

# AUTOMATIC REGISTRATION OF PANORAMIC 2.5D SCANS AND COLOR IMAGES

T. Abmayr<sup>1</sup>, F. Härtl<sup>1</sup>, G. Hirzinger<sup>2</sup>, D. Burschka<sup>3</sup> and C. Fröhlich<sup>1</sup>

<sup>1</sup> Zoller+Fröhlich GmbH, Wangen i.A., Germany,

<sup>2</sup> German Aerospace Center, Oberpfaffenhofen, Germany

<sup>3</sup> Technical University of Munich, München, Germany  
t.abmayr@zofre.de

**KEY WORDS:** Terrestrial Laser Scanning, Color Scanning, Sensor Calibration

## ABSTRACT:

Here we present an approach for a registration framework to fuse the 2.5D data from a laserscanner and the color information from a digital camera to 2.5D color images. The spatial resolution of the scanner is very dense, which emphasizes the need for precision during the calibration of both sensors as well as in the registration of the data streams. To achieve the same visual field for the digital camera, it is mounted on a vertical tilt unit. Overlapping color images are acquired by rotating the entire system horizontally and the tilt unit vertically by predefined angle increments. Using these angle increments as fixed input parameters, the calibration of the camera as well as the registration between both sensors can be simplified enormously. We use the Sift method and Ransac to find homologous points between the scanner and the camera data automatically and marker-free. In our results the marker-free approach is compared against marker-based registration.

## 1 INTRODUCTION

In this work we combine a high resolution array camera with a panoramic laserscanner. To achieve the same visual field of view for the camera it is attached to a vertical tilt unit. Once the tilt unit and the camera are mounted on the scanner device and the relation between the sensors is calculated, then both sensors can be regarded as one single system. Such a system combines the advantages of both sensors: The range information of the scanner delivers a highly accurate 2.5D point cloud of the environment, whereas the color data of the camera gives big benefits in better visualization, higher resolution and interpretation of the data. Aligned color and range data is the basis for continuative modeling tasks, e.g. for 3D reconstruction with a triangular mesh or 3D object recognition. By aligning different viewpoints, range data can be much more easily and stable combined to dense 3D models than single 2D camera data only. Finally from a measurement point of view, the combination of both sensors ideally completes each other: Camera sensors tend to blur edges less than scanners.



Figure 1: The Zoller+Fröhlich laser scanner Imager 5006 and a digital array camera mounted on it.

To date, several laser scanning systems with integrated cameras are available from a number of manufacturers (see reference for an overview about the technologies available in (Przybilla, 2006)). However, most of these systems lack a panoramic option (Riegl, 2007, Zoller+Fröhlich, 2007, Faro, 2007) as most of the scanners have no panoramic field of view or there is no tilt device for the camera. Two solutions (Leica, 2007) (Optech, 2007) are offered with a camera integrated into the system itself. These cameras, however, have low resolution and are primarily used to capture previews

and to set the field of view of the scanner. In addition, array cameras with tilt unit are available (Callidus, 2007), (Trimble, 2007),



Figure 2: A colored 3D point cloud.

but we have no information about their mapping technique or the accuracy of the mapping result. Beside this, the used cameras also have only a low resolution, which does not seem adequate for high accurate registration and data fusion.

In addition to these integrated systems, cameras can be used as external device (Kst, 2007), (SpheronVR, 2007). The scanner and the camera are not used as a combined system, but as two independent sensors and have the big advantage that there is no parallax between the optical centers of both sensors. These cameras are high resolution line-scan cameras, and - at least in case of (SpheronVR, 2007) - with a panoramic field of view. Several studies were carried out (Reulke et al., 2003, Schneider and Maas, 2004, Scheibe et al., 2004) to investigate the benefits of such a combination. However, the registration for every single viewpoint has to be recalculated as these cameras and scanners are not integrated into a single system. Another disadvantage when using such systems is that CCD-lines will cause the acquisition time to increase enormously in dark environments. This is due to the fact that every single line has to be exposed and no flashlight can be used. This is often the case in indoor applications.

The methods presented in this paper mainly focus on the geomet-

rical calibration and registration of the sensors but also its fusion to a dense 3D point cloud (see Fig. 1 and 2). Besides the key aspect of accuracy, the work focuses on the automation of the methods and their robustness.

## 2 HARDWARE AND ACQUISITION PROCESS

We start with a short description of the mechanical principle of the scanner and the acquisition setup and introduce some important definitions and notations.

### 2.1 Mechanical set-up of the Scanner

The scanner we use consists of a range measurement system in combination with a mirror deflection device (Zoller+Fröhlich, 2007). The deflection system points the laser beam into the direction of measurement, the laser beam is emitted and the reflected laser light is detected. A 3D scan is acquired by two rotations: First the mirror rotates around a horizontal axis ("elevation rotation axis") and thus deflects the laser beam in a vertical direction. The second rotation is around the vertical center axis ("azimuth rotation axis") of the system (see Fig. 1 and 3). The actual direction of the laser beam is measured by two encoders: The first describes the actual horizontal rotation and is adjusted at the center axis (azimuth encoder). The second encoder describes the mirror rotation and is adjusted along the mirror rotation axis. The zero position of this second encoder is located along the negative direction of the center axis (elevation encoder). These two encoders define a spherical coordinate system, which is called encoder coordinate system. The field of view of the scanner is  $360^\circ$  (azimuth) and  $320^\circ$  (elevation). In order to acquiring a panorama scan we rotate  $320^\circ$  in elevation and only rotate  $180^\circ$  in azimuth. In addition to range information the device also measures the reflectivity of the object-surface giving a photo like impression of the scanned area. There is a one to one correspondence between the reflectance and the range value of each sample with respect to the corresponding azimuth and elevation angles.

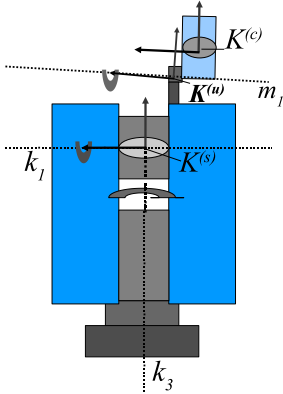


Figure 3: Coordinate systems and rotation axis

ing a photo like impression of the scanned area. There is a one to one correspondence between the reflectance and the range value of each sample with respect to the corresponding azimuth and elevation angles.

### 2.2 Camera-Setup and Acquisition Process

In our application a high resolution array camera is used. To achieve the same visual field of view as with the scanner, a vertical tilt unit is mounted on the scanning device with the camera attached to it. In order to acquire high resolution, well exposed pictures, data is not collected 'on the flight' in parallel with the scan, but consecutively after the scan. Overlapping color images are taken by rotating the system by predefined angle increments

around the azimuth rotation axis of the scanner and vertically around the tilt unit.

In our measurement concept, the scanner is assumed to be stable for a longer period, and can therefore be calibrated in the laboratory. The tilt unit and the camera can be removed from the scanner facilitating transportation, but have to be calibrated after reassembly with the scanner. This was taken into account for the calibration and registration step.

In order to describe the mapping between the scanner and the camera, a number of different coordinate systems are essential and are introduced in the next section.

## 3 COORDINATE SYSTEMS

An affine, orthogonal and right-handed coordinate system with the translation  $a$  and the basis vectors  $e_i$  is denoted throughout this paper by  $K := (a, e_1, e_2, e_3)$ . Based on this notation the following coordinate systems can be introduced (see Fig. 3): First, the **scanner coordinate system** (our reference coordinate system). Let  $k_1 \in \mathbb{R}^3$ ,  $\|k_1\| = 1$  be the elevation axis, and respectively take  $k_3 \in \mathbb{R}^3$ ,  $\|k_3\| = 1$  for the azimuth rotation axis. If we assume for now, that both axis intersect and  $\langle k_1, k_3 \rangle = 0$ , we can construct an affine, orthogonal and right-handed coordinate system

$$K^{(s)} := (0, k_1, k_2, k_3). \quad (1)$$

A second coordinate system

$$K^{(c)} := (w_0, w_1, w_2, w_3) \quad (2)$$

which describes the position of the camera with respect to the scanner is called **camera coordinate system**. If the camera can be described as an ideal pinhole camera, then the coordinate system holds the properties:

- The origin  $w_0$  is equal to the optical center of the camera.
- The third vector  $w_3$  is orthogonal to the image plane.
- $w_1$  is parallel to the horizontal border of the image plane, and  $w_2$  respectively to its vertical border.

Based on this we introduce a third coordinate system

$$K^{(u)} := (m_0, m_1, m_2, m_3). \quad (3)$$

which we call **camera-tilt-unit coordinate system**. This coordinate system has the following properties:

- The origin is equal to the center of the rotation axis of the tilt-unit.
- The direct of the first vector is along the direction of the rotation axis.

- The direction of  $m_3$  approximately corresponds to the direction of the of the z-axis  $k_3$  of the scanner coordinate system  $K^{(s)}$ .

We assumed several idealized properties for the geometry of the scanner as well as for the camera, e.g. orthogonal rotation axis. However, the real sensors deviate from these idealized models. Below we demonstrate how these models can be transformed to apply to real sensors and how the relation between the coordinate systems can be determined.

#### 4 MODELING THE SCANNER

Starting with the scanner, we first notice the similarity to theodolites for the mechanical set-up. Thus a simple but accurate enough model known in photogrammetry was used for the calibration (Abmayr et al., 2005), (Lichti and Franke, 2005), (Deumlich and Staiger, 2002).

Let  $l$  be the azimuth encoder angle,  $h$  be the elevation encoder angle and  $rg$  the range. Then the overall transformation  $\Phi(l, h, rg) : [0, \pi]^2 \times \mathbb{R} \rightarrow \mathbb{R}^3$  from un-calibrated encoder coordinate system into the scanner coordinate system  $K^{(s)}$  is defined through

$$\Phi(l, h, rg) = H \circ \Upsilon(l, h, rg) \quad (4)$$

with

$$\Upsilon(l, h, rg) = \left( l + \text{sign}(h - \pi) \left( \frac{b}{\sin(h)} + \frac{a}{\tan(h)} \right), h + c, rg \right) + \eta(l, h, rg). \quad (5)$$

$H$  describes the transformation from spherical - to Cartesian coordinates and is well known.  $\Upsilon_1$  corrects the non-orthogonality between the elevation rotation axis and the azimuth rotation axis as well as the non orthogonality between the laser beam and the elevation rotation axis. In photogrammetry these errors are called **trunnion axis error** and **collimation axis error** respectively. As described in section 2, the zero position of the elevation encoder must be equal to the negative horizontal rotation axis  $k_1$ . As this is usually not the case in real sensors due to mechanical inaccuracies, it has to be corrected by a constant vertical angle offset, which is described through  $\Upsilon_2$ . This error is called **vertical circle index error**. The unknowns  $a, b, c \in \mathbb{R}$  can be determined as described in (Abmayr et al., 2005).

$\eta$  is a term, which summarizes all additional calibration errors. As the focus is onto the calibration and registration of the camera we cannot go more into detail throughout this paper and refer to the literature (Lichti, 2007), (Rietdorf, 2005).

#### 5 MODELING THE CAMERA

The camera is mounted on a vertical tilt unit and fixed on the scanner. Mapping of different camera positions in relation to the scanner is therefore function of the horizontal rotation angle  $\alpha$  of the scanner and the vertical rotation angle  $\beta$  of the tilt unit. In the following paragraph we describe the relationship between camera and scanner.

#### 5.1 Projection

As the rotations of the scanner and the tilt-unit are highly accurate, we use these rotation angles as fixed input parameters for our model. For modeling the intrinsic parameters of the camera we use Tsai's camera model (Tsai, 1987), which is based on the pinhole camera of perspective projection, and is well known in Computer Vision. If we denote the horizontal rotation with the angle  $\alpha$  around the z-axis of the scanner with  $Z_\alpha$  and the vertical rotation with the angle  $\beta$  around the x-axis of the tilt-unit with  $X_\beta$  then the overall projection  $\Xi_{\alpha, \beta} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  from a point  $X := (x, y, z)$  of the scanner coordinate system  $K^{(s)}$  onto the pixel  $(u, v)$  in the color image can be written as

$$\Xi_{\alpha, \beta}(X) = \varphi_{\kappa, c_x, c_y, s} \circ \pi_f \circ T_{\alpha, \beta} \quad (6)$$

with

$$T_{\alpha, \beta} := M X_\beta \tilde{M} Z_\alpha. \quad (7)$$

$T_{\alpha, \beta}$  defines first the transformation from the *scanner coordinate system*  $K^{(s)}$  into the *camera-tilt-unit coordinate system*  $K^{(u)}$  and then the transformation from  $K^{(u)}$  into the *camera coordinate system*  $K^{(c)}$ . The perspective projection onto the image plane with the *focal length*  $f$  is described through  $\pi_f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  and the mapping from undistorted coordinates to distorted image coordinates with the *principle point*  $(c_x, c_y)$ , the parameter  $\kappa$  describing the *1st order radial lens distortion* and the *uncertainty scale factor*  $s$  is defined through  $\varphi_{\kappa, c_x, c_y, s} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . For detailed information see (Tsai, 1987).

According to the setup, the rotation angles  $\alpha$  and  $\beta$  define the actual position of the azimuth encoder of the scanner and the rotation angle of the tilt unit and are assumed to be known. By considering that  $M$  and  $\tilde{M}$  are homogenous matrices and thus can be described by 6 parameters each, we get 17 unknowns.

#### 5.2 Calibration and Registration

We now assume that we have taken  $I := \{(0, 0), \dots, (m, n)\}$  pictures, each index  $(i, j)$  corresponding to a horizontal angle  $\alpha$  and vertical angle  $\beta$  respectively. We further assume that there are  $k_{ij}$  corresponding points between scan and image  $(i, j)$ , which we denote with  $Q_{i, j, p}$  and  $q_{i, j, p}$ . As mentioned in section 2.1 each scanpoint  $Q_{i, j, p}$  is linked to a corresponding horizontal and vertical encoder value  $l_{i, j, p}$  and  $h_{i, j, p}$  respectively and a range value  $rg_{i, j, p}$ . It is important to note that points in the scanner can have correspondances in several images. We calculate at least two transformations to different camera positions by minimizing a least-squares error between corresponding points and then recover the transformations  $M$  and  $\tilde{M}$ . The method can be divided into three distinct steps:

- i.) As will be shown below *all* camera positions with the same vertical angle  $\beta$  can easily be transformed into the same camera position at a fixed angle  $\alpha_0$ , as they rotate on a circular path around the vertical rotation axis of the scanner. For more than six corresponding points in these images, Tsai's method can be used to get a transformation  $T_{\alpha_0, \beta}$  for a fixed camera position along this line. Tsai's method also solves start values for the intrinsic parameters.

ii.) Having recovered the intrinsics and at least four transformations  $T_{\alpha_i, \beta_j}$  for two different vertical angles  $\beta$ , then (7) can be transformed into an optimization problem which is called  $AX=XB$  problem (Park and Martin, 1994). Solving this problem gives a closed form solution for  $M$  and  $\tilde{M}$ .

iii.) The final step is a nonlinear optimization procedure for the fine adjustment of all 17 parameters.

We start with *i.*):

*To i.)* First the focus is on images taken at camera positions with a fixed angle  $\beta_0$ . It can easily be seen that for any point  $X_i \in \mathbb{R}^3$  (7) holds

$$T_{\alpha, \beta_0} X_i = T_{\alpha_0, \beta_0} (Z_{\alpha - \alpha_0} X_i). \quad (8)$$

Equ. (8) means geometrically, that if the vertical angle is fixed, then the camera rotates in a circle around the vertical rotation axis of the scanner. Furthermore this property implies that *all* camera positions along this line can be transformed into the same camera position at a fixed angle  $\alpha_0$ . So if there exist all together more than 6 corresponding points in these images, the transformation  $T_{\alpha_0, \beta_0}$  between scanner and camera at position  $(\alpha_0, \beta_0)$  can be calculated by using the 7 point algorithm used e.g. in Tsai's calibration method (Tsai, 1987).

*To ii.)* At least four transformations  $\{T_{\alpha_0, \beta_0}, T_{\alpha_1, \beta_0}, T_{\alpha_0, \beta_1}, T_{\alpha_1, \beta_1}\}$  are needed to solve the  $AX = XB$  problem: Assume that there exist two transformations  $T_{\alpha_0, \beta_0}$  and  $T_{\alpha_0, \beta_1}$ , with  $\beta_1 \neq \beta_0$  as described in *i.*). Then by using Equ. (8) two more transformations can easily be constructed through  $T_{\alpha_1, \beta_0} := T_{\alpha_0, \beta_0} Z_{\alpha_1 - \alpha_0}$  and  $T_{\alpha_1, \beta_1} := T_{\alpha_0, \beta_1} Z_{\alpha_1 - \alpha_0}$  for any angle  $\alpha_1$ .

Next it will be shown, how this can be transformed into a  $AX=XB$  problem: By transforming (7) it holds

$$T_{\alpha, \beta} := MX_{\beta} \tilde{M} Z_{\alpha} = MX_{\beta} \tilde{M} Z_{\alpha} \tilde{M}^{-1} \tilde{M}. \quad (9)$$

As described in section 3 the orientation of the z-axis of the scanner-coordinate system  $K^{(s)}$  and the tilt coordinate system  $K^{(u)}$  are constructed to be approximately the same and thus (9) can approximately be simplified to

$$T_{\alpha, \beta} = MX_{\beta} Z_{\alpha} \tilde{M}. \quad (10)$$

Given the two transformations  $T_{\alpha_i, \beta_i} := MX_{\beta_i} \tilde{M} Z_{\alpha_i}$  and  $T_{\alpha_j, \beta_j} := MX_{\beta_j} \tilde{M} Z_{\alpha_j}$  with  $(\beta_i \neq \beta_j)$  and applying (10) then both transformations can be solved for  $\tilde{M}$  and it holds

$$(X_{\alpha_j} Z_{\beta_j})(X_{\beta_i} Z_{\alpha_i})^{-1} M^{-1} = M^{-1} T_{\alpha_j, \beta_j} T_{\alpha_i, \beta_i}^{-1}. \quad (11)$$

With  $A_{i,j} := (X_{\alpha_j} Z_{\beta_j})(X_{\beta_i} Z_{\alpha_i})^{-1}$ ,  $B_{i,j} := T_{\alpha_j, \beta_j} T_{\alpha_i, \beta_i}^{-1}$  and  $X := M^{-1}$ , (11) can now be written as

$$A_{i,j} X = X B_{i,j}. \quad (12)$$

We refer to the literature for detailed explanation of how to solve  $X$  from (12), and which is now straight forward. This is a well known task in robotic manipulation and is e.g. shown by (Park

and Martin, 1994) in a very tidy way by using Li-Algebra. Finally  $\tilde{M}$  is derived from (10) through

$$\tilde{M} = (MX_{\beta_i} Z_{\alpha_i})^{-1} T_{\alpha_i, \beta_i}. \quad (13)$$

*To iii.)* Using the intrinsics from *i.*) and the extrinsics from *ii.*) with the equations from (4) and (6) a nonlinear least-squares error model can be built up through

$$\sum_{(i,j) \in I} \sum_{p \in \{0..k_{ij}\}} (\Xi_{\alpha_i, \beta_j} \circ \Phi(h_{i,j,p}, l_{i,j,p}, r_{g_{i,j,p}}) - q_{i,j,p})^2 \rightarrow \min \quad (14)$$

which can be solved by a standard approach (Stoer and Bulirsch, 1983).

## 6 HOMOLOGOUS POINTS

To optimize the error model introduced in section 5 homologous points between the scanner and the camera are needed. To date, this is usually done by using markers, as they can be detected with highest accuracy. The disadvantage is, however, that they have to be adjusted, and that they usually are not well distributed with respect to the system.

In this approach we study the Sift method to detect homologous points automatically and marker-free: In Computer Vision, Sift (Scale Invariant Feature Transform) (Lowe, 2004) is a very popular, feature based approach to find homologous points between overlapping images. In surveying, e.g. (Böhm and Becker, 2007) recently used the Sift method for the registration of 2.5D scan data. The Sift operator is scale-invariant, invariant to rotations and relatively stable to perspective distortions. The scale invariance is achieved by computing an image pyramid. Corners and edges are detected by calculating the difference of Gaussian (DoG) in each level. Candidates for homologous points are local maxima in these DoGs. The feature vector itself is constructed by a rotation invariant representation around the candidate point. Comparing the Euclidean distance of the feature vectors between the images results in a best match suggestion of homologous points. We refer to the literature for more information.

Using Sift usually results - depending on the parameters to adjust - in a high amount of suggested matches. On the other hand, it is obvious that these suggested matches contain a high percentage of false candidates. To detect these outliers we use the Ransac (Random Sample Consensus) method (Fischler and Bolles, 1981): Randomly, the minimal set of points needed to calculate the transformation between the images is extracted and all candidates are tested for consensus. This is repeatedly done until a break criteria is reached. The set with the largest consensus is selected for registration. Again we refer to the literature for more information.

In our approach we use instead of only camera images also the reflectance image of the scan as input data. This hardens the task, as scanners and cameras are quite different in their measurement characteristic (see Fig. 8 (a)). However, as the horizontal and vertical angle increments for each camera position is known, the corresponding regions between reflectance image and color images

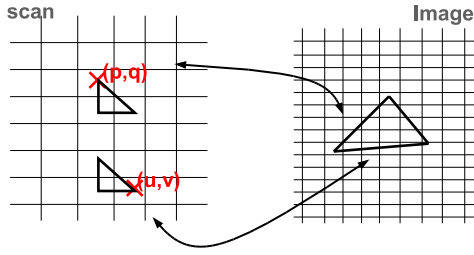


Figure 4: Two triangle-neighborhoods in the scan map to the same region in the image. To the neighborhood around the pixel  $(p, q)$  no color must be assigned (compare Fig. 5).

can be estimated. This is very helpful, as it restricts the search region enormously, and thus stabilizes the search for consensus significantly.

## 7 DATA FUSION

Now all sensors are assumed to be calibrated and the relation among them is known. Since this paper mainly focuses on the registration of the data streams we only give the basic ideas of the fusion process. The basic procedure is rather straightforward, as the mapping from a scanpixel  $(u, v)$  to image coordinates  $(m, n)$  is known. As introduced in section 2.1 each scanpixel  $(u, v)$  corresponds to encoder increments  $(h, l)$  and a range value  $rg$ . With (4) and (6) the mapping is defined through

$$(m, n) = \Xi_{\alpha, \beta} \circ \Phi(l, h, rg).$$

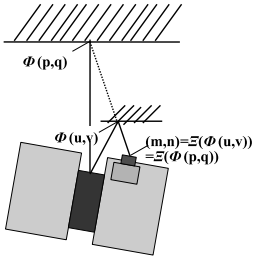


Figure 5: Ambiguity of the mapping because of occlusions due to parallax between both sensors.

The only problem which occurs however, is to take care of the occlusions due to parallax and the scene structure: mapping between scan and images in general is not one-to-one, and so mismatches occur, if objects are in the field of view of the scanner but not of the camera (see Fig. 5). To avoid those mismatches each pixel neighborhood in the scan is linked to its corresponding region in the images like indicated in Fig. 4. So if two neighborhoods in the scan map to the same region in the image, only this neighborhood can be assigned to the corresponding color values whose 3D values are closer to the camera viewpoint. This problem drafted is a classical z-buffer problem and is well known in computer graphics (Foley et al., 1996). The second point we would like to mention is the color adjustment between overlapping images: To get a homogenous color crossover in overlapping image regions, smoothing techniques for color adjustment must be performed. A simple blending technique we use is e.g. described in (Abmayr et al., 2005).

	Mean error in 10m
Targets vs. Groundtruth (no color)	1.1 mm
Targets vs. Groundtruth (with color)	3.3 mm

Figure 6: Mean error between markers extracted in the reflectance image and the ground truth in 10m distance and between the resulting color image and the ground truth respectively. In the mode used the scanner has a spatial point distance of 6.4 mm in 10m.

Marker-free 227 after Ransac (Total: 561)	Mean Error on Image Plane	Marker (Total: 23)	Mean Error on ImagePlane
Crosscheck vs „Marker Calibration“	2.20 pixel	Result „Marker Calibration“	1.1 pixel
Result „Sift/Ransac Calibration“	2.09 pixel	Crosscheck vs „Sift/Ransac Calibration“	1.25 pixel

Figure 7: Mean error between the homologous points projected on the image plane. The registration was calculated marker-free by using Sift and Ransac and marker-based.

## 8 EXPERIMENTAL SETUP AND RESULTS

As shown in a number of independent tests, the scanner proved to be stable for a long period of time, and could therefore be calibrated in the laboratory. The tilt unit and the camera were calibrated and aligned to the scanner each time after they had been assembled onto it. The calibration and registration process of the camera were carried out as explained in section 5. After these steps both systems could be regarded as one single system.

In the first experiment we use markers to prove the power and accuracy of the sensor models: We assembled a test scenario in a room with about 150 markers which were in varying positions in relationship to the system. The 3D positions of these markers were known highly accurate and were used as ground truth. The room was scanned and the images captured as explained in section 2. In the mode used the scanner exhibits a spatial point distance of 6.4 mm in 10m which yields a total of 50 million points for a full panoramic scan. In the first test, the markers were extracted and aligned versus the ground truth by the reflectance value of the scanner without fusing the scan and the color data. The result proved the high accuracy of the scanner itself (see Fig. 6). In the second test, the scan data was fused with the color data of the camera. Then the markers were extracted and aligned versus the ground truth. The result proved the accuracy of the joint system (see Fig. 6).

In the second experiment we investigated the automated, marker-free approach. In this test scenario we acquire data in the room shown in Fig. 2. To get homologous points between scan and camera we used the Sift and Ransac Method as described in section 6 (see Fig. 8(a)). Then we compared the registration obtained with the marker-free approach with the marker-based registration. For that purpose we crosschecked the registration of the marker based approach with the consensus set of the marker free approach and vice versa. The mean error of both registrations



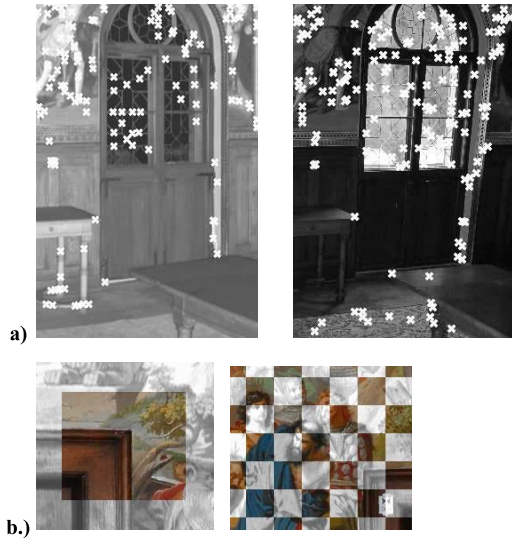


Figure 8: (a) Sift candidates in the reflectance image and one of the camera images. (b) Alternating presentation of the reflectance image and the resulting color image after the sensor fusion.

was quite similar as shown in Fig. 7. A close up after the data fusion is shown in Fig. 8(b): It exhibits an alternating presentation of the reflectance image and the transformed color images.

## 9 DISCUSSION AND OUTLOOK

This study shows how 2.5D data of a laserscanner can be aligned and fused with color images to yield highly accurate 2.5D panoramic color scans. Experiments showed that the calibration of the scanner was stable for a long period of time, and could therefore be calibrated off-line in the laboratory. Due to transportation conveniences the tilt unit and the camera however are designed to be removable and had to be recalibrated on the job. This was taken into account by the described registration and calibration methods. In this study we demonstrated that the alignment of the data sets could be automated and performed marker-free by searching homologous points with Sift and Ransac. To investigate the robustness and generality of the approach, our study has to be extended to a variety of different environments. To date, we only use homologous points between scan and camera images. However, the extra use of homologous points between camera images against each other additionally should improve the quality and robustness of the result.

## ACKNOWLEDGEMENTS

We thank M. Mettenleiter and the r&d team of Zoller + Fröhlich GmbH for their research work on the Imager 5006. We also thank A. Hildebrand and B. Neumann for hardware integration of the camera system and the tilt-unit into the Imager 5006. We also would like to thank H. Hirschmüller, M. Suppa and B. Strackebroch from the Institute of Robotics and Mechatronics at the German Aerospace Center and all other colleagues involved in

this project. This work was partly funded by the DFG project "Canopy Analyser".

## REFERENCES

- Abmayr, T., Dalton, G., Härtl, F., Hines, D., Liu, R., Hirzinger, G. and Fröhlich, C., 2005. Standardisation and visualization of 2.5d scanning data and rgb color information by inverse mapping. 7th Conference on Optical 3D Measurement Techniques, Vienna, Austria.
- Böhm, J. and Becker, S., 2007. Automatic marker-free registration of terrestrial laser scans using reflectance features. 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland I, pp. 338–344.
- Callidus, 2007. Laserscanner. <http://www.callidus.de>.
- Deumlich, R. and Staiger, R., 2002. Instrumentenkunde der vermessungstechnik. Herbert Wichmann Verlag; 9. Auflage, Hthig GmbH + Co. KG, Heidelberg.
- Faro, 2007. Laserscanner. <http://www.faro.com/>.
- Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM* 24 pp. pp. 381–395.
- Foley, J., van Dam, A., Feiner, S., Hughes, J. and Phillips, R., 1996. Introduction to computer graphics. Addison-Wesley.
- Kst, 2007. <http://www.kst-dresden.de/>.
- Leica, 2007. <http://www.leica-geosystems.com/hds/>.
- Lichti, D., 2007. Error modeling, calibration and analysis of an amcw terrestrial laser scanner system. *ISPRS Journal of Photogrammetry and Remote Sensing* 61(5), pp. 307–324.
- Lichti, D. and Franke, J., 2005. Self-calibration of the iqsun 880 laser scanner. 7th Conference on Optical 3D Measurement Techniques, Vienna, Austria.
- Lowe, D., 2004. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision* 60(2), pp. 91–110.
- Optech, 2007. <http://www.optech.ca/i3dtechoverview-ilris.htm>.
- Park, F. C. and Martin, B. J., 1994. Robot sensor calibration: Solving  $ax=xb$  on the euclidean group. *IEEE Transactions on Robotics and Automation* 10(5), pp. 717–721.
- Przybylla, H. J., 2006. Integration, fusion und kombination von terrestrischen laserscannerdaten und digitalen bildern. Workshop Anforderungen an geometrische Fusionsverfahren, DIN Deutsches Institut für Normung e.V. und Humboldt-Universität zu Berlin.
- Reulke, R., Wehr, A., Klette, R., Scheele, M. and Scheibe, K., 2003. Panoramic mapping using ccd-line camera and laser scanner with integrated position and orientation system. *Image and Vision Computing New Zealand*, editor D.G. Bailey pp. 72–77.
- Riegl, 2007. <http://www.riegl.com>.
- Rietdorf, A., 2005. Automatisierte auswertung und kalibrierung von scannenden messsystemen mit tachymetrischem prinzip, dissertation. Verlag der Bayerischen Akademie der Wissenschaften in Kommission beim Verlag C.H. Beck.
- Scheibe, K., Scheele, M. and Klette, R., 2004. Data fusion and visualization of panoramic images and scans. *Proceedings of the ISPRS working group V/1, Panoramic Photogrammetry Workshop, Part 5/W16, Dresden*.
- Schneider, D. and Maas, G.-H., 2004. Development and application of an extended geometrical model for high resolution panoramic cameras. *XXth ISPRS Congress, Istanbul, Turkey, Geo-Imagery Bridging Continents* pp. 366–371.
- SpheronVR, 2007. <http://www.spheron.com/>.
- Stoer, J. and Bulirsch, R., 1983. Introduction to numerical analysis. Springer Verlag.
- Trimble, 2007. <http://www.trimble.com/trimblegxs.asp>.
- Tsai, R., 1987. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-self tv cameras and lenses. *IEEE Journal of Robotics and Automation* 3(4), pp. 323–344.
- Zoller+Fröhlich, 2007. <http://www.zf-laser.com>.