Cartography and Dead Reckoning Using Stereo Vision for an Autonomous Car

Stefan K. Gehrig
Research Institute DaimlerChrysler AG
Stefan.Gehrig@DaimlerChrysler.Com

Fridtjof J. Stein
Research Institute DaimlerChrysler AG
Fridtjof.Stein@DaimlerChrysler.Com

Abstract

Our main objective in this paper is to perform a cartography of a road scene into a reference frame at rest, where 3D measurements delivered by on-board sensors serve as input. The main sensors of our autonomous vehicle are two CCD cameras. Their pictures are combined using stereopsis to generate 3D data.

We need dead reckoning to properly associate 3D data among the frames. This necessitates us to obtain a precise ego-motion estimation. Dead reckoning using only standard vehicle odometry can cause non-negligible errors. We use stationary points in the scene to support the determination of our ego-motion. Two types of stationary objects are used: Vertical landmarks such as traffic signs and lane markings are used to compensate positioning errors.

Preliminary results show that cartography as proposed in this paper is beneficial to detect stationary objects but needs further work for fast moving objects.

1 Introduction

In recent years tremendous progress has been made in the field of intelligent vehicles for regular traffic (e.g. [3, 5]). However, none of these approaches known to the authors tried to map the environment of the autonomous vehicle into a reference frame at rest.

How is this paper organized ? Section 2 gives an overview of related work both in the field of intelligent vehicles and robotics. In Section 3 the necessary transformations and the vehicle model for motion integration are briefly introduced. Section 4 describes the cartography based on 3D points. Dead reckoning techniques for an unknown environment are detailed in Section 5. Cartography results and simulation results of our dead reckoning algorithms are shown in Section 6. Conclusions and future work comprise the final Section.

2 Related Work

2.1 Related Work on Cartography

Cartography of range data is common for autonomous mobile systems, especially for systems equipped with unreliable range sensors (e.g. sonar sensors). Integrating sensor readings taken at different times/positions and superimposing them makes a compensation of the erroneous measurement readings feasible.

A popular approach for that, evidence grids, was pioneered by Moravec [10]. This mapping technique is also referred to as certainty grids or occupancy grids (see e.g. [1] and [9]). It models the free space around an autonomous system for navigation and obstacle avoidance.

One weakness of this mapping technique is its limited capability to map moving objects. Furthermore, discretization of the environment is necessary which might compromise managing the data if a sufficiently large area is modeled with suitable accuracy.

For the subsequent discussion we use the following convention: A global map is a map integrating information from different times and positions into a reference frame at rest. A local map, on the other hand, uses only current information to model the environment and has its origin at the sensor center of the moving vehicle.

2.2 Related Work on Dead Reckoning

Dead reckoning is needed in almost all mobile robot applications. Dead reckoning with local sensors in an unknown environment leads to the problem of ego-motion estimation. Integrating the motion of the ego-vehicle leads to proper localization in a reference frame at rest.

For cars, ego-motion estimation using vision clues has been investigated e.g. in [11] and [12]. In contrast to our research, they utilized only one camera.

3 Motion Integration

For the remainder of this paper, the following reference frames are used: For the local reference frame, the z-axis runs parallel along the vehicle longitudinal axis, the x-axis is perpendicular to the z-axis and parallel to the ground plane directed to the left. The y-axis protrudes upwards. The origin is located at the sensor center (in the middle between the two CCD cameras) projected onto the x-z-plane on the ground. The global reference frame is equivalent to the local coordinate system at the time of initialization and remains stationary.

In order to transform the 3D measurements into the global reference frame, two steps are performed. A flat road ($y = 0$) is assumed for all transformations. In the first step the coordinates of the ego-vehicle are transformed into the global reference frame. In the second step, the 3D measurements are transformed into the global reference frame knowing the orientation of the local map in the global reference frame.

For the underlying vehicle model in the coordinate transformations we use the kinematic Ackermann model extended by the self-steering gradient [14].

4 Cartography Based on 3D Points

4.1 Creating the Global Map

Our range sensor, a calibrated stereo camera system, delivers 3D measurement of significant points resulting in a sparse 3D point cloud (less than 500 3D measurements). The significant points of the left image are matched in the right image by correspondence analysis along the epipolar line similar to [4]. However, the algorithms described in the remainder of the paper are applicable to any range sensor delivering (sparse) range data.

The global map is created accumulating 3D points from the current and previous frames. To extract objects from 3D measurements, we apply a spatial clustering method to all 3D points except the ones on or below the ground. The list of 3D points is traversed and two 3D points are connected when their Euclidian distance is below a certain threshold. This cluster connectivity is protocoled using coloring schemes known from Graph theory [2].

One difficulty in combining range data from different frames is the matching between consecutive frames, also called the correspondence problem. How can we match data from one frame to the next? We match objects from different frames by requiring them to be close together in space. This procedure works well for small scene changes between consecutive frames.

4.2 Ageing of 3D Points

To avoid smearing of the moving objects in the global map, old 3D points have to be deleted from the map after a certain time period. 3D points that have other 3D points from a more recent frame in their vicinity are kept longer in the map than "loners". That implicitly removes outliers and accidentals from the map quickly. This approach is also able to handle clipping and occlusions automatically.

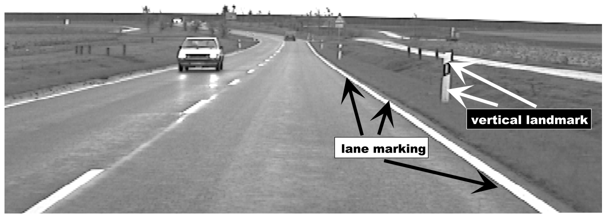

Figure 1: Typical traffic scene.

5 Ego-Motion Estimation

5.1 Introduction

We are trying to improve the imprecise velocity and steering angle data by using visual cues. This extends the ego-motion estimation to skidding situations.

For dead reckoning, our basic assumption is that all aligned 3D points close to the ground are stationary (**lane markings**). In addition, we assume that a group of 3D points vertically aligned alongside the road are stationary as well (**vertical landmarks**, see Figure 1), e.g. traffic lights, traffic signs, reflection posts, lampposts and trees. These primitives are input to our dead reckoning algorithm. Matching of the current frame information with previous ones is performed using the Extended Kalman Filter.

5.2 Introduction to the Kalman Filter

The Kalman filter [8] is a set of mathematical equations that provides an efficient solution to the discrete-data linear filtering problem. If the measurement or process relations are not linear, the Extended Kalman Filter is used to attack the problem. For the algorithms presented here, we follow the notation convention of [13] where the basic filter equations can also be found.

5.3 Dead Reckoning Using Vertical Landmarks

Finding Vertical Landmarks. Vertical Landmarks were chosen as reference objects for dead reckoning because they appear frequently along the road. Also, they have a unique and simple signature: Their 3D points are aligned in a vertical line. Searching through a list of 3D points and checking for points with similar x and z coordinates is highly discriminative in 3D.

System Description. We estimate the following states:

$$\vec{x} = \begin{bmatrix} x & z & \Phi & x_p & z_p \end{bmatrix}^T, \tag{1}$$

where x, z, and Φ are position and orientation of the ego-vehicle, respectively. x_p and z_p are the x and z-positions of the measured point. The measured point refers to the measured vertical landmark projected onto the x-z-plane.

Measurements are taken for $x_{p,l}$ and $z_{p,l}$, the measurement of the vertical landmark from the moving vehicle, generated by averaging over the x and z coordinates of the vertically aligned 3D points:

$$\vec{z} = \begin{bmatrix} x_{p,l} \\ z_{p,l} \end{bmatrix} \tag{2}$$

The steering angle δ constitutes the vector for the driving function.

$$u = [\delta] \tag{3}$$

Process description. Knowing the relation between radius r and steering angle δ from the extended Ackermann model, we can formulate the continuous system:

$$\dot{x} = v \cdot \sin\Phi, \tag{4}$$

$$\dot{z} = v \cdot \cos\Phi, \tag{5}$$

$$\dot{\Phi} = \frac{v}{r(\delta)}, \tag{6}$$

$$\dot{x_p} = 0, \tag{7}$$

$$\dot{z_p} = 0. \tag{8}$$

The control function must be linearized and hence the control matrix B appears as follows:

$$B = \left[\frac{v \cdot dt}{a+b}\right]. \tag{9}$$

Measurement Description. The measurement update step incorporates the new measurements. $x_{p,l}$ and $z_{p,l}$ can be expressed in terms of state variables:

$$h(\vec{x}) = \begin{bmatrix} x_{p,l} \\ z_{p,l} \end{bmatrix} = \begin{bmatrix} (x_p - x)\cos\Phi - (z_p - z)\sin\Phi \\ (x_p - x)\sin\Phi + (z_p - z)\cos\Phi \end{bmatrix}. \tag{10}$$

These equations and their derivatives w.r.t. the state vector are the input to the Kalman filter. Time update is performed at each time step. Measurement update steps are performed whenever new measurements are available. Variances are estimated using relations between pixel noise and distance/offset (z/x). The results are shown in Section 6.2.

5.4 Dead Reckoning Using Lanes

Finding Lane Markings. Finding lane markings is a standard procedure in intelligent vehicle applications (see e.g. [3]). For our purpose we use 3D points extracted with our stereo camera system and perform a Hough transform of all 3D points that lie close to the ground plane (flat road assumption). Only straight lane markings are considered in the current model.

We update line data into our map from the lane markings every frame using only the steering angle and the velocity in the first step (time update). In the second step, we match the old lane information in the map with our current lane information and correct our ego-position accordingly (measurement update).

Measurement Description. The system and process description of this system is equivalent to the system described in the previous section. Only x_p and z_p have to be interchanged with d and θ in the matrices, where d and θ stand for distance of the line to the global reference frame origin and orientation of the line w.r.t. the global reference frame, respectively.

We measure d_l and θ_l, the parameters of the lane marking in the local reference frame described by the line parameters

$$\begin{aligned} d_l &= d - x \cdot cos(\theta) - z \cdot sin(\theta), &(11)\\ \theta_l &= \theta - \Phi. &(12) \end{aligned}$$

The computation of the Jacobian of the measurement equations is straightforward. Measurement variances of d_l and θ_l are considered roughly constant since the variances are bound by the Hough transform parameters. In addition, both Kalman filters can be combined resulting in a state vector with seven states, which is expected to yield better results.

A more detailed analysis of these dead reckoning algorithms including more results can be found in [7].

6 Results

6.1 Cartography Results

Figure 2 depicts a sequence with only stationary objects. Figure 3 shows that superimposing several frames in the global map facilitates object detection tasks for stationary objects compared to using only the local map (see Figure 4). On the other hand, Figure 5 shows a sequence where clustering on the global map leads to heavy object smearing. A truck traveling at $25m/s$ is depicted there. The recognition with our stereo system based on the global map is shown in Figure 6 on the left which exhibits a severe smearing effect and which leads to a wrong distance determination. Figure 6 on the right shows the sequence with only using the local map. Note, that in both sequences

Figure 2: Stationary back court sequence.

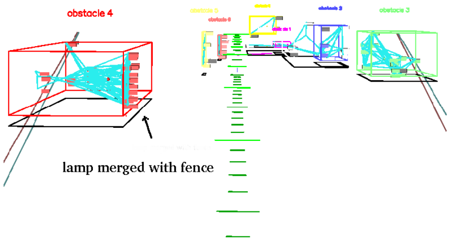

Figure 3: Global map representation of the back court sequence. Note that all objects protruding from the ground are detected.

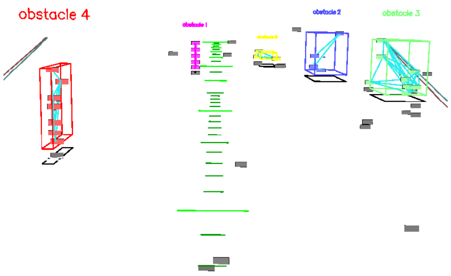

Figure 4: Local map representation of the back court sequence. Note that one lamppost and the fence on the left are not recognized (too few 3d measurements).

Figure 5: Dynamic freeway sequence with a truck.

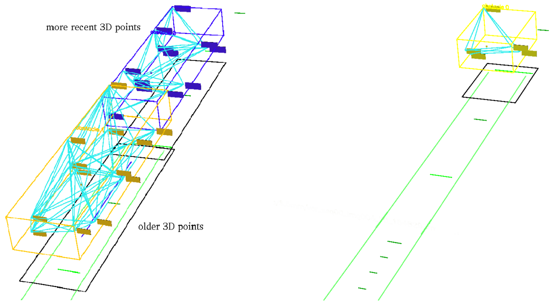

Figure 6: Global (left) and local (right) map representation of the freeway sequence. Due to old 3D points from previous frames the position is estimated too close to the ego-vehicle on the left.

all stereo features from 5 frames are accumulated in the global map.

For the reason of better visibility, the ageing of the 3D points is omitted. The global map also allows to extrapolate objects that disappeared out of the sensor range. In order to use the global map for control, the coordinates of the objects have to be transformed back to the local reference frame [6].

6.2 Dead Reckoning Simulation Results

Only simulated data can easily be compared to simulated ground truth to evaluate the Kalman filters.

Our simulation uses the vehicle model as described in Section 3 and allows for erroneous sensor readings of velocity, steering angle, and 3D measurements. In our simulated scenario, the ego-vehicle starts accelerating from 0 to $1m/s$ within one second and the velocity remains constant for the remainder of the simulation, whereas the steering angle takes values ranging from -8° to 8°.

Figure 7 shows a result for the lane-based Kalman filter. Here, the steering angle sensor delivered an offset of 2°. The measured lane marking is located $2m$ to the right with an orientation of 0°. The error on the distance parameter was set to $50cm$ and the orientation error to 2°(evenly distributed). The steering angle offset can easily be compensated and leads to a stable localization for the ego-vehicle (solid line). Pure motion integration leads to a large and increasing error (dashed line).

6.3 Dead Reckoning Real World Results

The dead reckoning algorithm using vertical landmarks has been run successfully in our research vehicle.

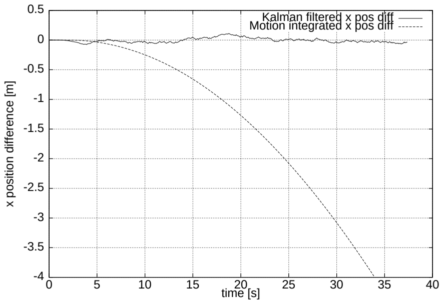

Figure 7: Deviation of the x position of the ego-vehicle compared to ground truth (simulated data). See text for details.

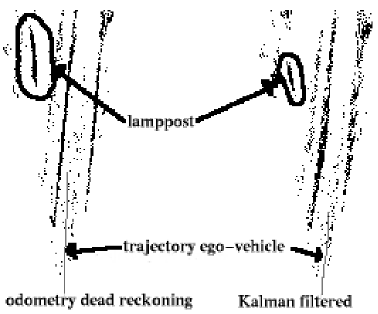

Figure 8: Stationary sequence with pure motion integration (left) and with Kalman filtering (right).

Superimposing local maps by pure motion integration causes the stationary object to smear (see Figure 8, left side). The x-z view in the figure shows all 3D measurements accumulated for a back court sequence with 100 frames. This "smearing" effect is reduced significantly with the dead reckoning algorithm using vertical landmarks (see Figure 8, right side).

7 Conclusions and Future Work

Cartography of the environment of an autonomous car is a beneficial procedure to superimpose vision clues from several frames, especially for stationary objects. More care must be taken for moving objects. The dead reckoning algorithms proposed in this paper improve the superposition of 3D points from different frames. Preliminary real world results indicate that motion integration errors are significantly reduced.

Future work includes tuning the algorithms and fusing the Kalman filters for vertical landmarks and lane markings. In addition, our proposed algorithm for deleting 3D points from the global map will be implemented in the near future.

References

[1] J. Borenstein, Y. Koren. *Real-Time Obstacle Avoidance for Fast Mobile Robots. IEEE Transact. Systems, Man, and Cybernetics*, 1989.

[2] T. H. Cormen, C. H. Leiserson, R. L. Rivest. *Introduction to Algorithms.* MIT Press, 1st Edition, 1990.

[3] E. D. Dickmanns, et al. *The Seeing Passenger Car VaMoRs-P.* In *Proceedings of the Intelligent Vehicles 94 Symposium*, 1994.

[4] U. Franke, et al. *Fast Stereo Object Detection for Stop and Go Traffic.* In *Proceedings of the Intelligent Vehicles 96 Symposium*, 1996.

[5] U. Franke, et al. *Autonomous Driving Approaches Downtown. IEEE Intelligent Systems and their Applications*, 1998.

[6] S. K. Gehrig, F. J. Stein. *A Trajectory-Based Approach for the Lateral Control of Vehicle Following Systems.* In *Proceedings of the Intelligent Vehicles 98 Symposium*, 1998.

[7] S. K. Gehrig, F. J. Stein. *Dead Reckoning and Cartography Using Stereo Vision for an Autonomous Car.* In *International Conference on Intelligent Robots and Systems*, 1999.

[8] R. E. Kalman. *A new approach to linear filtering and prediction problems. Transactions ASME J. of Basic Engineering*, 1960.

[9] L. Matthies, A. E. Elfes. *Integration of Sonar and Stereo Range Data Using a Grid-Based Representation.* In *Proceedings of the IEEE Conference on Robotics and Automation 88*, 1988.

[10] H. Moravec, A. E. Elfes. *High Resolution Maps from Wide Angle Sonars.* In *Proceedings of the IEEE Conference on Robotics and Automation 85*, 1985.

[11] K. Uchimura, Z. Hu. *Lane Detection and Tracking Using Estimated Camera Parameters for Intelligent Vehicles.* In *Proceedings of the Intelligent Vehicles 98 Symposium*, 1998.

[12] R. Wagner, K. Donner, F. Liu. *A "Half-Perspective" Approach to Robust Ego-Motion Estimation for Calibrated Cameras.* Technical report MIP-9718, University of Passau, 1997.

[13] G. Welch, G. Bishop. *An Introduction to the Kalman Filter.* Technical report, University of North Carolina at Chapel Hill, 1995.

[14] A. Zomotor. *Fahrwerktechnik: Fahrverhalten.* Vogel Buchverlag Würzburg, 1st edition, 1987.