Machine Vision and Applications (2012) 23:217–229
DOI 10.1007/s00138-011-0340-1

ORIGINAL PAPER

Three-dimensional point cloud alignment detecting fiducial markers by structured light stereo imaging

Sandro Barone · Alessandro Paoli · Armando Viviano Razionale

Received: 8 October 2009 / Revised: 7 July 2010 / Accepted: 18 April 2011 / Published online: 15 May 2011

Abstract In recent years, various methodologies of shape reconstruction have been proposed with the aim at creating Computer-Aided Design models by digitising physical objects using optical sensors. Generally, the acquisition of 3D geometrical data includes crucial tasks, such as planning scanning strategies and aligning different point clouds by multiple view approaches, which differ for user's interaction and hardware cost. This paper describes a methodology to automatically measure three-dimensional coordinates of fiducial markers to be used as references to align point clouds obtained by an active stereo vision system based on structured light projection. Intensity-based algorithms and stereo vision principles are combined to detect passive fiducial markers localised in a scene. 3D markers are uniquely recognised on the basis of geometrical similarities. The correlation between fiducial markers and point clouds allows the digital creation of complete object surfaces. The technology has been validated by experimental tests based on nominal benchmarks and reconstructions of target objects with complex shapes.

Keywords Reverse engineering · Stereo vision · Digital image processing · Multiple view scanning · Marker detection

S. Barone · A. Paoli (✉) · A. V. Razionale
Department of Mechanical, Nuclear and Production Engineering, University of Pisa, Largo Lucio Lazzarino n. 1, 56100 Pisa, Italy
e-mail: a.paoli@ing.unipi.it

S. Barone
e-mail: s.barone@ing.unipi.it

A. V. Razionale
e-mail: a.razionale@ing.unipi.it

1 Introduction

The reverse engineering techniques [1] are widely applied in many applicative fields, such as industrial manufacturing, automotive and aerospace, design of arts and bioengineering. The first activity of a reverse engineering process concerns shape acquisition by a 3D measurement system. Generally, 3D measurement systems can be classified into *point-by-point* and *full field* techniques.

Point-by point systems digitise physical surfaces by means of either a mechanical probe or a laser pointer integrated to a computerised numerical control machine. These measurements typically involve accurate, but slow acquisition processes to reconstruct complete objects.

Innovations in optical technologies have enabled the development of *full-field* techniques, which allow fast shape reconstructions by multi-data measurements. In particular, active methodologies using plane-light projection to generate 3D point clouds, have become one of the most experienced research areas in optical metrology [2].

Generally, the complete reconstruction of surface geometries by *full-field* techniques requires multiple view strategies, which have to be planned on the basis of dimensions of the target object, presence of occlusions (shadowing and obstructions), optical set-up and expected accuracy of the final model [3]. Clearly, some compromise must be defined between these conflicting requirements. Multiple views can be obtained moving a scanning system and/or a target object and aligning all the partial scans with reference to a common coordinate system. The major problem in combining multiple views is the automatic computation of the best transformation parameters (translation and rotation), which relate information in one view with information in other views [4].

In this paper, a fully automatic alignment procedure of 3D point clouds captured by a *full-field* technique, has been

developed on the basis of recognition of fiducial markers in 3D scenes. Point clouds and fiducial markers are contextually measured by an active stereo vision system based on structured lighting and encoded pattern analysis. Fiducial markers have a unique pattern, which is unambiguously and accurately detected in 2D images on the basis of automatic intensity-based analyses. Then, stereo correspondences are established by combining epipolar constraint and encoded pattern assumptions. Finally, a 3D frame of fiducial markers is created and used as reference for aligning point clouds captured by the structured lighting stereo system. The reference frame is continuously updated including new recognised markers on the basis of a coding approach using 3D geometrical similarities.

This methodology allows fast and accurate alignment of multiple views without requiring either particular textures on target objects or prior information regarding the optical sensor viewpoints.

In the following sections, the multiple view alignment process is overviewed. The proposed approach is described and the experimental results are presented and discussed.

2 Related work

3D *full field* techniques require point cloud alignments, which are typically based on two sequential steps: coarse alignment and global optimisation (i.e. *registration*). Coarse alignment consists of rough rigid transformations of 3D point clouds, whereas global optimisation is usually based on unattended iterations of the Iterative Closest Point (ICP) algorithm [5]. In this context, a good coarse matching is an essential prerequisite for a successful ICP convergence, which involves an iterative solution aimed at minimising the mismatch between corresponding points in overlapping regions of adjacent point clouds.

The coarse alignment can be based on interactive procedures to manually select three or more geometrical, or texture, references on overlapping areas. Fiducial markers can be used as references to compute the transformation parameters [6], though, their manual selection is time consuming.

In recent years, technical literature has proposed various multiple view techniques based on fully unmanned processes. In particular, the proposed methodologies can be classified into two main approaches: *tracking-based* and *shape-based* techniques [7–14].

Tracking-based techniques use independent devices that track relative motions between target object and scanner throughout the scanning process. Turntables can be used to move the object, whereas the scanner can be mounted on passive mechanical arms. More complex hardware can be used for fast tracking; for example, alignment procedures can be developed integrating a measurement device and a robot arm [7] or a coordinate measuring machine [8]. These approaches enable fast scans of target objects, even with complex shapes. However, the accuracy of alignments depends on the resolution of the tracking device. High precision translation/rotation controlled axes are not always affordable in economical terms and require an additional calibration with reference to the 3D vision system. Moreover, tracking systems can only support limited working volumes.

Shape-based techniques are based on automatic detection of invariant significant 3D morphological properties of the object surface [9–13]. In particular, correlations between partially overlapped adjacent point clouds can be established on the basis of geometric descriptors identifying physical variations of surface shapes. These methods do not require any prior knowledge of the original sensor view point, though information on scanning plan could be used to support the localisation of shape descriptors [14]. The greatest disadvantage of these methods are due to the requirement of surface features in the overlapped areas. Flat or uniform curvature geometries are not practically appropriate for shape-based alignments. Moreover, large percentage of computation time is usually required in pre-processing activity, which includes extraction of invariant shapes and organisation of the extracted primitives.

In this context, a further step toward the development of an efficient solution for the alignment of 3D point clouds could be based on a fully automatic recognition of fiducial markers fixed onto the target object.

Fiducial points are automatically detected in several applications: (i) Augmented Reality systems [15,16], (ii) development of indoor [17–19] or outdoor [20] real-time camera tracking systems, (iii) reconstruction and recognition of human faces [21,22] and (iv) medical imaging [23,24]. Fiducial targets have also been used to self-identify patterns for camera calibration purposes [25].

These applications require specific markers, which differ for shape, colour, material (passive as retro-reflective material or active using light emitting diodes) and encoded information (bar codes or special unique patterns) [26].

Although fiducial markers have been used in several contexts, their application to align 3D point clouds has not been fully evidenced in technical literature, but only some industrial solutions have commercially been proposed [27,28].

The aim of the present work was to develop an automatic procedure, which is able to accurately align point clouds measured by a 3D stereo vision system. The proposed methodology integrates a passive marker detection and a 3D full-field scanning process. The markers are uniquely recognised by an automatic procedure, which encodes 3D references on the basis of geometrical similarities.

This approach does not require expensive devices, such as tracker units, or long and tedious interactions, such as

manual detection of features localised in a scene. Moreover, the methodology is not constrained by geometrical attributes and texture peculiarities of the target object.

3 The proposed methodology

The procedure consists of a combination of a 3D stereo vision system based on a projection of encoded light stripes and an automatic multiple-view approach using a digital photogrammetric technique.

3.1 Projection of encoded light stripes

The stereo vision system is based on a structured lighting approach [6], which uses binary patterns in order to capture three-dimensional shapes. The system is composed of a standard XGA DLP multimedia projector and two digital monochrome cameras. The projector has a native resolution of 1024 × 768 pixels and is used to generate vertical and horizontal black and white striped light patterns. The two digital cameras have a resolution of 1280 × 960 pixels ($r_x = 1280$, $r_y = 960$) and are equipped with lenses having a focal length of 12 mm. The optical devices are fixed to an adjustable support and connected to a personal computer through a dual-head video board and an IEEE 1394 data acquisition board (Fig. 1).

The stereo vision approach is based on calibrating the optical configuration by evaluating the intrinsic (focal distances, co-ordinates of the principal points, radial and tangential distortions) and the extrinsic parameters [29]. The stripe projector is un-calibrated and not directly involved in measurement processes.

In this work, a double binary encoded light stripe approach is used for 3D shape recovery. In particular, a line correlation process is adopted by projecting a sequence of lines defined as crossing zones between white and black parallel stripes with periods progressively halved (Fig. 2). Each pixel in the camera images is characterised by a light intensity that can be either bright or dark depending on its localisation in the respective plane image. A binary code (0,1 with n bit) is assigned to each pixel, where n is the number of stripe patterns, and the values 0 and 1 are associated with the intensity levels, i.e. 0 = *black* and 1 = *white*. This encoding procedure provides $\ell = 2^n - 1$ encoded lines by using n patterns (Fig. 3).

Fig. 1 Stereo vision system integrated to a digital projector

The proposed methodology solves the problem of point stereo correspondence by projecting both horizontally and vertically striped encoded patterns [30]. A double code is assigned to the intersection points through the horizontal and vertical stripes. This yields an automatic and unique correspondence of conjugate points in the camera images. The resolution that can be obtained by the proposed procedure is $\ell_v \times \ell_h$ points, where ℓ_v and ℓ_h are the number of vertical and horizontal lines switched by the projector, respectively.

In this work, the maximum resolution of the depth measurement is obtained by shifting one of the stripe images three times by the width of one pixel. Taking into account the resolution of the projector integrated to the stereo system, this method provides 784.641 encoded points ($\ell_v = 1023$, $\ell_h = 767$) by projecting nine stripe patterns with periods progressively halved and three stripe patterns by shifting the finest stripes (the 9th stripe pattern). This approach allows the resolution of the projector to be exploited avoiding fine stripes, which may typically produce poor accuracy in line extraction.

The 3D coordinates of the scene points are then recovered by assuming the known image acquisition geometry and using triangulation of conjugate points.

3.2 Automatic multiple-view approach

The complete shape recovery generally requires the acquisition of multiple point clouds from different views, which have to be transformed into a common coordinate system. The alignment process is carried out referring each point cloud to a frame of fiducial markers detected in the grey-value images captured by the stereo vision system (Fig. 4).

The method is composed of the following main steps:

1. detection of fiducial markers in grey value images;
2. stereo correspondence and 3D point triangulation; and
3. 3D graph matching.

The procedure is automatically and progressively applied to each acquisition, allowing the operator to contextually define the best scanning strategy.

3.2.1 Marker detection in grey-value images

The markers used in this work are shown in Fig. 5a, b for white and dark surfaces, respectively. These patterns provide

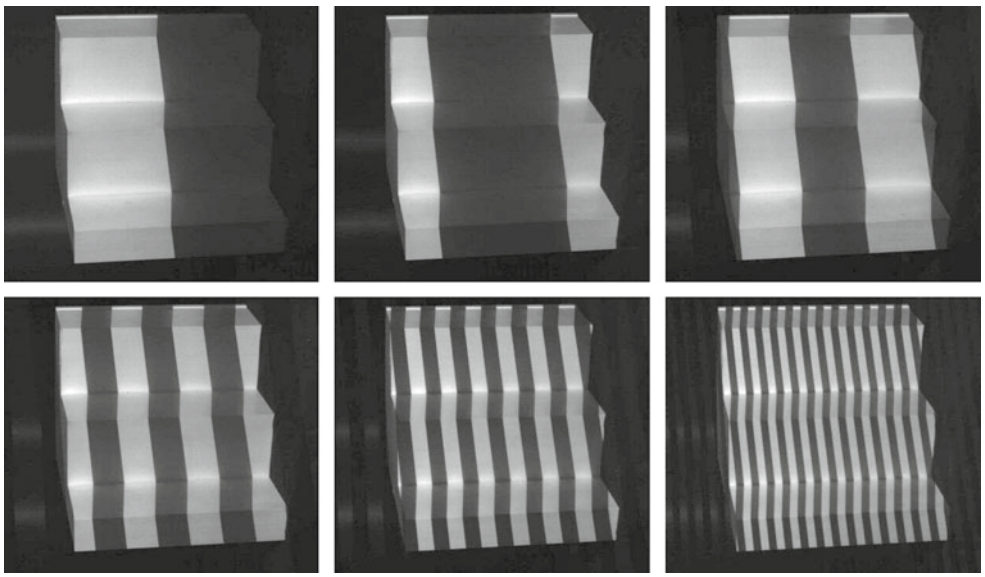

Fig. 2 Projection of vertical structuring light stripes

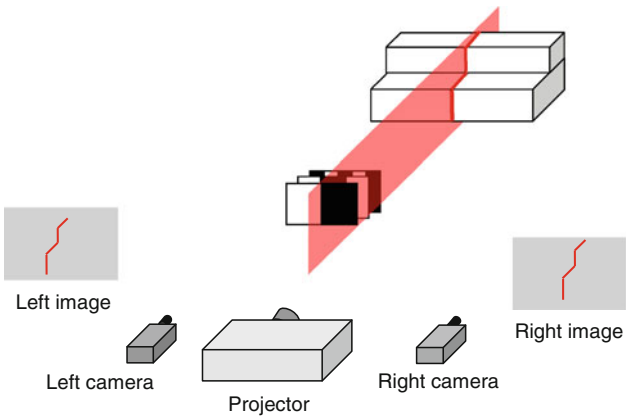

Fig. 3 Scheme of line correspondence of the stereo vision

a good contrast with all typologies of surrounding regions, allowing the detection of fiducial points with sub-pixel accuracy combining coarse localisation by image-processing techniques and fine adjustment by a corner finder algorithm.

The coarse fiducial detection is based on the calculation of the centroid of a circular marker evidenced by the exterior black ring. Circular markers are typically used since their centroids are easily measured with sub-pixel accuracy [20].

The original image is first filtered to outline the marker edges. A flood-fill operation is then performed in order to fill the holes delimited by the edges and to produce elliptical features. The result is binarised using a local adaptive threshold and the resulting connected components are labelled becoming candidates for fiducial detection. The final classification of the connected components is performed measuring a set of properties of the labelled regions and verifying their membership with respect to fixed threshold values.

Operatively, the first step of the marker detection procedure consists of the texture segmentation of the greyscale images. In particular, a statistical measurement has been used to characterise the local texture of the images. The intensity value at each pixel of the greyscale image is substituted by the value obtained computing the difference between the highest and the lowest intensity values in a neighbourhood window (*structuring element*) of 3×3 pixels centred at the given pixel. This procedure is implemented by independently applying the morphological operators of dilation and erosion [31,32] to the original images. The dilation of an image f by a structuring element b is denoted by $\delta_b(f)$ and the dilated value at a given pixel (x, y) is the maximum value of the image defined by the structuring element when its origin is at (x, y):

$$[\delta_b(f)](x, y) = \max\{f(x - x_0, y - y_0) | (x_0, y_0) \in D_b\} \tag{1}$$

where (x_0, y_0) represents an arbitrary pair of coordinates in the domain of D_b.

Similarly, the erosion of an image f by a structuring element b is denoted by $\varepsilon_b(f)$ and the eroded value at a given pixel (x, y) is the minimum value of the image in the window defined by the structuring element when its origin is at (x, y):

$$[\varepsilon_b(f)](x, y) = \min\{f(x - x_0, y - y_0) | (x_0, y_0) \in D_b\} \tag{2}$$

As example, Fig. 6a shows a motorcycle component captured by using the stereo vision system. Figure 6b shows the complement of the image obtained by applying the filtering operators to the image of Fig. 6a. The result consists of

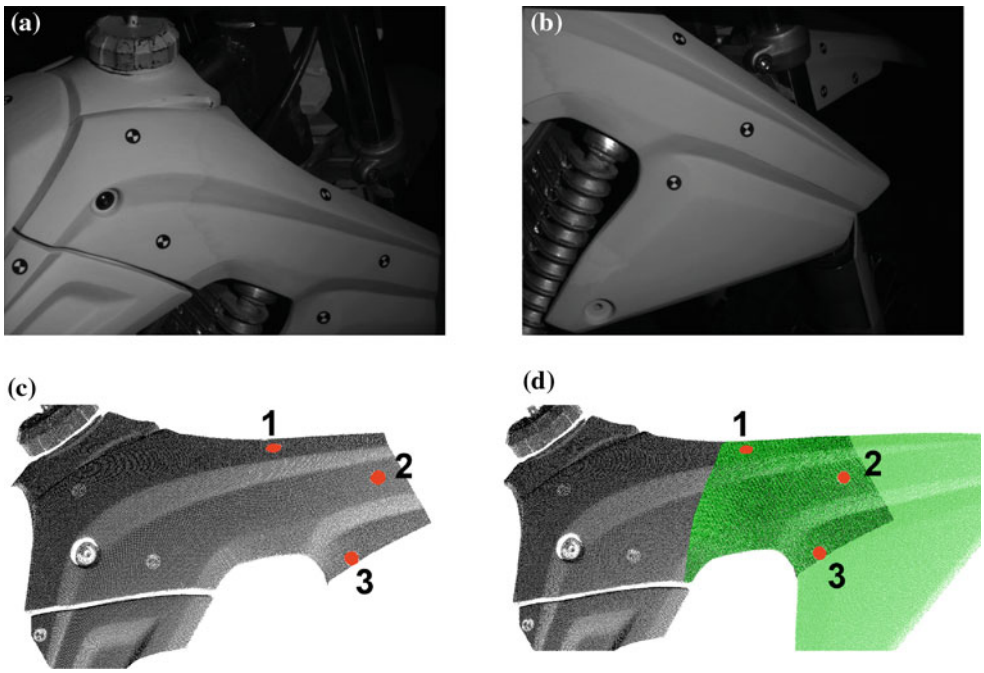

Fig. 4 Acquisition of a motorcycle component from two different views (**a**, **b**), and point cloud alignment using fiducial markers (**c**, **d**)

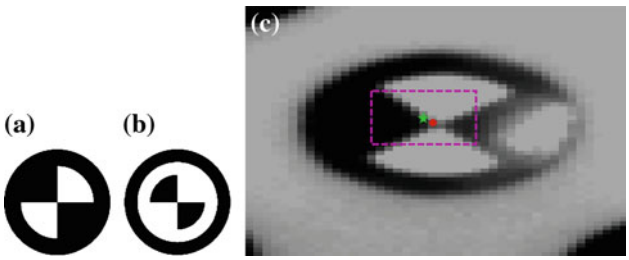

Fig. 5 Marker pattern for white objects (**a**), for dark objects (**b**) and marker detection (**c**) using (*circle point*) and not using (*star point*) a corner-finder refinement

enhancing the marker patterns within the original greyscale image.

A flood-fill operation is then applied on the filtered greyscale image (Fig. 6b) in order to fill the interior of each enhanced marker. The algorithm is based on a greyscale morphological reconstruction [32]. A new reference image (f_r) is determined as

$$f_r(x, y) = \begin{cases} 1 - f(x, y) & \forall\,(x, y) \in \text{border} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Then, the reference image is processed by repeated dilations using a 4-connected element. Each dilation is constrained to lie underneath the filtered greyscale image. When a dilation does not change the image, the process stops and the final dilation represents the reconstructed image (Fig. 6c). Practically, intensity values of bright areas, surrounded by darker areas, are brought to the same intensity value as surrounding pixels, so that fiducial areas are filled.

The detection of fiducial markers is then reduced to a filled ellipse recognition. The reconstructed image (Fig. 6c) is binarised using a local adaptive threshold to handle variations in scene lighting (Fig. 6d). The image is split into partially overlapping rectangular blocks having the same size ($r_x/10, r_y/10$). For each block the threshold value is determined by the Otsu's method [33]. A bilinear interpolation of the threshold values is then performed for each pixel. Then, all the connected components are recognised and labelled (Fig. 6e). Among all these connected regions, possible candidates to represent fiducial markers are extracted by considering a set of few attributes. Elliptical shapes are detected by calculating four different properties for each label: area (A), perimeter (P), eccentricity (e) and extent (*ext*, specify the proportion of the pixels in the bounding box that are also within the connected region). Maximum and minimum threshold values have been defined on the basis of image resolution for area and perimeter values. Eccentricity may vary between the degenerate cases 0 (circle) and 1 (line segment). A maximum of 0.9 has been chosen to discard markers, which are considered too oblique. The ellipse extent is theoretically expressed by $\pi/4$ and the adopted range takes into consideration the pixel discretisation and the border effect. 2D coordinates of the markers are preliminarily evaluated as intensity centroids of the connected components, whose properties verify at the same time all the above constraints.

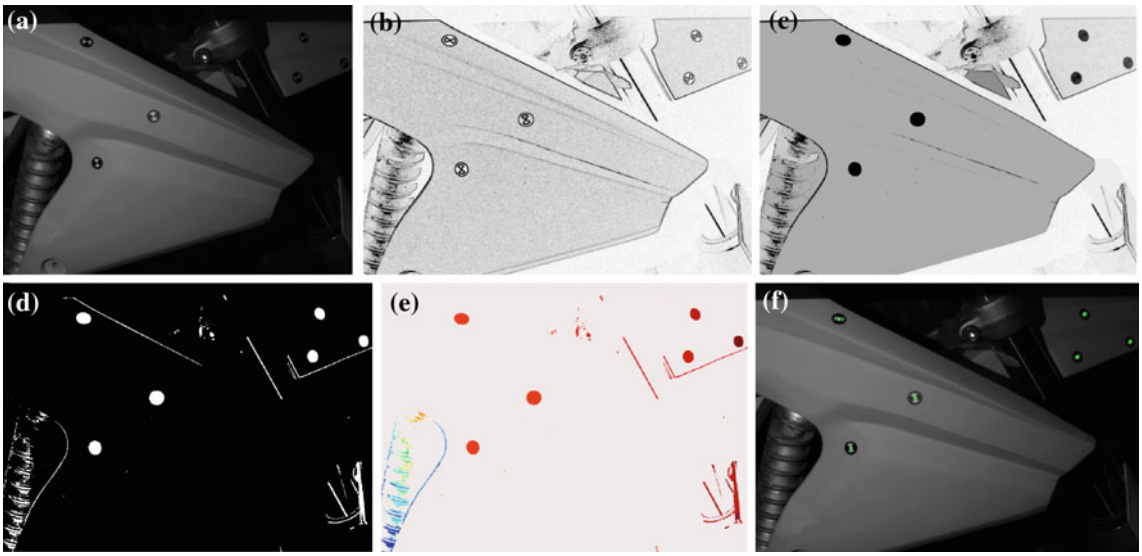

Fig. 6 Marker detection process: **a** original image, **b** complement of the filtered image, **c** result of the flood-fill operation, **d** binarisation, **e** labelling, **f** centroids extraction

Experimental tests have pointed out that the stereo correspondence algorithms are generally influenced by errors occurring in the marker detection process. Image processing introduces a little modification to the real marker shape, with the consequence of slightly moving the points detected from the exact cross centres. In this paper, a refinement process is applied using the Harris corner-finder algorithm [34], exploiting the particular pattern of the markers (Fig. 5a, b). An adaptive window is defined on the basis of the geometrical attributes (dimension, shape, orientation) of the given connected component and centred on the preliminary centroid (Fig. 5c).

3.2.2 Stereo correspondence constraints

In this paper, a robust stereo correspondence of markers is proposed by combining *epipolar line* and *encoded light stripe* constraints.

The epipolar geometry describes the relation between two images captured from two different views. Operatively, given a marker m_R detected on the image I_R (right camera), its conjugate m_L on the image I_L (left camera) belongs to a line, i.e. *the epipolar line*, defined as intersection of the image plane with the plane defined by m_R and the focuses of the two optical devices, i.e. *the epipolar plane*. The equation of the epipolar line (l_P) of m_R can be written as $l_P = F \times m_R$ on the basis of the perspective projection theory [3], where F is a 3×3 square matrix (*Fundamental Matrix*) depending on the epipolar parameters and the perspective projection matrices of the optical devices, which can be obtained by the calibration procedures.

The correspondence problem would then be straightforward considering that, each marker candidate m_R in the right image can only correspond to such point in the left image that lies on the relative epipolar line. Operatively, an algorithm of image rectification [35] has been used to simplify the correspondence search. However, false positives may erroneously occur due to the presence in the 3D scene of irregular details (Fig. 7). In this paper, an improvement of correspondence assignment is achieved by matching the codes extracted at each marker candidate area on the basis of the horizontal encoding procedure used for point cloud reconstruction. The correspondence candidates are correlated only if both the constraints, i.e. epipolar geometry and encoding light stripe, are satisfied. This approach avoids false positives occurring at irregular features, which generally have inaccurate light stripe code and remarkably reduce the inter-marker

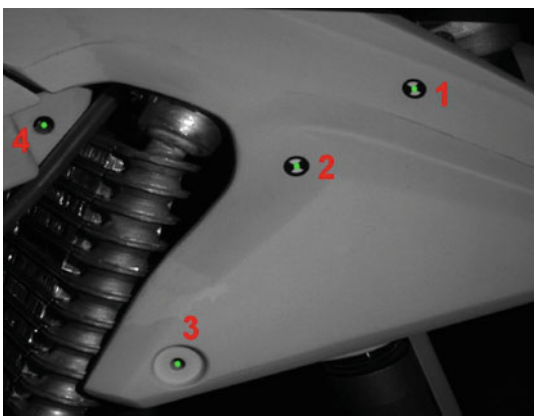

Fig. 7 False positives (*points 3* and *4*) in correspondence of circular features of the target object of Fig. 6a

confusion rate, since more than one marker candidate could lie on the same epipolar line.

The three-dimensional coordinates of the markers M are determined by a triangulation procedure given the 2D conjugate pairs m_R, m_L. The combination of two constraints, i.e. encoding light stripe and epipolar line, allows the matching problem to always have an unique solution.

3.2.3 3D graph matching

The last step of the point cloud alignment procedure consists of 3D matching of fiducial marker sets detected from different views (Fig. 8). The aim is at progressively creating a 3D skeleton of markers detected during the multiple-view scanning process.

The matching problem is efficiently reduced to a search of similar triangles within two data sets **p** and **q** constituted by M_p and N_q markers, respectively, captured in different measurements (with $p, q \geq 3$).

In particular, the algorithm can be structured in the following steps:

1. The $p \times p$ symmetric adjacency matrix $\mathbf{D}^p$ is associated to the set $\mathbf{p} = \{M_1 \ldots M_p\}$; the elements of the matrix are composed by the Euclidean distances $d^p_{i,j} = \|M_i - M_j\|$.
2. Analogously, the $q \times q$ symmetric adjacency matrix $\mathbf{D}^q$ is associated with the set $\mathbf{q} = \{N_1, \ldots N_q\}$; the elements of the matrix are composed by the Euclidean distances $d^q_{k,l} = \|N_k - N_l\|$.

Each column of the matrix $\mathbf{D}^p$ is compared with all the columns of the matrix $\mathbf{D}^q$; in particular, the differences of the distance values are computed and compared with a fixed tolerance ε, as

$$\left\| d^p_{i,j} - d^q_{k,l} \right\| < \varepsilon \quad \text{with } d^p_{i,j} \neq 0,\ d^q_{k,l} \neq 0 \tag{4}$$

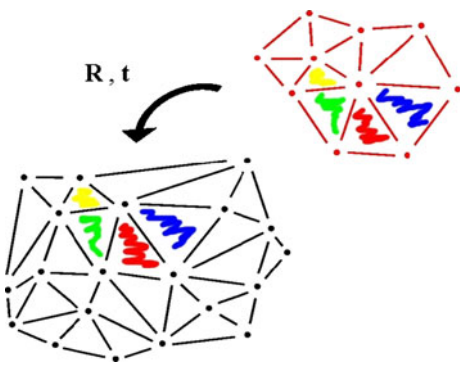

Fig. 8 3D matching between sets of markers (global and incoming frames)

If relation (4) is satisfied, the couple k, l is stored into an array of pointers representing a list of possible pairs of points (N_k, N_l) corresponding to (M_i, M_j). The search is iterated until relation (4) is satisfied at least twice for the same column: two corresponding pairs and, consequently, three common markers have been identified between the data sets **p** and **q**. The iterations are carried out varying the tolerance ε from a small value ($\varepsilon = 0$) to a fixed threshold ($\varepsilon = 0.5$ mm) by incremental steps of 0.01 mm.

The equivalence among all the remaining distances defined by matched marker pairs is then verified, allowing the final detection of similar triangles. Possible geometrical ambiguities, such as collinear configurations, are avoided checking all the angles between the matched markers. Given the set $\mathbf{c} = \{O_1 \ldots O_c\}$ of markers matched at the previous step, with c>=3, relation (5)

$$\min_{i=2\ldots c} \left\{ abs \left[\left(\frac{\overline{O_i O_{i-1}}}{\left\| \overline{O_i O_{i-1}} \right\|} \right) \cdot \left(\frac{\overline{O_i O_{i+1}}}{\left\| \overline{O_i O_{i+1}} \right\|} \right) \right] \right\} < \frac{\sqrt{3}}{2}$$
$$\text{where } c + 1 = 1 \tag{5}$$

ensures that exists at least one angle between 30° and 150°. If none of the angles satisfy to this condition, the iteration continues until another pair of corresponding markers has been detected.

Once the correspondent pairs of fiducial markers have been identified and geometrically verified, the translation **t** and rotation **R**, defining the relation between the two adjacent point clouds, are directly determined by applying the singular value decomposition method [36].

4 Experimental results

The proposed procedure has been tested reconstructing the shapes of nominal samples and free-form models regarding automotive industry.

A sphere and a cylinder (Fig. 9) have been used for initial testing. The sphere (Fig. 9a) has a radius of $r_{\text{sph\_mec}} = 44.513$ mm; this value has been obtained averaging five different measurements (standard deviation = 0.013 mm) by an

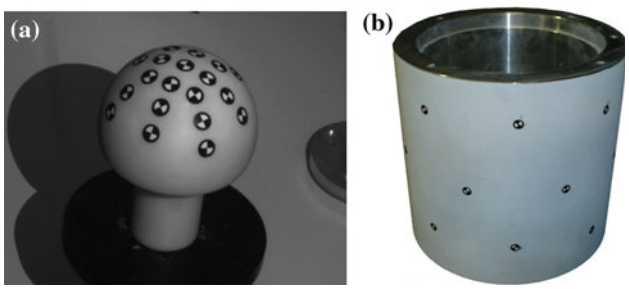

Fig. 9 Nominal samples used in the experimental tests: sphere (**a**) and cylinder (**b**)

independent mechanical probe. The same methodology has been used to measure the cylinder dimension (r_{cyl_mec} = 123.976 mm and standard deviation = 0.006 mm).

The stereo vision system has been configured for a working distance of 900 mm with a measuring area of 250 mm × 250 mm, lateral resolution of 0.2 mm and accuracy of 0.03 mm. The accuracy has been evaluated comparing data obtained measuring the cylinder surface by the mechanical probe and the structured lighting system. In particular, single-point clouds have been used to compute the radius of the cylinder through best fitting techniques.

Experimental tests have been carried out in order to evaluate the influence of the optical set-up on the marker detection algorithm.

Twenty-five fiducial markers have been disposed on the sphere surface along four different meridians varying the inclination angle (i.e. zenith angle), as shown in Fig. 9a. The sphere has been approximately disposed with the central marker normal to the viewing direction of the optical scanner, represented by the bisector of the angle between the two cameras. Figure 10a, b shows a pair of images captured by the stereo cameras; Fig. 10c shows the marker distribution obtained processing the stereo images by the proposed methodology.

Twenty different stereo images have been captured and processed by slightly and randomly rotating the sphere around the viewing direction. Thus, the influence of both marker orientations and mutual occlusions between the optical sensors have been analysed. Figure 11 shows the contour lines of the percentage values of the markers successfully detected by processing the 20 stereo images by using (Fig. 11a) and not using (Fig. 11b) the corner-finder refinement. The results have been displayed in a spherical coordinate system using the zenith angle and the *azimuth angle* measured on a reference plane through the sphere centre and orthogonal to the viewing direction (Fig. 10c). The experimental tests have pointed out a maximum zenith angle of 45° to obtain a full marker detection using the corner-finder refinement (Fig. 11a). The detection percentage decreases with the increasing of the zenith angle and becomes almost zero at 60°. Stereo occlusions generate elliptical shapes of the contour lines with the minor axis orientated along the baseline direction of the stereo vision system. This phenomenon reduces the detection percentages of the fiducial markers localised along the baseline direction with respect to those localised along the direction perpendicular the baseline.

Moreover, Fig. 11b outlines the effects occurring not using the corner-finder algorithm. The refinement in computing 2D fiducial coordinates allows the epipolar constraint to increase the stereo correspondences rate. For example, reflective spots, due to projector lighting (such as, the central marker showed in Fig. 10b), could provide incorrect shapes

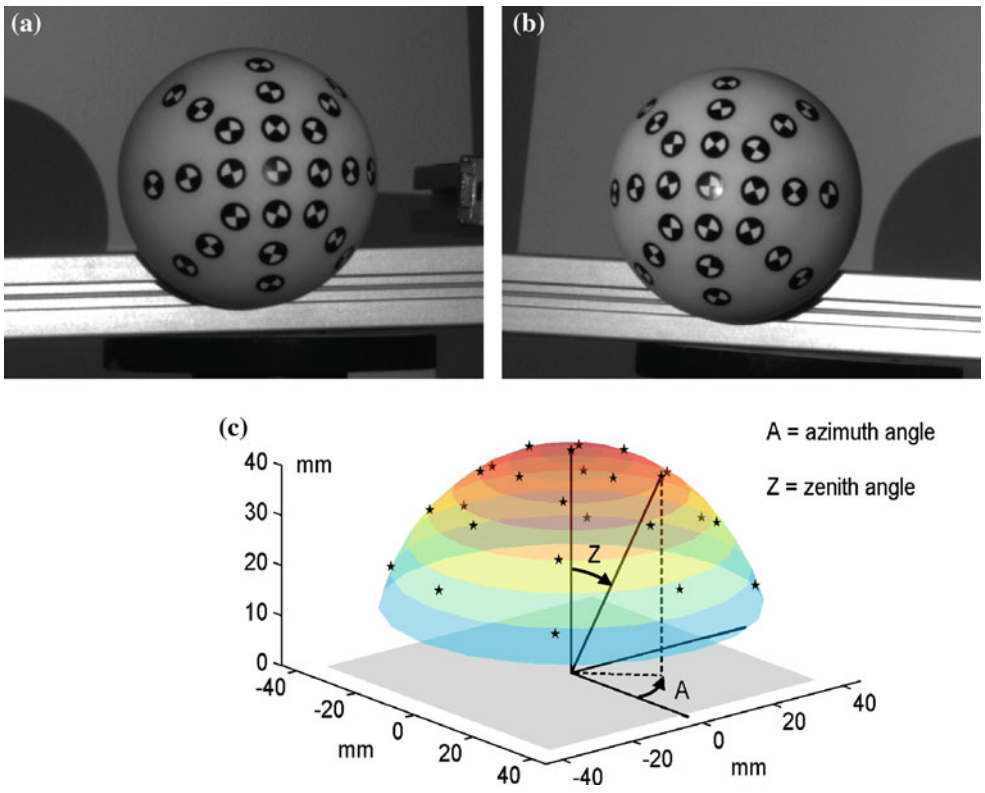

Fig. 10 *Left* (**a**) and *right* (**b**) images as captured by the stereo vision system and 3D marker coordinates disposed on the nominal sphere surface (**c**)

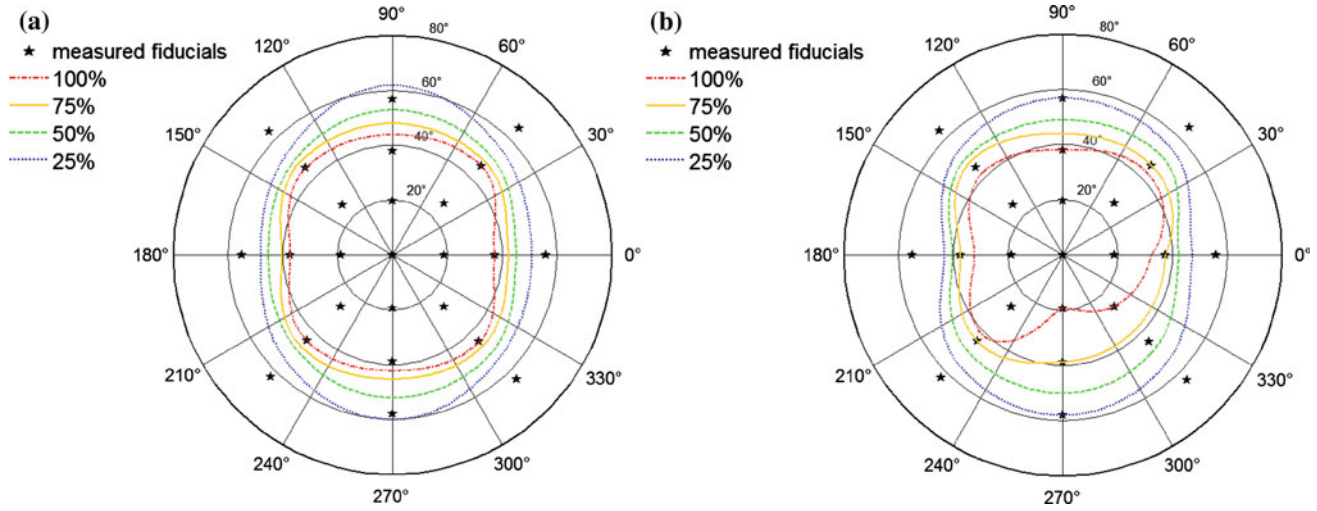

Fig. 11 Contour lines of the percentage values of the 3D markers correctly detected in the sphere of Fig. 9a by applying (**a**) or not (**b**) the corner-finder refinement

Table 1 Absolute values and standard deviations of the errors calculated as difference between the radius (r_{sph_mec}) measured by the mechanical probe and the radius (r_{sph_opt}) of the sphere obtained by best fitting the aligned point clouds

Test	Automatic with corner-finder		Automatic without corner-finder		Manual detection		Rotating table	
	$\|\|r_{sph\_mec} - r_{sph\_opt}\|\|$ (mm)	σ_{r_sph} (mm)	$\|\|r_{sph\_mec} - r_{sph\_opt}\|\|$ (mm)	σ_{r_sph} (mm)	$\|\|r_{sph\_mec} - r_{sph\_opt}\|\|$ (mm)	σ_{r_sph} (mm)	$\|\|r_{sph\_mec} - r_{sph\_opt}\|\|$ (mm)	σ_{r_sph} (mm)
1	0.023	0.043	0.225	0.181	0.247	0.183	0.01	0.054
2	0.011	0.07	0.169	0.25	0.19	0.196	0.028	0.052
3	0.074	0.054	0.091	0.169	0.243	0.17	0.034	0.050
4	0.07	0.057	0.069	0.116	0.397	0.226	0.035	0.055
5	0.045	0.052	0.277	0.182	0.131	0.116	0.034	0.056

affecting the stereo-correspondence technique (Fig. 5c). The use of a corner-finder algorithm minimises the impact of lighting disturbs in 3D scenes: practically, higher number of markers are framed as evidenced by the regular shapes of the contour lines shown in Fig. 11a with respect to those shown in Fig. 11b.

The multiple-view approach has been evaluated by capturing the complete shape of the sphere by the lighting stripe projection. The results have been compared with data obtained by manual selection of markers and by a tracking turntable.

Table 1 reports the results (errors and standard deviations) obtained from five independent tests. The absolute values of the errors have been calculated as difference between the radius (r_{sph_mec}) measured by the mechanical probe and the radius (r_{sph_opt}) of the sphere obtained by best fitting the aligned point clouds. The results have been obtained by applying the corner-finder refinement. Moreover, the effects occurring without applying the corner-finder algorithm are indicated to evidence the accuracy improvement. Figure 12 shows the distributions of the optical measurement deviations from the best fitting sphere for each of the alignment procedures, along with boxplot descriptions. The spread of the results obtained by manually selecting the fiducial markers is justified by the subjectivity of the approach. The proposed methodology and the turntable-based technique are almost comparable in terms of accuracy and processing time (few seconds to align ten point clouds).

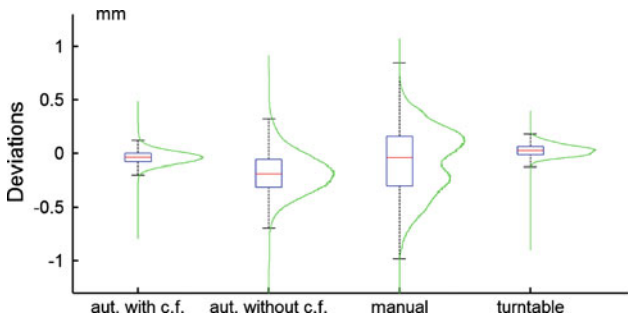

Fig. 12 3D deviations along with box-plots from the best fitting sphere for each of the alignment procedures (automatic with corner finder, automatic without corner finder, manual, turntable)

The developed methodology has been also validated reconstructing the shape of the cylinder shown in Fig. 9b. In particular, eight different scans have been aligned using a minimal number of markers, which have been distributed so that at least three references (not collinear) were in the overlapping areas. All the markers have been automatically recognised by the proposed procedure. Table 2 reports the results (error and standard deviations) obtained from five independent tests. The absolute values of the errors have been calculated as difference between the radius (r_{cyl_mec}) measured by the mechanical probe and the radius (r_{cyl_opt}) of the cylinder obtained by best fitting the aligned point clouds. A maximum error of 0.032 mm, with a standard deviation of 0.058 mm, has been evaluated.

Finally, a further test has been conducted by capturing the shape of a motorcycle helmet (Fig. 13) made of polyester reinforced with glass fibres and having a bounding box of 350 mm × 250 mm × 250 mm.

Figure 14a, b shows the green dots corresponding to the detected markers processing left and right grey intensity images. Some false positives are clearly visible in correspondence of particular features of the helmet (holes, button fastening). The application of both the epipolar constraint and encoded lighting projection has refined the detection procedure rejecting false positives. Figure 14c shows the red dots obtained solving the correspondence problem.

The entire object has been reconstructed using 18 different views. Table 3 reports the number of markers to be detected in all the multiple views and the fiducial markers recognised either using or not using the corner-finder algorithm to refine 2D marker detections. The analysis of these values points out that the 2D detection accuracy influences the successive stereo correspondence, reducing both false-positive and false-negative rates that are considered the main indicators of the robustness of a fiducial marker approach [16]. Figure 15a shows the 18 point clouds automatically aligned by the proposed procedure. The overall sample points (∼6 millions) have been reduced and post-processed by proper algorithms of smoothing and filtering (Fig. 15b). The final point cloud (∼0.5 million of sample points) has finally been triangulated to obtain the mesh representation as shown in Fig. 15c. The time required to capture each view scan has been about 10 s (including light stripe projection time), with a total time of about 3 min to digitise the entire object by the completely automatic multi-view process.

Table 2 Absolute values and standard deviations of the errors calculated as difference between the radius (r_{cyl_mec}) measured by the mechanical probe and the radius (r_{cyl_opt}) of the cylinder obtained by best fitting the aligned point clouds

Test	$\|\|r_{cyl\_mec} - r_{cyl\_opt}\|\|$ (mm)	σ_{r_cyl} (mm)
1	0.0320	0.058
2	0.0040	0.042
3	0.0120	0.039
4	0.0050	0.047
5	0.0040	0.051

5 Conclusions

In this paper, a procedure to acquire free form surfaces by combining a 3D stereo vision system and a fully automatic point cloud alignment procedure, based on fiducial markers detection, has been developed. The methodology consists of the acquisition of point clouds by a structuring light stereo vision system and the automatic detection of fiducial markers on overlapping areas of adjacent point clouds by processing the grey intensity images. Markers with a no-coded pattern are used, allowing large identification volumes, since the number of markers is not limited by the pattern uniqueness. The fiducial marker recognition is automatically performed on the basis of 3D geometrical similarities.

Experimental tests have been carried out reconstructing nominal geometries (spherical and cylindrical surfaces) and

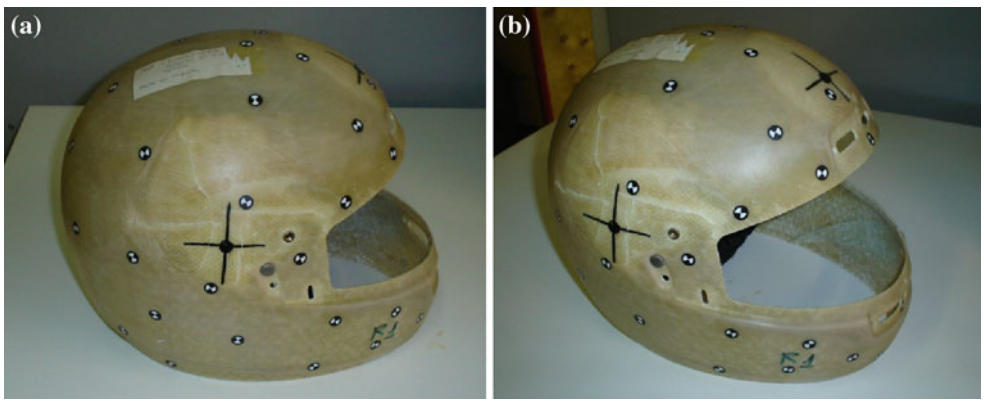

Fig. 13 Two different views of a motorcycle helmet used for testing

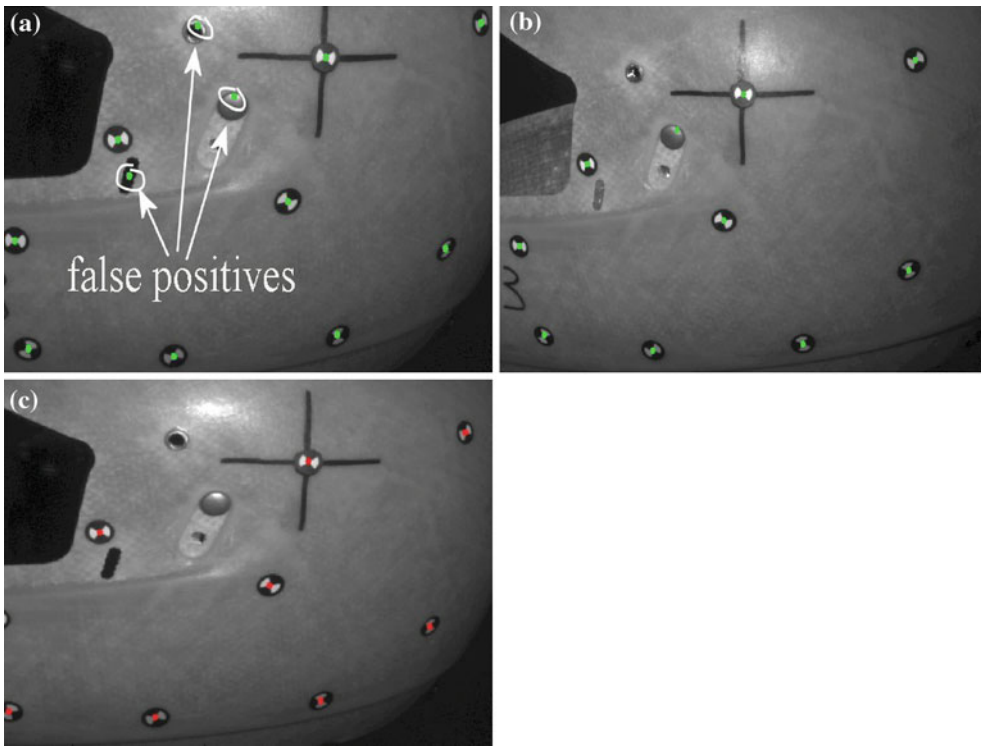

Fig. 14 2D detection of markers by processing *left* (**a**) and *right* (**b**) grey intensity images of the helmet shown in Fig. 13, and results obtained solving the correspondence problem (**c**)

Table 3 Fiducial markers and false positives obtained by processing 18 range views used for the complete reconstruction of the helmet

	Markers to be detected	Detected markers	False positives
Coarse detection	215	189	12
Refined detection	215	201	4

a motorcycle helmet model. The experimental tests have evidenced that the results can be influenced by the specific operative conditions in terms of configuration of the stereo vision system, multiple view strategy, texture peculiarities of target objects and environmental lighting. However, the proposed methodology does not require particular texture or geometrical restrictions. This enables rapid and high-precision alignment of multiple views, even with objects having large curvatures surfaces and uniform textures. The use of a corner-finder algorithm to adjust the fiducial 2D coordinates, combined with the double constraints of epipolar geometry and structured light encoding, provides a robust stereo correspondence methodology. This approach has demonstrated to reduce both false-positive and false-negative rates.

The robustness of the described methodology relies on an "incremental" approach rather than a "*do everything or fail*" approach. If a single-point cloud matching fails (i.e. at least three fiducial markers are not present), the shape reconstruction process continues and the failed alignment can be rescued as soon as the global marker skeleton has included at least three common markers. Thus, completely free scanning strategies are allowed during the measurement process.

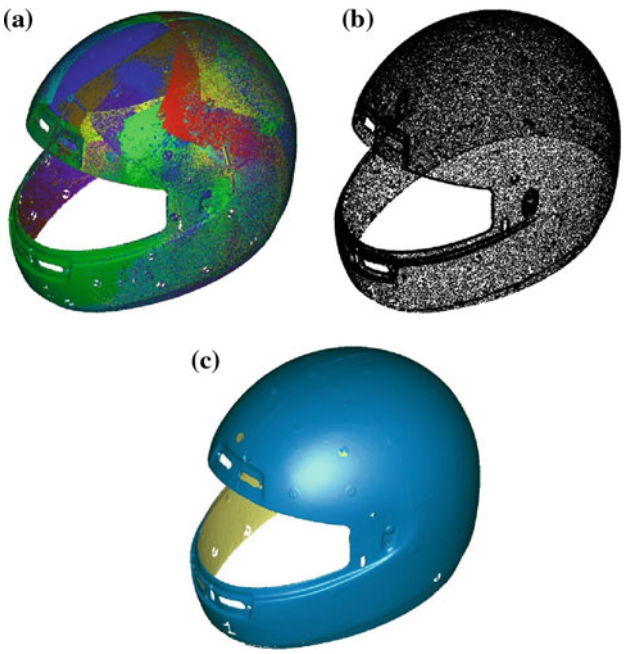

Fig. 15 Aligned point clouds (**a**), 3D points reduced by curvature sample (**b**), StL representation (**c**) of the motorcycle helmet shown in Fig. 13

References

1. Varady, T., Martin, R.R., Cox, J.: Reverse engineering of geometric models—an introduction. Comput. Aided Des. **29**(4), 225–268 (1997)
2. Gorthi, S.S., Rastogi, P.: Fringe projection techniques: whither we are? Opt. Lasers Eng. **48**(2), 133–140 (2010)
3. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision.133–140 Cambridge University Press, Cambridge (2000)
4. Bernardini, F., Rushmeier, H.: The 3D model acquisition pipeline. Comput. Graph. **21**(2), 149–172 (2002)
5. Besl, P.J., McKay, N.D.: A method for registration of 3D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
6. Barone, S., Curcio, A., Razionale, A.V.: A structured light stereo system for reverse engineering applications. In: XIII ADM—XV INGEGRAF, Cassino, Italy (2003)
7. Callieri, M., Fasano, A., Impoco, G., Cignoni, P., Scopigno, R., Parrini, G., Biagini, G.: RoboScan: an automatic system for accurate and unattended 3D scanning. In: 3DPVT'04, pp. 805–812, Thessaloniki, Greece (2004)
8. Sansoni, G., Patrioli, A.: Non contact 3D sensing of free-form complex surfaces. In: Proceedings of SPIE Videometrics and Optical Methods for 3D Shape Measurement, vol. 4309, pp. 232–239 (2000)
9. Gelfand, N., Mitra, N.J., Guibas, L.J., Pottmann, H.: Robust global registration. In: Proceedings of the Third Symposium on Geometry Processing, pp. 197–206 (2005)
10. Huber, D.F., Hebert, M.: Fully automatic registration of multiple 3D data sets. Image Vis. Comput. **21**(7), 637–650 (2003)
11. Li, N., Cheng, P., Sutton, M.A., McNeill, S.R.: Three-dimensional point cloud registration by matching surface features with relaxation labeling method. Exp. Mech. **45**(1), 71–82 (2005)
12. Beinat, A., Crosilla, F., Sepic, F.: Automatic morphological prealignment and global hybrid registration of LiDAR close range images. In: ISPRS Symposium on Image Engineering and Vision Metrology, Dresden, Germany (2006)
13. Rugis, J., Klette, R.: Surface registration markers from range scan data. In: 11th International Workshop on Combinatorial Image Analysis (IWCIA), pp. 430–444 (2006)
14. Pingi, P., Fasano, A., Cignoni, P., Montani, C., Scopigno, R.: Exploiting the scanning sequence for automatic registration of large sets of range maps. Comput. Graph. Forum **24**(3), 517–526 (2005)
15. Rekimoto, J., Ayatsuka, Y.: CyberCode: designing augmented reality environments with visual tags. In: DARE on Designing Augmented Reality Environments, vol. 1, pp. 1–10 (2000)
16. Fiala, M.: Designing highly reliable fiducial markers. IEEE Trans. Pattern Anal. Mach. Intell. **32**(7), 1317–1324 (2010)
17. Ababsa, F., Mallem, M.: A robust circular fiducial detection technique and real-time 3D camera tracking. J. Multimedia **3**(4), 34–41 (2008)
18. Naimark, L., Foxlin, E.: Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2002), vol. 1, pp. 27–36 (2002)
19. Nakazato, Y., Kanbara, M., Yokoya, N.: Localization of wearable users using invisible retro-reflective markers and an IR camera. In: SPIE, vol. 5664, pp. 563–570 (2005)
20. Claus, D., Fitzgibbon, A.W.: Reliable fiducial detection in natural scenes. In: 8th European Conference on Computer Vision (ECCV '04), vol. 4, pp. 469–480 (2004)
21. Galantucci, L.M., Percoco, G., Dal Maso, U.: Coded targets and hybrid grids for photogrammetric 3D digitisation of human faces. Virtual Phys. Prototyp. **3**(3), 167–176 (2008)
22. Bowyer, K.W., Chang, K., Flynn, P.: A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. Comput. Vis. Image Und. **101**(1), 1–15 (2006)
23. Audette, M.A., Ferrie, F.P., Peters, T.M.: An algorithmic overview of surface registration techniques for medical imaging. Med. Image Anal. **4**(3), 201–217 (2000)
24. Wyawahare, M.V., Patil, P.M., Abhyankar, H.K.: Image Registration Techniques: An Overview. Int. J. Signal Process. Image Process. Pattern Recognit. **2**(3), 11–28 (2009)
25. Fiala, M., Shu, C.: Self-identifying patterns for plane-based camera calibration. Mach. Vis. Appl. **19**, 209–216 (2008)
26. Owen, C., Xiao, F., Middlin, P.: What is the best fiducial? In: 1th IEEE International Augmented Reality Toolkit Workshop, Darmstadt, Germany (2002)
27. Creaform.: http://www.creaform3d.com. Accessed 25 June 2010
28. GOM mbH.: Optical measuring techniques. http://www.gom.com. Accessed 25 June 2010
29. Tsai, R.Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J. Robot. Autom. **RA-3**(4), 323–344 (1987)
30. Barone, S., Razionale, A.V.: A reverse engineering methodology to capture complex shapes. In: XVI INGEGRAF, Zaragoza, Spain (2004)
31. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer, Secaucus (2003)
32. Vincent, L.: Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Trans. Image Process. **2**(2), 176–201 (1993)
33. Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. Syst. Man Cybern. **9**, 62–66 (1979)
34. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of the Fourth Alvey Vision Conference, pp. 147–151 (1988)
35. Fusiello, A., Trucco, E., Verri, A.: A compact algorithm for rectification of stereo pairs. Mach. Vis. Appl. **12**(1), 16–22 (2000)
36. Eggert, D.W., Lorusso, A., Fisher, R.B.: Estimating 3-D rigid body transformations: a comparison of four major algorithms. Mach. Vis. Appl. **9**, 272–290 (1997)

Author Biography

Sandro Barone received his Master's Degree in Mechanical Engineering from the University of Palermo in 1990 and Ph.D. Degree in Mechanical Design in 1995. He joined the University of Pisa in 1996, and he is currently Full Professor of Design and Methods of Industrial Engineering at the Department of Mechanical, Nuclear and Production Engineering of the University of Pisa. His main research fields are: design methods by CAD/CAE systems and optical technologies for 3D scanning.

Alessandro Paoli received his Master's Degree in Aerospace Engineering from the University of Pisa in 2003 and Ph.D. Degree in Mechanical Engineering from the same University in 2008. He is currently working at the Department of Mechanical, Nuclear and Production Engineering of the University of Pisa as assistant researcher. His current research interests include structured light-based sensing, image processing, and 3D modelling processes for biomedical applications.

Armando Viviano Razionale received his Master's Degree in Aeronautical Engineering from the University of Pisa in 2001 and the Ph.D. Degree in Mechanical Engineering from the same University in 2005. He is currently working at the Department of Mechanical, Nuclear and Production Engineering of the University of Pisa as Assistant Professor in Design and Methods of Industrial Engineering. His current research interests include optical technologies for 3D scanning, and optical methods for mechanical design.