Tele-3D — Developing a Handheld Scanner Using Structured Light Projection

João Filipe Ferreira, Jorge Lobo, Jorge Dias
Institute of Systems and Robotics, Coimbra Pole
University of Coimbra
Coimbra, Portugal
jfilipe, jlobo, jorge@isr.uc.pt

Abstract

Three-dimensional surface reconstruction from two-dimensional images is a process with great potential for use on different fields of research, commerce and industrial production. In this article we will describe the evolution of a project comprising the study and development of systems which implement the aforementioned process, exploring several techniques with the final aim of devising the best possible compromise between flexibility, performance and cost-effectiveness. We will firstly focus our attention on past work, namely the description of the implementation and results of a fixed system involving a camera and a laser-stripe projector mounted on a pan-tilt unit which sweeps the surface vertically with a horizontal stripe. Then we will describe our current work on the development of a fully portable, handheld system using cameras, projected structured light and inertial/magnetic positioning and attitude sensors — the ***Tele-3D scanner****.*

1. Introduction

The problem of 3D spatial structure recovery and reconstruction has been the object of discussion in the past two decades for people involved in areas as diverse as the obvious computer, physical, or medical sciences, and as the perhaps less obvious, but nevertheless important, subjects as documentary and film production, down to all industries in general. We have, consequently, arrived to an era where this kind of technology is no longer only a part of science fiction [6].

This article presents the study and development of various 3D information recovery and reconstruction systems, commonly known as *3D scanners* or *3D digitisers* (although the latter is more encompassing and thus more correct, the first is more common and, for that reason, will be the designation to be used from now on). The goal of this research project is to aim for an "optimal solution" in 3D scanning, where low-cost components are mounted on an easy-to-use, light-weight device flexible enough to scan small to moderately sized scenes with satisfactory accuracy and resolution performances. With this objective in mind, we decided to resort to laser range scanning by triangulation technologies, since these have been proven to be a very robust and cost-effective solution for medium to high precision 3D scanning devices [11, 6, 3].

The traditional approaches using this kind of technologies on 3D scanning, however, have been reported to be severely limited in their usefulness by their constraints on the size of the scenes to be recovered and by the lack of flexibility and portability of the scanning devices thus constructed [12, 13, 10].

Therefore, we have endeavoured into a multiply phased research project, evolving from a somewhat cumbersome fixed laser-scanner system, to an exciting new approach, the *handheld 3D scanning systems*, with the final aim of addressing those, which we believe to be the main issues, that influence the performance of laser range scanning systems: *flexibility*, *portability*, *calibration* and *3D data integration*.

2. Past work — the fixed laser-scanner

In the course of this project, the first step we decided to take was the study and implementation of the classic, fixed set-up, laser-scanner. In the following subsections we will describe, consecutively, the system's mathematical, geometrical and physical background, and some of the results obtained by its use [6].

2.1. Theoretical background

The first step for the recovery and reconstruction of three-dimensional scenes from two-dimensional images is to determine how to extract the required three-dimensional data from the two-dimensional information available on those images.

Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02)
0-7695-1521-5/02 $17.00 © 2002 IEEE

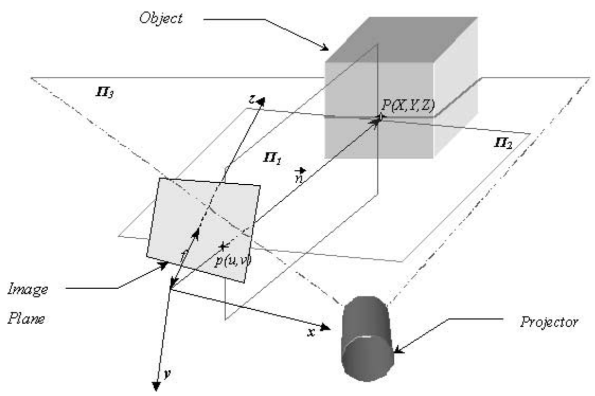

Figure 1. Generic structured light 3D scanner [6].

Figure 2. Baby-doll reconstruction subject.

The geometry involving a generic structured light 3D scanning system is presented in figure 1 [6].

It becomes obvious that the information extracted from the two-dimensional image-plane of a projective camera is not sufficient to completely reconstruct a three-dimensional scene. Thus, at least one additional restriction is needed to establish an univocal correspondence between the 3D point and its projection on the image-plane. There are several ways of achieving this goal, one of which being the use of *projected structured light*.

It is clear that the light source is in fact projecting a plane of light, which can be mathematically represented by equation 1.

$$\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = 0 \Leftrightarrow \mathbf{\Pi_3 P'} = 0 \qquad (1)$$

Using the geometry involved in the perspective projection of the world onto the camera's image-plane and the light-plane's restriction of equation 1, the 3D point in the scene can be uniquely related to its two-dimensional projection point in the image-plane of the camera by triangulation using:

$$\begin{cases} \mathbf{\Pi_1 P'} = 0 \\ \mathbf{\Pi_2 P'} = 0 \\ \mathbf{\Pi_3 P'} = 0 \end{cases} \qquad (2)$$

For a more complete explanation of these principles, please refer to [6].

2.2. Results

We will now present some of the system's reconstruction results. Figure 2 shows a baby-doll which was scanned by our device — a reconstruction example of the baby-doll subject is presented in figure 3.

One of the greatest downsides to this system is made clear by this figure; note the effects of severe occlusions suffered by the laser stripe on a substantial length of the subject's body. This means that relevant information is lost due to the otherwise essential depth-inducing baseline between the camera and the projector.

3. Current work — the handheld laser-scanner, Tele-3D

The following step we decided to take was the study and implementation of what we believe to be the cutting-edge of this field: a *portable, handheld laser-scanner*. This device will utilise a hybrid pose estimation procedure where both information intrinsic to the acquired data and extrinsic information measured from sensors are used to achieve data integration.

In the following subsections we will describe, consecutively, the handheld system's architecture, the involved mathematical background and the system's implementation.

3.1. System architecture

On figure 4, we present the handheld scanner's architecture. A rigid structure with a straight angle boomerang

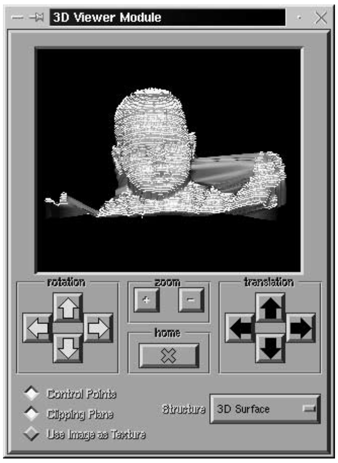

Figure 3. Baby-doll reconstruction with sampled points.

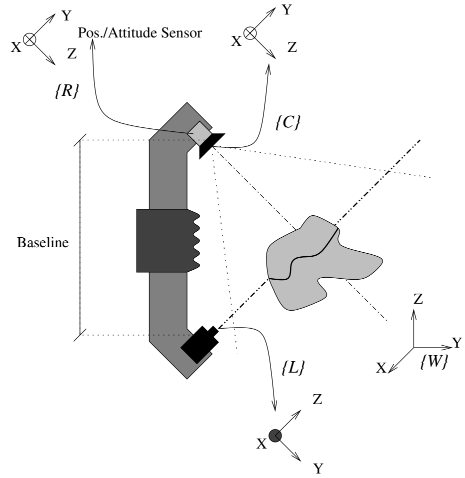

Figure 4. Handheld architecture schematics.

shape has the camera mounted on one side and the laser projector on the other; a handling grip is placed at the centre of mass of the structure. Coupled with the laser projector is the attitude and positioning device, using inertial and magnetic sensors. Also on figure 4, all referentials directly or indirectly involved on the scanning process are represented: the camera's referential, $\{C\}$, the laser projector's referential, $\{L\}$, the position and orientation measuring sensor's referential, $\{R\}$, and the world's absolute referential $\{W\}$.

controller connected to each and the positioning/attitude sensor - laser RS232 connections respectively.

3.2. Theoretical background

Consider an overview of our handheld system — a camera and a laser projector mounted in fixed position on a rigid structure that will be moved by hand so as to completely scan the 3D scene to recover. At any time, as can be inferred from figure 4, the camera and the laser projector's referentials are fixed relatively to each other and can thus be locally calibrated beforehand.

Image frames of the scene will be grabbed in sequence as the scan is performed and a set of 3D profiles will be determined from each intersection of the projected light-plane with the 3D surfaces on the scene. These profiles, designated as $\mathbf{P}(t_n)$ for each sampling time $t = t_n$, represent, in fact, a set of ordered points $\mathbf{x}(t_n)$ sampled from those surfaces. Contrary to what happened with the fixed positioned laser scanner of our previous work, the handheld scanner and its local referentials are moved and are rotated during the scanning process. Since no scaling or reflections are involved, each position and attitude change of the system between consecutive sampling instants can be modelled as a *rigid-body transformation*, a combination of a rotation with a translation expressed by [1, 5, 4]

$$\mathbf{x}(t_n) = \mathbf{R_{n-1,n}}\mathbf{x}(t_{n-1}) + \vec{\mathbf{t}}_{\mathbf{n-1,n}} \qquad (3)$$

where $\mathbf{x}(t_n)$ corresponds to any sampled point at sampling time $t = t_n$, with coordinates referred to any of the system's local referentials (for example, figure 4's $\{L\}$ or $\{C\}$).

Evidently, for complete integration of acquired data into a coherent reconstruction of the scene, computation of all such rigid transformations of any chosen system's local referential is needed so as to perform *3D registration*. The three-dimensional registration process has been defined in several ways depending on the implementation (check [8, 1, 2] for examples), but can be formalised in a more generic form as the *transformation of sets of three-dimensional measurements into a common coordinate system*.

A 3D registration technique can be described mainly by how it addresses 3 crucial problems [1, 2, 4]:

1. Choice of feature space (i.e., type of 3D measurements used: points, curves, planes, voxel images, etc.);

2. Choice of transformation (including rotation representation — by an orthonormal matrix, by quaternions, etc.) and, consequently, of the objective-function to be optimised and its parameter vector;

3. Feature matching (determining correspondences between features is mandatory, since registration will be impossible without them) and global optimisation.

A preferred registration method seems to be the *Iterative Closest Point* algorithm (ICP) and its variations, a technique devised by Besl and McKay and discussed in [2].

Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02)
0-7695-1521-5/02 $17.00 © 2002 **IEEE**

3.3. Present state of implementation

For our handheld 3D scanner project, the first issue we have addressed was the improvement of our calibration procedures: a robust camera calibration software package was developed using Intel's Open Source Computer Vision Library [9]. The OpenCV's calibration method, mainly based on [15], is an iterative algorithm applied to the pinhole model of the camera with added radial and tangential distortion coefficients. It uses a chessboard pattern to supply 3D points with well-known coordinates [9]. A procedure similar to what is described as *"Single Wall Phantom"* in [14] is then used to calibrate the laser-plane and the positioning sensor's respective coordinate systems simultaneously.

Several studies concerning the system's set-up, the registration methods and the positioning devices are being made. Some important issues have already become clear from these studies:

1. For the registration methods to work, *profiles must overlap*, i.e., a same region of the 3D scene must be scanned at least twice; since profiles are finite, these must also cross inside the scene's boundaries (see [7]) — this is due to the matching features constraint, that implies that, given two feature sets, one must be a subset of the other.

2. In order to simplify the model for registration and to adapt it for use with the pose measurement devices, the rigid-body transformation expression can be reformulated in a manner which decouples translation computation from that of rotations by referring the coordinates to the respective centroids of each point set (see [1]). The translation from sets A to B of matching points would then be given by:

$$\vec{\mathbf{t}}_{\mathbf{AB}} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{x}_{\mathbf{B_j}} - \mathbf{R}_{\mathbf{AB}}\frac{1}{N}\sum_{j=1}^{N}\mathbf{x}_{\mathbf{A_j}} \qquad (4)$$

3. The 3D scenes must be quasi-static for the system to work properly, making scanning of organic subjects, such as human bodies, difficult to achieve.

4. Discussion and Conclusions

Practical areas of use for 3D scanning systems, as described in this article's introductory section, will push towards a revolution in this field of research and small, portable, fast, reliable and cost-effective systems will be available as technologies advance. The current challenge is, thus, to optimise solutions to bring out the best of each implementation. This task, however, is anything but trivial.

It is thus the authors' conviction that a lot of experimental work in this field is still to come.

References

[1] M. A. Audette, F. P. Ferrie, and T. M. Peters. An Algorithmic Overview of Surface Registration Techniques for Medical Imaging. *Medical Image Analysis*, 4(3):201–217, September 2000.

[2] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[3] J. Davis and X. Chen. A Laser Range Scanner for Minimum Calibration Complexity. In *Proc. Third International Conference on 3D Digital Imaging and Modelling*. 3DIM 2001, 2001.

[4] O. D. Faugeras and M. Hebert. The Representation, Recognition and Locating of 3-D Objects. *The International Journal of Robotics Research*, 5(3):27–52, Fall 1986. Massachusetts Institute of Technology.

[5] J. Feldmar and N. Ayache. Rigid, Affine and Locally Affine Registration of Free-Form Surfaces. Research Report 2220, Institut National De Recherche En Informatique Et En Automatique (INRIA), March 1994. Project Epidaure.

[6] J. Ferreira and J. Dias. A 3D Scanner - Three-Dimensional Reconstruction From Multiple Images. In *Proc. Controlo 2000 Conf. on Automatic Control*, pages 690–695. University of Minho, Portugal, 2000. Student Forum.

[7] P. Hébert and M. Rioux. Toward a hand-held laser range scanner: Integrating observation-based motion compensation. In *Proceedings of IS&T/SPIE's 10th Annual Symposium on Electronic Imaging (Photonics West); Conference 3313: Three-Dimensional Image Capture and Applications*, San Jose, Ca, USA, January 1998.

[8] A. Hilton. Model Building from 3D Surface Measurements. Web page - http://www.ee.surrey.ac.uk/Research/VSSP/3DVision/model_building/model.html, December 1997. Tutorial.

[9] Intel Corporation, USA. *Open Source Computer Vision Library*, 1999-2001. Reference Manual.

[10] M. Levoy et al. The Digital Michelangelo Project: 3D scanning of large statues. In *SIGGRAPH 2000 Procedings*. New Orleans, LA, July 2000.

[11] S. J. Marshall and J. H. Gilby. New Opportunities in Non-Contact 3D Measurement. In *Proc. National Measurement Conf. 2001*, Harrogate, UK, 2001.

[12] B. C. McCallum, W. R. Fright, M. A. Nixon, and N. B. Price. A Feasibility Study of Hand-held Laser Surface Scanning. In *Proc. Image and Vision Computing New Zealand*, pages 103–108, Lower Hutt, August 1996.

[13] B. C. McCallum, M. A. Nixon, N. B. Price, and W. R. Fright. Hand-held Laser Scanning In Practice. In *Proc. Image and Vision Computing New Zealand*, pages 17–22. University of Auckland, October 1998.

[14] R. Prager, R. Rohling, A. Gee, and L. Berman. Automatic calibration for 3-D free-hand ultrasound. Technical Report CUED/F-INFENG/TR 303, Cambridge University Department of Engineering, September 1997.

[15] Z. Zhang. A Flexible New Technique For Camera Calibration. Technical report, Microsoft Research, 2000.

0-7695-1521-5/02 $17.00 © 2002 **IEEE**

IEEE COMPUTER SOCIETY