Kalman Filter based Detection of Obstacles and Lane Boundary in Monocular Image Sequences

Clemens Rabe, Christoph Volmer, and Uwe Franke

DaimlerChrysler AG, 71059 Sindelfingen, Germany
{clemens.rabe,uwe.franke}@daimlerchrysler.com
christoph.volmer@stud.tu-ilmenau.de

Abstract. This paper presents a system for monocular obstacle and lane boundary detection running in real-time. A Kalman Filter based depth from motion algorithm is used for the reconstruction of the three-dimensional scene. Using multiple filters in parallel the rate of convergence is significantly higher than in direct methods, especially if the vehicle drives slowly. In addition a pitch correction is introduced which improves the overall estimation in typical road scenarios. Real world examples illustrate the results of the proposed system.

1 Introduction

Intelligent cars of the foreseeable future will be equipped with a camera for tasks such as lane departure protection and Night Vision. However, several applications in robotics and driver assistance such as obstacle detection, obstacle avoidance, and parking require 3D information. Figure 1 displays a lane boundary detection based on the three-dimensional scene reconstruction algorithm presented in this paper as an example for an eligible application in driver assistance.

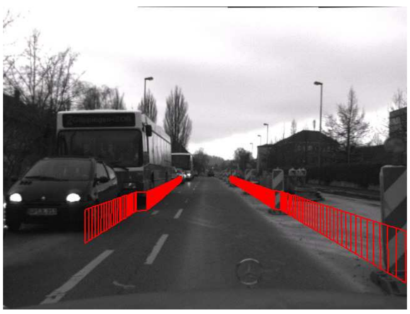

Fig. 1. Successfull lane boundary detection using a monocular image sequence.

If only one camera is available, depth must be estimated from the image sequence obtained while driving. Pollefeys [5] gives a comprehensive overview on structure from motion methods. These provide a very precise scene reconstruction but do not meet real-time requirements. Since the mentioned applications require 3D information for as many image points as possible in order not to overlook an important object, fast methods are needed.

In systems with very limited computational resources often a direct analysis of the optical flow is propagated. For example, Grünewald et al. calculate in [3] the flow in only one reference line and compare it to the expected flow of a ground plane. Every outlier is interpreted as part of an obstacle. Although this algorithm is fast and works well for the proposed scenarios, it highly depends on the quality of the optical flow algorithm. In case of a slow observer movement direct optical flow methods tend to have difficulties due to small displacement vectors in the image resulting in a small signal-to-noise ratio.

Therefore, an integration of multiple measurements over time promises more robust results with respect to measurement noise. Matthies et al. describe in [4] a Kalman Filter based method to estimate the depth of selected image features for a strict lateral observer movement. Carlsson extends this approach and presents in [1] an algorithm to estimate the observer motion as well as the scene structure using a single Kalman Filter. However, without any additional constraints on the motion or scene structure this can only be solved up to an unknown scale factor.

But if the camera motion is known, in vehicles we know at least speed and yaw rate, the 3D-position of stationary world points can be efficiently estimated using one Kalman Filter per image feature. This reduces the complexity of the calculation in contrast to one global Kalman Filter. In addition, the iterative refinement of the Kalman Filter is optimal with respect to computation time. This raises the hope that a real-time estimation of depth is possible which is robust with respect to measurement noise, especially in case of a slow observer movement.

In this paper we present a system for obstacle and lane boundary detection running in real-time. Image features are tracked using the Kanade-Lucas-Tomasi (KLT) tracker [6]. Kalman Filters estimate the 3d-positions under the assumption of a stationary world. In order to get a high rate of convergence we use a multiple filter approach [2]. Using the Kalman Filter prediction as an initialization of the tracker, its speed and accuracy are improved. To detect obstacles the heights of the 3D-points are analysed. As a pitching error directly induces a wrong height estimation, we added a pitch correction to gain more robust estimates.

The paper is organized as follows. Chapter 2 describes the basic principles of the Kalman Filter based depth reconstruction. The pitch correction is outlined in chapter 3. The obstacle and lane boundary detection is described in chapter 4. Results for real sequences are finally presented in chapter 5, followed by the conclusion in chapter 6.

2 Kalman Filter based 3D from Motion

Assuming that a tracked image feature corresponds to a stationary world point a Kalman Filter is used to estimate its 3d-position. In the following we use a right handed coordinate system with the origin at the road. The lateral x-axis points to the left, the height axis y points upwards and the z-axis representing the distance of a point is straight ahead. This coordinate system is fixed to the car, so that all estimated positions are given in the coordinate system of the moving observer. The camera is at $(x, y, z) = (0, height, 0)$.

2.1 System Model

The movement of a vehicle with constant velocity v and yaw rate $\dot{\psi}$ over the time interval Δt, measured by inertial sensors, can be described in this car coordinate system by the translation vector $\Delta \underline{x}_c$ and the rotational matrix R_y around the y-axis. The position of a static world point $\underline{x} = (X, Y, Z)^T$ after the time Δt can be described in the car coordinate system at time step k as

$$\underline{x}_k = R_y \left(\underline{x}_{k-1} - \Delta \underline{x}_c \right) . \tag{1}$$

This yields to the discrete system model equation

$$\underline{x}_k = A_k \underline{x}_{k-1} + B_k v + \underline{w}_{k-1} \tag{2}$$

with the state transition matrix $A_k = R_y$ and the control matrix

$$B_k = \frac{1}{\dot{\psi}} \begin{pmatrix} 1 - \cos\left(\dot{\psi}\Delta t\right) \\ 0 \\ -\sin\left(\dot{\psi}\Delta t\right) \end{pmatrix} . \tag{3}$$

The noise term $\underline{w}$ is assumed to be a gaussian white noise with covariance Q.

2.2 Measurement Model

Image coordinates u and v of a feature are measured using an appropriate point tracker. In our current implementation we use a Kanade-Lucas-Tomasi tracker [6]. Assuming a pin hole camera the nonlinear measurement equation for a point given in the camera coordinate system is

$$\underline{z} = \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{Z} \begin{pmatrix} X f_u \\ Y f_v \end{pmatrix} + \underline{\nu} \tag{4}$$

with the focal lengths f_u, f_v. The measurement noise term $\underline{\nu}$ is also assumed to be a gaussian noise with covariance R.

2.3 Initialization

The Kalman Filter given by the equations 2 and 4 estimates the world position of a static point in relation to the moving car. Before the filter can begin its work, it has to be initialized. Using the first two measurements an initial position can be calculated by triangulation. However, this is highly susceptible to measurement noise, especially if the measured displacement vector is small. Therefore, we get the initial position as the intersection of the focal ray with a pre-defined plane, e.g. the ground plane.

This initial guess will be refined by the filter over time. The more the first guess deviates from the correct value, the longer it takes until the estimate error is below a given threshold. To achieve a higher rate of convergence, multiple filters are used in parallel initialized using different pre-defined planes. Each filter represents a hypothesis in the state space of one world point. All hypothesis are then combined in a weighted average using the innovation errors as a goodness-of-fit criterium [2].

In a running system new features are continuously added. Under the asumption that nearby image features are laos closely related in the world an additional hypothesis can be formulated using the already existing estimation results of the neighbours. Therefore, a fixed area around the new feature is searched and an average world position is calculated. For each found feature its estimated position is weighted with the covariance and the distance to the new feature in the image.

3 Pitch correction

The Kalman Filter provides an estimate which is robust against gaussian noise in the measurement. However, a camera orientation change, such as pitch, yaw or roll, which is not measured by additional sensors, has a dramatic influence on the whole estimation process. Therefore, the correction of such errors is a fundamental step to improve the 3D-reconstruction.

In our scenarios the main component is pitching. Therefore we developed a pitch correction which takes place between the Kalman Filter prediction and measurement update step. Assuming a correct estimated 3D-point the predicted and the measured image position should be identical. However, any remaining innovation in vertical direction is interpreted as a pitch movement. To be robust against outliers we average the pitch deviations of all features. As the state covariance gives the uncertainty of the 3D-estimation, we use it as a weighting factor. The resulting pitch angle is used in the Kalman Filter measurement update step for all tracked features.

The benefit of the pitch correction can be seen in the right diagram of figure 2. It shows the estimated 3D-points in a top view. Without correction the pitching in this sequence causes wrong estimates, shown as green boxes. Especially points at a large distance tend to have serious problems. Here, the expected innovation caused by the forward motion is very small in contrast to the innovation due to pitching. On the other hand, the 3D-reconstruction using the pitch correction leads to stable estimation results, indicated by red crosses. This significantly improves obstacle and lane boundary detection.

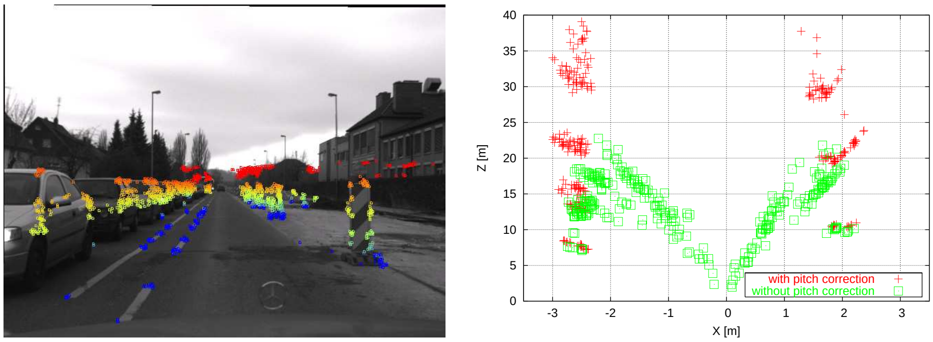

Fig. 2. 3d-estimation stabilized by pitch correction (left). The warmth of the color indicates the estimated height of the tracked image features. The improvement of the pitch correction is illustrated by the top view (right).

4 Obstacle and Lane Boundary Detection

Assuming a planar road all estimated world points with a significant height above the ground are interpreted as part of obstacles. For a robust object detection all these points are mapped into a two dimensional histogram according to their x-z-coordinates and a connected component analysis is performed to identify the obstacles.

The two dimensional histogram is also used to determine the lateral boundary of the free corridor in front of the vehicle. We start on the vehicles lateral position and search in each line for the left and right boundary respectively. Since we are interested to locate the first significant slope the derivative of this one dimensional distribution is calculated. The first maximas to each side in the derivative function now correspond to the lateral position of the boundary.

5 Real World Results

In this final section we discuss three real-live examples. First, a sequence of 60 frames containing a static obstacle is investigated. The sequence was taken from a truck driving at about 1 $\frac{\mathrm{m}}{\mathrm{s}}$ towards a standing person. Due to this very slow speed, the lengths of the optical flow vectors are small and the integration performed by the Kalman Filter is of paticular importance.

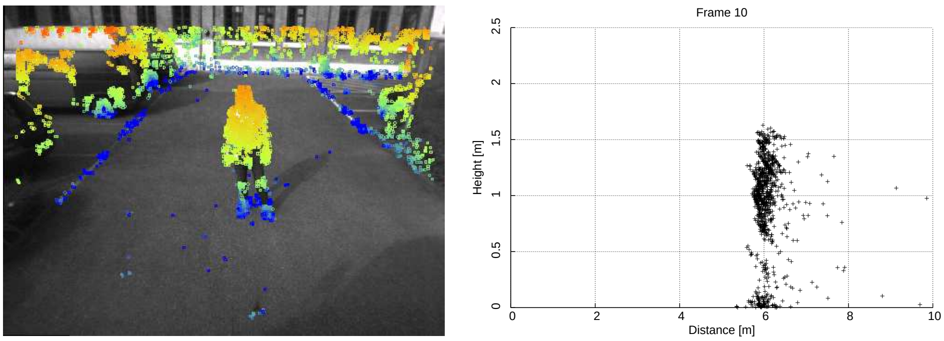

Fig. 3. Estimation result for the person at frame 10 (driven distance 40 cm). The warmth of the color encodes the estimated height.

Figure 3 shows on the left the estimated heights after 10 frames (i.e. 40 cm driving distance) qualitatively. We illustrate the height with the warmth of the color. The estimation of about 5000 image features was performed by three Kalman Filters per feature, initialized on the heights 0, 1 and 2 m. On the right the corresponding 3D-reconstruction result is given in a lateral view.

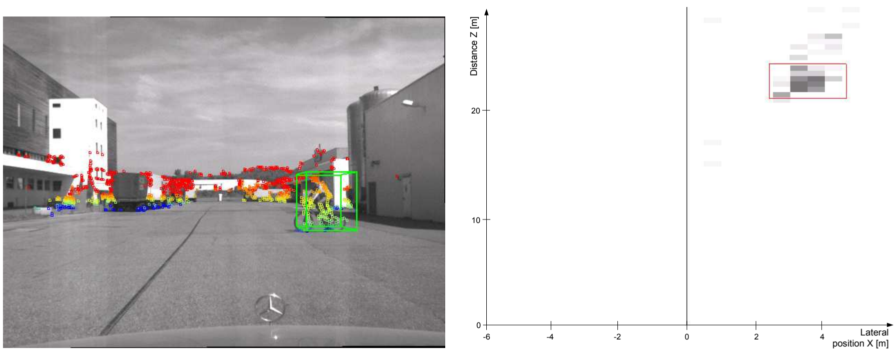

Fig. 4. Result of the obstacle detection. The two dimensional histogram on the right clearly shows the standing bicyclist as an obstacle.

The result of the obstacle detection method is illustrated by figure 4. Here we drive with a passenger car at about 3.5 $\frac{\mathrm{m}}{\mathrm{s}}$ towards a standing bicyclist. The purpose of this experiment is to stop the vehicle in case of an obstacle. The right figure shows the two dimensional histogram used for the obstacle detection with a quantisation of 0.5 m. The result of the closed component analysis is given by the red rectangle and is reprojected as a green box in the camera image.

The last example shown in figure 1 at the beginning of this paper is concerned with a typical construction site. Here, the actual lane is limited by beacons and other vehicles in a traffic jam. In contrast to the first examples, our car is driving at about $10\,\frac{\mathrm{m}}{\mathrm{s}}$.

After mapping the 3D-points into the two dimensional histogram, the 3D-lane boundary have been computed and reprojected into the image. The red fences show the results obtained by the presented depth-from-motion algorithm. This proves that 3D road boundary can be detected even if only one camera is installed in the vehicle.

6 Conclusion

A robust way to solve the 3D-from-motion problem is the usage of Kalman Filters. This paper shows a system for obstacle and lane boundary detection running in real-time. The proper combination of multiple filters initialized with different states as well as the proposed initialization based on the neighbourhood speed up the rate of convergence. The presented pitch correction stabilises the 3d-reconstruction on typical road scenes significantly. In our current implementation, the 3D-positions of 1100 points are calculated on a 3 GHz Pentium 4 at 15 Hz. This includes the image acquiring process, the tracking of the features on images with QVGA resolution, the Kalman Filter estimation using three filters per feature and the obstacle and lane boundary detection.

References

1. Carlsson, S.: Recursive Estimation of Ego-Motion and Scene Structure from a Moving Platform. Proceedings of the 7th Scandinavian Conference on Image Analysis, Aalborg, pp. 958–965, 1991
2. Franke, U. and Rabe, C.: Kalman Filter based Depth from Motion with Fast Convergence. Proceedings of the IEEE Intelligent Vehicles Symposium 2005, Las Vegas, NV, June 2005
3. Grünewald, M. and Sitte, J.: A Resource-Efficient Approach to Obstacle Avoidance via Optical Flow. Proceedings of the 5th International Heinz Nixdorf Symposium: Autonomous Minirobots for Research and Edutainment, Paderborn, pp. 205–214, October 2001
4. Matthies, L. and Kanade, T. and Szeliski, R.: Kalman Filter-based Algorithms for Estimating Depth from Image Sequences. International Journal of Computer Vision, Nr. 3, pp. 209–236, 1989
5. Pollefeys, M.: Self-Calibration and metric 3D reconstruction from uncalibrated image sequences. Leuven, 1999
6. Tomasi, C. and Kanade, T.: Detection and Tracking of Point Features. Technical Report CMU-CS-91-132, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, April 1991