AUTONOMOUS CALIBRATION OF MOVING LINE SCANNERS WITH CODED PHOTOGRAMMETRIC TARGETS RECOGNITION

Thomas Döring, Frederik Meysel and Ralf Reulke

German Aerospace Center (DLR) Rutherfordstr. 2 - 12489 Berlin (thomas.doering, frederik.meysel, ralf.reulke)@dlr.de

KEY WORDS: line scanner, calibration, position tracking, coded targets, convincedness, segmentation

ABSTRACT:

In this paper, we describe an automatic approach for a terrestrial line scanner calibration. The system is calibrated with prior knowledge of the exterior orientation of the camera in an unknown coordinate system. This data is acquired by tracking the moving sensor platform with an infrared camera tracking system (*ARTtrack2*) along a calibration scene. The calibration of a line scanner includes the determination of the transformation parameters to the known control point coordinate system (translation and rotation) and three parameters of the one-dimensional distortion model (simplified BROWN) for the sensor array. The focal length of the camera is assumed to be known. After a first rectification of the data with the tracking observations, photogrammetric targets are automatically detected and decoded. Finally, the distortion parameters of the camera and the transformation between world- and tracking coordinates are iteratively estimated.

1. INTRODUCTION

The recent years' growing activity in remote sensing applications for airborne and terrestrial applications allows for remote highaccuracy measurement of cultural heritage, architecture, buildings (Maresch, 1998), indoor and outdoor setups including traffic or crime scenes e.g. for 3D reconstruction purpose. Image data acquiring can be done with CCD-matrix and line sensors. Though line sensors provide a high resolution image, an additional sensor motion of the CCD-line is required to achieve the second image dimension. However, in order to use optical sensors for these purposes a calibration has to be done.

An overview about different calibration techniques can be found e.g. in (Grün, 2001). One approach is to determine the interior orientation of the camera with a given exterior orientation and known control points in a world coordinate system (Sehovic, 2001; Scheele, Krüger, and Schuster, 2005). Approaches that can do without *a priori* knowledge of absolute exterior orientation or control points are referred to as self-calibration techniques.



Figure 1. Automatic line scanner calibration

Thus, for a calibration knowledge about a defined calibration object such as a photogrammetric target field or geometric constraints of the scene is needed. This is done by measuring coordinates of features or special targets in both world- and imagecoordinates of the sensor. The same principle holds for linesensor cameras. In (Schneider and Maas, 2003; Parian and Grün, 2005) a line sensor calibration with a panoramic camera model is described. Below a calibration approach with an arbitrary but known motion trajectory with respect to an unknown coordinate system will be proposed.

Fig. 1 shows a overview of the used approach. The sections in this paper will be following that structure.



Figure 2. Line scanner mounted on platform

The first section describes the data acquisition and the resulting data basis. Next, the preprocessing for image data rectification is introduced, followed by section 3., dealing with target detection and identification of the targets' codes. Section 5. describes the parameter derivation and estimation of the global transformation between the used tracking system and world coordinates.

2. DATA ACQUISITION

For the Calibration three different types of data basis are used. Two of them are introduced in (Reulke, Wehr, and Griesbach, 2004). First, the *ARTtrack2* tracking system provides the positioning and orientation data of the used camera. This data is stored in ASCII data format (see figure 3).

Second, the data recorded by the line sensor camera consists of binary raw data in which three color channels of every line with



Figure 3. Tracking Data

10,298 pixels and 16 bits each are stored in band interleaved mode.

The third data basis is the control point data providing defined world coordinates.

The fact, that the camera position and orientation is known in the tracking coordinate system but unknown in the control point world coordinate system can be interpreted as a known relative orientation between the particular image lines but an unknown exterior orientation of the whole imaging set. With these known information the unknown parameters can be determined:

- coordinate and orientation offset between the two coordinate systems (PS and TS)
- · distortion parameters of the used camera
- optional: principle point of the used camera

To keep the number of unknown parameters low only the radial symmetric part of BROWN's distortion model (Brown, 1986) is accounted. Furthermore, the principle point coordinates are presumed to be zero.

After the imaging and tracking process the image coordinates of the control points are measured with an autonomous target recognition algorithm. Due to the motion disturbance of the camera platform (see Fig. 4 a) the image data has to be spatially rectified.

Preprocessing

Therefor the line image data is read in successively line by line. The recorded tracking data is used to project every pixel of every line on a defined reference plane which is parallel to the average image plane with no regard to any distortion using the collinearity equations as described in section 5.. The result is an *geometrically corrected image*. One constraint for the target recognition is that the control points are completely visible in the processed image. Thus, he resolution of the processed image has to be equivalent to the sampling distance of the camera.

As it can be seen in Fig. 4 b) the resulting picture is still distorted due to remaining error of the tracking data. However, the target recognition used in this approach can cope with distortion of moderate degree.

3. TARGET RECOGNITION

The recognition and identification of targets can be done in this preliminary rectified image. Two types of targets are used. Uncoded targets consist only of a white central disc on a black background. Additionally there exist different types and designs of



Figure 4. Camera data

coded targets. The used coded targets have a code on a concentric ring (van den Heuvel and Kroon, 1992; Schneider and Sinnreich, 1992) invariant to affine transformations, i.e. translation, scaling, skew, rotation, and are insensitive to image noise, speckle and even geometric distortions of moderate degree.

Fig. 5 shows the scheme of the recognition approach described below. After performing a global thresholding of the data, segments of a particular roundness are labeled and extracted as center disc candidates. Uncoded targets are verified easily by calculation of the centroid (surrounding black segment, no white neighbors inside and centroids very close), while coded targets have to be read out, the profile processed and the label of the target identified. Further readings are found in the references (Clarke, 1994).



Figure 5. Sequence of processing steps for target extraction

Segment based global thresholding

Thresholding has to be done for the segmentation of the image. A suitable thresholding value is essential for successful segmentation (Pavlidis, 1982).



Figure 6. Plots of a) histogram of input image and b) segment number results from thresholding

Unlike many histogram (Fig. 6 a) based algorithms to derive a global thresholding value the proposed approach uses additional knowledge about the scene, concerning reasonable number of segments found. It scales down the original image data, then performs a bundle of thresholding operations for different values. For each bimodal result, *the number of segments* in the image is counted and plotted with the thresholding values on the x-axis, as can be seen in Fig. 6 b).

Because not all results are suitable for segmentation a generous range was defined between 10 and 200 segments (all of the patterns used had a maximum of 150 segments for 50 targets) in a picture as a strong hint for a good thresholding result. The remaining task would be to chose a value from the *plateau* at about 90 segments in Fig. 6 b). Global or local minima turned out not to yield optimal thresholds. For the optimal value, we calculated the mean value of all threshold values that lead to good results. However, this approach may fail if the center of mass accidentally falls near to local maxima in the segment plot.

Region labelling

After binarization of the image the labelling of connected segments (also known as "blob coloring") is done. In our implementation, the standard algorithm with 4-connectedness is used. The resulting map has the same size as the original image, carrying the segment indices.

Coded Photogrammetric Targets

The targets used are 78 mm x 82 mm in size and have a a white center disc and a code ring around it on a black background, to provide the highest possible contrast in a scene. The radius of the center disc r is also the width of the code ring, which consists of 12 segments corresponding to a 12 bit code, running around the center between 2r and 3r from the centroid.



Figure 7. Coded photogrammetric target

Furthermore the targets provide a centroid mark in the middle of the center disc and a label for manual determination of the target number.

Uncoded targets are 38 mm x 41 mm in size with a wite center disc on a black background and a small label carrying an identification number for manual evaluation.

Subpixel accuracy determination of target center

A variety of techniques for the subpixel accuracy localization of target centers is known. A comparison of some techniques is given by (Shortis and Short, 1994). We used a greyscale centroid approach. In lower scale images, Gauss least square or ellipse fitting would be a better choice. The weighted algorithm is of

reasonable computational cost and highly precise because of the large target size in our close-range line scanner data. The centroid position of the segment is calculated by considering all pixels P_i belonging to the segment with

$$P_i = (x_i | y_i), \quad i = 1 \dots n \tag{1}$$

and n is the boolean area of the segment, i.e. the number of pixels in the segment. With the intensity of each pixel given as

$$I(P_i) = \text{intensity of } P_i \tag{2}$$

the centroid position $X_c = (x_c | y_c)$ with

$$x_{c} = \frac{\sum_{i=1}^{n} (x_{i} \cdot I(P_{i}))}{\sum_{i=1}^{n} I(P_{i})}$$
(3)

$$y_{c} = \frac{\sum_{i=1}^{n} (y_{i} \cdot I(P_{i}))}{\sum_{i=1}^{n} I(P_{i})}$$
(4)

is calculated as the sum of the x and y positions of all pixels, weighted and normalized with their pixel intensities and the overall intensity sum of the whole segment. The membership of pixels to the segment is decided in the thresholding step of the process.

Readout of the Identifier

Given the exact position of the target centroid P_c and the area A of the segment, a circular readout of the code ring is performed. From the segments' area A the radius results to be

$$r = \sqrt{\frac{A}{\pi}},\tag{5}$$

and the readout radius chosen for the circular scan is 2.5 times r since the codering runs between $2r \dots 3r$ around the center segment.



Figure 8. Codering readout a) before and b) after lock algorithm

Code extraction from brightness profile

When having extracted a brightness profile with the correct radius around the target center, spatial shifts and distortions, as well as blur and rotational mismatches occur. Several approaches can be used to extract a binary sequence from the brightness profile together with a number expressing the degree of convincedness a valid sequence has been found.

Algorithm 1 Profile shift analysis	
Initialization:	

set up variance sum array

for $\kappa = 0$ to 30° step 0.5° do - shift profile κ degrees for i = 1 to n do - calculate μ_{s_i} and $\sigma_{s_i}^2$ for all segments - save $\sum_n \sigma_{s_i}^2$ in array end for end for optimal shift value $\leftarrow \arg\min_{\kappa}(\sum_n \sigma_{s_i}^2)$ confidence value $\leftarrow C = \frac{\max(\sum_n \sigma_{s_i}^2)^2}{\min(\sum_n \sigma_{s_i}^2)^2}$

In this algorithm, the brightness profile is divided into n = 12 equidistant sections s_i , $i = 1 \dots n$, representing the 12 bits of data encoded into the target. For each section, the brightness information is collected and the variance $\sigma_{s_i}^2$ calculated. If the matching grid is not properly aligned with the code, high variance values will occur in some segments. In sparse codes, only few section values show activity. Summing up all variance values $\sum_{n} \sigma_{s_i}^2$ ensures that every derivation from the optimal lock will show up in the final value. Varying the rotational shift κ , changes the variance sum. The result will repeat itself for shifts above

the variance sum. The result will repeat itself for shifts above 30° (which is $\frac{360^{\circ}}{n}$). Therefor, the algorithm performs 60 steps at 0.5°, calculates all variance sums and chooses the shift value κ that yields the minimum value as the best shift.



Figure 9. Brightness profile of target center and corresponding locking mask

As a measure for the convincedness C of the algorithm, the final minimum value is brought into relation to the maximum variance sum found in the process:

$$C = \frac{\max(\sum_{n} \sigma_{s_i}^2)}{\min(\sum_{n} \sigma_{s_i}^2)^2}.$$
(6)

This value is used to determine if the target center and readout radius have been well chosen or if there is a target in the vicinity at all. The convincedness is not a sufficient condition, but a necessary one to confirm a successful readout and lock.

Code Matching

With a successful lock on the profile, the 12 segments are averaged and thresholded in the middle of the absolute minimum and maximum values of the profile. Having extracted a binary code string in that way, it has to be compared to the codes in the database. All coded targets in the scope of this work have codes with an invariant header suitable for detection of the beginning of the code.

target 3 label 4 bitcode 1000 xc 5443.324 yc 2009.482
target 4 label 14 bitcode 0110 xc 9002.387 yc 2012.384
target 5 label 8 bitcode 1100 xc 5439.016 yc 6103.905
target 6 label 0 bitcode 0000 xc 7138.328 yc 2010.384
target 7 label 26 bitcode 0011 xc 8999.184 yc 6111.176
target 8 label 0 bitcode 0000 xc 7253.837 yc 6099.394

Figure 10. Image coordinates of recognized targets data

The input code string is shifted until the code header leads the string. It is then compared to the codes in the database to extract the assigned target label. As the result of the target recognition, a list of targets, their labels and positions in image coordinates is produced, containing coded and uncoded photogrammetric targets.

4. STARTING VALUES OF THE ESTIMATION

The automatically determined pixel coordinates in the *geometrically corrected image* have to be transformed back to the image coordinate system of the camera. Each image pixel may correspond to multiple image coordinates of the camera. These relations are stored in the preprocessing. Thus, the image coordinates of the targets can easily be transformed back.

The estimation of the unknown parameters requires starting values. Preliminary derivations of these values occurs in two steps. The first step is to calculate the unknown Z-values for every measured control point r with more than one measurement with respect to the tracking coordinate system using spatial intersection (Kraus, 1996).

Algorithm 2 Derive exterior orientation starting values
for all measurements of every r_l do
compute X_{TS} and Y_{TS} using Equ. (7) and (8) with Z^r
end for
solve least-squares estimation for best fit parameters of the 6
parameter euclidian transformation $(PS) \rightarrow (TS)$ using all
calculated \mathbf{X}_{TS} and known \mathbf{X}_{PS}

In the second step approximations for the translation and rotation parameters of the unknown exterior orientation are calculated with an Gauss-Markov least-squares estimation.

The result of Alg. 2 are the starting values for t_X , t_Y and t_Z as well for ω , φ and κ . As for the distortion parameters the starting values are set to zero.

5. PARAMETER DERIVATION MODEL

The parameters which have to be derived should completely describe the transformation and projection respectively from the control points world coordinates to the image coordinates. The derivation of the particular parameters is based on the standard collinear equations.

$$X_{TS} = X_0 + (Z_{TS} - Z_0) \cdot \frac{Z_x}{N}$$
(7)

$$Y_{TS} = Y_0 + (Z_{TS} - Z_0) \cdot \frac{Z_y}{N}$$
(8)

with

$$\begin{bmatrix} Z_x \\ Z_y \\ N \end{bmatrix} = \mathbf{R}_{\omega,\varphi,\kappa}^{-1} \begin{bmatrix} x - x_0 \\ y - y_0 \\ -c \end{bmatrix}$$
(9)

where ω , φ and κ are the orientation angles of the camera with respect to the tracking coordinate system, X_{TS} , Y_{TS} and Z_{TS} the control point world coordinates in the tracking coordinate system, X_0 , Y_0 and Z_0 the center of projection coordinates (in the tracking coordinate system, too), x and y the rectified image coordinates and x_0 and y_0 the principle point.

The unknown transformation between the control points known in the control point coordinate system (PS) and the tracking coordinate system (TS) can be described with

$$\mathbf{X}_{TS} = \mathbf{Q}_{\alpha,\beta,\gamma} \cdot \mathbf{X}_{PS} + \mathbf{T}$$
(10)

The distortion model is given with

$$x = (x_d - x_0) \cdot f_d \tag{11}$$

$$y = (y_d - y_0) \cdot f_d \tag{12}$$

with

$$f_d = \left[1 + \left(r^2 K_1 + r^4 K_2 + r^6 K_3\right)\right]$$
(13)

and

$$r = \sqrt{(x_d - x_0)^2 + (y_d - y_0)^2}$$
(14)

where x_d and y_d are the distorted image coordinates and x_0 and y_0 the image principle point.

Thus, the projection model can be extended to

$$\begin{bmatrix} X_{PS} \\ Y_{PS} \\ Z_{PS} \end{bmatrix} = \mathbf{Q}_{\alpha,\beta,\gamma}^{-1} \begin{bmatrix} X_0 + (Z_{TS} - Z_0) \frac{Z_x}{N} - t_X \\ Y_0 + (Z_{TS} - Z_0) \frac{Z_y}{N} - t_Y \\ Z_{TS} - t_Z \end{bmatrix}$$
(15)

with

$$\begin{bmatrix} Z_x \\ Z_y \\ N \end{bmatrix} = \mathbf{R}_{\omega,\varphi,\kappa}^{-1} \begin{bmatrix} (x - x_0) \cdot f_d \\ (y - y_0) \cdot f_d \\ -c \end{bmatrix}$$
(16)

and Z_{TS} given by (10).

As it can be seen from (13) and (15) there are nine unknown parameters which have to be derived. These parameters are

$$\mathbf{p} = [t_X, t_Y, t_Z, \alpha, \beta, \gamma, K_1, K_2, K_3].$$

To determine these parameters the Gauss-Markov least-squares theorem

d

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{L}$$
(17)

will be used, where A is equal to the design matrix and L to the observations vector. In this article it is assumed, that the control point world coordinates are observations. That is because

it turned out that the use of the measured image coordinates as observation leads to an unstable geometry and thus to a highly multi-modal convergence set especially with respect to the Z coordinates. Consequently, the estimation model is not quite correct since Z_{TS} is calculated using these *observations*.

Since this model is based on linear equations (15) has to be linearized. Therefor the following Taylor series is proposed:

$$X_{PS} = F_X \left(\mathbf{p}^0 \right) + \frac{\partial X_{PS}}{\partial t_X} \cdot dt_X + \frac{\partial X_{PS}}{\partial t_Y} \cdot dt_Y + \frac{\partial X_{PS}}{\partial t_Z} \cdot dt_Z + \frac{\partial X_{PS}}{\partial \alpha} \cdot d\alpha + \frac{\partial X_{PS}}{\partial \beta} \cdot d\beta + \frac{\partial X_{PS}}{\partial \gamma} \cdot d\gamma + \frac{\partial X_{PS}}{\partial K_1} \cdot dK_1 + \frac{\partial X_{PS}}{\partial K_2} \cdot dK_2 + \frac{\partial X_{PS}}{\partial K_3} \cdot dK_3$$
(18)

and the same for Y_{PS} and Z_{PS} .

The design matrix \mathbf{A} is set up with

$$\mathbf{A} \in \mathbb{R}^{3m \times n}$$

with elements

$$a_{ij} = \frac{\partial o_i}{\partial p_j}, 1 \le i \le 3m, 1 \le j \le n$$
⁽¹⁹⁾

where m is the number of control points, n the number of unknown parameters – here n = 9 – and observation equations

$$\mathbf{o} \in \mathbb{R}^{3m}, \quad o_i = \begin{cases} X_{PS}|_k & \text{if } i = 3k - 2, \\ Y_{PS}|_k & \text{if } i = 3k - 1, \\ Z_{PS}|_k & \text{if } i = 3k. \end{cases}$$
(20)

The observation vector \mathbf{L} is given with

$$\mathbf{L} \in \mathbb{R}^{3m}, \quad L_{i} = \begin{cases} X_{PS}^{k} - F_{X}^{k} \left(\mathbf{p}^{0} \right) & \text{if } i = 3k - 2, \\ Y_{PS}^{k} - F_{Y}^{k} \left(\mathbf{p}^{0} \right) & \text{if } i = 3k - 1, \\ Z_{PS}^{k} - F_{Z}^{k} \left(\mathbf{p}^{0} \right) & \text{if } i = 3k. \end{cases}$$
(21)

where \mathbf{p}^0 is the approximation vector of the unknown parameters.

After the calculation of $d\hat{\mathbf{x}}$ the unknown vector \mathbf{p} will be updated with

$$\mathbf{p} = \mathbf{p}^0 + d\hat{\mathbf{x}} \tag{22}$$

and the approximation vector for the following iteration step will be $\mathbf{p}^0 = \mathbf{p}$.

6. RESULTS

To test the proposed algorithms a simulation program was implemented. This programm simulates the whole calibration process beginning with the generation of a defined grid of control points (using uncoded and coded targets). Furthermore the motion of the platform, the tracking of its position and orientation as well as the image recording is simulated.

To verify the quality of the calibration process the control point image measurements are transformed to the world coordinate system (PS) using the derived parameters and the known Z coordinate. After that, two values are calculated:

- (a) the standard deviation of the world coordinates with respect to its mean values and
- (b) the standard deviation with respect to the expectation values, the given control points.

The point of failure is the tracked position and orientation data accuracy. Thus, the position data (Fig. 11), the orientation data (Fig. 12) and both (Fig. 13) were randomly distorted with increasing standard deviations.



Figure 11. Calibration accuracy with declining position data accuracy



Figure 12. Calibration accuracy with declining orientation data accuracy



Figure 13. Calibration accuracy with declining position and orientation data accuracy

The achieved accuracy with *perfect* position and orientation data was about $\sigma_X = \sigma_Y = 0.2mm$. Fig. 11 shows the linear proportionality of the camera and the resulting point position data accuracy. This fact is obvious since (15) shows this linearity to the displacement vector, too. The visible peaks are due to failing iteration convergence. Fig. 12 is surprising at first. It seems that the results are less stable with a lower degree of orientation error than with a higher one. Further analysis of the derived parameters indicates that the oscillating accuracy is due to systematic geometric characteristics of the recorded scene. Angular errors leads to different but mathematically right results of the ambiguous solutions of the calibration problem. At a low level of this error the ambiguity applies more to the t_Z parameter which leads to two different intersection conditions (leading to two different result accuracy). The higher the orientation error the more the ambiguity applies to the t_Y parameter (and less to the t_Z) and accordingly to ω . This leads to a more random accuracy distribution.

The target recognition approach proved to be flexible for different lighting conditions, but not immune to local illumination dynamics. A badly illuminated calibration grid brings this approach to an abrupt end. The centroid location calculation worked accurate and quickly. The recognition and decoding of the targets' codes showed an impressing resistance to noise distortion and a moderate handling of the jittering. The image point accuracy depends on the provided image resolution and is in the range of a tenth of a pixel. In the upcoming work, we want to provide a comprehensive sensitivity analysis for variation of different experimental parameters (e.g. platform stability, tracking accuracy).

7. CONCLUSION AND OUTLOOK

We described a fully automatic self-calibration approach for a moving terrestrial line-scanner. The data is rectified with tracking observations. Control points for the parameter derivation are automatically recognized through photogrammetric targets decoding. All steps of this approach have been tested in a variety of changing conditions. An integrated setup and a comprehensive analysis of the implemented approach's performance and sensitivity to critical parameters is still to be done.

In further test the calibration will be evaluated with real data in a known environment to examine the convergence behavior and the stability of the iterations. Furthermore, the parameter set should be extended, to allow for a more accurate estimation of a better distortion model.

REFERENCES

Brown, J. F., 1986. Lense distortion for close-range photogrammetry. *Photogrammetric Engeneering and Remote Sensing* 52(1), 51–58.

Clarke, T., 1994. An analysis of the properties of targets used in digital close range photgrammetric measurement. *Videometrics*.

Grün, A., 2001. Calibration and orientation of cameras in computer vision. Springer.

Kraus, K., 1996. Photogrammetrie, 2. Band. Dümmler Verlag.

Maresch, M., 1998. *Linear CCD array based recording of buildings for digital models*. Ph. D. thesis, Technical University of Graz.

Parian, J. A. and Grün, A., 2005. Panoramic camera calibration using 3d straight lines. *Panoramic Photogrammetry Workshop, Proceedings of the ISPRS working group V/5* XXXVI-5/W8.

Pavlidis, T., 1982. Algorithms for graphics and image processing. Springer Verlag.

Reulke, R., Wehr, A., and Griesbach, D., 2004. High resolution mapping using ccd-line camera and laser scanner with integrated position and orientation system. *IAPRS* vol. XXXV, part B3.

Scheele, M., Krüger, S., and Schuster, R. a., 2005. New calibration scheme for panoramic line scanners. *Panoramic Photogrammetry Workshop, Proceedings of the ISPRS working group V/5* XXXVI-5/W8, 1–4.

Schneider, C.-T. and Sinnreich, K., 1992. Optical 3-d measurement systems for quality control in industry. *International Archives of Photogrammetry and Remote Sensing* 29(B5), 56–59.

Schneider, D. and Maas, H.-G., 2003. Geometric modelling and calibration of a high resolution panoramic camera. *Optical 3-D Measurement Techniques VI* Vol. II, 122–129. Institute of Geodesy and Photogrammetry, ETH Zürich.

Sehovic, A., 2001. Design und Kalibrierung einer CCD-Zeilenkamera für 3D Objektrekonstruktion. Dissertation.

Shortis, M.R., C. T. and Short, T., 1994. A comparison of some techniques for the subpixel location of discrete target images. *Videometrics*.

van den Heuvel, F. and Kroon, R., 1992. Digital close-range photogrammetry using artificial targets. *International Archieves of Photogrammetry and Remote Sensing* 29(B5), 222–229.