



TECHNISCHE UNIVERSITÄT
CHEMNITZ

Pixelbasierte Datenfusion zur Steigerung der Objekterkennungsleistung im Fahrzeugumfeld

von der Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Chemnitz

genehmigte

Dissertation

zur Erlangung des akademischen Grades
Doktoringenieur (Dr.-Ing.)

vorgelegt von

Dipl.-Ing. (FH) Jan Thomanek
geboren am 7. Oktober 1973 in Mittweida

eingereicht am 21. November 2013

Gutachter: Prof. Dr.-Ing. Gerd Wanielik
Prof. Dr. rer. nat. Madhukar Chandra

Tag der Verleihung: 23. Juli 2014

Forschungsberichte der Professur Nachrichtentechnik
herausgegeben von Prof. Dr.-Ing. Gerd Wanielik

Band 10

Jan Thomanek

**Pixelbasierte Datenfusion zur Steigerung der
Objekterkennungsleistung im Fahrzeugumfeld**

D 93 (Diss. TU Chemnitz)

Shaker Verlag
Aachen 2014

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Chemnitz, Techn. Univ., Diss., 2014

Copyright Shaker Verlag 2014

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-3140-9

ISSN 1610-1251

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

*Für
Linda und Cindy*

Vorwort

Diese Arbeit ist das Ergebnis meiner mehrjährigen berufsbegleitenden Forschungstätigkeit an der Professur für Nachrichtentechnik der Technischen Universität Chemnitz.

Mein besonderer Dank gilt deshalb dem Leiter dieser Professur Prof. Dr.-Ing. Gerd Wanielik für seine Unterstützung. Ich danke ihm für das in mich gesetzte Vertrauen, die Herausforderung einer nebenberuflichen Promotion zu bewältigen. Er ließ mir den Freiraum, anspruchsvolle Forschungsaufgaben eigenständig zu bearbeiten und half mir durch seine bereichernden Hinweise und Anregungen wiederholt, mein Augenmerk in neue thematische Bahnen zu lenken.

Herrn Prof. Dr. rer. nat. Madhukar Chandra danke ich sehr für das Interesse an der Thematik und für die Bereitschaft der Übernahme eines Gutachtens. Für die fachliche Betreuung der Arbeit, den damit verbundenen gemeinsamen Diskussionen und der freundlichen, über die Arbeit hinausgehende Zusammenarbeit gilt mein außerordentlicher Dank Dr.-Ing. Basel Fardi.

Bei der IAV GmbH bedanke ich mich für die Möglichkeit, einer über mehrere Jahre dauernden Teilzeitbeschäftigung. Mein besonderer Dank gilt dabei meinem Abteilungsleiter Abdulvahap User und meinem Teamleiter Uwe Lohs, die mir die Freiheit hinsichtlich einer flexiblen Arbeitszeitgestaltung einräumten und damit den Spagat zwischen Beruf und Promotion ermöglichten. Des Weiteren möchte ich mich bei meinem Projektleiter Dr. rer. nat. Hadj Hamma Tadjine für die Mitarbeit an innovativen Themen aus dem Bereich der Fahrerassistenzsysteme, sowie meinem Kollegen Dr.-Ing. Christian Lang für die Durchsicht des Manuskripts bedanken.

Ebenso geht mein Dank an die jetzigen und ehemaligen Mitarbeiter der Professur für Nachrichtentechnik. Deren Unterstützung und die freundschaftliche Zusammenarbeit haben sehr zum Gelingen dieser Arbeit beigetragen.

Mein persönlicher Dank gilt meiner Familie und meinen Eltern für ihre Unterstützung und insbesondere für ihre Geduld und ihr Verständnis für wissenschaftliches Arbeiten an Wochenenden und nach Feierabend.

Kurzfassung

Derzeit erfahren Systeme zur Kombination von Daten unterschiedlicher Sensoren für die Fahrzeugumfeldererkennung eine wachsende Bedeutung. Eine solche Fusion kann bei bildgebenden Sensoren bereits auf der Pixelebene erfolgen. Ein fusioniertes Bild bietet aufgrund des höheren Informationsgehaltes entscheidende Vorteile für die weiterführende Bildverarbeitung, wie z.B. bei der Objekterkennung und der Objektklassifikation.

Gegenstand der Arbeit ist die Entwicklung eines geeigneten Fusionschemas, welches die Daten zweier oder mehrerer bildgebender Sensoren auf Pixelebene zu einem gemeinsamen Mischbild kombiniert. Das Fusionsverfahren wird dabei so gewählt, dass das fusionierte Bild eine optimale Basis für die Merkmalsgewinnung eines mobilen Objekterkennungssystems darstellt und damit dessen Detektionsleistung steigert. Die praktische Anwendung und Validierung der entwickelten Fusionstechniken erfolgt anhand der Implementierung eines auf der Fusion einer Infrarot- und einer Farbkamera basierenden Fußgängererkennungssystems.

Das vorgestellte Fusionsframework überführt zunächst die Eingangsbilder mittels Wavelet-Transformation in eine Multiskalendarstellung, die es erlaubt, Bildmerkmale zu klassifizieren und damit das Gewicht bestimmter Strukturen in Abhängigkeit deren Relevanz für die Fusion zu verstärken oder abzuschwächen. Die Fusion der Daten erfolgt im transformierten Raum und fundiert auf einem wahrscheinlichkeitstheoretischen Ansatz, der die Entstehung unvollständiger und mit Rauschen behafteter Sensorbilder als Abbild einer realen Szene modelliert. Anhand verschiedener Experimente wird gezeigt, dass die pixelbasierte Datenfusion die Detektionsleistung eines Fußgängererkennungssystems gegenüber Einzelsensordlösungen und den bekannten Fusionsmethoden auf höheren Abstraktionsebenen steigert.

Inhaltsverzeichnis

Nomenklatur	ix
Abkürzungsverzeichnis	xix
1 Einleitung	1
1.1 Zielsetzung und Beitrag der Arbeit	4
1.2 Inhalt und Gliederung der Arbeit	7
2 Multisensorbasierte Szenenerfassung	9
2.1 Umfelderkennung im Automobil	9
2.1.1 Systeme und Sensorik zur lokalen Umgebungserfassung . . .	10
2.1.2 Sensordatenfusion	15
2.1.3 Fusion bildgebender Sensoren	20
2.2 Maschinelles Sehen	23
2.2.1 Bildsensor und Bildrepräsentation	23
2.2.2 Kameramodell und Kalibrierung	25
2.2.3 Stereoskopie	32
2.3 Pixelbasierte Datenfusion	38
2.3.1 Gegenstand der Bildfusion	38
2.3.2 Bildregistrierung	40
2.3.3 Methoden pixelbasierter Datenfusion	42
2.3.4 Fusion basierend auf Multiskalenzerlegung	44
2.4 Fußgängererkennungssysteme	54
2.4.1 Fußgängerschutz	54
2.4.2 Systemanforderungen und Systemarchitektur	55
2.4.3 Beschreibung des Funktionsprinzips	56
2.4.4 Multisensorbasierte Fußgängererkennung	63
3 Pixelbasierte Datenfusion in der Objekterkennung	67
3.1 Ein Framework zur Fusion bildgebender Sensoren auf Bildpunktebene	68
3.1.1 Anforderungen	69

3.1.2	Das Fusionsschema	70
3.2	Rektifikation der Sensorbilder als Form der Bild-Registrierung	72
3.2.1	Epipolargeometrie und Fundamentalmatrix	74
3.2.2	Schätzung der Rektifikationstransformationen	76
3.3	Multiskalenrepräsentation der Sensorbilder	80
3.3.1	Zerlegung der Bilder mittels verschiebungsinvarianter Wavelet-Transformation	81
3.3.2	Rauschunterdrückung	90
3.3.3	Multiskalenkantendetektion	96
3.4	Ein modellbasierter Ansatz zur Bildfusion	101
3.4.1	Das Bildentstehungsmodell	102
3.4.2	Fusion mittels Bayes-Schätzer	103
3.4.3	Schätzung der Modellparameter	105
3.4.4	Konsistenzprüfung und Zusammenfassung	109
4	Implementierung einer multisensorgestützten Fußgängererkennung	113
4.1	Systemüberblick	114
4.1.1	Funktionale Anforderungen	114
4.1.2	Systemarchitektur	115
4.2	Pixelbasierte Fusion einer Infrarotkamera und einer Videokamera	120
4.2.1	Rektifikation der Kamerabilder	120
4.2.2	Fusion der Kamerabilder	124
4.3	Formbasierte Fußgängererkennung	132
4.3.1	Hypothesengenerierung	132
4.3.2	Merkmalsextraktion	134
4.3.3	Definition des Trainings- und Evaluierungsdatensatzes	137
4.3.4	Fußgängerklassifikation	141
4.4	Filterung und Tracking	148
4.4.1	Aufbau des Tracking-Moduls	149
4.4.2	Zustandsschätzung mittels Kalman-Filter	152
5	Experimente und Evaluierung	159
5.1	Bewertungskriterien	159
5.1.1	Informationsgewinn der extrahierten Merkmale	160
5.1.2	Testdatensatz und Kreuzvalidierung	161
5.1.3	Fehlerrate, Konfusionsmatrix und ROC-Kurve	162
5.2	Betrachtung des Informationsgewinns durch die pixelbasierte Datenfusion	163
5.2.1	Inhalt des Experiments	164
5.2.2	Ergebnisse und Diskussion	165
5.3	Vergleich mit High-Level-Fusionsmethoden	169
5.3.1	Inhalt des Experiments	170

5.3.2	Ergebnisse und Diskussion	173
6	Zusammenfassung und Ausblick	181
6.1	Erbrachter wissenschaftlicher Beitrag	183
6.2	Offene Punkte und weiterführende Arbeiten	185
A	Anhang	187
A.1	Antisymmetrische Matrizen	187
A.2	Schätzung der Kameraparameter	188
A.3	Signale und Filter	190
A.4	Vektor- und Funktionenräume	191
	Literaturverzeichnis	193
	Abbildungsverzeichnis	207
	Tabellenverzeichnis	209
	Lebenslauf	211

Nomenklatur

Im Folgenden sind die in dieser Arbeit verwendeten mathematischen Symbole und deren Bedeutung aufgelistet. Sie sind entsprechend ihrer inhaltlichen Zuordnung gruppiert. Innerhalb einer Gruppe sind die Bezeichnungen und Symbole i. A. eindeutig. Es kann jedoch vorkommen, dass ein Ausdruck eine andere Bedeutung hat, dann ist dies jedoch klar aus dem Zusammenhang ersichtlich. Als Dezimalzeichen wird der Punkt verwendet.

Allgemeine Notation und Schreibweise

\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{Z}	Menge der ganzen Zahlen
\mathbb{P}	Menge der Bild- bzw. Pixelwerte
s	Skalar – klein und kursiv
\mathbf{v}	Vektor – klein und fett
\mathbf{M}	Matrix – groß und fett
\mathbf{X}	3D-Punkt – groß und fett
\mathbf{x}	2D-Punkt – klein und fett
$\tilde{\mathbf{x}}, \tilde{\mathbf{X}}$	2D-Punkt bzw. 3D-Punkt in homogenen Koordinaten
$\mathcal{E}\{\cdot\}$	Erwartungswert einer Zufallsgröße
$\mathcal{N}(\mathbf{x}, \mathbf{C})$	Multivariate Normalverteilung mit Erwartungswert \mathbf{x} und Kovarianzmatrix \mathbf{C}
$\exp\{\cdot\}$	Exponentialfunktion
$\langle f, g \rangle$	Internes bzw. Skalarprodukt
$p(A)$	Wahrscheinlichkeit des Ereignisses A
$p(A B)$	Bedingte Wahrscheinlichkeit des Ereignisses A , wenn das Eintreten des Ereignisses B bekannt ist

(x, y)	Position in der Bildebene in Form horizontaler x und vertikaler Pixelkoordinaten y
(u, v)	Position in der Wavelet-Domäne in Form horizontaler u und vertikaler Koeffizientenkoordinaten v

Projektive Geometrie, Stereoskopie und Rektifikation

(Abschnitte 2.2, 3.2 und 4.2.1)

B	Basislinie – Abstand der Projektionszentren zweier Kameras in einem Stereosystem
\mathbf{c}	Kamerahauptpunkt
\mathbf{C}	Ursprung des Kamerakoordinatensystems
d_x	Skalierungsfaktor zur Umrechnung der Disparität von Pixel in mm
\mathbf{D}	Matrix der euklidischen 3D-Transformation
\mathbf{e}	Epipol
\mathbf{E}	Essential-Matrix
f	Brennweite einer Kamera
f_x, f_y	Horizontale und vertikale Kamerakonstante
\mathbf{f}	Vektor der Elemente der Fundamentalmatrix \mathbf{F} mit $\mathbf{f} = (f_{11}, f_{12}, \dots, f_{33})^T$
\mathbf{F}	Fundamentalmatrix
\mathbf{H}	Homographiematrix mit den Spaltenvektor $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3]$
$\mathbf{H}_1, \mathbf{H}_2$	Rektifikationsterme zweier Kameras in Form von Homographiematrizen zur Transformation eines Bildpunktes von der originalen Bildebene in die rektifizierte Ebene
\mathbf{I}	Einheitsmatrix
\mathbf{J}	Jacobi-Matrix
k_x, k_y	Horizontaler und vertikaler Skalierungsfaktor zur Umrechnung in diskrete Pixelkoordinaten
\mathbf{K}	Intrinsische Kameramatrix
\mathbf{K}'	Interne Transformationsmatrix
$\mathbf{K}_{o1}, \mathbf{K}_{o2}$	Original-Kameramatrizen im Rektifikationsprozess
$\mathbf{K}_{r1}, \mathbf{K}_{r2}$	Kameramatrizen der virtuellen rektifizierten Kameras im Rektifikationsprozess
l	Epipolarlinie
\mathbf{O}	Ursprung des Weltkoordinatensystems
\mathbf{p}	Parametervektor zur Schätzung der Rektifikationsterme mittels Optimierungsalgorithmus mit $\mathbf{p} = (\Phi_1, \Psi_1, \Phi_2, \Theta_2, \Psi_2, f_1, f_2)^T$
\mathbf{P}	Projektionsmatrix

\mathbf{P}'	Perspektivische Projektionsmatrix
$\mathbf{P}_o, \mathbf{P}_r$	Projektionsmatrizen für Projektion eines Weltpunktes in die originale Bildebene bzw. in die rektifizierte Ebene
\mathcal{P}^n	Projektiver Raum der Dimension n
\mathbf{R}	Rotationsmatrix mit den Spaltenvektor $\mathbf{R} = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$
s	Skalierungsfaktor
\mathbf{t}	Translationsvektor
$\tilde{\mathbf{u}}_1$	Positionsvektor des Epipols im achsparallelen Stereosystem horizontaler Ausrichtung mit $\tilde{\mathbf{u}}_1 = (1, 0, 0)^T$
$\tilde{\mathbf{u}}_2$	Positionsvektor des Epipols im achsparallelen Stereosystem vertikaler Ausrichtung mit $\tilde{\mathbf{u}}_2 = (0, 1, 0)^T$
x_0, y_0	Horizontale und vertikale Komponente des Kamerahauptpunktes \mathbf{c}
\mathbf{x}	Bildpunkt in Pixelkoordinaten mit $\mathbf{x} = (x, y)^T$
$\tilde{\mathbf{x}}$	Bildpunkt in homogenen Pixelkoordinaten $\tilde{\mathbf{x}} = (x, y, 1)^T$
$\hat{\mathbf{x}}$	Gemessener Bildpunkt in Pixelkoordinaten als Projektion eines Weltpunktes
\mathbf{x}_o	Bildpunkt in der originalen Bildebene
\mathbf{x}_r	Bildpunkt in der rektifizierten Bildebene
\mathbf{x}'	2D-Punkt im Sensorkoordinatensystem mit den Koordinaten $\mathbf{x}' = (x', y')^T$
$\tilde{\mathbf{x}}'$	2D-Punkt im Sensorkoordinatensystem in homogenen Koordinaten $\tilde{\mathbf{x}}' = (x', y', 1)^T$
\mathbf{x}'_d	Durch die Linsenverzeichnung verzerrter 2D-Punkt im Sensorkoordinatensystem mit $\mathbf{x}'_d = (x'_d, y'_d)^T$
$\{\mathbf{x}_1^*, \mathbf{x}_2^*\}$	Ideales Bildpunktpaar, welches die Epipolarbedingung exakt erfüllt
\mathbf{X}	Vektor der Punktekorrespondenzen im Rektifikationsprozess mit $\mathbf{X} = (x_1, y_1, x_2, y_2)^T$
\mathbf{X}^*	Vektor der idealen Punktekorrespondenzen im Rektifikationsprozess mit $\mathbf{X}^* = (x_1^*, y_1^*, x_2^*, y_2^*)^T$
\mathbf{X}_W	3D-Punkt im Weltkoordinatensystem mit den Koordinaten $\mathbf{X}_W = (X_W, Y_W, Z_W)^T$
$\tilde{\mathbf{X}}_W$	3D-Punkt im Weltkoordinatensystem in homogenen Koordinaten $\tilde{\mathbf{X}}_W = (X_W, Y_W, Z_W, 1)^T$
\mathbf{X}_C	3D-Punkt im Kamerakoordinatensystem mit den Koordinaten $\mathbf{X}_C = (X_C, Y_C, Z_C)^T$
$\tilde{\mathbf{X}}_C$	3D-Punkt im Kamerakoordinatensystem in homogenen Koordinaten $\tilde{\mathbf{X}}_C = (X_C, Y_C, Z_C, 1)^T$

Δ	Disparität – Versatz zwischen den beiden Abbildungen \mathbf{x}_1 und \mathbf{x}_2 eines Weltpunktes \mathbf{X}_W
$\delta_{\mathbf{x}}$	Differenz der idealen und gemessenen Vektoren der Punktekorrespondenzen im Rektifikationsprozess
ϵ	Fehler der Epipolargleichung bei Anwendung der gemessenen Punktekorrespondenzen
Θ	Winkel der Rotation um die Y-Achse zur Überführung der Originalbildebene in die rektifizierte Ebene
κ_i	Radiale und tangentiale Verzerrungskoeffizienten mit $i = 1, 2, 3, 4, 5$
π	Ebene im Raum
ρ	Abstand eines Weltpunktes \mathbf{X}_W von der Kamera
Φ	Winkel der Rotation um die X-Achse zur Überführung der Originalbildebene in die rektifizierte Ebene
Ψ	Winkel der Rotation um die Z-Achse zur Überführung der Originalbildebene in die rektifizierte Ebene
ω	Bild des absoluten Konic
Ω_∞	Absoluter Konic

Pixelbasierte Datenfusion, Multiskalenzerlegung und Wavelet-Transformation

(Abschnitte 2.3, 3.3, 3.4 und 4.2.2)

$a_j^p(u, v)$	Activity Level im Subband p der Skala j an Position (u, v)
$\hat{A}(u, v)$	Kantenorientierungsbild
$A_d(u, v)$	Diskretes Kantenorientierungsbild
$c_j(k)$	Approximationskoeffizient der Skala j an Position k
$c_j(u, v)$	Approximationskoeffizient im zweidimensionalen Fall der Skala j an Position (u, v)
$c_j^x(u, v)$	Approximationskoeffizient des horizontalen Subbandes der Skala j an Position (u, v)
$c_j^y(u, v)$	Approximationskoeffizient des vertikalen Subbandes der Skala j an Position (u, v)
$C(y)$	Zu minimierende Kostenfunktion im Rahmen der MAP-Schätzung
$\mathbf{C}_{\mathbf{x}}$	Kovarianzmatrix der Wavelet-Repräsentationen \mathbf{x} gemäß dem Bildentstehungsmodell
$d_j(k)$	Detailkoeffizient der Skala j an Position k
$d_j(u, v)$	Detailkoeffizient im zweidimensionalen Fall der Skala j an Position (u, v)
$d_j^x(u, v)$	Detailkoeffizient des horizontalen Subbandes der Skala j an Position (u, v)

$d_j^y(u, v)$	Detailkoeffizient des vertikalen Subbandes der Skala j an Position (u, v)
$e(k)$	Impulsantwort bzw. Filterkern einer Glättungsfunktion
$e^x(u, v)$	Kantenrepräsentation (Edge Map) für vertikale Kanten
$e^y(u, v)$	Kantenrepräsentation (Edge Map) für horizontale Kanten
$f^\Psi(s, u, v)$	Kontinuierliche Wavelet-Transformierte einer zweidimensionalen Funktion $f(x, y)$ der Skala j an Position (u, v)
f	Rauschfreies Bildsignal
g	Verrauschtes Bildsignal
$h(k)$	Impulsantwort bzw. Filterkern eines digitalen Tiefpasses einer Analysefilterbank
$h'(k)$	Impulsantwort bzw. Filterkern eines digitalen Tiefpasses einer Synthesefilterbank
$H(z)$	Übertragungsfunktion eines digitalen Filters
$g(k)$	Impulsantwort bzw. Filterkern eines zu $h(k)$ komplementären digitalen Hochpasses einer Analysefilterbank
$g'(k)$	Impulsantwort bzw. Filterkern eines zu $h'(k)$ komplementären digitalen Hochpasses einer Synthesefilterbank
G	Zweidimensionale Median-Filtermaske
K	Justierparameter für die Multiskalenkantendetektion
$L^2(\mathbb{R})$	Hilbertraum als Vektorraum über die reellen Zahlen mit einem Skalarprodukt
$m_j^p(u, v)$	Matchwert im Subband p der Skala j an Position (u, v)
$M(u, v)$	Kantenbetragsbild
n	Gaußsches weißes Rauschen in der Wavelet-Domäne
n_i	Wavelet-Transformierte des Sensorrauschens
p	Ordnung eines digitalen Filters
p_n	Wahrscheinlichkeitsdichtefunktion des additiven weißen Rauschens
p_w	Wahrscheinlichkeitsdichtefunktion der rauschfreien Wavelet-Koeffizienten (a-priori-Verteilung)
$p_{w d}$	Bedingte Verteilungsdichtefunktion für das Vorkommen des rauschbefreiten Wavelet-Koeffizienten w beim gegebenen rauschbehafteten Koeffizienten d
p^x, p^y	Interskalenprodukt der horizontalen bzw. vertikalen Subbander
$r(k)$	Impulsantwort bzw. Filterkern einer Glättungsfunktion
R	Quadratische Filterregion
s	Parameter der Laplace-Verteilung
T	Schwellwert im Wavelet-Raum zur Rauschunterdrückung
v	Parameter der Laplace-Verteilung

\mathbf{v}	Eigenvektor einer Matrix
V_j	Funktionsraum der Skala j , wobei $f(x) \in V_{j+1}$ und $f(2x) \in V_j$ und damit $V_{j+1} \subset V_j$ gilt
W_j	Komplementärer Funktionsraum zu V_j , so das gilt $V_j = V_{j+1} \oplus W_{j+1}$ und $V_{j+1} \perp W_{j+1}$
\mathcal{W}	Lokale Bildregion in der Wavelet-Repräsentation mit 3×3 bzw. 5×5 Koeffizienten
w	Rauschbefreites Wavelet-Signal
\hat{w}	Schätzung der rauschfreien Wavelet-Koeffizienten
$w_j^p(u, v)$	Gewicht eines Wavelet-Koeffizienten im Subband p der Skala j an Position (u, v)
x_i	Gemessene Wavelet-Repräsentation des realen Sensorbildes i
\mathbf{x}	Vektor der gemessenen Wavelet-Repräsentationen der realen Sensorbilder $\mathbf{x} = (x_1, x_2, \dots, x_K)^T$
y	Wavelet-Repräsentation eines idealen Sensors bzw. einer realen Szene
\hat{y}	Schätzung der Wavelet-Repräsentation der realen Szene
y_0	Mittelwert der Wavelet-Verteilung der realen Szene
\mathbf{z}	Gaußsches weißes Rauschen in der Bilddomäne
α	Lipschitz-Exponent
β_i	Beobachtungsoperator, der die Wavelet-Repräsentation eines idealen Sensors in die Wavelet-Darstellungen des realen Sensors i überführt
$\boldsymbol{\beta}$	Vektor der Beobachtungsoperatoren mit $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_K)^T$
$\Gamma(x)$	Gamma-Funktion mit $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$
$\eta_h(d)$	Schwellwertfunktion der Hard-Thresholding-Methode in Abhängigkeit der Wavelet-Koeffizienten d
$\eta_s(d)$	Schwellwertfunktion der Soft-Thresholding-Methode in Abhängigkeit der Wavelet-Koeffizienten d
$\theta(x)$	Eindimensionale Glättungsfunktion
$\Theta(x, y)$	Zweidimensionale Glättungsfunktion
Θ	Fusionsregel zur Kombination von Bilddaten auf Pixelebene
λ	Eigenwert einer Matrix
$\mu_{4,w}$	Wölbung (viertes Moment) der Wavelet-Koeffizienten-Verteilung
$\boldsymbol{\mu}_{\mathbf{x}}^{mod}$	Mittelwert innerhalb der Wavelet-Repräsentationen gemäß dem Bildentstehungsmodell
$\boldsymbol{\mu}_{\mathbf{x}}^{mess}$	Gemessener Mittelwert innerhalb der Wavelet-Repräsentationen

σ	Rauschpegel
σ_n	Standardabweichung des Weißen Rauschens im Wavelet-Raum
σ_w	Standardabweichung der Wavelet-Koeffizienten
$\sigma_{\mathbf{x}}$	Standardabweichung innerhalb der gemessenen Wavelet-Repräsentationen \mathbf{x}
σ_y	Standardabweichung der Koeffizienten innerhalb der Wavelet-Repräsentation eines idealen Sensors
$\Sigma_{\mathbf{n}}$	Kovarianzmatrix des transformierten Sensorrauschens
$\phi(x)$	Skalierungsfunktion
$\hat{\phi}(\xi)$	Fourier-Transformierte der Skalierungsfunktion
$\hat{\phi}(\omega)$	Fourier-Transformierte der Skalierungsfunktion $\phi(x)$
$\phi_{j,k}(x)$	Skalierungsfunktion mit der Skalierung 2^j und der Translation $k \cdot 2^j$
$\phi(x, y)$	Zweidimensionale Skalierungsfunktion
$\Phi^x(x, y)$	Separable zweidimensionale Skalierungsfunktion mit $\Phi^x(x, y) = \phi(x)\theta(y)$
$\Phi^y(x, y)$	Separable zweidimensionale Skalierungsfunktion mit $\Phi^y(x, y) = \theta(x)\phi(y)$
$\psi(x)$	Mutter- bzw. Basis-Wavelet
$\psi_{a,b}(x)$	Wavelet mit dem Skalierungsfaktor a und dem Verschiebungsfaktor b mit $a, b \in \mathbb{R}$
$\psi_{j,k}(x)$	Wavelet einer dyadischen Familie mit der Skalierung 2^j und der Translation $k \cdot 2^j$
$\psi^D(x, y)$	Zweidimensionale Wavelet-Funktion mit diagonaler Orientierung
$\psi^H(x, y)$	Zweidimensionale Wavelet-Funktion mit horizontaler Orientierung
$\psi^V(x, y)$	Zweidimensionale Wavelet-Funktion mit vertikaler Orientierung
$\Psi^x(x, y)$	Separables zweidimensionales Wavelet mit $\Psi^x(x, y) = \psi(x)\theta(y)$
$\Psi^y(x, y)$	Separables zweidimensionales Wavelet mit $\Psi^y(x, y) = \theta(x)\psi(y)$
Ψ	Transformationsvorschrift zur Überführung eines Bildes vom Ortsraum in den transformierten Raum

Maschinelles Lernen, Merkmalsextraktion und Klassifikation

(Abschnitte 2.4 und 4.3)

D	Gesamtdetektionsrate eines Kaskadenklassifikators
-----	---

D_k	Detektionsrate einer Stufe k des Kaskadenklassifikators
$f_n(\mathbf{x})$	Testfunktion im Knoten n eines Entscheidungsbaumes basierend auf den Eingaberaum \mathbf{x}
E_n	Entropie innerhalb eines Knotens n eines Entscheidungsbaumes vor der Aufspaltung
E'_n	Entropie innerhalb eines Knotens n eines Entscheidungsbaumes nach der Aufspaltung
$E_{n,L}$	Entropie des linken Pfades nach der Aufspaltung im Knoten n
$E_{n,R}$	Entropie des rechten Pfades nach der Aufspaltung im Knoten n
F	Gesamtfehlalarmrate eines Kaskadenklassifikators
F_k	Fehlalarmrate der Stufe k des Kaskadenklassifikators
$h_C(d)$	Histogrammwert für die Zelle C für die Gradientenrichtung d
$h_t(\mathbf{x})$	Klassifikationsergebnis eines Basisklassifikators anhand eines Merkmalsvektors \mathbf{x}
$H(\mathbf{x})$	Klassifikationsergebnis basierend auf der gewichteten Mehrheitsentscheidung mehrerer Basisklassifikatoren
$II(u, v)$	Integralbild
k	Korrekturfaktor zur Einstellung der Schwelle zwischen den beiden Objektklassen
K	Normierungsfaktor für die Berechnung der HoG
L_n	Linker Pfad nach Aufspaltung des Eingaberaumes am Knoten n eines Entscheidungsbaumes
p_n^y	Wahrscheinlichkeit des Auftretens eines Beispiels der Objektklasse y im Knoten n eines Entscheidungsbaumes
R_n	Rechter Pfad nach Aufspaltung des Eingaberaumes am Knoten n eines Entscheidungsbaumes
Q_n	Differenz der Entropien im Knoten n eines Entscheidungsbaumes vor und nach der Aufspaltung
\mathcal{T}	Trainingsdatensatz
\mathcal{V}	Validierungsdatensatz
V	Gesamtergebnis basierend auf mehreren Klassifikationsergebnissen im Rahmen eines Voting-Verfahrens
v_j	Klassifikationsergebnis des Klassifikators j
$w_{t,i}$	Gewicht eines Beispiels i für das Training des Basisklassifikators h_t
\mathbf{x}_i	Gesamtmerkmalsvektor eines Trainingsbeispiels i
\mathbf{x}^m	Merkmalsvektor eines Meta Feature m
\mathcal{X}	Gesamtdatensatz
y_i	Objektklasse des Beispiels i mit $y_i \in \{0, 1\}$

α_t	Gewicht einer Entscheidung eines Basisklassifikators h_t zur Berechnung der Gesamtaussage
β_j	Gewicht eines Klassifikatorergebnisses im Rahmen eines Voting-Verfahrens
ϵ_t	Trainingsfehler des Basisklassifikator h_t
θ_n	Schwellwert der Testfunktion f_n im Knoten n des Entscheidungsbaumes

Tracking

(Abschnitt 4.4)

A	Systemmatrix bzw. Zustandsübergangsmatrix
B	Steuermatrix
H	Messübergangsmatrix
K	Kalman-Verstärkung
P	Kovarianzmatrix des Schätzfehlers des Systemzustandsvektors
Q	Kovarianzmatrix des Prozessrauschens
R	Kovarianzmatrix des Messrauschens
ΔT_k	Zeitintervall zwischen zwei Messzyklen
$\mathbf{u}(k)$	Steuervektor bzw. Eingangsvektor zum Zeitpunkt k
$\mathbf{v}(k)$	Eingangsstörung bzw. Prozessrauschen zum Messzeitpunkt k
$\mathbf{w}(k)$	Messrauschen zum Zeitpunkt k
$\mathbf{x}(k)$	Wahrer Systemzustandsvektor zum Messzeitpunkt k
$\hat{\mathbf{x}}(k)$	Schätzung des Systemzustandes zum Messzeitpunkt k
$\hat{\mathbf{x}}(k k-1)$	Prädiktion des Systemzustands zum Zeitpunkt k basierend auf den Informationen bis zum Zeitpunkt $k-1$
$\tilde{\mathbf{x}}(k)$	Schätzfehler des Systemzustandsvektors
$\mathbf{z}(k)$	Messvektor zum Zeitpunkt k
$\hat{\mathbf{z}}(k k-1)$	Prädizierter Messwert für den Zeitpunkt k auf der Basis der Informationen bis zum Zeitpunkt $k-1$
$\mathbf{v}(k)$	Innovation im Messraum
$\sigma_{a_x}, \sigma_{a_y}$	Standardabweichung der Beschleunigung in X- und Y-Richtung
σ_u, σ_v	Unsicherheit des Fußpunktes eines Detektionsfenster im Bild

Abkürzungsverzeichnis

ABS	Antiblockiersystem
ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance Systems
AEB	Automated Emergency Braking
BMBF	Bundesministerium für Bildung und Forschung
CAN	Controller Area Network
CCD	Charge-coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CV	Computer Vision
CV	Constant Velocity
DLT	Direkte Lineare Transformation
DR	Detektionsrate
DWT	Diskrete Wavelet-Transformation
EKF	Extended Kalman Filter
ESP	Elektronisches Stabilitätsprogramm
FAS	Fahrerassistenzsystem
FIR	Fern-Infrarotkamera
FN	False Negative
FOV	Field of View
FP	False Positive
FPR	False Positive Rate
FR	Fehlerrate
FWT	Fast Wavelet Transform
HD	High Definition
HDR	High Dynamic Range
HMI	Human Machine Interface
HoG	Histogram of Oriented Gradients
HoSC	Histogram of Shape-Contexts
IEEE	Institute of Electrical and Electronics Engineers
KF	Kalman-Filter
LBP	Local Binary Pattern

LVDS	Low Voltage Differential Signaling
MAP	Maximum-a-posteriori
ML	Maximum Likelihood
NCAP	New Car Assessment Program
NIR	Nah-Infrarotkamera
NPV	Negative Prediction Value
PCA	Principal Component Analysis
PDF	Probability Density Function
PF	Partikelfilter
PMD	Photonic Mixer Device
PPV	Positive Prediction Value
ROC	Receiver Operating Characteristics
ROI	Regions of Interest
SNR	Signal-Noise-Ratio
SVD	Singular Value Decomposition
SVM	Support Vector Machines
TN	True Negative
ToF	Time-of-Flight
TP	True Positive
TTC	Time-to-Collision
VGA	Video Graphics Array
XML	Extensible Markup Language