Targetless Calibration of a Lidar - Perspective Camera Pair

Levente Tamas
Technical University of Cluj-Napoca
Baritiu st. 24, 400118, Cluj-Napoca, Romania
Levente.Tamas@aut.utcluj.ro

Zoltan Kato
University of Szeged
P.O. Box 652, H-6701 Szeged, Hungary
kato@inf.u-szeged.hu

Abstract

A novel method is proposed for the calibration of a camera - 3D lidar pair without the use of any special calibration pattern or point correspondences. The proposed method has no specific assumption about the data source: plain depth information is expected from the lidar scan and a simple perspective camera is used for the 2D images. The calibration is solved as a 2D-3D registration problem using a minimum of one (for extrinsic) or two (for intrinsic-extrinsic) planar regions visible in both cameras. The registration is then traced back to the solution of a non-linear system of equations which directly provides the calibration parameters between the bases of the two sensors. The method has been tested on a large set of synthetic lidar-camera image pairs as well as on real data acquired in outdoor environment.

1. Introduction

One of the most challenging issue in robotic perception applications is the fusion of information from several different sources. Today the majority of the platforms include range (2D or 3D sonar/lidar) and camera (color/infrared) sensors that are usually work independently, although the information from different sources can be used in a complementary way.

In order to fuse the data coming from these independent devices, the measured data has to be transformed into a common coordinate frame. This is achieved by either extrinsic or intrinsic-extrinsic calibration, depending on whether the internal camera parameters are available or not. In case of the extrinsic parameter estimation for a range-camera sensor pair the 3D rigid translation between the two coordinate systems is determined.

The case of extrinsic parameter estimation for 2D/3D lidar and perspective camera has been performed especially for environment mapping applications, however this problem is far from being trivial. Due to the different ways of functionality of the lidar and camera, the calibration is often performed manually, or by considering special assumptions like artificial markers on images, or establishing point matches. These procedures tend to be laborious and time consuming, especially when calibration has to be done more than once during data acquisition. In real life applications, it is often desirable to have a flexible one step calibration for systems which are not necessarily containing sensors fixed to a common platform.

1.1. Related work

The need for registration between heterogeneous sensor data is common to multiple research fields including aerial remote sensing [13, 11], medical images processing [18] or mobile robotic applications [5, 22, 17, 20]. Different approaches are tackling the non-trivial 2D-3D registration problem: point or line correspondence finding between the two domains [11, 21], intensity image based correlation [17], use of specific artificial land-mark [6, 7, 9, 1, 14, 28, 23] or mutual information extraction and parameter optimization[22, 26, 25].

The extrinsic calibration of 3D lidar and low resolution color camera was first addressed in [23] which generalized the algorithm proposed in [30]. This method is based on manual point feature selection from both domains and it assumes a valid camera intrinsic model for calibration. A similar manual point feature correspondence based approach is proposed in [20].

There are also extensions to the simultaneous intrinsic-extrinsic calibration presented in the work [12] which used the intensity information from lidar to find correspondences between the 2D-3D domains. Other works are based on the fusion of IMU or GPS information in the process of 2D-3D calibration [15], mainly in initialization phase of the calibration [22].

Recently there has been an increasing interest in various calibration problem setups ranging from high-resolution spatial data registration [11] to low-resolution, high frame rate depth commercial cameras such as Kinect [7, 29], or in the online calibration during different measurements in time such as in case of a traveling mobile robot [17, 10].

1.2. Contributions

This paper proposes a novel region based calibration framework for 2D and 3D sensors emerging from different sources in a single step automatic manner. The main contribution is the formulation of the calibration problem as a general 2D-3D registration problem which works without special targets or established point matches. The registration is accomplished by solving system of nonlinear equations based on the idea of [4]. However, the equations are constructed in a different way here due to the different dimensionality of the lidar and camera coordinate frames. This correspondence-less registration framework in the context of the calibration is well suitable for both extrinsic and intrinsic-extrinsic calibration. The concept was proven to be viable on a large scale synthetic dataset as well as on real data. The robustness against segmentation error has also been demonstrated and comparative tests confirmed the advantages of the proposed method over state-of-the-art.

2. Region-based calibration framework

Consider a lidar camera with a 3D coordinate system having its origin $\mathbf{O}$ in the center of laser sensor rotation, x and y axes pointing to the right and down, respectively, while z is pointing away from the sensor. Setting the world coordinate system to the lidar's coordinate system, we can always express a 3D lidar point $\mathbf{X}$ with its homogeneous world coordinates $\mathbf{X} = (X_1, X_2, X_3, 1)^T$. The perspective camera sees the same world point $\mathbf{X}$ as a homogeneous point $\mathbf{x} = (x_1, x_2, 1)^T$ in the image plain obtained by a perspective projection $\mathbf{P}$:

$$\mathbf{x} = \mathbf{P}\mathbf{X} \tag{1}$$

where $\mathbf{P}$ is the 3×4 camera matrix, which can be factored into the well known $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}|\mathbf{t}]$ form, where $\mathbf{I}$ is the identity matrix, $\mathbf{K}$ is the 3×3 upper triangular *calibration* matrix containing the camera intrinsic parameters:

$$\mathbf{K} = \begin{pmatrix} f_x & & o_x \\ & f_y & o_y \\ & & 1 \end{pmatrix}, \tag{2}$$

while $\mathbf{R}$ and $\mathbf{t}$ are the rotation and translation, respectively, aligning the camera frame with the world coordinate frame. Therefore we have 4 intrinsic and 6 extrinsic parameters. Note that in $\mathbf{K}$ we implicitly assume rectangular pixels (*i.e.* no skew), f_x and f_y denotes the number of pixels per world coordinate units along the x (resp. y) axes, and $\mathbf{o} = (o_x, o_y)^T$ is the principal point of the camera on the image plane.

A classical solution of the calibration problem is to establish a set of 2D-3D point matches using a special calibration target [23, 14, 1, 7, 12], and then solve for $\mathbf{P}$ via a system of equation based on (1) or the minimization of some error function. When a calibration target is not available, then solutions typically assume that the lidar points contain also the laser reflectivity value (interpreted as a gray-value), which can be used for intensity-based matching or registration [17, 20, 11, 14].

However, in many practical applications (*e.g.* infield mobile robot), it is not possible to use a calibration target and most lidar sensors will only record depth information. Furthermore, lidar and camera images might be taken at different times and they need to be fused later based solely on the image content. Therefore the question naturally arises: what can be done when neither a special target nor point correspondences are available? Herein, we propose a solution for such challenging situations. In particular, we will show that by identifying a single planar region both in the lidar and camera image, the extrinsic calibration can be solved. When two such non-coplanar regions are available then the full calibration can be solved. Of course, these are just the necessary minimal configurations. The more such regions are available, a more stable calibration is obtained.

Our solution is based on the 2D shape registration approach of Domokos *et al*. [4], where the alignment of nonlinear shape deformations are recovered via the solution of a special system of equations. Here, however, the calibration problem yields a 2D-3D registration problem, which requires a different technique to construct the system of equations: Since correspondences are not available, (1) cannot be used directly. However, individual point matches can be integrated out yielding the following integral equation:

$$\int_{\mathcal{D}} \mathbf{x} d\mathbf{x} = \int_{\mathbf{P}\mathcal{F}} \mathbf{z} d\mathbf{z}, \tag{3}$$

where $\mathcal{D}$ corresponds to the region visible in the *camera* image and $\mathbf{P}\mathcal{F}$ is the image of the *lidar region* projected by the camera matrix $\mathbf{P}$. The above equation corresponds to a system of 2 equations only, which is clearly not sufficient to solve for all parameters of the camera matrix $\mathbf{P}$. Therefor we adopt the general mechanism proposed in [4] to construct new equations. Indeed, (1) remains valid when a function $\omega : \mathbb{R}^2 \to \mathbb{R}$ is acting on both sides of the equation

$$\omega(\mathbf{x}) = \omega(\mathbf{P}\mathbf{X}), \tag{4}$$

and the integral equation of (3) becomes

$$\int_{\mathcal{D}} \omega(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{P}\mathcal{F}} \omega(\mathbf{z}) d\mathbf{z}. \tag{5}$$

Adopting a set of nonlinear functions $\{\omega_i\}_{i=1}^{\ell}$, each ω_i generates a new equation yielding a system of ℓ independent equations. Hence we are able to generate sufficiently many equations. The parameters of the camera matrix $\mathbf{P}$ are then simply obtained as the solution of the nonlinear system of equations (5). In practice, an overdetermined system is constructed, which is then solved by minimizing the algebraic

error in the *least squares sense* via a standard *Levenberg-Marquardt* algorithm.

Note that computing the integral on the right hand side of (5) involves the actual execution of the camera projection $\mathbf{P}$ on $\mathcal{F}$, which might be computationally unfavorable. However, choosing power functions for ω_i:

$$\omega_i(\mathbf{x}) = x_1^{n_i} x_2^{m_i}, \quad n_i \leq 3 \text{ and } m_i \leq 3 \tag{6}$$

and using a triangular mesh representation $\mathcal{F}^{\triangle}$ of the lidar region $\mathcal{F}$, we can adopt an efficient computational scheme. First, let us note that this particular choice of ω_i yields the 2D geometric moments of the projected lidar region $\mathbf{P}\mathcal{F}$. Furthermore, due to the triangular mesh representation of $\mathcal{F}$, we can rewrite the integral adopting ω_i from (6) as

$$\int_{\mathcal{D}} x_1^{n_i} x_2^{m_i} d\mathbf{x} = \int_{\mathbf{P}\mathcal{F}} z_1^{n_i} z_2^{m_i} d\mathbf{z} \approx \sum_{\forall \triangle \in \mathcal{F}^{\triangle}} \int_{\triangle} z_1^{n_i} z_2^{m_i} d\mathbf{z}. \tag{7}$$

The latter approximation is due to the approximation of $\mathcal{F}$ by the discrete mesh $\mathcal{F}^{\triangle}$. The integrals over the triangles are various geometric moments which can be computed using *e.g.* the following formula for $x^p y^q$ [2]:

$$2 \sum_{k=0}^{p} \sum_{l=0}^{q} \frac{(-1)^{k+l} \binom{p}{k} \binom{q}{l} \nu_{kl}}{k+l+2} x_0^{p-k} y_0^{q-l} \tag{8}$$

where

$$\nu_{kl} = \sum_{i=0}^{k} \sum_{j=0}^{l} \frac{\binom{k}{i} \binom{l}{j}}{k-i+l-j+1} (x_0 - x_1)^i (x_1 - x_2)^{k-i} (y_0 - y_1)^j (y_1 - y_2)^{l-j} \tag{9}$$

with the notation x_i and y_i, $i = 0 \ldots 2$ for the vertices of the triangles.

The summary of the numerical implementation of the proposed method is presented in Algorithm 1. Note that normalization is critical to ensure a numerically stable solution (see [4] for details). For extrinsic calibration, $\mathbf{K}$ is known a priori, while for intrinsic-extrinsic calibration, it is initialized with $f_x = f_y = 800$ (corresponding to a normal 65 degree field of view) and $\mathbf{o}$ is set to the center of the image.

3. Evaluation on synthetic data

For the quantitative evaluation of the proposed method, we generated a benchmark set containing 29 different shapes and their transformed versions, a total number of 2D-3D data pairs exceeding 2500 samples divided into different test setups. The 3D-2D image pairs were generated in

Algorithm 1 The proposed calibration algorithm

Input: 3D point cloud and 2D binary image representing the same region, and the calibration matrix $\mathbf{K}$

Output: Parameters of the camera matrix $\mathbf{P}$

1: Normalize 3D points into the unit cube and 2D point into the unit square centered in the origin.
2: Triangulate the region represented by the 3D point cloud.
3: Construct the system of equations of (7) with the polynomial ω_i functions of (6).
4: Initialize the camera matrix as $\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}]$.
5: Solve the nonlinear system of equation in (7) using the Levenberg-Marquardt algorithm
6: Unnormalize the solution.

Figure 1. Synthetic contour noise examples (left to right): original shape, 5%, and 10% contour noise

the following way: The 3D image were generated by placing 2D planes in the 3D Euclidean space and the shapes were placed on these planes, whose size was normalized into a 1m^3 cube, with a random initial rotation in the range of $-\frac{\pi}{16}, \ldots, \frac{\pi}{16}$ and a depth of 10m. For images with 2 and 3 planes, a similar procedure was applied and the placement of planes relative to each other was initially perpendicular followed by a random translation of $5 - 15$m and rotation of $-\frac{\pi}{16}, \ldots, \frac{\pi}{16}$. The initial placement of the 3D planes was considered by taking into account common urban structural properties, in which environment the real data experiments were performed.

The synthetic 2D images of the 3D data were generated with a camera being rotated in range of $-\frac{\pi}{4}, \ldots, \frac{\pi}{4}$ and the random displacement of $2 - 10$m. The camera calibration matrix $\mathbf{K}$ used for projection had a random focal length f_x, f_y in the range of $400 - 1600$ with a 5% difference between f_x and f_y, and the principal point $\mathbf{o}$ was set to the center of the 1024×768 image plane with an added 5% random variation. Thus random 2D projections were obtained with the so defined random camera matrix $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}|\mathbf{t}]$.

In practice, the planar regions used for calibration are segmented out from the lidar and camera images. In either case, we cannot produce perfect shapes, therefore robustness against segmentation errors was also evaluated on simulated data (see samples in Fig. 1): we randomly added or removed squares uniformly around the boundary of the shapes, both in 3D and 2D, yielding an error around the

contour of 5% and 10% of the original shape. Using these corrupted images, we tested the robustness with respect to 3D errors using the 3D contour noise corrupted images and corresponding noise-less 2D projections as well as robustness with respect to 2D image errors using 2D contour noise corrupted images and their noise-less 3D images.

The algorithm was implemented in Matlab and all test cases were run on a standard quad-core PC. Calibration errors were characterized in terms of the percentage of non-overlapping area of the reference and registered images (denoted by δ), the Frobenius norm of the difference between the found and the true camera matrices, as well as differences in the external parameters.

3.1. Extrinsic parameter estimation

In this test case the results for extrinsic parameter estimation using a virtual camera with known $\mathbf{K}$ is presented. The plots contain information about the test cases with 1, 2, and 3 planes used for calibration, as well as the robustness test results with corrupted 3D and 2D regions.

In Fig. 2 the result for translation error between the camera and lidar coordinate frames is presented. The translation error is not the same on the different axes: the largest is on the Y axis, as in this direction the resolution of the camera was lower than along the X axis. The best results were achieved with the 3 plane set as this is the most constrained setup for the camera-lidar position. For the majority of the cases the error was less than 3cm for a range of 10m. Segmentation errors increase these errors, but it is robust enough up to 10% error level.

In Fig. 3 the result for rotation error between the camera and lidar coordinate frames is presented. The results are not the same for the different axes. The most robustness is manifested along the Z axis rotation: this is not surprising as this axis is perpendicular to the observation plane, hence rotation causes only a minimal distortion on the image. Other axes are more affected by error, but for more than 70% of the cases the error was lower than 0.5 degrees on each axis, wich is quite good for a 45 degrees rotation range.

The registration error δ and the runtime together with the Frobenius norm of the projection matrix is shown in Fig. 4. The contour noise on the 3D template, which can be interpreted as depth noise, causes a larger error than segmentation errors in the 2D camera image. This is visible in the δ error plot as well as in the Frobenius norm. The runtime for the majority of the cases is less than 2min.

3.2. Intrinsic-extrinsic parameter estimation

In this section the results for intrinsic-extrinsic parameter estimation are summarized. Since we need at least two non-coplanar regions for calibration, hence we only used the 2 and 3 plane datasets. Since the translation and rotation evaluation for the individual axes gave similar results

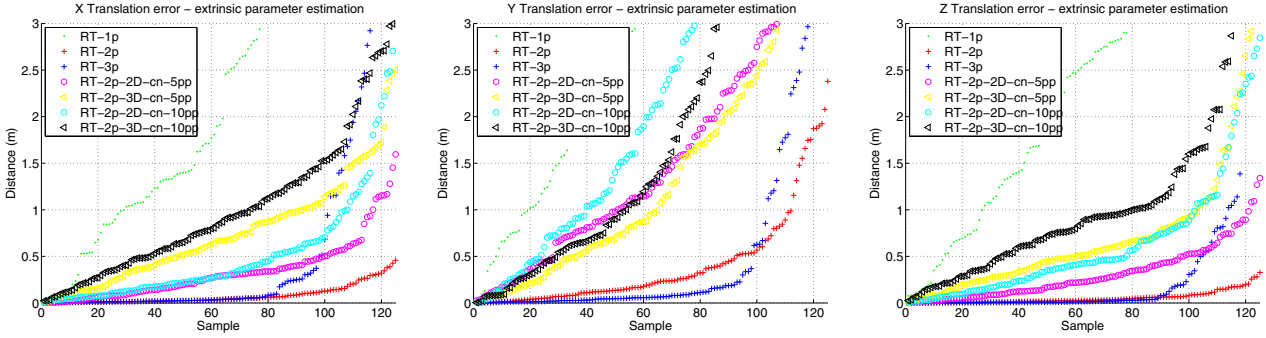

Figure 2. Extrinsic calibration results (from left-to-right) the translation errors along X, Y and Z axis.

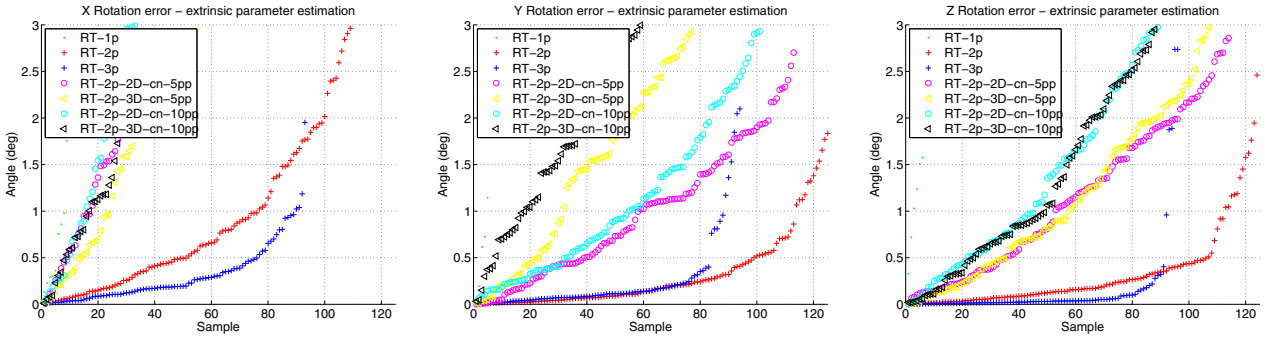

Figure 3. Extrinsic calibration results, including (from left-to-right) the rotation errors around X, Y and Z axis.

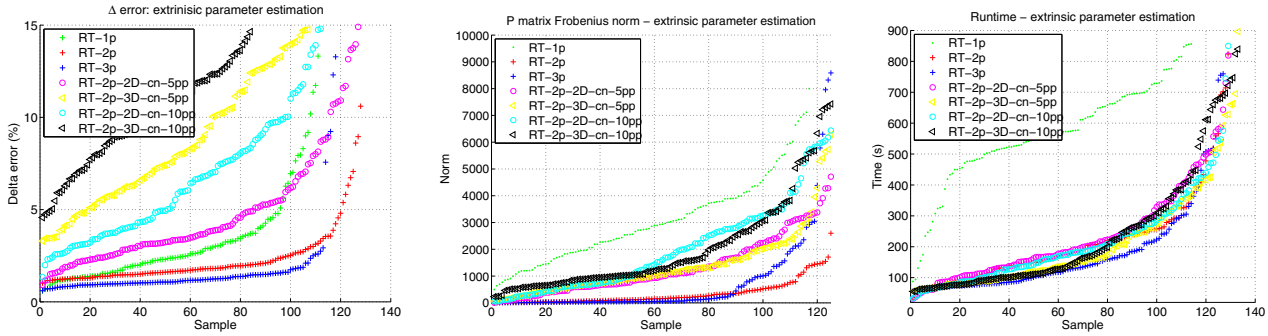

Figure 4. Synthetic data calibration results for extrinsic parameter estimation, including the δ error, the Frobenius norm and the runtime

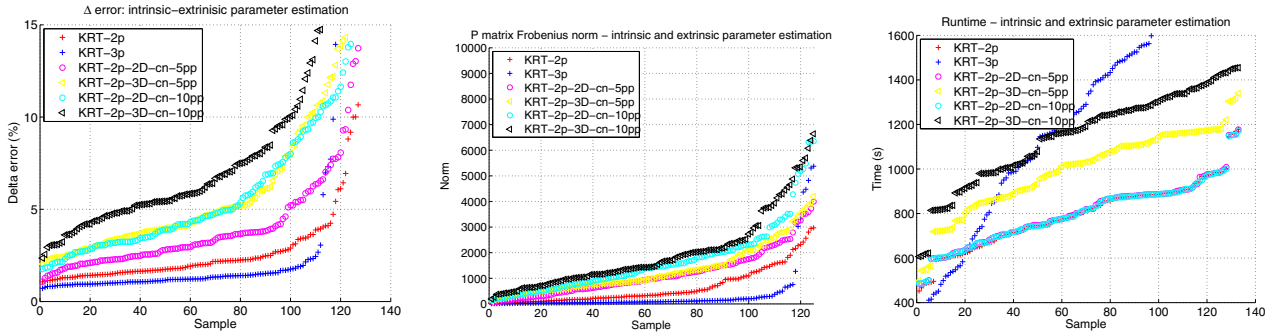

Figure 5. Intrinsic-extrinsic calibration results (from left-to-right) the δ error, the Frobenius norm and the runtime.

as in the case of the extrinsic parameter estimation, only the overall error metrics are highlighted.

The δ registration error in this test unit was the lowest for the templates containing 3 non-coplanar planes as shown in Fig. 5. The intrinsic-extrinsic registration for contour noise corrupted images seemed to be more robust than the case with only extrinsic registration. This can be explained by considering the higher degrees of freedom for the optimizer during the parameter estimation. This is also visible on the Frobenius norm plot compared to the case with the extrinsic parameter estimation.

3.3. Comparison with other methods using synthetic data

The comparison with other calibration methods could be performed only in a limited manner due to the basic differences of the proposed method with the existing ones in this field of research. The already existing methods differ either by the idea of using artificial landmarks as in the case of [6, 7, 9, 1, 14, 23] or by the use of multiple image-depth pairs like in [17, 10].

The approaches which are based on high resolution airbone lidar-camera information registration are often used for a very narrow transformation range between the two sensors [11] in contrast to the proposed algorithm which can be used for almost arbitrary transformations with the restriction mentioned before.

The low-resolution depth camera such as the projected pattern infrared cameras (like the Kinect) calibrations usually deal also with the internal model estimation for the depth camera, which are part of the global multi-phase optimization [7].

The class of algorithms which are built on reflectivity values such as the ones presented in [17, 26] are not an option for lidar without reflectivity information such as the tilted LMS200 based Sick Laser scanner used in our real data sets.

Methods using some special target, like a checkerboard calibration pattern [23] or other fiducial markers [1] were tested on an appropriate dataset for extrinsic calibration using the code provided by the authors. For both cases more than 2 different 2D-3D data pairs are required for the external calibration, hence the test was performed on 3 and 6 regions. Each of the methods was implemented in Matlab and the tested on Linux 32-bit machine, with 2D images having 640×480 resolution and the depth in the range of $1 - 10m$. The results of the tests are presented in the Table 1.

The method proposed by Alismail *et al.* [1] is quite sensitive to the size of the fiducial marker, which was a circle placed in the Euclidean space. While the method is computationally efficient, already a 2% error in the radius of the 3D point cloud caused noticeable error. Since this method requires at least six different planes and a specific marker for the calibration, real data test were not available.

Method	δ error (%)	Runtime (s)
HSA2012 - 6 regions [1]	7.4	8.1
RU2005 - 3 regions [23]	8.3	15.9
RU2005 - 6 regions [23]	1.6	32.7
Proposed - 3 regions	0.9	54.2
Proposed - 6 regions	0.8	49.6

Table 1. Comparative results with [1] and [23] for 3 and 6 regions.

The work of Unnikrishnan and Hebert [23] proved to be scalable in terms of the number of planes. It is suggested by the authors that at least 10 image-depth pairs should be considered for calibration, although it gives reasonable results for fewer planes. The performance of this method was comparable to the proposed one for 6 planes, but it is suitable only for extrinsic calibration with specific markers.

4. Real data experiments

Next, we present some calibration results on real data. The planar regions are simple rectangular regions segmented in the 2D and 3D domains. The lidar and camera images are taken from different viewpoints.

4.1. Experiment setup

In our experimental setups a custom lidar mounted on a mobile robot provided the 3D data and different perspective cameras took either an IR or color image of the same scene. The scanning of the environment with the *P3* type mobile robot was performed using custom laser tilted range finder configuration. This custom lidar data has a $0.5\,\mathrm{deg}$ angular resolution and a $1cm$ depth resolution in a $80m$ depth range.

Color images were taken by a standard camera of 1024×768 resolution with prior calibration and radial distortion removal. The IR image is taken by an industrial Flir camera of 240×240 resolution with unknown intrinsic parameters. The images were captured at night time in order to reduce the effect of solar heat on the building facades.

It is important to highlight that the conventional internal calibration for such a camera is not trivial, calibration of IR and depth data is rather cumbersome, often special setups are needed both for intrinsic and extrinsic calibration [3]. Thus we applied the intrinsic-extrinsic calibration for the IR experiment.

4.2. Region segmentation

In order to make the calibration user friendly the region selection both in 2D and 3D was automated with efficient segmentation algorithms. Considering the correspondence establishment between the segmented 2D and 3D regions as minimal one-click user intervention, this aspect represents the only human interaction in the calibration process.

There are several automated or semi-automated 2D segmentation algorithms in the literature including clustering, histogram thresholding, energy based or region growing variants [24]. In this work we used the region growing algorithm which gave the closest segmentation to the regions which were also segmented in the 3D space [19] extended with a contour extraction assisted registration [27].

A number of sparse 3D point cloud segmentation methods have been developed recently including robust segmentation [16], difference of normals based segmentation [8]. The region growing proved to give stable results in urban environment which could be segmented in 2D domain too.

Figure 6. Calibration example with real outdoor data (left-right): 3D data with the segmented region (yellow); color 2D images with corresponding regions (green); color information overlaid on 3D data using intrinsic - extrinsic parameter estimation for two images.

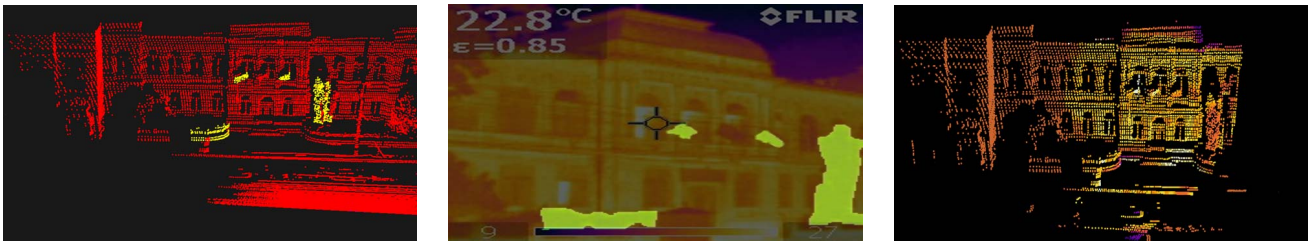

Figure 7. Calibration example with real outdoor data (left-right): 3D data with the segmented region (yellow); infrared 2D data with the segmented region (brighter); infrared information overlaid on 3D data using intrinsic - extrinsic parameter estimation for infrared camera

Figure 8. Calibration example with real data of an outdoor environment (left-right): 3D data with the selected region (yellow); color 2D data with the corresponding region (green); color information overlaid on 3D data using the extrinsic parameter estimates

4.3. Algorithm evaluation on custom data

For the laser scan in Fig. 6 with ranges up to 30m, the point density was rather low at larger distances. As it is also mentioned in [17], in an outdoor environment the accuracy is highly depending on the distances of the regions used for calibration: a larger distance yields less accurate calibration. In order to emphasize the general applicability of the proposed method, we used two color images from very different point of views for the same lidar scan. The result of the projected color data on the 3D data is shown in Fig. 6.

The most interesting calibration experiment was performed with an IR camera which has a rather limited resolution and a narrow field of view. In spite of the large displacement (approximately 10m away and with a 20 degrees rotation) and unknown internal parameters, good result was achieved (see Fig. 7) by our method. Lacking precise ground truth and artificial landmarks, the reprojected data can be examined at the higher intensity parts from the flags on the image which match their 3D point pairs.

4.4. Comparison using public datasets

For the performance evaluation of the proposed method a calibration on the public KITTI dataset (`http://www.cvlibs.net/datasets/kitti/`) was considered, which contained also the ground truth for the camera-lidar devices. In Fig. 8, the extrinsic calibration of a color camera with known $\mathbf{K}$ matrix and sparse 3D lidar data from the drive $nr = 5$ is shown.

In order to evaluate the accuracy of the registration, the transformation parameters were compared against the ground truth values. Also the calibration test was performed using the mutual information extraction described in [22] for 3D data with intensity and normal information. The calibration results are summarized in the Table 2. The results of the proposed method (Prop.) proved to be sensitive to the segmentation accuracy, nevertheless the registration both visually and numerically was accurate in the range of few millimeters translation and around 1 degree rotation. The mutual information based method with lidar intensity (Int.) data gave smaller absolute errors but this method using only depth data (Norm.) became quite sensitive to local minimals and the final calibration error was larger. Also the runtime GPU implementation of the mutual information method is with an order of magnitude slower than the CPU implementation of the proposed algorithm.

Tf.	t_x	t_y	t_z	roll	pitch	yaw
Prop.	0.011	0.029	0.38	1.4	1.9	1.5
Norm. [22]	0.014	0.036	0.41	1.5	2.3	1.6
Int. [22]	0.007	0.026	0.18	0.8	1.2	0.9

Table 2. Comparative results with the proposed method (Prop), normal based MI(Norm)[22] and intensity based MI (Int)[22].

5. Conclusions

A nonlinear explicit correspondence-less calibration method was proposed in this work. The calibration is based on the 3D-2D registration of a common lidar-camera region. The proposed method uses minimal information (only depth data and shape of regions) and is general enough to be used in a great variety of applications. It has been tested on a large synthetic dataset. The algorithm was also validated in real life experiments with different cameras and with both extrinsic and intrinsic-extrinsic calibration experiments.

6. Acknowledgement

This research was supported by the EU and the State of Hungary, co-financed by the European Social Fund through project FuturICT.hu(grant no.: $TAMOP - 4.2.2.C - 11/1/KONV - 2012 - 0013$) and in the framework of "National Excellence Program" $TAMOP - 4.2.4.A/2 - 11 - 1 - 2012 - 0001$ as well as SCIEX-NMS-CH project no. 12.239.

References

[1] H. S. Alismail, L. D. Baker, and B. Browning. Automatic calibration of a range sensor and camera system. In *Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission*, pages 286–292, Zurich, Switzerland, October 2012. IEEE.

[2] G. C. Best. Helpful formulas for integrating polynomials in three dimensions (in Technical Notes and Short Papers). *In-*

ternational Journal of Mathematics and Computer Science, 18(86):310–312, April 1964.

[3] D. Borrmann, H. Afzal, J. Elseberg, and A. Nüchter. Thermal 3d modeling of indoor environments for saving energy. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012, Vilamoura, Algarve, Portugal, October 7-12*, pages 4538–4539. IEEE, 2012.

[4] C. Domokos, J. Nemeth, and Z. Kato. Nonlinear shape registration without correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):943–958, 2012.

[5] A. Geiger, F. Moosmann, O. Car, and B. Schuster. Automatic camera and range sensor calibration using a single shot. In *International Conference on Robotics and Automation*, pages 3936–3943. IEEE, 2012.

[6] X. Gong, Y. Lin, and J. Liu. Extrinsic calibration of a 3d lidar and a camera using a trihedron. *Optics and Lasers in Engineering*, 51(4):394 – 401, 2013.

[7] D. Herrera C, J. Kannala, and J. Heikkila. Joint depth and color camera calibration with distortion correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):1–8, 2012.

[8] Y. Ioannou, B. Taati, R. Harrap, and M. Greenspan. Difference of normals as a multi-scale operator in unorganized point clouds. *ArXiv e-prints*, Sept. 2012.

[9] S. Krause and R. Evert. Remission based improvement of extrinsic parameter calibration of camera and laser scanner. In *Control Automation Robotics Vision (ICARCV), 2012 12th International Conference on*, pages 829–834, 2012.

[10] J. Levinson and S. Thrun. Automatic online calibration of cameras and lasers. In *Proceedings of Robotics: Science and Systems*, Berlin, Germany, June 2013.

[11] A. Mastin, J. Kepner, and J. W. F. III. Automatic registration of lidar and optical images of urban scenes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2639–2646, Miami, Florida, USA, June 2009. IEEE.

[12] F. M. Mirzaei, D. G. Kottas, and S. I. Roumeliotis. 3d lidar-camera intrinsic and extrinsic calibration: Identifiability and analytical least-squares-based initialization. *The International Journal of Robotics Research*, 31(4):452–467, 2012.

[13] R. Mishra and Y. Zhang. A review of optical imagery and airborne lidar data registration methods. *The Open Remote Sensing Journal*, 5:54–63, 2012.

[14] O. Naroditsky, E. P. Iv, and K. Daniilidis. Automatic alignment of a camera with a line scan lidar system. In *International Conference on Robotics and Automation*, pages 3429–3434, Shanghai, China, May 2011. IEEE.

[15] P. Nunez, P. Drews, R. Rocha, and J. Dias. Data fusion calibration for a 3d laser range finder and a camera using inertial data. In *European Conference on Mobile Robots*, pages 31–36, Dubrovnik, Croatia, September 2009.

[16] A. Nurunnabi, D. Belton, and G. West. Robust segmentation in laser scanning 3d point cloud data. In *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*, pages 1–8, 2012.

[17] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice. Automatic targetless extrinsic calibration of a 3d lidar and camera by maximizing mutual information. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 2053–2059, Toronto, Canada, July 2012.

[18] J. Pluim, J. Maintz, and M. Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[19] M. Preetha, L. Suresh, and M. Bosco. Image segmentation using seeded region growing. In *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*, pages 576–583, 2012.

[20] D. Scaramuzza, A. Harati, and R. Siegwart. Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In *IEEE International Conference on Intelligent Robots and Systems*, pages 4164–4169, San Diego, USA, October 2007. IEEE/RSJ, IEEE.

[21] I. Stamos and P. K. Allen. Automatic registration of 2-d with 3-d imagery in urban environments. In *International Conference on Computer Vision*, pages 731–737, 2001.

[22] Z. Taylor and J. Nieto. A mutual information approach to automatic calibration of camera and lidar in natural environments. In *Australian Conference on Robotics and Automation*, pages 3–5, Wellington, Australia, December 2012.

[23] R. Unnikrishnan and M. Hebert. Fast extrinsic calibration of a laser rangefinder to a camera. Technical report, Carnegie Mellon University, 2005.

[24] S. R. Vantaram and E. Saber. Survey of contemporary trends in color image segmentation. *Journal of Electronic Imaging*, 21(4):1–28, 2012.

[25] P. Viola and W. M. Wells, III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, Sept. 1997.

[26] N. Williams, K.-L. Low, C. Hantak, M. Pollefeys, and A. Lastra. Automatic image alignment for 3d environment modeling. In *Computer Graphics and Image Processing, 2004. Proceedings. 17th Brazilian Symposium on*, pages 388–395, 2004.

[27] Y. Xiaohan, J. Yla-Jaaski, O. Huttunen, T. Vehkomaki, O. Sipila, and T. Katila. Image segmentation combining region growing and edge detection. In *Pattern Recognition, 1992. Vol.III. Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on*, pages 481–484, 1992.

[28] A. Zhang, S. Hu, Y. Chen, H. Liu, F. Yang, and J. Liu. Fast Continuous 360 Degree Color 3D Laser Scanner. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36, pages 409–415, Beijing, China, July 2008. ISPRS.

[29] C. Zhang and Z. Zhang. Calibration between depth and color sensors for commodity depth cameras. In *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo*, pages 1–6, Barcelona, Catalonia, Spain, 11-15 July 2011. IEEE.

[30] Q. Zhang. Extrinsic calibration of a camera and laser range finder. In *International Conference on Intelligent Robots and Systems*, pages 2301 – 2306, Sendai, Japan, September 2004. IEEE.