2018 18th International Conference on Control, Automation and Systems (ICCAS 2018)
Oct. 17–20, 2018; YongPyong Resort, PyeongChang, GangWon, Korea

Towards Automated 3D reconstruction in SME factories and Digital Twin Model generation

Mathias Minos-Stensrud, Ole Henrik Haakstad, Olav Sakseid, Baard Westby, and Alex Alcocer*

Department of Mechanical, Electrical, and Chemical Engineering
Oslo Metropolitan University
Pilestredet 35, 0166, Oslo, Norway
Email: alex.alcocer@oslomet.no * Corresponding author

Abstract: This paper presents preliminary results towards the development of digital twin models for Small to Medium Enterprise (SME) factories in a partially automated and cost-effective manner. In many cases it is desirable to obtain a 3D model of a factory floor and machinery, that can be used for visualization of Digital Twin models. Current commercial 3D reconstruction solutions involve the use of high-end LiDAR sensors which increase the cost of the 3D scanning process and suppose a barrier for SME factories on their path towards Industry 4.0. The paper presents a comparison of 3D reconstruction results using low-cost sensors including a Zenfone AR mobile phone, an Intel RealSense ZR300 and a Kinect v2. The small size and weight of the sensors make it possible to be mounted on small unmanned aerial vehicles and enable future 3D reconstruction in an autonomous manner. The data was processed using an open source Simultaneous Localization and Mapping (SLAM) library RTAB-Map. The results were compared with a professional 3D scan using a GeoSLAM LiDAR. Experimental results from a scanning of a university research lab with a small simulated production line and two UR3 industrial manipulators is presented. The obtained 3D model was used to generate a simple Digital Twin model that can be visualized using a VR headset.

Keywords: Digital Twin, SLAM, Autonomous Systems, Industry 4.0, 3D reconstruction, Unmanned Aerial Vehicles, RGB-D cameras

1. INTRODUCTION

One of the central elements of industry 4.0, Cyber-physical systems, and the digitalization of manufacturing processes are the concept of digital twin [1–5]. A digital twin is a virtual copy of a physical system that can be used for a number of purposes including visualization, process monitoring, process optimization, and predictive maintenance. An important characteristic of digital twin is that it is connected to the real physical system that it represents.

Digital twin models have been used for many years in large enterprises, but their application has been very limited in small to medium enterprise (SME) factories due to the high cost associated with their development. This paper presents preliminary results of a research project DigiFab *Automated SME factory digitization and road map to Industry 4.0* that aims at providing SME factories with affordable tools to develop digital twin models and progress towards industry 4.0 [6].

In many cases, it is necessary to obtain a 3D model of a factory floor and its machinery in order to develop a digital twin. Current solutions involve the use of LiDAR scanners, in a process that is costly and cumbersome. This paper researches the use of drones to perform 3D scanning in factories in an autonomous manner. The goal is to explore the use of low-cost sensors such as structured light depth cameras [7] as found in ZenFone AR mobile phone, an Intel RealSense ZR300 and Microsoft Kinect v2 that are small and light enough to be mounted on a small drone. By using Simultaneous Localization and Mapping (SLAM) [8] it is possible to envisage a system that allows small drones to simultaneously localize themselves and perform 3D reconstruction in factory environments in a fully autonomous manner. This could be accomplished by using readily available open source libraries for SLAM such as RTAB-Map [9] [10].

This paper presents results that explore different aspects of the road towards automated 3D-reconstruction and generation of digital twins. In particular, it concentrates on three parts: (i) Position control of a drone using a feedback-loop with a motion capture system; (ii) 3D-reconstruction using three cost-effective and mobile sensors; and (iii) A demonstration of a simple digital twin of OsloMet ARIS research lab containing a small production line with 2 Universal Robots UR3 industrial manipulators, and which can be visualized in virtual reality (VR).

2. EXPERIMENTS WITH DRONE INDOOR POSITION CONTROL

Outdoor 3D-reconstruction using drones is becoming standard practice in commercial contexts. It is cheap, efficient and accurate. Techniques like photogrammetry and LiDAR approaches are widely used. However, they involve large amounts of manual control along the way and typically use large drones to be able to carry LiDAR payloads. The objective of this paper is to explore the use of indoor autonomous drone(s) in performing 3D recon-

Fig. 1 The Intel Aero Ready To Fly drone together with the Asus Zenfone AR [11].

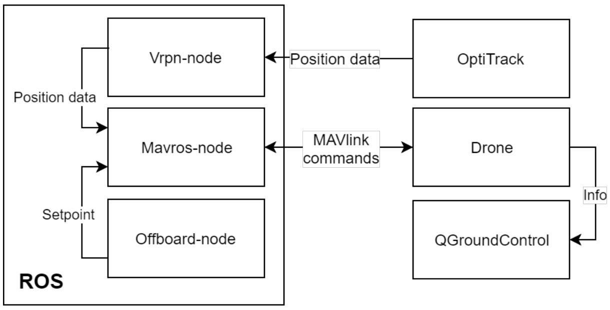

Fig. 2 A flowchart of the position-control system

struction of SME factories.

In the absence of GPS measurements, this paper explores the use of SLAM in order to both provide position information for the closed-loop control of the drone as well as a 3D point cloud. As a preliminary step towards this goal, an experimental setup was made in a research laboratory equipped with an OptiTrack motion capture system that provides real-time position information. The goal is to use these external position measurements in the development and benchmarking of algorithms.

The preconditions for this subtask was the Intel Aero Ready To Fly drone (Figure 1) and an OptiTrack motion capture system. The Intel Aero is a UAV development platform equipped with the Intel Aero Compute Board running Linux on a Quad-core CPU, an STM32 microcontroller running PX4 autopilot on NuttX, and the Intel vision accessory kit. This allowed for use of the Robotic Operation System (ROS) as a framework which facilitates for easy further development in the future.

Figure 2 shows a flowchart of the final implemented solution which was based on a guide from PX4 [12]. The OptiTrack system sends the pose data of a defined rigid body of the Intel Aero drone over the VRPN-protocol. This is received by the ROS-node vrpn_client_ros [13]. The desired setpoint is generated by a ROS-node made of a provided example by PX4 [14]. Both the orientation data and the desired setpoint is sent to the PX4 autopilot over the MAVLink protocol using the ROS-node mavros [15]. The autopilot, in turn, adjusts the position of the drone based on these two variables.

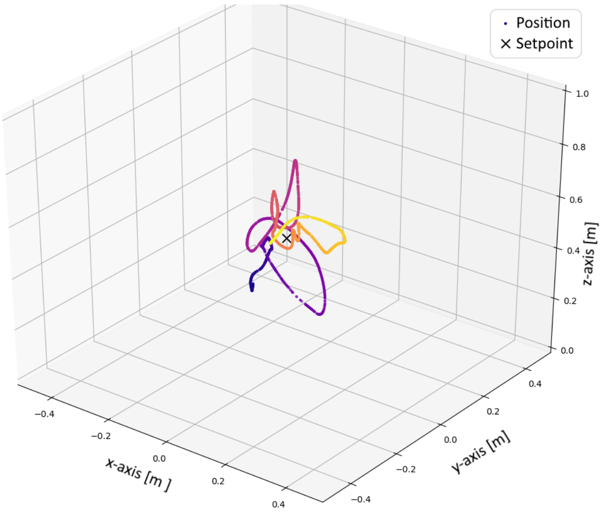

Fig. 3 The position of the drone during a test-flight. 2500 points during 30s plotted in a cube of 1 m^3, with the setpoint (0, 0, 5) m.

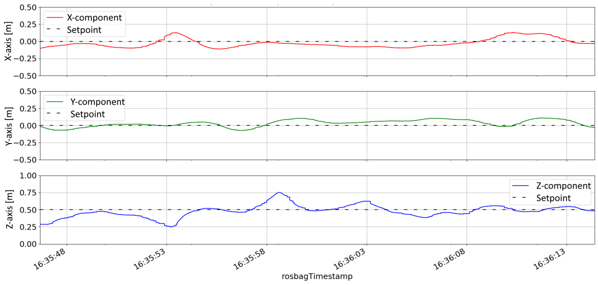

Fig. 4 The position of the drone plotted component-wise during a test-flight.

While the motion capture system can provide an accuracy of 0.2 mm, an analysis of one of the test flights showed a mean error in distance of 12.2 cm over 30 seconds of flight, as seen in figures 3 and 4. The same test flight showed a mean absolute deviation of 4-6 cm in every axis.

An approach to the use of position measurements from a motion capture system has been given. For accurate benchmarking it is still necessary to improve the system. The viability of using drones for 3D-reconstruction is still uncertain and needs further research. While the 3D-reconstruction would not depend directly on the quality of the flight, variables as indoor turbulence and limited space available suggests the necessity of 3D-sensors and algorithms with high tolerance to sudden movements and with excellent position tracking.

3. TOWARDS AUTOMATIC 3D-RECONSTRUCTION

LiDAR laser scanners are widely used in 3D scanning and 3D reconstruction. They provide excellent accuracy and cloud density, but also very expensive and usually stationary. One example is the Riegl VZ-400i with a range up to 800m, accuracy up to 5mm and data acquisition up to 500,000 measurements per second, but with a price tag of over 100 000 USD (Table 1).

It was desirable to look at cheaper alternative 3D sensors and to what extent these would be useful in the context of 3D scanning done autonomously with a drone. Furthermore it was desirable to look into the road from a point cloud to a 3D model both manually and semi-automatically.

Three 3D sensors where investigated, the Intel Realsense ZR300, the Asus Zenfone AR and finally the the Microsoft Kinect v2 (Figure 5). They were compared to the ZEB-REVO hand-held laser scanner from GeoSLAM.

Fig. 5 From the left: Intel RealSense ZR300, Microsoft Kinect v2, Asus ZenFone AR, ZEB-REVO. Not in correct sizes.

3D reconstruction was done for all the three 3D sensors using RTAB-Map (Real-Time Appearance-Based Mapping), an RGB-D Graph-Based SLAM approach based on an incremental appearance-based loop closure detector. The loop closure detector uses a bag-of-words approach to determinate how likely a new image comes from a previous location or a new location. When a loop closure hypothesis is accepted, a new constraint is added to the maps graph, then a graph optimizer minimizes the errors in the map. A memory management approach is used to limit the number of locations used for loop closure detection and graph optimization, so that real-time constraints on large-scale environments are always respected. [9] [10].

Intel RealSense ZR300 uses Active Stereo IR Technology for 3D sensing. This is in principle a stereo IR vision technique which also utilizes an IR projector projecting a known pattern to counter low texture cases [21] [22]. The ZR300 also have an IMU-sensor integrated and Visual-Inertial Odometry (VIO) was implemented for this sensor in RTAB-Map and resulted in very robust 3D scanning; The sensor did not lose track of its position during the SLAM-process.

Microsoft Kinect v2 uses the Time Of Flight principle for 3D sensing and measures the time the light emitted takes to return and its phase shift. Kinect v2 depends solely on visual odometry and only with enough computing power was this a workable solution for the initial problem, autonomous 3D scanning.

Asus ZenFone AR uses three different cameras and has called this for the TriCam system. An HD camera with 23 megapixels, a camera dedicated to motion tracking, and an IR depth sensor [17]. Zenfone also makes use of Visual Inertial Odometry, but since all computing is done onboard the phone, it lacks computing power and robustness to shaking.

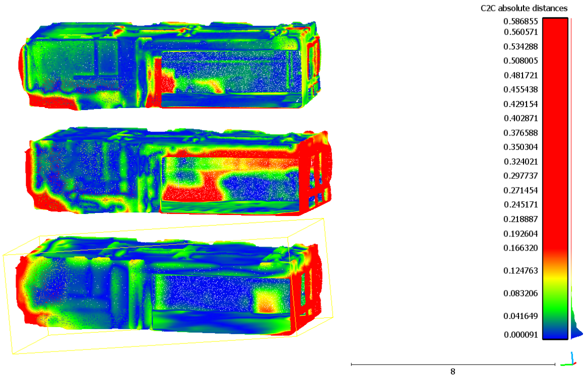

Fig. 6 From the top: Asus Zenfone AR, Intel RealSense ZR300, Microsoft Kinect v2. The color indicates the magnitude of error, ranging from 0.05 m - 0.5 m.

Two techniques were considered moving from point cloud to 3D model. Semi-automatic generation of meshes and manually drawn 3D models. Semi-automatic mesh generation was done using multiple filters in RTAB-Map and Meshlab, such as point cloud filtering, smoothing and surface reconstruction. Autodesk Revit was used to manually draw 3D models, and the point cloud generated by ZEB-REVO was used as a template. The final result can be seen i figure 7.

All three point clouds produced by the sensors was compared to a point cloud produced by the ZEB-REVO from GeoSLAM, as seen in figure 6. The figure is the ZEB-REVO 3D model with the magnitude of the error in each comparison colorized. The comparison does not provide sufficient information on which sensor which provides the best reconstruction, since the result is highly dependent on optimal implementation, calibration and a comparable reconstruction process. However, the Intel RealSense ZR300 showed the greatest potential for further work. It was lightweight and featured robust 3D reconstruction thanks to the possibility of using visual-inertial odometry. The new Intel Realsense D415 and D435 should also be considered, featuring, among other things, greater depth-sensing range.

Table 1 Comparison of 3D sensors

	Intel RealSense ZR300 [16]	ASUS Zenfone AR [17]	Microsoft Kinect v2 [18]	GeoSLAM ZEB-REVO [19]	RIEGL VZ-400i [20]
Measurement technique	Active stereo IR Technology	TriCam system IR projector	Time of Flight	LiDAR	LiDAR
Min. range	0.55 m	Not specified	0.5 m	Not specified	0.5-1.5 m
Max. range	2.8 m	Not specified	4.5 m	15-30 m	120-800 m
Data acquisition rate	30-60 frames/sec	Not specified	30 frames/sec	43 200 meas./sec	42 000-800 000 meas./sec
Horizontal FoV	59°	Not specified	70°	360°	360°
Vertical FoV	46°	Not specified	60°	270°	100°
Accuracy	±5% of meas. dist.	Not specified	∼1 cm	1-3 cm	5 mm
Price (as of 27/08/18)	∼400 USD	∼ 700 USD	∼ 100 USD	∼ 70 000 USD	∼ 100 000 USD

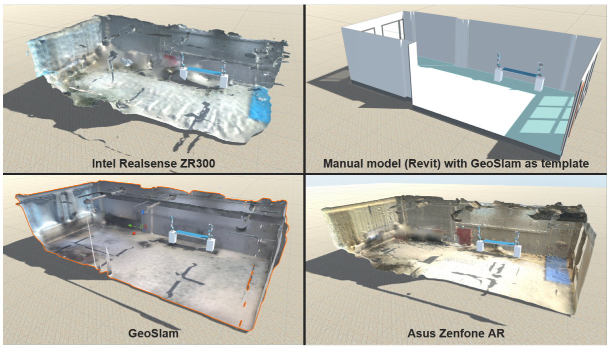

Fig. 7 The generated 3D models from Intel Realsense ZR300, Asus Zenfone AR, ZEB-REVO and the manually drawn model as a comparison

4. TOWARDS DIGITAL TWIN MODEL GENERATION

A digital twin is an extensive concept. It can be used to understand how a product performs, through data analysis, and how the product will perform in the future through simulations. It can be used to identify potential errors or troubleshoot situations remotely with visualization of real-time data. It might become a key feature in the training of employees in environments involving risk or downtime from production.

The purpose of this subtask was to use the produced 3D reconstructions, as described in section 3., in development of a simplified digital twin. This would in turn give insight on how to automate the next step after 3D reconstruction: Generation of digital twins.

At hand was an Oculus Rift system, two Universal Robots UR3-robots and an OptiTrack motion capture system. A real time connection was made with the UR3-robots using the Real-Time Data-Exchange interface (RTDE-interface) from Universal Robots. This provided the joint angles of each robot which in turn was read by Unity and attributed to the two 3D models of the UR3-robots. The OptiTrack software offers Unity integration and this was used to get the position of the four boxes which the UR3-robots moved. The 3D models (See section 3). was imported to simulate the environment. Figure 8 and 9 shows the final result.

Fig. 8 3D models displayed side by side, along with the robots inside, in Unity.

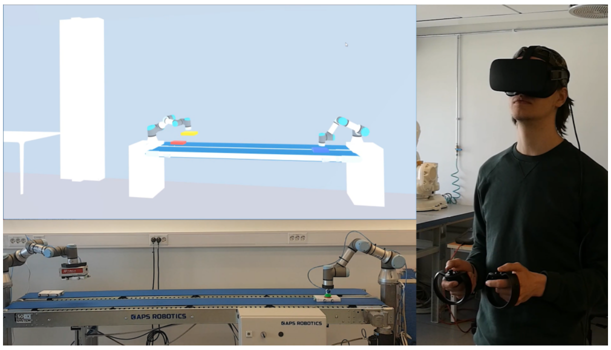

Fig. 9 Digital Twin model demonstration using Oculus VR headset.

A working simplified digital twin was developed where the 3D robot was moving visually indistinguishable from the real UR3-robot. It became clear that an equivalent process, 3D modeling and real-time communication between interfaces, had to be made for every different robot a digital twin should include.

5. CONCLUSION AND FUTURE WORK

The purpose of this paper has been to investigate and demonstrate some of the aspects of the road towards automated 3D reconstruction and Digital Twin model generation.

For further work, the Intel Aero drone's built-in 3D-sensor could be used for SLAM on-board, on-the-go. This data could further be used for both real-time 3D reconstruction and exploration algorithms for autonomous 3D scanning.

From point clouds, a surface-finding algorithm could detect walls, ceilings, and floors and generate appropriate flat surfaces which compared to the 3D-mesh presented in this paper would be less computational intensive in a VR environment.

In this project, the RTDE-interface was used only for a visual representation of robots. For further work, much more data could be collected and used for real-time monitoring and analysis. The RTDE-interface could also be used for remote controlling: programming the robots remotely in VR.

6. ACKNOWLEDGEMENETS

The research was supported by Norwegian Research Council through project DigiFab, and OsloMet lighthouse project Autonomic Systems. The authors would like to thank Espen Olsen for his help in experiments carried out at OsloMet ARIS lab.

REFERENCES

[1] A. W. David B. Cameron and T. M. Komulainen, "Oil and Gas digital twins after twenty years. How can they be made sustainable, maintainable and useful?" in *Proceedings of the 59th International Conference of Scandinavian Simulation Society, SIMS2018*, accepted.

[2] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567 – 572, 2015, 15th IFAC Symposium onInformation Control Problems in Manufacturing.

[3] G. Schroeder, C. Steinmetz, C. E. Pereira, I. Muller, N. Garcia, D. Espindola, and R. Rodrigues, "Visualising the digital twin using web services and augmented reality," in *2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, July 2016, pp. 522–527.

[4] F. Tao and M. Zhang, "Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing," *IEEE Access*, vol. 5, pp. 20 418–20 427, 2017.

[5] S. Boschert and R. Rosen, "Digital twin–the simulation aspect," in *Mechatronic Futures*. Springer, 2016, pp. 59–74.

[6] OsloMet, "The road towards industry 4.0," Internet http://www.hioa.no/eng/News/Veien-mot-Industri-4.0, [05/21/2018], 2018.

[7] P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo, *Operating Principles of Structured Light Depth Cameras*. Cham: Springer International Publishing, 2016, pp. 43–79.

[8] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[9] M. Labbe and F. Michaud, "Appearance-based loop closure detection for online large-scale and long-term operation," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 734–745, 2013.

[10] ——, "Online Global Loop Closure Detection for Large-Scale Multi-Session Graph-Based SLAM," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2014, pp. 2661–2666.

[11] PX4, "Px4 - intel aero," Internet: https://docs.px4.io/en/flight_controller/intel_aero.html, [05/15/18].

[12] ——, "External position estimation," Internet: https://dev.px4.io/en/ros/external_position_estimation.html, [05/08/18].

[13] P. Bovbel, "vrpn_client_ros - ros wiki," Internet: http://wiki.ros.org/vrpn_client_ros, [05/08/18].

[14] PX4, "Mavros offboard," Internet: https://dev.px4.io/en/ros/mavros_offboard.html, [05/09/18].

[15] V. Ermakov, "Mavlink to ros," Internet: https://github.com/mavlink/mavros, [05/09/18].

[16] Intel, "Intel realsense 3d camera zr300 data sheet," Jan 2017.

[17] ASUS, "Zenfone ar tech specs," Hentet 16/05/2018 fra https://www.asus.com/Phone/ZenFone-AR-ZS571KL/Tech-Specs/.

[18] T. Sugiura, "Kinect v2 introduction and tutorial," Hentet 18/05/2018 fra https://www.slideshare.net/SugiuraTsukasa/kinect-v2-introduction-and-tutorial, 2014.

[19] GeoSLAM, "Zeb-revo datasheet," Hentet 28/08/2018 fra https://geoslam.com/wp-content/uploads/2018/04/GeoSLAM-ZEB-REVO-Solution_v9.pdf, 2017.

[20] RIEGL, "Riegl vz-400i data sheet," Internet: http://www.riegl.com/uploads/tx_pxpriegldownloads/RIEGL_VZ-400i_Datasheet_2017-12-18.pdf, [05/20/18].

[21] Intel, "Intel realsense 3d camera zr300," Online: http://click.intel.com/media/ZR300-Product-Datasheet-Public-002.pdf, [11/05/2018], 2017.

[22] W. Jang, C. Je, Y. Seo, and S. W. Lee, "Structured-light stereo: Comparative analysis and integration of structured-light and active stereo for measuring dynamic shape," *Optics and Lasers in Engineering*, vol. 51, no. 11, pp. 1255–1264, 2013.