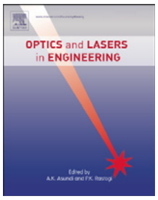
Optics and Lasers in Engineering 51 (2013) 116–127

Contents lists available at SciVerse ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng



Multiple alignments of range maps by active stereo imaging and global marker framing

Sandro Barone, Alessandro Paoli*, Armando Viviano Razionale

University of Pisa, Department of Mechanical, Nuclear and Production Engineering, Largo Lucio Lazzarino, n.1–56126 Pisa, Italy

ARTICLE INFO

Article history:
Received 25 July 2012
Received in revised form
30 August 2012
Accepted 6 September 2012
Available online 25 September 2012

Keywords:
3D optical scanning
Stereo-vision
Multi-view measurements
Marker detection
Global refinement

ABSTRACT

Three-dimensional shapes can be digitised by using active imaging techniques, which reconstruct entire objects by capturing multiple range maps from different viewpoints. Multiple-view reconstructions require the computation of translation and rotation parameters to transform each range map with reference to a global coordinate frame.

In this paper, an automatic method has been developed to efficiently align 3D range maps acquired by an active stereo vision system. The methodology is based on referring range maps to a global frame of fiducial markers captured by the stereo system. The procedure includes a refinement of the marker frame in order to globally minimise the misalignment errors.

The methodology optimises the overall accuracy of 3D reconstructions regardless of the scanning strategies, even processing large data sets.

The proposed approach has been experienced and validated by measuring both nominal shapes and industrial models.

1. Introduction

In the past few years, 3D scanning technologies have been extensively used within various engineering fields with the aim at capturing shapes of physical objects. The most popular technologies are based on active triangulation systems, which use lighting devices and vision sensors [1,2]. Typically, imaging data from two calibrated views are triangulated to provide a range map of 3D points corresponding to the surface viewed by the vision devices.

The reconstruction of complex shapes requires multiple range maps to be captured from different views and aligned into a common coordinate system.

The range map alignment is a non-linear problem, which typically involves a two-step solution: (1) estimation of a local coarse alignment by an approximate transformation; (2) global registration by Iterative Closest Point (ICP) algorithms to minimise mismatches between overlapping surfaces [3–5].

During the last decades, the technical literature has proposed various coarse alignment approaches, which can be classified as follows:

(1) detection of tie points [6–7];
(2) mechanical tracking [8–9];
(3) shape-feature matching [10–21].

Tie points of range maps can be interactively identified on overlapping surfaces of adjacent views [6]. This approach is suitable for all applications, though based on time-consuming procedures. Alternatively, marker-based methods [7] consist of detecting fiducial points distributed on target surfaces. This approach allows automatic measurements within preliminarily planned strategies.

In high-end systems, coarse alignment may be performed by mechanical tracking devices, such as a coordinate measurement machine [8], or a robotic arm [9]. The tracking procedure can be automatised by means of a calibration process, which spatially relates the optical probe to the motorised axes. These approaches provide alignment errors that depend on both the accuracy of the single tracking devices and the calibration routines. Moreover, the reconstruction procedure is affected by restrictions on size and complexity of target objects.

Shape-feature matching techniques are based on defining shape descriptors and detecting relative correspondences, which are used to match the range maps under some rigidity constraints. These methods might not guarantee proper robustness if significant geometric features are missing (*i.e.*, objects with flat or uniform curvature regions).

The coarse alignment procedures provide lower accuracy than single view measurements due to accumulation of local misalignments [22], which can be reduced by ICP algorithms. In general,

---

* Corresponding author. Tel.: +39 50 221 8174; fax: +39 50 221 8065.
E-mail addresses: s.barone@ing.unipi.it (S. Barone), a.paoli@ing.unipi.it (A. Paoli), a.razionale@ing.unipi.it (A.V. Razionale).

0143-8166/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.optlaseng.2012.09.003

the final accuracy of an alignment process depends on: (i) number of range maps, (ii) local accuracy of the coarse alignments, (iii) topological nature of the target surfaces (open or closed shapes) and (iv) amount of three-dimensional features onto overlapping areas between adjacent range maps. Although, some interesting solutions have been proposed, the development of methodologies capable to efficiently handle all these aspects, still represents a matter of research.

In this paper, an alignment method has been developed in order to accurately align a high number of range maps captured by a structured lighting system based on a stereo vision set-up. Individual range maps are referred to a 3D frame of fiducial markers spatially distributed onto the target surface and detected processing images acquired by the stereo vision rig. Fiducial markers are progressively addressed within the 3D frame as follows:

(1) 2D detection by intensity analyses;
(2) 3D triangulation and weight assignment based on viewing orientation;
(3) 3D framing by searching for distance similarities.

The proposed approach has been developed on the basis of a marker-guided methodology presented in [7]. The preliminary proposal was limited to *local coarse alignments* of range maps. In this paper, the methodology is further enhanced by globally minimising the misalignment errors between 3D marker coordinates through a weight-based refinement method applied to the 3D marker frame. The novel contribution significantly outperforms the previous approach by strengthening the overall methodology in terms of accuracy, even processing large data sets and regardless of the adopted multi-view strategy.

The benefits of the proposed method have been validated by measuring nominal geometries and reconstructing industrial models. Moreover, the effectiveness of the refinement process has been experimentally demonstrated by evidencing the advantages in reverse engineering applications.

2. Range map acquisition

An active stereo vision system has been developed to capture range maps by projecting encoded light stripes. The system set-up (Fig. 1a) includes a DLP projector (1024 × 768 pixels) and two 8-bit monochrome CCD cameras (1280 × 960 pixels) equipped with lenses having a focal length of 16 mm.

The stereo vision rig is calibrated by evaluating the intrinsic (focal distances, co-ordinates of the principal points, radial and tangential distortions) and the extrinsic parameters of the digital cameras [23].

An encoding approach is used for 3D shape recovery by projecting vertical and horizontal light stripes defined as crossing zones between fringes with periods progressively halved [24]. In particular, a binary code (0,1 with n bit) is assigned to each camera pixel, where n is the number of stripe patterns, and the values 0 and 1 depend on intensity levels, *i.e.*, 0 = *black* and 1 = *white*.

A double code is then obtained by projecting both horizontal and vertical stripes (Fig. 1b). This methodology solves the problem of point stereo correspondence by encoding $\ell_v \times \ell_h$ points, where $\ell_v (\ell_v = 1023)$ and $\ell_h (\ell_h = 767)$ are the numbers of vertical and horizontal lines switched by the projector, respectively.

In this work, the maximum resolution is obtained by projecting 9 stripe patterns with periods progressively halved and three stripe patterns by 1-pixel shifting the 9th pattern (Fig. 2). This procedure avoids a further stripe halving, which may typically produce poor accuracy in line extraction.

3D coordinates are finally calculated by using the midpoint triangulation technique for stereo corresponding points [25].

3. 2D marker detection

In this work, markers are textured by a unique pattern (Fig. 3a), whose centroid is detected by a preliminary image localisation and a corner finder refinement. In particular, the detection algorithms are applied to the digital images captured by the stereo vision system used for the range map acquisition.

Initially, the original grey-scale images I_O are filtered by applying the morphological operators of dilation and erosion [26,27]. The intensity value at pixel (x_i,y_i) is substituted by the difference between the highest and the lowest grey intensities in a neighbourhood window of 3 × 3 pixels (x,y) centred at the given pixel as follows:

$$I_F(x_i,y_i) = \max\{I_0(x_i-x,y_i-y)\} - \min\{I_0(x_i-x,y_i-y)\} \qquad (1)$$

As example, Fig. 3a shows several fiducial markers located on different background grey level intensities. Fig. 3b shows the image obtained by applying the filtering operators to the image of Fig. 3a. The result consists of enhancing the marker patterns within the original grey-scale image.

The filtered grey-scale image (Fig. 3b) is then binarised using a threshold value τ (Fig. 3c). The grey intensity probability distribution (Fig. 4a) of the filtered image presents a substantial positive skeweness due to the preponderance of dark over bright pixels. For this reason, the threshold value τ is determined as the grey intensity value corresponding to 99% of the maximum value of the cumulative distribution function (Fig. 4b).

A flood-fill operation is then applied to the binarised image (Fig. 3c) in order to fill the interior of each enhanced marker.

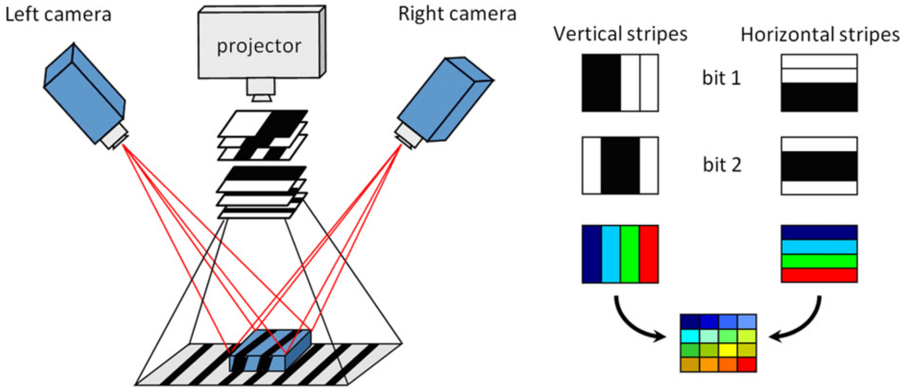


Fig. 1. (a) Projection and acquisition scheme of structuring light stripes and (b) double coding procedure.