Photometric Laser Scanner to Camera Calibration for Low Resolution Sensors

Johannes Gräter[1], Tobias Strauss, Martin Lauer *

August 1, 2017

Abstract

In modern automated systems the usage of sensors with orthogonal measurement principles is indispensable for assuring a redundant system. The combination of laser scanners and cameras is particularly prevalent. To profit from the potential of fusing information from these sensors, the basic step is sensor registration. State of the art cross-calibration methods for laser scanners and cameras establish geometric correspondences for example by detecting checker-boards in size and form. However these approaches suffer from ambivalent features and are only usable if the laser scanner has a high scanning density. Therefore, we present a reconstruction-free and vision-based extrinsic calibration algorithm with distinct features, which was developed with special focus on low-density laser scanners and high measurement noise. Moreover, we contribute an evaluation methodology demonstrating the capability of our approach. Along with this paper, we publish the associated code[1].

1 Introduction and state of the art

As automated driving turns from a field of research to applied science, safety and reliability of automated systems come into focus. Hereby, sensors with orthogonal measurement principles are particularly interesting because their fusion leads to considerably better results in reliability, quality and cost.

- **Reliability**: in case of failure of one sensor, another can bridge the missing data.
- **Quality**: by fusing data from different sensors the scenery can be described more accurately.
- **Cost**: whereas a laser scanner that can model the whole scenery is expensive, the fusion of a monocular camera and a low resolution laser scanner can lead to similar results.

In order to fuse information from a ranging sensor and a camera at an early stage, sensor registration is indispensable. Once the intrinsic and extrinsic calibration of all sensors are known, we can locate the laser spots in the camera image and associate the depth information with camera pixels, as shown in Fig. 1 and Fig. 2. Consequently,

*[1]The authors are with the Department of Measurement and Control, Karlsruhe Institute of Technology, contact `johannes.graeter@kit.edu`

[1]Source code available on GitHub: *https://github.com/KIT-MRT/PLCC*

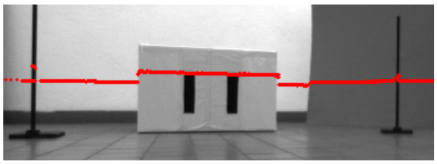


Figure 1: Projection of laser scanner measurements (red) from a dense scan into the camera image after calibration. Vertical shifts of the projections are caused by changes in depth of the observed points. We can observe their exact correspondence of edges of objects placed in the foreground.

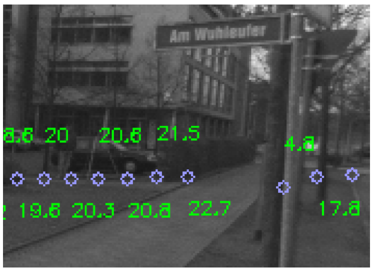


Figure 2: Projection of laser scanner measurements (blue) from a sparse scan into the camera image after calibration. Depth values are shown in green.

further processing of the data such as combined laser scanner and camera odometry from Zhang et al. [?] can be affected.

The calibration of high quality laser scanners to cameras has been studied extensively ([?], [?], [?], [?], [?], [?]). A widely used laser scanner is the Velodyne HDL-64E, which provides data from 360° and 64 scanning planes with a horizontal resolution of 0.2° and depth noise around 1.5 *cm* standard deviation. With this sensor a detailed three dimensional point cloud of the scene can be produced.

Most calibration algorithms associated with this sensor reconstruct 3d points from the camera images and register them to the corresponding point cloud from the laser scanner ([?], [?]). The reconstruction of the structure seen by the monocular camera is possible using checker-boards. Points of interest are found by a corner detector and reconstructed with the knowledge of board and tile size. In a second step, the resulting point cloud is registered to the point cloud obtained by the laser scanner. Popular registration methods use point-to-point correspondences to apply the well-known ICP algorithm [?] or align their normals([?], [?]). Zhang et al. [?] established a similar method by first reconstructing a calibration plane from the camera by the aid of checker-boards and minimizing the point to plane distance from the laser scanner data in order to estimate the extrinsic calibration of a high-resolution laser scanner possessing only a single scanning plane.
To negate the necessity of calibration boards, Scaramuzza et al. [?] developed a method that calibrates a rotating single line laser scanner and a camera automatically. In order to recover the three dimensional structure of the environment, they use a structure-from-motion approach and compare the resulting points with the laser scanner point cloud.
However, the reconstruction of three dimensional information from cameras is always

prone to errors. This problem becomes particularly prominent if we possess only few measurements and cannot counterbalance a lack of quality with quantity as is the case with low resolution sensors. Therefore, reconstruction-free calibration methods aim to recover the registration of the sensors without triangulating the viewing rays of the camera. Li et al.([**?**], applied in [**?**]) refrain from registering a scene reconstruction but detect edges in the camera image and a single-plane ranging sensor of a custom-made calibration object. By minimizing the back-projection error, they obtain the extrinsic calibration of the sensors.

However, all of these methods rely on a high-quality model of the surroundings, which requires an expensive high-quality laser scanner with both high ranging accuracy and high horizontal resolution. Particularly the calculation of normals in the point cloud ([**?**], [**?**], [**?**]) requires a dense and accurate point cloud, which low-cost laser scanners cannot supply. Also edges as used by Li et al. [**?**] can only be detected accurately for laser scanners with high angular resolution. The result of these methods applied on low resolution laser scanners will be poor since the edges cannot be located accurately.

For low-cost laser scanners, which are particularly of interest for mass-production, the established methods are not applicable. Since the specifications of mass-production laser scanners (e.g. [**?**]) are more than one order of magnitude lower than high-resolution laser scanners, state-of-the-art methodologies are bound to fail.

Therefore, we present a reconstruction-free approach to register a low-cost laser scanner and a monocular camera. Taking advantage of the fact that most imaging sensors are sensitive to the infra-red laser-light emitted by the laser scanner, the sensors are registered by minimizing an error metric applicable for a variety of cameras including cameras with a very large field of view or even without a single view point. Moreover, we establish a novel feature which is based on the visually observable laser spots in the image, enabling the registration of laser scanners with less than 50 points per scan and a depth noise with up to 0.5 m standard deviation.

Moreover, we propose a methodology to quantify the error of any laser scanner to camera calibration, using edge correspondences on an easily reproducible object.

2 Methodology

2.1 Problem formulation

In the following we assume that the camera and the laser scanner are both calibrated internally, that means there exist two known mappings:

$$\pi_{laser,i}(d_i) \rightarrow \mathbf{x}_i \quad , d_i \in \mathbb{R},\ \mathbf{x}_i \in \mathbb{R}^3,\ i = 1...n \tag{1}$$

from a depth measurement d_i of laser beam i to a point $\mathbf{x}_i$ in three dimensional space and

$$\begin{aligned} &\pi_{camera}(\mathbf{p}_i) \rightarrow \mathbf{v}_i \cdot s + \mathbf{t}_i \\ &,\ s \in \mathbb{R},\ \mathbf{p}_i \in \mathbb{R}^2,\ \mathbf{v}_i,\ \mathbf{t}_i \in \mathbb{R}^3,\ \|\mathbf{v}\|_2 = 1 \end{aligned} \tag{2}$$

where $\mathbf{t}_i$ denotes the view point, $\mathbf{v}_i$ the viewing direction and s the distance of point $\mathbf{p}_i$ from the camera. In order to establish correspondences between pixels in the camera frame and depth measurements of the laser beams, the extrinsic calibration given as a $SE3$ transformation ΔP between the camera and the laser scanner has to be determined. The most crucial part for estimating the correct calibration is the choice of the features that we use for quantifying the correctness of a calibration hypothesis. Since

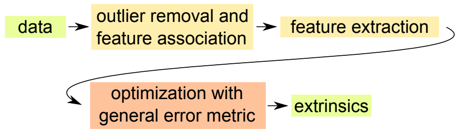


Figure 3: Flow chart of the calibration procedure.

our approach shall be applicable to laser scanners with low resolution and high noise in the depth measurements, classic geometric constraints such as the shape of calibration objects or correspondences between reconstructed points on a checker-board are not applicable. Instead we take advantage of the fact that most camera sensors are sensitive to the infra-red light that laser scanners emit. Hence, we can observe in the camera image the point at which the laser beam meets an object as shown in Fig. 4.

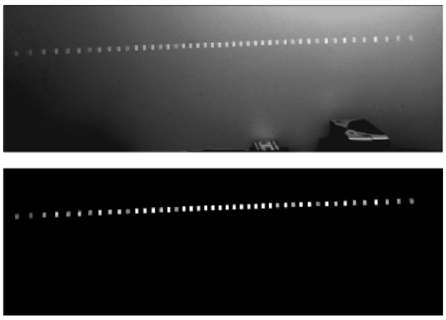


Figure 4: Spots of the laser scanner on a wall seen in the camera image in a darkroom. On top we see the scene with lights on, in the lower image the lights are turned off, with extracted laser spots. Note that the exposure time of the camera control must be tuned so that the laser points are visible.

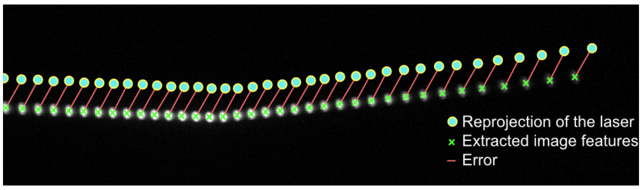
Moving our setup to an environment with low illumination, we can detect laser spots in the camera image. Once the feature points have been extracted as described in section 2.3 and section 2.4, we can derive $\Delta P = f(\alpha, \beta, \gamma, t_x, t_y, t_z)$ by solving the minimization problem in equation 3. Therein, α,β,γ indicate the rotation angles, t_x,t_y,t_z the translation vector and $f(\ldots)$ a mapping from a pose given in three angles and translation to $SE3$. The steps of the calibration procedure are illustrated in Fig. 3.

$$\Delta \hat{P} = \underset{\Delta P = f(\alpha,\beta,\gamma,t_x,t_y,t_z)}{\operatorname{argmin}} \sum_i err_i(\pi_{camera}(\mathbf{p}_i), \Delta P \cdot \pi_{laser,i}(d_i)) \tag{3}$$

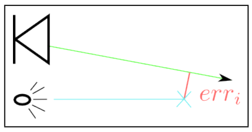
The function $err_i(\ldots,\ldots)$ returns a quantitative error between the viewing ray and the laser beam. The choice of this error function is dependant on the camera model used which is described in section 2.2.

2.2 Using different camera models

Depending on the distortion of the camera lens, we have to formulate the function $err_i(\ldots,\ldots)$ appropriately. The most general way of describing the calibration of a camera is the mapping from pixel to viewing ray given in equation 2. In that case, we formulate the error as the sum of the distances from the viewing rays to the laser scanner measurements and the minimization problem results in the following equation



(a)



(b)

Fig. 5a: Error (red) used for pinhole camera models. The error is the Euclidean distance between the projection of the laser beam (cyan) in the image and the extracted feature in image coordinates (green).
Fig. 5b: Error $err_i = \mathbf{n}^T(\Delta P\mathbf{x}_i - \mathbf{t}_i)$ used for general camera models. The normal $\mathbf{n}$ to the viewing ray is shown in red. The error is the minimal Euclidean distance between the laser scanner measurement (cyan) in $\mathbb{R}^3$ to the viewing ray resulting from the extracted feature (green).

$$\Delta\hat{P} = \underset{\Delta P = f(\alpha,\beta,\gamma,t_x,t_y,t_z)}{\operatorname{argmin}} \sum_i \left\| (\Delta P\mathbf{x}_i - \mathbf{t}_i) \times \mathbf{v} \right\|_2^2 \tag{4}$$

This error is illustrated in Fig. 5b.
Applying the error metric, even cameras that have to be described by highly non-linear camera-models such as fisheye cameras or even cadadioptric cameras with a non-single-view-point can be used in the calibration setup.
For pinhole cameras the back-projection error is a more adequate choice as described by Hartley and Zissermann [**?**]. In case of a pinhole camera model, a unique mapping from $\mathbf{x}_i \in \mathbb{R}^3$ to $\mathbf{p}_i \in \mathbb{R}^2$ exists, which is $\pi^{-1}_{camera}(\mathbf{x}_i)$. Therefore, we can project the estimated position of the laser measurement into the camera frame and minimize the squared distance to the observed position p_i in the camera , see Fig. 5a. The minimization problem for a pinhole camera model is formulated in equation 5.

$$\Delta\hat{P} = \underset{\Delta P = f(\alpha,\beta,\gamma,t_x,t_y,t_z)}{\operatorname{argmin}} \sum_i \left\| \pi^{-1}_{camera}(\Delta P\mathbf{x}_i) - \mathbf{p}_i \right\|_2^2 \tag{5}$$

In order to represent the observable measurement space as good as possible, it is important to generate measurements at varying distances. This can be achieved easily by putting a cardboard in front of the sensor and varying its distance.

2.3 Outlier removal and feature association

A crucial step for the success of this method is the proper extraction of visual measurements and their association with the corresponding laser beam. Especially at high distances between sensors and cardboard, the measured points in the image can have a very low brightness, whereas at small distances the points meld as shown in Fig. 6. To overcome this phenomenon, we overlap the camera images from different points in time, which results in a fan-like image, representing the epipolar geometry of the laser scanner beams in the image. Subsequently we extract these lines manually and extract pixels with maximum brightness in their proximity. The overlapped image is shown in Fig. 7. In that way the association of camera measurements to the corresponding laser beams can be established very robustly and outliers can be negated.

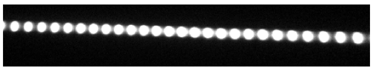


Figure 6: When the distance between sensor and cardboard is low, the sensor tends to overflow. As a result the measurements meld into each other.

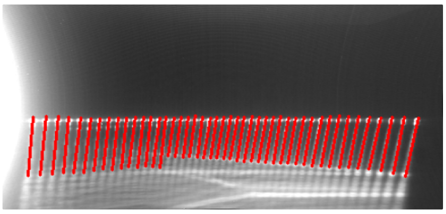


Figure 7: Overlapped images of the infra-red light of the laser scanner at different distances. In red we see the manually marked lines used for measurement extraction.

2.4 Feature extraction

The most important criterion for a high quality calibration is the precision of the extracted features. Due to the absence of background illumination in our setting, the localization of the measurement points can be done very precisely. Using the extracted epipolar lines as described in section 2.3, we guide the extraction of the features, as shown schematically in Fig. 8.
As preprocessing we apply simple Gaussian smoothing to reduce noise. Henceforth, for each epipolar line, we search for the pixel with maximum luminance located in its proximity. In that way features are extracted precisely and associated in one single step.

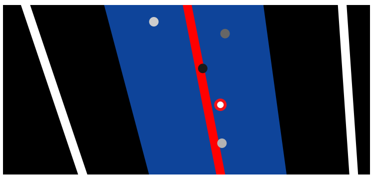


Figure 8: Sketch of the feature extraction methodology. The red line is the epipolar line of interest. The visual measurement point (red) corresponding to this line is the point with maximum luminance in its area of interest (blue).

3 Results

In order to demonstrate the applicability of this method to laser scanners with both high and low range, we apply the proposed algorithm to the calibration of two setups, each with a different laser scanner.

1. A prototype $Spies\ RMS4/90-106B$ [?](Fig. 9a) with a depth accuracy of 50 cm, a maximum range of 100 m and a horizontal resolution of $2°$.

2. An off-the-shelf $Pepperl\&Fuchs\ R2000$ [?](Fig. 9b) with a depth accuracy of

(a) *Spies RMS*4/90 − 106*B* and camera

(b) *Pepperl*&*Fuchs R*2000 and camera

1.4 *cm* and a maximum range of 10 *m*. Its maximum horizontal resolution is $0.07°$, but it will be tuned to $2°$ for the calibration.

As camera, we use an off-the-shelf $PointGreyFlea2(FL2-14S3M)$, with global shutter and 1.4 megapixels at 15 frames per second.

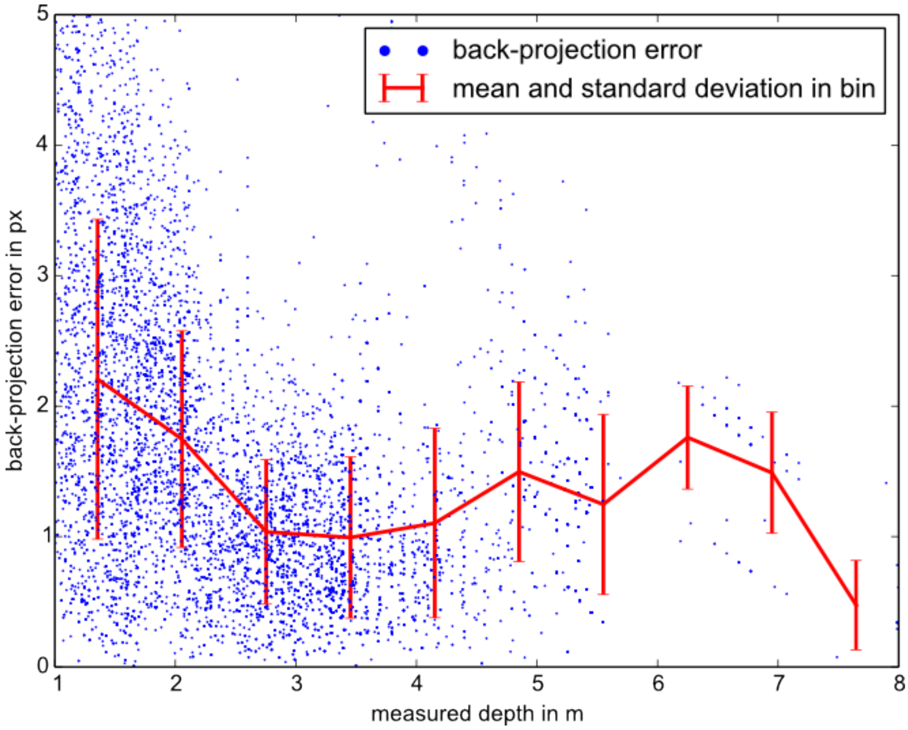
The measurements for the calibration are elaborated in a darkroom without background illumination. In order to describe the quality of the extracted features we calculated the back-projection error as defined in equation 5. Since the depth noise of the sensor dominates the uncertainty of the back-projection error for small distances, the mean error is reciprocal to the depth. Therefore, we evaluated it at different distances, as shown in Fig. 10, demonstrating the convergence of the algorithm. Some qualitative results after calibration of the $SpiesRMS4/90-106B$ can be seen in Fig. 2 and for the $Pepperl\&Fuchs\ R2000$ in Fig. 1.
Although the back-projection error supplies a rough idea about the calibration accuracy, a qualitative evaluation requires a procedure that is independent from the calibration procedure itself in order do avoid overfitting effects, which is described in the subsequent section 4.

4 Evaluation

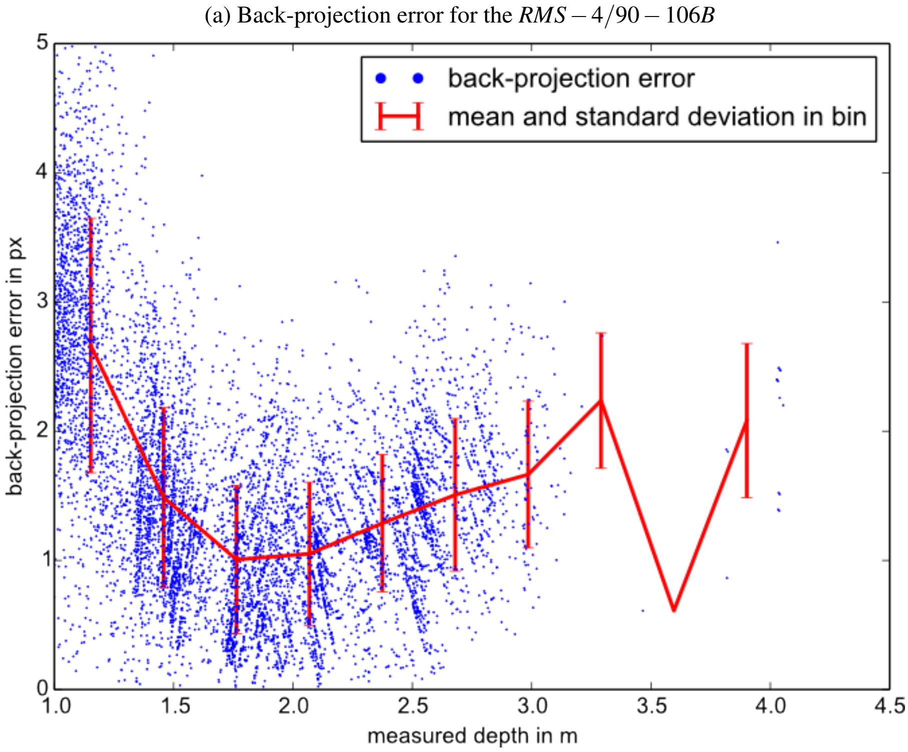
A valuable evaluation of the calibration error, must be based on a set of features with higher accuracy than the one being optimized. Therefore, we will take advantage of the fact that the $Pepperl\&Fuchs\ R2000$ possesses a maximum horizontal resolution of $0.7°$ enabling the use of features that are more precise. In that way the full precision of our method is demonstrated exemplarily for this sensor.
In the following, we apply a methodology similar to established laser scanner to camera calibration by registering edges in the laser scan and in the image. Natural environments can deliver these features, however their quality might vary according to the scene and the features are not reproducible. Moreover their extraction can hardly be automatized.
Consequently, we use a custom made test object, which can be reproduced easily, shown in figure 11.

Basically it consists of a card box with two slits on its front. Its face is covered with white paper, its inside is coloured black. Except for the two frontal slits, the card box is closed so that no light can enter.
With this setup the edges in the laser scan and in the image can be extracted and compared by the following methodology:

1. (a) Cluster the point-cloud by the depth gradient.



(a) Back-projection error for the $RMS-4/90-106B$



(b) Back-projection error for the $R2000$

Figure 10: Back-projection errors for $Pepperl\&Fuchs$ $R2000$ and the prototype $Spies$ $RMS-4/90-106B$. Whereas the $R2000$ possesses a very good depth precision of less than 1.4 cm for a 6 sigma-range, its maximum range is one order of magnitude lower than for the $RMS-4/90-106B$. As a result the $R2000$ does not emit enough light to detect features at distances bigger than 3 m. Our method is capable of calibrating both of these very different laser scanners, since both, precise near features as well as less precise far features, minimize the error metric.

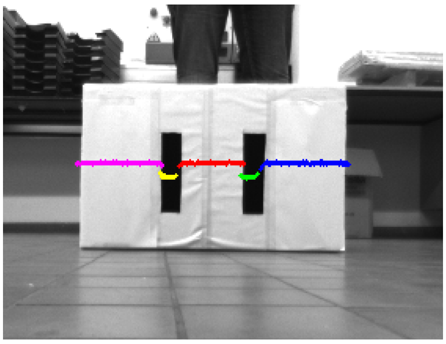


Figure 11: Card box with back-projected points from the laser scanner.

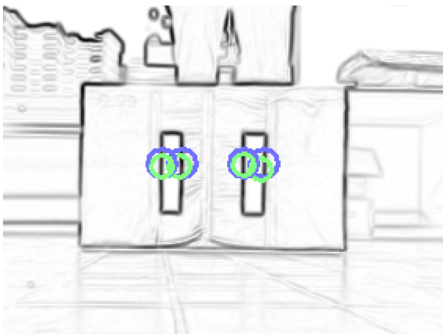


Figure 12: Inverted gradient image with points on the edges found in the camera (blue) and the corresponding projected points from the laser scanner (green). The euclidean distance between both describes the quality of the calibration.

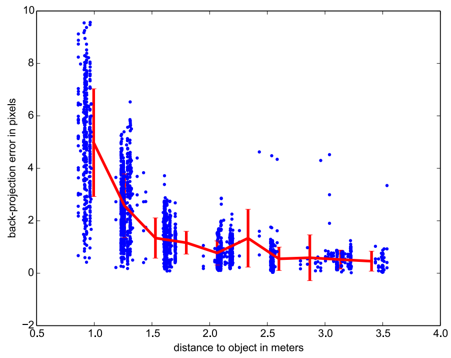


Figure 13: Result of the laser scanner to camera calibration for the $Pepperl\&Fuchs\ R2000$. The calculated errors are shown in blue with mean and standard deviation shown in red. The back-projection error on the edges of the card box slits is reciprocal to the distance to the card box, since the depth noise of the laser scanner becomes dominant for small distances.
For distances over 2 m, the calibration error of around 0.5 px with a standard deviation of 0.25 px becomes visible.

(b) Find the pattern of clusters that fits the geometry of the card box, as shown in Fig. 11. The length of the box as well as its depth in the slits is matched.

(c) Extract the edges in the laser scan. The edge points of the clusters corresponding to the slit edges are projected in the image using the calibration between laser scanner and camera.

2. (a) Calculate the gradient image $G = \sqrt{G_x^2 + G_y^2}$, where G_x and G_y denote the sobel filtered image in directions x and y.

   (b) Project all points on the front of the card box to the image and fit a line through it.

   (c) Extract the edges in the camera image. We extract the image point in G on the fitted line. We use binning to suppress local maxima which do not correspond to edge points.

3. For each edge of the slit calculate the back-projection error as defined in equation 5, see Fig. 12.

Again, we evaluate the error dependent on the depth to the box as presented in figure 13 for these precise features.
This additional error metric serves as a quantification of the accuracy of our method.
Whereas at small distances the depth noise of the laser scanner dominates the back-projection error, the error of the calibration can be observed at higher distances. The mean back-projection error drops below 1 px at a distance of 2 m converging to 0.5 px and a standard deviation of 0.25 px. This error is in the same order as the error of the edge detector, demonstrating the high precision of our calibration method.

5 Conclusion

Cross-calibration approaches for cameras and laser scanners are important to be able to combine both sensor types in mobile applications. However, most of the existing calibration approaches require high resolution and highly accurate sensors, which is not affordable in many applications.
Different from these approaches, we have introduced a new calibration method in this paper which is applicable also for low resolution laser scanner systems and for sensors with moderate measurement accuracy. Our approach does not require full three dimensional reconstruction of the image points, but is based on the minimization of the back projected error of laser scanner points in the camera image. We make use of the fact that the points sensed by the laser scanner become visible in the camera image if the background illumination is low.
The new method has been empirically evaluated for two different laser scanner systems to show the broad applicability of the approach. For both setups the calibration procedure returned high quality results with back-projection errors in the order of magnitude of only one pixel. An independent, quantitative evaluation of the calibration accuracy based on a slitted card box has confirmed the results.
The full source code as well as sample data is freely available[2].
Disposing of cross-calibration with an error of only one pixel enables the application of a camera-laser-scanner combination in many tasks like obstacle detection, sparse world reconstruction, segmentation, object tracking, among others. Therefore, our new method allows to integrate successfully low-cost low resolution laser-scanner-systems for many applications in mobile robotics, surveillance, advanced driver assistance systems and automated driving.

References

[1] GEIGER, A., MOOSMANN, F., CAR, Ö., AND SCHUSTER, B. Automatic camera and range sensor calibration using a single shot. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on* (2012), IEEE, pp. 3936–3943.

[2] HARTLEY, R., AND ZISSERMAN, A. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[3] IBEO. Ibeo mass production laser scanners, may 2016.

[4] LI, G., LIU, Y., DONG, L., CAI, X., AND ZHOU, D. An algorithm for extrinsic parameters calibration of a camera and a laser range finder using line features. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on* (2007), IEEE, pp. 3854–3859.

[5] LIN, K.-H., CHANG, C.-H., DOPFER, A., AND WANG, C.-C. Mapping and localization in 3d environments using a 2d laser scanner and a stereo camera. *Journal of information science and engineering 28*, 1 (2012), 131–144.

[6] PANDEY, G., MCBRIDE, J., SAVARESE, S., AND EUSTICE, R. Extrinsic calibration of a 3d laser scanner and an omnidirectional camera. In *7th IFAC symposium on intelligent autonomous vehicles* (2010), vol. 7.

[2]Source code available on GitHub: *https://github.com/KIT-MRT/PLCC*

[7] PEPPERL&FUCHS. Pepperl&fuchs, r2000, may 2016.

[8] RUSINKIEWICZ, S., AND LEVOY, M. Efficient variants of the icp algorithm. In *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on* (2001), IEEE, pp. 145–152.

[9] SCARAMUZZA, D., HARATI, A., AND SIEGWART, R. Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on* (2007), IEEE, pp. 4164–4169.

[10] SPIES. Engineers spies, may 2016.

[11] TAYLOR, Z., AND NIETO, J. Motion-based calibration of multimodal sensor arrays. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on* (2015), IEEE, pp. 4843–4850.

[12] ZHANG, J., AND SINGH, S. Visual-lidar odometry and mapping: Low-drift, robust, and fast. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on* (2015), IEEE, pp. 2174–2181.

[13] ZHANG, Q., AND PLESS, R. Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on* (2004), vol. 3, IEEE, pp. 2301–2306.