From Design To Demolition: Robotics Holds a Promise to Revolutionize Every Stage Of A Building's Life Cycle

There is little doubt that the construction business is evolving through building information modeling (BIM [1]). Having a true digital representation of a building as it is being built or renovated to compare to the design CAD allows for better design, planning and collaboration as well as avoiding costly errors. The virtualization of operations is improving efficiency to the point of being rightfully called a revolution [2].

As soon as the proverbial cornerstone is laid, the building starts to deviate, to some degree, from the design intent captured in its digital form as a CAD model. Within generally accepted tolerances, those deviations are forgivable; however, they can still amount to important and unnecessary costs or lost revenue. For example, the consequences of even a slight dimensioning error can be eye watering when constructing luxury condominiums [3]. Furthermore, if important errors remains undetected until late in construction, the building's structure might be compromised, making corrections very expensive.

To avoid this very problem (the late detection of the deviation from design intent or nominal values), some industries adopted metrology as a cornerstone of their production process. For example, during production, car frames and doors are routinely inspected before installation to insure a good fit, harmonizing the production chain, reducing delays, increasing productivity and normalizing costs. Borrowing from the automotive industry, the construction and renovation processes should continuously monitor and measure the building being built, updating the BIM accordingly to avoid unnecessary costs and loss of productivity.

Design and construction are, however, only the very first steps of the building's life cycle. Later stages such as operations, maintenance, renovation, repurposing, energy efficiency and even demolition and disposal could benefit just as much from BIM as the construction, especially if some degree of autonomous response is provided.

For example, once built and in operation, the building starts to deviate from its initial state : partitioning changes might be necessary to accommodate new tenants, damages caused by structural strains, water infiltrations and others will accumulate, security risks due to changes might appear and so on. All of those changes are of obvious interest to insurance companies and should be (and currently only sometimes are) thoroughly documented. Indicators of building health could also be precisely monitored: the air quality, noise levels, cleanliness, elevator inspections, presence and good functioning of security equipment (defibrillators, first aid kits, fire extinguishers) and level of consumables (from light bulbs to soap in bathrooms) are only a few examples.

Accomplishing all of those tasks require constant attention and updates to the BIM. In short, it is useful only to the degree it represents reality faithfully; otherwise, BIM might do the opposite of its intent and cause costly errors and damages rather than operational excellence. Traditionally, there are two ways of providing relevant data : fixed sensors and manual inspections.

Fixed sensors (eg. surveillance cameras, movement sensors, smoke detectors, etc.) are usually deployed in relatively small numbers in the most critical parts of the building to gather live data. One reason behind this sparseness is cost : in addition to buying standalone sensors, their installation, wiring and maintenance quickly add up to an expensive system. Moreover, complexity rapidly explodes, creating an important number of failure points. Installing a large network of fixed sensors to collect dense information throughout the building is thus unreliable and prohibitively expensive regardless whether it is wired or wireless system.

Manual inspections involve a human patrolling the area to measure with specialized equipment. The types of data that can be gathered are much more diverse than with static sensors : some examples are thermal and 3D geometry scanning, structure and elevator inspections, equipment inventory, security rounds and reading gauges that are not connected to the building numerical system. Creating a dense and complete data set is no problem either : a human can be as thorough as the task requires.

However, performing such inspections is expensive in both manpower and money. Thus, while daily manual inspections are beneficial, the cost would quickly pass beyond the profitability point of the building, leading to infrequent (at best) monitoring. Moreover, the specialized equipment used can be costly (e.g. good quality 3D scanners currently are beyond most means). For large projects, the sporadic use of such equipment (and required specialist workforce) is acceptable and profitable. However, despite having the same needs, small contractors and homeowners will find such equipment financially out of reach.

In short, both fixed sensors and manual inspections cannot provide at an economically acceptable cost the complete up-to-date information required to keep BIM truly useful throughout the life of a building. We believe that robotics will provide the answer, combining the frequency of readings from fixed sensors to the versatility and thoroughness of manual scanning all the while being cheap and simple to operate.

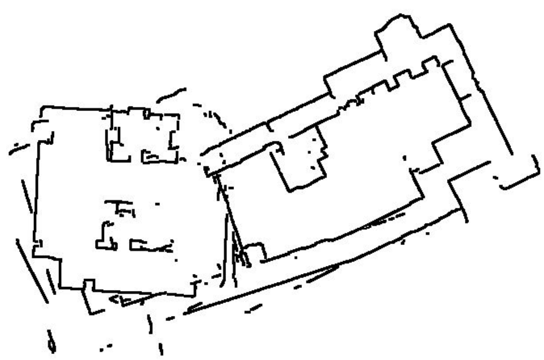
[1] Wikipedia, Building Information Modeling.
https://en.wikipedia.org/wiki/Building_information_modeling, May 2019
[2] Renata Barradas Gutierrez, The Digital Revolution in Construction.
http://blog.hexagongeosystems.com/the-digital-revolution-in-construction/, April 2017.
[3] Excelsior Measuring, About Lidar Scanning. https://excelsiorlevel.com/sample-page/lidar/.

Cost-Effective Map Building

Introduction to SLAM

Every homeowner, building manager and construction foreman should have precise, detailed and up-to-date maps of their buildings. Due to cost and manpower constraints, we believe that autonomous robots will be the cornerstone in creating those maps [1]. To do this, the computational problem of modelling an unknown environment and tracking the robot's position within it at the same time must be solved; this problem is usually referred to as SLAM (simultaneous localisation and mapping).

SLAM is something of a chicken-or-the-egg problem. First, the robot needs a perfect map to precisely track its position using visual sensors such as cameras, sonars, lidars, etc. Second, the robot needs to know it's position and orientation perfectly to add new information correctly to the map it is building. Any imprecision in one of those two steps will cascade down to the other, resulting in an increasingly inaccurate map and robot pose. The Figure 1 illustrates this phenomenon.



(a) (b)

Figure 1 : a) Error accumulation during the movement of the robot leads to a progressive drift of the modeled map. b) Using proper SLAM techniques, the same dataset can be used to create accurate maps.

There are two components to the mapping system: the sensors used to gather information and the SLAM algorithm itself. Both of those components can be optimized to minimize error accumulation, but we'll concentrate here on the former.

Better sensors mean that the robot position and orientation can be better estimated as it moves along in both dead reckoning (inertial navigation, akin to walking with your eyes closed) and with feedback from the environment (comparing the sensor's readings to the map). Good visual sensors also mean that the sensor information added to the map is more precise, leading to a virtuous circle of better pose estimation, and thus better mapping.

Metrologic Mapping Requirements

Knowing that high quality sensors are important to the SLAM process and that we intend our robots to provide building information modeling (BIM) to a large array of customers and in a variety of environmental conditions, our sensor and algorithm selection needs to take into account the following constraints :

- **Accuracy:** BIM requires maps of metrological quality; as such, our goal is creating maps accurate within 0.1% of the true distances.
- **Cost:** In a business to business (B2B) setting, the cost of a metrology device is more forgiving than when selling to consumers (B2C). Ideally however, a single platform should be usable for both. We thus believe that a bare bones mapping robot should cost below $500.
- **Ambient light:** Robots should be able to work in the dark: the building in which the robots operate might be unfinished, without electricity or simply unoccupied at night. Conversely, robots should also be able to work in places in which only ambient light is available (such as unoccupied buildings during the day). This last environment is surprisingly harsh, the luminosity varying greatly from one spot to the next and direct sunlight being far brighter (roughly 1000x [5]) than common office lighting.
- **Range:** In commercial spaces, 30 meters is the maximal expected distance between the robot and every directly observable surfaces from its position.
- **Computational complexity:** Regardless the size of the environment and the amount of data collected, computations should not slow down as the amount of data or the size of the map increases.
- **Power consumption:** The robot should be able to map 10’000 m² (the size of an average department store [2]) on a single charge. Minimisation of the power requirements allows greater autonomy or using a smaller battery pack.
- **Reliability:** The robot design should minimize the number of fragile or mobile components.

Choosing a visual sensor

Visual sensors are the backbone of a robot’s ability to produce good quality maps; as discussed previously, they directly affect the quality of the SLAM as they gather the data added to the map and indirectly as they help track the robot’s position. Choosing the right sensor is thus critical to achieve a metrological robot. There are many options when it comes to sensing devices that might be used for SLAM. Those options are usually split into two categories : passive and active sensors.

Passive sensors simply “take in” the environment; some examples extracting 3D information from a scene are passive stereo and feature tracking cameras. Advantages of this class of sensor include their relative physical simplicity, low power consumption and low cost.

Unfortunately, since one of our requirements is to be able to function at long ranges in complete darkness, such sensors are unsuitable as they rely on ambient visible light.

Active sensors, such as active stereo (e.g. structured light sensors), lidars, sonars and radars, produce easily identifiable signals and then measure their reflection on the environment. Obviously, this self-illumination solves the darkness constraint; however, sensors of this class suffers from other problems.

For instance, stereo cameras intrinsically have a very limited range due to their error growing exponentially with the distance [4]; the Figure 2 illustrates the phenomenon for a camera configuration similar to the Intel Realsense series [5]. From this, it is clear that at the required range of 30 m, such a system would incur significant imprecisions. Lidars generally do not suffer from this problem : given sufficient return signal strength, they have a constant error regardless of the measured distance (except for very short ranges).

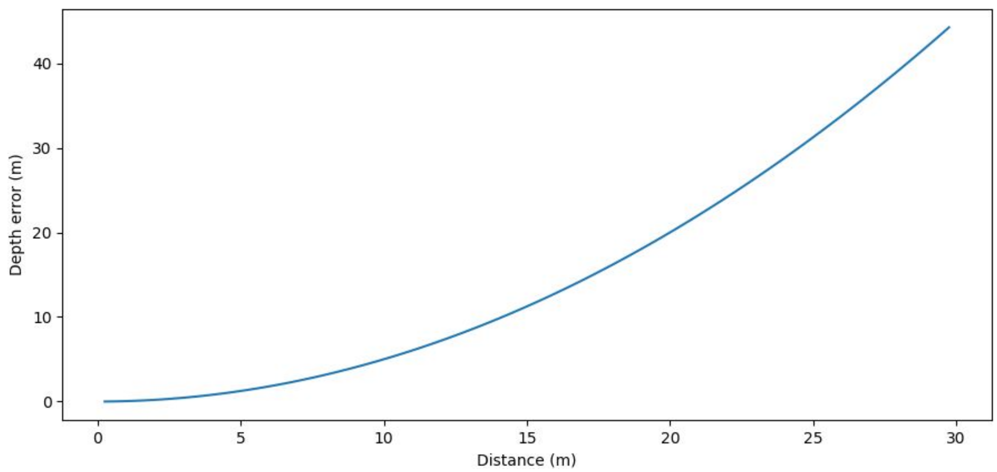


Figure 2 : Error buildup of a stereoscopic camera setup with baseline of 5 cm and resolution of 1280 x 1024 as a function of the distance of observed feature

A flash lidar would therefore seem like a natural choice for our application. However, they present a power consumption problem : the power of light arriving at the surface pointed at by a lidar is inversely proportional to the distance between the two squared (unless the light is collimated). This makes lightning the whole scene at 30 m distance an incredibly power hungry task, which is unacceptable given battery costs.

We therefore limit ourselves to a single laser beam sensor that uses time of flight (ToF) technology. There seems to be many options when it comes to this technology due to recent developments in the autonomous automotive industry [6]. Indeed, many of the industry players believe that lidars is the way to go, leading to massive investments in several dozen tech companies. Low cost, indoor solutions, on the other hand, are fewer. A typical example of a rotating single beam sensor are Hokyo sensors, their UTM-30LN [7] model costing over $6,000 for the required range. Rotating the lidar along 2 axis to achieve 3D scanning, Leica’s BLK360 [8] is currently their entry level metrological 3D scanner at the prohibitive cost of roughly

$16,000. At the other end of the price spectrum, sensors like the $100 RPLidar A1M8 [9] are too short range for our needs and raises serious durability concerns. This leaves us with a seemingly impossible task of finding the sensor that fits all our criteria. It turns out, however, that there is a satisfactory off the shelf solution: LIDAR-Lite [10], a single point ToF sensor produced by PulsedLight (recently acquired by Garmin).

The LIDAR-Lite is marketed as a ranging solution for drone, robot, or unmanned vehicle applications. With a retail price of $150, it sports the following specifications:

Table 1 : LIDAR-Lite v3HP Specifications

Current consumption	85 mA during an acquisition
Range (70% reflective target)	40 m (131 ft)
Resolution	+/- 1 cm (0.4 in.)
Accuracy < 2 m	±5 cm (2 in.) typical*
Accuracy ≥ 2 m	±2.5 cm (1 in.) typical
Mean	±1% of distance maximum
Ripple	±1% of distance maximum
Update rate (70% Reflective Target)	Greater than 1 kHz typical
*Nonlinearity present below 1 m (39.4 in.)	

Using LIDAR-Lite to create precise maps

Using the LIDAR-Lite for our application (autonomous robotic mapping) might not be an obvious choice. After all, being a single point fixed sensor, it lacks the ability to produce scans (i.e. contours or curves of the environment around the sensor). The lack of precision is also obvious : with measurements having a bias of 1% off the mark (*mean*) and the measurement noise being equivalent to 1% of the distance measured (*ripple*), this sensor does not attain our metrological grade objective of 0.1% error out of the box. Nevertheless, software can correct what the LIDAR-Lite lacks in hardware capabilities.

The first missing element, the scanning capability, can be easily provided by the robot. Indeed, making the robot turn on itself around the focal point of the sensor provides the contour of the visible environment. To achieve this, the robot provides angular position, time-synchronized with the sensor's distance reading. In addition to the angular position synchronization, our proprietary auto-calibrating technology allows for accurate reading without knowing the exact angular speed. Overall, this technique provides us with a robustness and significant cost advantage over mounting the sensor on a standalone motor.

The second missing element is accuracy: the 1% error bias and 1% noise standard deviation are more than an order of magnitude worse than metrology grade accuracy. Our processing algorithms solves this through the integration (averaging) of data. To illustrate how this works (Figure 3), consider multiple measurements of the same point corrupted with an arbitrarily large noise. If the error is of zero-mean (i.e. the mean of an infinite number of samples stands on the true value), then the true measure can be obtained with arbitrary accuracy by integrating a sufficient number of samples.

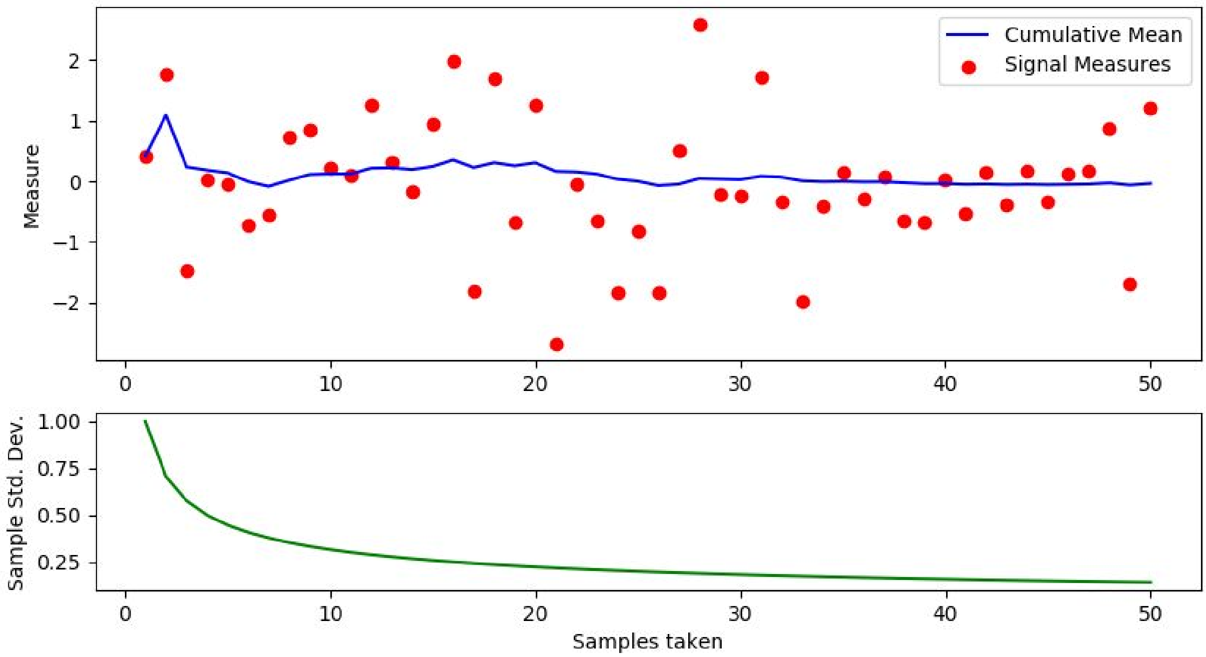


Figure 3 : Effect of averaging data on measurement quality

Therefore, an important step is to calibrate the LIDAR-Lite to make it a zero-mean error sensor. This is accomplished via a reference range finder (in our case, a slow but precise Bosch 50cx rangefinder) mounted on the robot. By capturing the simultaneous measurements made by the two sensors for a range of distances (0.25 m – 25 m) and then using a polynomial model to represent the difference between the two sensors, the LIDAR-Lite bias can be corrected. Other than this calibration, only a shape preserving filtering is applied to the LIDAR-Lite data to make it ready for SLAM.

The process described above reduces the noise found in individual LIDAR-Lite measurements; it is however not the only method used to smooth out noisy data. First, as described previously, individual range measurements are organised into 360° scans of the environment through the movement of the robot. Knowing the relationship between measurements, it is possible to filter out noise. Second, when two such scans overlap, the overlap can be used to refine the estimated pose of the robot during the scans (effectively eliminating the effect of noise on it). The overlap region can also be averaged, further reducing the noise in the map model. Consequently, this allows us to incrementally build the map whilst keeping the error low.

Results

The procedure described above produces very impressive results for a very low cost autonomous mapping robot. The accuracy obtained by measuring a part of a multi-room environment of approximately 30 m x 20 m is up to 0.006% of the longest measured distance, 0.085% in average and 0.147% in the worst case. The results illustrated in Figure 4 and tabulated in Table 2 were obtained by using three different robots for a total of 8 times at random starting locations. The robots mapped the space without human intervention.

An interesting feature of those trials was the discovery of inaccuracies in the official CAD of the office space in Figure 4. One flagrant example is the rightmost wall of the open space : a double wall was installed during construction without any updates to the plans, reducing the advertised space length by about 30 cm (this was confirmed using the precise Bosch 50cx rangefinder). Other inaccuracies stem from furniture occupying the space; unfortunately, since the described robot setup only produces 2D curves, those items are impossible to distinguish from walls. This is where inexpensive 3D cameras could come to the rescue.

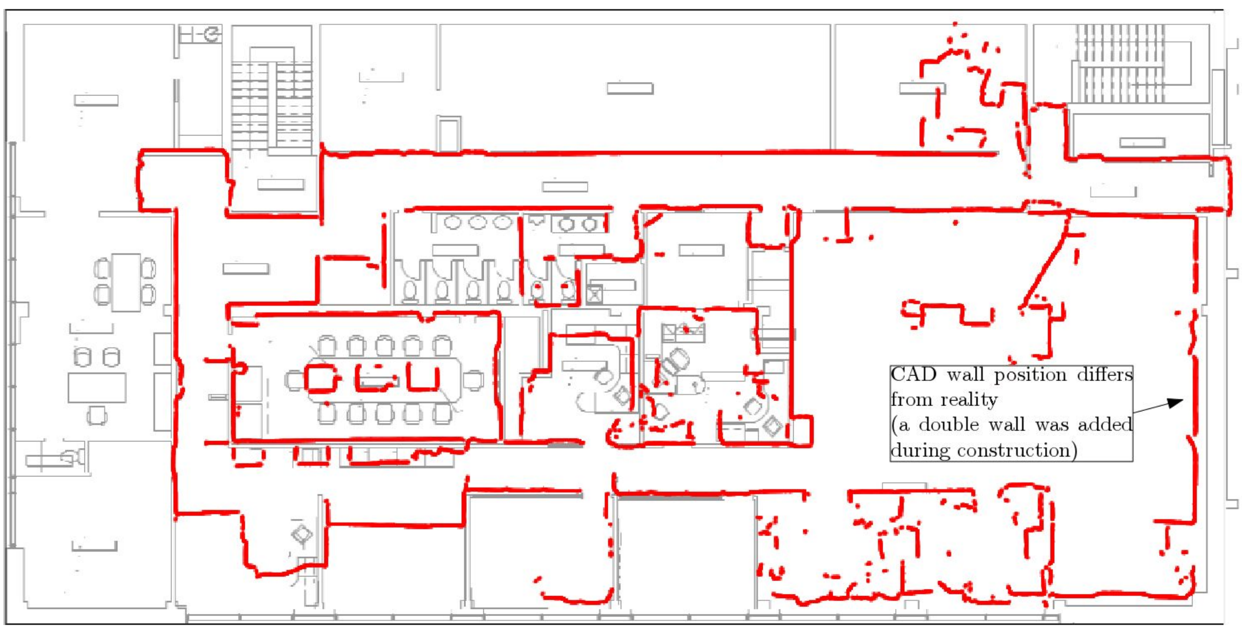


Figure 4 : Closeup example of a map created by the robot superimposed over the CAD of the building.

Table 2 : Longest distance measurement error made by the robots during the mapping phase

Sample	Distance (mm)	Error (mm)	Error (%)
Ground Truth	32 606		
Robot 1, Trial 1	32 651	45	0.137
Robot 1, Trial 2	32 654	48	0.147
Robot 1, Trial 3	32 575	31	0.095
Robot 2, Trial 1	32 633	27	0.081
Robot 2, Trial 2	32 629	23	0.070
Robot 2, Trial 3	32 645	39	0.119
Robot 3, Trial 1	32 608	2	0.006
Robot 3, Trial 2	32 598	8	0.024
Average	32 624	28	0.085

References

[1] LINK TO THE FIRST WHITE PAPER
[2] Wikipedia, Walmart. https://en.wikipedia.org/wiki/Walmart, May 2019.
[3] Wikipedia, Lux. https://en.wikipedia.org/wiki/Lux, May 2019.
[4] Rajesh Rao. *CSE 455 : Computer Vision / Lecture 16 - Stereo and 3D Vision.* https://courses.cs.washington.edu/courses/cse455/09wi/Lects/lect16.pdf, September 2019.
[5] Intel, Intel Realsense product page. https://www.intelrealsense.com/stereo-depth.
[6] Timothy B. Lee, How 10 companies are trying to make powerful, low-cost lidar. https://arstechnica.com/cars/2019/02/the-ars-technica-guide-to-the-lidar-industry, February 2019.
[7] Hokuyo, UTM-30LX product page. https://www.hokuyo-usa.com/products/scanning-laser-rangefinders/utm-30lx.
[8] Leica, BLK360 product page. https://leica-geosystems.com/products/laser-scanners/scanners/blk360.
[9] Slamtec, RPLidar A1 product page. http://www.slamtec.com/en/lidar/a1.
[10] Garmin, LIDAR-Lite v3HP product page. https://buy.garmin.com/en-US/US/p/578152/pn/010-01722-10.

Filling in the 3D Details

As mentioned previously [1], we strongly believe that giving a robot the ability to autonomously map its environment provides valuable information for Building Information Modeling (BIM). For example, our first tests [8] successfully detected out of date building plans, a common occurrence following remodelling (Figure 1).

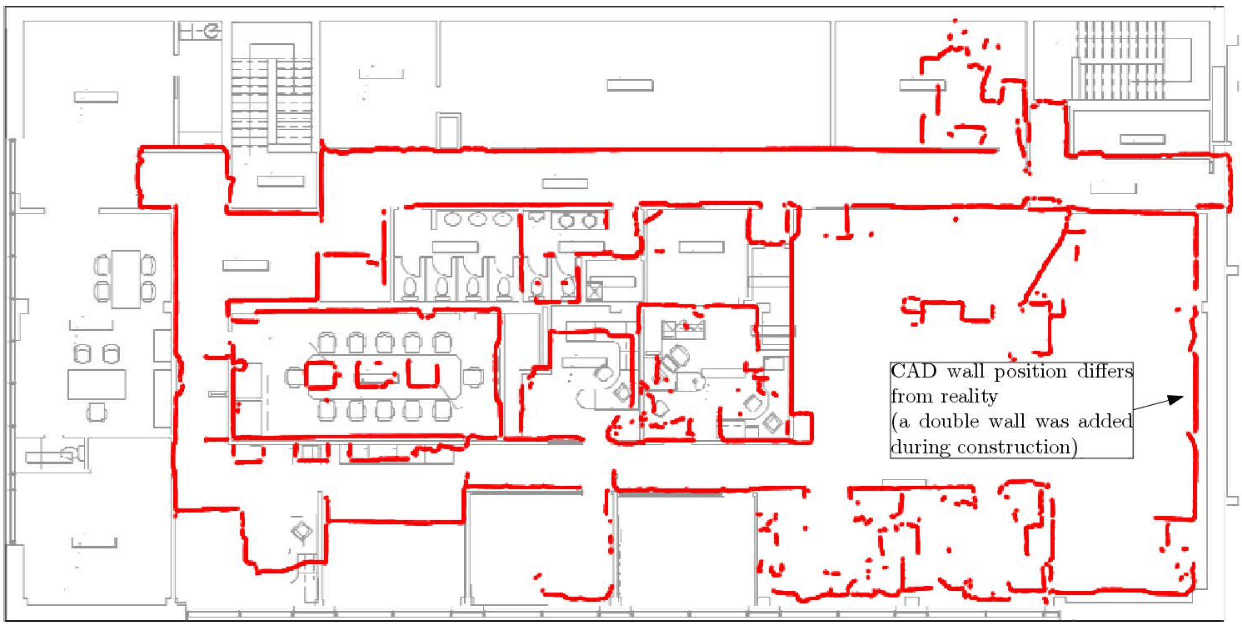


Figure 1 : Building's CAD error detection through robot 2D scanning

However impressive, the displayed map is still limited; while it is internally consistent, fine details could not be properly captured using a bare-bones robot equipped only with an inexpensive lidar. Moreover, the data produced is a horizontal slice of the world at the height of the robot (~20 cm high), which makes it hard to understand for a human used to living high up (~175 cm). Both of those problems can be fixed using 3D scanning (Figure 2).

Figure 2 : 3D scan of a building’s interior

Traditionally, a full 3D scan of a building's interior involves a human lugging around a 3D scanner throughout its rooms [2]; this process turns out to be very expensive for a variety of reasons. First, in order to achieve metrological quality, the used scanner has to be of metrological quality itself : fast, accurate and long range, hence expensive. Second, humans are needed to create the scans : to select the best scanning spots, to move the equipment, to launch and validate scans and finally to stitch together the scans into a cohesive map. A lot of those steps require specialist knowledge, which is also expensive. Finally, to limit the time (and thus expenditures) required to complete operations, only the bare minimum of scans required to cover the space are performed. This can lead to underwhelming results as details are missed.

An alternative method that we support relies on short range, low cost scanners. Sensors in this category can be exceptionally precise : for example, at 30 cm standoff a small baseline stereo pair can achieve micrometer accuracy [3]. The problems with using them to scan buildings are twofold :

1. They are shortsighted, thus the number of scans necessary to completely model a building is an order of magnitude greater than with typical large 3D sensors. While the amount of work and precision required might make such an approach impractical for humans, robots are well suited to it.
2. The maps they produce are imprecise at long distances, meaning that weaving individual scans into a cohesive whole without accumulating errors is very hard. The traditional

ways to correct this use scale bars or other calibration methods [7]; in addition to requiring extra material, this presents the disadvantage of being time consuming. Luckily, the whole error accumulation problem can be entirely avoided by precisely positioning the sensor within an exact 2D map such as the one in Figure 1 (even if this map lacks fine details).

One such short range scanner is the Intel Realsense D435 [4], an active stereo camera that projects a pseudo-random light pattern onto the scene in order to recover 3D point cloud using the triangulation principle. While its theoretical maximum range is 10 m, this sensor suffers from the typical stereo camera exponential accuracy drop with measured distance [6]. Indeed, at 4 m, the expected reading's RMS error is about 60 mm (or 180 mm peak-to-peak). Worse, in practice, this error can be considerably higher depending on lighting conditions and the type of surface being measured. Since we aim to produce metrologic quality models (accurate within 0.1% of the true value), the Intel Realsense D435 might appear to be a poor choice; as we'll discuss in an upcoming paper, its shortcomings can be corrected through the smart use of already available data.

For the moment, to illustrate the viability of short range and low cost 3D scanning, we mounted a RealSense D435 camera on a robot in addition to a LIDAR-Lite sensor [5] for long range positioning. The procedure used to generate a full 3D scan of a building is simple. First, an accurate (but lacking in details) outline of the space is built using the technique described in our previous paper [8]. Using this contour, the robot automatically generates a map refinement and scanning trajectory (Figure 3). At each stop along this path, the robot captures a panorama of colored 3D points. By progressively stitching those panorama together, a full model is built (Figure 4). No human is involved in the whole process (other than to click “Go” and open doors that is).

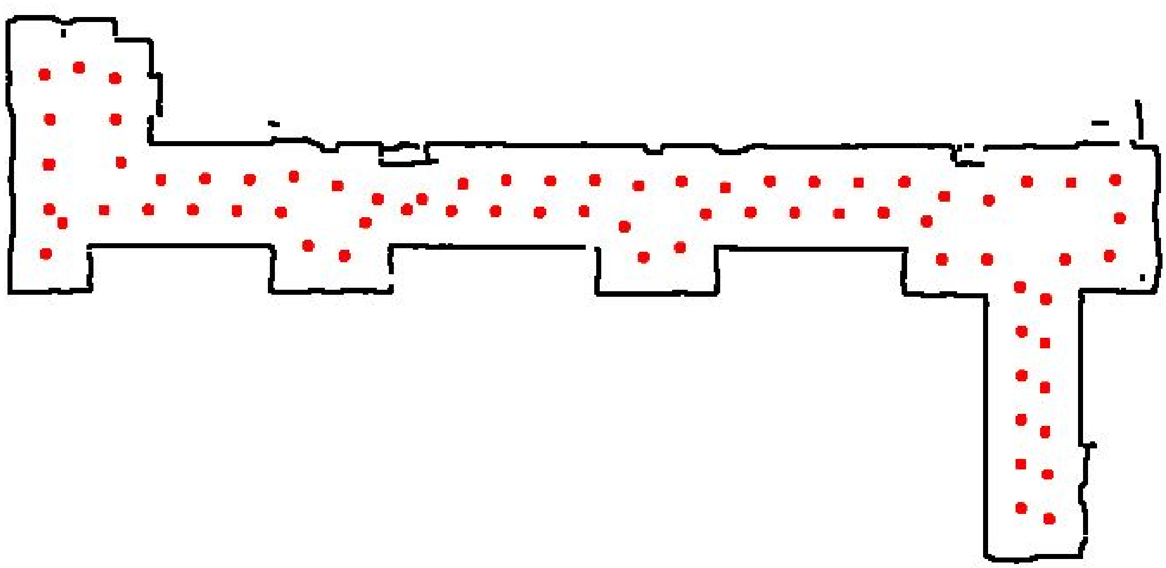


Figure 3 : Automatically generated map refinement positions

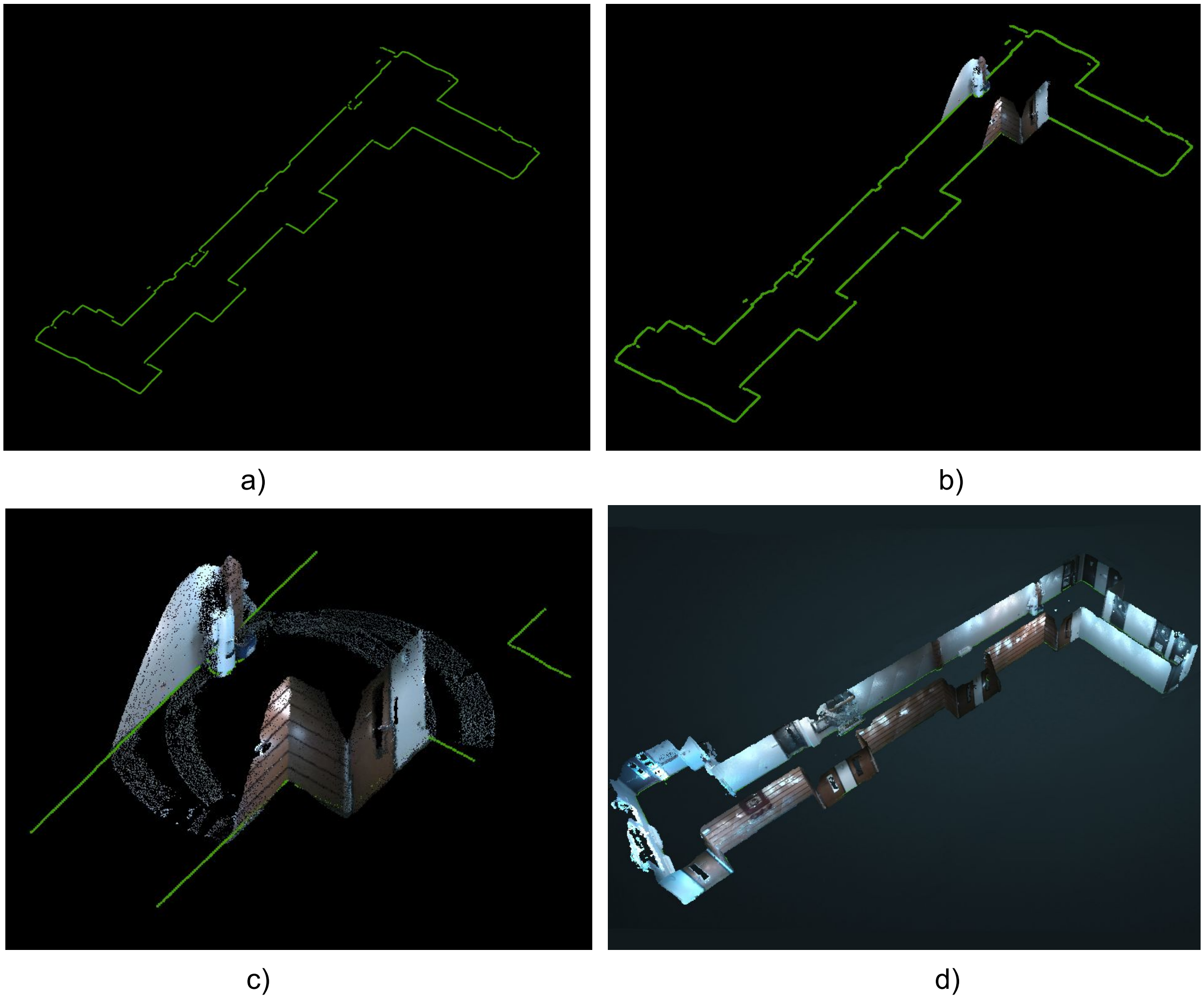


Figure 4 : Progressive construction of a building's 3D model. a) Exact 2D map obtained using the technique described in [8] b) First two 3D panoramas applied on the model c) Closeup of the first two panoramas d) Completed model

The scans in Figure 4 are raw : no post-treatment was applied to fuse the data, filter it or correct the colors and illumination artefacts. As such, positions that are far from the D435 during an individual scan (such as the top of the walls) are noticeably wavier because of increased noise. However, since the robot position is known precisely through the use of the building outline, the effect of these errors is not accumulated. In other words, by limiting the D435 produced data to 2.5 m measurements the expected error is (in our experience) 50 mm or 2% when measured from the sensor; when compared to the size of the building, this error is within a 0.2% threshold. As mentioned previously, this noise phenomenon can be corrected through clever treatment of the collected information, the subject of our next paper.

In short, we've shown that creating a complete 3D colored model of a building is possible only using a short range 3D sensor and a 2D map serving as a guide. Such models have many uses for BIM, from structure inspection during building construction to planning renovations or repurposing passing and passing by many others. The best part of the presented system is cost : whilst 3D scanners are traditionally expensive to buy and operate, the combination of a robot with two inexpensive sensors lets us create accurate 3D models at a fraction of the price.

References

[1] LINK TO FIRST PAPER
[2] Matterport, Company main page. https://matterport.com.
[3] Creaform, HandySCAN 3D product page. https://www.creaform3d.com/en/metrology-solutions/portable-3d-scanner-handyscan-3d.
[4] Intel, Intel Realsense Depth Camera D435 product page. https://click.intel.com/intelr-realsensetm-depth-camera-d435.html.
[5] Garmin, LIDAR-Lite v3HP product page. https://buy.garmin.com/en-US/US/p/578152/pn/010-01722-10.
[6] Anders Grunnet-Jepsen, Tuning D435 and D415 Camera for Optimal Performance. https://realsense.intel.com/wp-content/uploads/sites/63/BKM-For-Tuning-D435-and-D415-Cameras-Webinar_Rev3.pdf?language=en_US.
[7] Chris Aher, Automatic Accuracy : Getting the most out of DotProduct data with the accuscale-DP scale bar kit. https://www.dotproduct3d.com/blog/accuscaledp.
[8] LINK TO SECOND PAPER