

A refined range image registration technique for multi-stripe laser scanner

Carles Matabosch^a, Joaquim Salvi^a, David Fofi^b and Fabrice Meriaudeau^b

^aInstitut d'Informatica i Aplicacions, University of Girona, Campus Montilivi, 17017, Girona, Spain;

^bLaboratoire Electronique, Informatique et Image, University of Burgundy, rue de la fonderie 12, 71200, Le Creusot, France

ABSTRACT

Nowadays, visual inspection is very important in the quality control for many industrial applications. However, the complexity of most 3D objects constrains the registration of range images; a complete surface is required to compare the acquired surface to the model. Range finders are very used to digitalize free form shape objects with large resolutions. Moreover, one single view is not enough to reconstruct the whole surface due to occlusions, shadows, etc. In these situations, the motion between reconstructed partial views are required to integrate all surfaces in a single model. However, the use of positioning systems is not always available or adequate due mainly to the size of the objects or the environmental conditions imposed by the precise mechanics which suffer from vibrations present in the industry. In order to solve this problem, a 3D hand sensor is developed to reconstruct 3D objects that let us to compare them with respect the original one.

Keywords: Global model reconstruction, Registration, Laser projection

1. INTRODUCTION

3D reconstruction is very important in computer vision. There are several applications that use this information, as reverse engineering, modelling, robot navigation, visual inspection, etc. In order to obtain these surfaces, there are different kind of 3D sensors based on laser projection¹ or coded structured light² among others. In general, laser projection techniques are based on the use of a laser emitter coupled to a cylindrical lens that spread the light forming a plane that is projected to the measured surface.

The projection of a laser plane only lets us to reconstruct a profile of the measuring surface. So, in most cases a mechanical system is added that permits a scanning. In some applications: a) the laser plane is projected onto a rotating mirror and reflected towards the surface; b) the laser beam is attached to a moving worm gear; c) the laser beam keeps motionless while the object is placed on a rotating table. All these techniques permit the reconstruction of a whole surface with high resolution. However, the accuracy strongly depends on the mechanical system used so that potential vibrations present in the environment produce misalignments and consequently the accuracy is considerably influenced. Furthermore, the sequence of images that are captured in the scanning process forces the object to be motion controlled reducing the number of applications, i.e. industrial conveyors can not be considered.

Besides, this kind of sensors can only reconstruct part of the object due to object occlusions and the limited field of view of the sensor. In order to obtain a full reconstruction, several images have to be fused using Registration algorithms. The goal of these techniques is to determine the Euclidean motion between range images. Theoretically, the best way to register several views is a Multi-view registration technique, that registers all views simultaneously avoiding the drift or error propagation problem. Moreover, in Multi-view registration

Further author information: (Send correspondence to C.M.)

C.M.: E-mail: cmatabos@eia.udg.es, Telephone: 0034 972419812

J.S.: E-mail: qsalvi@eia.udg.es, Telephone: 0034 972419812

D.F.: E-mail: d.fofi@iutlecreusot.u-bourgogne.fr, Telephone: +33 (0)385731126

F.M.: E-mail: fabrice@iutlecreusot.u-bourgogne.fr, Telephone: +33 (0)385731077

an initial estimation is required which is not previously known in our case. Then an initial rough alignment must be also considered.

This rough alignment can be found using a coarse registration technique, which determine the motion between two views without prior information. Several methods have been proposed such as Spin Image,³ Point signature⁴ and RANSAC-Based Darces⁵ among others. However, coarse registration techniques usually require lots of time to converge and in some situations the quality obtained is not good enough.

In our application, the motion between views is relative small (16 frames/s can be acquired by the camera), so that the identity matrix can be used as an estimation between two consecutive views, and this estimation can be then refined using a Pair-wise algorithm. The most common technique is Iterative Closest Point (ICP) introduced by Besl.⁶ This algorithm finds the motion by minimizing the distance between temporal point correspondences known as Closest Points. However, this algorithm suffers the problem of local minima. Independently of Besl, Chen⁷ proposed a similar approach of ICP by minimizing distances between scene points in the first surface and a local planar approximation of the second surface. Other approaches are presented that improves these techniques by introducing features in the points,⁸ thresholds to remove false correspondences^{9,10}, selecting the best points to register,¹¹ etc. A survey of these techniques has been recently presented.¹²

Using one of these techniques, the motion between two consecutive surfaces can be obtain. However, after several views the accumulation of the error can be important. To avoid this problem, we propose to refine this motion by using a multi-view registration algorithm between the views contained in a loop. Several authors refine a loop by distributing the error in the loop between all views,¹³ besides, others refine all views simultaneously.¹⁴ Although the first one is the fastest solution, the second one is preferred to obtain better results.

Summarizing, the goal of this paper is to present a one-shot 3D sensor that lets us to acquire several surfaces of the same object by using a hand sensor without any reference to a world coordinate system. When reconstruction is obtained, a registration technique is applied to align all views. However, as the error registration is propagated in each view, final errors can not be despicable. Then, a technique to minimize the error propagation is used.

The paper is structured as follows. First, section 2 presents the one-shot 3D sensor used. The calibration of this system is presented in section 3. Section 4 presents the reconstruction step. Registration process is detailed in section 5. Finally, some experimental results are presented and the article ends with conclusions.

2. ONE-SHOT 3D SENSOR

Nowadays, there exist a wide sort of commercial 3D scanners based on laser projection and coded structured light, among others. However, only few of them are suitable to obtain a 3D surface using a unique acquired image (without scanning), which are basically based on coded structured light projection.¹⁵ However, the projectors used in structured light are big, expensive and resolution limited so that they are usually not considered in industry.

To solve this problem, a laser emitter is used. Currently, there exist a considerably amount of lenses which can be coupled to a laser emitter which spreads the light forming a given pattern: planes, circles, dots and stripes. Besides, it has been demonstrated that stripe patterns are the most suitable in measuring processes because of the easy segmentation and the use of subpixel techniques in the detection of the stripe peaks.¹

The number of projected stripes is directly related with the final resolution. A compromise between resolution and segmentation is required because the difficulties in the segmentation process increase with the number of laser stripes projected. Hence, the use of several laser profiles can be classified as an uncoded structured light so that the Stripe Indexing problem needs to be solved. The proposed solution is commented in the following sections.

3. CALIBRATION

The process of calibration consists of finding a relation between 3D points on the measuring surfaces with the projection of these points in the acquired image. This relation can be linearly approximated by the following

equation:

$$\begin{bmatrix} sX \\ sY \\ sZ \\ s \end{bmatrix} = {}^W T_L \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (1)$$

Once ${}^W T_L$ is known, 2D points in the image frame can be directly transformed to 3D points in the world reference frame. Obviously, the parameters t_{ij} of matrix T should be estimated as precisely as possible in order to maximize the accuracy in the reconstruction. According to equation 1, the expressions for sX , sY , sZ and s are obtained and shown in equation 2

$$\begin{aligned} sX &= t_{11} \cdot u + t_{12} \cdot v + t_{13} \\ sY &= t_{21} \cdot u + t_{22} \cdot v + t_{23} \\ sZ &= t_{31} \cdot u + t_{32} \cdot v + t_{33} \\ s &= t_{41} \cdot u + t_{42} \cdot v + t_{43} \end{aligned} \quad (2)$$

Arranging the terms and grouping, a homogeneous system of three equations with 12 unknowns (t_{11} to t_{43}) is obtained as shown in equation 3

$$\begin{aligned} t_{11} \cdot u + t_{12} \cdot v + t_{13} - t_{41} \cdot u \cdot X - t_{42} \cdot v \cdot X - t_{43} \cdot X &= 0 \\ t_{21} \cdot u + t_{22} \cdot v + t_{23} - t_{41} \cdot u \cdot Y - t_{42} \cdot v \cdot Y - t_{43} \cdot Y &= 0 \\ t_{31} \cdot u + t_{32} \cdot v + t_{33} - t_{41} \cdot u \cdot Z - t_{42} \cdot v \cdot Z - t_{43} \cdot Z &= 0 \end{aligned} \quad (3)$$

In order to search for correspondences, the complete quadrangle is used.¹⁶ The original method has been adapted to calibrate the set of 19 planes obtaining the 19 transformation matrices which describes the geometry of the sensor. For each laser plane, the following steps are processed:

- Detection of the points of the laser profile in the image plane,
- Find the correspondences between points in the image plane and 3D points in the calibrating plane,
- and Compute the T matrix using the correspondences given by the previous step.

Now, the description of these steps is detailed in the following sections.

3.1. Points in the laser profile

When a unique plane is projected to the scene, the peak detection with subpixel accuracy can be determined with high accuracy using a FIR filter approach.¹⁷ However, when more planes are projected, the derived curve of the profile is high influenced by the neighborhood. In some situations, the derived curve does not cross to zero at the maximum value of the intensity profile. To solve this problem, an adapted methodology that is based on a previous work related to coded structured light is used.¹⁵ First of all, the first derivative is computed using the convolution of each row with the vector $[-1 -1 -1 0 1 1 1]$. Then, the second derivative is computed obtaining the enhancement of the peaks compared to the intensity image. A threshold is finally used to segment the stripes as follows:

$$\begin{cases} 0 & \text{if } f_i'' < \text{threshold} \\ 255 & \text{otherwise} \end{cases} \quad (4)$$

where f is the intensity profile curve and f_i'' is the second derivative in each pixel of the row. The interval of each peak can be found easily analyzing all the pixels in a consecutive order. For each interval, the central value is computed as an approximation of the position of each maximum. Then, a local derivative is computed in each estimated peak as follows:

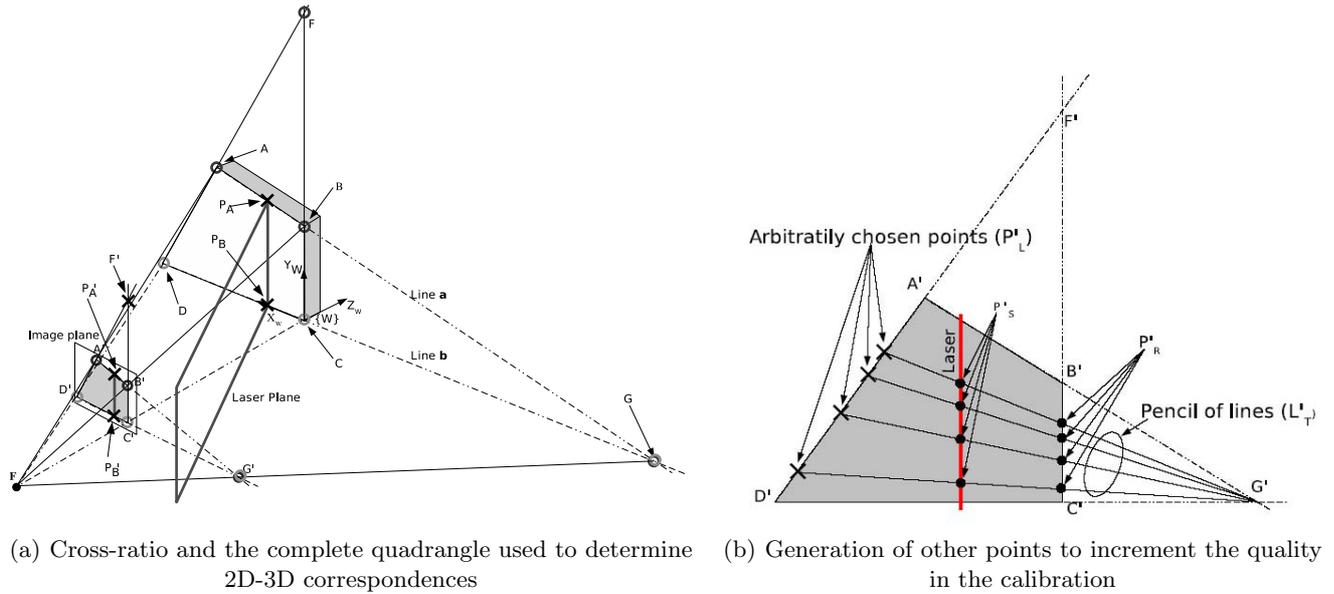


Figure 1. Calibration process

$$f'_L = conv([1 \ -1 \ -1 \ 0 \ 1 \ 1 \ 1], [f(p_i - 3) : f(p_i + 3)]) \quad (5)$$

where $conv$ is the convolution, and $f(p_i)$ is the value of the intensity profile in the i th estimated peak. The zero-crossing of the f'_L function gives us the sub-pixel position of the peak of each laser stripe. Furthermore, if the intensity value of these points is less than a threshold, this peak is not considered.

3.2. Correspondences between points in the image and 3D points

The methodology is based on the *complete quadrangle*.¹⁸ The principle of this method is the cross-ratio between the complete quadrangle and the acquired image of this quadrangle (see fig. 1a)).

$$\frac{\overline{A'P'_A}}{\overline{A'G'}} = \frac{\overline{AP_A}}{\overline{AG}} \quad (6)$$

As A, B are known 3D points, and A', B' and P'_A can be found analyzing the acquired image, P_A can be determined by the cross-ratio principle. The same principle is applied with point P_B . If the quadrangle is moved along the Z-axis, a set of 2D-3D correspondences can be found for each Z position. Using this set of correspondences, eq. 1 can be solved determining the transformation matrix.

In general, only two points are used for every plane position. Note that calibration accuracy is related directly to the number of correspondences used. In order to improve the accuracy, a set of points along the laser stripe are selected. To do this, arbitrary points (P'_L) are selected in the quadrangle (see fig. 1b)). The pencil of lines that joint these points with point G' are created. The intersection of these lines with the laser stripe gives us the auxiliary points of the calibration. The process to determine the 3D correspondence points is the same as in the first situation. More details are presented in.¹⁶

3.3. Compute T matrix using known correspondences

Now the transformation matrix can be obtained by minimizing eq. 7 which has been easily obtained arranging eq. 3, where t_{ij} 's are the parameters of the ${}^W T_L$ matrix, u_i and v_i are the pixel coordinates and X_i, Y_i and Z_i are the coordinates of the 3D position. The solution is obtained from the computation of the vector θ that minimizes equation $A \cdot \theta = 0$. A good estimation using Total Least Square technique is computed from the eigenvector corresponding to the smallest eigenvalue of matrix $A^T \cdot A$.

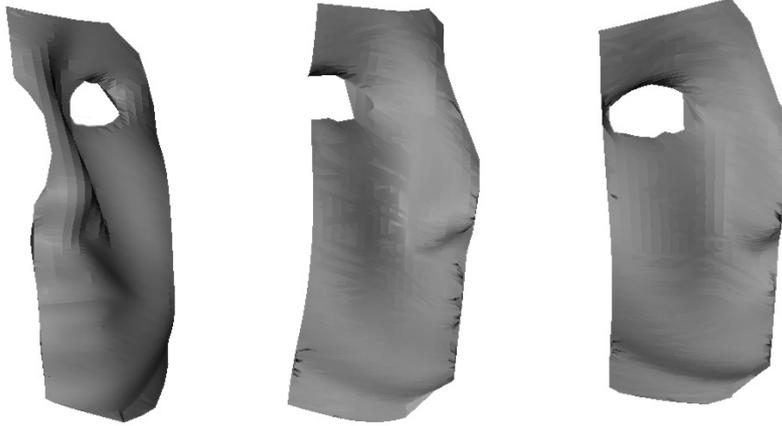


Figure 3. Some of the reconstructed partial views

between two (or more) range images. If the motion between them is very small or an initial estimation is known, *Fine registration* techniques can be directly applied. This group of techniques minimizes iteratively the distances between point correspondences as shown in eq 8:

$$f = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\vec{m}_i - R\vec{p}_i - \vec{t}\|^2 \quad (8)$$

where m_i and p_i are points of the second and the first range image respectively, and R and t are the initial motion parameters. However, when the motion between views is unknown, a first step is required. The goal of this step, known as *Coarse Registration*, is to determine a poor estimation of the motion searching for correspondences between both surfaces.

In our work, as we acquired images consecutively, the motion between a surface and its consecutive is very small, and a Fine registration method can be directly applied. The motion between all consecutive images is computed using a Point-to-plane variant of Iterative Closest Point algorithm,⁷ improved with a normal space sampling approach¹¹. This method is a Pair-wise registration algorithm because only two images are registered simultaneously.

Theoretically, when the coarse registration is done all views can be represented in the same coordinate frame, and the final model can be built. However, as the motion is not exactly computed, error is propagated and after some views this error can be considerable, especially in objects with smooth shape. To solve this problem, a refinement step is needed. Although some authors use Multi-view registration algorithms, these algorithms can only be used with a limited number of views for several reasons. First of all, if the surface contains a lot of views, the position of final views can be far from the real one due to the propagation error. Secondly, if lots of views are registered simultaneously, computing time can be very intensive. To cope with these problems, we propose to determine cycles and refine only views contained in these cycles.

To determine a cycle, the relative position between each view is computed, and if the norm of the translation vector between all the already registered views and the last acquired view is less than a threshold, a cycle is considered between both views. Then, the motion between the views of the cycle can be refined using a Multi-view approach^{13,14}. After this refinement, the process goes on acquiring a new image and registering it with respect to the last view already registered.

6. EXPERIMENTAL RESULTS

A set-up consisting of an off-the-shelf camera, a 635 nm laser emitter and an optical lens which spreads the laser beam into 19 planes has been arranged conforming the imaging system. A calibrating quadrangle has been

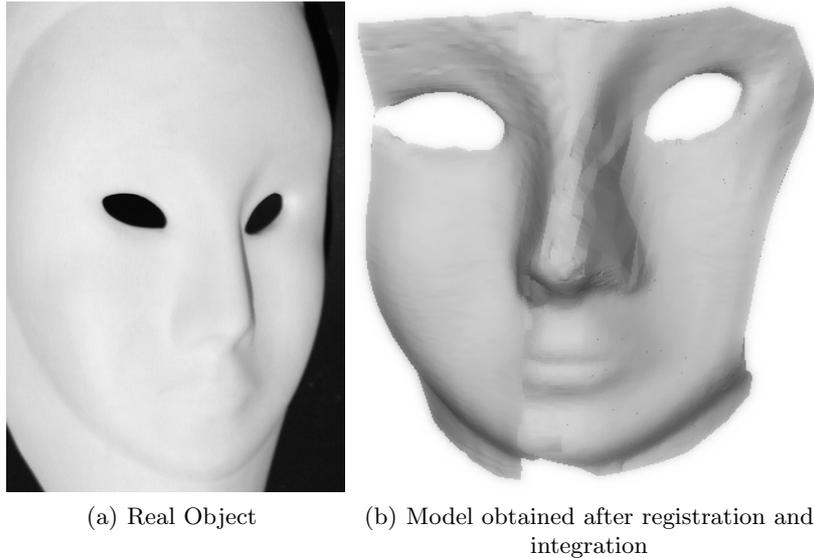


Figure 4. Final results

located at several distances from the system in increments of 2 mm. The closest plane is located approximately at 20 cm from the imaging system. For every quadrangle position, two images are acquired: a) the first is an image of the quadrangle; b) the second is the projection of the laser on the quadrangle. The first image is used to determine the parameters of the quadrangle while the second one is used to determine the geometry of the laser. Then, every laser stripe is determined by a sequence of n correspondences which are used to compute the transformation matrix for each stripe. Where n represents the number of planes used in the calibration.

The accuracy of this system is computed from the discrepancy between the reconstructed 3D points and the 3D points used in the calibration process. The reconstruction of the calibration planes is obtained with a RMS error of 0.5 mm and a variation of less than a 0.2 mm.

When the calibration parameters have been determined, 10 range images of the same object are reconstructed, some of which are presented in Fig. 3 and finally registered obtaining as a sample the object present in Fig. 4b).

7. CONCLUSIONS

This paper presents a new method to build 3D models. A one-shot imaging system based on a single off-the-shelf camera and a stripe laser pattern is used to acquire surfaces. After that, a registration step is used to register accurately all the acquired surfaces in order to generate a complete 3D model of the measuring object. The system benefits from one-shot techniques to recover the 3D shape of surfaces in non-controlled motion environments or even in the presence of vibrations because in any case the 3D view is obtained from a single image without any scanning. Registration is used to align every 3D acquisition with respect to a world coordinate system obtaining a complete reconstruction of the measuring object. The calibration benefits from the use of the complete quadrangle and image processing from the use of a nice stripe peak detector with subpixel accuracy.

The refinement process lets us to obtain a global model removing the misalignments produced by error propagation. Experimental results show that the accuracy obtained in the reconstruction step is quite acceptable (less than 0.5 mm in the centered area) and the reconstruction of real objects is also provided with both quality and qualitative good results.

ACKNOWLEDGMENTS

Work founded by Spanish Project TIC2003-08106-C02-02.

REFERENCES

1. J. Forest and J. Salvi, "An overview of laser slit 3d digitasers," in *International Conference on Robots and Systems*, pp. 73–78, (Lausanne), October 2002.
2. J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition* **37**, pp. 827–849, April 2004.
3. A. Johnson and M. Hebert, "Surface registration by matching oriented points," in *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pp. 121–128, May 1997.
4. C. J. R. Chua, "Point signatures: A new representation for 3d object recognition," *International Journal of Computer Vision* **25**, pp. 63–85, October 1997.
5. C.-S. Chen, Y.-P. Hung, and J.-B. Cheng, "A fast automatic method for registration of partially overlapping range images," in *International Conference on Computer Vision*, pp. 242–248, (Bombay), January 1998.
6. P. Besl and N. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, pp. 239–256, February 1992.
7. G. Chen, Y. ad Medioni, "Object modeling by registration of multiple range images," in *IEEE International Conference on Robotics and Automation*, pp. 2724 –2729, April 1991.
8. G. Sharp, S. Lee, and D. Wehe, "Icp registration using invariant features," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**, pp. 90–102, January 2002.
9. E. Trucco, A. Fusiello, and V. Roberto, "Robust motion and correspondences of noisy 3-d point sets with missing data," *Pattern Recognition Letters* **20**, pp. 889–898, September 1999.
10. T. Zinsser and H. Schnidt, J. Niermann, "A refined icp algorithm for robust 3-d correspondences estimation," in *International Conference on Image Processing*, pp. 695–698, September 2003.
11. S. Rusinkiewicz and M. Levoy, "Efficient variant of the icp algorithm," in *3rd International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152, 2001.
12. C. Matabosch, J. Salvi, D. Fofi, and F. M. F., "Range image registration for industrial inspection," in *Machine Vision Applications in Industrial Inspection*, (San Jose, California, USA), January 2005.
13. G. Sharp, S. Lee, and D. Wehe, "Mutliview registration of 3d scenes by minimizing error between coordinate frames," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, pp. 1037–1050, August 2004.
14. T. Masuda, "Generation of geometric model by registration and integration of multiple range images," in *Third International Conference on 3-D Digital Imaging and Modeling*, pp. 254 –261, May 2001.
15. J. Pages, J. Salvi, C. Collewet, and J. Forest, "Optimised de bruijn patterns for one-shot shape acquisition," *Image and Vision Computing* **23**, pp. 827–849, April 2005.
16. J. Forest, *New methods for triangulation-based shaped acquisition using laser scanners*. PhD thesis, University of Girona, 2004.
17. J. Forest, J. Salvi, and C. Cabruja, E.and Pous, "Laser stripe peak detector for 3d scanners. a fir filter approach," in *International Conference on Pattern Recognition*, **3**, pp. 646 – 649, (Cambridge, United Kingdom), August 2004.
18. C. Chen and A. Kak, "Modelling and calibration of a structured light scanner for 3d robot vision," in *IEE conference on robotics and automation*, pp. 807–815, 1987.
19. A. Robinson, L. Alboul, and M. Rodrigues, "Methods for indexing stripes in uncoded structured light scanning systems," *Journal of Winter School of Computer Graphics* **12**(1-3), 2004.