Tracking und Bewegungsvorhersage von Fahrzeugen in komplexen Innenstadtszenarien
(Tracking and motion prediction of vehicles in complex urban traffic scenes)

Autoren

Julian Einhaus[1,2], Christoph Hermes[2], Markus Hahn[1], Christian Wöhler[1], Franz Kummert[2]

[1]Daimler AG, Group Research and Advanced Engineering, P. O. Box 2360, D-89013 Ulm
[2]Applied Informatics, Faculty of Technology, Bielefeld University, Universitätsstr. 25, D-33615 Bielefeld

Telefonnummern und E-Mail-Adressen

J. Einhaus:	+49-731-505-2120	jeinhaus@techfak.uni-bielefeld.de
C. Hermes:	+49-731-505-4865	chermes@techfak.uni-bielefeld.de
M. Hahn:	+49-731-505-2108	markus_hahn@gmx.de
C. Wöhler:	+49-731-505-2148	christian.woehler@web.de
F. Kummert:	+49-521-106-2929	franz@techfak.uni-bielefeld.de

Acknowledgements

This work was carried out within the research initiative AKTIV-AS supported by the German Bundesministerium für Wirtschaft und Technologie (grant no. 19S6011A).

Abstract

The detection and tracking of vehicles in urban traffic and the long-term prediction of their positions and motion states are indispensable skills of advanced driver assistance systems. Many object tracking systems rely on Kalman filters [1] or particle filters [2,3]. In such systems, motion models are typically restricted to rather simple approaches, such as constant curve radius and constant velocity or acceleration, which often turn out as inadequate for prediction intervals exceeding one second.

In this contribution we describe an object tracking system which relies on a vehicle-based stereo camera. A sparse scene flow field is computed based on stereo and optical flow information extracted using computationally efficient feature-based techniques [4]. In a subsequent clustering stage, the scene flow field is segmented into object hypotheses using a graph-based clustering stage [5]. For each cluster a model is generated which is given by the histogram of the grey values in both camera images. The current object position is determined with the Mean-Shift algorithm [6,7,8] applied to the 3D point cloud. The 3D points are reprojected into the images and the corresponding grey values are used to weight the 3D points in the Mean-Shift scheme based on their relative frequency in the model histogram, which is thus interpreted as a probability of the 3D point to belong to the object.

Given a series of measurements, i.e. the object trajectory up to the current time step, the object state at a specific point in time in the future is predicted based on a Bayesian framework in which the probability distribution of the motion hypotheses is represented by a set of samples (particles) which are propagated in time using a particle filter [9]. The likelihood to observe the measured trajectory, given the model trajectory, is obtained by the quaternion-based rotationally invariant longest common subsequence (QRLCS) metric [10].

The experimental evaluation of our system on complex real-world urban traffic scenes shows that it allows an early detection and robust tracking of vehicles (cars and bicycles) in the presence of cluttered background, partial occlusions, and even temporary full occlusions. Furthermore, we demonstrate that the long-term (2–3 seconds ahead) prediction behaviour of our particle filter framework is superior to that of the standard approach assuming constant acceleration and curve radius, especially regarding the predicted yaw angle and yaw rate.

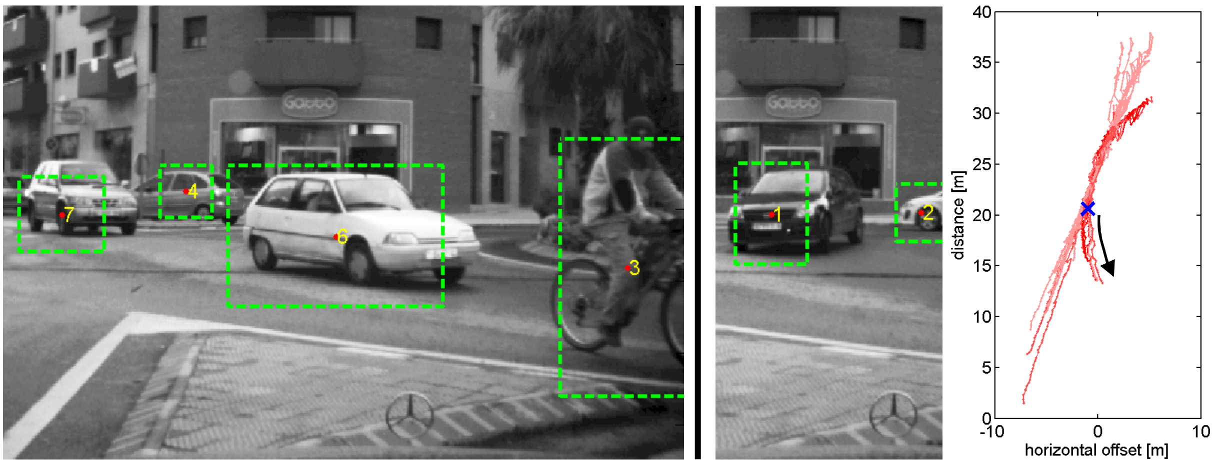

Fig. 1: Left: Typical object detection results. Right: Particle trajectories associated with vehicle 1 following a circular path (intensity denotes particle likelihood, current position is marked by a blue cross).

[1] Barth, A., Franke, U., 2008. Where will the oncoming vehicle be the next second? Proc. IEEE Intelligent Vehicles Symposium.

[2] Montemerlo, M., Whittaker, W., Thrun, S., 2002. Conditional particle filters for simultaneous mobile robot localization and people-tracking. Proc. Int. Conf. on Robotics and Automation.

[3] Beuter, N., Swadzba, A., Schmidt, J., Sagerer, G., 2009. 3D-Szenenrekonstruktion in dynamischen Umgebungen. Proc. Oldenburger 3D-Tage.

[4] Stein, F., 2004. Efficient Computation of Optical Flow Using the Census Transform. Proc. DAGM, pp. 79–86.

[5] Bock, H. H., 1974. Automatische Klassifikation. Vandenhoeck & Ruprecht, Göttingen.

[6] Fukunaga, K., Hostetler, L. D., 1975. The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition. IEEE Trans. on Information Theory 1, 32–40.

[7] Bradski, G. R., 1998. Real time face and object tracking as a component of a perceptual user interface. Proc. 4th IEEE Workshop on Applications of Computer Visios, pp. 214–219.

[8] Comaniciu, D., Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Analysis and Machine Intelligence 24(5), 603–619.

[9] Black, M. J., Jepson, A. D., 1998. A Probabilistic Framework for Matching Temporal Trajectories: CONDENSATION-Based Recognition of Gestures and Expressions. Proc. Europ. Conf. on Computer Vision, pp. 909–924.

[10] Hermes, C., Wöhler, C., Schenk, K., Kummert, F., 2009. Long-term Vehicle Motion Prediction. Proc. IEEE Intelligent Vehicles Symposium.