Dead Reckoning and Cartography Using Stereo Vision for an Autonomous Car

Stefan K. Gehrig
Research Institute DaimlerChrysler AG
FT3/AB, HPC T 728
70546 Stuttgart, Germany
Phone: ++49-711-17-41484
Fax: ++49-711-17-47054
Stefan.Gehrig@DaimlerChrysler.Com

Fridtjof J. Stein
Research Institute DaimlerChrysler AG
FT3/AB, HPC T 728
70546 Stuttgart, Germany
Phone: ++49-711-17-41880
Fax: ++49-711-17-47054
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 (velocity and steering angle) can cause non-negligible errors, especially in situations where side slip or skidding occurs. We use stationary points in the scene to support the determination of our ego-motion. Two types of stationary objects are used: Firstly, stationary vertical landmarks such as traffic signs are used to compensate errors in our localization prediction. Secondly, lane markings measured in consecutive frames are used to compensate orientation errors.

Preliminary results show that dead reckoning using stationary objects can vastly improve self-localization.

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.

Our approach matches 3D-data from one frame to the next and generates object data out of this map. Cartography is a common procedure in the field of robotics, mostly used for partially known indoor environments. The road scenario differs from the indoor scenario in several ways.

- The environment around a vehicle on the road is completely unknown.
- Many fast moving objects appear in the scene.
- Cartography should be performed significantly faster than nowadays video frame rate ($40ms$).
- For vision based sensors extremely varying illumination conditions must be handled.

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 introduced. Section 4 describes the cartography based on 3D points. Dead reckoning techniques for an unknown environment are detailed in Section 5. Simulation results of our dead reckoning algorithms and some preliminary cartography results 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 [9]. This mapping technique is also referred to as certainty grids or occupancy grids (see e.g. [1, 8]). It basically 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 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 in [10] and [11]. In contrast to our research, they utilized only one camera.

3 Motion Integration

3.1 Transformation into a Reference Frame at Rest

For the subsequent analysis, 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:

$$\begin{pmatrix} x_n \\ z_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ z_{n-1} \end{pmatrix} + M_{rot}(\Phi) \begin{pmatrix} 0 \\ v_n \end{pmatrix} \cdot \Delta t. \quad (1)$$

Here Φ denotes the current orientation of the ego-vehicle w.r.t. the reference frame at rest, $\Delta\Phi$ denotes the change in orientation:

$$sign(\delta) \cdot \frac{v_n \cdot \Delta t}{r} = \Delta\Phi_n, \quad \Phi = \Sigma\Delta\Phi_i, \quad (2)$$

where n refers to the current time step, $n-1$ to the previous one. r is the radius of curvature of the ego-vehicle corresponding to the current steering angle. δ denotes the steering angle at the wheel of the vehicle. v_n is the velocity of the ego-vehicle at time step n. Φ is set to 0 at the time of initialization.

In the second step, the 3D measurements are transformed into the global reference frame:

$$\begin{pmatrix} x_m \\ z_m \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} + M_{rot}(\Phi) \begin{pmatrix} x_{m,l} \\ z_{m,l} \end{pmatrix}, \quad (3)$$

where the subscript m refers to measured 3D coordinates and l refers to the measurement in the moving reference frame.

3.2 The Vehicle Model

The radius of curvature r is derived from the steering angle δ using the kinematic Ackermann model [13]. In order to make the vehicle model also valid for higher velocities, the self-steering gradient of the ego-vehicle is taken into account:

$$x_c = \frac{a + b + SSG \cdot v^2}{\tan(\delta)}, \quad (4)$$

$$r = \sqrt{x_c^2 + b^2}. \quad (5)$$

SSG denotes the self-steering gradient in $[^\circ \cdot s^2/m]$, v is the velocity of the ego-vehicle. See Figure 1 for an explanation of the other used quantities. The Ackermann model with the extension mentioned above is limited to small lateral accelerations.

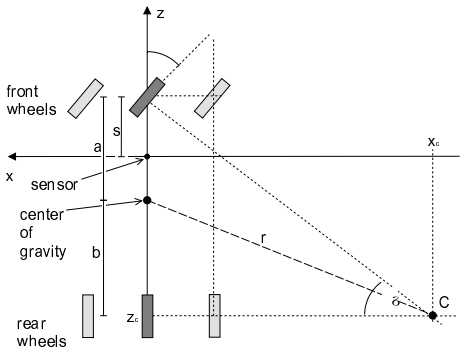

Figure 1: Ackermann Model (w: width of car, a: distance center to front axle, b: distance center to rear axle).

4 Cartography Based on 3D Points

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.

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.

For dead reckoning, our basic assumption is that all 3D points close to the ground are stationary (e.g. moving shadows on the ground would violate this assumption). In addition, we assume that a group of 3D points vertically aligned alongside the road are stationary as well. This holds true for traffic lights, traffic signs, reflection posts, and trees, but a skinny pedestrian walking alongside the road might be mistaken as stationary. These primitives are input to our dead reckoning algorithm explained in Section 5.

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.

5 Ego-Motion Estimation

5.1 Introduction

We do neither use an absolute positioning system such as GPS nor a gyroscope to measure the yaw rate which leaves us with only the steering angle and the velocity to integrate the ego-motion.

The ego-motion estimation algorithms of [10] and [11] rely heavily on good statistics since they need at least 50 significant points in a scene to produce a good estimation result. Hence, calculation times of these algorithms with a fast processor are very high. However, we need to perform an ego-motion estimate very fast. Using the sensor data for velocity and steering angle (see Section 3), our largest uncertainty lies in the estimation of the yaw angle and of small velocities.

We decided to use stationary points in the scene to determine our ego-motion. Two types of stationary objects are used (see Figure 3).

Vertical Landmarks: The first type of object are vertical landmarks such as traffic signs, which are used to compensate errors in our localization prediction.

Lane markings: The second type are lane markings, which are used to compensate orientation errors.

Matching of the current frame information with previous ones is performed using the Extended Kalman Filter. This method is explained in the following section.

5.2 Introduction to the Kalman Filter

The Kalman filter [7] 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 [12]. We assume for simplicity that our measurements are independent from each other. The Kalman filter algorithm consists of two steps: A time update step ("prediction") which uses only the system dynamics for prediction and a measurement update ("correction") step which incorporates the new measurements.

5.3 Dead Reckoning Using Vertical Landmarks

5.3.1 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 vertically aligned. Searching through a pre-sorted list of 3D points and checking for points with similar x and z coordinates is highly discriminative in 3D and very fast. Typical vertical landmarks that are suitable for dead reckoning purposes are traffic signs, reflections posts and traffic lights. Trees along the road are also used for that purpose.

5.3.2 System Description

We estimate the following states:

$$\vec{x} = \begin{bmatrix} x & z & \Phi & x_p & z_p \end{bmatrix}^T, \tag{6}$$

where x, z, and Φ are described in Section 3.1. 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 3D-points of the local reference frame, i.e. the raw measurement of the vertical landmark from the moving vehicle:

$$\vec{z} = \begin{bmatrix} x_{p,l} \\ z_{p,l} \end{bmatrix}. \tag{7}$$

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

$$u = [\delta]. \tag{8}$$

5.3.3 Process description

Using the equations derived in Section 3, we can formulate the continuous system:

$$\dot{x} = v \cdot \sin\Phi \quad (9)$$

$$\dot{z} = v \cdot \cos\Phi \quad (10)$$

$$\dot{\Phi} = \frac{v}{r(\delta)} \quad (11)$$

$$\dot{x_p} = 0 \quad (12)$$

$$\dot{z_p} = 0 \quad (13)$$

From that we can derive the Jacobian $A_{continuous}$ of the process functions

$$A_{continuous} = \begin{bmatrix} 0 & 0 & v \cdot \cos\Phi & 0 & 0 \\ 0 & 0 & -v \cdot \sin\Phi & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (14)$$

By means of discretization one obtains the transition matrix A. The control function must be linearized and hence the control matrix B appears as follows:

$$B = \begin{bmatrix} \frac{v \cdot dt}{a+b} \end{bmatrix}. \quad (15)$$

To compute the predicted state $\vec{x}_k^-$ the exact equations 10 to 13 are used in discretized form:

$$\vec{x}_{k+1} = f(\vec{x}_k, u_k). \quad (16)$$

To compute the error covariances, the transition matrix A is used:

$$P_{k+1}^- = A_k P_k A_k^T + Q_k, \quad (17)$$

where P^- is the error covariance projected ahead and Q is the system covariance matrix describing the system uncertainty.

5.3.4 Measurement Description

The measurement update step incorporates the measurements. We measure $x_{p,l}$ and $z_{p,l}$, the coordinates of the vertical landmark. These quantities can be expressed in terms of state variables (see Figure 2):

$$h(\vec{x}) = \begin{bmatrix} x_{p,l} \\ z_{p,l} \end{bmatrix} = \begin{bmatrix} \Delta x \cos\Phi - \Delta z \sin\Phi \\ \Delta x \sin\Phi + \Delta z \cos\Phi \end{bmatrix}, \quad (18)$$

where Δx and Δz are shorthands for $(x_p - x)$ and $(z_p - z)$, respectively The Jacobian H of the measurement equations is straightforward to compute. These equations and their derivatives w.r.t. the state vector are the input to the Kalman filter. The Kalman gain K is computed using

$$K_k = P^- H_k^T (H_k P^- H_k^T + R_k)^{-1}, \quad (19)$$

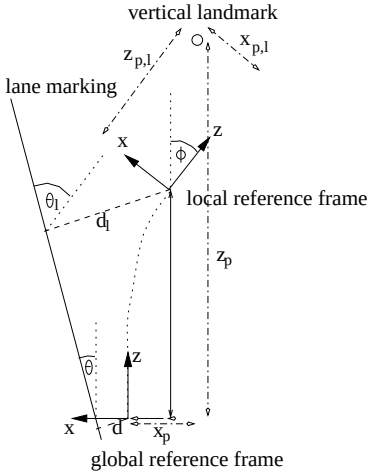

Figure 2: Vertical Landmark and Lane Marking measured from the global and the local frame.

where R is the measurement covariance matrix describing the measurement uncertainty. The corrected state estimate is derived from

$$\vec{x}_k = \vec{x}_k^- + K_k(\vec{z}_k - h(\vec{x}_k^-)), \quad (20)$$

and the corrected error covariance matrix is

$$P_k = (I - K_k H_k) P_k^-. \quad (21)$$

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). Simulation results are shown in Section 6.2.

5.4 Dead Reckoning Using Lanes

5.4.1 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).

5.4.2 Measurement Description

The system and process description of this system is equivalent to the system described in the previous sec-

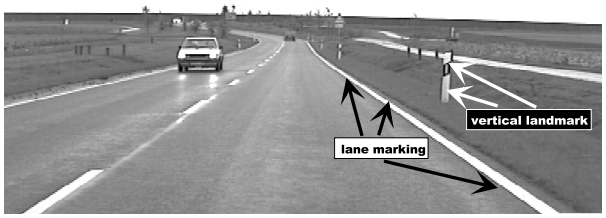

Figure 3: Typical traffic scene.

tion. 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

$$d_l = d - x \cdot cos(\theta) - z \cdot sin(\theta), \tag{22}$$

$$\theta_l = \theta - \Phi. \tag{23}$$

Consult Figure 2 for the geometrical relations. The computation of the Jacobian H 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.

6 Results

6.1 Cartography Results

Figure 3 depicts a typical traffic scene with both lane markings and vertical landmarks available for dead reckoning. The superposition of the local maps with pure motion integration is accurate for a few frames. With an increasing number of frames entering the global map, smearing of the objects in the map occurs due to positioning errors. In order to use this 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 3.2 and allows for erroneous sensor readings of velocity, steering angle, and 3D measurements.

In our simulated scenario the ego-vehicle takes the path depicted in Figure 4. 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 5 gives an example of the

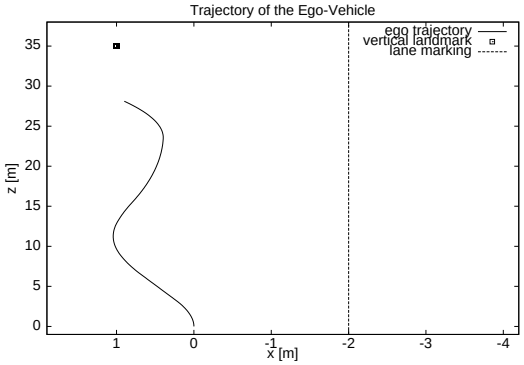

Figure 4: Profile of the simulated path (solid line). The square indicates the measured landmark, the dashed line shows the measured line.

dead reckoning performance. In the simulation, the velocity sensor always delivers a velocity 5% larger than the actual velocity. In addition, a 2% measurement error on the x and z-position of the vertical landmark was added (evenly distributed). In the simulated sequence, a vertical landmark was approached from a $35m$ distance on a zig-zag path. The solid line shows the Kalman filtered result of the z position of the ego-vehicle with deviations of less than $60cm$ at all times. Without the measurement feedback, pure motion integration leads to a much larger and increasing error (dashed line). Figure 6 shows a result for the lane-

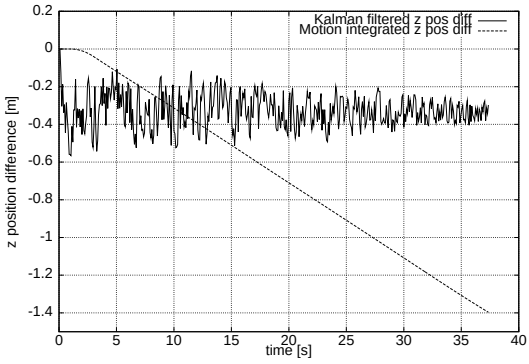

Figure 5: Deviation of the z position of the ego-vehicle compared to ground truth (simulated data). See text for details.

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). All other simulation parameters and the ego-vehicle path

were kept the same. 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).

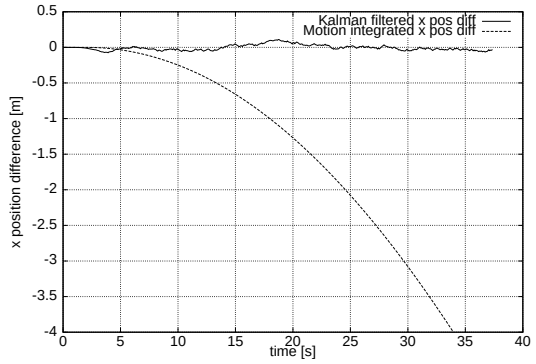

Figure 6: Deviation of the x position of the ego-vehicle compared to ground truth (simulated data). See text for details.

6.3 Dead Reckoning Real World Results

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

Superimposing local maps by pure motion integration causes the stationary object to smear (see Figure 7, 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 7, 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. 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.

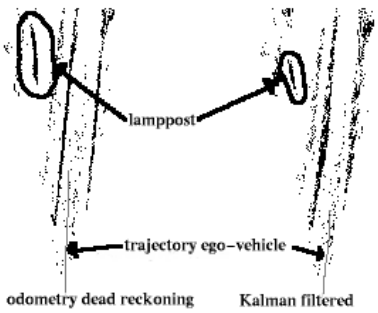

Figure 7: Stationary Scene with pure motion integration (left) and with Kalman filtering (right).

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

[8] 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.

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

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

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

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

[13] A. Zomotor. *Fahrwerktechnik: Fahrverhalten.* Vogel Buchverlag Würzburg, 1987.