

F. Gaßmann, St. Müller:

**Hintergrundschätzung für Aufgaben der
Verkehrsflussanalyse
mittels digitaler Bildverarbeitung**

**48. Internationales Wissenschaftliches Kolloquium
Technische Universität Ilmenau
22.-25. September 2003**

Inhalt

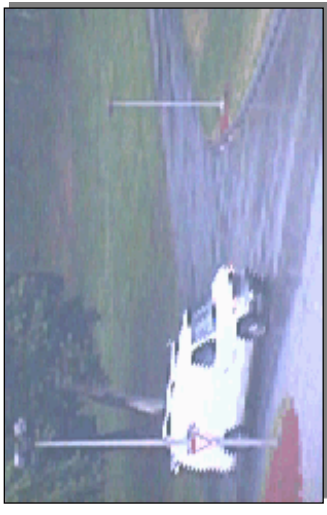
- ▶ **Zielstellung / Einordnung der Arbeiten**
- ▶ **Problematik der Objektsegmentierung in
Outdoorszenen**
- ▶ **Verfahren zur Objektsegmentierung**
- ▶ **Multi-Gauß-Ansatz**
- ▶ **Anpassung / Erweiterungen des Verfahrens**
- ▶ **Ergebnisse und Ausblick**

Projekt / Verbundprojekt

- **Verbundprojekt OIS - Optische Informationssysteme für die Verkehrsszenenanalyse und Verkehrslenkung -**
 - ▶ gefördert vom Bundesministerium für Bildung und Forschung
 - ▶ Projektlaufzeit: 09 / 2001 - 12 / 2002
- ▶ Projektträger: DLR – Deutsches Zentrum für Luft- und Raumfahrt e.V.
- ▶ Projektziel:
 - ▶ Entwicklung eines Systems zur automatischen Auswertung von Bildfolgen aus Straßenverkehrsszenen
 - ▶ Gewinnung aussagekräftiger Verkehrsdaten direkt aus Kreuzungsszenen
 - ▶ Interpretation der Verkehrsdaten zum Zweck der Verkehrsprognose und Knotenpunktsteuerung

Projekt / Aufgabenstellung

- ▶ Unterauftragnehmer: Zentrum für Bild- und Signalverarbeitung (ZBS) e.V.
- ▶ Aufgabenstellung: automatische Segmentierung verkehrsrelevanter Objekte in Bildfolgen durch Hintergrundmodellierung
- ▶ Teilaufgaben:
 - ▶ automatische Initialisierung der Hintergrundmodellierung
 - ▶ Objektsegmentierung (Fahrzeuge, Fußgänger, keine Klassifikation)
 - ▶ Modelladaption an Bildstream (Reaktion auf Änderungen des Hintergrunds)



Projekt / Randbedingungen

- pro Kreuzung: mehrere Kameraknoten mit jeweils bis zu 4 Kameras (GW, RGB oder IR)
- pro Kameraknoten: Objekterkennung (Segmentierung, Tracking und Klassifikation), Hardware: PC, (optional: Auslagerung von Teilalgorithmen auf FPGA)
- pro Kamera: Bildgröße: ca. 400 x 500 Pixel, Bildrate: 200 ms, Objekterkennung in Echtzeit erforderlich

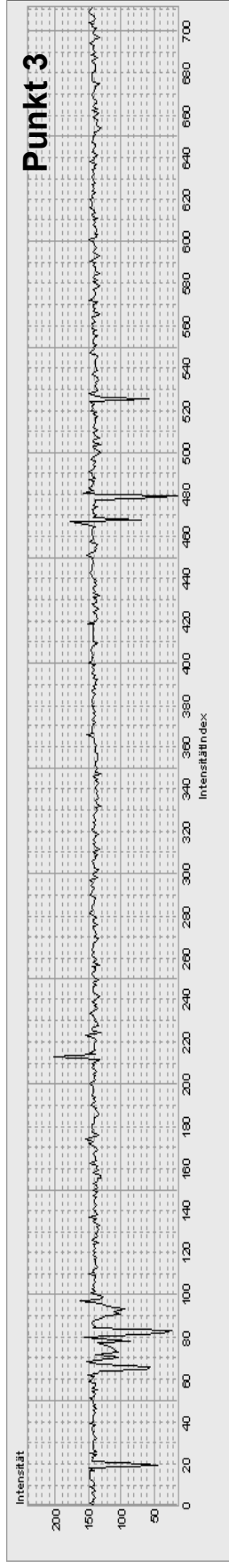
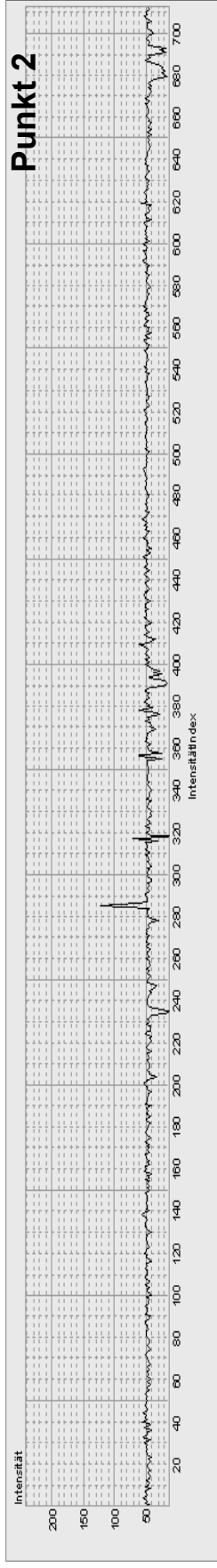
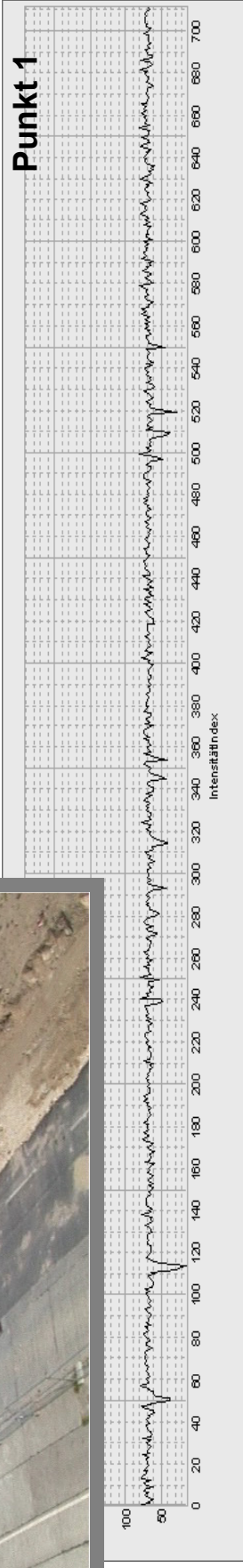


Probleme in Verkehrsszenen

- plötzliche globale Änderungen der Wetter- bzw. Umweltbedingungen
 - Sonne / Wolken: Helligkeit, Schatten (FZ, Gebäude, ...)
 - Regen: Helligkeit, Direktreflexionen (gesamte Szene, nasse Fahrbahn)
 - Schneefall: gesamte Szene, Schnee / Schneematsch
- Problematik der langen Stillstandszeiten von Objekten (Ampelschaltung, Hintergrund ist lange und gleichbleibend verdeckt und selten sichtbar)
- Bewegung von Bäumen, Sträuchern, Gras, Wasser, ... durch Wind
- hintergrundidentische FZ-Bereiche (gleichfarben)
- Überlappungen von FZ
- Schattenwurf von FZ
- Direktreflexionen bei Sonne
- Kamerabewegung durch Wind
- Besonderheiten bei Dunkelheit (völlig veränderte Beleuchtungsverhältnisse)

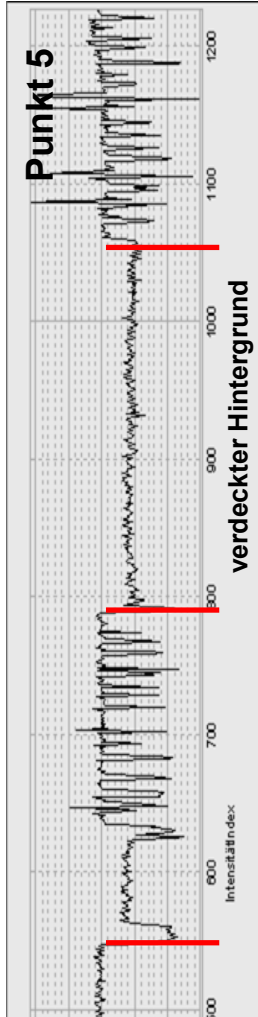
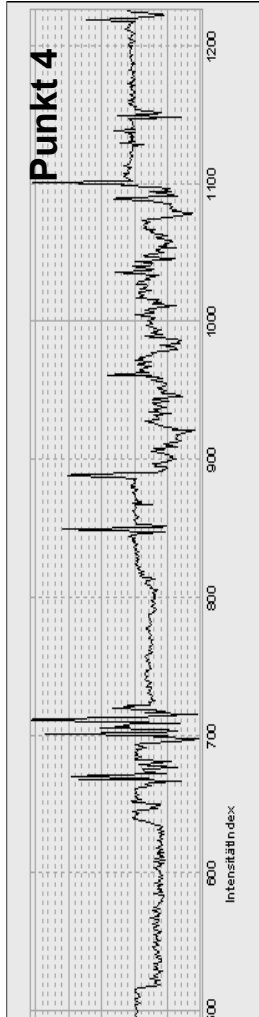
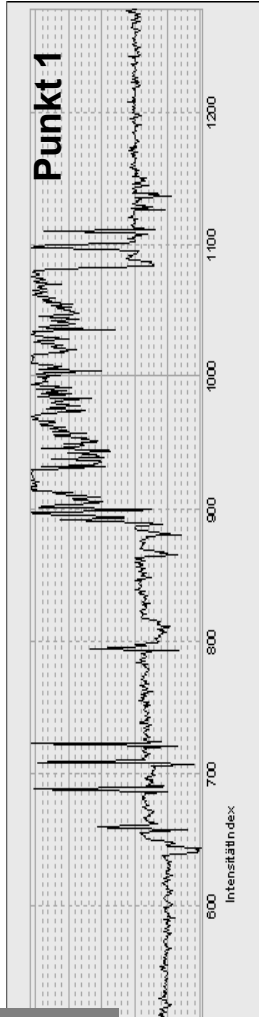


Typischer Grauwertverlauf an Punkten auf „normaler“ Straße (keine stehenden Objekte)

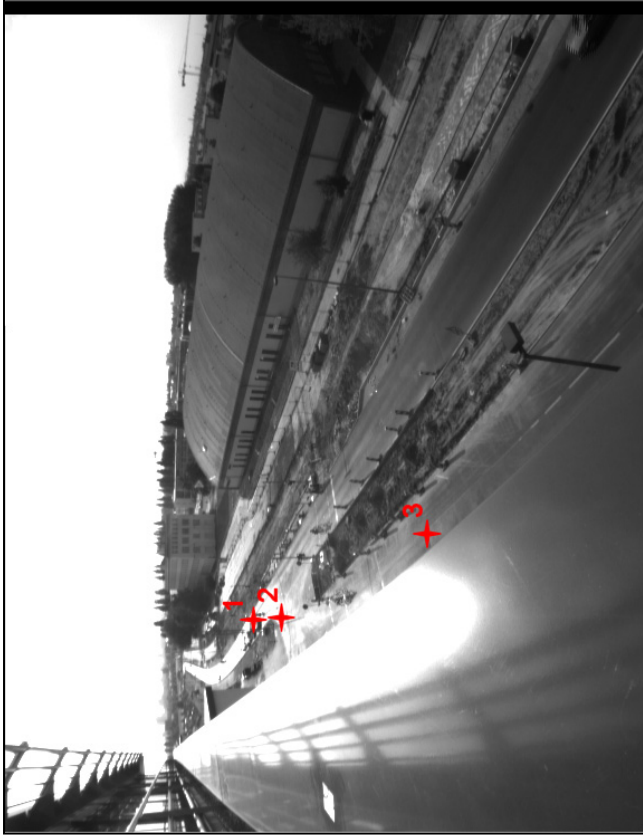




Grauwertverlauf an Punkten vor der Ampel (mit stehenden Objekten)

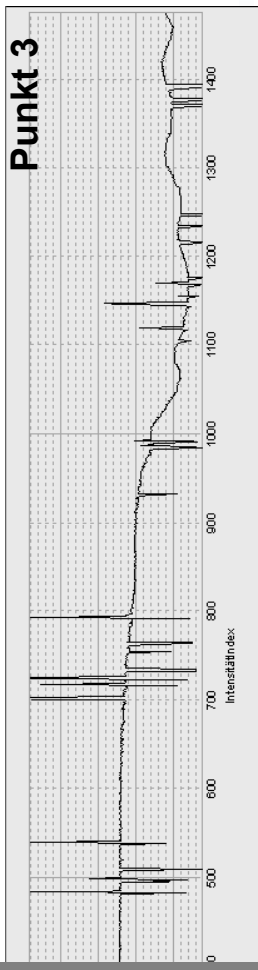
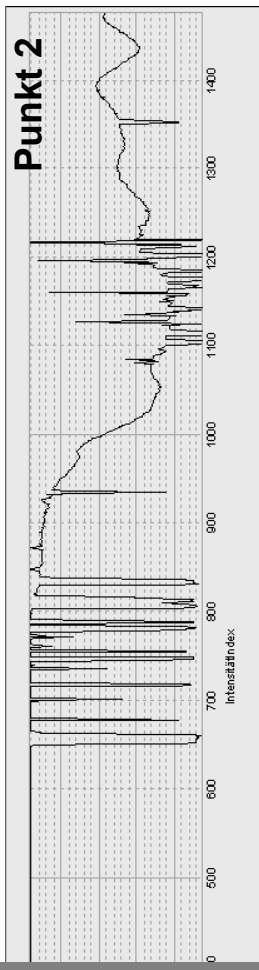
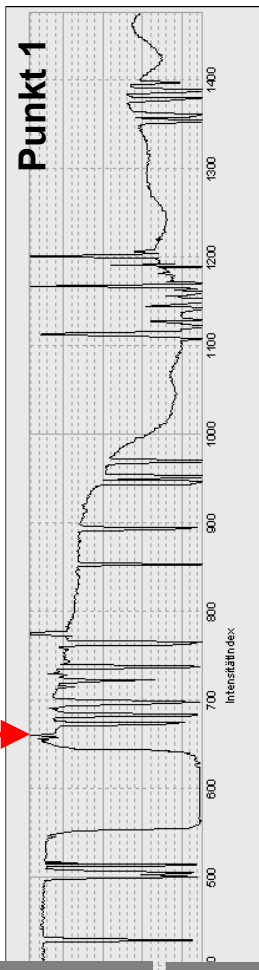


Hintergrundniveau



Beleuchtungsschwankungen durch sich verändernde Bewölkung (führt in Punkt 1 und 2 zur Invertierung von HG und VG)

Beleuchtungsänderung ab hier



Ansätze aus der Literatur



seit mehr als 10 Jahren: große und steigende Anzahl von Veröffentlichungen zur Problematik „Verkehrsanalyse durch Bildverarbeitung“

- **Differenzbildanalyse / einfache statistische Verfahren**
- **Hintergrundmodellierung durch Gaußverteilungen**
 - einzelne Verteilung (Single Gauß)
 - Kombination von Verteilungen (Multi Gauß)
- **Sonstige**
 - Eigenbilder
 - Hidden Markov Modelle
 - Optischer Fluss
- **Hybride Ansätze (Kombination von Verfahren)**

Ansätze und vergleichende Bewertung

alle Verfahren: unzureichende Segmentierungsqualität
(Diskrepanz zwischen schnellen globalen Änderungen der Hintergrundhelligkeit und langen Stillstandszeiten von Fahrzeugen)

Ansatz	Bewertung
Differenzbildakkumulator	2)
Kalman-Schätzer	2)
Single-Gauß	2)
Multiple-Gauß	
Eigenbildanalyse	2) 3) 4)
Hidden Markov Modelle	2)
Optischer Fluss	2) 4)

- 2) keine Modellierung periodischer Schwankungen der Hintergrundhelligkeit
(Windbewegung von Bäumen, Sträuchern, Gras, Wasser, E-Leitungen, Fahnen, ..., andere bewegte Objekte z.B. Werbetafel)
- 3) hoher Speicherbedarf
- 4) lange Rechenzeit

Multi-Gaußverteilung (1)

- ▶ einfaches Verfahren
- ▶ breiter Spielraum für Modifikationen



**die weiteren Arbeiten
konzentrierten sich
deshalb auf diesen
Ansatz**

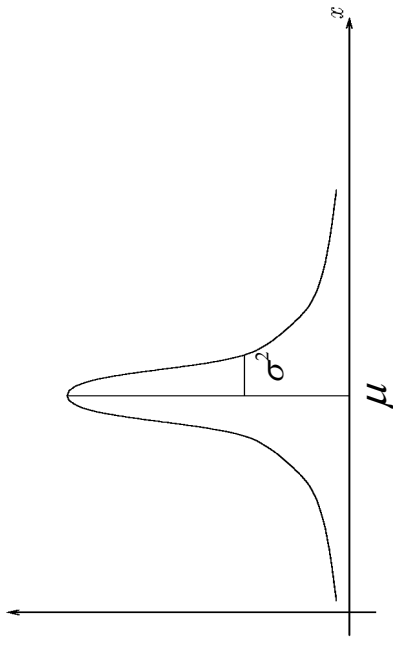


Differenzbildanalyse / einfache statistische Verfahren

- ▶ **Verfahren:**
 - Differenzbildanalyse:
 - Be- und Verrechnung von Differenzbildern der Bildserie
 - einfache statistische Hintergrundmodelle:
 - Akkumulation der Differenzen aufeinander folgender Bilder oder Filter über eine Anzahl von zurückliegenden Bildern z.B. Median, Mittelwert oder Modalwert
 - Segmentierung von bewegten Objekten durch Schwellwertverfahren: Änderungen werden detektiert, Schwellen heuristisch (auch adaptiv) gewählt
- ▶ **Vorteile:**
 - geringer Speicherbedarf, schnell, einfach zu implementieren
- ▶ **Nachteile:**
 - schlechte Segmentierungsqualität (Detektion von Änderungen des Hintergrunds, unvollständige Objektsegmentierung, ...)

Einzelne Gaußverteilung (1)

- ▶ **Verfahren:**
 - pro Bildpixel: Repräsentation des Hintergrunds durch Mittelwert und Streuung einer Gaußverteilung
 - Segmentierungsschwellwerte aus Streuung
 - Adaption des Hintergrunds mittels Updatefunktion:
 - Kalmanfilter
 - gewichtetes Mittel
- ▶ **Vorteile :**
 - geringer Speicherbedarf
 - gut parallelisierbar
 - einfacher, schneller Algorithmus
- ▶ **Nachteile :**
 - bewegter / wechselnder Hintergrund resultiert in sehr breiter Streuung -> Segmentierungsqualität sinkt deutlich



Einzelne Gaußverteilung (2)

► **Update:**

$$p_{x_t} = \frac{1}{\sqrt{2\pi}\sigma_{t-1}} e^{-\frac{1}{2} \frac{(x_t - \mu_{t-1})^2}{\sigma_{t-1}^2}}$$

$$\mu_t = \mu_{t-1} + p_{x_t} (x_t - \mu_{t-1})$$

$$\sigma_t^2 = \sigma_{t-1}^2 (1 - p_{x_t}) + p_{x_t} (x_t - \mu_{t-1})^2$$

► **Klassifikation:**

- **Hintergrund** falls gilt: $\frac{|x_t - \mu_{t-1}|}{\sigma_{t-1}} < \delta$
- **Vordergrund** sonst

Kalmanfilter (1)

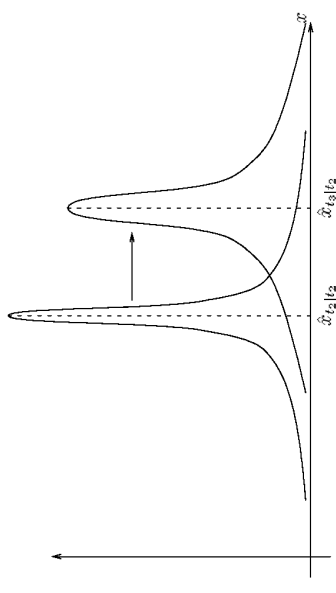
► Verfahren:

- Kalmanfilter ist ein Schätzer, der aus einer Folge von Messwerten eine Aussage über den zukünftigen Systemzustand trifft
- Schätzung erfolgt mittels Rekursionsfunktion für jeden neuen Messwert:
 - aus Varianz des letzten Schätzwertes und der Varianz des gemessenen Wertes wird KalmanGain berechnet
 - Differenz des neuen Messwertes zur Prognose wird mit KalmanGain gewichtet zum alten Schätzwert addiert
 - Wahl des KalmanGain abhängig von Klassifikationsergebnis (Vordergrund: langsame Adaption, Hintergrund: schnelle Adaption)

- Segmentierung VG / HG durch Betrachtung der Differenz von vorhergesagtem und tatsächlichem Wert

► Vorteile / Nachteile:

- siehe unimodale Gaußverteilung



vorhergesagter Wert für Zeitpunkt t_3 aus Filter zum Zeitpunkt t_2 berechnet

Kalmanfilter (2)

► Update:

$$\begin{pmatrix} \tilde{s}_t \\ \dot{\tilde{s}}_t \end{pmatrix} = A \begin{pmatrix} s_{t-1} \\ \dot{s}_{t-1} \end{pmatrix}$$

Vorhersage

$$A = \begin{pmatrix} 1 & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix}$$

Systemmatrix

$$\begin{pmatrix} s_t \\ \dot{s}_t \end{pmatrix}_t = \begin{pmatrix} \tilde{s}_t \\ \dot{\tilde{s}}_t \end{pmatrix} + K_t \left(x_t - H \begin{pmatrix} \tilde{s}_t \\ \dot{\tilde{s}}_t \end{pmatrix} \right)$$

Updategleichung

$$H = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Messmatrix

$$K_t = \begin{pmatrix} k_{t1} \\ k_{t2} \end{pmatrix}$$

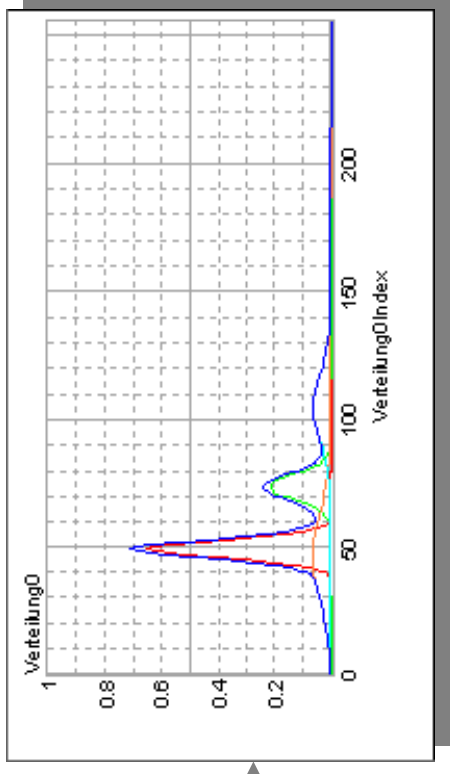
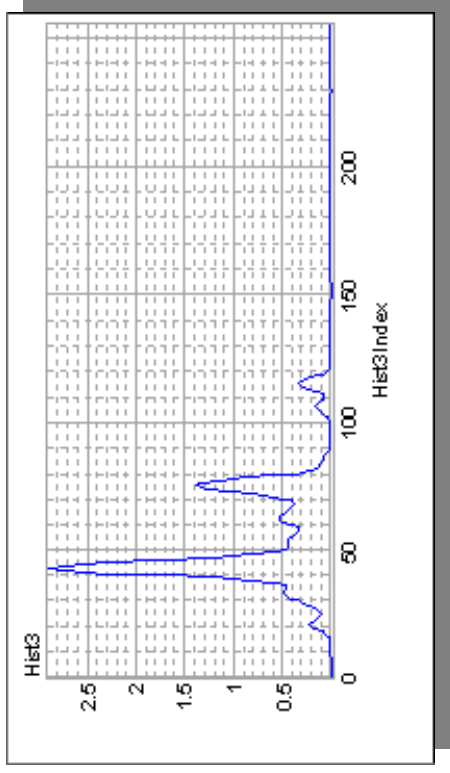
Kalmangain

(vereinfacht durch zwei konstante Werte, die in Abhängigkeit vom Klassifikationsergebnis gewählt werden)

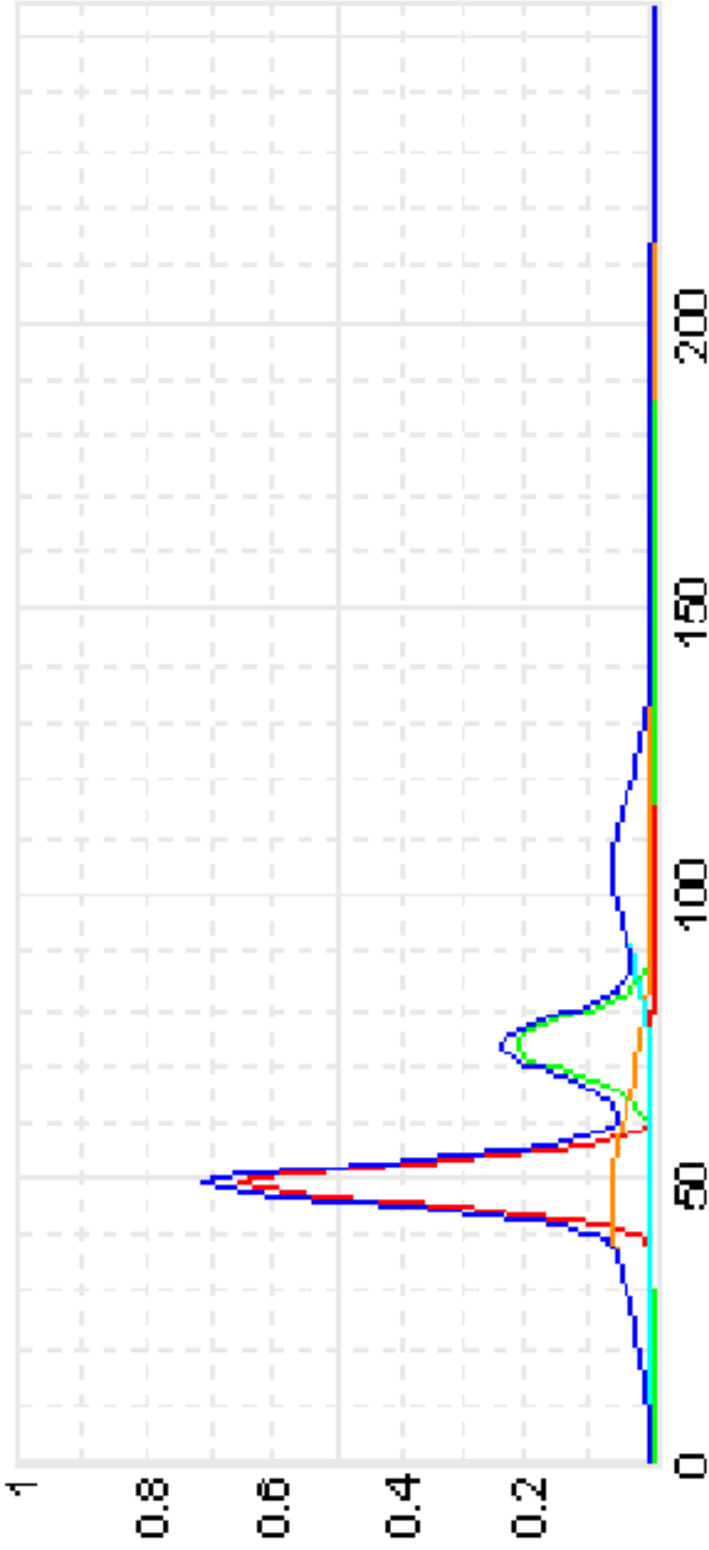
$$k_{t1} = k_{t2} = \begin{cases} \alpha, & \text{falls } |\tilde{s}_t - x_t| > th \\ \beta, & \text{sonst} \end{cases}$$

Multi-Gaußverteilung (2)

- **Prinzip:**
 - Verteilung jedes Pixels wird durch eine gewichtete Summe von Gauß-Verteilungen modelliert
 - am stärksten gewichtete Verteilungen repräsentieren Hintergrund
- **Vorteile:**
 - länger zurückliegende Zustände bleiben bekannt
 - gute Modellierung sich periodisch ändernden Hintergrunds
 - Initialisierung unkritisch
 - großer Spielraum für Modifizierungen des Basisverfahrens



Multi-Gaußverteilung (3)



- Rot, Orange, Grün, Zyan: 4 einzelne Gauß-Verteilungen
- Dunkelblau: Wahrscheinlichkeitsdichte der Mischverteilung

Multi-Gaußverteilung (4)

- **Updateprinzip** (EM-Algorithmus):
 - fällt ein Wert in die Nähe des Mittelwertes einer Verteilung, so wird diese adaptiert und ihr Gewicht erhöht
 - wenn nicht: Ersetzung der am wenigsten relevanten Verteilung
 - die Gewichte der nicht adaptierten Verteilungen werden verringert

$$p_{x_t} = \frac{1}{\sqrt{2\pi}\sigma_{t-1}} e^{-\frac{1}{2} \frac{(x_t - \mu_{t-1})^2}{\sigma_{t-1}^2}}$$

Wahrscheinlichkeitsdichte:

$$w_{k,t} = (1 - \alpha)w_{k,t-1} + \alpha M_{k,t} \quad M_{k,t} = \begin{cases} 1 & \text{Treffer} \\ 0 & \text{sonst} \end{cases}$$

Verteilungsgewichte:

$$\begin{aligned} \sigma_t^2 &= (1 - \rho)\sigma_{t-1}^2 + \rho(x_t - \mu_t)^T (x_t - \mu_t) \\ \mu_t &= (1 - \rho)\mu_{t-1} + \rho x_t \end{aligned} \quad \text{mit} \quad \rho = \alpha p_{x_t}$$

Streuung:

Mittelwert:

Unzulänglichkeiten des Multi-Gaußverfahrens (1)

- **Konflikt zwischen langen Standzeiten von Fahrzeugen (Stauraum vor Ampel) und schnellen Beleuchtungsänderungen**

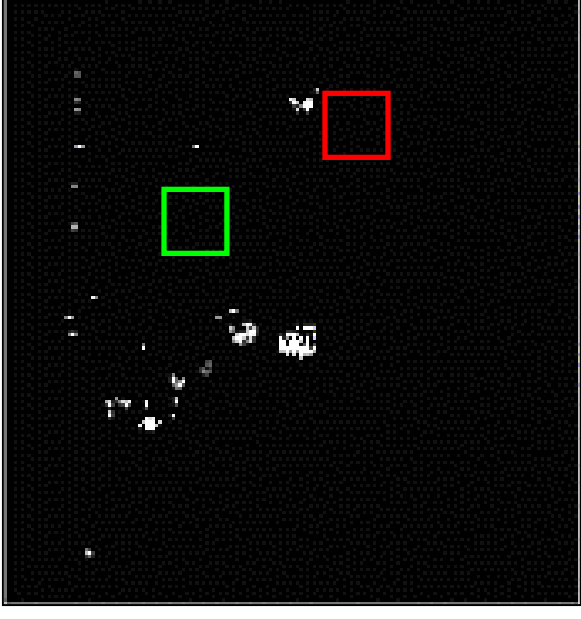
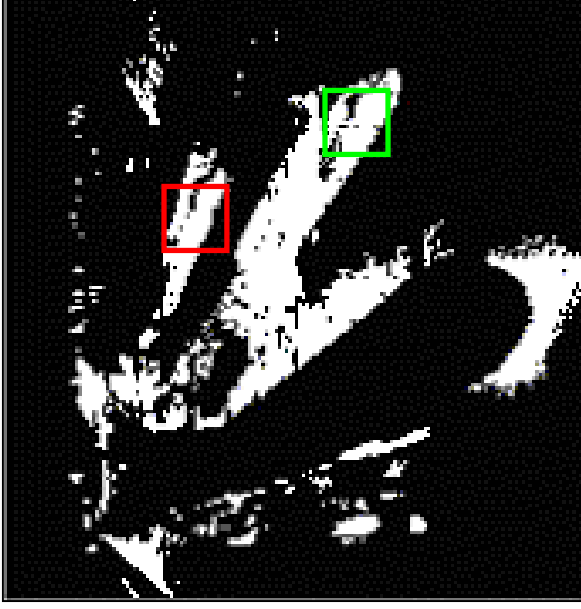
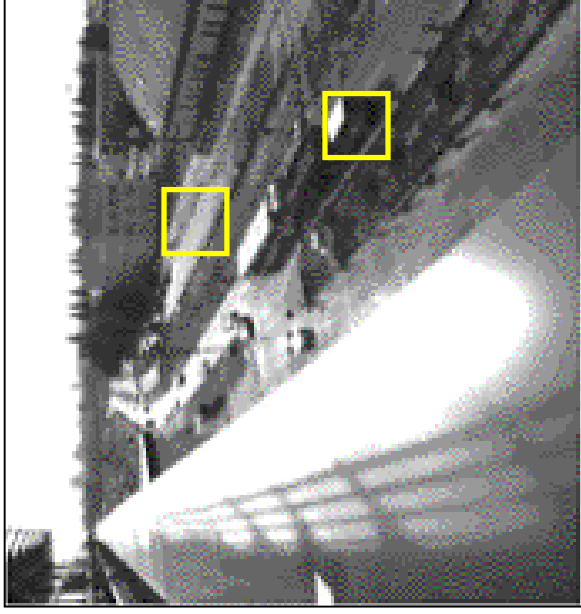
Hintergrundstatistik soll
schnellen
Beleuchtungsänderungen
nachgeführt werden



über einen längeren Zeitraum
unveränderte Objekte sollen nur
sehr langsam adaptiert werden

- problematisch sind sehr langsame Bewegungen von Fahrzeugen, dadurch gehen die Randpixel der Objekte auch kontinuierlich von Hintergrundfarbe zu Objektfarbe über und werden adaptiert
- **Hintergrund unter stehenden Objekten wird bei Beleuchtungsänderungen nicht adaptiert**
 - Änderung des Hintergrundes kann aus dem Verhalten der nichtverdeckten Umgebung abgeleitet werden.

Unzulänglichkeiten des Multi-Gaußverfahrens (2)



links: Originalbild

mitte: Lernrate zu klein, Beleuchtungsänderungen werden nicht adaptiert

rechts: Lernrate zu groß, Fahrzeuge werden adaptiert

Modifikationen

- **Splitting der Adaptionsgeschwindigkeit von Mittelwert / Streuung und Gewichten:**
 - dadurch wird eine Adaption an Beleuchtungsänderungen möglich, ohne dass stehende Fahrzeuge adaptiert werden (große Lernrate für Mittelwert und kleine Lernrate für Gewichte)
- **Änderung der Updatefunktion:**
 - Verzicht auf Exponentialfunktion \Leftrightarrow schnellere Adaption stärkerer Abweichungen, weniger Rechenaufwand
 - wird möglich, da nur Werte adaptiert werden, die auch in der Nähe des Mittelwertes liegen, und da die Werte selbst eine gaußähnliche Statistik aufweisen
- **Nachführung verdeckter Hintergrundpixel durch Referenzpixel:**
 - Auswahl von Referenzpixeln in der Umgebung verdeckter Pixel
 - Auswahlkriterium ist ähnlicher Grauwert

Splitting der Lernraten / Änderung der Updatefunktion

Gewichte:

$$W_{k,t} = \begin{cases} W_{k,t-1} + \alpha & \text{Treffer} \\ W_{k,t-1} & \text{sonst} \end{cases}$$

anschließend Normierung der Gewichte
auf Summe == 1

Mittelwert:

$$\mu_t = (1 - \beta)\mu_{t-1} + \beta x_t$$

Streuung:

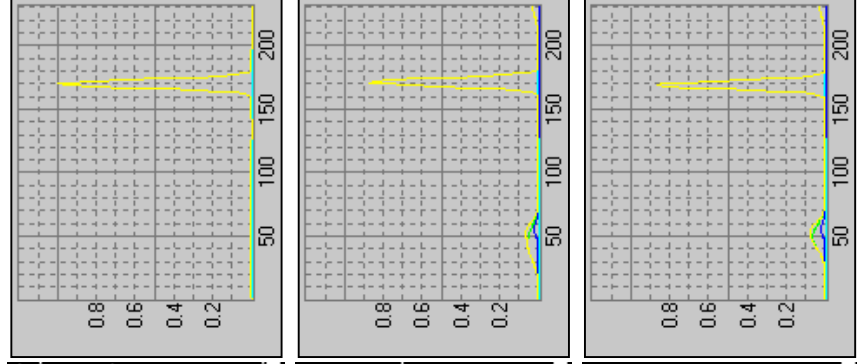
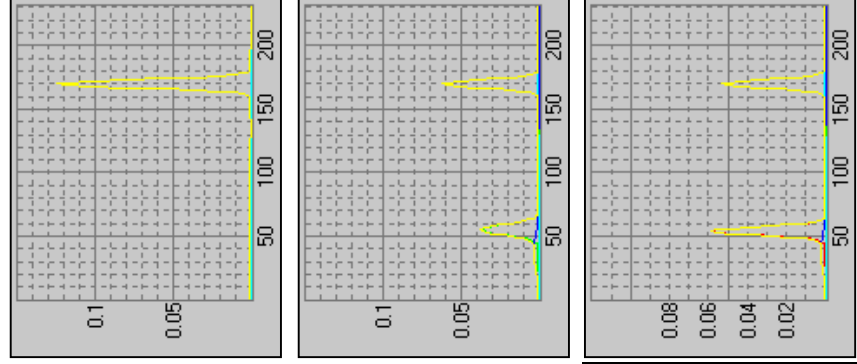
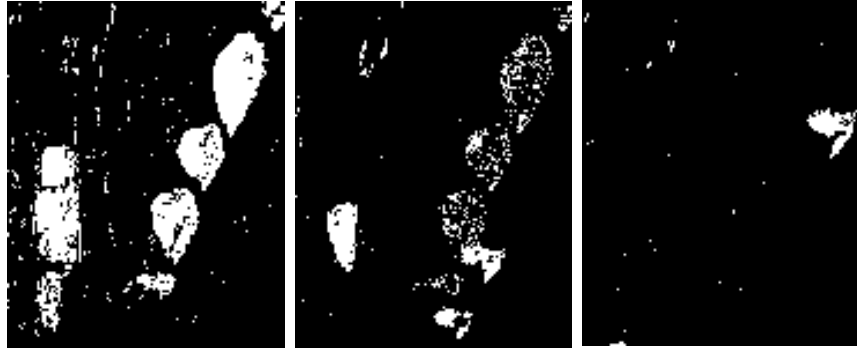
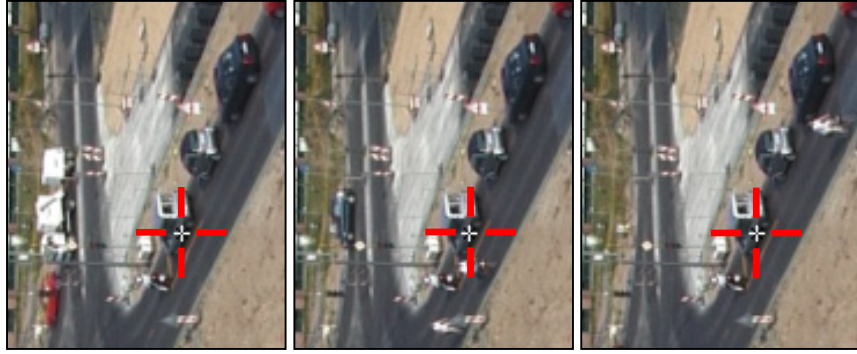
$$\sigma_t^2 = (1 - \beta)\sigma_{t-1}^2 + \beta(x_t - \mu_t)^T(x_t - \mu_t)$$

unabhängige Lernraten α und β

Splitting der Lernraten

Spalten: Originalbild, Multi-Gauß-Originalansatz, Multi-Gauß modifiziert (jeweils Segmentierungsergebnis und Verteilungen eines Einzelpixels (rote Markierung))

Zeilen: Zustand zum Zeitpunkt Bild 0, Bild 60, Bild 80



Nachführung verdeckter Pixel (1)

- **jedem Pixel werden Referenzen aus seiner Umgebung zugewiesen**
- **wenn Objekt detektiert ist:**
 - durchschnittliche Mittelwertänderung der unverdeckten Referenzen auf die Verteilung anwenden, die das größte Gewicht besitzt (aktuelle Hintergrundverteilung)
- **schnelle Helligkeitsänderungen führen nicht mehr zu Geisterobjekten, wenn sich stehende Objekte wieder bewegen**
- **Zuordnung von Referenzen:**
 - Referenzpixel werden an zufälligen Positionen (gaußverteilt mit aktuellem Pixel als Zentrum) ausgewählt
 - Auswahlkriterium ist Ähnlichkeit zum Grauwert am betrachteten Ort (Annahme: Pixel mit ähnlichen Grauwerten sind ebenfalls HG-Pixel und haben auch ähnliches Verhalten)
 - Verwerfen und Neusuchen von Referenzen, wenn der Grauwertabstand zum Zentralpixel zu groß wird

Nachführung verdeckter Pixel (2)

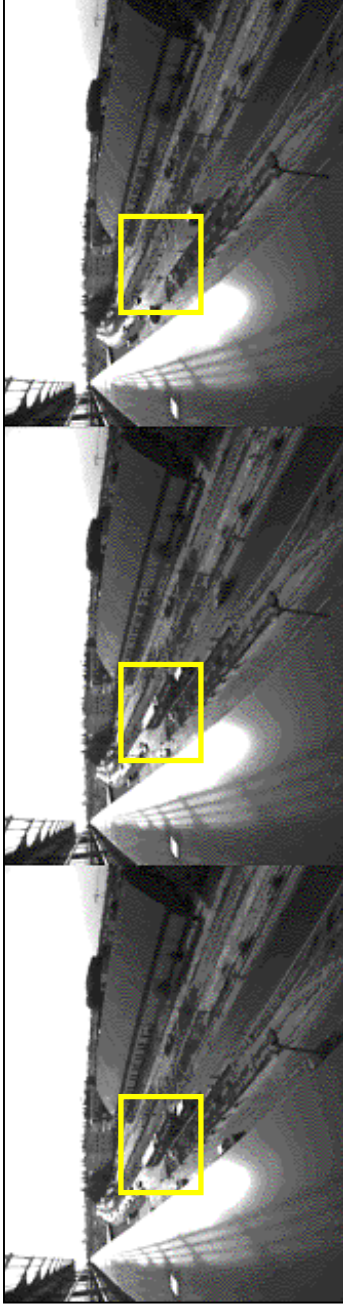


blau: Suchradius,
grün: Zentralpixel,
rot: gefundene
Referenzen
mit ähnlichem
Grauwert

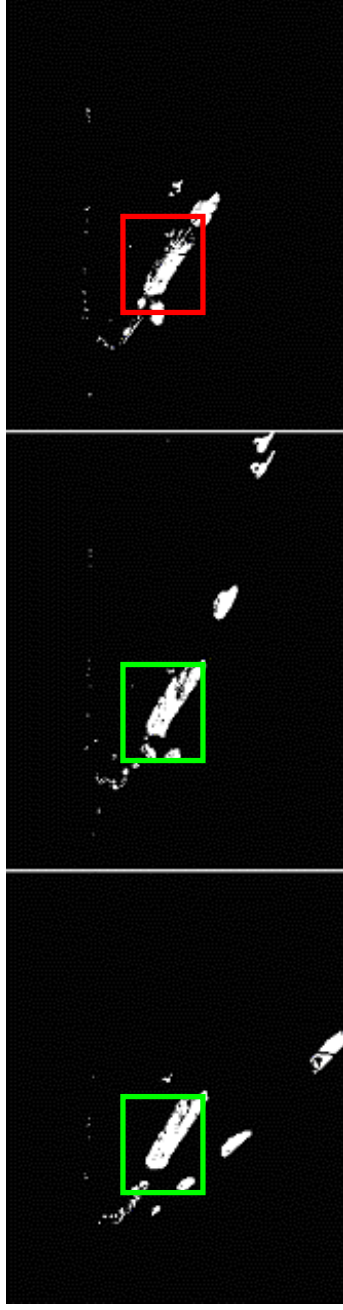
Beispiel für Pixelkorrespondenzen zur Nachführung verdeckter Pixel

Nachführung verdeckter Pixel (3)

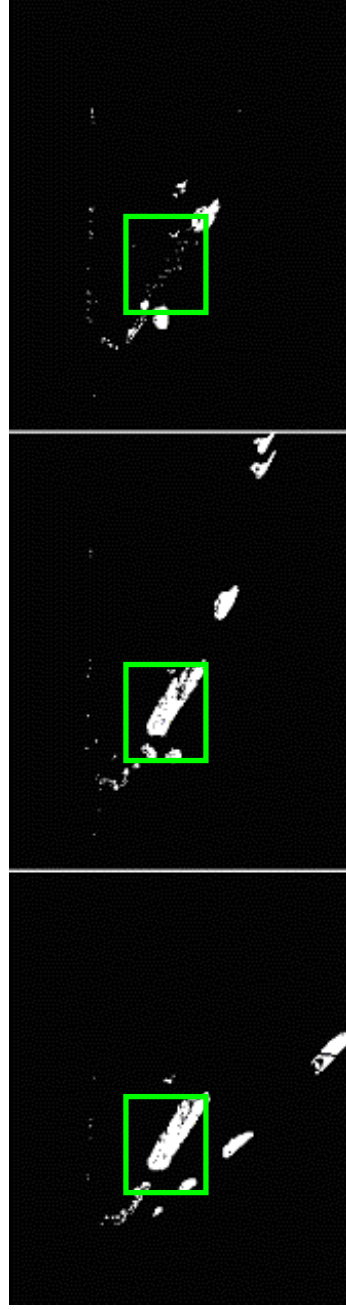
Spalte 1 und 2: Stillstand an Signalanlage, Spalte 3: Weiterbewegung in der Grünphase



Zeile 1: Bildserie mit Beleuchtungsänderung (Bildabstand 60 Bilder)



Zeile 2: keine Nachführung der Beleuchtungsänderung: nach der Weiterbewegung des Fahrzeugs verbleiben fehlsegmentierte Bereiche



Zeile 3: Nachführung der Beleuchtungsänderung, keine fehlsegmentierten Bereiche nach der Weiterbewegung der Fahrzeuge

Verhalten bei extremem Beleuchtungswechsel (1)

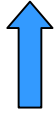


sehr schneller und starker Beleuchtungswechsel:

**blauer Himmel,
schnell vorbeiziehende
und scharf konturierte
Wolken,
z.B. Rückseitenwetter,
Zeitkonstanten: < 5s
bei 5 Bilder / s**

Verhalten bei extremem Beleuchtungswechsel (2)

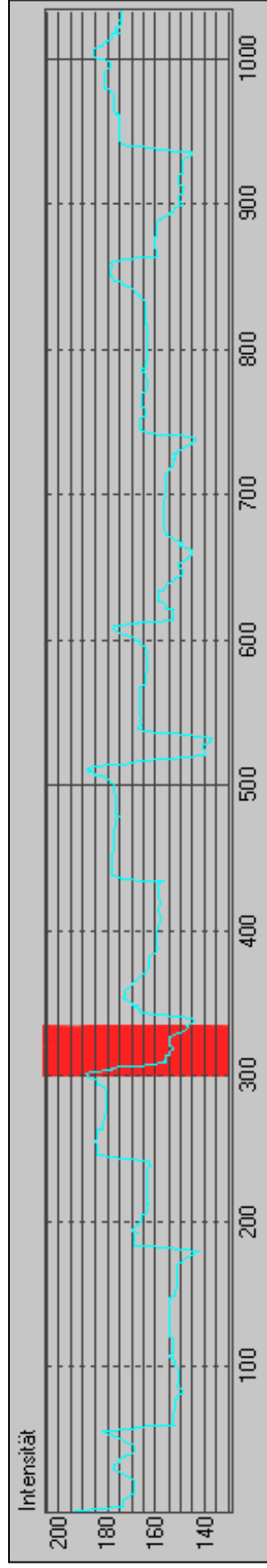
sehr schneller und starker
Beleuchtungswechsel



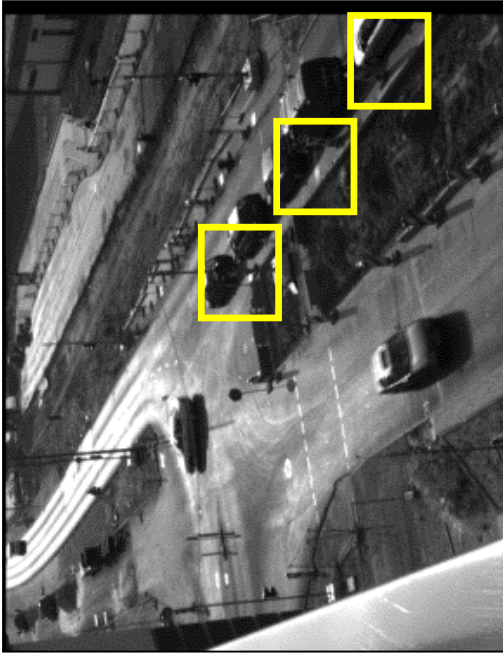
Grauwertverläufe von
Einzelpixeln bei
Beleuchtungswechsel und
langsam fahrenden
Fahrzeugen sind nicht
mehr unterscheidbar



Beispielbilder und Intensitätsdaten entstammen den rot markierten Bereichen

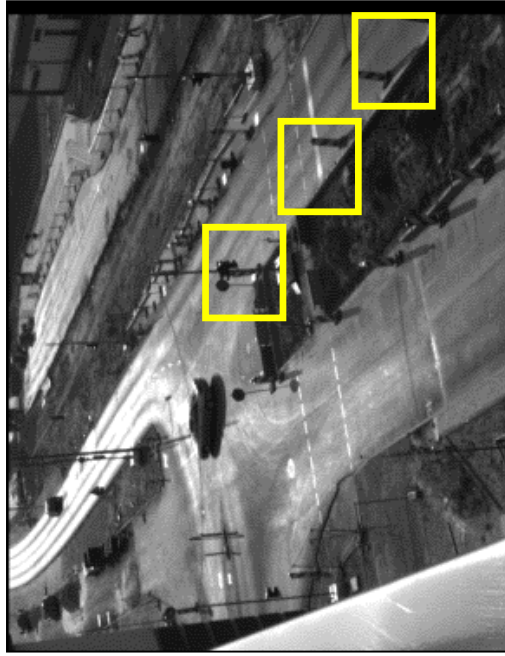


Adaptionerscheinungen an Objektändern



bei sehr langsamfahrenden Fahrzeugen an geraden Objektändern, die parallel zur Fahrtrichtung ausgerichtet sind:

die Grauwerte im Übergangsbereich zwischen Fahrzeug und Hintergrund ändern sich



langsam und kontinuierlich und werden adaptiert, verlässt die gerade Objektkante die betreffenden Pixel dann wird der Sprung vom adaptierten Grauwert zum Hintergrundgrauwert als Objekt interpretiert: Geisterkanten

Lösungsmöglichkeiten

Verhalten bei extremem Beleuchtungswechsel

- ▶ Steuerung der Lernrate anhand globaler Beleuchtungsänderung im aktuellen Bild (Erhöhen der Lernrate nur bei kritischem Wetter)
- ▶ dafür werden im Bild in unkritischen Bereichen ‚Sensoren‘ platziert, unkritisch sind Bereiche in denen sich keine potentiellen Objekte bewegen

Adaptionerscheinungen an Objekträndern

- ▶ nichtverdeckte Pixel an Objekträndern werden wie verdeckte Pixel behandelt: keine Adaption sowie Nachführung von Beleuchtungsänderungen

Ergebnisse / Ausblick

- ▶ **Realisierung eines Verfahrens zur Segmentierung von Verkehrsszenen: modifizierter Multi-Gauß-Ansatz mit folgenden Eigenschaften:**
 - **selbstadaptives Verfahren (Lernphase automatisch), Beherrschen periodischer Änderungen des Hintergrunds**
 - **Bewältigung unterschiedlicher Zeitkonstanten (Stausituation vor Ampeln, Beleuchtungsänderungen in Outdoorszenen)**
 - **Adaption auch bei verdecktem Hintergrund**
 - **Laufzeit (AMD Athlon, 2.1 GHz, Bildgröße 748 x 576): 100ms**
- ▶ **Ausblick / Erfordernisse:**
 - **Lösung der Problematik extrem schneller Beleuchtungsänderungen**
 - **Vermeidung von Adaptionerscheinungen an Objekträndern**