3D Reconstruction of Real World Scenes Using a low-cost 3D range scanner

Paulo Dias
IEETA - Dep. de Electrónica e Telecomunicações
paulo.dias@ieeta.pt

Miguel Matos
Departamento de Engenharia Mecânica
a13821@alunos.mec.ua.pt

Vítor Santos
TEMA - Departamento de Engenharia Mecânica
vsantos@mec.ua.pt

Universidade de Aveiro
Campus Universitário de Santiago
3810-193 Aveiro

Abstract

This paper presents a 3D reconstruction technique for real world environments based on a traditional 2D laser range finder modified to implement a 3D laser scanner. The paper describes the mechanical and control issues addressed to achieve physically the 3D sensor as well as the adaptation of some previously developed techniques used to merge range and intensity data and illustrates the potential of such a unit. The result is a promising system for 3D modelling of real world scenes at a commercial price 10 or 20 times lower than current commercial 3D laser scanners.

Keywords

3D Reconstruction, Laser scanner, Range data, Camera Calibration, Data fusion.

1. INTRODUCTION

3D reconstruction of objects and environments is becoming an important topic of research with applications in many areas such as virtual museums, game and entertainment, architecture description and modelling, virtual reality, robot navigation, archaeology, inspection, cultural heritage and many industrial applications like reverse engineering. In the architecture industry, 3D reconstruction is useful for like archiving buildings texture and geometry before major modifications or modelling and modifying 3D models of an existing site to see the effect of the changes before any modification. Commercial solutions already exist for small dimension objects; most of them are based on vision systems that triangulate the position of the points by analysing the distortion of projected grids. This solution is not suitable for open spaces with larger dimensions (depths of tens to hundreds of meters) and in these cases many issues are not completely solved. Modelling of 3D large scenes presents many problems due to acquisition of large-scale data, complexity of the geometry and difficulties to cope with reflectance properties of the objects and variations in the lighting. The most advanced solutions are based on commercial laser range finders developed by a few companies (*Riegl Laser Measurement Systems*, *Cyra Technologies*, *Zoller & Frölich*, *Callidus Precision Systems*, among others) but the price of these equipments is still prohibitive (from 30 000 € to 100 000 €). A few groups are using these sensors and produced interesting results such as in Sequeira *et al.* [Sequeira99], El-Hakim *et al.* [El-Hakim98] and Stamos and Allen [Stamos00]. In addition to the range data, these works also combine the geometry information with digital pictures in order to "paint" the models with real colour information from the scenes and achieve realistic models. Unfortunately the high cost of laser range finder limits the use of this solution to only few groups despite the promising results obtained up to now.

An obvious solution is to adapt a traditional, much less expensive, 2D range finder (3-4 k€) to increase its capability and turn it into a Three-dimensional sensor. The aim is to create a sensor that can be used for 3D reconstruction at an affordable price. The idea is not new, as can be seen in Surman *et al.* [Surman01] and its references but the concept was initially developed without the authors being aware of other approaches and its main initial application was robotic navigation, since no colour information was acquired.

The adaptation of a 2D sensor to a 3D system needs to solve two main issues: develop a mechanically stable tilt unit that would rotate the base 2D laser range finder, and synchronize the pan and tilt information to produce a spherical representation of points that would further be processed for navigation or modelling. The laser unit used in this project is a SICK LMS 200 (see Table 2), which is an indoor version, but has, nonetheless, operated well also in outdoors, as shown later.

This paper describes briefly the solutions adopted for the mechanical unit, its controller and the communication issues. In section 3 range data processing is presented whereas section 4 addresses the problem of registration between digital photographs and range data. A few examples of indoor and outdoor scenes acquired at the campus of the University of Aveiro illustrate the whole reconstruction process in the results section. Finally open issues and conclusions are presented at the end of the paper.

2. THE TILT UNIT

2.1 Mechanical concerns

For the mechanical structure to support the laser sensor, two main approaches were possible: setting the laser either in tilt or pan position, that is, scanning preferably in horizontal or vertical planes. Figure 1 shows both approaches as developed by Surman *et al.* [Surman01], and by Batavia *et al.* [Batavia02].

Figure 1. Two possible approaches for the base scanner, by Surmann (left) and Batavia (right).

Since navigation was the primary purpose of the approach, a solution of the type proposed by Surmann was chosen because it allows faster and more efficient detection of vertical structures and obstacles, such as doors, chairs, tables, etc. Hence, the mechanical structure to support the laser range finder should have adequate robustness and stiffness and also it should not interfere with any mechanical part in the final device.

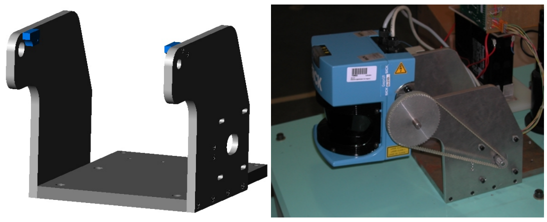


Figure 2. The mechanical structure model (left), and the real unit lying on top of a mobile robot (right).

Besides these concerns of mechanical motion, the geometry of the supporting structure should not occlude the laser beam for a wide range of tilt angles. This means, for example, scanning downwards should be possible, as well as looking backwards in upside down position. The final result is the structure shown in Figure 2.

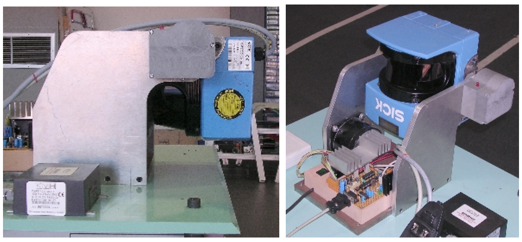


Figure 3. The laser pointing downwards (left) and pointing backwards (right).

The materials used for the construction are ordinary steel (for the base and shafts) and aluminium (used on the lateral supports). With such unit the laser range finder is able to point along a range of about 270º as shown in Figure 3, although the authors did not yet take advantage of the full range of the system.

2.2 The Control Unit

The control unit should be able to be interfaced with a standard communication link, such as RS232, USB or other. On the other hand, to actuate the tilt unit, either a servomotor or a step motor is an interesting solution due to the nature of position control of this system. For several reasons (availability in stock, motor torque, mechanical gearings, etc.) the usage of a step motor appeared as more advantageous. To perform the control and interface to external systems a PIC micro controller (PIC16F876) from Microchip was chosen. It interfaces by the serial port, issues displacement commands to the stepper motor and reads the value of the potentiometer for feedback positioning of the tilt angle. A classical scheme based on the L297 and L298 chips for power control drives the step motor. Figure 4 illustrates the main blocks of the control unit and Figure 5 illustrates the PCB implementation of the controller prototype.

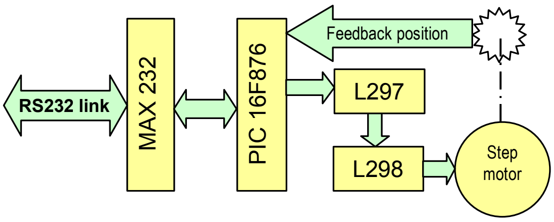


Figure 4. Functional block of control unit

One of the many advantages of the PIC micro controller, especially the flash versions, concerns future updates of the unit firmware, which can be done easily by means of the serial port without the usage of complex programmers or other hardware.

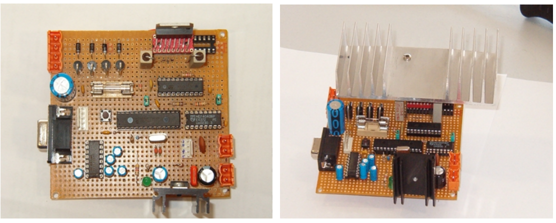


Figure 5.Control/power unit PCB; with cooler (right)

The control of the position is made in a close loop with the use of a potentiometer attached to one of the shafts that supports the sensor. With this process we can have the precise position of the sensor up to 10 bits (the PIC's ADCs resolution).

2.3 Communications

There are two components concerning communications. One is to obtain range data from the laser ranger, and the other is to interface the tilt unit, either to program its actions or to obtain its tilt angle. For maximal data throughput, the laser unit generates continuously a data stream at about 500 kbits per second, yielding a full 180º scan with 0.5º resolution each 26 ms, or, stated in another form, 38 full scans per second. That is done using a dedicated RS422 board addressed under a Linux operating system. Dedicated low-level protocols were developed to access measurements in the data stream since the vendor did not offer this software at programmer level, nor was it easily available from other research groups that use a similar unit.

The other communication issue concerns the tilt controller itself. A RS232 based serial interface at 9600 baud was developed for the unit to accept acquisition parameters, such as velocity and upper and lower scanning limits, and also to send continuously data regarding the instantaneous orientation of the laser range finder. The communication system is depicted in Figure 6.

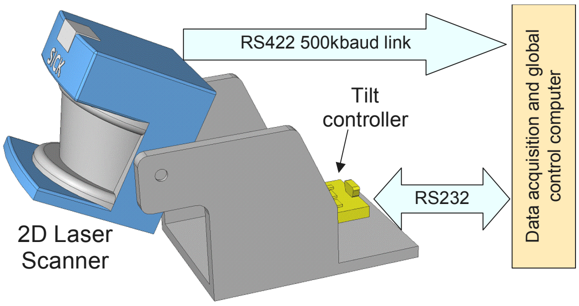


Figure 6. Communications during data acquisition.

Communications to the central computer were done asynchronously, but at a rate that ensures very little discrepancy between data. A full range scan (361 values) is embedded in a 732-byte data frame that takes circa 26 ms or less to transfer, and the position read at the potentiometer usually took about 3 ms to reach the central computer; this results in an overall cycle of less than 30 ms that under an angular velocity of 30 º/s would result in maximum tilt misalignment of less than 1º between the pan extremes of 0º and 180º. In practice these problems were not observed as relevant.

Table 1. Messages accepted by the control unit

Message	Description
1 1 1 0 0 0 0 0	Stop
0 0 0 0 0 0 0 0	Start immediate scan
0 0 0 x x x x x	Start scan with initial delay (0-31)
0 0 1 x x x x x	Set velocity of scan (0-31)
0 1 0 x x x x x	Upper scan limit in factors of 10º (0-31)
0 1 1 x x x x x	Upper scan limit in factors of 0.5º (0-19)
1 0 0 x x x x x	Lower scan limit in factors of 10º (0-31)
1 0 1 x x x x x	Lower scan limit in factors of 0.5º (0-19)
1 1 0 1 1 1 1 1	Returns current system configuration
1 1 0 0 0 0 0 0	Continuous scan between limits

A custom protocol has been developed to enable communications with the control unit. The messages are very simple and are all one byte long. The three most significant bits define the message and the 5 least significant bits define the argument, should the message have one (Table 1).

Table 2. Feature summary of the 3D laser unit

Property	Value
Tilt velocity (adjustable)	~1.5 º/s to 140 º/s
Tilt limits (adjustable)	–90º to +180º
Maximal covered solid angle in a single scan (pan × tilt)	180º × 270º
Tilt position best resolution	~0.3º
Original SICK specifications for LMS200	
Pan resolution	0.25º/0.5º (100º/180º pan range)
Maximal range	20 meters
Minimal range	0.15 meters
Range accuracy	±5 cm (was better in practice)
Measurement rate	13800+ measurements/s

The main features of the developed system are summarised in Table 2. To the knowledge of the authors, some features are not found in other existing systems, such as the adjustable scanning speed and adjustable tilt range within such a wide solid angle. Table 3 shows some features of three high-end 3D scanning systems costing about one hundred thousand Euros or more; naturally, their accuracy, angular resolution and linear range are unbeatable by the system proposed, but some other features such as angular coverage and data rate are comparable and, in some cases, better.

Table 3 - Some commercial 3D scanning systems

Property	Riegl LMSZ210	Z&F Imager 5003	Cyrax 2500
Maximal covered solid angle	80° x 333°	310° x 360°	40° x 40°
Angular Resolution	~0.072°	~0.01°	~0.04°
Maximal range	120 meters	53.5 meters	100 meters
Minimal range	2 meters	0.4 meters	1.5 meters
Range accuracy	15 mm	< 5 mm	< 6 mm
Measurements per second	9333	125000	1000

3. RANGE DATA PROCESSING

This section presents the reconstruction process used to compute 3D triangulated models from the cloud of points provided by the developed sensor. Two main issues are to be considered when computing 3D models from range points, first how to triangulate the points in order to obtain surfaces, and second how to register the several range images necessary to cover the whole model. These issues are addressed in the following paragraphs.

Figure 7 shows two range images (cloud of points) acquired with our laser unit. The first is from a laboratory at the mechanical department, and the second is an external view from the IEETA (Institute de Engenharia Electrónica e Telemática de Aveiro) building, both located at the University of Aveiro - Portugal. The range images size are respectively 361 by 125 and 361 by 127. The acquisition time was in both cases less than 5 seconds.

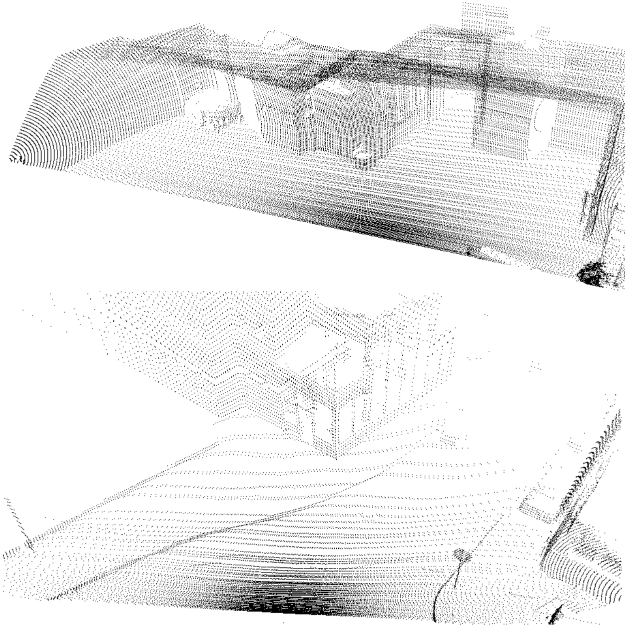


Figure 7. Two range images acquired with our laser unit.

3.1 Edge detection and Triangulation of single range image

For the triangulation of the cloud of point, a 2D Delaunay triangulation is used. This solution leads to models with lots of redundant information and large size if all range points are used in the process. Many systems try to detect areas where the geometry is more complicate in order to increase the number of triangle in these areas and decrease the number of points/triangles in areas with poor geometry information (such as planes or walls).

The 3D reconstruction software we developed takes this problem in account by analysing discontinuities in the range data. Two types of edges are detected: jump edges (that corresponds to occlusions where there is a jump in the measured distance) and roof edges (areas where there is a change in the orientation of the data). The edge detection process we used was based on the work from Jiang [Jiang99]. The algorithm fits a function along the scan lines to detect discontinuities in the range data acquired. Figure 8 present the results of the edge detection process with the laboratory image.

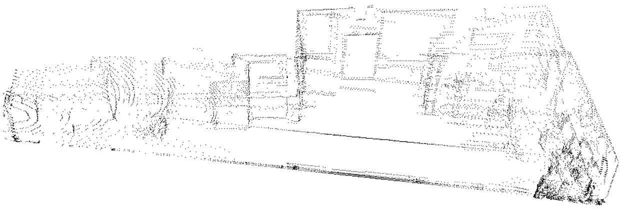
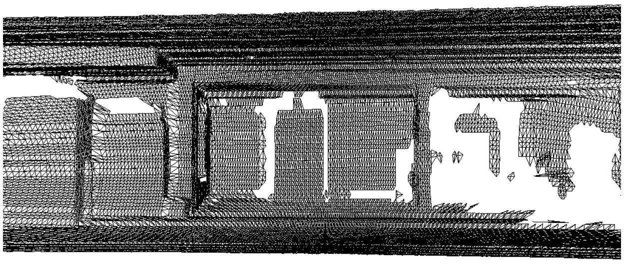


Figure 8. Edges detected in the Laboratory image.

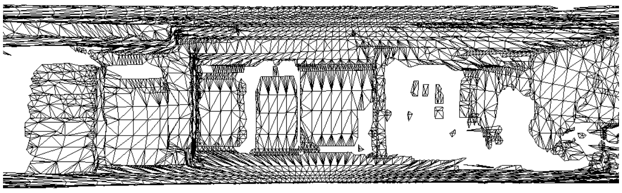
The edge information is used in two ways during the computation of the 3D triangulated models. First the edge points are maintained in the model, whereas non-edge points can be sub-sampled at a fixed rate. This operation results in models of reduced dimensions but with the main edges and discontinuities maintained.

The jump edges are also used as a filter to remove triangles over discontinuities, since there are normally located at occlusions. In Figure 9, we present the results of the triangulation of the laboratory image with two different sub-sampling. In the first model all the range point were used in the triangulation, resulting in smaller triangles but also a much larger model (4210kb). In the second model, non-edge points were sub-sampled with a rate of 5, resulting in a reduced model with a size of only 871kb, but with the main discontinuities still present.

The implementation code we use is based on the 2D Delaunay triangulation provided in the VTK Toolkit software.



(a)

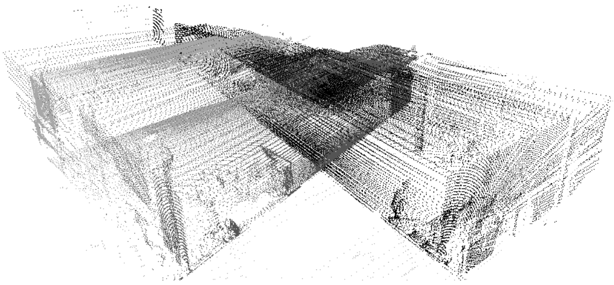


(b)

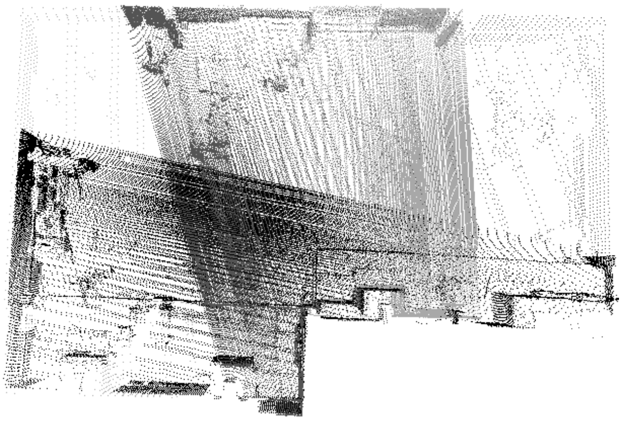
Figure 9. Triangulated models of the laboratory image. (a) with all the data (size: 4210kb) and (b) with sub-sampling of 5 (871kb).

3.2 Registration of several images

A major problem in 3D reconstruction is related to occlusions: it is impossible to have a complete description of a complex environment with only a single image. To achieve complete models always require several range images taken from different viewpoints. It is then necessary to register these images spatially. This issue is well known in the 3D reconstruction community and is normally solved using the Iterative Closest Point algorithm presented for the first time in [Besl92]. In our implementation, a user interface gives the possibility to select 3 corresponding points (or more) between two range images to feed the iterative closest point algorithm that minimizes the distance between the range images using the user selected points as initial guess for the 3D rigid transform between the data. Figure 10 presents the result of the fusion of three clouds of points in the laboratory scene.



(a)



(b)

Figure 10. Two views of the registration of 3 range images of the laboratory.

4. FUSION OF RANGE AND INTENSITY DATA

Range data gives a geometric description of the scenes, but most laser scanners do not provide any colour information. An approach to achieve 3D models of real environments consists in combining the geometric information from the laser range finders with digital photographs. The geometry of the scene is recovered from the 3D cloud of points, whereas texture information is extracted from high-resolution digital photographs as seen in Sequeira *et al.* [Sequeira99], El-Hackim *et al.* [El-Hakim98], Stamos and Allen [Stamos00], Smith and Elstrom [Smith99] and Kurazame *et al.* [Kurazume02]. With such techniques, it is possible to achieve photo-realistic 3D models of real scenes. In this section we apply these techniques to the range data obtained with our developed sensor to validate its capabilities in acquiring coherent 3D range data from a real scene. More details about the method used to register range and intensity data can be found in Dias [Dias03].

4.1 Range and intensity registration

For each scene, along with the range data, several digital photographs were also acquired with a hand held Canon EOS 300D digital camera (resolution of 3072 x 2048 pixels). Two of these digital photographs of the laboratory and the IEETA scenes are presented in Figure 11.

The digital photographs are registered with the range data using a camera calibration process. The used model for the camera is the one proposed by Tsai [Tsai87] and the implementation is the one by Wilson [Wilson94]. The method requires correspondences between 3D points in range data and 2D pixels in intensity images to compute the camera model.

Figure 11. Digital photographs of the laboratory and the IEETA scenes.

Most commercial Laser Range Finders provide the distance and a reflectance value that gives information about the amount of light reflected by the objects. This information provides a kind of black and white image that is normally used to select correspondences in the range data. Since our laser does not provide directly this reflectance information, a 3D interface based in OpenGL was developed in order to allow a user to select interactively pixels in the photographs and 3D points in the range data. In addition, if more than the minimum eleven points are selected, a RANSAC technique, as seen in Fischler and Bolles [Fischler81], is used in order to reject outliers and increase the robustness of the process.

4.2 Texture mapping

The camera model is the starting point to compute a texture map for the three dimensional models since it gives the possibility to link 3D position in range data with 2D pixels in the photographs. It is then possible to re-project the 2D pixels in the intensity images in order to create a new image that is fully registered with the 3D data as shown in Figure 12. This final image is useful to evaluate the quality of the registration and can be used directly to texture map the 3D models, giving a much more realistic impression.

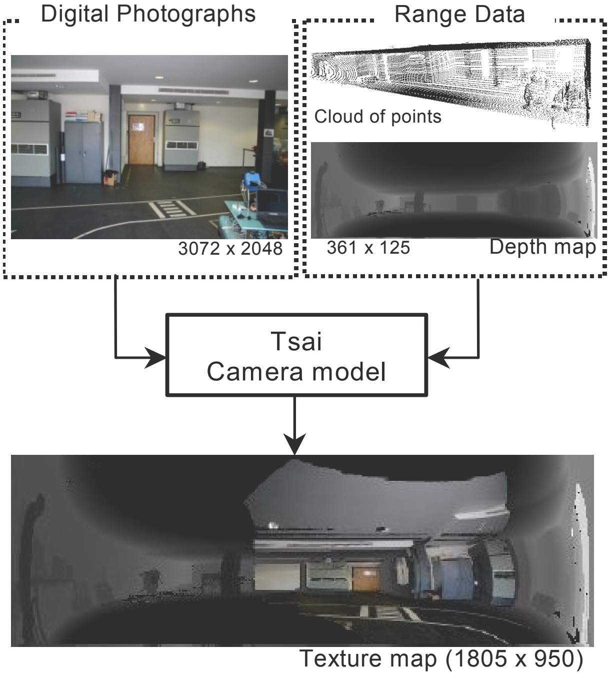


Figure 12. The re-projection process.

During the re-projection process, the range data is re-sampled at a higher resolution and bilinear interpolation used to compute the "extra" 3D positions. This interpolation makes possible to take advantage of the higher resolution of the photographs for the texture maps. Projecting the pixels can also lead to occlusion errors. To solve this problem, a Z-buffer is used to ensure that occluded areas will not be re-projected in the texture map.

The field of view of the range sensor is larger than a normal camera. In this situation there is the need to use several images to cover the whole range data. To allow this possibility, the re-projection can be repeated for several images to provide a texture map that covers the whole scene. In a final step, the images are blended together into a final texture map. An example is presented in Figure 13 for the laboratory scene. The blending technique used in this case consists simply in averaging the images in the common areas.

Figure 13. Average blending of three images

Figure 14 presents three different snapshots from the VRML model of a single image of the laboratory textured with the image resulting from the blending.

Figure 14. Three snapshots of a single textured image of laboratory model.

5. RESULTS

In this section we present some indoor and outdoor results obtained combining all the methods presented in the paper to illustrate the potentialities of the developed laser scanner. We present some results with three different test scenes, one indoor and two outdoors, all acquired within the campus of the University of Aveiro.

- The laboratory model (Figure 15). This is the final model of the laboratory used along the paper to illustrate our methods. The final model was compute from 3 range scans and the texture is based on 6 digital photographs.

Figure 15: Model of Laboratory (3 range images).

- The IEETA model (Figure 16). The resulting model includes data coming from 12 different range scans, and the texture information was computed from the merging of 12 digital photographs. The holes in the model are due to the large windows in the entrance of the building pointing out a limitation of Range scanners that cannot measure distance to windows.

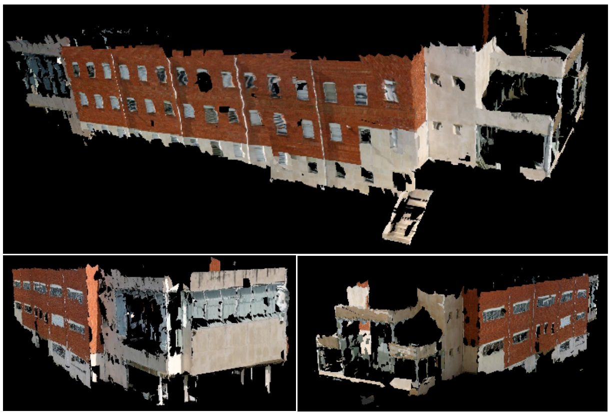


Figure 16: Model of IEETA (12 range images).

- The SACA model (Figure 17). This is a partial model of a complex building in the campus (Secção Autónoma de Comunicação e Arte). In this model 11 range images and 14 digital photographs were used.

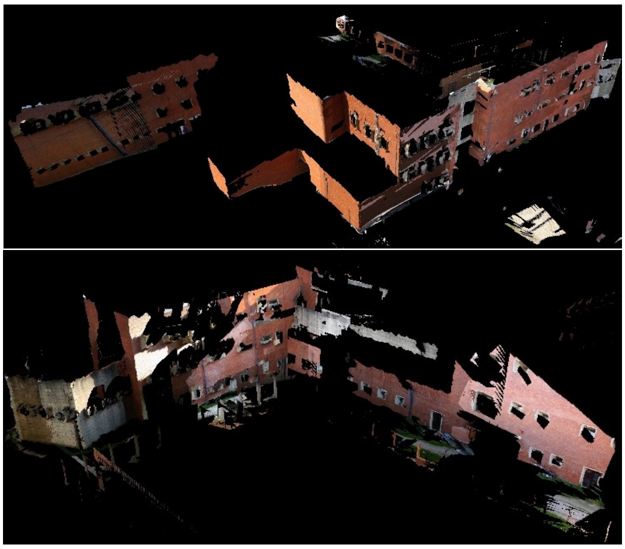


Figure 17: Partial model of SACA (11 range images).

6. OPEN ISSUES AND CONCLUSIONS

When compared to commercial systems for 3D scanning, the proposed solution shows up some limitations namely in range and linear accuracy; indeed, some commercial systems have millimetre precisions, which can even be excessively high and not necessary to common navigation or modelling applications. On the other hand, the proposed device has a cost much lower than commercial solutions and uses a quite common range finder, which is almost an established industrial component, therefore easy to find and purchase. Moreover, the developed unit presents a few features and advantages not found in any commercial solutions, such as the large coverage of the environment in a single scan (180º x 270º) and its full versatility in adjusting tilt range and velocity. A major limitation of the sensor is related to the absence of reflectance image. The sensor only gives geometric information making difficult the registration when the scenes have little geometry. In these situations a reflectance image could help by giving additional information about the reflected light.

Among possible improvements on the sensor, we can underline the possibility to use an outdoor laser (we use an indoor version of the SICK) that would increase the range (we experimentally measure around 20 metres) and robustness.

As far as software and automatic reconstruction tasks are concerned, many issues have not been yet considered in this paper. Indeed, the reconstruction software is still in an early stage and many questions have not been considered. Many of the tools can easily be improved, for example the use of a constrained Delaunay triangulation to optimise the edges in the final models. Blending and optimisation of the final merged model based on several range and digital pictures also has to be taken into account.

Range data must also undergo a minor correction procedure due to the fact that the tilt rotation axis does not cross exactly the centre of the laser emitter resulting in slight distortions for larger tilt angles, especially at further pan positions.

Independently of the needed improvements just mentioned, the proposed device has shown good capabilities for a fast low cost 3D perception of space and showed also enough accuracy to allow the modelling of large environments When merged with digital photographs, the resulting models are photo-realistic and can be used in a wide variety of application all this at a reduced cost when compared with state of the art laser scanners.

7. REFERENCES

[Sequeira99] Sequeira, V., Ng, K. , Wolfart, E., Gonçalves, J. G. M., Hogg, D. (1999). Automated Reconstruction of 3D Models from Real Environments, *ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier).* vol. 54, pp. 1-22, 1999.

[El-Hakim98] El-Hakim, S. F., Brenner, C., Roth, G. (1998). A multi-sensor approach to creating accurate virtual environments, *ISPRS Journal for Photogrammetry and Remote Sensing*, Vol. 53(6). pp. 379-391, December 1998.

[Stamo00] Stamos, I., Allen, P. K. (2000). 3-D Model Construction Using Range and Image Data, *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 531-536 (volume I). South Carolina, June 2000.

[Surman01] Surman, H., Lingemann, K., Nüchter, A., Hertzberg, J. (2001). Fast acquiring and analysis of three dimensional laser range data, *6th International Workshop on Vision, Modelling, And Visualization, VMV 2001*, pp. 59-66, Stuttgart, 21-23 Nov 2001, Germany.

[Batavia02] Batavia, P. H., Roth, S. A., Singh, S. (2002). Autonomous Coverage Operations In Semi-Structured Outdoor Environments, *Proceedings of the Intl. Conference on Intelligent Robots and Systems*, EPFL, Lausanne, Switzerland, October 2002.

[Jiang99] Jiang X., Bunke H. (1999). Edge Detection in Range Images Based on Scan Line Approximation, *Computer Vision and Image understanding*, Vol.73, No. 2m February, pp 183-199.

[Besl82] Besl,P.J, McKay N.D, A method for registration of 3D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No 2, pp. 239-256 , February 1992.

[Smith99] Smith, P. W., Elstrom, M. D. (1999). Stereo Based Registration of Multi-Sensor Imagery for Enhanced Visualization of Remote Environments, *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1948-1953, May 1999.

[Kurazume02] Kurazume, R., Nishino, K., Zhang, Z., Ikeuchi, K. (2002). Simultaneous 2D images and 3D geometric model registration for texture mapping utilizing reflectance attribute, *Proceedings of 5th Asian Conference on Computer Vision (ACCV).* Vol. I, pp. 99-106, January 2002

[Dias03]Dias, P. (2003). Three dimensional Reconstruction of Real World Scenes Using Laser and Intensity Data, *PhD thesis, University of Aveiro*, September 2003.

[Tsai87] Tsai, R.Y. (1987). A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, *IEEE Journal of Robotics and automation*, Vol. RA-3, No. 4, pp. 323-344, August 1987.

[Wilson94] Wilson, R. G. (1994). Modeling and Calibration of Automated Zoom Lenses, *PhD thesis*, Department of Electrical and Computer Engineering, Carnegie Mellon University, January 1994.

[Fischler81] Fischler, M. A., Bolles, R.C. (1981). Random Sample Consensus: a paradigm for model fitting with application to image analysis and automated cartography, *Communications of the ACM 24 (6)*, pp. 381-395.