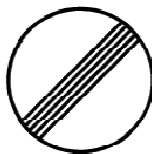
- Trainieren eines Klassifikators (Offline) mit einer manuell klassifizierten Trainingsmenge von Verkehrszeichen.

Die Klassifizierung der Farbe wird häufig für alle möglichen (RGB) Werte vorberechnet und in einer Look-Up Tabelle gespeichert.

Die Bestimmung der Zusammenhangskomponenten kann man „Region Growing“ verwenden (<http://www.ideo.columbia.edu/res/pi/4d4/region-growing.html>).

Probleme Farbsegmentierung(1)

- Schwellwerte für Farbsegmentierung schwer zu bestimmen
- Manche Schilder sind farblos

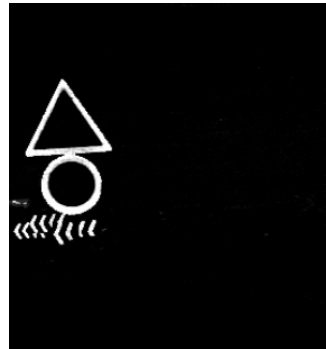


Vor allem die unterschiedlichen Lichtbedingungen machen die Klassifizierung sehr schwer.

Außerdem ist auch die immer wieder die mangelnde Dynamik von Digitalkameras bei stark unterschiedlich beleuchteten Aufnahmen problematisch (z.B. direkter Sonneneinfall oder schattige Unterführungen)

Probleme Farbsegmentierung(2)

- Zusammenfallende Regionen





Formbasierte Verfahren

- In der Regel: Vorverarbeitung
 - Kanten-/Eckenerkennung
 - Ergebnis: binäres Feature-Bild
- Formerkennung im Feature-Bild
 - z.B. mit
 - template matching
 - Hough-Transformation

Vorteil von formbasierten Ansätzen: Eine Graubildkamera reicht aus (Aber: Klassifizierung schwieriger).



Hough-Transformation(1)

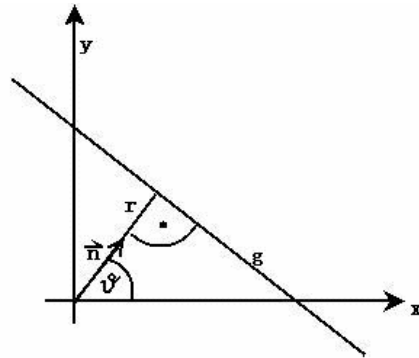
- Ursprünglich zur Erkennung von Linien
- Generalisierbar für beliebige parametrisierbare Formen
- Grundprinzip: Jedes Feature-Pixel „stimmt“ für Formen die es beinhalten.
Formen mit vielen Stimmen kommen in Frage

Hough-Transformation(2)

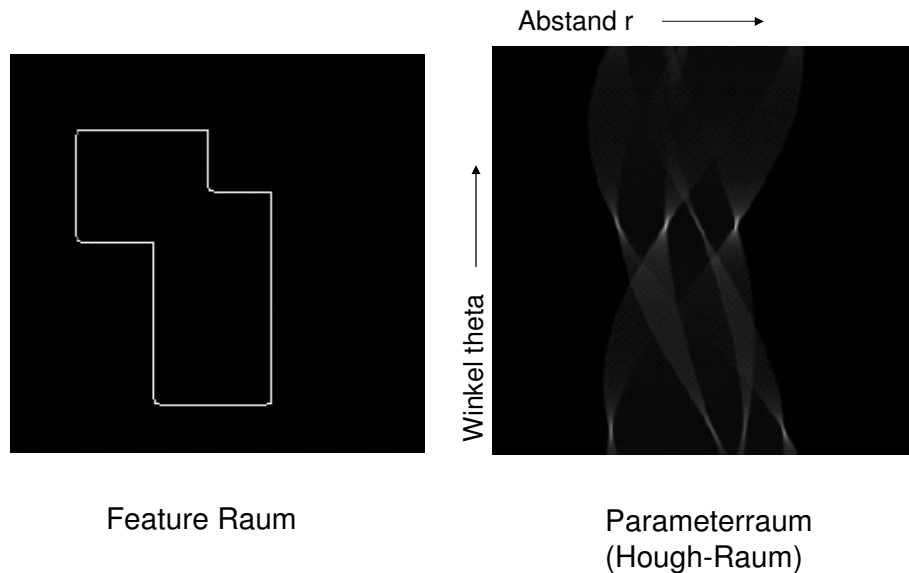
- Gerade g darstellbar durch Hessesche Normalenform:

$$\begin{pmatrix} x \\ y \end{pmatrix} \cdot \vec{n} = r$$

- r : Abstand zum Ursprung
- φ : Winkel des Normalenvektors zur x -Achse



Hough-Transformation(4)



Die Bereiche im Hough-Raum von hoher Intensität sind Linien die viele Feature-Pixel enthalten.

Es sind 8 solche Bereiche zu sehen die den 8 Linien im Feature Raum entsprechen.

Die Normalen der 4 senkrechten Linien haben den Winkel 0. Die waagerechten Linien haben den Winkel $\pi/2$.

Deshalb sind auch jeweils 4 Bereiche im Hough-Raum auf einer waagerechten angeordnet (selber Winkel).

Die Intensität der Bereiche ist ein direkter Hinweis auf die Länge der jeweiligen im Feature-Raum.



Probleme bei Formerkennung

- Echtzeitanwendung kritisch
- teilweise verdeckte Schilder
- Rotation und Verzerrung

Häufig sind Optimierungen erforderlich um Echtzeitkriterium zu erfüllen.
Deshalb auch häufig im Einsatz wenn nur eine Form gesucht ist (z.B. bei Ampelerkennung).



Vergleich

- Formbasierte Ansätze robuster gegen unterschiedliche Lichtbedingungen
- Formerkennung rechenintensiver
- Farbbasierter Ansatz invariant zu Skalierung und Blickwinkel

Modul: Klassifizierung (1)

- bekommt Kandidaten vom Detektor (Bildausschnitt)
- bestimmt:
 - welches Verkehrszeichen der Kandidat ist.
 - oder ob es kein Verkehrszeichen ist.
- Herausforderung: robust false-positives ablehnen.

Ca. 95% der Kandidaten sind false-positives (aus: [1] Robust traffic signal recognition)

Richtwert: max. 1% false-positives zulassen, 10% der relevanten Objekte ablehnen.

Die Fehler die hier entstehen werden (hoffentlich) vom Tracker bei der endgültigen Entscheidung ignoriert.



Klassifizierung (2)

- Bekommt sequenziell jeden Kandidaten.
- Kandidat wird vorverarbeitet.
- Kandidat wird Verkehrszeichen zugeordnet oder verworfen.
- Tracker trifft endgültige Entscheidung

Tracker hilft endgültige Entscheidung per Mehrheitswahl über alle Kandidaten einer Track.

Sobald die Wahrscheinlichkeitsmasse hoch genug wird kann eine vorzeitige Entscheidung getroffen werden.

Einzelne Fehlentscheidungen des Klassifizierers können so kompensiert werden.

Vorverarbeitung

- Skalierung auf NxN Pixel
 - duplizieren bzw. überspringen von Pixeln.
- Ausstanzen des Zeichens.
- Bei Bild als Eingabe:
 - Kontraste normalisieren.
 - Farbraum reduzieren.
- Bei Merkmalen als Eingabe:
 - Merkmale invariant zur Skalierung (z.B. Kompaktheit, Hauptkomponenten)

Neben den Hauptkomponenten und der Kompaktheit gibt es unter anderen auch noch folgende geeignet Merkmale (hier nicht weiter erläutert / zur eigenen Recherche):

- unscaled spatial moment
- unscaled central moment
- normalized central moment

Die ersten 40 - 50 Hauptkomponenten waren in einem Experiment von DaimlerChrysler zur Unterscheidung von 12 verschiedenen Verkehrszeichen nötig, um eine akzeptable Erkennungsquote zu erlangen ([2] Hypothesis verification based on classification at unequal error rates)

Mehr Infos zu Linearer Interpolation:

- <http://www.computerbase.de/lexikon/Interpolation>
- http://de.wikipedia.org/wiki/Lineare_Interpolation



Klassifizierungstechniken

- Allgemein: mehrschichtige Neuronale Netze
 - Radial Basis Function – Netzwerke
 - Polynomial Classifier
- Alternativ: Kombination von Detektion und Klassifizierung

Einführende Informationen zu Radial Basis Function – Netzwerken, Polynomial Classifier und Simulated Annealing finden Sie auf den Zusatzfolien.

Mehr Informationen zu RBF-Netzen:

- <http://www.anc.ed.ac.uk/~mjo/intro/intro.html>

Trainieren des Klassifizierers (1)

- Variationen im internationalen Standard müssen berücksichtigt werden.
- Umweltbedingungen sorgen für Variationen



Leider gibt es keine Nicht-kommerziellen Datenbanken für Verkehrszeichen - Bilder.

Das größte Archiv an Aufnahmen stellt DaimlerChrysler den von ihnen geförderten Forschungseinrichtungen zur Verfügung.

Trainieren des Klassifizierers (2)



- Bestimmte Muster aus dem Verkehrs-Alltag sehen bekannten Verkehrszeichen sehr ähnlich und führen zu false-positives.
- Klassifizierer trainiert explizite „Garbage“-Klasse die immer abgelehnt wird.



Trainieren des Klassifizierers (3)

- Vergrößern der Trainingsmenge (Bootstrapping):
 - Schräglage des Bildes simulieren.
 - Noise/Verunreinigung künstlich hinzufügen.
 - Pixelverschiebungen erzeugen.

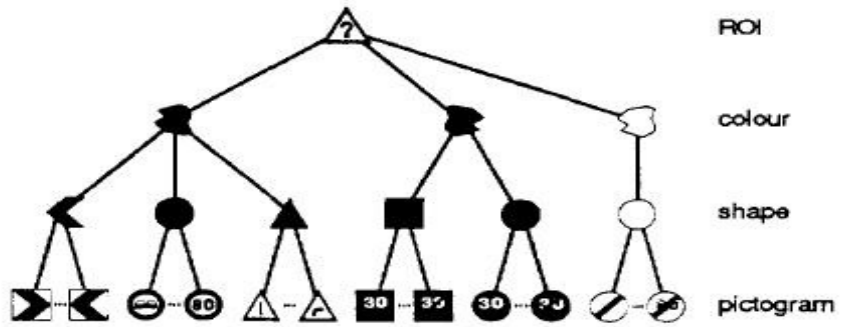
z.B. für Schräglage: 5 Varianten von jedem Bild (1° - 5° Neigung)



Hierarchisches Klassifizieren

- Erst wird „grob“ in wenige Unterklassen klassifiziert.
- Spezialisierte Klassifizierer für jede Unterklasse.
- Vorteile:
 - Wenn ein Klassifizierer die Eingabe ablehnt ist die grobe Klasse u. U. noch bekannt.
 - Klassifizierer müssen in weniger Unterklassen einteilen.

Hierarchisches Klassifizieren





Modul: Tracking

- gruppiert Kandidaten einer Bildsequenz, die dem selben Objekt in der echten Welt entsprechen
- Stabilisierung der Klassifizierung
- Endgültige Klassifizierung durch (einfache) Mehrheitswahl über mehrere Frames einer Track.

Zur eigenen Recherche:

Der Tracker nutzt zum Gruppieren häufig so genanntes „Correlation Tracking“ oder einen Kalman-Filter.

Implementierungen(1)

- 1999, DaimlerChrysler
- runde/dreieckige Schilder
- Hierarchisches template matching
- RBF Klassifizierer
- 90% Erkennungsrate, <5% false positives
- 6-8Hz: 266 Mhz Dual Pentium II

Graubildkamera mit Auflösung von 360x288 Pixel

Klassifizierer:

16x16 Pixel Bilder als Eingabe

50-250 Referenzvektoren pro Klasse

Verwendetes Fahrzeug: E-Klasse T-Model

Quelle: [3]



Implementierung(2)

- 2004, DaimlerChrysler
- Ampelerkennung mit Farbkamera
- Detektor: Farbsegmentierung
- Klassifizierer: Feed-Forward Neural Network
- GPS + erweiterte digitale Karten
- 95% Erkennungsrate, 1 fpos/hour
- 25 Hz, 3 GHz Pentium IV

Es wurden 95% aller Ampeln in bis zu 100m Entfernung erkannt.
Sensor: 10 Bit VGA High Dynamic Range Camera, 16mm Linse
Quelle: [1]

Quellen(1)

- [1] **Robust Recognition of Traffic Signals**
Linder, Kressel, Kaelberer; DaimlerChrysler 2004
- [2] **Hypothesis Verification Based on Classification at Unequal Rates**
Kressel, Lindner, Wöhler, Linz; DaimlerChrysler 1999
- [3] **Traffic Sign Recognition Revisited**
D.M. Gavrila, DaimlerChrysler AG 1999
- [4] **From Door to Door – principles and applications of computer vision and driver assistant systems**
Franke, Gavrila, Gern, Görzig, Jannssen Paetzold, Wöhler, DaimlerChrysler
- [5] **Autonomous Driving approaches Downtown**
Franke, Gavrila, Linder, Wöhler, Görzig, Paetzold; 1999, DaimlerChrysler



Quellen(2)

- [6] **Hybrid Approach for Traffic Sign Recognition**
Janssen, Ritter, Stein, Ott; 1993 DaimlerChrysler
- [7] **Road Sign Classification using the Laplace Kernel Classifier**
Paclik, 1999 Czech Technical Univ.
- [8] **Road Traffic Sign Detection and Classification**
Escalera, Mereno, Salichs, Armignol
1997, Universidad Carlos III de Madrid
- [9] **Traffic Sign Detection for Driver Support Systems**
Escalera, Armingol, Salichs
Universidad Carlos III de Madrid
- [10] **A Model-Based Road Sign Identification System**
Lauzière, Gingras, Ferrie
2001, McGill University, Canada