Real-time 3D Mobile Mapping for the Built Environment

Jingdao Chen[a], Yong K. Cho[b]

[a] School of Electrical and Computer Engineering, Robotics Institute, Georgia Institute of Technology
[b] School of Civil and Environmental Engineering, Georgia Institute of Technology
E-mail: jchen490@gatech.edu, yong.cho@ce.gatech.edu

Abstract –

Laser scanned point clouds are relevant for geometric analysis, building maintenance and energy modeling, yet obtaining fully registered point clouds is a labor-intensive process when targets or common registration features are absent. Thus, we propose a versatile mobile platform that forms an incremental 3D map of the environment in real time using an orthogonal pair of Lidar (Light Detection and Ranging) devices. The horizontal scanner aims to estimate the robot position and orientation with SLAM (Simultaneous Localization and Mapping) techniques whereas the vertical scanner recovers the building structure in the vertical plane. We also developed a real time point cloud visualization tool that allows an operator to track the mapping progress. The method was evaluated with walk-through laser scans of a complete building floor.

Keywords

3D, Lidar, point cloud, SLAM, real-time

1 Introduction

Three-dimensional models of building components in the as-built condition are important tools in building maintenance and restoration applications. The acquisition of data for creating building models has traditionally been done with tools such as measuring tapes and distometers [1]. However, advances in recent technology has led to use of a host of techniques such as robotic total station, stereo photogrammetry, depth cameras, Light Detection and Ranging (Lidar), as well as hybrid setups to survey and map various aspects of building infrastructure. The acquired data in the form of visual or range data can be incorporated into a single three-dimensional point cloud representation. The generated point cloud can be further post-processed to create a Building Information Model (BIM) [2], which describe relational characteristics and attributes of building elements. The generated model can also be used to monitor defects and estimate energy usage characteristics [3].

Static scanning solutions such as fixed camera setups and ground-based Lidar are widely used in the data collection step of infrastructure modelling. However, they suffer from the limitation that the registration step to combine data collected from multiple scan times is labor-intensive. For example, images taken in multiple locations for photogrammetry have to be combined in a feature point matching process. For the case of ground-based Lidar, multiple scans have to be registered together in a post-processing step using either Global Positioning System (GPS) information or iterative optimization techniques. These data collection methods offer limited feedback to the operator in terms of the final registered output during the scan process, so the generated data is often vulnerable to problems such as incomplete scans or occluding structures.

Thus, in this study we propose a mobile scanning solution that generates a registered three-dimensional point cloud output in real-time. The proposed method involves two Lidar devices mounted in an orthogonal position on a mobile platform. The two devices work concurrently in a Simultaneous Localization and Mapping (SLAM) setting to estimate the current position and map of the environment so that each scan can be automatically registered. The combined point cloud is presented to the operator through a visualization tool to determine scan progress. The proposed method aims to achieve a more efficient and accurate scanning process compared to conventional scan techniques.

2 Literature Review

Most instances of 3D mapping have been achieved by using photogrammetry, laser scanners or even a combination of both methods with some also using additional technologies such as sonar and infra-red (IR) [4]. Since most research has been focused on the photogrammetric and Lidar-based SLAM, we discuss some common practices of each approach and how close they are to solving the SLAM problem. We also discuss the pros and cons of each approach.

2.1 Application of SLAM

Simultaneous Localization and Mapping (SLAM) [5] is a significant problem in mobile robotics where map building and localization is carried out without a priori information. SLAM algorithms collect environment features and learn a mapping between the robot location and the environment. Thus, SLAM algorithms provide a useful tool to automatically obtain the scan origin and register sensor data onto a global map.

2.1. Photogrammetric Approach

In [6], Davison et al. were able to recover the 3D trajectory of a monocular camera moving through a previously unknown scene. Here the nature of the map depends on the initial and current positions of the camera, 3D position of the center of the feature and the orientation of its local surface. This algorithm is unable to handle sudden jerky movements and while it uses modest processing power, it has limited application in getting a detailed map of the environment. In [7], sixteen advantages are provided on the advantages of the photogrammetric 3D workflow over the directly measured laser point clouds with significant points being the inexpensive hardware available for vision based techniques and the density of surface points being higher but these require multiple cameras and hardware to run expensive algorithms to achieve real time SLAM. In [8], a Kinect sensor is used to generate a 3D model of a building and the result is compared to another model generated from a laser scanner. While the paper does make a case for detecting discontinuity points using a vision based system, the authors concluded that the measurements are highly influenced by the material and lighting conditions. As a consequence, the generated point cloud is noisy and not uniform. Khoshelham et al. [9] expounded on this in detail in their paper where they note that the random error of depth measurements increases quadratically with increasing distance from the sensor. Their technique also required accurate stereo calibration of the IR camera and visual camera.

Thus we can summarize the advantages of using visual based SLAM as being relatively less expensive, lighter and consuming less power per unit instrument than using laser scanners. Visual based systems are also able to detect colors and other features/objects which cannot be detected by Lidar in a cluttered environment. These approaches also have their disadvantages: 1) They can accumulate large errors if there are large distances, 2) They are very susceptible to even slight changes to lighting conditions, 3) They can get confused in a dynamic environment and 4) They are difficult to use at night [10].

2.2 Lidar Based Approach

There have been several attempts in the literature to map an indoor environment with laser scanners. Shohet and Rosenfeld [11] talk about two main factors affecting the precision of the map: 1) orientation of the carriage on which the scanner is mounted and 2) distance between the sensor and the walls being scanned. However advancements in Lidar technology have rendered the last point irrelevant with commercial Lidars now available with a range of 120 m [12]. Current research is focused on improving the accuracy of maps generated from laser scanners. Jung et al [13] proposed a kinematic 3D laser scanning system for indoor mapping which imposes some constraints on the environment where main structures are formed from straight lines and all structures are parallel or perpendicular. Another approach to solving the SLAM problem is having multiple robots on the field and then integrating the pose estimations of the robots and individual maps to create a more accurate model of the environment [14]. This approach, while successful, does not allow the operator to have a real-time view of the map as the off-line optimization step may require the robots to stop exploration and perform the optimization process before they resume the exploration task.

Kohlbrecher and von Stryk [15] proposed a 2D SLAM system known as Hector SLAM which performs laser scan matching to obtain a pose estimate and planar map of the environment. The SLAM system is flexible, scalable, and has been deployed on unmanned ground vehicles (UGV), unmanned surface vehicles (USV), and small indoor navigation systems. LOAM (Lidar Odometry and Mapping in Real time) [16] is another approach for real-time SLAM but requires a spinning Lidar which is more expensive to achieve than a planar Lidar approach like Hector SLAM. There have been several studies which built on Hector SLAM to produce more accurate maps of indoor environments using external sensors [17-20]. Khan et al [20] proposed a data driven method, built on Hector SLAM, which models laser intensities and measures the surface reflectivity to augment a geometric model of the surrounding environment.

In this study, the method of [15] was built upon to obtain full 3D Lidar mapping. The goal is to achieve a more mapping-focused SLAM compared to navigation [19] or exploration [17]. This study also utilizes less sensory requirements compared to [13] and [16].

3 Methodology

The hardware setup of our experiments consists of 2 Lidar devices as shown in Figure 1. The larger device is a SICK Lidar LMS511 which scans at a 0.5° angular resolution and 50 Hz sample rate. The smaller device is a SICK Lidar LMS151 which scans at a 0.25° angular resolution and 25 Hz sample rate. Each Lidar device only scans along a single plane so it is necessary to stack the Lidar devices in order to capture full 3D information. The Lidar devices are mounted orthogonally to each other on the mobile platform where the larger Lidar is used for scanning in the horizontal plane and the smaller Lidar is used for scanning in the vertical plane. The Lidar devices are connected to an NVIDIA Jetson TK1 processor running Ubuntu Linux. The Robot Operating System (ROS) [21] software suite is used to carry out sensor data input and output, coordinate timing, and perform frame transformation between the two Lidar devices. Point cloud processing modules from the Point Cloud Library (PCL) [22] are also used for point cloud filtering, storage and visualization.

Figure 1: Lidar devices used in the hardware setup

The following sections provide details on each step of the scan process. Each step is processed continuously to generate a real time scan output in the form of a point cloud.

3.1 Lidar Data Collection

The data collection process involves walk-through scans with the mobile platform equipped with two Lidar devices. The two sites utilized for this study are a laboratory room and corridors along a selected building floor since they contain features of interest to qualitatively evaluate the scan output. A laboratory room setting is challenging because it is relatively unstructured and contains occlusion in the form of furniture whereas mapping a series of corridors presents a challenge in terms of scanning long distances and mapping around corners.

3.2 Motion Estimation

The Lidar data from the horizontal device is processed with a SLAM algorithm to estimate the motion of the mobile platform on the horizontal plane. The SLAM algorithm used in this study is Hector SLAM [15], which performs scan matching between the current Lidar input and an incrementally built map. The SLAM algorithm estimates the optimal Lidar pose including translation and rotation to match the current scan with a map of the environment. The SLAM algorithm also outputs an occupancy grid map which indicates the grid locations that are likely to be obstacles and the grid locations that are unobstructed. An example of the generated planar map is shown in Figure 2 for the building corridor experiment. The map shows a series of grid lines representing the areas scanned by the horizontal Lidar and a single thick curve representing the trajectory of the mobile platform. Light areas of the map correspond to open space that may be traversed by the mobile platform whereas dark areas correspond to building structures. The trajectory is updated for each input scan from the horizontal Lidar with the calculated current position and orientation of the mobile platform.

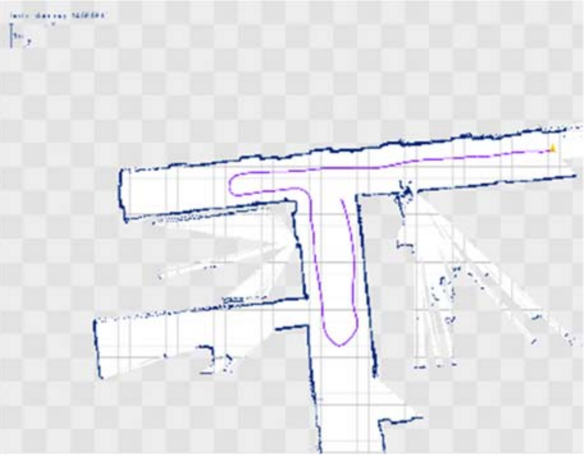

Figure 2: Two dimensional map on the horizontal plane with trajectory estimate of the mobile platform. Blue grid lines indicate scanned areas while the purple curve indicates the trajectory of the mobile platform.

3.3 Scan Registration

Based on information gathered about the mobile platform position from the previous step, the vertical scan input from the second Lidar device can then be registered onto a global three-dimensional map. For each vertical scan collected, the corresponding timestamp is obtained from the sensor and matched with the nearest timestamp from the horizontal sensor. Linear interpolation is then used to obtain the 2D position of the mobile platform at the time when the vertical scan is collected. In terms of position, the current estimate is obtained by fitting a linear motion trajectory through the preceding position estimates. In terms of rotation, the current estimate is obtained through Spherical Linear Interpolation (SLERP) [23] of the neighboring rotation quaternions. The vertical scan is added to a global map by horizontal translation with the calculated 2D vector. A three-dimensional map is thus incrementally built up through the stacking of multiple planar scans.

3.4 Scan Visualization

We also develop a visualization tool to aid the operator of the mobile platform in tracking the scan progress. Figure 3 shows an example of a scan in progress rendered using OpenGL. The visualization displays a set of larger points representing the current horizontal scan and a set of smaller points representing the registered vertical scans in 3D world coordinates. The visualization tool enables the operator to identify unscanned areas and detect potential occlusions during the scan process. This ensures a higher quality and more complete point cloud in the final output.

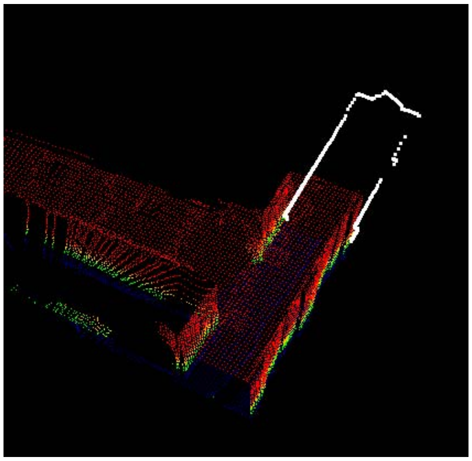

Figure 3: Mapping progress shown in user display. Bold white points indicate horizontal scans while small colored points indicate vertical scans.

4 Results

The proposed technique was applied to two separate scenarios and the scan results in the form of 3D point clouds are shown below. The first experiment was carried out in an indoor laboratory room as shown in Figure 4. The ceiling component of the point cloud was filtered out to allow a clearer view of the room contents. The scan trajectory can also be observed in the middle section of the point cloud. The scan result demonstrates the ability of our technique to capture shape profiles from furniture such as desks and cabinets even in a relatively cluttered environment.

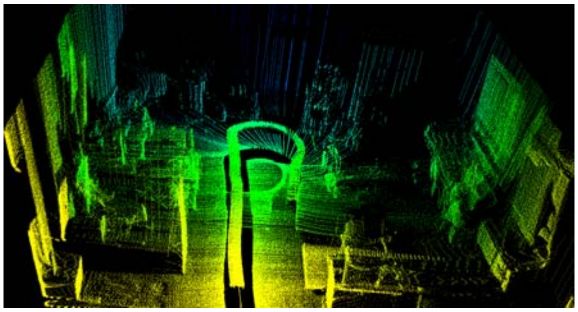

Figure 4: Indoor lab mapping result

On the other hand, the second experiment was carried out over an entire building floor as shown in Figure 5. The result consists of scans along multiple corridors on the same floor. The scan result is largely extensive with a few missing areas corresponding to areas not accessible by the mobile platform. This demonstrates the ability of our technique to capture scans in a large area and register them in a single map.

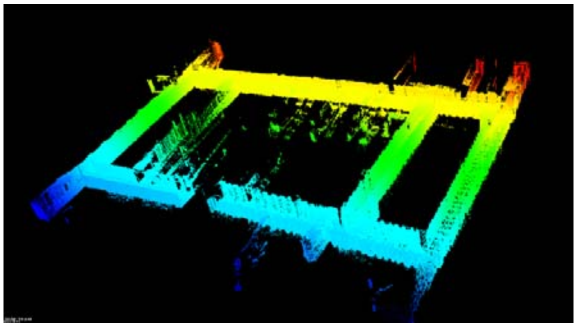

Figure 5: Corridor mapping result

The accuracy of the generated point cloud map was validated against a commercially available laser scanning system [25]. The static Lidar was placed in 4 different scan points along the corridor shown in Figure 5 and the fully-registered point cloud was used as a reference. The reference point cloud is registered using manual coarse registration and Iterative Closest Point (ICP) fine tuning. Since the scans were taken in an indoor environment, the accuracy metric used is the Euclidean Root Mean Squared Error (RMSE) of point distances to three selected best-fit planes. The calculated Euclidean distance error of the point cloud data is shown in Table 1. The estimated scanning time involved is shown in Table 2. The results indicate that the proposed method achieves reasonable accuracy in the generated point cloud data but involves much faster scanning time.

Table 1: Euclidean distance error of point cloud

Plane	RMSE (single Lidar) (m)	RMSE (our method) (m)
1	0.03511	0.03534
2	0.03948	0.03883
3	0.04233	0.04215

Table 2: Scanning time comparison between different methods

Method	Scanning Time (minutes)
Single Lidar	30
2 Lidar mobile platform (our method)	5

5 Discussion

This section discusses the strengths and weaknesses of the proposed technique for generating a three-dimensional point cloud map of the built environment along with the associated challenges. From the scan results obtained from two test scenarios, we found that the hardware and software setup is capable of performing scan localization and mapping in real time with an update rate of 25Hz. The visualization tool is also effective in providing the operator with an informative view of the scan progress.

The proposed technique is able to generate detailed point clouds due to the high angular resolution of the Lidar devices. However, our methodology heavily relies on accurate localization of the mobile platform at each scan time. The sensor pose estimate may occasionally be erroneous due to jerky motion of the mobile platform which causes a mismatch in the scan registration process. One strategy is to employ a median filter to smoothen the estimated trajectory by eliminating outliers. We can also enforce geometric consistency conditions among vertical scans in a local region such as planarity constraints along the floor and walls. Another limitation of the proposed technique is non-uniform point cloud resolution in the generated 3D map. The point cloud tends to be denser for regions where the mobile platform is moving fast and smoothly and sparser for regions where the mobile platform is moving quickly or undergoing a rapid rotation. To overcome this problem, the operator can reduce the velocity when moving around corners and obstacles while obtaining feedback from the visualization tool to ensure a high resolution in the point cloud output.

Table 3 shows a comparison between our proposed mapping technique with alternative techniques. For example, generating a 3D map using pure visual input involves a less expensive hardware setup and is able to output color information. However, visual scene reconstruction is prone to scale inaccuracies because the process of translating between pixel values and metric units is complex and rely on accurate camera calibration. A depth camera, such as the Microsoft Kinect [24], is able to output both color and depth information but operates at lower ranges compared to a Lidar device. A single ground-based Lidar, such as the FARO Focus laser scanner [25], is capable of generating point clouds at a high resolution, but suffers from the limitation of requiring an additional step to register scans taken from multiple locations. Our proposed technique involving two Lidar devices has the ability to perform automatic registration of scan outputs in real time, but involves an expensive hardware setup.

Table 3: Comparison between scanning techniques

Technique	Advantages	Disadvantages
Visual Camera	Inexpensive hardware, has color information	Sensitive to lighting, vulnerable to scale inaccuracies
Depth camera	Has color and depth information	Limited range
Single Lidar	High resolution output	Require additional registration step
2 Lidar mobile platform (our method)	Real time, automatic registration	Expensive hardware

6 Conclusion

In conclusion, the results of this study demonstrate the viability of the proposed technique for creating three-dimensional maps of the built environment in real time using an orthogonal pair of Lidar devices mounted on a mobile platform. The proposed technique leverages state-of-the-art SLAM algorithms and scan registration to generate dense and accurate point cloud outputs. The main advantages to the proposed technique are the high update rate, high output resolution and effective progress visualization. The proposed technique has the potential to be utilized in infrastructure surveying applications such as geometric analysis, building maintenance and energy modeling. For future work, we hope to improve the accuracy of scan registration on the global three-dimensional map. This can conceivably be achieved through better handling of sensor error and identifying a more precise motion model for the mobile platform. We also consider including color information in the point cloud output by adding a visual sensor, which could be useful in providing a better visualization for the point cloud output as well as helping to identify color features in the post-processing stage.

7 Acknowledgement

This material is based upon work supported by the National Science Foundation (Award #: CMMI-1358176). Any opinions, findings, and conclusions or recommendations expressed on this material are those of the authors and do not necessarily reflect the views of the NSF.

References

[1] C. Eastman, P. Teicholz, R. Sacks, K. Liston, "BIM Handbook – A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors", *John Wiley & Sons, Inc*. (2008)

[2] L. Barazzetti, F. Banfi, R. Brumana, G. Gusmeroli, M. Previtali, G. Schiantarelli, "Cloud-to-BIM-to-FEM: Structural simulation with accurate historic BIM from laser scans", *Simulation Modelling Practice and Theory*, Volume 57, September 2015, Pages 71-87,

[3] C. Wang, and Y. Cho. "Automated 3D Building Envelope Recognition in Point Clouds." *ASCE, Proceedings of Construction Research Congress* 2012, West Lafayette, IN, pp. 1155-1164

[4] Akshay Kumar Shastry, Sanjay Anand, V. Chaitra, K. Uma Rao, and D. R. Akshay, " SONAR Validation for SLAM," *International Journal of Innovation, Management and Technology* vol. 4, no. 5, pp. 498-501, 2013

[5] J. J. Leonard, J. J., and H. F. Durrant-Whyte. "Simultaneous map building and localization for an autonomous mobile robot." in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'91)*, pp. 1442–1447, New York, USA.

[6] A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052-1067, June 2007.

[7] F. Leber, A. Irschara, T. Pock, P. Meixner, M. Gruber, S. Scholz, A. Wiechert. *Photogrammetric Engineering and Remote Sensing.* Vol 76, No.10, pp 1123-1134, 2010

[8] D. Roca, S. Lagüela, L. Díaz-Vilariño, J. Armesto, P. Arias, "Low-cost aerial unit for outdoor inspection of building façades", *Automation in Construction*, Volume 36, December 2013, Pages 128-135, ISSN 0926-5805

[9] Khoshelham K, Elberink SO. "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications." *Sensors* (Basel, Switzerland).2012;12(2):1437-1454. doi:10.3390/s120201437..

[10] Fuentes-Pacheco, J., et al. "Visual simultaneous localization and mapping: a survey." *Artificial Intelligence Review* 43(1): 55-81

[11] Igal M. Shohet, Yehiel Rosenfeld, "Robotic mapping of building interior—precision analysis", *Automation in Construction*, Volume 7, Issue 1, December 1997, Pages 1-12, ISSN 0926-5805,

[12] Velodyne Lidar. http://velodynelidar.com/hdl-64e.html

[13] Jung J, Yoon S, Ju S, Heo J. "Development of Kinematic 3D Laser Scanning System for Indoor Mapping and As-Built BIM Using Constrained SLAM". Passaro VMN, ed. *Sensors (Basel, Switzerland)*. 2015;15(10):26430-26456.

[14] H. Jacky Chang, C. S. George Lee, Y. Charlie Hu and Yung-Hsiang Lu, "Multi-robot SLAM with topological/metric maps," *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, San Diego, CA, 2007,pp.1467-1472.

[15] S. Kohlbrecher, J. Meyer, O. von Stryk, U. Klingauf. "A flexible and scalable SLAM system with full 3D motion estimation." in Proc. of the IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR'2011), pp. 50–55. Kyoto (2011)

[16] J. Zhang and S. Singh. "LOAM: Lidar Odometry and Mapping in Real-time". *Robotics: Science and Systems Conference* (RSS). Berkeley, CA,

July 2014

[17] F. Hoeller, A. Konigs and D. Schulz, "Autonomous reconnaissance and surveillance in urban structures - Eurathlon 2013," *Autonomous Robot Systems and Competitions* (ICARSC), 2014 IEEE International Conference on, Espinho, 2014, pp. 223-228.

[18] Hoeller, F.; Konigs, A.; Schulz, D. "Autonomous reconnaissance and surveillance in urban structures - Eurathlon 2013", Autonomous Robot Systems and Competitions (ICARSC), 2014 IEEE International Conference on, On page(s): 223 – 228

[19] Schueftan, D.S.; Colorado, M.J.; Mondragon Bernal, I.F. "Indoor mapping using SLAM for applications in Flexible Manufacturing Systems", *Automatic Control* (CCAC), 2015 IEEE 2nd Colombian Conference, page(s): 1 – 6

[20] S. Khan, D. Wollherr and M. Buss, "Modeling Laser Intensities For Simultaneous Localization and Mapping," in *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 692-699, July 2016.

[21] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng, “ROS: An open-source robot operating system,” in *Proc. Open-Source Software Workshop Int. Conf. Robotics and Automation*, Kobe, Japan, 2009

[22] R. B. Rusu and S. Cousins, “3D is here : Point Cloud Library (PCL),” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[23] K. Shoemake, "Animating rotation with quaternion curves." in *Computer Graphics: Proceedings of SIGGRAPH* ’85, 245–254. San Francisco: ACM

[24] Microsoft Kinect. http://www.xbox.com/en-us/kinect/

[25] FARO Focus 3D X. http://www.faro.com/en-us/products/3d-surveying/faro-focus3d/overview