

Verteilte Multiobjekt-Multisensorfusion mit dem PHD-Filter

Marco Kruse und Fernando Puente León

Karlsruher Institut für Technologie,
Institut für Industrielle Informationstechnik,
Hertzstraße 16, D-76187 Karlsruhe

Zusammenfassung Da Fahrerassistenzsysteme in Zukunft voraussichtlich verstärkt auf Kommunikation zwischen Fahrzeugen zurückgreifen können, wird in diesem Beitrag der Grundstein für ein Sensorfusionsframework für diese Anwendung gelegt. Für die Fusion auf Objektebene wurden Multiobjekt-Verfolgungsfilter als geeignetes Werkzeug gewählt, wobei zunächst das Probability-Hypothesis-Density-Filter aufgrund seiner gezielten Auslegung für die Verfolgung mehrerer Objekte betrachtet wird. Dieses Filter wird für den Einsatz mit multiplen Sensoren, die sich auch an zu verfolgenden Objekten befinden können, erweitert. Erste Simulationen werden vorgestellt und daraus notwendige Änderungen abgeleitet.

1 Einleitung

Die Motivation für die Kombination mehrerer Informationsquellen durch Sensorfusion sind u. a. die Verringerung der Messunsicherheit durch zusätzliche, statistisch unabhängige Messungen, die Gewinnung qualitativ neuer Information aus gleichartigen Sensoren (z. B. Tiefendaten aus zwei Kameras) oder die Kompensation der Schwächen einzelner Sensorprinzipien durch weitere heterogene Sensoren.

In modernen Fahrerassistenzsystemen werden bereits oftmals die Daten verschiedener Sensoren (v. a. Kameras/Radar) miteinander fusioniert, um die Schwächen der einzelnen Sensoren bezüglich Auflösung oder Zuverlässigkeit zu vermindern. Jedoch kann auch diese kombinierte Information in komplexen Szenarien unzureichend sein, wenn der beschränkende Faktor das Sichtfeld des Verkehrsteilnehmers ist. Dann ist

es nötig, auch die Daten von anderen Verkehrsteilnehmern in die Umfeldwahrnehmung miteinzubeziehen.

Während innerhalb eines Fahrzeugs die Sensoren, aufgrund der hohen erreichbaren Datenraten bei kabelgebundener Kommunikation, direkt auf Sensordatenebene fusioniert werden können, ist anzunehmen, dass die erzielbaren Datenraten bei einem Austausch zwischen Verkehrsteilnehmern auf absehbare Zeit keine Übertragung von Sensor-Rohdaten zulassen werden und damit nur eine Fusion von Objekthypothesen möglich sein wird.

Ein gut erforschtes Werkzeug zur Sensorfusion auf Objektebene stellen Multiojekt-Verfolgungsfiler dar. Die drei Hauptansätze hierbei sind

- Einzelobjekt-Verfolgungsfiler mit einem Datenassoziationsschema,
- Multi-Hypothesen-Tracking und
- auf der Statistik zufälliger endlicher Mengen basierende Filter.

Der vorliegende Beitrag beschreibt das Grundgerüst eines auf der Statistik zufälliger endlicher Mengen beruhenden Multiojekt-Verfolgungsfilters, welches in der Lage ist, Daten von verschiedenen Sensoren, welche an Fahrzeugen angebracht oder fest installiert sein können, zu einem Gesamtlagebild zu fusionieren. Den Kern des Verfahrens bildet zur Zeit das Probability-Hypothesis-Density-Filter (PHD-Filter) [1].

Der Beitrag gliedert sich wie folgt: Das PHD-Filter und seine Varianten werden in Abschnitt 2 kurz vorgestellt. Die Verallgemeinerung auf mehrere Sensoren beschreibt Abschnitt 3. Die durchgeführten Simulationen und deren Ergebnisse finden sich in Abschnitt 4, während Abschnitt 5 mit einer Zusammenfassung und einem Ausblick auf Erweiterungen abschließt.

2 PHD-Rekursion

Bei der Erweiterung der Einzel- auf die Multiobjekt-Verfolgung vollzieht man den Übergang von der Betrachtung zufälliger Vektoren zur Betrachtung zufälliger endlicher Mengen. Eines der ersten technisch realisierbaren Filter auf Basis der Statistik zufälliger endlicher Mengen war das Probability-Hypothesis-Density-Filter, welches Mahler 2003 vorstellte [1]. Anstelle der Multiobjekt-Dichte $f(\mathbf{X})$ wird deren erstes Moment, also die Intensität oder Probability-Hypothesis-Density $v(\mathbf{x})$, verfolgt. Sie

ist durch folgende Gleichung definiert:

$$E \{|S|\} = \int |\mathbf{x} \cap S| f(\mathbf{x}) \delta \mathbf{x} = \int_S v(\mathbf{x}) \, d\mathbf{x}, \tag{19.1}$$

wobei E den Erwartungswertoperator, S eine beliebige Untermenge des Zustandsraumes eines einzelnen Objektes und $|A|$ die Kardinalität der Menge A bezeichnen.

Wie man sieht, ist die Intensität auf dem Raum eines einzelnen Objektes definiert und kann unter der Annahme, dass die vorkommenden zufälligen endlichen Mengen durch Poisson-Prozesse beschrieben werden können, näherungsweise durch folgende Gleichungen propagiert werden:

- Prädiktion:

$$v_{k|k-1}(\mathbf{x}_k) = b(\mathbf{x}_k) + \int p_S(\mathbf{x}_{k-1}) f(\mathbf{x}_k|\mathbf{x}_{k-1}) v_{k-1|k-1}(\mathbf{x}_{k-1}) \, d\mathbf{x}_{k-1} \tag{19.2}$$

- Korrektur:

$$v_{k|k}(\mathbf{x}_k) \approx v_{k|k-1}(\mathbf{x}_k) \times \tag{19.3}$$

$$\left((1 - p_D(\mathbf{x}_k)) + \sum_{\mathbf{z} \in Z} \frac{p_D(\mathbf{x}_k) f(\mathbf{z}|\mathbf{x}_k)}{c(\mathbf{z}) + \int p_D(\mathbf{x}_k) f(\mathbf{z}|\mathbf{x}_k) v_{k|k-1}(\mathbf{x}_k) \, d\mathbf{x}_k} \right).$$

Dabei bezeichnen $p_S(\mathbf{x})$ und $p_D(\mathbf{x})$ die zustandsabhängigen Überlebens- bzw. Detektionswahrscheinlichkeiten sowie $b(\mathbf{x})$ und $c(\mathbf{z})$ die Intensitäten der neu erscheinenden Objekte bzw. der Fehldetektionen (Clutter). Ohne weitere Annahmen an die einzelnen Größen ist auch diese Rekursion bisher nur durch Partikelrepräsentationen verfolgbar (vgl. z. B. [2–4]). Unter ähnlichen Annahmen wie beim klassischen Kalman-Filter (u. a. lineares Bewegungs- und Beobachtungsmodell, konstantes p_S und p_D) lassen sich aber auch hier analytische Lösungen finden wie z. B. in [5]. Diese können dann auch analog zu Extendend- und Unscented-Kalman-Filter erweitert werden, um schwache Nichtlinearitäten zu handhaben.

Im Rahmen dieses Beitrags wurde ein Unscented-Kalman-PHD-Filter eingesetzt, d. h. die PHD $v(\mathbf{x})$ wird durch eine Summe von gewichteten Normalverteilungen beschrieben. Die Nichtlinearitäten der verwendeten

Bewegungs- und Beobachtungsmodelle werden durch die Verwendung der Unscented-Transformation [6] approximiert. Die Überlebens- und Detektionswahrscheinlichkeiten werden als konstant $\forall x$ angenommen.

Als Objekthypothesen des aktuellen Zeitschritts werden die Mittelwerte derjenigen Summanden der PHD genommen, deren Gewicht $> 0,5$ ist.

3 Erweiterung auf verteilte Sensoren

Um das Verfahren auf verteilte Sensoren zu erweitern, sollte zunächst geklärt werden, welche Annahmen getroffen werden. Sensoren werden in drei Kategorien unterteilt:

- Stationäre Sensoren, welche die Umgebung erfassen (z. B. Kameras an unübersichtlichen Kreuzungen o. ä.),
- Sensoren an Fahrzeugen, welche die Umgebung erfassen (Kameras, Radar, Lidar etc.),
- Sensoren an Fahrzeugen, welche interne Daten erfassen um den Zustand des Fahrzeugs zu bestimmen (Inertialsensoren, GPS etc.).

Es wird angenommen, dass jedes Fahrzeug, welches seine erstellten Objekthypothesen an andere kommuniziert, auch über Sensorik verfügt, die es ihm ermöglicht, die eigene Position fortlaufend zu schätzen, und dass diese Schätzung mit hoher Präzision erfolgt. Da alle von Sensoren gelieferten Objekthypothesen kommuniziert werden, kann die eigentliche Fusion zentral gelöst werden, da alle Fahrzeuge auf die gleiche Datenbasis zurückgreifen.

Die Verarbeitung einer neuen (mengenwertigen) Messung Z^k eines Sensors erfolgt wie in Abbildung 19.1 gezeigt, wobei die einzelnen Blöcke im Folgenden näher beschrieben werden.

3.1 UK-Schritt

Die einzelnen Fahrzeuge schätzen ihren eigenen Zustand fortwährend mit. Dies geschieht aufgrund des gewählten, nichtlinearen Bewegungsmodells (s. Abschnitt 4.1) mit Hilfe eines dedizierten Unscented-Kalman-Filters pro Sensorfahrzeug.

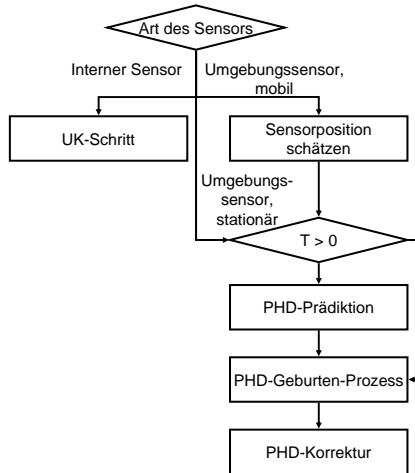


Abbildung 19.1: Verarbeitung einer neuen Messung.

3.2 Sensorposition schätzen

Liefert der Sensor eines Fahrzeugs neue Objekthypothesen, sollen diese über einen PHD-Filterschritt mit dem bisherigen Lagebild fusioniert werden. Da die Messungen z relativ zum Fahrzeug erfolgen, beeinflussen die Position und Orientierung des Sensorfahrzeugs die Likelihood $f(\mathbf{z}|\mathbf{x})$ und müssen daher zunächst bestimmt werden. Bei der Berechnung der Likelihood wird für die Egoposition \mathbf{x}_{ego} des Fahrzeugs die Schätzung des zugehörigen UK-Filters (siehe vorigen Unterabschnitt) herangezogen. Eine Schätzung, welche außerdem die aktuelle Messung Z^k und die prädiizierte PHD $v_{k|k-1}(\mathbf{x}_k)$ zum Messzeitpunkt berücksichtigt, sollte aber in Zukunft erarbeitet werden.

Da momentan die Detektionswahrscheinlichkeiten der Sensoren als konstant angenommen werden, kann das Sichtfeld der Sensoren nicht mit modelliert werden. Der Messung Z^k muss daher noch die Pseudomessung \mathbf{x}_{ego} hinzugefügt werden, damit die Information über die Existenz und Position des Sensorfahrzeugs nicht verloren geht, d. h. $Z^{k'} = Z^k \cup \mathbf{x}_{\text{ego}}$. Beim entsprechenden Summanden in Gleichung (19.4) müssen dann auch einige Anpassungen vorgenommen werden. So wird die Likelihood des Sensors durch eine Normalverteilung mit den Parametern \mathbf{x}_{ego} als Mittelwert

und der zugehörigen, vom UK-Filter geschätzten, Kovarianz ersetzt. Die Werte p_D bzw. $c(\mathbf{z})$ werden zu eins bzw. null gesetzt, da die Egoschätzung weder ausfallende Messungen noch Fehldetektionen hat.

Bei ortsfesten Sensoren wird angenommen, dass ihre Position vorab sehr genau bestimmt wurde. Da sie nicht an Objekten befestigt sind, deren Zustand ebenfalls mitgeschätzt werden soll, ist es in diesem Fall nicht nötig, eine Pseudomessung und einen Pseudosensor zu erzeugen.

3.3 PHD-Prädiktion

Je nachdem, ob seit dem letzten Eingang von Sensordaten Zeit vergangen ist oder nicht, wird die letzte PHD mit Hilfe von Gleichung (19.2) bis zum aktuellen Zeitpunkt prädiziert, wobei die Intensität der neu hinzukommenden Objekte $b(\mathbf{x}_k)$ im nächsten Teilabschnitt gesondert geschätzt wird.

3.4 PHD-Geburten-Prozess

Es wird angenommen, dass neue Objekte in der Nähe von vergangenen Messungen entstehen. Daher wird die Intensität der neu hinzukommenden Objekte $b(\mathbf{x}_k)$ in Abhängigkeit von der letzten Messung Z^{k-1} bestimmt.

3.5 PHD-Korrektur

Als Ausgangsdichte für den Korrekturschritt nach Gleichung (19.4) wird die Summe $v_{k|k-1}(\mathbf{x}_k)$ aus prädizierter PHD und geschätzter PHD neuer Objekte benutzt. Die gesonderte Behandlung der Pseudomessung aus Abschnitt 3.2 ist zu beachten.

4 Experimente

Um die Funktionsfähigkeit des vorgeschlagenen Verfahrens zu überprüfen, wurden in Simulationen zufällig verrauschte Messdaten erzeugt und diese dann gefiltert. Eine Bewertung der Filterergebnisse erfolgt dabei mit der OSPA-Metrik [7] als Abstandsmaß zwischen der Menge der geschätzten und der tatsächlich vorhandenen Objekte. Dabei wird die

OSPA-Metrik in den Lokalisations- und den Kardinalitätsterm aufgetrennt. Als Abstandsmaß wird die gewichtete euklidische Distanz mit dem Gewichtungsvektor $[\frac{2}{3} \frac{2}{3} \frac{6}{\pi} \frac{2}{3} \frac{18}{\pi} 3, 6]^T$ verwendet, um den Einfluss der Größen des Zustandsvektors (vgl. Abschnitt 4.1) anzugleichen. Die Parameter der Metrik werden zu $p = 2$ und $c = 50$ gewählt.

4.1 Bewegungsmodell

Für die Simulation wurde ein zweidimensionales Bewegungsmodell mit konstanter Beschleunigung und Gierrate verwendet, d. h. der Zustandsvektor eines Objektes $\mathbf{x} = [x \ y \ \theta \ v \ \omega \ a]^T$ setzt sich aus der absoluten Position in kartesischen Koordinaten x, y , der Geschwindigkeit in polarer Darstellung θ, v sowie der Gierrate ω und der Beschleunigung a entlang der Bewegungsrichtung zusammen. Die Dynamik lässt sich somit über

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{v} \\ \dot{\omega} \\ \dot{a} \end{bmatrix}^T = [v \cdot \cos(\theta), v \cdot \sin(\theta), \omega, a, 0, 0]^T \quad (19.4)$$

beschreiben.

4.2 Sensormodelle

Die für die Simulationen verwendeten Sensormodelle sollen kurz vorgestellt werden.

Ego-Sensorik

Sämtliche Sensoren zur Bestimmung der eigenen Position werden zu einem Pseudosensor abstrahiert, dessen Beobachtungsraum identisch mit dem Zustandsraum der Objekte ist:

$$\mathbf{z} = \mathbf{x} + \mathbf{n} \quad (19.5)$$

mit $\mathbf{n} \sim \mathcal{N}(\mathbf{0}; \text{diag}[\sigma_x^2, \sigma_y^2, \sigma_\theta^2, \sigma_v^2, \sigma_\omega^2, \sigma_a^2])$. Die Parameter wurden dabei wie folgt gewählt: $\sigma_x = \sigma_y = 0,05$ m, $\sigma_\theta = 0,1^\circ$, $\sigma_v = 1$ km/h, $\sigma_\omega = 0,1^\circ/\text{s}$ und $\sigma_a = 0,1$ km/h/s.

Lidar

Ein Lidar liefert den Winkel β eines Objektes relativ zur Sensororientierung und dessen Abstand r zum Sensor:

$$\mathbf{z} = [\beta \ r]^T = \left[\arctan\left(\frac{y}{x}\right), \sqrt{x^2 + y^2} \right]^T + \mathbf{n} \quad (19.6)$$

mit $\mathbf{n} \sim \mathcal{N}(\mathbf{0}; \text{diag}[\sigma_\beta^2, \sigma_r^2])$. In den Simulationen wurden $\sigma_\beta = 0,25^\circ$ und $\sigma_r = 0,1$ m gesetzt.

Radar

Das Radar liefert, wie beim Lidar, eine Positionsschätzung in Polarkoordinaten – und zusätzlich noch die radiale Geschwindigkeitskomponente des Objektes:

$$\mathbf{z} = [\beta \ r \ v_r]^T = \left[\arctan\left(\frac{y}{x}\right), \sqrt{x^2 + y^2}, v \cdot \cos(\theta - \beta) \right]^T + \mathbf{n} \quad (19.7)$$

mit $\mathbf{n} \sim \mathcal{N}(\mathbf{0}; \text{diag}[\sigma_\beta^2, \sigma_r^2, \sigma_{v_r}^2])$, wobei $\sigma_\beta = 0,5^\circ$, $\sigma_r = 0,5$ m und $\sigma_{v_r} = 0,25$ km/h gewählt wurden.

Stereo-Kamera

Als Messgrößen der Stereo-Kamera werden die Pixelpositionen des Objektes in beiden Kameras angenommen:

$$\mathbf{z} = [\mathbf{k}_l \ \mathbf{k}_r]^T = \left[\frac{y_l}{x_l} \cdot \frac{f}{\Delta y} - \frac{1}{2}, \frac{y_r}{x_r} \cdot \frac{f}{\Delta y} - \frac{1}{2} \right]^T + \mathbf{n}. \quad (19.8)$$

Die Koordinaten x_l, y_l und x_r, y_r sind die Koordinaten des Objektes relativ zur linken bzw. rechten Kamera. Der Kameraparameter $\frac{f}{\Delta y}$ wird durch die Brennweite f der Kamera und den Abstand Δy zweier Pixel auf der Sensorfläche bestimmt und durch Kalibration als bekannt vorausgesetzt ($\frac{f}{\Delta y} \approx 439,75$ in den Simulationen, entspricht einer Kamera mit 640 Pixeln pro Zeile und 36° Öffnungswinkel). Das Messrauschen wird durch $\mathbf{n} \sim \mathcal{N}(\mathbf{0}; [\sigma_k^2, -0,99\sigma_k^2; -0,99\sigma_k^2, \sigma_k^2])$ mit $\sigma_k = \frac{1}{10}$ charakterisiert. Durch die hohe negative Korrelation der beiden Pixelfehler wird eine verrauschte Disparität simuliert.

4.3 Szenario

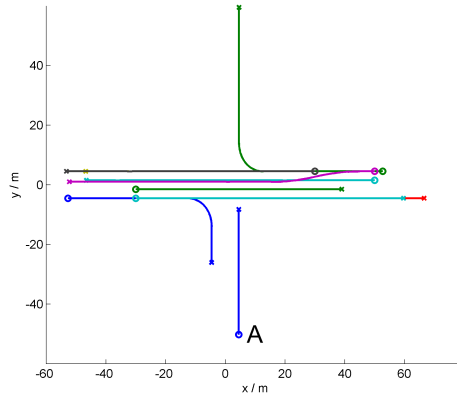


Abbildung 19.2: Simulierte Kreuzungsszene (\circ markiert den Anfang und \times das Ende einer Trajektorie).

Als Simulationsszenario wurde eine Kreuzungsszene gewählt. Die Trajektorien der Objekte sind in Abbildung 19.2 zu sehen. Im Wesentlichen fahren die Autos horizontal mit konstanter Geschwindigkeit (ca. 50 km/h) über die Kreuzung, während das Fahrzeug A auf die Kreuzung zufährt und bremst. Die Fahrzeuge erscheinen/verschwinden immer ca. 50 m von der Kreuzung entfernt. Für die Simulationen wurden folgende Sensoren verwendet:

- ein Radar an A,
- eine Stereo-Kamera an A,
- eine Stereo-Kamera, die fest an der linken oberen Ecke der Kreuzung steht und zur Kreuzungsmitte ausgerichtet ist,
- ein Lidar an einem der Fahrzeuge, die von rechts kommen.

Die Detektionswahrscheinlichkeiten p_D der Sensoren wurden einheitlich zu 1 gesetzt, d.h. es fallen keine Messungen aus. Die Abtastzeiten der Sensoren wurden alle zu $T = 0,2\text{s}$ gesetzt, wobei die Werte um $0,05\text{s}$ versetzt geliefert werden. Das Szenario wurde jeweils mit jedem Sensor

einzel simuliert und anschließend unter Verwendung aller Sensoren. Die Ergebnisse sind in den Abbildungen 19.3 und 19.4 gezeigt.

Es fällt auf, dass in der Zeit, in der neue Objekte erscheinen (bis 3 s), der Kardinalitätsfehler nach Kamera-Updates durch Fusion stark absinkt. Es ist zu erwähnen, dass das Filter nach einiger Zeit stets in der Lage ist, den Zustand – bis auf kurze Kardinalitätsschwankungen – gut zu schätzen.

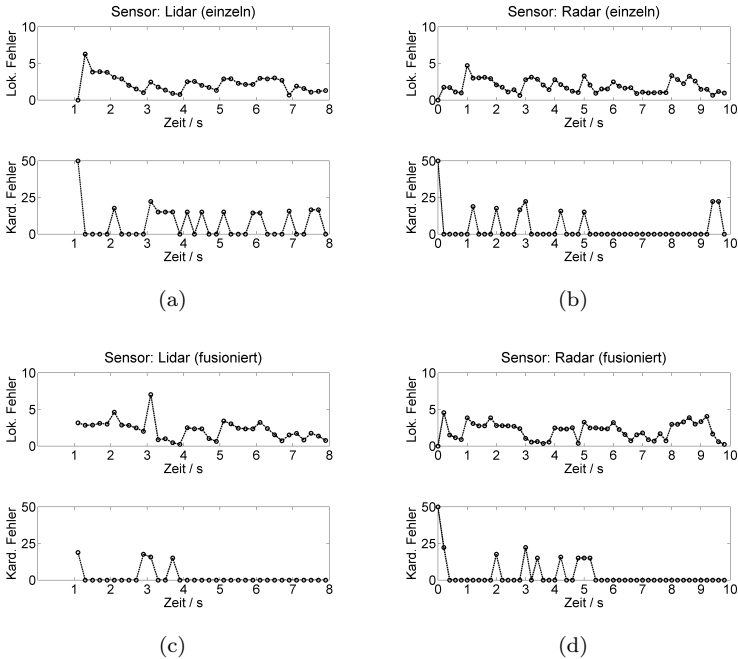


Abbildung 19.3: OSPA-Werte nach Radar- und Lidar-Updates bei Verwendung eines einzigen Sensors (a), (b) oder aller Sensoren (c), (d).

5 Zusammenfassung und Ausblick

Die Simulationen haben gezeigt, dass Sensorfusion mit dem PHD-Filter möglich ist. Weitere Versuche mit ausfallenden Messwerten haben aber

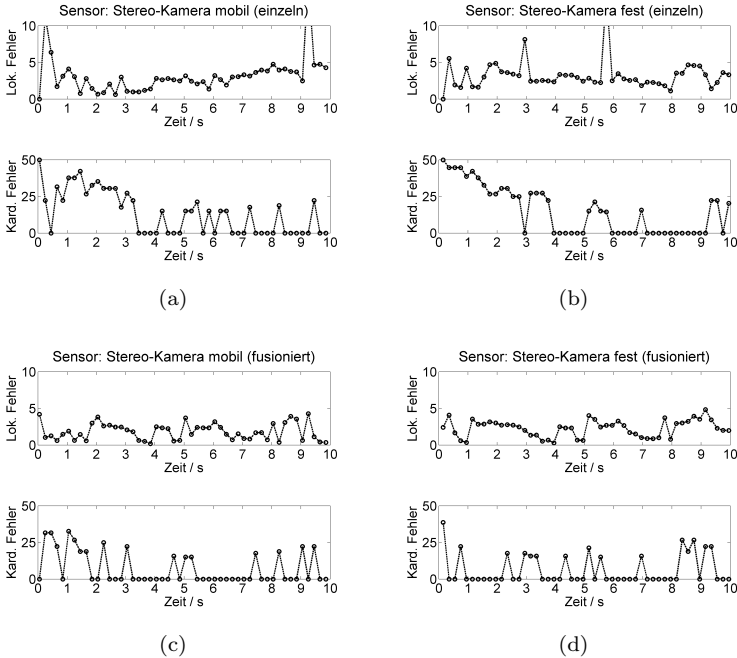


Abbildung 19.4: OSPA-Werte nach Updates der Kameras an Fahrzeug A bzw. an der Kreuzung bei Verwendung eines einzigen Sensors (a), (b) oder aller Sensoren (c), (d).

gezeigt, dass sich das Filter sehr stark auf die Korrekturterme in Gleichung (19.4) verlässt, wodurch ein ausbleibender oder sehr schlechter Messwert das Fusionsergebnis verschlechtern kann.

Um die vorliegenden Ergebnisse zu verbessern, sind mehrere Maßnahmen angedacht. Zum Einen soll das PHD-Filter als Grundlage der Fusion durch das CBMeMber-Filter [8] ersetzt werden. Dieses nimmt den mengenwertigen Systemzustand nicht als Poisson-Prozess an und verfolgt dessen Intensität wie das PHD-Filter, sondern verwendet einen Multi-Bernoulli-Prozess und verfolgt dessen Wahrscheinlichkeitsdichte. Um sowohl das Sichtfeld der Sensoren als auch nicht-normalverteiltes Sensorrauschen realistisch zu modellieren, sollen Partikelrepräsentationen ge-

nauer untersucht werden. Insbesondere deterministische Partikelapproximationen wie z. B. in [9] sollen hierbei in Betracht gezogen werden, da sie mit deutlich weniger Partikeln als herkömmliche Partikelfilter auskommen.

Literatur

1. R. Mahler, „Multitarget Bayes filtering via first-order multitarget moments“, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 39, Nr. 4, S. 1152–1178, Oct. 2003.
2. D. Clark, I. Ruiz, Y. Petillot und J. Bell, „Particle PHD filter multiple target tracking in sonar image“, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 43, Nr. 1, S. 409, 2007.
3. B. Vo, S. Singh und A. Doucet, „Sequential Monte Carlo implementation of the PHD filter for multi-target tracking“, in *Proc. Int. Conf. on Information Fusion*. Citeseer, 2003, S. 792–799.
4. T. Zajic und R. Mahler, „Particle-systems implementation of the PHD multitarget-tracking filter“, in *Proc. SPIE*, Vol. 5096, 2003, S. 291.
5. B. Vo und W. Ma, „The Gaussian mixture probability hypothesis density filter“, *IEEE Trans. Signal Processing*, Vol. 54, Nr. 11, S. 4091, 2006.
6. S. Julier, „The scaled unscented transformation“, in *Proc. American Control Conference*, Vol. 6, 2002, S. 4555–4559 vol.6.
7. D. Schuhmacher, B. Vo und B. Vo, „A consistent metric for performance evaluation of multi-object filters“, *IEEE Trans. Signal Processing*, Vol. 56, Nr. 8, S. 3447–3457, 2008.
8. B. Vo, B. Vo und A. Cantoni, „The cardinality balanced multi-target multi-Bernoulli filter and its implementations“, *IEEE Trans. Signal Processing*, Vol. 57, Nr. 2, S. 409–423, 2009.
9. V. Klumpp und U. Hanebeck, „Dirac mixture trees for fast suboptimal multi-dimensional density approximation“, in *Proc. 2008 IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, 2008.