3D Reconstruction of Real World Scenes

Using a low-cost 3D range scanner

Paulo Dias*

Department of Electronic and Telecommunications / IEETA, University of Aveiro, Portugal

And

Miguel Matos, Vítor Santos

Department of Mechanical Engineering / TEMA, University of Aveiro, Portugal

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. The use of such a system can simplify measurements of existing buildings and produce easily 3D models and ortophotos of existing structures with minimum effort and at affordable price.

* Corresponding author. paulo.dias@ieeta.pt
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 and civil engineering industry, 3D reconstruction is useful for 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 real and expensive modification is realized. 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 & Fröhlich, Callidas 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. (1999), El-Hakim et al. (1998) and Stamos and Allen (2000). 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. (2001) 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 pho-
tographs 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. (2001), and by Batavia et al. (2002).

FIGURE 1

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

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 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 270° 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 microcontroller (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.

FIGURE 4

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.

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 180° scan with 0.5° 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 5.

FIGURE 5

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 °/s would result in maximum tilt misalignment of less than 1° between the pan extremes of 0° and 180°. In practice these problems were not observed as relevant.

TABLE 1

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

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

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 6 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 (Instituto 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 6

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 (1999). The algorithm fits a function (a line in our case) along the scan lines to detect discontinuities in the range data acquired, and mark as edges the points where the distance between the function and the point is maximum and above a threshold distance. Figure 7 present the results of the edge detection process with the laboratory image. The edge detection in this model took 266ms in a P4 at 2.0GHz.

**FIGURE 7**

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 8, 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. In this case the final model has 44976 vertexes (4210kb), and the triangulation algorithm took 13 minutes (P4 at 2.0GHz). In the second model, non-edge points were sub-sampled with a rate of 5, resulting in a reduced triangulated model with a size of only 5394 vertexes (871kb) computed in 43 seconds (P4 at 2.0GHz), 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.

**FIGURE 8**

### 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 by Besl
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 iteratively the distance between the range images using the user selected points as initial guess for the 3D rigid transform between the data. Figure 9 presents the result of the fusion of three clouds of points in the laboratory scene. In this example, two views were first registered based on 3 user-selected points, the registration took 53 seconds and 8 iterations. The resulting registered image (with two views) was then registered with the third view in 7 iterations and 95 seconds (P4 at 2.0GHz).

FIGURE 9

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. (1999), El-Hackim et al. (1998), Stamos and Allen (2000), Smith and Elstrom (1999) and Kurazame et al. (2002). 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 (2003).

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). 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 (1987) and the implementation is the one by Wilson (1994). The method requires correspondences between 3D points in range data and 2D pixels in intensity images to compute the camera model.

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 (1981), is used in order to reject outliers and increase the robustness of the process.

The results of the camera calibration process for the laboratory image (3 range images and 7 digital photographs) are presented in Table 4. For each digital photograph, we present in the table the number of point correspondences selected by the user in the interface. The average calibration error is the average distance in pixels between the selected points and the points re-projected using the computed camera model in an image normalized to 512x512 pixels. We also present for each calibration the processing times in a P4 at 2.0GHz.

**TABLE 4**

4.2 Texture mapping

Based on the camera model is we compute a texture map for the three dimensional models since we have 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 10. 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 10**

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 process re-project each 3D points into the photograph and if more than two points are projected at the same pixel location, it ensures that only the closest one will be considered when computing the final texture map. The processing times of the full re-projection process (Z-buffering and Re-projection) are presented in Table 4.

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 11 for the laboratory scene. The blending technique used in this case consists simply in averaging the images in the common areas and took 1.26 seconds in this example.

**FIGURE 11**
Figure 12 presents three different snapshots from the VRML model of a single image of the laboratory textured with the image resulting from the blending.

FIGURE 12

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 13). 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 7 digital photographs.

FIGURE 13

- The IEETA model (Figure 14). 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 14

- The SACA model (Figure 15). 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 15

6. OPEN ISSUES AND FUTURE WORK

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 triangulation done with a constrained 2D Delaunay triangulation can be easily improved using a 3D triangulation. 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.

In the near future, experiments in a virtual reality environment are also planned. A Head Mounted Display as well as a Three Degree of Freedom sensor have been acquired (see Figure 16) in order to allow the user to experience a full immersion in the 3D models created using these techniques. The idea is to give the user a feeling of “being there” as well as making the navigation in the 3D world easier since The Yaw, Pitch and Roll movements will be directly connected to the movements of the user’s head by the Intersense sensor. In this situation, the usual interface, mouse or keyboard, will be only used to control the X, Y and Z movements in the 3D world.

7. 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 (180° x 270°) 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.

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.
8. REFERENCES


<table>
<thead>
<tr>
<th>Message</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 1 0 0 0 0 0</td>
<td>Stop</td>
</tr>
<tr>
<td>0 0 0 0 0 0 0 0</td>
<td>Start immediate scan</td>
</tr>
<tr>
<td>0 0 0 x x x x x</td>
<td>Start scan with initial delay (0-31)</td>
</tr>
<tr>
<td>0 0 1 x x x x x</td>
<td>Set velocity of scan (0-31)</td>
</tr>
<tr>
<td>0 1 0 x x x x x</td>
<td>Upper scan limit in factors of 10º (0-31)</td>
</tr>
<tr>
<td>0 1 1 x x x x x</td>
<td>Upper scan limit in factors of 0.5º (0-19)</td>
</tr>
<tr>
<td>1 0 0 x x x x x</td>
<td>Lower scan limit in factors of 10º (0-31)</td>
</tr>
<tr>
<td>1 0 1 x x x x x</td>
<td>Lower scan limit in factors of 0.5º (0-19)</td>
</tr>
<tr>
<td>1 1 0 1 1 1 1 1</td>
<td>Returns current system configuration</td>
</tr>
<tr>
<td>1 1 0 0 0 0 0 0</td>
<td>Continuous scan between limits</td>
</tr>
</tbody>
</table>
Table 2. Feature summary of the 3D laser unit

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tilt velocity (adjustable)</td>
<td>~1.5 º/s to 140 º/s</td>
</tr>
<tr>
<td>Tilt limits (adjustable)</td>
<td>~90º to +180º</td>
</tr>
<tr>
<td>Maximal covered solid angle in a single scan (pan × tilt)</td>
<td>180º × 270º</td>
</tr>
<tr>
<td>Tilt position best resolution</td>
<td>~0.3º</td>
</tr>
<tr>
<td><strong>Original SICK specifications for LMS200</strong></td>
<td></td>
</tr>
<tr>
<td>Pan resolution</td>
<td>0.25º/0.5º (100º/180º pan range)</td>
</tr>
<tr>
<td>Maximal range</td>
<td>20 meters</td>
</tr>
<tr>
<td>Minimal range</td>
<td>0.15 meters</td>
</tr>
<tr>
<td>Range accuracy</td>
<td>±5 cm (was better in practice)</td>
</tr>
<tr>
<td>Measurement rate</td>
<td>13800+ measurements/s</td>
</tr>
</tbody>
</table>
Table 3 - Some commercial 3D scanning systems

<table>
<thead>
<tr>
<th>Property</th>
<th>Riegl LMSZ210</th>
<th>Z&amp;F Imager 5003</th>
<th>Cyrax 2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal covered solid angle</td>
<td>80° x 333°</td>
<td>310° x 360°</td>
<td>40° x 40°</td>
</tr>
<tr>
<td>Angular Resolution</td>
<td>~0.072°</td>
<td>~0.01°</td>
<td>~0.04°</td>
</tr>
<tr>
<td>Maximal range</td>
<td>120 meters</td>
<td>53.5 meters</td>
<td>100 meters</td>
</tr>
<tr>
<td>Minimal range</td>
<td>2 meters</td>
<td>0.4 meters</td>
<td>1.5 meters</td>
</tr>
<tr>
<td>Range accuracy</td>
<td>15 mm</td>
<td>&lt; 5 mm</td>
<td>&lt; 6 mm</td>
</tr>
<tr>
<td>Measurements per second</td>
<td>9333</td>
<td>125000</td>
<td>1000</td>
</tr>
</tbody>
</table>
Table 4. Results of the camera calibration in laboratory data.

<table>
<thead>
<tr>
<th>Range Image</th>
<th>Photograph</th>
<th>Nº of points used for calibration</th>
<th>Average Calibration Error (pixels)</th>
<th>Processing Time (seconds)</th>
<th>Reprojection Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range 1</td>
<td>Photo 1</td>
<td>15</td>
<td>8.08</td>
<td>14.57</td>
<td>12.13</td>
</tr>
<tr>
<td></td>
<td>Photo 2</td>
<td>14</td>
<td>4.82</td>
<td>16.26</td>
<td>14.91</td>
</tr>
<tr>
<td></td>
<td>Photo 3</td>
<td>12</td>
<td>4.65</td>
<td>20.11</td>
<td>12.39</td>
</tr>
<tr>
<td>Range 2</td>
<td>Photo 1</td>
<td>17</td>
<td>15.28</td>
<td>13.03</td>
<td>15.97</td>
</tr>
<tr>
<td></td>
<td>Photo 2</td>
<td>13</td>
<td>3.91</td>
<td>23.33</td>
<td>14.48</td>
</tr>
<tr>
<td>Range 3</td>
<td>Photo 1</td>
<td>12</td>
<td>17.19</td>
<td>12.56</td>
<td>10.56</td>
</tr>
<tr>
<td></td>
<td>Photo 2</td>
<td>16</td>
<td>13.17</td>
<td>10.97</td>
<td>11.80</td>
</tr>
</tbody>
</table>
Fig. 1. Two possible approaches for the base scanner, by Surmann (left) and Batavia (right).
Fig. 2. The mechanical structure model (left), and the real unit lying on top of a mobile robot (right).
Fig. 3. The laser pointing downwards (left) and pointing backwards (right).
Fig. 4. Functional block of control unit
Fig. 5. Communications during data acquisition.
Fig. 6. Two range images acquired with our laser unit.
Fig. 7. Edges detected in the Laboratory image.
Fig. 8. Triangulated models of the laboratory image. (a) with all the data (size: 4210kb) and (b) with subsampling of 5 (871kb).
Fig. 9. Two views of the registration of 3 range images of the laboratory.
Fig. 10. The re-projection process.
Fig. 11. Average blending of three images
Fig. 12. Three snapshots of a single textured image of laboratory model.
Fig. 13: Model of Laboratory (3 range images).
Fig. 14: Model of IEETA (12 range images).
Fig. 15: Partial model of SACA (11 range images).
Fig. 16: (a) Head Mounted Display and
(b) 3DOF Sensor for future Virtual Reality applications