Handheld Simultaneous Localization and Mapping Devices for 3D Scanning

M. Froehlich [a] and S. Azhar [a]

[a] McWhorter School of Building Science, College of Architecture, Design and Construction, Auburn University
E-mail: maf0032@auburn.edu, salman@auburn.edu

Abstract – 3D Scanning has allowed for the scanning of spaces serving as a valuable tool for verifying existing conditions. The cost, setup and processing times associated with procuring this activity limits its use within an active construction site. With setup consisting of multiple scans from triangulated views of one another, achieving an accurate scan can be troublesome. Recent developments of Simultaneous Localization and Mapping (SLAM) technologies have enable indoor tracking of spaces through an array of sensors. The Google® Tango Tablet incorporates SLAM technology within a handheld device. Paired with the tablets depth sensors and software, 3D scanning can be achieved quickly and at substantially lower costs. This paper presents findings of a research study which investigated the 3D scanning applications of Handheld SLAM devices. This research helps us to explore an emerging technology that could allow for significant costs and time saving for the construction industry. This research will be a great aid in determining the value and best applications of handheld SLAM devices.

Keywords –
3D Scanning Technologies; IR Depth Sensors; SLAM;

1 Introduction

There are increasingly a large number of 3D scanners available for many different markets and applications covering, Art, Reverse Engineering, Automotive, Archaeology, Product Design, Architecture, Fashion Design, Body Measurement, BIM, Civil Engineering, Metrology, Forensics, and National Security. There are limitations, advantages and disadvantages to each scanner such as resolution, scan area, mobility, data capturing speed, tolerances [1].

Over the last twenty years, laser-scanning technology has seen a significant shift in applications across multiple disciplines [2]. The use of scanning is growing in this area thanks to steadily declining services costs and better informed clients. Although this area has historically not been quite as lucrative as the civil, infrastructure, industrial plant, and facilities applications, there are numerous successful projects and service providers in this space today [3].

In a case study by C. Clark and J. Liu conducted field testing for drywall installation using laser scanning and BIM. They determined that field scanning requirements for simple rooms such as offices are minimal requiring only one medium resolution scan with image capture for reference; a twenty-minute investment for each office. The accuracy proved that the integration of laser scanning and BIM technology can be used in a specific area without having to make any field measurements. The study additionally concluded that on most construction projects, the window for conducting scans is relatively short. Suggesting a faster process for scanning of an area to further be researched [2].

Becerik-Gerber, et. al. performed research on data acquisition errors caused by target setup, acquisition, and reorientation. They explored how different target types and target layouts affect registration accuracy[4]

Additionally, 3D laser scanning requires stillness in the scene that is captured. Thus, ongoing production activities can prove troublesome. If the geometries in the scene are geometrically complex and numerous, more data captures are required to get comprehensive spatial data of the entire scene [5].

The progression of low cost technology continues with the recent development of the Google® Tango Tablet. Combining Infrared (IR) depth sensors and cameras paired with Simultaneous Localization And Mapping (SLAM) technology allows for scanning on the jobsite quickly with minimal training and prep time by any construction personnel or trade.

The purpose of this study is to test the feasibility of a 3D scanning tablets and their applications for use in construction practice. The research will focus on prep time, scanning time, and accuracy of the scan.

2 Literature Review

Despite recent technological advances, developing a functional mobile AR system to be used in an unprepared environment is a considerable challenge. Solutions require high-accuracy data and precise positioning of the user and their relation to the hardware to ensure correct registration between reality and the data to be augmented [6]. Accurate registration and positioning of virtual objects in the real environment require accuracy in tracking the user's head as well as sensing the locations of real objects in the environment. The biggest single obstacle to building effective AR systems is the requirement of accurate, long-range sensors and trackers [7]. A common approach to indoor localization is to make use of WiFi or Bluetooth sources in the environment [8] [9] [10]. These techniques tend to be sensitive to the WiFi source placement, the accuracy of the signal strength map and the amount of clutter and dynamics in the environment. Additionally, the number of WiFi sources in the environment plays an important role [11]. Zollmann finds tracking currently consist of sensor-based registration techniques that incorporate inertial sensors, camera and Global Positioning System (GPS), and remote localization techniques that only make use of a camera, but depend on a 3-D point cloud [12].

A different approach, widely used in robotics, provides an alternative. So-called simultaneous localization and mapping (SLAM) methods build and update a map of an unknown environment while the same time keeping track of the tracking device's location within the environment (the learned map). SLAM methods can work with various sensors but often optical sensors like cameras, depth cameras or laser range scanners are used. Therefore, SLAM methods combined with mobile hardware, for sensing and mapping the 3D structure of an environment, allow for self-contained tracking of a user's viewpoint [13].

Determining a camera's position and orientation is a key requirement in augmented reality applications, a crucial step in simultaneous localization and mapping (SLAM) algorithms, and overall a major area of focus in computer vision [14][15][16][17]. Absolute pose methods utilize 2D-3D correspondences between image pixels and 3D points in a known scene to determine the full 6-degree-of freedom (d.o.f.) camera pose. These methods are often more efficient than relative pose methods and have been demonstrated to improve accuracy and reduce camera pose jitter in SLAM[18].

The Google® Tango tablet features a color camera and a motion and depth sensing camera. The device's fisheye camera a large field-of-view is observed in each im- age, which significantly simplifies and speeds up the capturing of larger scenes compared to standard lenses [19]. It includes software for SLAM which builds and/or updates a 3D map of the environment. The resulting map is essentially a mesh which can be stored for later use (i.e. later tracking in a known environment) or shared with other users [13].

2.1 Google® Tango Tablet

The recent developments in mobile technology allow advanced interfaces such as AR to be moved from the laboratory into the field. Their ability to present information on site, where it is needed, holds many advantages for professional application [12]. The "Project Tango®" prototype is an Android smartphone-like device which tracks the 3D motion of the device, and creates a 3D model of the environment around it. The Tango® Tablet was developed around, motion tracking, depth perception, and area learning. Project Tango® devices contain customized hardware and software designed to track the full 3D motion of the device, while simultaneously creating a map of the environment. These sensors allow the device to make over a quarter million 3D measurements every second, updating its position and orientation in real-time, combining that data into a single 3D model of the space around you [20].

Figure 1. Google® Tango Tablet

Table Google® Tango Tablet

Specification	
Screen	7.02" 1920x1200 HD IPS display (323 ppi) Scratch-resistant Corning glass
Dimensions	119.77 x 196.33 x 15.36mm
Weight	0.82 lbs (370g)
Cameras	4 MP 2mm RGB-IR pixel sensor 1 MP front facing, fixed focus
OS	Android 4.4 KitKat
Wireless	4G LTE

	Dual-band Wi-Fi (2.4GHz/5GHz) WiFi 802.11 a/b/g/n NFC(reader and peer to peer modes)
Audio Output	Dual stereo speakers 3.5mm audio connector (OMTP standard)
Memory	128 GB Storage 4 GB RAM
Ports	Micro HDMI USB 3.0 Micro SD card Nano SIM slot
Processor	NVIDIA® Tegra K1 w/ 192 CUDA cores
Sensors	Motion tracking camera 3D depth sensing Accelerometer Ambient Light Barometer Compass GPS Gyroscope

2.2 Limitations

The tango tablet alike 3D laser scanners are 3D laser scanners are sensitive to the laser light reflected from the scanned object surface. When scanning reflective surfaces with complex geometry, the scanners may record extensive outliers due to undesirable specular reflections. It has been reported that outliers caused by specular reflections can be formed according to the developed mixed reflection and multi-path reflection models [21].

Additionally, the hardware of focus in this research, the Google® Tango Tablet at the current time of this paper is still in the development phase.

3 Research Methodology

The aim of this paper is to determine the feasibility of IR Depth Sensor with SLAM tablets for area scanning. The process of laser scanning requires significant planning and time, of which is not conducive of many construction project schedules. The release of the Google® Tango Tablet has the potential to allow for a faster process for scanning of areas. Preliminary testing of the tablet and its accuracy is carried out though scanning as-built conditions. The as-built area scanned, a small conference room in the Miller Gorrie Center on the campus of Auburn University was chosen. The produced scans were then compared to the as-built drawings and also the as-built Autodesk® Revit model. Time to prep, conduct the scan, and time to process the scan were additionally assessed during this preliminary study.

4 Details of Study Extent

The SLAM technology of the Google® Tango Tablet, and its ability to compute its location, independently, without the the aid of GPS required minimal prep time. Additionally, there is no need for markers to be placed as required by laser scanning. Unlike a fixed scanner, i.e. laser scanners, the Google® Tango Tablet utilizes built-in motion tracking. Of which allows for the user to sweep the area from different angles scanning untethered within a space.

The Google® Tango Constructor application developed for the tablet was used to conduct the scans. Upon starting the tablet calibrates itself requiring the user to extend the tablet out in front of then in a steady position for a few seconds. Once the tablet is calibrated the user has full range to scan spaces within 0.5 to 4 meters (1.5 to 13 feet) from the tablets sensors. Currently the tablet is designed to work best indoors at moderate distances [8]. The application supports functions allowing the user to pause, resume, or reset a scan in progress. This allows scanning to stop or pause and resume scanning of the same area. Additional options allow for the scan to be saved in the native file type for the application, as will as an .OBJ file time.

The results of a scan produce a 3D mesh which can be saved in the .OBJ file format and transferred to a computer for further processing. The file was then imported into Autodesk® Memento Software and exported as an Autodesk® Recap Point Cloud File. The converted scan was then imported into Autodesk® Revit as a point cloud file. Providing the same functionality as laser scanned produced point cloud.

5 Results

During the duration of this research multiple operating (OS) system and application updates were released. With each iteration, both the OS and application significantly improved in functionality as well as overall stability. Scans were repeated to obtain uniformity of the produced 3D models and point clouds. The multiple scans allowed for comparison and analysis in terms of accuracy, preparation time, scan time, processing time, and usability of the produced 3D mesh with regards to its applications.

5.1 Conference Room Scan

The size of the chosen conference room was approximately 13ft x 18ft. With a conference table and chairs centrally located in the room. To test the accuracy of the produced 3D models from the scan, two starting points, point A and point B were determined in the conference room, as marked in figure 2. The scanning path followed the perimeter of the room, clockwise from each starting point.

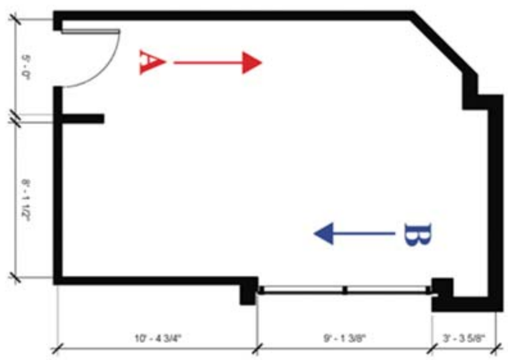

Figure 2: Starting Points of Scan

A total of ten scans were conducted, five scans with a start at point A, and five scans with a start at point B. A live view of the 3D Mesh can be seen on the tablet screen as it scans the area. The view shows the mesh model simultaneously being constructed refining itself as the the tablet sweeps the area. The user has the following options for viewing on the screen: first person, third person or top view. First person viewing was chosen as it displays in scanning in real-time as the user moves the tablet over areas to be scanned.

The multiple scans of the small conference room were able to be completed in a relatively short amount of time. The times of all ten scans were completed within the time range of 5-8 minutes. The scanning times and results are dictated by the user and their movements through the space. Figure 3 illustrates the results of a scan initiated at starting point A, scanning clockwise around the room. Figure 4 illustrates the results of a scan initiated at the starting point B, also scanning clockwise around the room. Figure 5 illustrates the results of a scan also initiated at starting point B.

The scans were conducted in the order as mentioned. There were mixed results produced by this initial test. The 3D models and point clouds varied in quality as well as in accuracy, with similarities shared by scans of the same starting points. Additionally, quality improved as the scans progressed, likely the result of the user's level of comfort in handling the tablet during scanning.

.

Figure 3: Scan Result Example 1

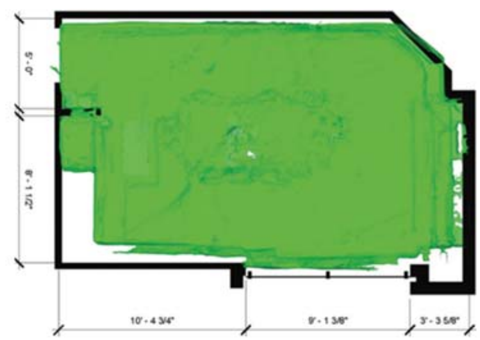

Figure 4: Scan Result Example 2

Figure 5: Scan Result Example 3

All the scans initiating at starting point A and starting point B exhibited improved quality and accuracy within the initial trajectory line among commencing from the starting points. Scanning accuracy of the overall model deteriorated with time and its movement along the path. This was due to drifting of the tablet and its ability to maintain its known location in the space. The portions affected by the drift maintain a greater level accuracy based of its new coordinate location within the space.

5.2 Mechanical Room Scan

Next the mechanical room within the Gorrie Center was chosen for scanning. The room is filled with mechanical equipment, piping, and ducts. The small space would be a challenge for laser scanning, and its multiple views required to develop a 3D scan of the space. Figure 6 shows a point cloud of the Mechanical Room produced from the tablet. Figures 6-8 illustrate sections through the room.

The scan of the Mechanical Room gave the tablet more surfaces and objects to reference its to. Overall there was only minor drifting of which occurred along the bottom wall in figure 7. The location of the drifting was poorly lit as well as having only flat surfaces. Both of which are the likely cause of the drifting. The scan the room was completed in 19 minutes. The complexity and tight spaces of the room required additional attention in order to get a usable model.

The existing Revit Model did not include the modelling of the pumps, pipes, ducts and equipment. The point cloud produced from the tablets scan shows the location of those items. The location of these items within the room can easily be modelled from the point cloud if needed.

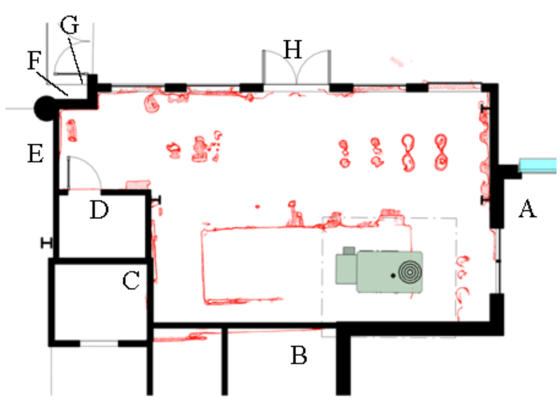

Figure 6: Mechanical Room Plan

Table 1. Mechanical Room scan analyses of As-Built Dimensions and Tablet scanned Dimensions

Wall	As Built Model (Meters)	Tablet Scan Range (Meters)
A	6.28	6.07-6.49
B	9.22	8.79-9.65
C	3.62	3.43-3.81
D	2.5	2.11-2.89
E	2.22	1.29-3.15
F	1.04	1.01-1.07
G	0.46	0.41-0.51
H	10.68	10.46-10.9

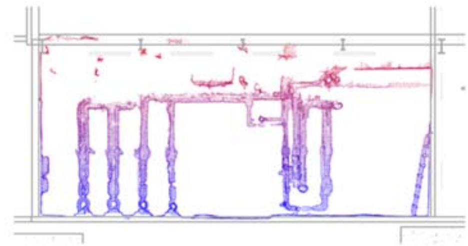

Figure 7: Mechanical Room Section A

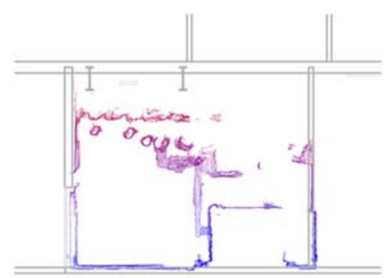

Figure 8: Mechanical Room Section B

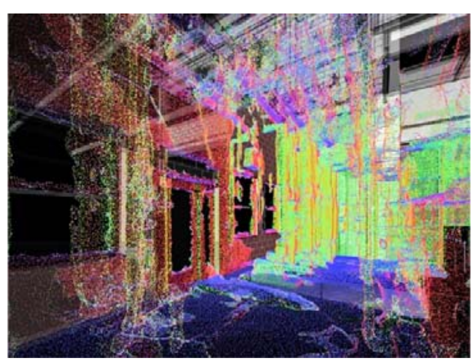

Figure 9: Mechanical Room Point Cloud

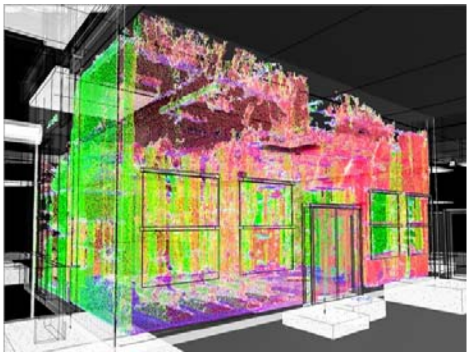

Figure 10: Mechanical Room Point Cloud

5.3 Large Space Scans

Expanding the scope of areas scanned, an office and connecting hallway were scanned on the second floor of the Gorrie Center. The initial start location is marked by the red dot in Figure 10. The starting office was full of desk, shelves, and equipment, providing ample surfaces for the tablet to reference itself to. Little to no drifting of the scan occurred initially within this first space. Shown in purple of figure 10 is the point cloud aligned to this first space.

As the scan continued into the hallway, the first major drift occurred as seen, in Figure 10. Shown in green is the point cloud aligned to the tablets new coordinate system. As the scan continued down the hallway a third drift of the tablets location occurred, shown in orange, the point cloud is aligned to its location as well. There was a significant amount of drift between the initial point cloud and the drifted aligned point clouds. While there was a significant drift from beginning to end, each segment produced usable scan information. Consistency (the true position is always located inside the estimated uncertainty) cannot be guaranteed in such cases [22]. The longer the trajectory is, the more important the drift becomes [23].

It can also be determined the greater amount of surfaces in a space to be scanned has greater potential in producing a more accurate model or point cloud.

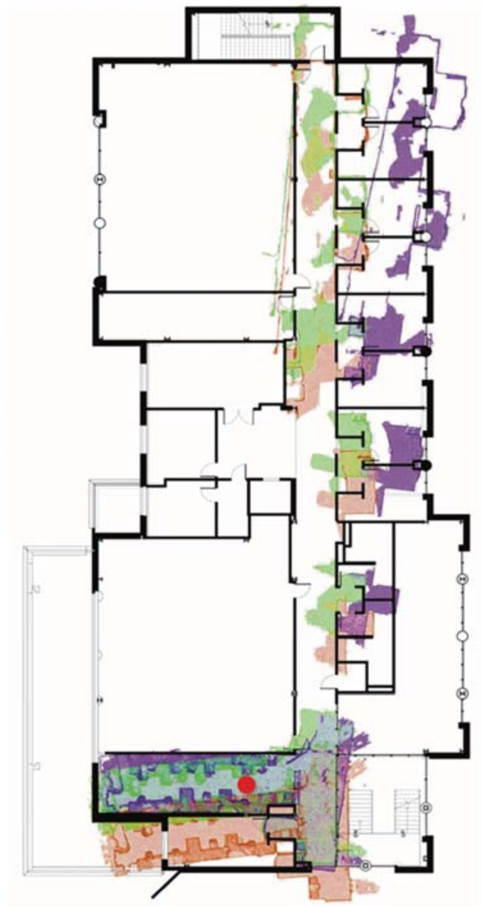

Figure 11: Overlay Illustrating Drifting producing any scans.

6 Discussion and Recommendations

6.1 Area Learning

The Tablet has a feature "Area Learning" in which the tablet familiarizes itself or "localizes" know locations within a space. Every time the tablet scans the area it registers any new points and modifies the "learned area" of that space. At the time of this research, our scans utilizing the constructor app we not able to benefit from this feature. This would significantly benefit and aid to the accuracy of indoor location tracking, as well as producing any scans.

6.2 Integrity of Areas to be scanned

This research has determined there to be a wide range of factors that influence the integrity of a produced point cloud. For any area in order to be scanned must have surfaces capable of being registered by the tablets sensors. Translucent or Glass material cannot be read by the tablet, as well as shiny or surfaces with a gloss finish. The tablets sensors aren't able to get an accurate depth reading due to the scattering of the transmitted laser and IR signals.

Lighting is also essential to the scan and its ability to scan an area. Too much light, or not enough light caused drifts in the tablets location within the space as well as holes in the produced model or point cloud. Direct sunlight on surfaces, specifically cause holes, as they are not able to be scanned, similar to glass or windows. As this research was conducted indoors in an as-built environment, lighting was detrimental to the overall quality of a scan.

The user's method of scanning a space, in regards to how they handle the tablet, their speed through the space, and ability to thoroughly scan around the space from multiple perspectives. Further research to determine the best practices for scanning using the tablet could be beneficial.

Additionally, through this research it can be determined that the quantity and close proximity of surfaces to one another can produce a more accurate model. The more surfaces the tablet is able to identify and register, the more accurate it can identify its own location within an area to be scanned. While cluttered and/or narrow environments provide many optical features and simplify the use of SLAM methods, they reduce the free walking experience as the resulting resets would disturb the user experience [13].

The accuracy of the scans completed in this research, those of the conference room and of the mechanical room produced different levels of accuracy. With the mechanical room have more surfaces though out the space, the tablet was able to consistently track its own location. Within the mechanical room scan the area of drift occurred in an area consisting of flat surfaces (wall and air handing unit). There were no distinguishable surfaces of which could be registered by the tablet.

6.3 Conclusions

This paper presented results of the use of handheld SLAM for 3D scanning. The results consisted of an array of as-built conditions. The scans (point clouds) were analysed and compared against as-built Building Information Models to test accuracy. Dimensions were analysed in comparison to the tablet scans. Multiple scans were completed giving a range of accuracy as shown in Table 1. While these scans are not of the same quality or high accuracy of terrestrial laser scans, they can serve as a beneficial aid during construction. Time and skill required to complete scans on site are minimal and can serve as a beneficial tool. Additionally, there are significant costs and time savings in comparison to the cost of traditional fixed laser scanners.

Future Research

Incorporating SLAM technologies into user friendly devices has a wide range of applications as well as aid in higher levels of communication amongst those involved in building construction. SLAM allows for a device to know where it is within a space, untethered, deriving location from a array of sensors. The Google® Tango Tablet while still in development shows great promise. This research tested the applications of area scanning for use in the construction industry and has determined the tablet has the great potential for fast area scanning. While scanning of large spaces isn't nearly as accurate as scanning smaller ones, they both offer advantages.

Future research will build upon this study, continuing research with the Google® Tango Tablet, as well as the Microsoft® HoloLens. The HoloLens is a mixed reality headset of which utilizes SLAM technology to overlay information within a space. The headset will release the user of handling a tablet, allowing for a hands-free experience. Like the tablet the headset scans the space around the user, continuously scanning as the user moves around. SLAM technology is an innovative approach that has yet to be fully realized.

References

[1] 3D scanners comparison. Aniwaa 2016. http://www.aniwaa.com/comparison/3d-scanners/ (accessed April 3, 2016).

[2] Caitlin T. Clark, Junshan Liu. A Case Study of Integrating 3D High Definition Laser Scanning Technology and BIM into Drywall Installation. 50th ASC Annual International Conference Proceedings, 2014.

[3] Geoff Jacobs. Uses in Building and Architectural Surveys. Professional Surveyor Magazine 2005.

[4] Becerik-Gerber B, Jazizadeh F, Kavulya G, Calis G. Assessment of target types and layouts in 3D laser scanning for registration accuracy. Automation in Construction 2011;20:649–58. doi:10.1016/j.autcon.2010.12.008.

[5] Humayun Kabir Biswasa D, Boschéa F, Suna M. Planning for Scanning Using Building Information Model: A Novel Approach with Occlusion Handling. Symposium on Automation and Robotics in Construction and Mining (ISARC 2015), vol. 15, 2015, p. 18.

[6] Schall G, Zollmann S, Reitmayr G. Smart Vidente: advances in mobile augmented reality for interactive visualization of underground infrastructure. Personal and Ubiquitous Computing 2013;17:1533–49.

[7] Wang X, Kim MJ, Love PED, Kang S-C. Augmented Reality in built environment: Classification and implications for future research. Automation in Construction 2013;32:1–13. doi:10.1016/j.autcon.2012.11.021.

[8] Biswas J, Veloso MM. Wifi localization and navigation for autonomous indoor mobile robots 2010.

[9] Chintalapudi K, Padmanabha Iyer A, Padmanabhan VN. Indoor localization without the pain. Proceedings of the sixteenth annual international conference on Mobile computing and networking, ACM; 2010, p. 173–84.

[10] Thrun S, Burgard W, Fox D. Probabilistic robotics. Cambridge, Mass: MIT Press; 2005.

[11] Winterhalter W, Fleckenstein F, Steder B, Spinello L, Burgard W. Accurate indoor localization for RGB-D smartphones and tablets given 2D floor plans. Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, IEEE; 2015, p. 3138–43.

[12]S. Zollmann, C. Hoppe, S. Kluckner, C. Poglitsch, H. Bischof, and G. Reitmayr *Augmented Reality for Construction Site Monitoring and Documentation* Proceedings of the IEEE, vol. 102, no. 2, pp. 137–154, Feb. 2014.

[13] Nescher T, Zank M, Kunz A. Simultaneous Mapping and Redirected Walking for ad hoc Free Walking in Virtual Environments, IEEE Virtual Reality Conference 2016, USA; 2016.

[14] Sweeney C, Flynn J, Nuernberger B, Turk M, Hollerer T. Efficient Computation of Absolute Pose for Gravity-Aware Augmented Reality, IEEE; 2015, p. 19–24. doi:10.1109/ISMAR.2015.20.

[15] Landy MS, Movshon JA, editors. Computational models of visual processing. Cambridge, Mass: MIT Press; 1991.

[16] Maybank S. Theory of reconstruction from image motion. Berlin: Springer; 1993.

[17] Kneip L, Scaramuzza D, Siegwart R. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE; 2011, p. 2969–76.

[18] Gauglitz S, Sweeney C, Ventura J, Turk M, Hollerer T. Model Estimation and Selection towardsUnconstrained Real-Time Tracking and Mapping. IEEE Transactions on Visualization and Computer Graphics 2014;20:825–38. doi:10.1109/TVCG.2013.243.

[19] Schops T, Sattler T, Hane C, Pollefeys M. 3D Modeling on the Go: Interactive 3D Reconstruction of Large-Scale Scenes on Mobile Devices, IEEE; 2015, p. 291–9. doi:10.1109/3DV.2015.40.

[20] Google. Project Tango – Google 2015. https://www.google.com/atap/project-tango/hardware/ (accessed September 18, 2015).

[21] Wang Y, Feng H-Y. Modeling outlier formation in scanning reflective surfaces using a laser stripe scanner. Measurement 2014;57:108–21. doi:10.1016/j.measurement.2014.08.010.

[22] Julier SJ, Uhlmann JK. Building a million beacon map. In: McKee GT, Schenker PS, editors., 2001, p. 10–21. doi:10.1117/12.444158.

[23] Bresson G, Aufrère R, Chapuis R. Making visual SLAM consistent with geo-referenced landmarks. Intelligent Vehicles Symposium (IV), 2013 IEEE, IEEE; 2013, p. 553–8.