

Personendetektion in Grauwertbildern eines Stereo-Fahrzeugkamarasystems in Nachtsituationen mit variierenden Beleuchtungsverhältnissen

ALEXANDER HANEL¹, HORST KLÖDEN², LUDWIG HOEGNER¹ & UWE STILLA¹

Zusammenfassung: In Fahrzeugen der neuesten Generation werden Stereobilder vom Fahrweg des Fahrzeugs bei Tageslicht zur Detektion von Personen im Fahrweg genutzt. In diesem Beitrag wird untersucht, ob ein derartiger Detektionsalgorithmus für Bilder aus der Nacht geeignet ist. Dazu wird für verschiedene Beleuchtungssituationen je ein separates Detektormodell angelernt und ausgewertet. Weiter wird in diesem Beitrag eine Methode vorgestellt, mit der Falschdetektionen anhand der Disparitätskarte erkannt und verworfen werden können. Dazu wird die metrische Größe jeder Detektion bestimmt und die Plausibilität anhand realer Größenangaben von Personen geprüft. In einer Messfahrt wurden für verschiedene Beleuchtungssituationen am Tag und in der Nacht Stereobilder des Fahrwegs aufgezeichnet. Es hat sich gezeigt, dass der Detektor durch separate Modelle für Beleuchtungssituationen bei Nacht eine hohe Detektionsgüte mit einer mittleren Präzision über 65% erreichen kann. Dies ist wahrscheinlich auf die geringe Varianz des Erscheinungsbildes von Personen bei Nacht zurückzuführen. Die geringe Detektionsgüte bei Tag mit unter 30% lässt sich durch Plausibilitätsprüfung auf annähernd 80% steigern.

1 Objektdetektion im automobilen Umfeld

Heutige Fahrzeuge besitzen eine Vielzahl von Fahrerassistenzsystemen zur Verhinderung von Unfällen oder zur Verringerung ihrer Auswirkungen. Wesentliche Voraussetzung für solche Systeme ist es, zuverlässige Informationen über Objekte im Fahrweg eines Fahrzeugs zu besitzen. Insbesondere Fußgängern als leicht verletzliche Verkehrsteilnehmer kommt bei der Auslegung solcher Systeme eine hohe Bedeutung zu. Deswegen werden häufig Detektionsalgorithmen verwendet, die Personen in Bildern des Fahrwegs detektieren. Die neueste Generation von Fahrzeugen sind zur Erfassung des vorausliegenden Fahrzeugumfelds dazu oft mit einem Stereokamerasystem ausgestattet (Abb. 1).

Heute verwendete Detektionsalgorithmen sind in der Regel für das Erscheinungsbild von Personen bei Tageslicht ausgelegt. Bei Nacht stehen mit derartigen Algorithmen arbeitende Fahrerassistenzfunktionen dem Fahrer nicht zur Verfügung. Weiterhin können Bilder analysiert werden, um die Beleuchtungssituation der abgebildeten Straßenverkehrsszene zuverlässig zu bestimmen (HANEL et al. 2015). Dadurch wird es möglich, Detektionsalgorithmen durch verschiedene Parametersätze an die vorherrschende Beleuchtungssituation anzupassen und somit auch für Nachtfahrten auszuliegen.

¹ Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstraße 21, D-80333 München, E-Mail: [alexander.hanel, ludwig.hoegner, stilla]@hcu-hamburg.de

² BMW Forschung und Technik GmbH, Hanauer Straße 46, D-80992 München, E-Mail: horst.kloeden@bmw.de



Abb. 1: Unfallvermeidung durch Detektion eines Fußgängers im Fahrweg des Autos mit Hilfe eines Stereokamerasystems am Innenspiegel [CONTINENTAL, 2011]

Zur Einschätzung des Kollisionsrisikos mit einem Fußgänger spielt dessen Abstand vom Auto eine wesentliche Rolle. Immer leistungsfähigere Prozessoren ermöglichen es heutzutage, statt Monobilddaten auch Stereobilddaten während der Fahrt zu verarbeiten. Daraus können mit einer Disparitätskarte Personen mit geringem Abstand vor dem Fahrzeug detektiert werden. Noch effizienter kann die Entfernung bestimmt werden, wenn die Disparitätskarte zur Nachverarbeitung von Detektionen im Grauwertbild verwendet wird.

2 Arbeiten im Umfeld

2.1 Bilddeskriptoren

Zur Detektion von Objekten in Bildern existiert eine große Bandbreite an Verfahren. Diese Verfahren beinhalten typischerweise einen Deskriptor, um Merkmale von zu detektierenden Objekten aus dem Bild abzuleiten, und einen Klassifikator, um Merkmale verschiedener Objekttypen voneinander zu unterscheiden. Ein Beispiel für ein solches Verfahren ist der HOG-SVM-Detektor (DALAL & TRIGGS 2005). Der Merkmalsvektor des HOG-Deskriptors (engl. *histogram of oriented gradients*) setzt sich aus gewichteten Richtungsangaben der dominierenden Grauwertkanten der rasterförmig angeordneten Bildteile zusammen. Eine zeiteffiziente (KITTI 2016) Erweiterung des HOG-Deskriptors sind ACF-Merkmale (engl. *aggregated channel features*), die mit AdaBoost trainiert werden (DOLLAR et al. 2014). ACF-Merkmale aggregieren sich aus mehreren Kanälen des Bildes. Dazu zählen die LUV Farbkanäle, sechs HOG-Merkmale und die Gradientenstärke.

2.2 Klassifikationsverfahren

Klassifikationsverfahren lassen sich in überwachte und unüberwachte Verfahren unterscheiden. Bei überwachten Verfahren muss beim Anlernen des Klassifikatormodells die tatsächliche Klasse der Trainingsdaten bekannt sein (z. B. Bayes-Klassifikator, FUKUNAGA 1990), während unüberwachte Verfahren aus den Trainingsdaten Cluster bilden (z. B. Expectation Maximization, BILMES 1998). Zur Objektdetektion wird häufig die Support-Vektor-Maschine (JOACHIMS 1999) als überwachtes Klassifikationsverfahren eingesetzt. Sie erlaubt die Bestimmung der Trennebene

zur Unterscheidung von Merkmalsvektoren verschiedener Objekttypen auch für eine große Menge Trainingsdaten in akzeptabler Zeit.

Zur Beurteilung von Klassifikationsverfahren können aus der Zahl der Korrekt-, Falsch- und Fehlklassifikation Precision und Recall berechnet werden (SOKOLOVA & LAPALME 2009). Precision setzt die Zahl der korrekt detektierten Objekte ins Verhältnis zu allen detektierten Objekten und bewertet somit Falschklassifikation. Recall hingegen setzt die Zahl der korrekt detektierten Objekte ins Verhältnis zu allen tatsächlichen Objekten (Ground Truth) und bewertet somit Fehlklassifikation.

2.3 Stereosehen

Algorithmen zur Berechnung von Tiefeninformationen aus Bildern eines Stereokamerasystems sind Stand der Technik (z. B. SCHARSTEIN & SZELISKI 2002). Typischerweise werden zunächst die Bilder beider Kameras mit den Parametern der Stereokamerakalibrierung rektifiziert. Über Punkt-zu-Punkt-Korrespondenzen eines zeitsynchron aufgenommenen Bildpaars kann die Disparitätskarte berechnet werden, aus der Tiefenwerte abgeleitet werden können.

3 Personendetektion in Bildern des Straßenverkehrs bei Nacht

In diesem Beitrag wird ein Personendetektor nach dem HOG-SVM-Ansatz (Kapitel 2) verwendet. Dieser eignet sich zum Beispiel für Bilder, die den Fahrweg eines Fahrzeugs aus der Perspektive einer Kamera am Innenspiegel (vgl. Abb.1) zeigen. Dieser Abschnitt gliedert sich in zwei Teile. Im ersten Teil wird die Vorgehensweise zur Untersuchung der Eignung des Personendetektors für Nachtaufnahmen unterschiedlicher Beleuchtungssituationen beschrieben. Im zweiten Teil wird eine Methode vorgestellt, mit der Detektionen von Personen im Bild anhand der Disparitätskarte auf Plausibilität überprüft werden können. Prinzipiell lassen sich die folgenden Beschreibungen auch auf andere Objekttypen übertragen.

3.1 Spezifische Klassifikatormodelle für jede Beleuchtungssituation

Der Begriff Beleuchtungssituation bezeichnet in diesem Beitrag die Beleuchtung einer Szene (HANEL et al. 2015). Mehrere Beleuchtungsquellen unterschiedlicher Position, Richtung und Stärke tragen zur Beleuchtungssituation bei. Verschiedene Beleuchtungssituationen werden unterschieden („Beleuchtungsklassen“). Für jede Beleuchtungsklasse wird ein separates Klassifikatormodell des Personendetektors angelehrt. Zur Detektion von Personen in einem Bild wird dieses rasterförmig abgetastet und jeder entnommene Bildausschnitt mit dem Detektor ausgewertet. Für jeden Bildausschnitt wird zuvor die Beleuchtungsklasse bestimmt (HANEL et al. 2015), um das passende Klassifikatormodell anwenden zu können.

Die Eignung des Personendetektors wird durch einen Vergleich von in Testbildern detektierten Personen mit Ground Truth Daten bestimmt. Der Detektor berechnet dazu für jede detektierte Person die Eckkoordinaten eines umschließenden Rechtecks (Abb. 2a). Jede Detektion weist einen Sicherheitswert auf, der den Abstand des Merkmalsvektors von der Klassengrenze im Merkmalsraum beschreibt (LIBSVM 2016). Liegt der Sicherheitswert über einer festgelegten absoluten Sicherheitsschwelle, wird eine Detektion akzeptiert. Als Ground Truth werden im Bild durch ein umschließendes Rechteck manuell markierte Personen verwendet. Als

Korrektklassifikation wird eine Detektion gewertet, sobald sich mehr als ein festzulegender Anteil der Rechtecksfläche von Detektion und Ground Truth überlappen. Um unterschiedlich groß abgebildete Personen detektieren zu können, wird jedes Testbild in eine Bildpyramide zerlegt.



Abb. 2: Bildausschnitte zum Anlernen eines Personendetektors: a) positiver Bildausschnitt mit Person, b) negativer Bildausschnitt ohne Person.

Zur Beurteilung der Detektionsgüte werden für einen Testdatensatz die mittlere Precision sowie die Precision-Recall-Kurve (PR-Kurve) abgeleitet. Jeder Punkt in dieser Kurve repräsentiert den Precision- und Recall-Wert für eine bestimmte Sicherheitsschwelle. Die Sicherheitsschwelle wird dazu in einem festzulegenden Wertebereich variiert. Je höher die Sicherheitsschwelle, desto weniger Falschklassifikationen, desto mehr Fehlklassifikationen. Die PR-Kurve kann somit genutzt werden, um für die produktive Anwendung des Detektors eine optimale Sicherheitsschwelle zu wählen, bei der Precision und Recall möglichst hoch sind.

3.2 Plausibilitätsprüfung von Detektionen im Bild

Im Folgenden wird eine Methode zur Überprüfung der Plausibilität einer Detektion im Bild mit Hilfe der Disparitätskarte beschrieben. Diese Methode erlaubt es, Falschklassifikationen zu erkennen und zu verwerfen. Dadurch kann für den Detektor eine niedrige Sicherheitsschwelle mit wenigen Fehlklassifikationen gewählt und somit die Detektionsgüte insgesamt gesteigert werden. Voraussetzung zur Plausibilitätsprüfung ist ein Stereokamerasystem, bei dem eines der beiden rektifizierten Bilder zur Personendetektion verwendet wird und die Disparitätskarte abgeleitet werden kann (Abb. 3).

Zur Plausibilitätsprüfung wird für jede zuvor im Bild detektierte Person die metrische Höhe des umschließenden Rechtecks ermittelt. Liegt dieser Wert im Vergleich mit realen Größenangaben von Personen außerhalb eines Akzeptanzintervalls, so wird diese Detektion verworfen. Da Falschdetektionen oft eine unplausible Höhe des umschließenden Rechtecks aufweisen, reduziert sich deren Anzahl.



Abb. 3: Beispielaufnahme (aufgehellt) eines Fahrzeug-Stereokamerasystems in der Nacht. Links: Rektifiziertes Grauwertbild, rechts: Disparitätsbild.

Zunächst werden die Eckkoordinaten des umschließenden Rechtecks vom rektifizierten Stereobild in die Disparitätskarte projiziert. Im Fall einer korrekten Detektion zeigt der Bildausschnitt innerhalb des umschließenden Rechtecks die detektierte Person sowie an jedem Seitenrand Objekte aus dem Bildhintergrund. Um die metrische Größe der Person abzuschätzen, muss deren Entfernung von der Kamera anhand der Disparitätskarte berechnet werden. Dazu wird innerhalb des umschließenden Rechtecks der Median aller Disparitätswerte, die zu den Pixeln der abgebildeten Person gehören, gebildet und anhand der Stereokalibrierparameter in einen Entfernungswert umgerechnet. Für die Person und den Bildhintergrund werden im Histogramm der Disparitäten dazu unterschiedliche, aber dominante Werte angenommen. Durch Bestimmung des markantesten Minimums im Histogramm zwischen diesen Maxima (Peakiness-Test, vgl. HOEGNER et al. 2014) wird eine absolute Schwelle bestimmt, anhand derer zum Bildhintergrund gehörende Pixel erkannt werden können (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Die Höhe des umschließenden Rechtecks wird mit Hilfe des berechneten Entfernungswertes ermittelt und mit Größenangaben realer Personen verglichen. Für die Größenangabe wird eine untere sowie obere Schwelle gewählt, wobei berücksichtigt wird, dass die Höhe des umschließenden Rechtecks ebenfalls den Bildhintergrund beinhaltet. Detektionen mit nicht plausibler metrischer Höhe des umschließenden Rechtecks werden verworfen.

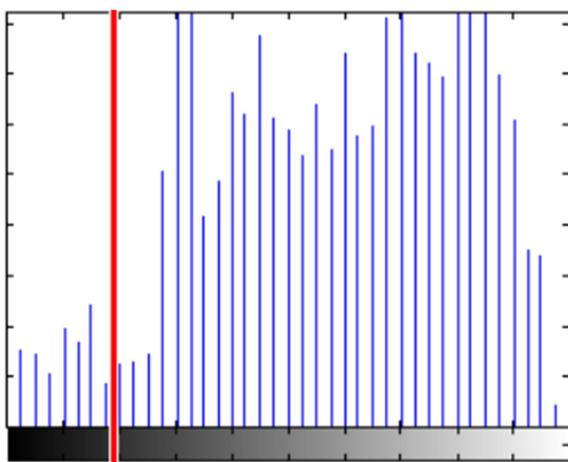


Abb. 4: Histogramm der Disparitätswerte des umschließenden Rechtecks einer Detektion. Disparitäten durch 8bit-Grauwerte codiert auf horizontaler Achse dargestellt. Markantestes Minimum zwischen Disparitätsmaximum von Person und Bild-hintergrund in rot.

4 Experiment

Es wurden am Tag und in der Nacht Messfahrten auf städtischen und ländlichen Straßen durchgeführt. Ein Stereokamerasystem hat Grauwertbilder vom Fahrweg vor dem Fahrzeug durch die Windschutzscheibe aufgezeichnet. Zum Anlernen des Klassifikators werden für jede Beleuchtungsklasse aus den Bildern etwa 10.000 positive und 30.000 negative Bildausschnitte mit beziehungsweise ohne Einzelpersonen extrahiert (Abb. 2). Das ungleiche Verhältnis zwischen positiven und negativen Stichproben ist typisch für Personendetektoren (DALAL & TRIGGS 2005). Positive Bildausschnitte sind durch manuelles Markieren der Personen im Bild vorgegeben, während negative Bildausschnitte zufällig aus Bildern des Fahrzeugkamerasystems ohne abgebildete Personen selektiert wurden. Die Bildausschnitte wurden per Algorithmus absteigend in vier Beleuchtungsklassen (HANEL et al. 2015) eingeteilt: *Tag*, *Nacht-Stadtzentrum* (Ego-Fahrzeug-, Gegenverkehr- und Straßenbeleuchtung), *Nacht-Wohngebiet* (Ego-Fahrzeug- und Straßenbeleuchtung) und *Nacht-Landstraße* (nur Ego-Fahrzeugbeleuchtung).

Die aufgenommenen Stereobilder wurden rektifiziert sowie die Disparitätskarte abgeleitet. Der Personendetektor wurde für jede Beleuchtungsklasse angelemt. Die Eignung des Detektors wird für die Beleuchtungsklasse *Tag* durch 204 Bilder mit 105 abgebildeten Personen, für *Nacht-Wohngebiet* durch 278 Bilder und genauso vielen Personen, für *Nacht-Stadtzentrum* durch 225 Bilder mit 200 Personen und für *Nacht-Landstraße* durch 208 Bilder mit 130 Personen evaluiert. Anschließend wird für jede Detektion die Plausibilitätsprüfung durchgeführt.

5 Ergebnisse und Diskussion

In diesem Abschnitt werden zunächst die Ergebnisse der Personendetektion mit HOG-SVM für die unterschiedlichen Beleuchtungsklassen diskutiert. Grundlage bildet für jede Beleuchtungsklasse die mittlere Precision (Tab. 1) sowie die PR-Kurve (Abb. 5).

Tab. 1: Mittlere Precision des Personendetektors für jede Beleuchtungsklasse vor (mittlere Spalte) und nach (rechte Spalte) der Plausibilitätsprüfung.

Beleuchtungsklasse	Mittlere Precision	Mittlere Precision nach Plausibilitätsprüfung
Wohngebiet	83,80%	87,04%
Stadtzentrum	65,19%	81,89%
Landstraße	65,05%	73,79%
Tag	22,97%	78,82%

Es zeigt sich, dass die mittlere Precision des Personendetektors bei Tagaufnahmen deutlich geringer ist als bei Nachtaufnahmen. Während die Werte der mittlere Precision der Nachtaufnahmen weit über 50% liegen, befindet er sich für Tagaufnahmen weit darunter. Ursache für diesen Unterschied kann sein, dass bei Nachtaufnahmen weniger Falschklassifikationen auftreten. Der durch die Dunkelheit nahezu schwarze Bildhintergrund weicht stärker vom Erscheinungsbild von Personen im Bildvordergrund ab, als der Bildhintergrund bei Tagaufnahmen

mit verschiedenen abgebildeten Objekten. Dadurch sinkt bei Nachtaufnahmen der Sicherheitswert von Detektionen in solchen Bildbereichen, wodurch diese verworfen werden.

Bei Vergleich der Nachtaufnahmen untereinander lässt sich feststellen, dass der Personendetektor für die Beleuchtungsklasse *Nacht-Wohngebiet* eine um 20 Prozentpunkte höhere mittlere Precision aufweist, als für die beiden anderen Nacht-Klassen. In der Klasse *Nacht-Stadtzentrum* variiert das Erscheinungsbild von Personen stärker als bei *Nacht-Wohngebiet*. Während in Wohngebieten die Straßenbeleuchtung regelmäßig ist, variiert die Beleuchtung im Stadtzentrum häufiger (Straßenbeleuchtung, Schaufenster, andere Fahrzeuge). Variierende Beleuchtung beeinflusst das Erscheinungsbild von Personen im Bild, was sich als negativer Einfluss auf die Detektionsgüte äußert. Die Bilder in *Nacht-Landstraße* hingegen sind nur durch den Fahrzeugscheinwerfer beleuchtet, wodurch sich Personen in dunklen Bildteilen befinden können und somit nicht detektierbar sind, was Fehlklassifikationen verursacht.

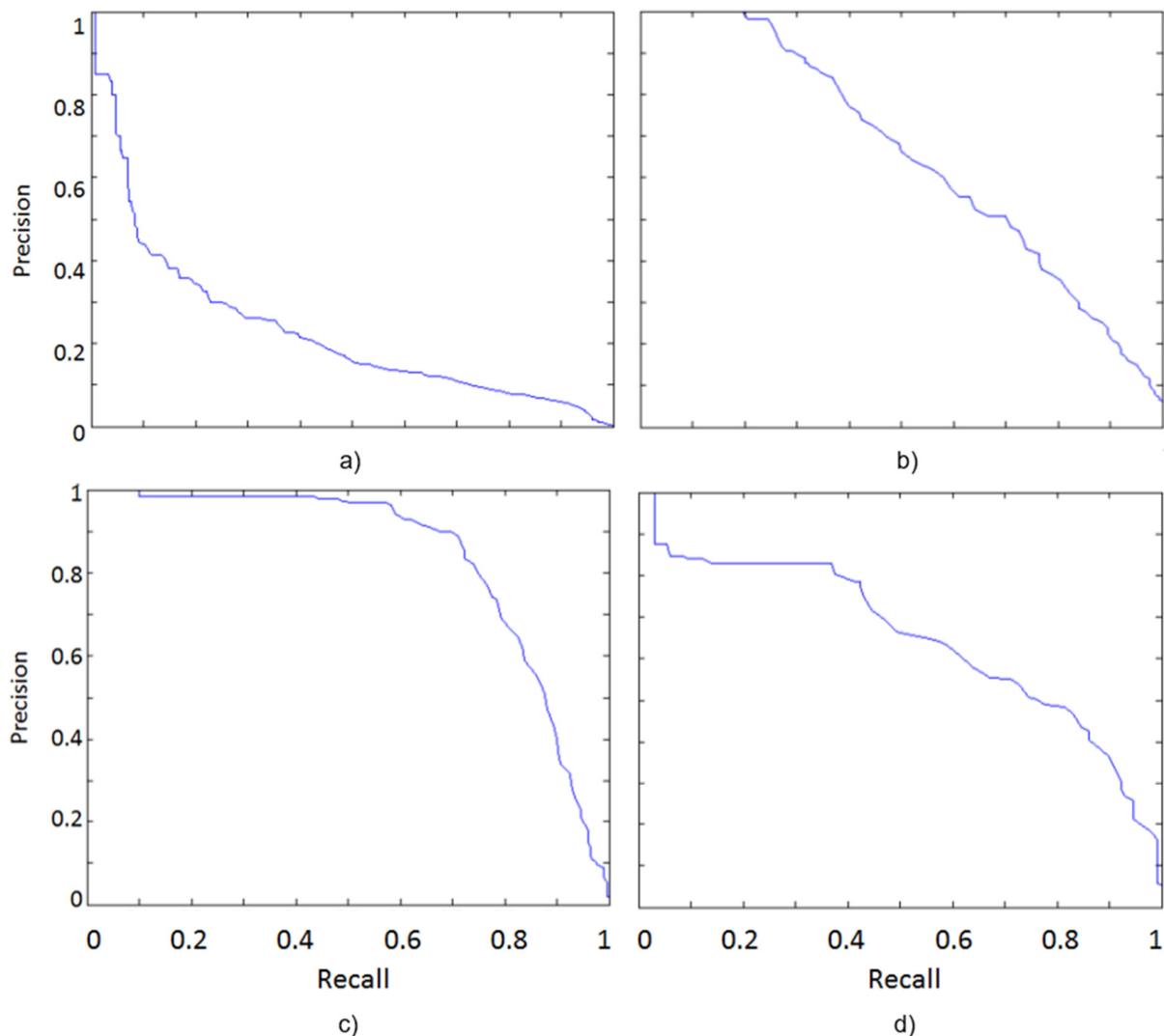


Abb. 5: PR-Kurven für den Personendetektor, a) Tag, b) Stadtzentrum, c) Wohngebiet, d) Landstraße.

Die PR-Kurve verläuft für *Nacht-Wohngebiet* am nächsten an den Asymptoten $Precision = 1$ beziehungsweise $Recall = 1$. Ersteres bedeutet, dass der Detektor für diese Beleuchtungsklasse die beste Precision aufweist, letzteres lässt die gleiche Feststellung für Recall zu. Es gibt somit bei der Wahl der Sicherheitsschwelle des Detektors ein Intervall der Recallwerte (0 bis ca. 0,6), für das keine oder kaum Falschklassifikationen in Kauf genommen werden müssen. Die optimale Sicherheitsschwelle ergibt sich jedoch an der Stelle mit dem größten Abstand der PR-Kurve vom Ursprung ($P \sim 0,8$; $R \sim 0,7$), wobei sowohl Fehl- als auch Falschklassifikationen in Kauf genommen werden. Etwas weniger asymptotisch verläuft die PR-Kurve für die beiden anderen Nacht-Beleuchtungsklassen. Es treten somit mehr Falsch- und Fehlklassifikationen auf. Am schlechtesten ist der Verlauf der PR-Kurve für Tag zu beurteilen, da sie am wenigsten asymptotisch verläuft. Es ist offensichtlich, dass in diesem Fall keine Sicherheitsschwelle gewählt werden kann, bei der nur eine der beiden Fehlerarten auftritt.

Bei den Nachtaufnahmen der Klassen *Wohngebiet* und *Landstraße* muss kritisch betrachtet werden, dass in den Trainings- und Testdaten aufgrund der geringen Passantenfrequenz nachts immer wieder dieselben Personen abgebildet sind. Es ist somit nicht auszuschließen, dass der Personendetektor nur für diese Personen eine hohe Güte liefert. Dieser Argumentation steht entgegen, dass die Güte für *Nacht-Stadtzentrum* vergleichbar ist, in deren Aufnahmen aber viele verschiedene Passanten zu sehen sind. Weiter muss noch angemerkt werden, dass neben der Beleuchtungsklasse auch die Wahl der Parameter des HOG-Deskriptors und des SVM-Klassifikators die Detektionsgüte beeinflussen.

Untersuchungen haben gezeigt, dass zur Personendetektion verschiedene Beleuchtungsklassen für Nachtaufnahmen unterschieden werden müssen. Ein einheitlicher Detektor für alle Nachtaufnahmen liefert schlechtere Resultate als der Detektor für Tagaufnahmen. Zusammengefasst zeigen die vorangegangenen Feststellungen, dass ein HOG-SVM-Personendetektor bei einer ausreichenden Zahl an Beleuchtungsklassen eine hohe Detektionsgüte für Nachtaufnahmen liefern kann.

Im Weiteren werden die Ergebnisse der Plausibilitätsprüfung besprochen. Für alle vier Beleuchtungsklassen steigt die mittlere Precision durch die Plausibilitätsprüfung (Tab. 1). Am deutlichsten ist der Anstieg bei der Klasse *Tag*, am geringsten bei *Nacht-Wohngebiet*. Das zeigt, dass gerade bei geringer mittlerer Precision die Plausibilitätsprüfung eine deutliche Verbesserung der Detektionsgüte bewirkt. Zurückzuführen ist diese Feststellung auf die gesunkene Zahl an Falschdetektionen (Abb. 6 und Abb. 7).

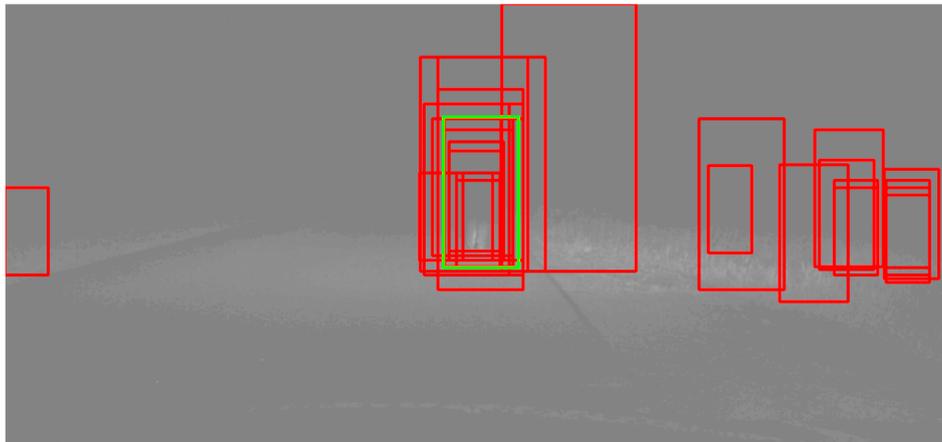


Abb. 6: Detektionen (rot) eines Bildes der Beleuchtungsklasse Nacht-Landstraße. Ground Truth einer Person in grün.

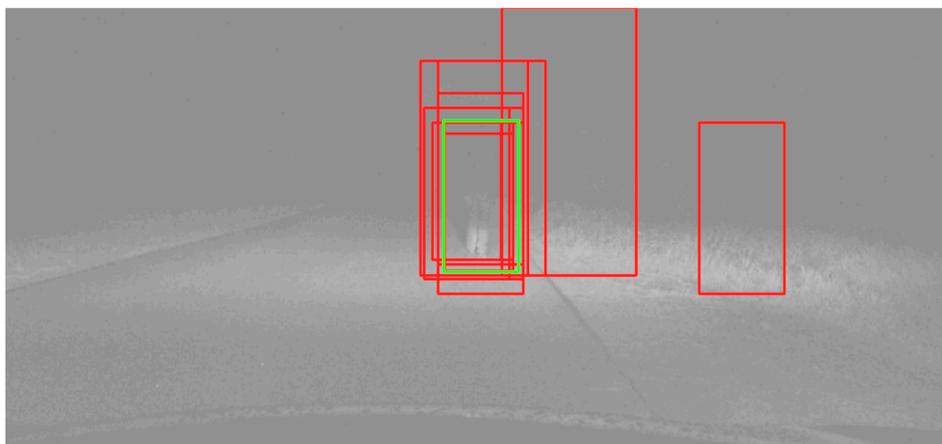


Abb. 7: Detektionen (rot) nach Plausibilitätsprüfung anhand der Disparitätskarte basierend auf den Detektionen aus Abb. 6.

Diese Beobachtung zeigt sich auch an den PR-Kurven. Für Tagaufnahmen verläuft die PR-Kurve (Abb. 8a) nun deutlich asymptotischer als ohne Plausibilitätsprüfung. Selbst für hohe Recall-Werte nimmt die Precision kaum ab. Es ist aber auch ersichtlich, dass immer ein paar wenige Falschdetektionen in Kauf genommen werden müssen (Precision erreicht nie 1). Bei der optimalen Sicherheitsschwelle sind Precision und Recall ($P \sim 0,75$; $R \sim 0,9$) besser als ohne Plausibilitätsprüfung. Eine ähnliche Feststellung lässt sich für *Nacht-Stadtzentrum* treffen (Abb. 8b). Es lässt sich ebenfalls eine Sicherheitsschwelle mit besserer Precision und Recall wählen ($P \sim 0,75$; $R \sim 0,75$). Im Gegensatz zu den Tagaufnahmen treten für diese Beleuchtungsklasse bei niedrigen Recall-Werten überhaupt keine Falschklassifikation auf. Ähnliche Feststellungen lassen sich für *Nacht-Landstraße* treffen. Bei *Nacht-Wohngebiet* sind die Auswirkungen der Plausibilitätsprüfung gering, da bereits vorher eine hohe Detektionsgüte vorgelegen hat. Die Plausibilitätsprüfung erlaubt somit – im Gegensatz zu einer Änderung der Sicherheitsschwelle – eine Reduktion der Falschklassifikationen, ohne zusätzliche Fehlklassifikationen in Kauf nehmen zu müssen.

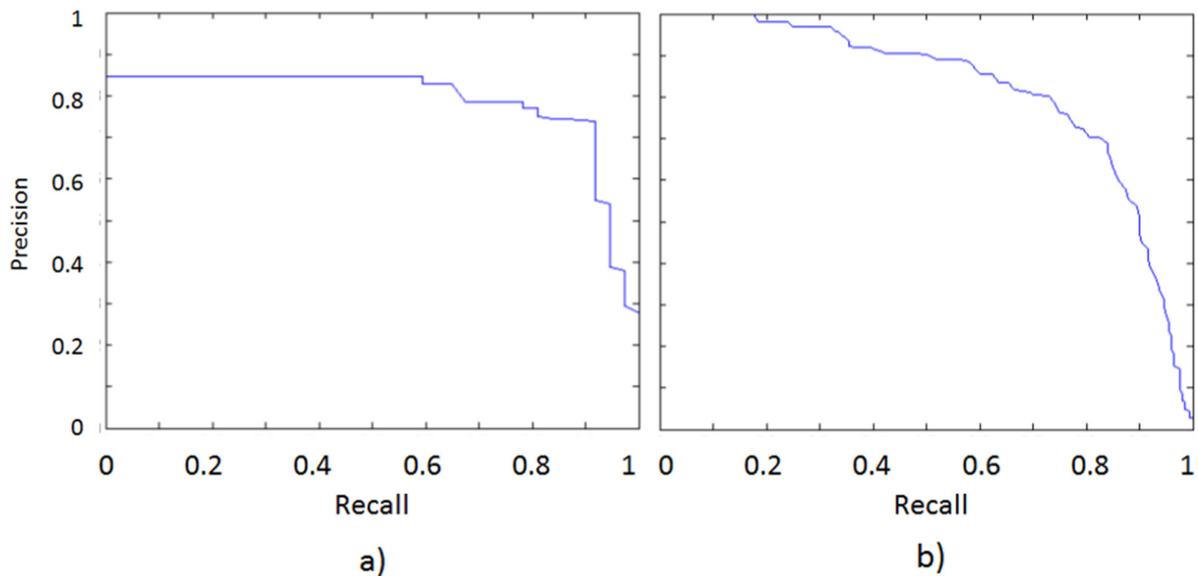


Abb. 8: Precision-Recall-Kurve nach Plausibilitätsprüfung, a) Tag, b) Stadtzentrum.

6 Fazit und Ausblick

Dieser Beitrag gliedert sich in zwei Teile. Zunächst wird ein Verfahren beschrieben, mit dem die Eignung eines HOG-SVM-Personendetektors für Nachtaufnahmen einer Fahrzeugkamera untersucht werden kann. Dazu werden verschiedene Beleuchtungssituationen der Straßenszene unterschieden und für jede Beleuchtungsklasse ein separates Klassifikatormodell angelernt. Die Analyse der Eignung geschieht anhand der mittleren Precision des Detektors sowie der Precision-Recall-Kurve. Im zweiten Teil wird ein Verfahren vorgestellt, mit dem Falschdetektionen in Bildern reduziert werden können. Aus der Disparitätskarte eines Stereokamerasystems wird die metrische Größe einer Detektion bestimmt und mit Größenangaben realer Personen verglichen. Mit einem Stereokamerasystem wurden Bilder vom Fahrweg eines Fahrzeugs bei Tag und Nacht für unterschiedliche Beleuchtungsklassen aufgenommen. Die Ergebnisse zeigen, dass für nächtliche Beleuchtungsklassen eine höhere Detektionsgüte erreicht werden kann als für Beleuchtungsklassen bei Tageslicht. Eine Unterscheidung verschiedener Beleuchtungsklassen bei Nacht ist notwendig, um eine hohe Detektionsgüte sicherzustellen. Die Plausibilitätsprüfung kann insbesondere für die Tagklasse mit zuvor schlechter Detektionsgüte eine deutliche Verbesserung auf das Niveau der Nachtklassen bewirken.

In künftigen Arbeiten kann untersucht werden, ob eine weitergehende Unterteilung der Beleuchtungsklassen die Personendetektion verbessern kann. Insbesondere die schlechte Detektionsgüte der Tagklasse lässt Optimierungsbedarf vermuten. Durch Tracken einer Person über mehrere Bilder einer Bildfolge können ebenfalls bessere Ergebnisse erwartet werden. Ein Stereo-Thermalkamerasystem könnte gerade bei Nacht zu einer verbesserten Disparitätsberechnung und dadurch besserer Plausibilitätsprüfung führen. Überdies kann es zusätzliche temperaturabhängige Merkmale zur Detektion von Personen liefern.

7 Literaturverzeichnis

- BILMES, J., 1998: A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report, University of California, Berkeley, Department of Electrical Engineering and Computer Science, Computer Science Division.
- CONTINENTAL, 2011: Mit der zweiten Kamera sieht man besser – die Stereokamera. Website, http://www.continental-corporation.com/www/presseportal_com_de/themen/pressemitteilungen/3_automotive_group/chassis_safety/press_releases/pr_20110504_stereokamera_de.html, abgerufen am 28.04.2016.
- DALAL, N. & TRIGGS, B., 2005: Histograms of oriented gradients for human detection. *IEEE Conference on Computer Vision and Pattern Recognition* **1**, 886-893.
- DOLLAR, P., APPEL, R., BELONGIE, S. & PERONA, P., 2014: Fast Feature Pyramids for Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36** (8), 1532-1545.
- HANEL, A., 2014: Bestimmen der Beleuchtungsverhältnisse und Optimieren der Fußgängerdetektion aus Bildern eines Fahrzeugkamarasystems. Masterarbeit, Technische Universität München, Ingenieur fakultät Bau Geo Umwelt, Fachgebiet Photogrammetrie und Fernerkundung.
- HANEL, A., KLÖDEN, H., HOEGNER, L. & STILLA, U., 2015: Erkennung von Szenenbeleuchtungen aus Bildern eines Fahrzeugkamarasystems zur Verbesserung der Objektdetektion. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **24**, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 254-262.
- HOEGNER, L., HANEL, A., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U., 2014: Towards people detection from fused time-of-flight and thermal infrared images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3), 121-126.
- FUKUNAGA, K., 1990: *Introduction to Statistical Pattern Recognition* (2nd Ed.). Academic Press Professional, Inc.
- JOACHIMS, T., 1999: Making large scale SVM learning practical. Technischer Bericht, Universität Dortmund.
- KITTI, 2016: Object Detection Evaluation 2012. Webseite, http://www.cvlibs.net/datasets/kitti/eval_object.php, abgerufen am 28.04.2016.
- LIBSVM, 2016. LIBSVM FAQ. Webseite, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/...faq.html#f4151>, abgerufen am 28.04.2016.
- SCHARSTEIN, D. & SZELISKI, R., 2002: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision* **47**(1), 7-42.
- SOKOLOVA, M. & LAPALME, G., 2009: A Systematic Analysis of Performance Measures for Classification Tasks. *Inf. Process. Manage.* **45** (4), 427-437.