Robust Monocular Egomotion Estimation Based on an IEKF

Frank Pagel
Fraunhofer Institute for Information and Data Processing
Autonomous Systems and Machine Vision
76131 Karlsruhe, Germany
frank.pagel@iitb.fraunhofer.de

Abstract

In this contribution a robust approach for the estimation of the camera motion is presented. For this purpose, features from a monocular image sequence are extracted and evaluated so that the threedimensional path of a moving camera can be calculated. The algorithm gives robust results even in the presence of noise and independently moving objects. The two different categories of constraint equations used in the proposed algorithm are the epipolar constraint and the trilinear constraints. The optimization of the constraints with respect to the motion parameters is implemented as a robust Iterated Extended Kalman Filter. Test results are presented from real data, captured from a moving vehicle in an urban scenario.

1. Introduction

Egomotion estimation is widely used in a variety of tasks. Several approaches exist already in literature about how to calculate the motion parameters given only the data of an image sequence, e. g. [1]. Once the motion parameters between the single frames are known one can perform a 3D reconstruction ([2]) and generate depth maps by fusing the single depth images ([3]).
[4] presented a stratified approach to metric self calibration. Although the results looked good it is very difficult to implement this algorithm on a standard PC hardware to run in real-time.
Very promising results were presented recently by [5]. The core of this approach, called MonoSLAM, is the online creation of a sparse but persistent map of natural landmarks within a probabilistic framework. The algorithm runs very fast with about 30Hz on standard PC hardware.

Especially in monocular motion estimation tasks the difficulty of independently moving objects appears. This may lead to wrong estimates. Therefore it is important to find a solution that handles outliers reliably. Outliers in this paper are feature correspondences that belong to either an independently moving object or that are simply wrong temporal feature matches. These outliers should be detected and not be considered in the further estimation process. There are mathematical techniques like the epipolar and trifocal geometry that directly relate the motion parameters and the 2D image data (see [6]). These geometric equation can be used to constrain the motion parameters. The constraints are embedded into a robust and well-known state tracker. An Iterated Extended Kalman Filter (IEKF) with a robust innovation step is used. Using an IEKF, additional sensor data, e. g. odometry data, can be easily integrated. Even more, an extension of the parameter vector can be done, e. g. the scales of the motion trajectories of several cameras. A similar IEKF was also successfully used in [7]. There, additional to the epipolar and trifocal constraints, a bundle adjustment was implemented ([8]). The drawback of the bundle adjustment is the increase of state parameters as for each scene point the depth value is additionally estimated. Furthermore earlier experiments have shown that in cases of a bad parameterization the bundle adjustment tends to converge to zero. Although in [7] an active stereo rig was continously calibrated , the basic principles could also be used for a monocular motion estimation.
The motion tracker in this paper is designed to estimate the camera motion reliably in real world environments even in the presence of independently moving objects. The IEKF can easily be varied by integrating other constraint functions or additional sensor information. Furthermore, it is a well known feature of the Kalman Filter to provide the uncertainty of the estimated state vector. This is an important feature for the automatic analysis of the tracker results.

2 Motion Model

Assuming a static background the motion of the camera is considered as an Euclidean motion (fig 3). Therefore the motion between two time frames is modeled via a rotation

978-0-7695-3651-4/09 $25.00 © 2009 IEEE
DOI 10.1109/CRV.2009.12

matrix $\mathbf{R}(\alpha, \beta, \gamma)$ with $\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$ and a translation vector $\mathbf{t} = (t_x, t_y, t_z)^T$, so that

$$\mathbf{X}_{k+1} = \mathbf{R}^T\mathbf{X}_k - \mathbf{R}^T\mathbf{t}, \tag{1}$$

where $\mathbf{X}_k \in \mathbb{R}^3$ is a point in the coordinate system of $\mathbf{c}_k$.

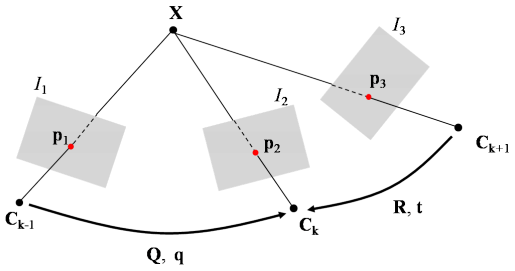

Figure 1. Motion model for three consequtive frames. R, t are the Euclidean motion parameters to be estimated. Q, q are the motion parameters from the previous time step.

3 Constraint Functions

In this approach the only available information is the image data. Therefore sparse optical flow vectors between three consecutive frames are used ([9], see fig. 2). As these flow vectors depend on the camera motion itself, the measured image motion has to satisfy some restrictions that will be used in a Kalman Filter. Here the well-known epipolar (for two cameras) and trifocal constraint (for three cameras) are used ([6]).

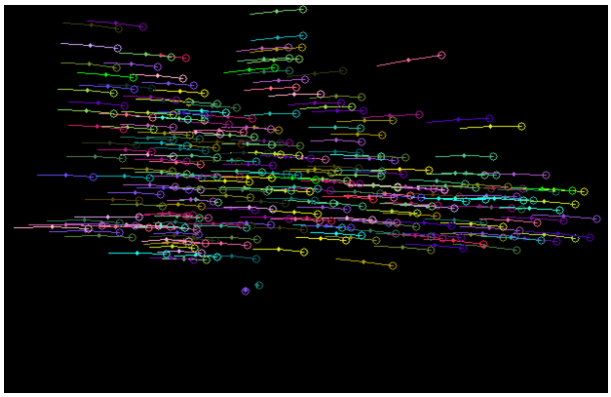

Figure 2. The measured data for the IEKF: Optical flow for three consecutive frames.

3.1 Epipolar Constraint

Assuming constant intrinsic camera parmeters, the fundamental matrix F and the essential matrix E are given by

$$\mathsf{F} = \mathbf{A}^{-T}\mathbf{R}[\mathbf{t}]_\times\mathbf{A}^{-1} \tag{2}$$

respective

$$\mathsf{E} = \mathbf{A}^T\mathsf{F}\mathbf{A}. \tag{3}$$

$\mathbf{A}$ is the intrinsic camera matrix containing the focal length and image center, $\mathbf{R}$ and $\mathbf{t}$ are the motion parameters between two consecutive frames and $[\mathbf{t}]_\times$ is the singular, skew-symmetric matrix of $\mathbf{t}$. For any two corresponding image points $(\mathbf{p}_{1_i}, \mathbf{p}_{2_i}) = \mathbf{z}_i$ between two frames the epipolar constraint

$$\mathbf{p}_{1_i}^T\mathsf{F}\mathbf{p}_{2_i} = 0 \tag{4}$$

is satisfied. Here the fundamental matrix can be seen as a function of the extrinsic and intrinsic camera parameters $\mathbf{x} = (f, \alpha, \beta, \gamma, t_x, t_y, t_z)^T$, so the constraint function

$$h_E(\mathbf{x}, \mathbf{z}_i) := \mathbf{p}_{1_i}^T\mathsf{F}(\mathbf{x})\mathbf{p}_{2_i} \stackrel{!}{=} 0 \tag{5}$$

can be defined. The interpretation of this constraint is that a pair of corresponding points have to lie within the same epipolar plane and therefore on its relating epipolar lines (see fig. 3).

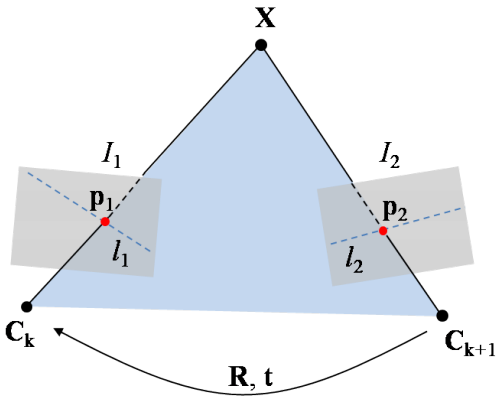

Figure 3. Projection of a 3D point X onto the image planes I_1 and I_2 after an Euclidean motion R, t. The epipolar plane is given by the three points X, C_1 and C_2. The epipolar lines l_1 and l_2 are the intersections of the epipolar plane with the image planes.

3.2 Trifocal Constraint

The trifocal constraint is a constraint for corresponding points between three cameras. In this section the basic

equations that are used in the Kalman Filter are derived. For a detailed derivation of the equation please see [6].
The projection of a 3D point $\mathbf{X} = (X, Y, Z)^T$ to an image point $\mathbf{p} = (x, y)^T$ is given by

$$\tilde{\mathbf{p}}_j \cong \mathbf{A}[\mathbf{R}_j|\mathbf{t}_j]\tilde{\mathbf{X}} = \mathbf{P}_j\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{P}_j^1 \\ \mathbf{P}_j^2 \\ \mathbf{P}_j^3 \end{bmatrix} \tilde{\mathbf{X}}, \tag{6}$$

where $\cong$ means equal up to a scalar factor. $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{p}}_j$ are homogenous coordinates. $\mathbf{R}_j, \mathbf{t}_j$ respective $\mathbf{P}_j$ are the motion parameters between two consecutive time frames at time j, $\mathbf{p}_j$ is the projectied image point at time j. From eq. 6 follows

$$\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} \mathbf{P}_j^1\tilde{\mathbf{X}}/\mathbf{P}_j^3\tilde{\mathbf{X}} \\ \mathbf{P}_j^2\tilde{\mathbf{X}}/\mathbf{P}_j^3\tilde{\mathbf{X}} \end{bmatrix} \Leftrightarrow \tag{7}$$

$$\begin{bmatrix} x_j\mathbf{P}_j^3 - \mathbf{P}_j^1 \\ y_j\mathbf{P}_j^3 - \mathbf{P}_j^2 \end{bmatrix} \tilde{\mathbf{X}} =: \begin{bmatrix} \mathbf{B}_j^1 \\ \mathbf{B}_j^2 \end{bmatrix} \tilde{\mathbf{X}} = \mathbf{0} \Rightarrow \tag{8}$$

$$\begin{bmatrix} x_1\mathbf{P}_1^3 - \mathbf{P}_1^1 \\ y_1\mathbf{P}_1^3 - \mathbf{P}_1^2 \\ x_2\mathbf{P}_2^3 - \mathbf{P}_2^1 \\ y_2\mathbf{P}_2^3 - \mathbf{P}_2^2 \\ x_3\mathbf{P}_3^3 - \mathbf{P}_3^1 \\ y_3\mathbf{P}_3^3 - \mathbf{P}_3^2 \end{bmatrix} \tilde{\mathbf{X}} =: \begin{bmatrix} \mathbf{B}_1^1 \\ \mathbf{B}_1^2 \\ \mathbf{B}_2^1 \\ \mathbf{B}_2^2 \\ \mathbf{B}_3^1 \\ \mathbf{B}_3^2 \end{bmatrix} \tilde{\mathbf{X}} =: \mathbf{D}_{6\times4}\tilde{\mathbf{X}} = \mathbf{0}. \tag{9}$$

$\mathbf{B}_j^1$ is the plane that contains the camera center at time j as well as the line given by $x_j = const$. Analogously the plane $\mathbf{B}_j^2$ is defined by the camera center and the line $y_j = const$. The homogenous eq. 9 has non-trivial solution if and only if the columns of $\mathbf{D}$ are linearly independent. This is given if the 4×4 minors of $\mathbf{D}$ are equal to zero. This leads us to the constraints

$$h_{T1}(\mathbf{x}, \mathbf{z}_i) := det([\mathbf{B}_1^1, \mathbf{B}_1^2, \mathbf{B}_2^1, \mathbf{B}_3^1]^T) \stackrel{!}{=} 0 \tag{10}$$

and

$$h_{T2}(\mathbf{x}, \mathbf{z}_i) := det([\mathbf{B}_1^1, \mathbf{B}_1^2, \mathbf{B}_2^1, \mathbf{B}_3^2]^T) \stackrel{!}{=} 0. \tag{11}$$

These constraints shall guarantee that the four planes $\mathbf{B}_1^1, \mathbf{B}_1^2, \mathbf{B}_2^1, \mathbf{B}_3^1$ respective $\mathbf{B}_1^1, \mathbf{B}_1^2, \mathbf{B}_2^1, \mathbf{B}_3^2$ intersect in one single point. Fig. 4 shows the geometric interpretations of the last equations.

4 Robust Iterated Extended Kalman Filter

In this paper an Iterated Extended Kalman Filter (IEKF) is used as described in [10]. Therefore all noise is assumed as normal distributed noise.

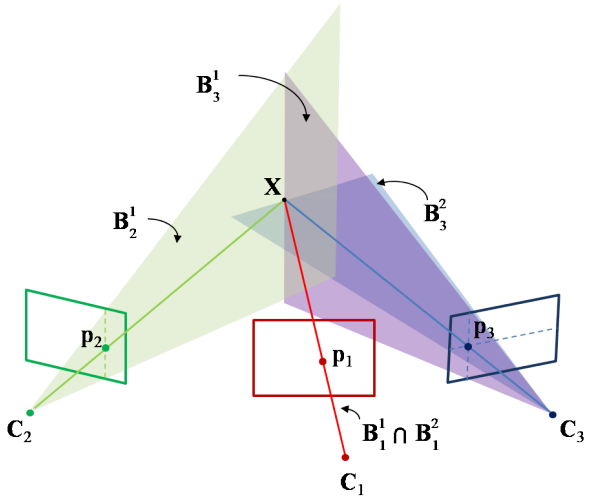

Figure 4. Geometric interpretation of eq. 10 and 11.

4.1 State Parameters

The motion parameters to estimate are the three translation parameters t_x, t_y, t_z and the three rotation angles α, β, γ. Furthermore the focal length f is estimated. To calculate the trifocal constraint the motion parameters of the previous time steps are needed, too (see fig. 1). So these parameters are simply added as additional arguments instead of estimating them again by adding them to the state vector. By reducing the dimension of the state vector the robustness of the estimation process increases as well as reduces computational time.
This approach has the advantage that the absolute translation norm w. r. t. the norm of the translation from the previous time step can be calculated (but still only up to scale as nothing is known about the scene). This is enforced by the trifocal constraint. If the parameters $\mathbf{Q}$ and $\mathbf{q}$ are already wrong, the estimation of $\mathbf{R}$ and $\mathbf{t}$ will be probably wrong, too. However, as the Kalman Filter provides measures to evaluate the estimation result, like e. g. the residuals of the constraint functions or the covariance matrix of the Kalman Filter, such cases can be caught and some algorithmic counteraction can be started. So, for example, a second Kalman Filter based only on the epipolar constraint could be run to estimate $\mathbf{R}$, $\mathbf{t}$ without the data from the previous time step.

4.2 System and Measurement Model

There is a state model

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}) \text{ with } \mathbf{w}_k \sim N(\mathbf{0}, \mathbf{\Sigma}_\mathbf{w}), \tag{12}$$

a measurement model

$$\hat{\mathbf{z}}_k = \mathbf{z}_k + \mathbf{e}_k \text{ with } \mathbf{e}_k \sim N(\mathbf{0}, \mathbf{\Sigma_e}) \tag{13}$$

and the constraint function

$$\mathbf{h}(\mathbf{x}, \mathbf{z}) = \mathbf{0} \tag{14}$$

at time k. For each measurement triple $\mathbf{z}_i = (\mathbf{p}_{1_i}, \mathbf{p}_{2_i}, \mathbf{p}_{3_i})^T \in \mathbb{R}^6$, $i = 1, ..., N$, N is the number of measurements and $\mathbf{z} = (..., \mathbf{z}_i^T, ...)^T$, eq. 5, 10 and 11 are used to define

$$\mathbf{h}(\mathbf{x}, \mathbf{z}) = \begin{bmatrix} \vdots \\ \mathbf{h}_i(\mathbf{x}, \mathbf{z}) \\ \vdots \end{bmatrix} \text{ with} \tag{15}$$

$$\mathbf{h}_i(\mathbf{x}, \mathbf{z}) = \begin{bmatrix} \mathbf{h}_E(\mathbf{x}, \mathbf{z}_i) \\ \mathbf{h}_{T1}(\mathbf{x}, \mathbf{z}_i) \\ \mathbf{h}_{T2}(\mathbf{x}, \mathbf{z}_i) \end{bmatrix} \tag{16}$$

Furthermore the assumptions

$$\mathbb{E}\{\mathbf{w}_i {\mathbf{w}_j}^T\} = \mathbb{E}\{\mathbf{e}_i {\mathbf{e}_j}^T\} = \mathbb{E}\{\mathbf{e}_i {\mathbf{w}_i}^T\} = \mathbf{0} \tag{17}$$

are made with $i, j \in \mathbb{N}, i \neq j$.
The prediction step is performed using a nonlinear function

$$\hat{\mathbf{x}}_k^- = \mathbf{f}(\hat{\mathbf{x}}_{k-1}^+, \mathbf{0}) \tag{18}$$

$$\mathbf{P}_k^- = \mathbf{F}_\mathbf{x} \mathbf{P}_{k-1}^+ \mathbf{F}_\mathbf{x}^T \tag{19}$$

with

$$\mathbf{F}_\mathbf{x} = \partial \mathbf{f} / \partial \mathbf{x}|_{\hat{\mathbf{x}}_{k-1}^-, \mathbf{0}}. \tag{20}$$

The iterated innovation step is performed as follows:
First set $\check{\mathbf{x}}_{k,0} := \hat{\mathbf{x}}_k^-$, $\check{\mathbf{z}}_{k,0} := \hat{\mathbf{z}}_k$.
For $l = 0...L-1$ do

$$\check{\mathbf{x}}_{k,l+1} = \hat{\mathbf{x}}_k^- - \mathbf{K}_{k,l} \mathbf{r}_{k,l} \tag{21}$$

$$\check{\mathbf{z}}_{k,l+1} = \hat{\mathbf{z}}_k - \mathbf{\Sigma_e} \mathbf{H}_{\mathbf{z},l}^T [\mathbf{H}_{\mathbf{z},l} \mathbf{\Sigma_e} \mathbf{H}_{\mathbf{z},l}^T]^{-1} \mathbf{r}_{k,l} \tag{22}$$

with

$$\mathbf{H}_{\mathbf{x},l} = \partial \mathbf{h} / \partial \mathbf{x}|_{\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}} \tag{23}$$

$$\mathbf{H}_{\mathbf{z},l} = \partial \mathbf{h} / \partial \mathbf{z}|_{\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}} \tag{24}$$

$$\mathbf{K}_{k,l} = \mathbf{P}_k^- \mathbf{H}_{\mathbf{x},l}^T [\mathbf{H}_{\mathbf{x},l} \mathbf{P}_k^- \mathbf{H}_{\mathbf{x},l}^T + \mathbf{H}_{\mathbf{z},l} \mathbf{\Sigma_e} \mathbf{H}_{\mathbf{z},l}^T]^{-1} \tag{25}$$

$$\begin{aligned} \mathbf{r}_{k,l} = \mathbf{h}(\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}) + \\ \mathbf{H}_{\mathbf{z},l} \cdot (\hat{\mathbf{z}} - \check{\mathbf{z}}_{k,l}) + \\ \mathbf{H}_{\mathbf{x},l} \cdot (\hat{\mathbf{x}}_k^- - \check{\mathbf{x}}_{k,l}) \end{aligned} \tag{26}$$

After L iterations set

$$\hat{\mathbf{x}}_k^+ := \check{\mathbf{x}}_{k,L} \tag{27}$$

$$\mathbf{P}_k^+ := [\mathbf{I} - \mathbf{K}_{k,L-1} \mathbf{H}_{\mathbf{x},L-1}] \mathbf{P}_k^-. \tag{28}$$

In cases when the residual or the covariance of the previous Kalman step is too big, the previous state estimation is likely to be wrong. Therefore the parameters **Q**, **q** cannot be reliably used for the estimation of the current parameters **R** and **t** (see fig. 1). Point corresondences of the current frame can be used to estimate the fundamental matrix (eq. 2) using a RANSAC-based estimation algorithm. The intrinsic parameters are assumed to change only in slow rates and therefore to be constant over a small period of time. So they can be used to calculate the essential matrix (eq. 3) and hence the rotation and translation parameters by factorizing E ([11]), e. g. using a quaternion based technique as in [12]. As the translation vector after the factorization is scaled to 1, the translation parameters are scaled with the norm of **q**. The factorization can also be applied for an initialization of the motion parameters.

4.3 Robust Innovation

The epipolar and trifocal constraint functions are very sensitive to outliers. Outliers may be bad correspondences but also correspondences from independently moving objects. For such correspondences the constraints 5, 10 and 11 are not satisfied anymore so that such outliers have to be detected and rejected. The following robust innovation is implemented according to [7]:
From all constraint functions $\mathbf{h}_i$ Λ subsamples $S_m, m = 1, ..., \Lambda$, are randomly selected with λ data samples (correspondence triples) each. For each subsample S_m an innovation step as in eq. 27 and 28 is executed. The results are Λ state and covariance candidates $\hat{\mathbf{x}}_{k,m}, \mathbf{P}_{k,m}$. Then for each candidate the Mahalanobis distance can be calculated

$$\begin{aligned} \delta_m^i = & \mathbf{h}_i(\hat{\mathbf{x}}_{k,m}, \hat{\mathbf{z}}_{k,i})^T \cdot \\ & [\mathbf{H}_{\mathbf{x}_i} \mathbf{P}_{k,m} (\mathbf{H}_{\mathbf{x}_i})^T + \mathbf{H}_{\mathbf{z}_i} \mathbf{\Sigma}_{\mathbf{e}_i} (\mathbf{H}_{\mathbf{z}_i})^T]^{-1} \cdot \\ & \mathbf{h}_i(\hat{\mathbf{x}}_{k,m}, \hat{\mathbf{z}}_{k,i}). \end{aligned} \tag{29}$$

The best subsample is given by the index m^* which has the smallest median for all corresponding distances $\delta_{m^*}^i$:

$$m^* = \text{argmin}_{m=1...\Lambda} \{\text{median}_{i=1...N} \quad \delta_m^i\} \tag{30}$$

δ follows a χ^2 distribution. All measurements $\hat{\mathbf{z}}_i$ that do not satisfy the condition

$$\delta_{m^*}^i \leq \epsilon \tag{31}$$

are rejected, where one can get ϵ from a χ^2-table with significance level σ=1%. Finally all remaining measurement data are used to re-estimate the *a posteriori* state $\hat{\mathbf{x}}_k^+$ by performing the IEKF innovation step.

5 Experimental Results

To show the test results of the proposed system two sequences of real data are considered in this paper. First,

the camera-carrying vehicle executes a left turn (`Moltke`, see fig. 9), and second, a double bend turn is performed (`Bahnhof`, see fig. 7). The camera looks to the front of the car. Both sequences contain independently moving objects just like cyclists, pedestrians or other moving vehicles.
To get more robust flow vectors, the optical flow is calculated with 20Hz whereas the motion is estimated only with 4Hz. For the tests the number of iterations for the innovation step of the IEKF was set to $L = 3$. The noise parameters were set to $\mathbf{\Sigma_w} = 0.1\mathbf{I}$ and $\mathbf{\Sigma_e} = \mathbf{I}$.

5.1 Path Reconstruction

To calculate the motion path in a world coordinate system the current orientation of the camera must be accumulated. For this purpose the rotation matrix $\mathbf{R}_w(\varphi, \theta, \psi)$ is initially set to $\mathbf{R}_w(\varphi, \theta, \psi) = \mathbf{I}$. The global 3D position $\mathbf{X}_w$ is initially set to $(0, 0, 0)^T$. At each time step j the current orientation and position is updated

$$\mathbf{X}_w := \mathbf{X}_w - \mathbf{R}_w(\varphi, \theta, \psi)\mathbf{R}(\alpha_j, \beta_j, \gamma_j)_j^T \mathbf{t}_j \quad (32)$$

$$(\varphi, \theta, \psi)^T := (\varphi, \theta, \psi)^T - (\alpha_j, \beta_j, \gamma_j)^T. \quad (33)$$

5.2 Results with Real Data

Fig. 5 and 6 show image stills from the test scenes. The results are shown in fig. 8 and 11 respective 10 and 12. Fig. 8 and 10 show the reconstructed paths. The abscissas in 11 respective 12 show the time stamps from the Kalman Filter (4Hz update rate).
The camera carrying vehicle in fig. 8 is performing a double bend turn, that is a slight left turn immediately followed by a right turn. In the `Moltke`-sequence the ego vehicle performs a left turn. In fig. 10 the camera first moves straight forward, turns left and then goes straight forward again. It can be seen that even the angle of the trajectory could be estimated very well. The estimated tracks in fig. 8 and 10 obviously correspond to the reference paths in the real scenes in fig. 7 and 9. Notice that the translation parameters t_x, t_y, t_z in fig. 11 and 12 have no unit as the absolute scale of the motion of the camera cannot be determined.
The angle β can roughly be interpreted as the steering angle. In fig. 11 the estimation of β clearly shows the left (negative angle, time stamp 0-30) and right (positive angle, time stamp 30-66) steering of the vehicle. One can also see in fig. 12 when the driver of the car begins steering left (time stamp 33-77), shortly countersteers (time stamp 77-97) and then drives straight forward again. There is only little translational motion in x-direction but a big rotational motion around the y-axis.

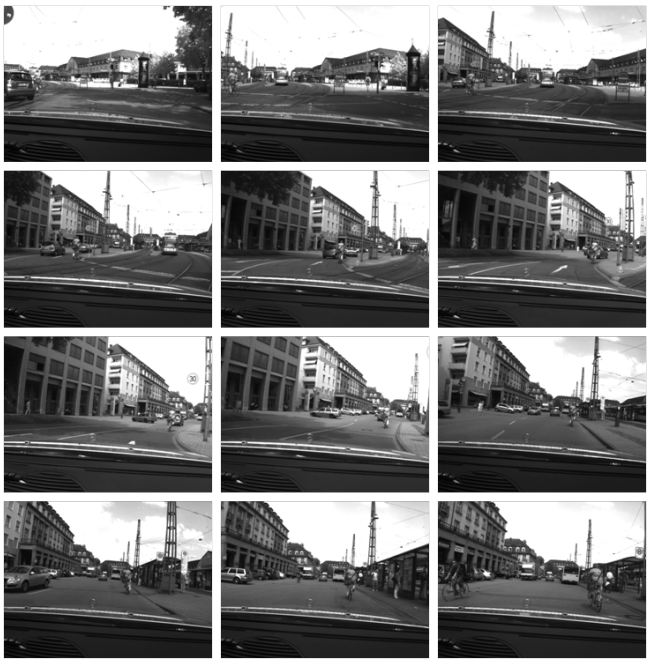

Figure 5. Stills from the `Bahnhof`-sequence (double bend turn).

Figure 6. Stills from the `Moltke`-sequence (left turn).

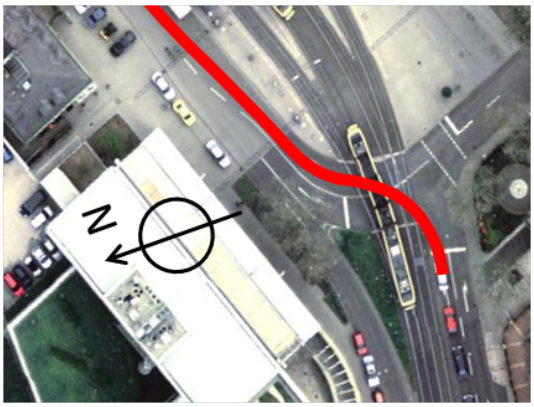

Figure 7. Top view of the `Bahnhof`-scene with reference trajectory.

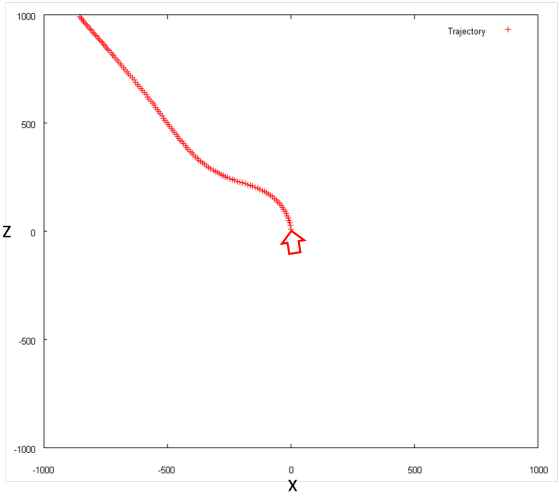

Figure 8. Top view of the estimated 3D motion trajectory of the `Bahnhof`-sequence. The 3D trajectory is projected into the XZ-ground plane of a world coordinate system.

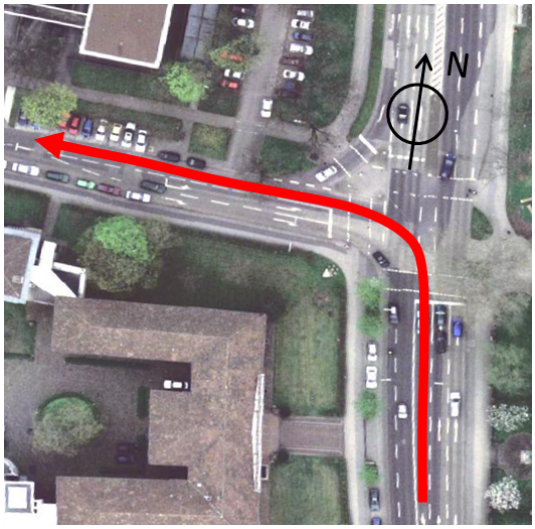

Figure 9. Top view of the `Moltke`-scene with reference trajectory.

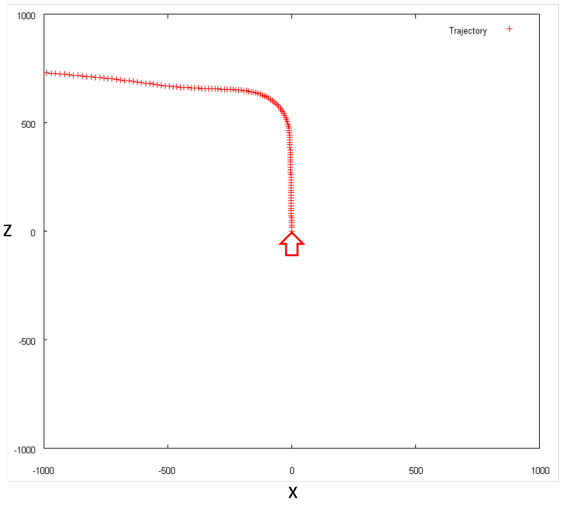

Figure 10. Top view of the estimated 3D motion trajectory of the `Moltke`-sequence. The 3D trajectory is projected into the XZ-ground plane of a world coordinate system.

6 Conclusion

An algorithm was proposed to perform a robust motion estimation of a single camera up to a scale factor. Therefore standard techniques are used as epipolar and trifocal geometry as well as a robust Iterated Extended Kalman Filter.
More tests with real data and investigations of the filter with respect its behavior with different parameter constellations will be done (e. g. estimation of the current and the previous motion parameters instead of estimating only the current ones). There are also other constraint functions that could be considered. Further investigations are planned using odometric data like absolute velocity and steering angle to stabilize the system.
A lot of future tasks base on the current work. One will be the implementation of a motion-based 3D reconstruction with multiple cameras on a moving platform. But also safety relevant applications in advanced driver assistant systems, e. g. collision warning, could be possible continuative developments of this approach.

References

[1] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," in *Nature*, 1981.

[2] R. Koch, M. Pollefeys, and L. Van Gool, "Multi viewpoint stereo from uncalibrated video sequences," in *ECCV*, 1998.

[3] P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, D. Nister, and M. Pollefeys, "Real-time visibility-based fusion of depth maps," in *ICCV*, 2007.

[4] M. Pollefeys, "Self-calibration and metric 3d reconstruction from uncalibrated image sequences," Dissertation, 1999.

[5] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "Monoslam: Real-time single camera slam," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 1052–1067, 2007.

[6] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[7] T. Dang and C. Hoffmann, "Tracking camera parameters of an active stereo rig," in *DAGM-Symposium*, 2006, pp. 627–636.

[8] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, *Bundle Adjustment - A Modern Synthesis*. London, UK: Springer-Verlag, 2000.

[9] J. Shi and C. Tomasi, "Good features to track," in *ICCVPR*, 1994.

[10] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Non-linear Approaches*. John Wiley & Sons, 2006.

[11] J. Weng, T. S. Huang, and N. Ahuja, "Error analysis of motion parameters from image sequences," in *ICCV*, 1987.

[12] O. D. Faugeras, G. Toscani, and F. Lustmann, "Motion and structure from motion from point and line matches," in *ICCV*, 1987.

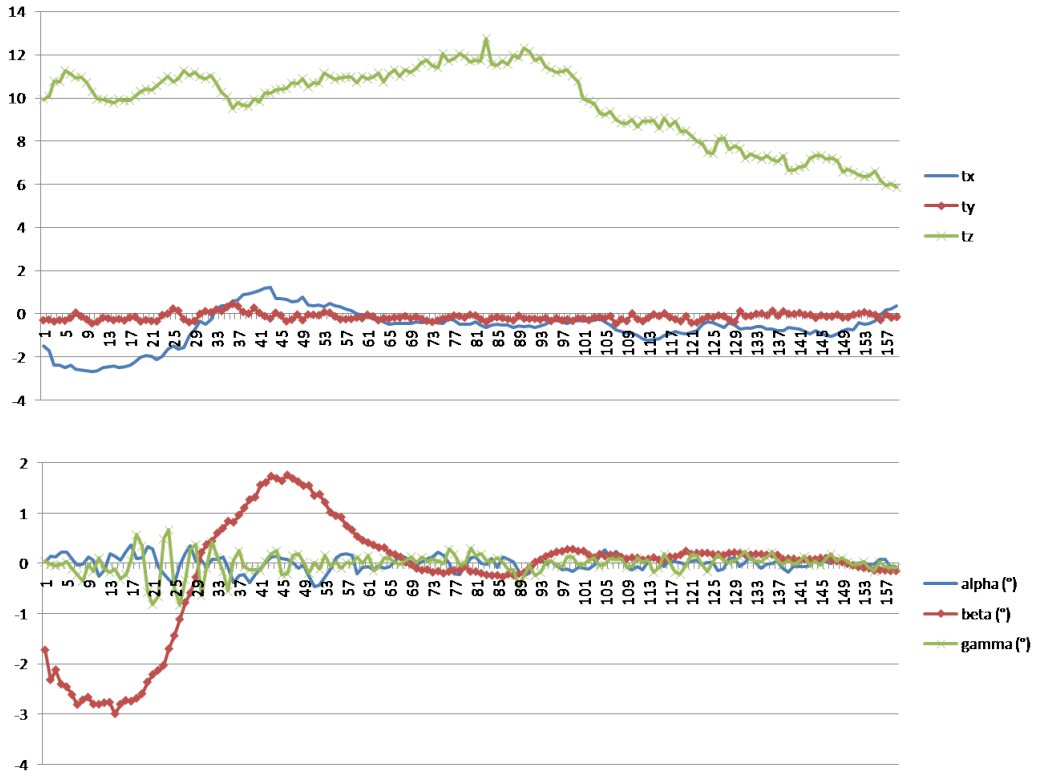

Figure 11. Diagrams of the estimated motion parameters of the `Bahnhof`-sequence.

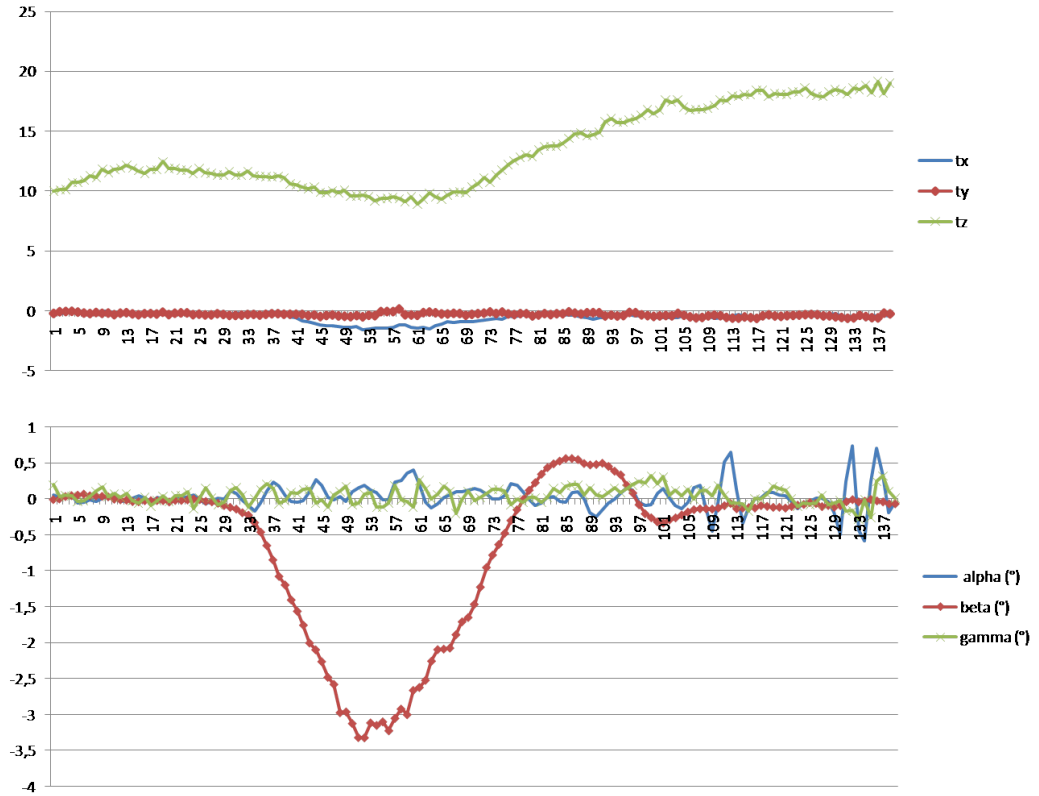

Figure 12. Diagrams of the estimated motion parameters of the `Moltke`-sequence.