

3D Non-Contact Building Survey Technique

Levente Tamas, Gheorghe Lazea

Robotics Research Group, Dorobantilor st 71-75, Cluj-Napoca, 400609, RO

Levente.Tamas, Gheorghe.Lazea@aut.utcluj.ro

Keywords: 3D laser scanner, pointcloud processing, segmentation, scan registration.

Abstract: This paper presents a non-contact 3D environments mapping technique using mobile robots with different perception devices such as stereo camera, structural light camera or custom made 3D laser scanner. The custom developed non-contact scanning device is suitable for making building surveillance both in small and large scales. These measurements can be further used for urban region planning as well as for navigation purposes for autonomous agents such as mobile vehicles.. The large maps can be built from successive scan merging in an iterative manner. For such an approach is essential the initial matching between two measurements, which in our proposed method was performed using special extracted features from the measured data sets. The registered maps can be further used for perceptual purposes, including object segmentation or planning maps for the mobile agents. For both types of applications, there are given experimental results obtained from the proposed registration and segmentation algorithms.

Introduction

The 3D perception problem represents a emerging research for the robotics community. There are several potential applications for this domain mainly from the urban surveillance, path planning or cultural heritage conservation fields. [1].

Each application requires specific data handling acquired even from the same measurement device. A part of the problems emerging from the mobile robotics domain require real time data processing, like dynamic perception and planning experiments, while for urban mapping often the post processing of the data is sufficient [2].

Different sensors can be used for the data acquisition including stereo cameras, laser range finders, time-of-flight cameras, or the recently adopted structured light sensors [3]. These devices have their own special characteristics in terms of precision, range or speed. Thus the way in which these sensors are selected depends on the specific requirements of the measurement problem to be solved [4].

The large area perceptions, such as the urban area measurements require several different measurements to be aligned in the same coordinate frame. This kind of problem is well studied in the 2D mainly in the image processing domain. Although these 2D algorithms can be adopted for the 3D data registration, they need special adaptation for the spatial data. Also the data characteristics such as range, noise or data distribution have large influence on the algorithms used for 3D data processing [5].

For the data merging different algorithms can be used including keypoint and feature descriptor extractors, nonlinear correspondence estimators or odometry based approaches [6]. Although the data merging can be performed based only on the odometry information, this registration is prone to errors due to the error integrator characteristics of the odometers [7]. Hence a more robust method is applied for the initial alignment phase based on an extracted keypoint-feature data set [8]. A similar version of this approach was also adopted for the paper at hand.

In this paper several types of data from different sensors were tested for registration purposes with different key-points and feature descriptors. The main scope of the paper was to determine a suitable setup for the map registration using different type of data from various sensors and make a segmentation of the obtained maps.

In first part of the paper the hardware and software details regarding the 3D perception devices including stereo camera, structural light camera and custom 3D laser scanner are presented.

The next section deals with the pointcloud processing, including data filtering, keypoint-feature extraction, segmentation and registration issues. We conclude the paper with experimental results both from indoor and outdoor environments.

Common 3D Perception Techniques

There are more possibilities to acquire 3D information from the surrounding environment. The measurement methods can be divided into 3 major categories based on applied sensor and sensing technology: stereo vision (with two or more cameras) [9], active triangulation [10] and time of flight measurements. One of the most precise time of flight measurement systems is based on the laser scanners; however it is the most expensive one. A cheaper variant is the stereo camera, with less precision in the depth estimations [11].

Stereo Imaging. Using two or more cameras in the same scene can be used to estimate the depth information from the scene. The third coordinate information can be achieved by comparing at least two overlapping images from the same scene, the differences at pixel level between the two images yield to the depth information.

This method is known in the literature as stereo image processing, and there are several variants for the depth estimation including the popular sum of differences or the sum of absolute differences methods [12].

A schematic view about a stereo camera and its main parameters are presented on Fig. 1, where C_1 and C_2 are the coordinates of the left and right cameras, f is the focal distance, B is the distance between the cameras, Z is 3rd coordinate along which the depth is estimated, and finally x_1 , x_2 stores the projection of the points from the world coordinate frame on the camera x axes. The depth information can be extracted from the scene by with the formula

$$Z = \frac{f \times B}{d} \quad (1)$$

where the difference of the regions in pixels at the two images is stored in d [13].

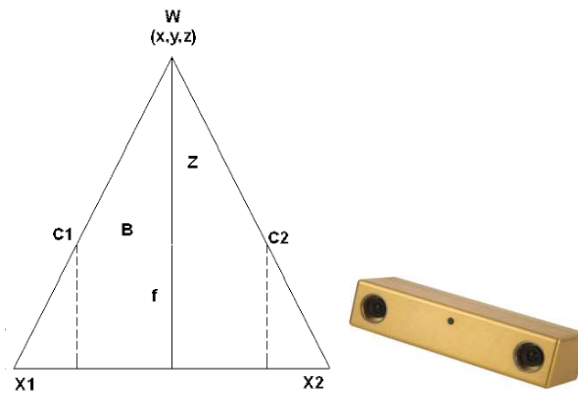


Fig. 1 Stereo geometry and a commercial stereo camera used for the data acquisition

In order to enhance the robustness of the depth estimation, the distances of the corresponding pixels at the two camera images are replaced by corresponding windows with sizes usually ranging from 1...64. For these regions can be successfully applied the same sum of absolute differences (SAD) algorithm, which will return a more robust depth estimation. This can be expressed with the following double sum:

$$\min_{d=d_{\min}}^{d_{\max}} \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{m}{2}}^{\frac{m}{2}} |R(x+i, y+j) - L(x+i+d, y+j)| \quad (2)$$

where d_{\min} and d_{\max} represent the range on which the differences are considered to be valid, m is the window size which is used to compare the regions on R and L image pair. After this step, the depth information Z can be expressed the formulae (1).

In such a way at relative high sample rates 3D colored information can be obtained about the surrounding environment in a non-contact type measurement. For an indoor scenario, an example image with RGB-D is given in Fig. 2.

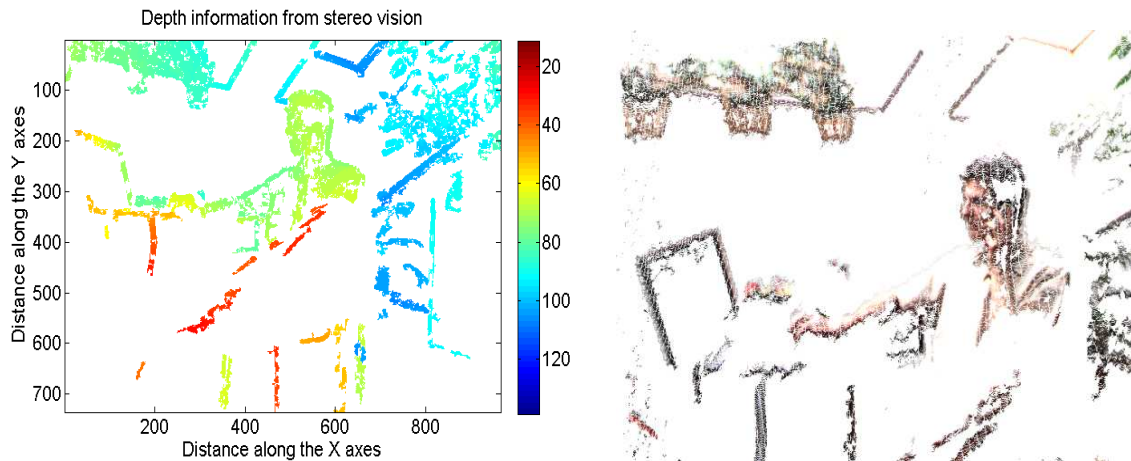


Fig. 2 Typical stereo camera image in indoor environment using the SAD algorithm

Projected Structural Light Devices The projected structural light cameras such as the low cost Kinect RGB-D camera captures infrared and RGB data from the same scenario. It has also an infrared emitter, which is used for structural light projection on the scene as this is described by the inventors. The projected pattern is captured by the infrared camera and is spatial correlated with the reference one.

The effective distance measurement is done on the principle of computing the distance between the reference points from the emitter and the ones received at the infrared sensor via a triangulation using the known baseline distance between the emitter and the receiver [14].

An indoor scan example using the Microsoft Kinect camera is presented in Fig. 3. At the left hand side the fused RGB-D information is shown, while on the right hand side the extracted depth information from the IR light is shown. As it can be seen, there are rather large spatial discontinuities in this type of data, although the frame rate at which the data feed can be capture is rather high compared to the stereo camera or the 3D laser.

Even though this type of camera can be used only in indoor environment, the frame rate of this device is rather high, comparable with the frame rate of a stereo imaging device. Usually the RGB and depth image fusion is performed after the data acquisition, and needs rather high computational effort.

Further on the depth image and the RGB data can be fused by using a calibration based on the fixed parameters of the camera. The main error sources for this device are related to the lighting conditions of the measured scene, as for strong lighting (e.g. sun) the projected infrared pattern have low contrast. This is also the main reason for the indoor usage limitation of this type of sensor. An alternative for the outdoor 3D perception on longer ranges is the laser range finder.



Fig. 3 Typical structural light RGB-D data and the IR image associated with the same scan from Kinect sensor

Lidar Based Techniques The use of laser scanners in the spatial perception is becoming popular in several applications starting from autonomous vehicles to urban environment mapping [15].

In the case of the custom developed 3D scanner, the main part is a two dimensional planar laser range finder which was mounted on a custom mechanical actuator ensuring an additional degree of freedom, thus yielding to a complete 3D sensing device [16]. There are several possibilities for rotating the planar laser around the axes. In the proposed custom setup a the rotation along the pitch axes was considered, as this is the most common for laser ranges mounted on autonomous vehicles. The complete setup is visible on Fig. 7.

The actuator was driven with a servo motor with a sufficient high torque and an auxiliary rotation encoder was mounted in order to measure precisely the rotation of the planar laser around the pitching axes. These components as well as the embedded controller used for interfacing with the PC are low cost commercial products.

The spatial information about the environment can be recovered by measuring a distance ρ and the associated pitch and yaw angles. Thus the direct kinematic transformation from the laser scanner coordinate frame to the world is given by:

$$p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{pmatrix} \times \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \\ 0 \end{pmatrix}. \quad (3)$$

In the Eq. (3) it was explicitly assumed that the two rotation axes have the same common origin point, i.e. there is no translation between the two rotations like transformations. Other sources of error are discussed in more details in [7].

A typical indoor scan in an office environment is presented on Fig. 4 where the color encoding is showing the high of the laser scan line in the room.

As it can be seen on the Fig. 4 the field of view of the laser is 180° horizontally and adjustable vertically according to the application. Also the pointcloud density varies from the edges to the center of the image, as the projection of the points from the spherical coordinats into the Euler ones is a nonlinear transformation.

In case that the target application requires that the density of the measured pointcloud to be distributed along another ax, the whole scanner can be rotated and thus obtained a different point density in the 3D space.

Also the requiremenets of the application impose the sensor type used for the perception task based on the advantages/disadvantages of these ones in terms of speed, accuracy and robustness.

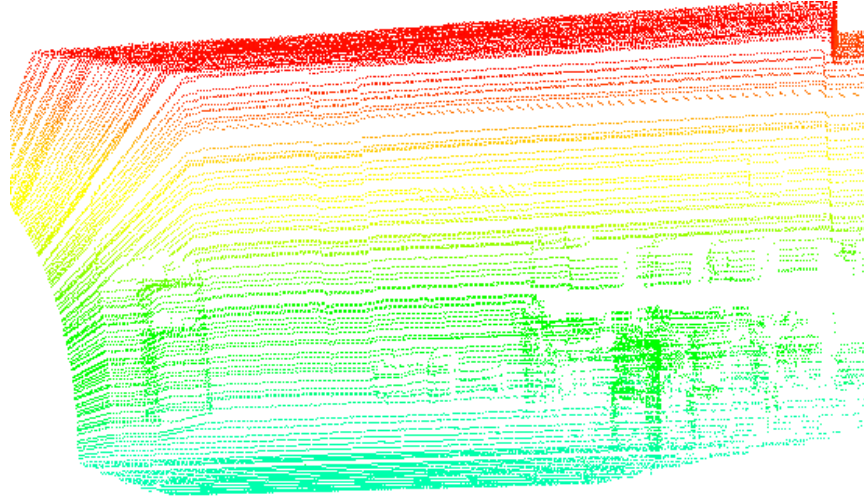


Fig. 4 A typical indoor 3D laser scan with the color handler computed as a function of the height

3D Data Processing Algorithms

In this section are summarized the algorithms used for 3D data processing, including the filtering, segmentation and registration ones.

Data Preprocessing. This part includes mainly pass-through and voxel grid filtering [13] and data structure ordering [17]. In the proposed experiment design, these filtering techniques were adopted due to their robustness and low computational complexity.

Scan Segmentation. The main role of the segmentation is to obtain regions that have some common features, e.g. they lay in the same plane, as their normal have a similar value [17].

In the proposed algorithm (1), the region based segmentation approach [13] is presented. This is based on a set composition idea, i.e. an initial starting point is selected from the original pointcloud P , and iteratively the set is completed with neighbor points which have similar normal N (within a certain threshold limit c_θ), or their curvature is less than a specified curvature threshold c_t . The resulting segmented union of sets Ω contains the region of common points.

Algorithm 1 Region Based Segmentation

Input: $P, N, c, \Omega(), c_t, \theta_t$

```

1:  $R \leftarrow \emptyset$ ;
2:  $A \leftarrow \{1, \dots, |P|\}$ ;
3: while  $\text{do}\{A\}$  is not empty
4:    $\{R_c\} \leftarrow \emptyset$ ;
5:    $\{S_c\} \leftarrow \emptyset$ ;
6:   for  $i = 0 \rightarrow \text{size}(S_c)$  do
7:      $P_j \leftarrow B_c\{j\}$ ;
8:     if  $(\cos^{-1} \prec N_{S_c(i)}, N_{S_c(j)} \succ \leq \theta_t)$  AND  $(c_{P(j)} \leq c_t)$  then
9:        $R_c \leftarrow R_c \cup P_j$ ;
10:       $A \leftarrow A \setminus P_j$ ;
11:       $S_c \leftarrow S_c \cup P_j$ ;
12:    end if
13:  end for
14: end while
15:  $R \leftarrow R \cup R_c$ 

```

Output: R

Algorithm 1 – Region based segmentation for pointclouds with depth information

An example of the algorithm output for an indoor dataset is presented in Fig. 5. As it is visible, the main plains (floor, ceiling, walls) are in a separate segment. The furniture in the room are placed in a common segment.

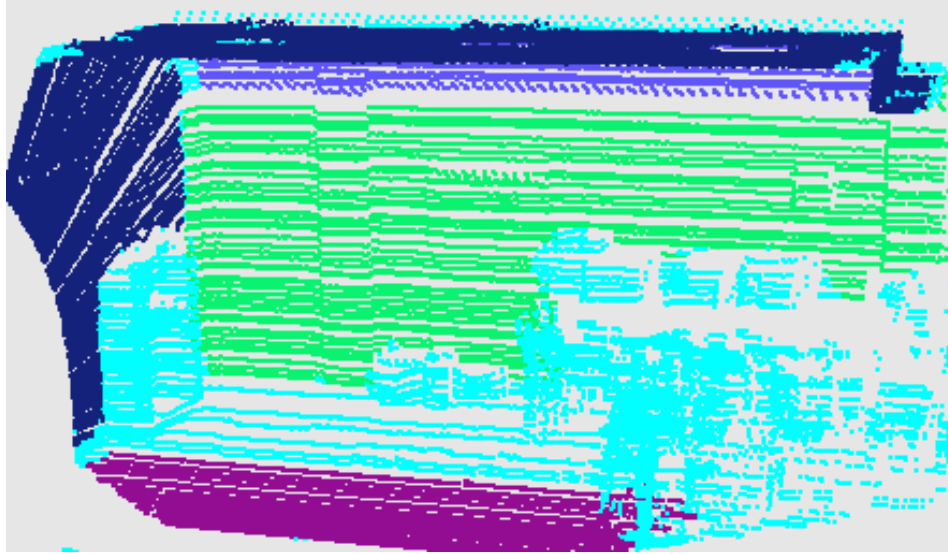


Fig. 5 A region based segmentation output for a typical indoor scan with furniture using Alg. 1

This kind of segmentation technique can be further more refined, by reducing the noise in the corners of the regions, which often are misclassified. Also the dimensions for which the objects are considered in separated regions can be adaptively selected according to the scanned scene's property.

Iterative Registration Variants. The pointcloud merging problem can be solved as with a minimization algorithm that computes the distance between the two scans [18].

In the proposed method the Iterative Closest Points (ICP) algorithm was used for the scan registration merging [19].

This solves the optimization problem by computing the difference of the template data set and the transformed data in the space [20]. This transformation is decomposed in two consecutive motions, a rotation \mathbf{R} and a translation \mathbf{t} .

During the iterative scan for similar point pairs in the M in the dataset D is performed in a trial-error manner. In case of a valid transformation \mathbf{R} and \mathbf{t} forms a valid transformation pair which minimizes the distance given by the following formulae:

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N_m} \sum_{j=1}^{N_d} w_{i,j} \left\| m_i - (\mathbf{R}d_j + \mathbf{t}) \right\|^2. \quad (3)$$

where $w_{i,j}$ is assigned 1 if a valid correspondence is found between the i^{th} point from M denoted with m_i and the j^{th} point from D denoted with d_j . Further details regarding this type of registration can be found in [21].

Keypoint-feature Based Scan merging. A common approach for boosting the ICP robustness is the augmentation of the points with additional features like point color, geometric features or point histograms [22]. This transposes the optimization problem in a higher order dimensional space search.

Our approach for the data registration is based on the correspondence estimation for the extracted keypoint features. In this paper the Normal Aligned Radial Feature (NARF) [23] keypoints were adopted for the extraction of interest points from range images. This type of keypoint takes into account the information about the borders and surfaces, ensures the detection from different perspectives and the stability for the descriptor computation [21].

One of the key components of the normal aligned feature is the disk size, i.e. the size of the zone in which the keypoint is computed [23]. For the proposed method different tunings were tested and the most suitable around 15cm was considered to be implemented in the registration.

The pseudo code of our approach is posted in the Algorithm 2, which describes briefly the main steps of the data merging for laser scans.

Algorithm 1 ICP with initial alignment

Require: P_s, P_t

- 1: $F_s = \text{ComputeFPFH}(P_s)$;
 - 2: $F_t = \text{ComputeFPFH}(P_t)$;
 - 3: $(t^*, A_f) = \text{InitialAlignment}(F_s, F_t)$;
 - 4: **while** ($\text{error}_{diff} < \epsilon$) **or** ($\text{iter} < \text{iter}_{max}$) **do**
 - 5: $A_d = \text{getClosestPoints}(t^*, P_s, P_t)$;
 - 6: $t^* = \text{argmin} \left(\frac{1}{|A_d|} \sum_{j \in A_d} w_j |t(p_s) - p_t|^2 \right)$;
 - 7: **end while**
 - 8: **return** t^*
-

Algorithm 2 The ICP based registration algorithm

At the beginning of the algorithm the feature are computed and the initial alignment guess is perfumed based on their relative transformation. Further on, this is refined with the ICP matching.

Experimental Methodology

The aligned map for the indoor environment is presented in Fig. 6, while results for the outdoor registration results are shown in Fig. 7. In both cases the registration was performed using a pair alignment approach and the FPFH descriptors for the computed NARF keypoints. The initial alignment of the scans was performed based on the filtered correspondences of the FPFH descriptors. This alignment was then used for the ICP refinement, computed on the last pair of data in the alignment loop.

For both scenarios the error convergence of the ICP algorithm was monotonically decreasing, a suitable registration error was achieved in less than 100 iterations. This scenario was obtained by considering the maximum distance between two neighbor points to be less than 1m.

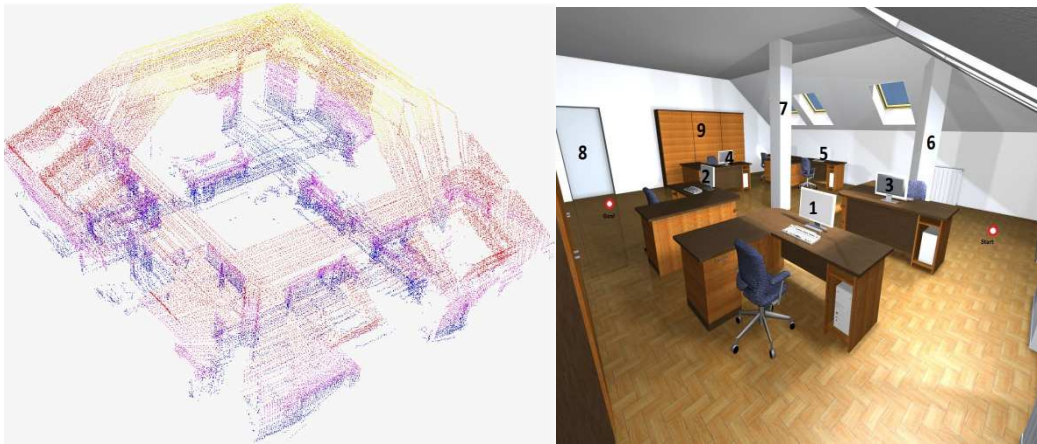


Fig. 6 A typical indoor registration result and its virtual CAD design

For the indoor scenario presented in Fig. 6, the total map was obtained from four consecutively merged parts. Further on, from this kind of 3D representation, a CAD model for the room was built easily on scale, presented on the right hand side of the figure.



Fig. 7 A typical outdoor scan registration results; the photo of the same scenario; the custom 3D laser scanner mounted on the P3 type robot

The outdoor map was registered on a total length of 40m path. The main aim at this part was to segment the merged map in such a way, that the main traversable path for the robot to be obtained. As it can be seen on Fig. 7, the output of the segmentation Algorithm (1), successfully discriminated between the main land plain and the obstacles on the way like the trash-been, or larger rocks. In our experiment we used the P3 mobile robot from ActiveMedia, on which the custom 3D laser scanner was mounted.

Summary

In this paper there were presented non-contact building 3D perception techniques using stereo-imaging, structural light cameras or tilted laser scanners. For the data acquired from these devices the preprocessing phase including filtering and segmentation is essential. The different measurements were merged in a common coordinate frame during the registration phase. For this different registration approaches were used including the ICP based and the keypoint-feature based ones. The results of these algorithms are showed both for indoor and outdoor data sets.

Further on we plan to validate these investigations in real life applications including archeological and architectural reconnaissance applications.

Acknowledgement

This paper was supported by the project "Development and support of multidisciplinary postdoctoral programs in major technical areas of national strategy of Research - Development - Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectorial Operational Program Human Resources Development 2007-2013.

References

- [1] F. Maurelli, D. Droschel, T. Wisspeintner and H. M. S., "A 3D Laser Scanner System for Autonomous Vehicle Navigation," Proceedings of the International Conference on Advanced Robotics (ICAR), pp 1-6, 2009.
- [2] J. Surmann, "A 3D Laser Range Finder for Autonomous Mobile Robots," in Proceedings of the International Symposium on Robotics, Zurich, 2001.
- [3] J. W. Weingarten, G. Gruener and R. Siegwart, "A State-of-the-Art 3D Sensor for Robot Navigation," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2004, pp 2155-2160.
- [4] O. & W. B. Wulf, "Fast 3D-Scanning Methods for Laser Measurement Systems," in International Conference on Control Systems and Computer Science, Bucharest, 2005.

- [5] D. Chetverikov and D. a. P. K. Stepanov, "Robust euclidean alignment of 3D point sets: the trimmed iterative closest point algorithm," *Image and Vision Computing*, 2005.
- [6] M. Magnusson, T. Duckett and A. J. Lilienthal, "Scan Registration for Autonomous Mining Vehicles Using 3D-NDT," *Journal of Field Robotics*, vol. 24, pp. 803-827, 2007.
- [7] A. Howard, D. F. Wolf and G. S. Sukhatme, "Towards 3D Mapping in Large Urban Environments," *International Conference on Intelligent Robots and Systems (IROS 2004*, pp 419-426.
- [8] A. Zhang, S. Hu, Y. Chen, H. Liu, F. Yang and J. Liu, "Fast Continuous 360 Degree Color 3D Laser Scanner," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* vol. 36, pp. 409-415, 2008.
- [9] B. Morisset, R. B. Rusu, A. Sundaresan, K. Hauser, M. Agrawal, J.-C. Latombe and M. Beetz, "Leaving Flatland: Toward Real-Time 3D Navigation," *IEEE International Conference on Robotics and Automation (ICRA)*, pp3786-3792, 2009.
- [10] K. Ohno, S. Tadokoro, K. Nagatani, E. Koyanagi and T. Yoshida, "Trials of 3-D Map Construction Using the Tele-operated Tracked Vehicle Kenaf at Disaster City," *International conference on Intelligent robots and systems (IROS)*, 2010.
- [11] R. Kaushik, J. Xiao, W. Morris and Z. Zhu, "3D laser scan registration of dual-robot system using vision," *International conference on Intelligent robots and systems (IROS)*, 2009.
- [12] A. Birk, N. Vaskevicius, K. Pathak, S. Schwertfeger, J. Poppinga and H. Buelow, "3D Perception and Modeling: Motion Level Teleoperation and Intelligent Autonomous Functions," *IEEE Robotics and Automation Magazine (RAM)*, vol. 6, pp. 53-60, 2009.
- [13] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," *International Conference on Robotics and Automation (ICRA)* pp1-4, 2011.
- [14] P. Henry, M. Krainin, E. Herbst, X. Ren and D. Fox, "RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments," *Advanced Reasoning with Depth Cameras Workshop in conjunction with the RSS conference*, 2010.
- [15] N. Vandapel, D. F. Huber, A. Kapuria and M. Hebert, "Natural Terrain Classification Using 3-D Ladar Data", *IEEE International Conference on Robotics and Automation (ICRA)* pp 5117-522, 2004.
- [16] Henriksen, L. & Krotkov, E. Natural Terrain Hazard Detection with a Laser Rangefinder *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1997, 968-973 .
- [17] J. J. Leger and P. C. Biesiadecki, "Tradeoffs Between Directed and Autonomous Driving on the Mars Exploration Rovers," *The International Journal of Robotics Research*, vol. 26, pp. 91-104, 2007.
- [18] H. Surmann, K. Lingemann, A. Nuchter and J. Hertzberg, "A 3D Laser Range Finder for Autonomous Mobile Robots," *Robotics and Autonomous Systems*, pp 181-192, 2001.
- [19] A. Nuechter, *3D Robotic Mapping*, Springer, p. 210, 2009.
- [20] Z. Zhang, *Iterative Point Matching for Registration of Free-Form Curves*, Technical Report No. RR-1658, 1992.
- [21] L. Tamas and L. C. Goron, "3D Map Building with Mobile Robots," in *IEEE MED Conference*, pp 615-621, 2012.
- [22] G. C. Sharp, S. W. Lee and D. K. Wehe, "ICP Registration Using Invariant Features," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 24, pp. 90-102, 2002.
- [23] B. Steder, R. B. Rusu and K. Konolige, "Point Feature Extraction on 3D Range Scans Taking into Account Object Boundaries," *IROS IEEE Conference*, 2011.