AUTOMATIC 3D LANE MARKING RECONSTRUCTION USING MULTI-VIEW AERIAL IMAGERY

C.-Y. Sheu[*2], F. Kurz[1], P. Angelo[1]

[2] University Stuttgart, Germany -
chunyu.sheu@gmail.com
[1] German Aerospace Center (DLR), Remote Sensing Technology Institute, Wessling, Germany -
franz.kurz@dlr.de, pablo.angelo@dlr.de

Commission I, WG I/2

KEY WORDS: Aerial Image Sequences, 3D Line-features Reconstruction, Lane markings

ABSTRACT:

The 3D information of road infrastructures are gaining importance with the development of autonomous driving. The exact absolute position and height of lane markings, for example, support lane-accurate localization. Several approaches have been proposed for the 3D reconstruction of line features from multi-view airborne optical imagery. However, standard appearance-based matching approaches for 3D reconstruction are hardly applicable on lane markings due to the similar color profile of all lane markings and the lack of textures in their neighboring areas. We present a workflow for 3D lane markings reconstruction without explicit feature matching process using multi-view aerial imagery. The aim is to optimize the best 3D line location by minimizing the distance from its back projection to the detected 2D line in all the covering images. Firstly, the lane markings are automatically extracted from aerial images using standard line detection algorithms. By projecting these extracted lines onto the Semi-Global Matching (SGM) generated Digital Surface Model (DSM), the approximate 3D line segments are generated. Starting from these approximations, the 3D lines are iteratively refined based on the detected 2D lines in the original images and the viewing geometry. The proposed approach relies on precise detection of 2D lines in image space, a pre-knowledge of the approximate 3D line segments, and it heavily relies on image orientations. Nevertheless, it avoids the problem of non-textured neighborhood and is not limited to lines of finite length. The theoretical precision of 3D reconstruction with the proposed framwork is evaluated.

1. INTRODUCTION

The availability of large-scale, accurate high resolution 3D information of roads with lane markings and road infrastructure plays an important role towards autonomous driving. Aerial imagery is a valuable database to derive 3D information of roads even in areas difficult to access, like on motorways. Compared to optical satellite data, acquiring large-scale 3D lane markings by optical aerial imagery is more efficient and has higher accuracy and spatial resolution. In view of the fact, that in Germany exists no area-wide, high resolution 3D information of the road surfaces including lane markings, new methods to derive this information are demanded.

The standard workflow with aerial images would be to project the images onto a Digital Surface Model (DSM) and to derive the information in the projected imagery, but the generation of DSM from stereo images is challenging in the regions with low textures. The lane markings, for example, are the most visible texture on asphalt roads useful for 3D reconstruction. Thus, it is desired to improve the quality of the DSM on the road surfaces by exploiting the line character of the lane markings.

In this paper, a framework to automatically detect the lane markings in the unprojected aerial imagery, and refine the 3D information of the road surface by exploiting the line character of the lane markings is proposed.

The unprojected aerial images with their bundle-adjusted orientations and the DSM are the inputs of the algorithm. The aerial image dataset are acquired with special flight configuration at both sides of the motorway. Some standard pre-processing steps and a standard line detection algorithm for automatic lane marking detection in image space are applied. By sliding a window of reasonable length and width through the curved long lane lines, all line segments in all covering images are collected assuming the lane markings to be straight in each sliding window.

Similar to Taylors idea on minimization of the objective function (Taylor and Kriegman, 1995), the use of orthogonal regression to optimize the 3D position of each line segments in object space so that its back projection would best fit the detected 2D line in all the covering views is investigated, i.e. the position and height of each 3D lane marking segments will be refined in one optimization step. The proposed approach addresses the challenging (quasi) infinite and curved properties of lane markings in the 3D reconstruction.

The framework will be tested on aerial imagery from the German highway A9 taken on 29th March 2017 from the DLR helicopter BO-105.

2. RELATED WORK

For 3D line reconstruction, (Schmid and Zisserman, 1997, Bay et al., 2005, Wang et al., 2009) have tried to match line segments based on their appearances or some additional geometry constraints.

(Schmid and Zisserman, 1997) exploit the epipolar geometry of line segments and the one-parameter family of homographies to
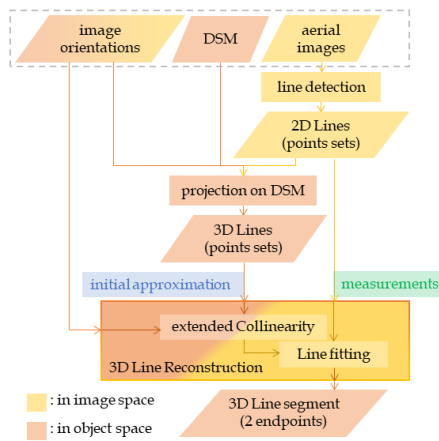
---

*Corresponding author

Figure 1. The work flow

Figure 2. Line types of lane markings

(a) Continuous line

(b) Dashed line

(c) Continuous and dashed double lines

(d) Continuous double lines

(e) Dashed double lines

	motorways	other roads
narrow lines	0.15 [m]	0.12 [m]
wide lines	0.30 [m]	0.25 [m]

Table 1. Widths of lane markings

	motorways	other roads	
		in town	out of town
line / gap	6 [m] / 12 [m]	3 [m] / 6 [m]	4 [m] / 8 [m]

Table 2. Lengths of dashed lane markings with ratio 1:2

provide point-wise correspondences, allowing cross-correlation of patches around line segments along the candidate lines in the epipolar-beam-region for matching scores evaluation.

In the cases of poorly textured or shape-changing neighborhood of line segments in different views, line segments are barely comparable using classical correlation patches yet the color neighborhood along this line segment undergoes only slight changes. Based on color histogram rather than textures, (Bay et al., 2005) exploit the appearance similarity of line segment pairs and their topological layout to iteratively increase the correct matches. While color provides a very strong cue for discrimination, it may fail in the case where color feature is not distinctive, e.g. gray images. Besides, although matching groups of line segments takes more geometric information into account for disambiguation, the disadvantage is the increased computational complexity.

Without resorting to any other constraints or prior knowledge, (Wang et al., 2009) propose a purely image content-based line descriptor MSLD (Mean-Standard deviation Line Descriptor) for automatic line segments matching. Adapting SIFT-like strategy, MSLD is highly distinctive and robust against image rotation, illumination change, image blur, viewpoint change noise, JPEG compression and partial obstruction.

In order to create 3D models without the need of explicit line matching, (Jain et al., 2010) generate all possible hypothetical straight 3D line segments by triangulating all the detected straight 2D line segments from different views. They then keep the one whose back projection on the gradient images of neighboring views has the highest score, assuming that line features correspond to high gradient areas in images. Built upon the same principles whilst applying epipolar constraint on the end-points of line segments, (Hofer et al., 2013) generate less hypothetical 3D line segments and thus increase performance significantly while still creating accurate results. However, both approaches are barely possible in the case of infinite line reconstruction, where the detected 2D lines in different views do not exactly correspond to the same part of a 3D line.

3. METHODOLOGIES

The following sections introduce the principles of 3D lane marking reconstruction method of this work, based on the work flow shown in Figure 1.
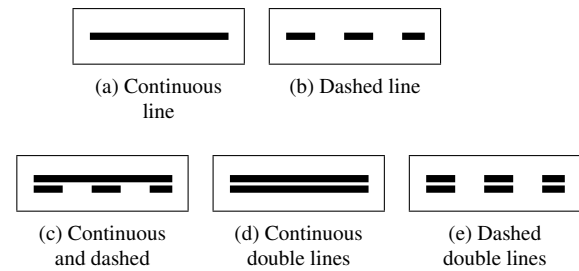
3.1 Lane markings Properties and Automatic Extraction

The appearance of lane markings on German roads including line type, color and width is specified depending on the road type. Different line types of lane markings are shown in Figure 2 and their line widths are defined in Table 1. As shown in Table 2, the dashed lane markings on motorways have 6 meter length.

Because of the appearance, the problem of lane marking detection can be treated as a line detection problem. We restrict the proposed framework to lane markings with single white lines (dashed or continuous) of 0.3 meter width. Other types like in restricted zone, double lines, parking areas, temporal yellow lines in construction sites etc, are excluded.

In this work, the principle to extract line features is to firstly derive the line direction for each pixel by using partial derivatives of a Gaussian smoothing kernel. Pixels that have a local maximum in the second directional derivative perpendicular to the line direction are marked as line points. By thresholding their second directional derivative values, the accepted line points are then linked and connected (Steger, 1998). The resulting connected points which compose a line are of sub-pixel precision. Figure 3 shows the extracted lines on part of the masked original image.

3.2 Line Fitting

Line fitting is the process of constructing an infinite straight line that has a best fit to a 2D dataset. To minimize the perpendicular distances from the data points to the regression line, and for a combination with point-wise extended collinearity equation, a orthogonal regression model with line equation in two-point form is derived in Section 3.2.1.

Regarding the fact that the extracted lines lie mainly in column direction in image space, the functional relation between variables x and y will be setup as $x = f(y)$ to avoid weakly solvable equations system.
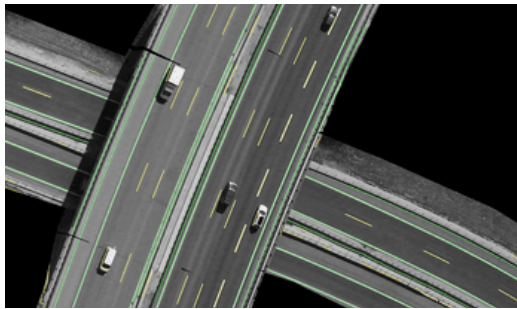
Figure 3. The extracted continuous lane-lines are marked in green and the dashed in yellow. Both cases are reconstructed into 3D with the same framework; different colors here are only for illustration.
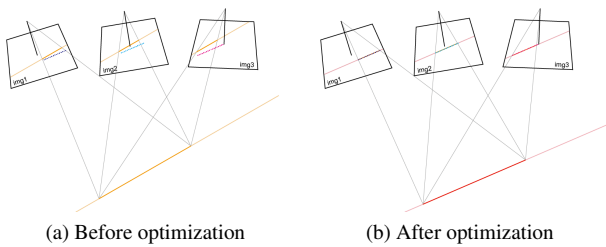

(a) Before optimization (b) After optimization

Figure 4. After optimization, the back-projection of the reconstructed line segment should best fit the extracted 2D lines in all the covering images.

3.2.1 Orthogonal Regression in Two-point Form Let the unknown coordinates of two different points on the regression line in 2D space be (x_1, y_1) and (x_2, y_2) where $y_2 \neq y_1$, and the observed 2D points be $\{x_i, y_i\}_{i=1}^{n}$ with measurement errors e_{x_i} and e_{y_i} in both variables. The orthogonal regression model in two-point form is:

$$x_i - e_{x_i} = (x_1 - \frac{(x_2 - x_1)}{(y_2 - y_1)} \times y_1) + \frac{(x_2 - x_1)}{(y_2 - y_1)} \times \bar{y}_i \quad (1)$$

$$y_i - e_{y_i} = \bar{y}_i \quad (2)$$

To express (1) and (2) shortly, a function $\mathcal{F}$ is defined as

$$\hat{\mathbf{p}} = \mathcal{F}(\mathbf{p}_s, \mathbf{p}_e, y) \quad (3)$$

which takes 2D coordinates of a start-point $\mathbf{p_s}(x_s, y_s)$ and an end-point $\mathbf{p_e}(x_e, y_e)$ that define an infinite line, and takes the measured y-coordinate y of an image point $\mathbf{p}(x, y)$, and returns the estimated image coordinates $\hat{\mathbf{p}}(\hat{x}, \hat{y})$ which lies on the infinite line $\overline{\mathbf{p_s} \mathbf{p_s}}$.

3.3 3D Line Reconstruction with Nonlinear LS Adjustment

In this section, the process of refining the position of a 3D line segment based on the extracted 2D lines on images and the viewing geometry is described, as illustrated in Figure 4.

The observation equations for LS adjustment are set up in Section 3.3.1 and constraint equations in Section 3.3.2. They describe the fitting of a straight line to the extracted lines in all covering images, where the fitting lines on different images are transformed from a single 3D straight line segment through
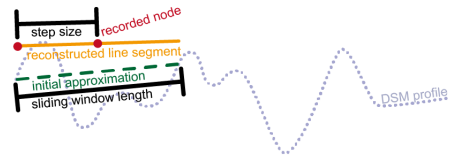

(a) The first line segment of "sliding window length" is reconstructed, with its starting point and its middle point of "step size" from the starting point being recorded.
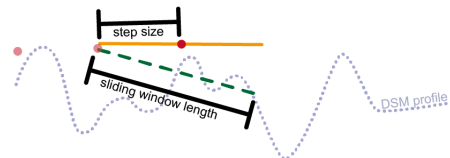
(b) Starting from the recorded node of previous process, another line segment is reconstructed, i.e. the sliding window has moved "step size" forward. Its middle point is then recorded and the step is repeated.

Figure 5. 3D reconstruction of a lane marking by a sliding window.
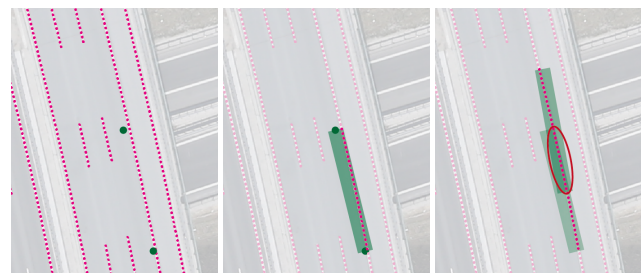
(a) Pink points represent the extracted lines. The greens are the endpoints of the back projected approximate 3D line segment.

(b) The points in the buffering area are collected as the measurements for LS adjustment.

(c) The red circle points out the reconsidered measurements in successive sliding windows.

Figure 6. collection of affected lane marking points

the extended collinearity equation. Regarding the fact that the collinearity is a point-wise condition, a line segment is represented by its two endpoints. In Section 3.3.3 it is further linearized and the substitute linear LS model is estimated.

To simplify the problem, a long lane-marking segment is partially reconstructed through a sliding window in the object space. Each segment is approximated by a straight line, taking into account the maximum curvature of the highway. The sliding scheme is illustrated in Figure 5.

The measurements for each reconstruction process are collected correspondingly, as shown in Figure 6 by back-projecting the initial line segment into image space and buffering 10 pixels width on each side. By this, all the extracted 2D line segments in this region are collected. As shown in Figure 6c, the reconsideration in overlapping region of successive sliding windows makes the reconstruction more robust.

3.3.1   Observation Equations   Given a start-point $\mathbf{P}_s$ and an end-point $\mathbf{P}_e$ of a line segment L in the object space and the camera parameters $\mathbf{q}^j$ of camera j. Consider the case where there are J images covering this line segment. With the extended collinearity equations, the start- and end-points of this line segment's back-projection in image j have the image coordinates $\mathbf{p}_s^j(x_s^j, y_s^j)$ and $\mathbf{p}_e^j(x_e^j, y_e^j)$:

$$\begin{aligned} \mathbf{p}_s^j &= \mathcal{G}(\mathbf{q}^j, \mathbf{P}_s) \\ \mathbf{p}_e^j &= \mathcal{G}(\mathbf{q}^j, \mathbf{P}_e) \end{aligned} \qquad \forall j = 1, 2, ...J \qquad (4)$$

Let l^j be the corresponding line segment of L being extracted (observed) on image j. Given a dataset $\{x_{l,i}^j, y_{l,i}^j\}_{i=1}^{N_l^j}$ of N_l^j points on line segment l^j, their estimated image coordinates $\hat{\mathbf{p}}_{l,i}^j$ on the infinite line $\overline{\mathbf{p}_s^j, \mathbf{p}_e^j}$ computed from the orthogonal regression model (3) are:

$$\hat{\mathbf{p}}_{l,i}^j = \mathcal{F}(\mathbf{p}_s^j, \mathbf{p}_e^j, y_{l,i}^j) \qquad \forall i = 1, 2, ...N_l^j \qquad (5)$$

Combining (4) with (5) gives function $\mathcal{H}$:

$$\begin{aligned} \hat{\mathbf{p}}_{l,i}^j &= \mathcal{F}(\mathcal{G}(\mathbf{q}^j, \mathbf{P}_s), \mathcal{G}(\mathbf{q}^j, \mathbf{P}_e), y_{l,i}^j) \\ &= \mathcal{H}(\mathbf{q}^j, \mathbf{P}_s, \mathbf{P}_e, y_{l,i}^j) \qquad \forall i = 1, 2, ...N_l^j, \quad \forall j = 1, 2, ...J \end{aligned}$$
$$(6)$$

which takes camera interior and exterior parameters $\mathbf{q}^j$, object coordinates of $\mathbf{P}_s$ and $\mathbf{P}_e$ which define a line $\overline{\mathbf{P}_s, \mathbf{P}_e}$, and the observed y-coordinate of the point $\mathbf{p}_{l,i}^j$ in image space, and returns the estimated image coordinates $\hat{\mathbf{p}}_{l,i}^j$ on the back projected line of $\overline{\mathbf{P}_s, \mathbf{P}_e}$.

Being written in the structure of the *Gauss-Markov model*, they are expressed as:

$$\boldsymbol{l} + \widehat{\boldsymbol{v}} = \boldsymbol{f}(\widehat{\boldsymbol{x}}): \quad \hat{\mathbf{p}}_{l,i}^j + \hat{\boldsymbol{v}}_{l,i}^j = \mathcal{H}(\mathbf{q}^j, \mathbf{P}_s, \mathbf{P}_e, y_{l,i}^j) \qquad (7)$$

with the amount of observations $N = 2 \times \sum_{j=1}^{J} N^j$.

The unknown parameters in the *Gauss-Markov model* are

$$\boldsymbol{x}: \quad \begin{bmatrix} X_s & Y_s & Z_s & X_e & Y_e & Z_e & y_1^1 & ... & y_{N^J}^J \end{bmatrix}^\mathsf{T} \quad (8)$$

with the amount of unknowns $U = 6 + \sum_{j=1}^{J} N^j$.

3.3.2   Constraint Equations   There are three constraints on the unknown parameters used in this work:

Fixing the X-, Y-coordinates of the start-point using the approximate values:

$$\hat{X}_s - X_s{}^0 = 0 \qquad (9)$$
$$\hat{Y}_s - Y_s{}^0 = 0 \qquad (10)$$

Fixing the length of the line segment (i.e. constraining the relative location of the end-point):

$$\sqrt{(\hat{X}_s - \hat{X}_e)^2 + (\hat{Y}_s - \hat{Y}_e)^2 + (\hat{Z}_s - \hat{Z}_e)^2} - S = 0 \quad (11)$$

Written in the structure of the *Gauss-Markov model with constraints* $\boldsymbol{h}(\widehat{\boldsymbol{x}}) = \boldsymbol{0}$:

$$\begin{bmatrix} \hat{X}_s - X_s{}^0 \\ \hat{Y}_s - Y_s{}^0 \\ \sqrt{(\hat{X}_s - \hat{X}_e)^2 + (\hat{Y}_s - \hat{Y}_e)^2 + (\hat{Z}_s - \hat{Z}_e)^2} - S \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$(12)$$

with the amount of constraints $H = 3$.

3.3.3   Least-Squares Estimation for 3D Line Reconstruction
The nonlinear equation system is approximated to be locally linear with small step size of the unknown quantities. The linearized form is expressed as:

$$\widehat{\Delta \boldsymbol{l}} = \Delta \boldsymbol{l} + \widehat{\boldsymbol{v}} = \underset{N \times U}{\mathsf{A}} \widehat{\Delta \boldsymbol{x}} \qquad (13)$$

$$\boldsymbol{c}_h = \underset{H \times U}{\mathsf{H}^\mathsf{T}} \widehat{\Delta \boldsymbol{x}} \qquad (14)$$

where
the $N \times U$ design matrix is the Jacobian of the function evaluated at the approximate values of the unknown parameters

$$\mathsf{A} = \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}}\bigg|_{\boldsymbol{x} = \widehat{\boldsymbol{x}}^a}$$

the $U \times H$ constraint matrix is the Jacobian of the constraints

$$\mathsf{H} = \left(\frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}} \right)^\mathsf{T}\bigg|_{\boldsymbol{x} = \widehat{\boldsymbol{x}}^a}$$

and the residual constraints are

$$\boldsymbol{c}_h = -\boldsymbol{h}(\widehat{\boldsymbol{x}}^a)$$

with the corrections

$$\Delta \boldsymbol{l} = \boldsymbol{l} - \boldsymbol{f}(\widehat{\boldsymbol{x}}^a) =: \widehat{\boldsymbol{v}}^a \qquad (15)$$

$$\widehat{\Delta \boldsymbol{x}} = \widehat{\boldsymbol{x}} - \widehat{\boldsymbol{x}}^a \qquad (16)$$

where $\widehat{\boldsymbol{x}}^a$ is the approximate values for the estimates of the unknown parameters.

In the *linearized substitute model* as shown in (13) and (14), it turns to solve for the increments of unknowns $\Delta \boldsymbol{x}$ instead of the unknowns themselves.

As the lines were extracted independently and with the same procedure, the measurements are assumed to be independent from each other and equal-weighted. That is, the weight matrix $\mathsf{W}_{ll} = $ is an identity matrix.

The unknown parameters $\widehat{\Delta \boldsymbol{x}}$ of the linearized model can be determined from the extended normal equation system

$$\begin{bmatrix} \mathsf{A}^\mathsf{T}\mathsf{W}_{ll}\mathsf{A} & \mathsf{H} \\ \mathsf{H}^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} \widehat{\Delta \boldsymbol{x}} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathsf{A}^\mathsf{T}\Delta \boldsymbol{l} \\ \boldsymbol{c}_h \end{bmatrix} \qquad (17)$$

With the iteration index ν and the approximate values in the first iteration $\widehat{\boldsymbol{x}}^{(1)} = \widehat{\boldsymbol{x}}^{(a)}$, we have

$$\widehat{\boldsymbol{x}}^{(\nu+1)} = \widehat{\boldsymbol{x}}^{(\nu)} + \widehat{\Delta \boldsymbol{x}}^{(\nu)} \qquad (18)$$
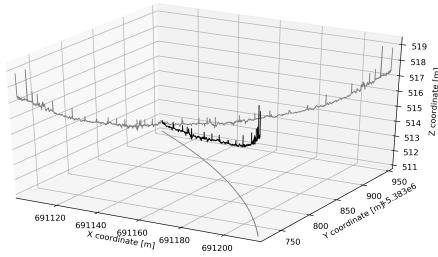
Figure 7. The projected line on DSM.

By updating the parameters using Equation (18), the LS estimation is applied iteratively until convergence is achieved.

The redundancy of the problem is

$$R = N + H - U = \sum_{j=1}^{J} N^j - 3$$

Two kinds of singular cases can happen. First, a configuration defect in object space can happen: as the 3D reconstruction approach still relies on intersection of multiple projection rays from different views, either the cases where there is only one image covering the targeted line segment, or when the targeted line segment lies on all of the epipolar planes of any of its two covering images, in these cases the problem is not solvable. Second, a configuration defect in image space can happen: as mentioned in Section 3.2, when the extracted line lies (nearly) in row direction on all the covering images, the problem is also not solvable. In the cases that the targeted line segment lies only on some of the stereo pairs' epipolar planes, the problem is still solvable whereas those stereo pairs are not contributing to the solution. Or similarly, when only in some of the images the extracted line segments lie (nearly) in row direction, the problem is solvable whereas those images are not contributing measurements to the estimation.

3.4 Line Projection on the DSM (Determination of Initial Parameter Estimates)

An initial approximation of the unknowns is required for convergence to the correct solution. To provide an initial 3D line segment, the extracted line features derived in Section 3.1 can be projected onto a DSM based on the bundle adjusted exterior and interior orientations. An example is shown in Figure 7.

Given image coordinates $\mathbf{p}(x, y)$ of a point and bundle-adjusted image orientations $\mathbf{q}$, there is still one degree of freedom in extended collinearity equation on solving object point $\mathbf{P}$. Combined with the usage of a DSM, which provides the height information Z_{DSM} given a position (X, Y), the corresponding object coordinates can be solved iteratively until the increment ΔZ is small enough, i.e. convergence achieved. The iterative scheme is illustrated in Figure 8.

Considering that the DSM is raster (discrete) whereas X and Y have continuous numerical values, the DSM height Z_{DSM} for the given point (X, Y) is bilinear interpolated.

4. EXPERIMENTAL RESULTS AND EVALUATIONS

Section 4.1 provides information about the used dataset. Section 4.2 describes the chosen parameters for line extraction, line
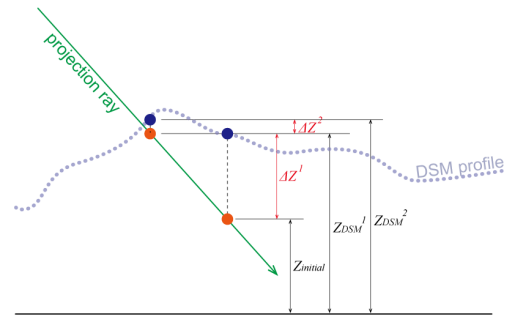

Figure 8. Iterative scheme of single projection ray intersecting DSM.

Canon EOS 1D-X	
Lenses	Zeiss Makro Planar f/2.0 50mm
Sensor / Pixel size	Full frame CMOS / 6.944 μm
Image size	5184×3456 pixel, ratio 3:2 (17.9 MPix)

Table 3. Oblique camera properties

	RGB, 50mm lens
Oblique angle	$\pm 15°$
FOV	$\pm 34°$ across strip, $\pm 13°$ along strip
Coverage @500m	780 m × 230 m
GSD @500m	6.9 cm (nadir)

Table 4. Viewing geometry

projection on DSM, and window sliding processes. Section 4.3 evaluates the true data results. The result of single continuous lane-marking is presented and the theoretical precision is evaluated.

4.1 Input Data

Aerial Images For real-time mapping applications during disasters, mass events and traffic monitoring scenarios, the German Aerospace Center (DLR) has developed a new optical sensor system – the 4k system – operated on a helicopter from DLR. The oblique aerial images used in this work are acquired from a Canon EOS 1D-X camera, one of the three non-metric cameras in the 4k system, with an oblique viewing angle τ of 15°. The pixel size is around 6.9 m, with the combination of focal length 50 mm and flying height H_{flight} around 500 m above ground, leading to a GSD of ~6.9 cm.

As described in (Fischer et al., 2017), this aerial imagery could be improved to have an absolute geolocation accuracy of better than 30 centimeters if TerraSAR-X geodetic points are included as reference points.

Table 3 lists the properties of this camera, and Table 4 provides the viewing geometry information.

The images used in this work are acquired with a special flight configuration at both sides of the motorway which guarantees a continuous stereo view perpendicular to the lane marking direc-
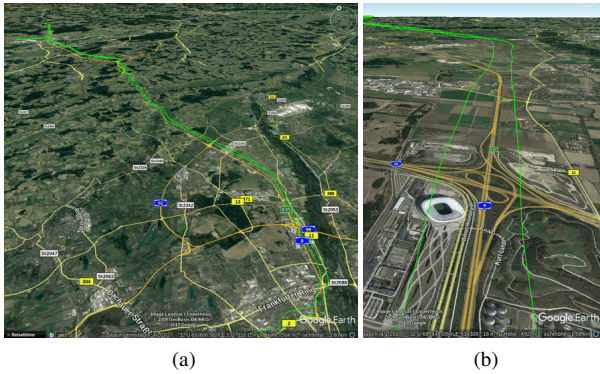
(a)   (b)

Figure 9. Flight trajectory of DLR helicopter visualized on the Google Earth platform. The green polyline shows the flight trajectory. *Source:* **Google Earth** *04/01/2017*

position precisions	[meter]	attitude precisions	[degree]
σ_{north}	0.055	σ_{roll}	0.002
σ_{east}	0.035	σ_{pitch}	0.002
$\sigma_{altitude}$	0.069	σ_{yaw}	0.005

Table 5. Precisions of Exterior Orientations

Interior Orientations		precisions		unit
focal length c	51 [mm]	σ_c	6.9	[μm]
principal point pp_x	-42.259	σ_{pp_x}	0.167	[μm]
principal point pp_y	115.384	σ_{pp_y}	0.799	[μm]
radial distortion A_1	-24.07	σ_{A_1}	0.01	-
radial distortion A_2	13042.30	σ_{A_2}	22.73	-

Table 6. Interior Orientations and their precisions

tion.[1] This is realized by flying at the right-hand side with respect to flying direction along the motorway, with the left oblique camera looking left-down to the motorway, in both forward and backward trips. The flight configuration is shown in Figure 9 on the Google Earth platform.

There are 15 images used in this work: 8 images in forward trip (first strip) and 7 in backward (second strip). The length of flight strips is around 890 meter. Besides, the forward overlap is around 70%, and all the lane markings are covered by both strips, whereas the side overlap depends on the distance of flight strips, which is a result of the pilots navigation ability and other influences, like wind. Nevertheless, the motorway in its entire width was covered by the two flight strips. Altogether, this results in approximately eightfold coverage in road areas.

Exterior and Interior Orientations The aerial images are geo-referenced by Global Navigation Satellite System GNSS/Inertial system IGI AEROcontrol-IId and further improved by Satellite Positioning Service of the German National Survey (SAPOS) correction. Additionally, a global terrain model (from Shuttle Radar Topography Mission (SRTM)) was introduced as pass information in the bundle adjustment, to improve the estimation of the focal length and boresight alignment. Self-calibrating bundle adjustment is applied to calibrate Interior Orientation (IO) parameters and to refine Exterior Orientation (EO) parameters. The self-calibration bundle adjustment was executed with 2107 images of one camera and 242000 tie points based on two overlapping flight strips. Table 5 and Table 6 show the precision of IO and EO parameters after self-calibrating bundle adjustment. To provide an overall quality on the interior orientations: from the calibration result of interior orientations (involving lens distortion), the residuals appear non-systematic and the biggest residual $r_{max,IO}$ is around 1 pixel.

The precision of the interior orientation parameters like focal length and principal point is quite high, but still realistic, as for each of the 242000 tie points a light-weighted height from the SRTM DEM was introduced into the bundle adjustment, which reduces the correlations significantly. Tangential distortion parameters are not taken into account, as they differ not significantly from zero. More details about the calibration can be found

in (Kurz et al., 2012). Without changing the sensor and viewing configuration, one calibration set can be used for different image acquisitions.

To evaluate the influence of the uncertainties in exterior and interior parameters on positioning precision in object space, the maximum values for each component based on the flight configuration was calculated. The quality of interior and exterior orientation parameters set would have a maximum impact in object space for around 16 [cm] in X,Y-direction:

- caused by inaccurate camera position:
$$\sqrt{\sigma_{north}^2 + \sigma_{east}^2} \approx 0.065 \text{ [meter]}$$

- caused by inaccurate camera attitude:
$$H_{flight} \times \sqrt{(\tan(\sigma_{roll} + \tau) - \tan\tau)^2 + (\frac{\tan\sigma_{pitch}}{\cos\tau})^2}$$
$$\approx 0.026 \text{ [meter]}$$

- caused by inaccurate Interior Orientations:
$$r_{max,IO} \times GSD = 0.069 \text{ [meter]}$$

and around 6.9 [cm] in Z-direction (caused by inaccurate camera position).

The above information tells the positioning precision in object space with measurements on a single image. With corresponding measurements from multi-view stereos, which allows the intersection of multiple projection rays from different directions, the positioning precision is expected to be improved. The final accuracy of object points of 4k images was validated in (Fischer et al., 2017).

DSM The DSM used in this experiment is generated by SGM based on multi-view stereos. As the asphalt road surfaces on which the lane markings are located are poorly textured, the SGM generated DSM is especially noisy in such poorly textured areas. However, such high-resolution DSM gives a good starting point for lane marking refinement. In other words, the DSM will be used only for setting up the initial values of the work flow, and will not influence the final results of 3D lane marking reconstruction.

The DSM has 20 cm grid spacing. Figure 10 shows a part of the DSM.
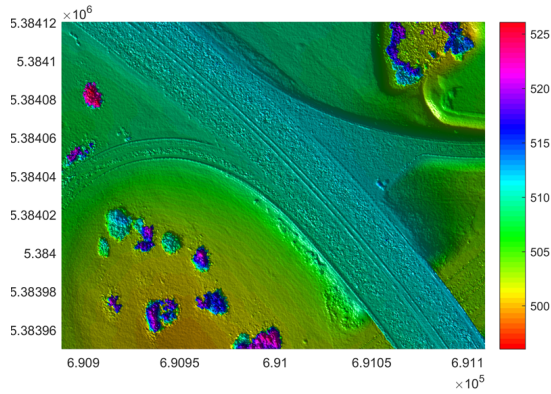
---

[1]The classical photogrammetric approach on flight planning is to have several straight flight lines which cover the whole motorway in a stereo view. This would be possible in this project yet would require more flight costs and would produce many more images.

Figure 10. Part of the DSM in road area. It is noisy in the center of motorway.

Figure 11. Masked Image

Road Masks Road segments are masked out from original images based on OpenStreetMap (OSM) data with 25 meter buffer width around road axes, as shown in Figure 11.

4.2 Parameter Selection

In lane marking extraction process, the σ value for Gaussian smoothing is set to 1.8 to slightly suppress the noise in images. A length threshold on the extracted lines is applied to reduce false positives. Regarding the fact that a dashed lane-line is around 6 meter long which corresponds to 62–87 pixels in image space (when parallel to one of the coordinate axes or in 45° direction), the extracted lines whose length is less than 65 pixels are rejected.

The length of the sliding window should be decided base on the expected curvature of the targeted line and the robustness of the reconstruction model. In the cases of continuous lane markings on motorways, the sliding window length was fixed to 16 m, or for the last segment it might be up to 24 m long, as a compromise between optimization robustness and the minimized systematic errors arisen from straight-approximated curvature. As to the cases of dashed lane-lines (6 m long), a sliding window is as long as its targeted approximating line, i.e. 6 m length. The step size of sliding window was set to half of the sliding window length.

The adapted lens distortion model is none shape-preserving, i.e. a 3D straight line is no more straight in image. However in every independent reconstruction process, only a short 16 m line segment is reconstructed, whose lens distortion correction along this line segment is no larger than one pixel. Thus the bending of a line segment on an image, arose from lens distortion, is ignored in this work, but could be easily removed from the images by rectifying the images (i.e. calculating non distorted images).
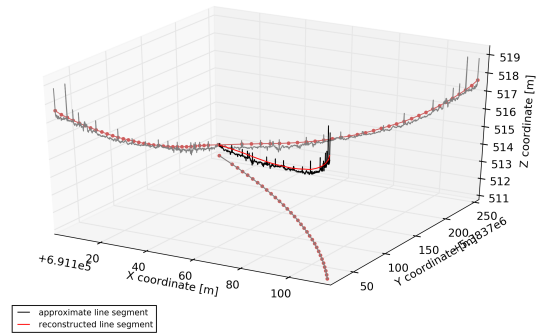

Figure 12. The reconstructed line segments and the unrefined DSM profile in UTM 32N coordinate system.

4.3 Results

A continuous lane marking of 258.7 meters length is reconstructed. Figure 12 shows the reconstructed line segments and the DSM profile. The distances from the reconstructed line segment to the DSM profile are collected along the reconstructed line segments with 0.2 meter spacing (considering the DSM grid of 0.2 meter), resulting in sample size of 1268. The sample mean is -0.162 [m] and the sample standard deviation is 0.119 [m].

Assuming DSM height profile being significantly lower than the reconstructed line segments for more than 15 centimeters, A lower-tailed Z-test is adopted. As a result, with 95% confidence we can claim that the DSM profile is in average, statistically and significantly lower than the reconstructed line segments for at least 15 centimeters in this region.

The posterior standard deviation of measurements is better than 0.7 pixel, the theoretical precision of the reconstructed lines is within 2.5 cm in vertical direction and within 5 mm in horizontal direction. The reconstructed lane-markings in the test area are shown in Figure 13 and Figure 14. They are further exploited to improve the DSM at lowly textured road surfaces, as shown in Figure 15.

5. CONCLUSION

We proposed a framework for automatic 3D lane marking reconstruction using multi-view aerial imagery. Standard line detection algorithms are applied to extract lane markings on the aerial images. By exploiting the use of linear regression in image space and with the combination of collinearity condition, lane markings are reconstructed based on their detected positions in images and the viewing geometry. Without the utilization of the neighboring textures, the approach requires initial approximate 3D lines and is highly dependent on quality of pre-known image orientations. Nevertheless, it is robust to partial occlusions of the targeted lines on images and is applicable to the (quasi-)infinite line features as well as in cases of lowly textured neighboring and it improves the DSM at lowly textured road surfaces.

From experimental results of true data in Section 4.3, some conclusions are given:

- The DSM profile, i.e. the initial 3D line approximation, is significantly and systematically dozens of centimeters far away from the reconstructed line segments. This indicates the necessity of reconstrucion based on detected 2D lines in

(a)
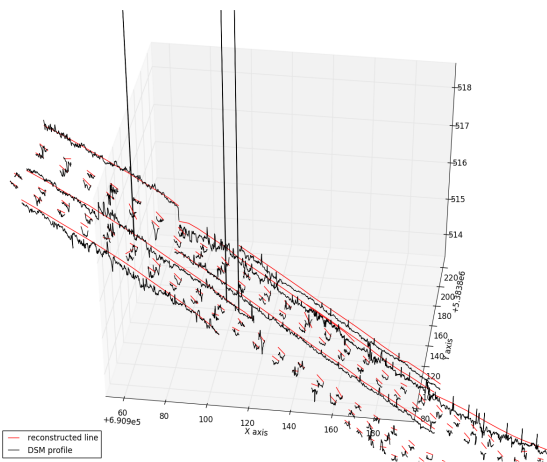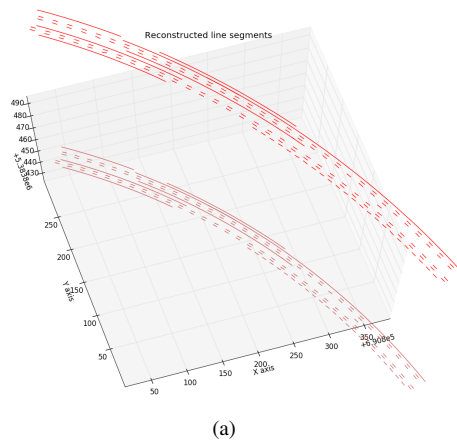
(b) Zoom in to show the difference between the DSM profiles and the reconstructed lines.

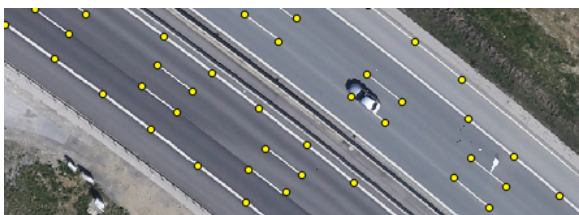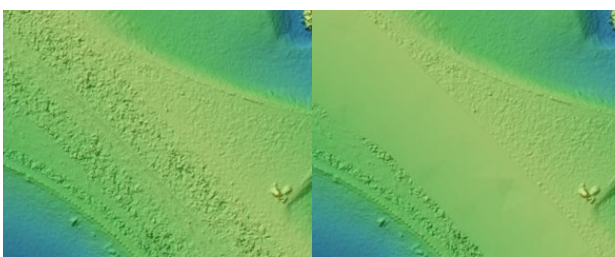Figure 13. All the reconstructed lane-lines in the test area.

Figure 14. The reconstructed line-nodes on aerial image, in UTM 32N coordinate system.

(a) The original DSM surface.   (b) The refined DSM surface.

Figure 15. With the reconstructed lane-lines, the DSM quality on road surfaces is improved.

image space and the viewing geometry, instead of simply reducing DSM noise by applying mean filter or such.

- The LS-estimated precision of the measurements, involving the lane marking extraction quality as well as the quality of image orientation parameters, is better than 1 pixel.

- The configuration defect barely happens: 1.when the linear regression functional model between image coordinates x and y are properly set up according to the characteristics of the detected lane-lines in image space. 2.when the flight configuration guarantees stereo views whose base-lines are perpendicular to lane-lines orientations in 3D.

- Theoretical precision of reconstructed lines is within 2.5 cm in vertical direction and within 5 mm in horizontal direction.

Some general conclusions are:

- Image configuration plays an important roll in 3D reconstruction. In the case of this work, being covered by more views whose base-lines are more perpendicular to lane-lines directions, would improve the reconstruction result.

- Lane markings can be exploited to provide information on refining SGM-generated DSM.

- The proposed reconstruction workflow relies on initial approximation with enough accuracy, i.e. a global terrain model like SRTM would not be sufficient as starting height, as the error could be several meters in particular at roads.

REFERENCES

Bay, H., Ferrari, V. and Gool, L. V., 2005. Wide-baseline stereo matching with line segments. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, pp. 329–336.

Fischer, P., Plaß, B., Kurz, F., Krauss, T. and Runge, H., 2017. Validation of hd maps for autonomous driving. *International Conference on Intelligent Transport Systems in Theory and Practice, mobil.TUM*.

Hofer, M., Wendel, A. and Bischof, H., 2013. Line-based 3d reconstruction of wiry objects. *Proceedings of the 18th Computer Vision Winter Workshop* pp. 78–85.

Jain, A., Kurz, C., Thormählen, T. and Seidel, H., 2010. Exploiting global connectivity constraints for reconstruction of 3d line segments from images. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* pp. 1586–1593.

Kurz, F., Türmer, S., Meynberg, O., Rosenbaum, D., Runge, H., Reinartz, P. and Leitloff, J., 2012. Low-cost systems for real-time mapping applications. *Photogrammetrie Fernerkundung Geoinformation* pp. 159–176.

Schmid, C. and Zisserman, A., 1997. Automatic line matching across views. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, pp. 666–671.

Steger, C., 1998. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(2), pp. 113–125.

Taylor, C. and Kriegman, D., 1995. Structure and motion from line segments in multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(11), pp. 1021–1032.

Wang, Z., Wu, F. and Hu, Z., 2009. Msld: A robust descriptor for line matching. *Pattern Recognition* 42(5), pp. 941–953.