# Indoor Point Cloud Segmentation for Automatic Object Interpretation

**LAVINIA S. RUNCEANU[1], SUSANNE BECKER[1], NORBERT HAALA[1] & DIETER FRITSCH[1]**

*Summary: The paper presents an algorithm for the automatic segmentation of point clouds from low cost sensors for object interpretation in indoor environments. This algorithm is considering the possible noisy character of the 3D point clouds and is using an iterative RANSAC approach for the segmentation task. For evaluating the robustness, it is applied on two indoor datasets, acquired with the Google Tango tablet and with the NavVis M3 trolley. The realized evaluation reveals the potential of the two systems for delivering data suitable for automatically interpreting indoor structures.*

## 1 Introduction

A large variety of systems and applications are offering mapping, localization and navigation services for outdoor environment. However, people spend most of their time indoor, where there is a lack in digital maps and where conventional GPS services do not work. Despite this need in indoor navigation applications, the developments in the field of augmented and virtual reality, including also game industry, and Building Information Modeling (BIM), are also requiring intelligent indoor models. In order to obtain the needed models, innovative equipment is required to replace the traditional systems, which are mostly expensive and sometimes inconvenient to use. Recently, a variety of systems, designed for this purpose, were made available on the market. In order to increase the mapping efficiency and in the same time to reduce the mapping costs, these systems integrate laser scanners, cameras and sometimes inertial measurement units. Also, they adapt their design to the indoor space, being built as a trolley or as a backpack (e.g. NavVis M3 Trolley, Leica Pegasus). Another step further is made by the availability of devices integrating depth cameras, at a consumer-level. These platforms are using low cost sensors which made them affordable to the general public (e.g. Phab 2 Pro Phone, Google Tango Tablet, Microsoft Kinect). All this progress enables unexperienced users to contribute to the indoor mapping request, but in the same time this rises new challenges which need to be overcome. The acquired 3D data needs to be automatically interpreted in order to obtain the models for the aforementioned applications. For the automatic 3D data interpretation, segmentation is needed. Point cloud segmentation is a subject of research for many years, however point clouds coming from low cost sensors, rise new challenges for the segmentation and interpretation process, which are addressed by this work.
The contributions of the paper are:

- a robust algorithm for automatic extraction of wall structures in indoor scenarios;
- analysis of the potential of different sensors to provide interpreted indoor structures with a data-based algorithm.

[1] University of Stuttgart, Institute for Photogrammetry (ifp), Geschwister-Scholl-Straße 24D, D-70174 Stuttgart, E-Mail: [Forename.Lastname]@ifp.uni-stuttgart.de

The paper is structured as follows: Section 2 provides an overview of the related work focused on the available indoor mapping systems and on the segmentation and reconstruction methods using point clouds. The used systems and the obtained indoor data are presented in section 3. The definition of the proposed interpretation algorithm is given in section 4. Section 5 gives an accuracy analysis of the input data, as well as of the automatically interpreted structures. Conclusions are given in section 6.

## 2 Related Work

The related work is presented separately considering the available indoor mapping systems and the point clouds reconstruction methods.

### 2.1 Indoor mapping systems

Being well suited for mapping textureless surfaces, often present in indoor environments, active systems are the most used for satisfying the indoor mapping request. Improvements in sensor design and technology as well as in the used algorithms have contributed to the large variety of systems used for 3D data collection today. A lot of research work is addressing the available data acquisition solutions for indoor modeling. However, in this paper, the focus is on the recent ones using Indoor Mobile Mapping Systems (IMMS) commercially available and low cost consumer-grade range cameras will be mentioned here.

Initially introduced as *Zebedee* by BOSSE et al. (2012), the ZEB1 from GeoSLAM (GEOSLAM 2017) is composed of a 2D laser scanner, a low-cost IMU and two robotic systems springs. THOMSON et al. (2013) are comparing it and the iMS 3D from Viametris (VIAMETRIS IMS 3D 2017) with Terrestrial laser scanner (TLS) in terms of accuracy. Both IMMS systems delivered centimeter-accuracy, but the iMS 3D, composed of three 2D laser scanners and a Point Grey Ladybug spherical camera, proved to generate higher quality point clouds. BASSIER et al. (2015) investigated data acquisition techniques and workflows considering the transitory tendency from TLS to IMMS. Despite the previous mentioned systems, they considered also the M3 trolley from NavVis in their comparison. Another solution offered by Leica Geosystems is the integration of the needed mapping sensors in the Pegasus Backpack. KURIAN & MORIN (2016) used the 3D point cloud generated by it for developing a method of minimizing the computation cost and data storage for real-time mapping applications. Although range cameras were available for several years in the game industry, the research interest grew with the release of Microsoft Kinect (MSDN KINECT 2017) in 2010. Being firstly based on structured light and later on time-of-flight (ToF) principle, the Microsoft Kinect became one of the most affordable device for 3D mapping. BÖHM (2014) includes it in his investigations, where he checked the performance of some structured light sensors in regard to accuracy and repeatability. Continuously improvement in the hardware and software made available low-cost range camera devices, aiming among others at indoor data acquisition, like it is the case of DPI-8 (DOTPRODUCT DPI-8 2017) and Google Tango tablet (WIKIPEDIA TANGO 2017). The capabilities of the last mentioned one, Tango tablet, have been subject of research for indoor scanning by DIAKITÉ & ZLATANOVA (2016).

Due to the indoor scene complexity and to sensors limitations (like battery life), normally the acquired data, is composed of multiple sessions (sometimes from different viewpoints), which have to be aligned and geo-referenced, before other processing is performed. While there is the option to manually define the point correspondences and to apply an algorithm like Iterative Closest Point (ICP) (BESL & MCKAY 1992), this problem is mostly solved automatically with Simultaneous Localization and Mapping (SLAM) (LEONARD & DURRANT-WHYTE 1991) approaches. Each of the before mentioned systems are using a type of SLAM implementation for partially or totally solving the registration problem. Being commercial products, the used algorithm is not always good documented. However, some research works are addressing the SLAM problem specifically to the device, like: BOSSE et al. (2012) for *Zebedee*, NEWCOMBE et al. (2011) for Kinect, LASKAR et al. (2016) for Tango tablet.

### 2.2 Segmentation and reconstruction methods using point clouds

The resulted point clouds after the registration task are normally not used as an end product. They are usually modelled as surfaces and volumes with the use of a variety of reconstruction methods. A lot of research work was realized in this regard. However, most of it is making use of the Manhattan World constraints and it is dealing with high accurate data, coming from TLS. JEKE et al. (2009) are making use of the Manhattan World constraints and a graph structure for fitting cuboids to a point cloud, which are further merged to rooms and corridors. BUDRONI & BÖHM (2009) are performing a plane sweeping algorithm for identifying walls in a Manhattan World scenario. The approach used by VALERO et al. (2012) for wall segmentation is based on the work done by OKORN et al. (2010) and is finally delivering a boundary representation model. PREVITALI et al. (2014) are presenting an automatic reconstruction algorithm from TLS point clouds to semantically enriched models using terrestrial laser scanner data. BECKER et al. (2015) are proposing a grammar-based approach for automatic reconstructing 3D interiors from laser scanner point clouds. In his PhD thesis, KHOSRAVANI (2016) proposed an approach for obtaining topological correct indoor model from Kinect measurements. However, this solution is limited to small scale data (a room, a hallway, etc.), while we aim at reconstructing an entire floor. Also, part of the point clouds used within our investigations are less accurate than the point clouds provided by Kinect. Therefore, the algorithm from section 4 is proposed.

## 3 Indoor Mapping Systems and Delivered Data

In this work, two different systems, Google Tango tablet and NavVis M3 trolley, were used for the indoor mapping task. Their characteristics are presented in the followings.

### 3.1 Google Tango tablet

Being firstly introduced by Google in 2014, the Tango tablet (WIKIPEDIA TANGO 2017) is a development kit with 3D motion detection and depth measurement capabilities. Among sensors normally available for mobile devices, like accelerometer, gyroscope, GPS, etc., it also integrates a 3D depth sensor, composed by an infrared (IR) projector and IR sensor, and a wide-angle motion tracking camera (fisheye camera) (Fig. 1 (a)).

The depth is perceived with the help of infrared structured light (FOFI et al. 2004), giving the device the possibility of measuring the distance to the surrounding objects in a range of 0.5 – 4 m, with approximately 1% accuracy. The resulted depth data can be obtained in form of a point cloud or a

textured mesh, its quality being directly influenced by the light source lighting the objects and by their reflectivity (GOOGLE TANGO 2016a).
The device can track its own position and orientation in space (its pose) by making use of visual-inertial odometry (LI & MOURIKIS 2012). Hence, the device's pose combines the change in position, determined by tracking the features in the images from the motion tracking camera, and the rotation and acceleration changes, coming from the inertial motion sensors. However, this concept has limitations, over time the device poses being affected by drift (GOOGLE TANGO 2016b). In order to reduce the aforementioned drift and to estimate the device's position within the already measured areas, the device has area learning capabilities (GOOGLE TANGO 2017) which performs Simultaneous Localization and Mapping (THURN & LEONARD 2008).
As a mobile device used for indoor scanning, the Tango tablet is very easy to handle and it gives to the user the flexibility of moving through the space, which a fixed scanner cannot offer. Nevertheless, it has limited resources which influence also the space extend which can be measured in a session.

### 3.2 NavVis M3 trolley

The M3 trolley is a 3D mapping solution, being released in 2014 by the company NavVis (NAVVIS M3 TROLLEY 2017). It integrates three Hokuyo UTM-30LX laser scanners, a HDR panoramic camera head composed of six cameras, an inertial measurement unit (IMU), WiFi sensors, a magnetometer and an on-board computer (Fig. 1 (b)).
Despite the 3D point clouds, acquired in a range between 0.1 - 30 m, with an accuracy of approximately 3 cm, the M3 trolley delivers 360° panorama images with the possibility of virtually navigate through them afterwards. Also, these panorama images are used, during post-processing, to generate coloured point clouds, with different level of details up to 5 mm resolution (NAVVIS DEMO DATA 2017).
During the data acquisition, quality maps, as 2D floor plans, are generated for each measurement session, by using a graph-SLAM algorithm (NAVVIS M3 TROLLEY 2017). This algorithm is based not only on the 2D laser input, but also on the panoramic images, IMU, WiFi and magnetic field measurements (BASSIER et al. 2015), and also enables the M3 trolley to centimeter-accurate estimate its position within the 2D map. This feature helps the user to be aware of the space still unmapped and to improve the data quality in terms of completeness.
Even if the M3 trolley allows to maximum measure 45 minutes in a session, its 1.98 m height together with the normally low-height of the door openings is forcing the operator to pass it from room to room in compact form, which implies the need of a new calibration for the IMU sensors and so the start of a new session. Target points, previously set up, before the measurement took place, are called anchor points and are considered to be constrains for the SLAM back-end approach (GRISETTI et al 2010), enabling, in post-processing, the 2D maps optimization and the automatic registration of the point clouds. For each measurement session at least 2 anchor points have been used. Being designed especially for indoor scanning, the NavVis M3 trolley, together with the embedded algorithms, succeeded to exceed classical fixed scanner solutions in terms of efficiency and flexibility. However, it may not be such a suitable solution for staircases, very narrow spaces, or with very low ceiling, due to its construction.

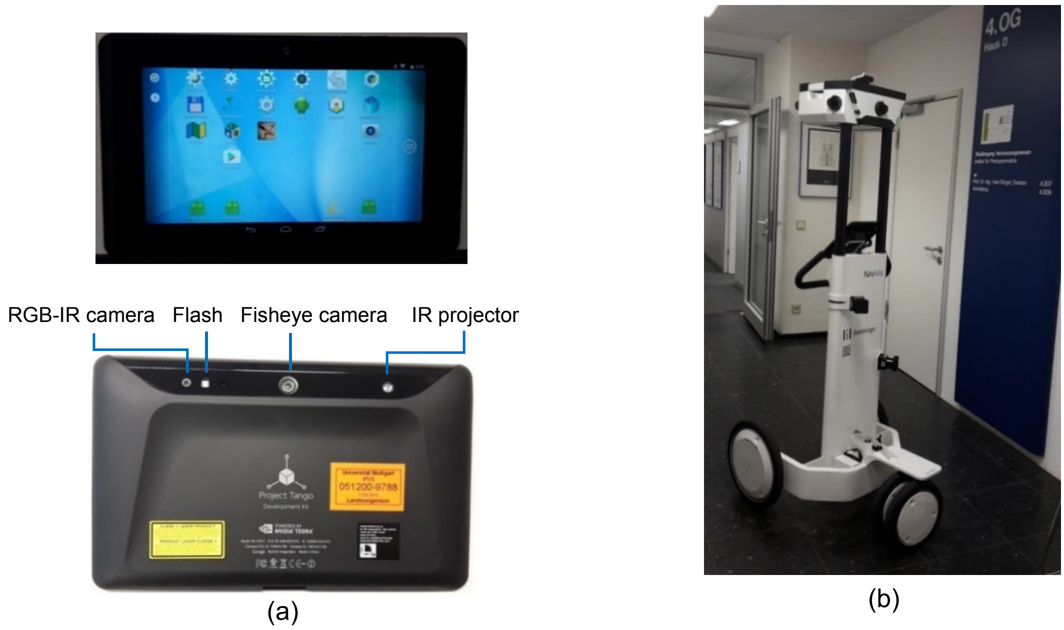



Fig. 1: (a) Google Tango tablet; (b) NavVis M3 trolley

## 4 Our Algorithm for 3D Indoor Interpretation

The goal of our approach is to automatically interpret 3D indoor point clouds. We decided to work with point clouds instead of meshes (which also implies the normal computation), in order to have a flexible algorithm, usable for a variety of indoor mapping systems, which usually deliver point clouds as a raw data. Therefore, the proposed algorithm consists of: (1) a 3D data pre-processing step and (2) a segmentation and semantic interpretation step. For enabling comparison between different measurement systems, the input data consists of point clouds of the same floor of one office building. In order to describe the algorithm, an indoor dataset acquired with Google Tango tablet is considered.

### 4.1 3D data pre-processing

The pre-processing step is preparing the 3D data for the segmentation. The floor and the ceiling are firstly removed by applying a height filtering. Considering that the errors coming with the data, acquired with low cost sensors, are strongly influencing the later processing, a sparse outlier removal algorithm is applied. This algorithm is considering for every point the mean distances to $k$ -nearest neighbours (in our case $k = 20$) and is computing their mean $\mu$ and standard deviation $\sigma$. All the points having the mean distance outside the interval $\mu \pm \alpha \cdot \sigma$ are considered outliers. The standard deviation multiplier threshold is set to be $\alpha = 1$ after multiple experiments which proved its applicability. Fig. 2 (a) and (b) are showing the effect of applying the outlier filtering. For making the further process more efficient, a downsampling method based on octree voxel grid filtering, is used. This filtering is replacing all the points bounded by one voxel with the corresponding centroid. The selected size for the voxel grid was chosen to be 5 cm, considering the overall noise of the point cloud. This way, the size of the point cloud is reduced with a factor

of 5. Some data irregularities are eliminated with the help of a resampling algorithm, which implies a Moving Least Squares surface reconstruction method, firstly introduced by LANCASTER & SALKAUSKAS (1981). This method smooths and recreates the missing parts of a point cloud by using a higher order polynomial interpolation between the neighbouring points. Through this, part of the registration errors is eliminated and also the curvature of flat surfaces is reduced by local plane fitting, like Fig. 2 (c) and (d) are showing.
The 3D data pre-processing was implemented in C++ using the Point Cloud Library (PCL) (RUSU & COUSINS 2011).

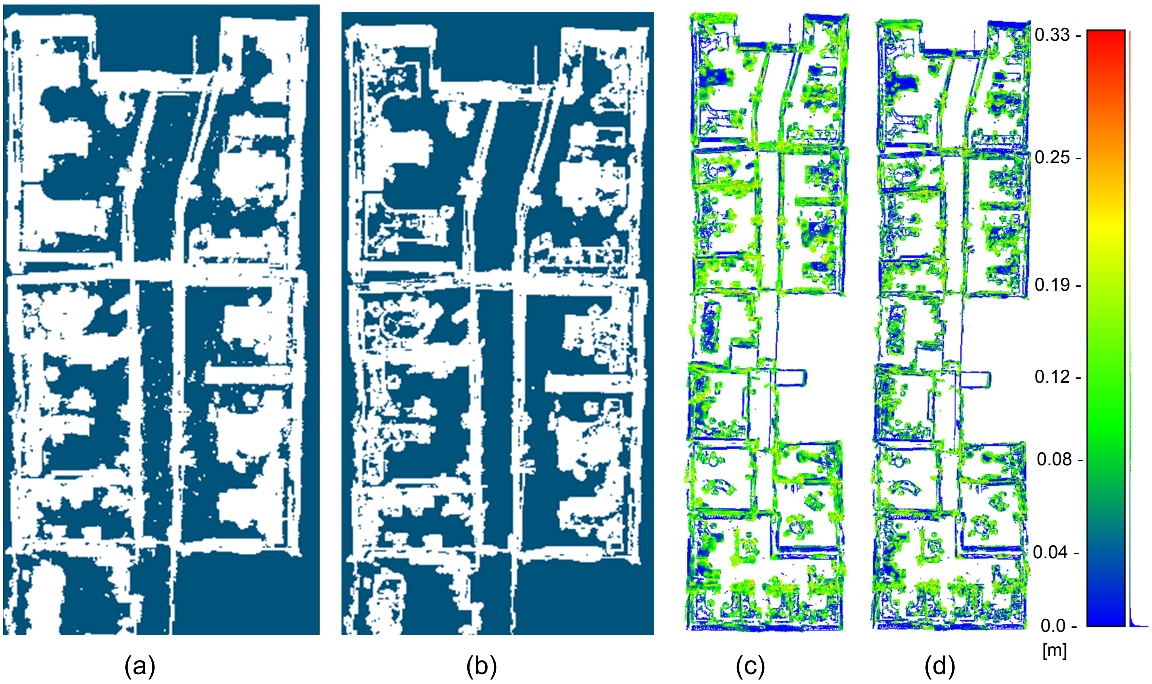



Fig. 2: (a) Original point cloud; (b) After sparse outlier removal; (c) Point cloud curvature after downsampling; (d) Point cloud curvature after Moving Least Square surface reconstruction [m]

## 4.2 Segmentation and semantic interpretation

Interpreting the 3D data semantically implies to distinguish between a large variety of indoor objects, such as walls, chairs, tables, cupboards, etc., and for this reason different algorithms are normally used, adapted to the object to be detected.
In this current work and as a first step in modelling the 3D interior, walls are aimed to be detected. Therefore, similar to KHOSRAVANI (2016), the 3D points are filtered in a height range in order to remove the furniture and to obtain the wall structure. The height thresholds should be selected according to the typical height of the furniture, which is in our case less than 1.7 m, and the ceiling decorations, which are approximately 0.5 m under the ceiling. Therefore the 3D points between 1.7 – 2.3 m were considered for further investigations. This filtering can be iteratively applied, for semantically interpreting different objects. After the walls are detected, the corresponding 3D points should be removed from the dataset and a new filtering range should be applied according to the new object aimed to be detected. For example, in a range of 0.5 m down from the previous maximum (2.3 m in this case) tall cupboards and bookshelves can be detected.

After the filtering, the points extracted in a height range between 1.7 m and 2.3 m are clustered by applying a model-based segmentation algorithm. We propose an iterative RANSAC method, which detects 3D lines describing the walls (see Fig. 3). The lines are then passed through a clipping and filtering process in order to retain only the line segments describing the reality. This is realized by following steps:

- Step 1: The filtered 3D point cloud and the 3D lines are transformed into a 2D map grid and 2D lines, respectively, by projecting them onto a horizontal plane, parallel with the floor. Each cell value represents the number of points falling into that cell. The grid cell dimensions are established according to the point sampling distance and the accuracy conditions.
- Step 2: Each projected 2D line is segmented by checking the corresponding grid values, in order to detect possible line breaks (segments of lines, which are not overlapping 3D point regions) and to save the continuous ones (Fig. 4 (a)). It is considered that the data may have interruptions, due to occlusions or low-reflectance surfaces, and for this reason, the continuous segments are accepted to have discontinuities less than 0.5 m.
- Step 3: The resulted segments are filtered by removing the ones smaller than 1 m resulted from the clipping process, as it can be seen in Fig. 4 (b).
- Step 4: The remaining segments are clustered according to their orientation and distance to the other segments (Fig. 4 (c)).
- Step 5: One line segment is fitted for each cluster in order to have the individual walls represented by individual line segments, like it is shown in Fig. 4 (d).
- Step 6: In order to divide the indoor space into rooms, the wall segments are intersected with the neighbouring ones and according to the specific situations they are extended or trimmed (Fig. 4 (e)).

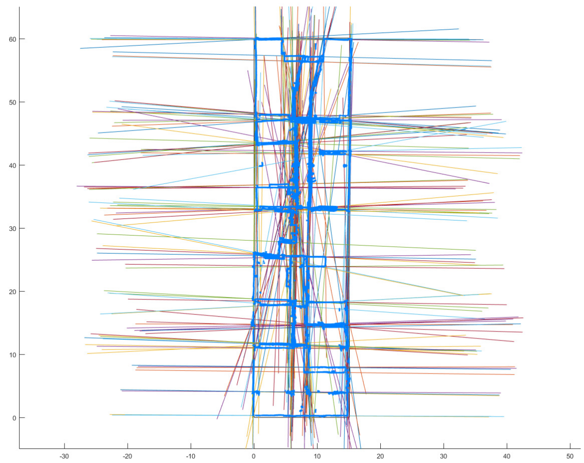
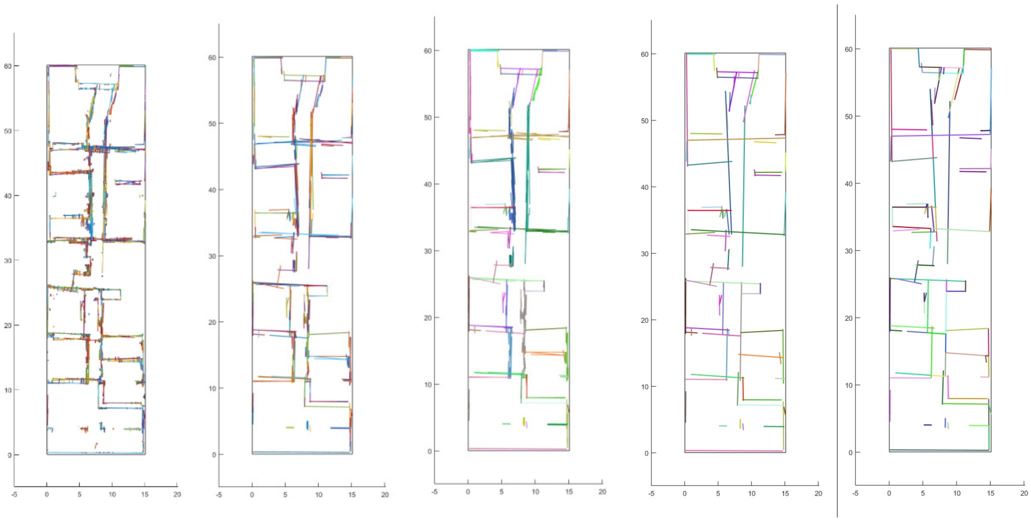



Fig. 3: 2D projection of the 3D point cloud (blue points) and the 3D lines resulting from the iterative RANSAC



(a) (b) (c) (d) (e)

Fig. 4: Line clipping and filtering process for obtaining the wall segments (random colours are associated to different lines): (a) line segmentation; (b) filtering small segments; (c) segment clusters; (d) line fitting; (e) segment intersections

The proposed algorithm is very flexible, not being restricted to the Manhattan World constraints, and accurate enough to detect the main wall structures. However, some problems, addressed in section 5, appear due to the noisy and incomplete character of the data. By imposing topological constraints, the results could be further improved.

## 5 Accuracy evaluation and comparison

The algorithm presented in section 4 is designed for detecting wall structures in 3D point clouds coming from low cost sensors. A comparison between the 3D point clouds coming from different systems, i.e. the Tango tablet and the NavVis M3 trolley, reveals the challenges which were needed to be overcome and to which extend was this succeeded.

### 5.1 Tango tablet versus NavVis M3 trolley for obtaining indoor models

For automatically interpreting 3D indoor data, it is considered an indoor dataset acquired with Google Tango tablet and it is analysed and interpreted in parallel with data, of the same indoor space, coming from NavVis M3 trolley.

Automatic point cloud registration was not subject of research of this work. However, a significant aspect is that the Tango data used for these investigations, representing the 4th floor of our office building, is composed of multiple small datasets, limited by the device overheating. Each individual dataset was captured by using simultaneous localization and mapping, integrated in the mode Area Learning provided by Google Tango (GOOGLE TANGO 2017). This enabled a relatively accurate capture of each individual indoor section. Nevertheless, their registration manually realized, depending on the overlapping regions, could have caused errors, resulting in double walls and object displacements. Even though, the NavVis M3 Trolley acquired also the data in different

sessions, for the individual rooms, the integrated graph-based SLAM algorithm (GRISETTI et al 2010) made possible the automatic registration for all the datasets.
Though both systems are delivering 3D data useful for a large variety of applications, Fig. 5 is showing their noise with respect to a reference dataset, coming from the laser scanner Leica HDS3000. In the Tango tablet case (Fig. 5 (a)), the computed absolute distances reveal noise, (up to 0.3 m) also for the flat surfaces of non-moving objects, like walls or tall cupboards, which is not the case for the M3 trolley. It can also be observed, that the non-reflecting surfaces, like windows, are causing for both datasets missing parts in the point clouds, affecting especially the further reconstruction of the exterior walls of the building.

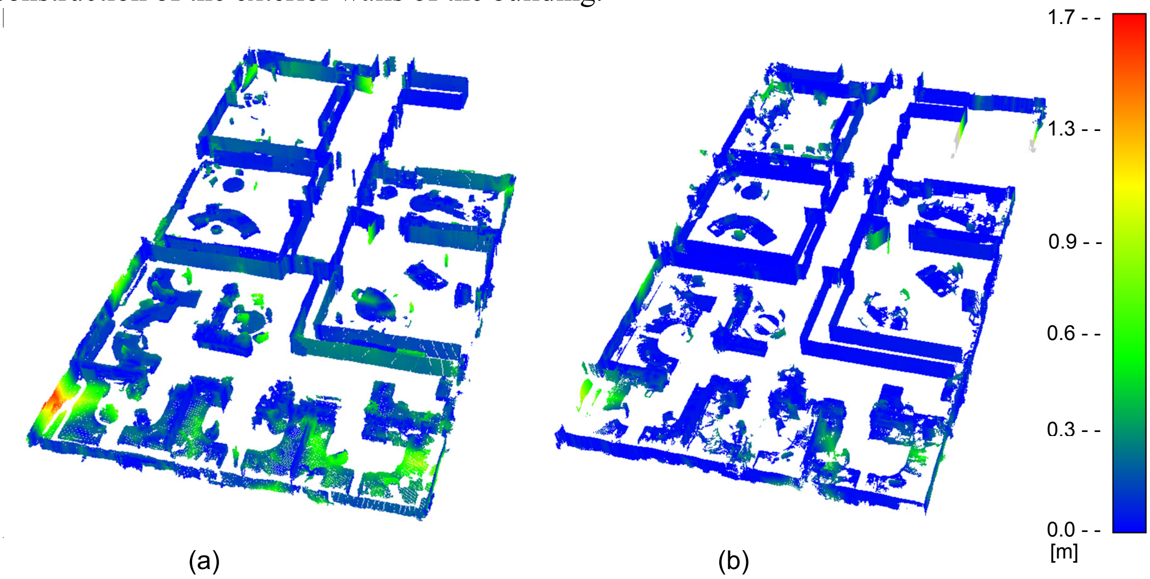



Fig. 5: Cloud to Cloud absolute distance in [m]: (a) between Tango and TLS point clouds; (b) between M3 trolley and TLS point clouds

The distance range of the sensors is a property which influences the accuracy and completeness of the acquired 3D data and therefore also the quality of the interpreted structures. By having the range distance up to 4 m, the Tango tablet is limiting the obtained point cloud to this range, while the NavVis M3 Trolley allows an acquisition up to 30 m. Therefore, for automatically detecting the wall structures in rooms with very high ceiling, one must consider during the processing that considerable big parts of those structures can be missing. This aspect was handled by the proposed algorithm from section 4, by filtering out the points over 2.3 m.

### 5.2 Quality evaluation of the interpreted wall structures

After applying the proposed algorithm, the resulted wall structures from both datasets are compared with a reference dataset coming from the laser scanner Leica HDS3000, acquired a couple of years ago. In this way it was possible to realize the influence of data quality onto the resulted interpreted structures. In Fig. 6, it is shown the 2D projection of the corresponding 3D point cloud, as blue points, and the detected wall segments, as red lines. It is considered that a wall is detected only when the corresponding segment is complete, from one wall junction to the other. Thus, 44 from 55 segments (80%) were detected for the Tango tablet dataset, and 47 from 55 segments (85.5%) for the M3 trolley. However, the main majority of the undetected wall segments are located on the building shell, being mainly made of windows. As it was previously shown in section 5.1., non-reflecting surfaces, like windows are affecting the reconstruction process. This issue can be overcome by adding some knowledge, like the ground plan of the building. Furthermore, some wall segments could be misplaced due to some object parts, different than walls, still remained after the pre-processing step.

Both datasets made possible to automatically detect the main wall structures, but in some cases they were shifted from the reference position. For example, Fig. 7 shows a closer detail of Fig. 6, where, in the case of Tango dataset, the wall segments are shifted in centimetre range. This is visible only if they are compared with a reference dataset, coming from TLS. Thus, it is proved that the displacement is not influenced by the proposed algorithm, but of the noisy character of the data and possible remaining registration errors.

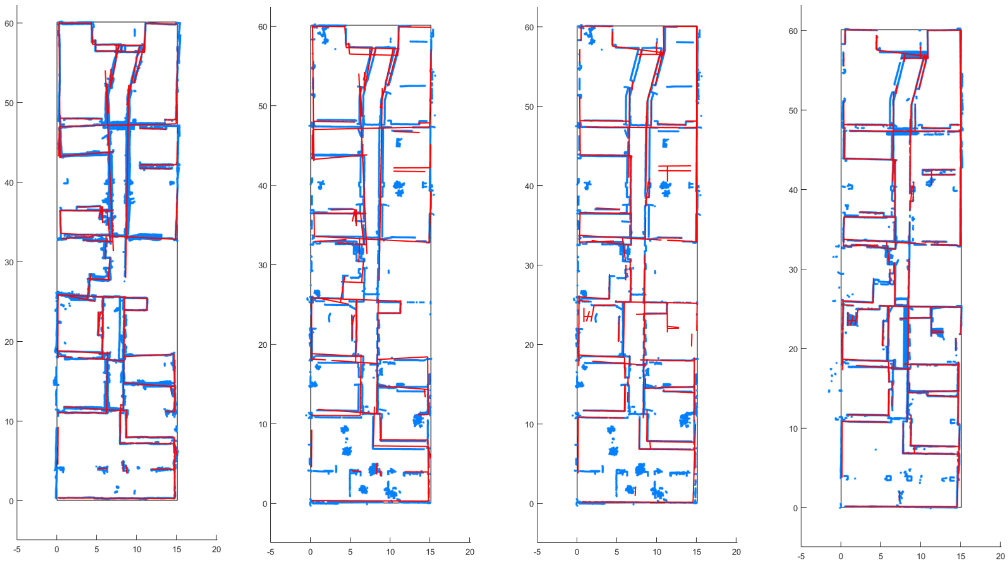


Fig. 6: Detected wall segments from: (a)Tango tablet and Tango tablet point cloud; (b) Tango tablet and TLS point cloud; (c)M3 trolley and TLS point cloud; (b) M3 trolley point cloud and M3 trolley point cloud

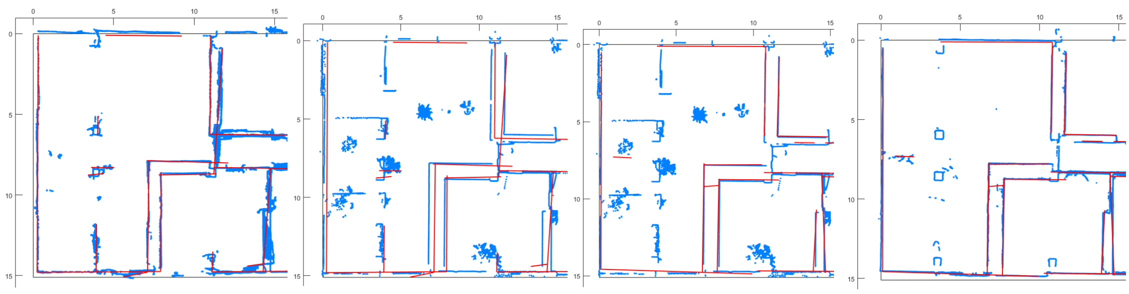


Fig. 7: Detail of detected wall segments from: (a)Tango tablet and Tango tablet point cloud; (b) Tango tablet and TLS point cloud; (c)M3 trolley and TLS point cloud; (b) M3 trolley point cloud and M3 trolley point cloud

## 6 Conclusions

We presented an evaluation of an IMMS system (NavVis M3 trolley) and a low-cost device (Google Tango tablet). Also, it was investigated their potential for delivering interpreted indoor structures. For doing that, we proposed a wall extraction algorithm, which proved to be robust and flexible, not being restricted to Manhattan World constraints. By comparing the results with a TLS reference dataset, it was proved that the main wall structures were detected, the missing ones being corresponding for surfaces which do not reflect IR light, like windows. Only few detected segments, coming from Tango tablet dataset have been displaced in centimetre range, due to possible registration errors. On the one side, there are applications which require accurate indoor models, like mapping indoors of industrial facilities and for this, a system like NavVis M3 trolley will be suitable. On the other side, for some other applications, like indoor navigation, a mobile device with Google Tango's capabilities will successfully perform the task. Furthermore, a way to overcome these possible displacement problems is to use a formal grammar for modeling the indoor environment, which is a subject of further work.

## 7 Acknowledgements

This work is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), grants FR 823/25-3 and RO 1086/17-3, and the European Social Fund (ESF) as well as the Ministry Of Science, Research and the Arts Baden-Württemberg.

## 8 References

BASSIER, M., YOUSEFZADEH, M. & VAN GENECHTEN, B., 2015: Evaluation of data acquisition techniques and workflows for Scan to BIM. Proceedings of Geo Bussiness.

BESL, P. J. & MCKAY, N. D., 1992: Method for registration of 3-D shapes. Robotics-DL Tentative, 586-606.

BOEHM, J., 2014: Accuracy investigation for structured-light based consumer 3d sensors. Photogrammetrie Fernerkundung Geoinformation, **2014**(2), 117-127.

BOSSE, M., ZLOT, R. & FLICK, P., 2012: Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping. IEEE Trans. Robot. **28**(5), 1104-1119.

BUDRONI, A. & BÖHM, J., 2009: Toward Automated Reconstruction of Interiors from Laser Data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38**(5).

DIAKITÉ, A. A. & ZLATANOVA, S., 2016: First experiments with the Tango tablet for indoor scanning. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **III-4**, 67-72.

DOTPRODUCT DPI-8, 2017: https://www.dotproduct3d.com/DPI8.php, last access 16.01.2017.

FOFI, D., SLIWA, T. & VOISIN, Y., 2004: A comparative survey on invisible structured light. Electronic Imaging 2004, International Society for Optics and Photonics, 90-98.

GEOSLAM, 2017: http://geoslam.com/hardware-products/zeb1/, last access 16.01.2017.

GOOGLE TANGO, 2016a: https://developers.google.com/tango/overview/depth-perception (last updated 07.12.2016), last access 16.01.2017.

GOOGLE TANGO, 2016b: https://developers.google.com/tango/overview/motion-tracking (last updated 10.06.2016), last access 16.01.2017.

GOOGLE TANGO, 2017: https://developers.google.com/tango/overview/area-learning (last updated 05.01.2017), last access 16.01.2017.

GRISETTI, G., KUMMERLE, R., STACHNISS, C. & BURGARD, W., 2010: A tutorial on graph-based SLAM. IEEE Intelligent Transportation Systems Magazine **2**(4), 31-43.

JENKE, P., HUHLE, B. & STRASSER, W., 2009: Statistical reconstruction of indoor scenes. In Proc. of 17. Int. Conf. in Central Europe on Comp. Graphics, Vis. and Comp. Vision (WSCG '09).

KHOSRAVANI, A. M., 2016: Automatic Modeling of Building Interiors Using Low-Cost Sensor Systems. Deutsche Geodätische Kommission, Reihe C, Nr. **767**, München.

KURIAN, A. & MORIN, K. W., 2016: A fast and flexible method for meta-map building for ICP based SLAM. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLI-B3, 273-278.

LANCASTER, P. & SALKAUSKAS, K., 1981: Surfaces generated by moving least squares methods. Mathematics of computation **37**, 141-158.

LASKAR, Z., HUTTUNEN, S., HERRERA, D., RAHTU, C. E. & KANNALA, J., 2016: Robust loop closures for scene reconstruction by combining odometry and visual correspondences. IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2603-2607.

LEONARD, J. J. & DURRANT-WHYTE, H. F., 1991: Mobile robot localization by tracking geometric beacons. IEEE Transactions on Robotics and Automation **7**, 376-382.

LI, M. & MOURIKIS, A. I., 2012: Improving the accuracy of EKF-based visual-inertial odometry. IEEE International Conference on Robotics and Automation (ICRA), 828-835.

MSDN KINECT, 2017: https://msdn.microsoft.com/en-us/library/jj131033.aspx, last access 16.01.2017.

NAVVIS DEMO DATA, 2017: http://www.navvis.com/products/demodata/, last access 16.01.2017.

NAVVIS M3 TROLLEY, 2017: http://www.navvis.com/products/m3-trolley/, last access 16.01.2017.

NEWCOMBE, R. A., DAVISON, A. J., IZADI, S., KOHLI, P., HILLIGES, O., SHOTTON, J., MOLYNEAUX, D., HODGES, S., KIM, D. & FITZGIBBON, A., 2011: KinectFusion: Real-time dense surface mapping and tracking. 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 127-136.

OKORN, B., XIONG, X., AKINCI, B. & HUBER, D., 2010: Toward automated modeling of floor plans. Proceedings of the symposium on 3D data processing, visualization and transmission **2**.

PREVITALI, M., BARAZZETTI, L., BRUMANA, R. & SCAIONI, M., 2014: Towards automatic indoor reconstruction of cluttered building rooms from point clouds. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **II**-5, 281-288.

RUSU, R. B. & COUSINS, S., 2011: 3D is here: Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation (ICRA), 1-4.

THOMSON, C., APOSTOLOPOULOS, G., BACKES, D. & BOEHM, J., 2013: Mobile Laser Scanning for Indoor Modelling. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **II-5**/W2, 289-293.

THRUN, S. & LEONARD, J. J., 2008: Simultaneous localization and mapping. Springer handbook of robotics, Springer, 871-889.

VALERO, E., ADÁN, A. & CERRADA, C., 2012: Automatic method for building indoor boundary models from dense point clouds collected by laser scanners. Sensors **12**, 16099-16115.

VIAMETRIS IMS 3D, 2017: http://viametris.info/iMS3D/EN/, last access 16.01.2017.

WIKIPEDIA TANGO, 2017: https://en.wikipedia.org/wiki/Tango (platform), last access 16.01.2017.