A new structured light calibration method for large surface topography

Sotoca JM*, Buendía M*, Iñesta JM**.
*Unidad de Biofísica. Dpt. Fisiología. Universidad de Valencia.
Av. Blasco Ibañez, 15, E-46010 Valencia, Spain.
**Dpt. Informática.Universitat Jaume I. Campus Riu Sec, E-12071 Castellón, Spain
e-mail: jomarso@alumni.uv.es, mateo.buendia@uv.es, inesta@inf.uji.es

Abstract. A new strategy for object surface recovery using a structured light technique is presented. The calibration method is based on the projection of a net on two different planes: a base position and a position displaced forwards to a distance enough for the problem objects to be contained between them. These images are used for the system to be calibrated. This is suitable to be applied to large objects and needs not to know the geometry of the whole system, but only the relative distances of the calibration planes. It will be validated with the shape of human backs, with the aim to evaluate its usefulness as a diagnosis technique for back deformation pathologies.

1 INTRODUCTION

Using structured light, the deformation of a light pattern projected on an object surface provides data about its topography [13]. This way, a method of range retrieval is available as an alternative to stereo imaging. The 3D position of a point is univocally determined by the intersections of straight lines with a plane. Using trigonometry it is possible to establish relations between the positions where the light beams are reflected on the object and the positions of the light spots in the camera image plane [1].

Structured light is a particular case of stereo vision, in which one of the cameras is replaced by a light source that projects a known light pattern on the scene. Thus, the landmarks are artificially created by the illumination pattern, from known positions in the image plane of the light source. This technique simplifies the correspondence problem of stereo-vision changing the problem of seeking similar points in both images by that of solving the correspondences between the projected pattern and the imaged one. On the other hand, this scheme is only valid in controlled environments, because its nature makes it unsuitable for outdoor contexts, natural illumination or any changing domain where either the world reference system or the camera-projector relative positions could change without control.

A number of techniques have been used for analysing the imaged pattern, trying to attain more accuracy [4], related to the orientation of the surface and relative depth [12], or dealing with the illumination and light reflection problem [7]. Nevertheless, the problems to which more attention are paid in the structured light literature are calibration and the identification of the fringes in the image plane. This ambiguity in fringe identification (the actual correspondence problem in structured light) is named *the indexation problem* by a number of authors [2,11], that could be enunciated as the problem of mapping the imaged fringes to their positions in the projected grid. If the projected pattern is not fringes but a net, the indexation is applied to the nodes of the net, looking for their mapping into integer indices in the original grid.

Different approaches exist to solve the indexation problem through some kind of coded structured light, related to any method to pattern local characterisation, like colour fringe coding [2] with several proposed operations implemented in real-time by hardware, or using a binary coding of the light pattern fringes [11].

For the calibration problem, the focal distances, and the determination of the translation and orientation of the co-ordinate imaging system with respect to the global co-ordinate system is usually the problem to be solved in the literature, and different approaches are proposed to it [1,3].

Our system proposes a different approach to the indexation problem for a square grid, based on a previous calibration method that imposes 3D constraints to the projected grid solution. The method is based on geometric considerations over the information provided by the projected pattern and two calibration projections, avoiding colour coding and complex patterns that make the system more expensive and compromise its efficiency.

The performance of the proposed method will be tested in a real problem: the reconstruction of the human back surface with the aim to obtain some results that confirm the applicability of the system to retrieve 3D information from large objects. This fact allows taking this methodology out of its usual domain (robotics and artificial vision laboratories) and its most well known applications (object location and characterisation, and range retrieval), to extend it to the medical imaging and assisted diagnosis fields. It could be used for preventive study of backbone pathologies related to raquidic deformations, particularly scoliosis.

2 METHOD DESCRIPTION

Using a directional illumination device, a square grid is projected over the studied zone, taking the image of the projected net with a digital camera. During the calibration phase, this process is repeated twice, as detailed below, without changing the camera-projector setting: firstly, the grid is projected on a screen placed at the front of the zone where the problem objects will be positioned. The grid projected on that plane will be named “front net”. Secondly, the screen is moved backwards, to a distance enough to make room for the problem objects to be placed between the two parallel planes. The grid is projected again and this new projection will be named “back net”. With these two images the system is calibrated as described below in such a way that it can predict the light beam trajectories coming from the projector and traversing the “calibrated space region”. The experimental setting is presented in Fig. 1, where the positions of the screen during the calibration phase are shown. Also an object is displayed between both planes illuminated by the pattern in order to get the distorted “object net”. During operation, the back plane must remain in its calibration position.

Once the system is calibrated, if we establish position relations between the points in both reference planes and their correspondence in the camera image plane using triangulation methods, the 3D co-ordinates of the landmarks of the projected pattern can be derived from the analysis of the grid projection. These co-ordinates permit the reconstruction of the surface topography of any object placed between both planes, with *z* values measured as distances to the back plane.

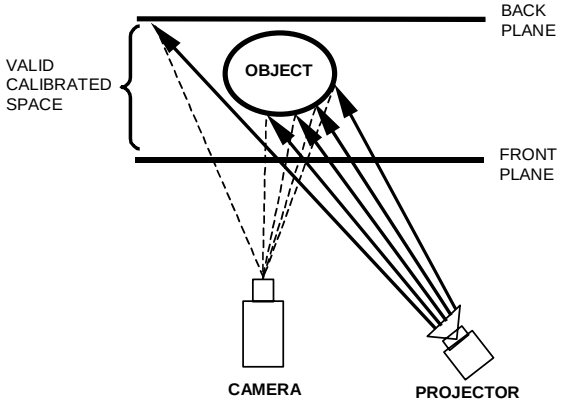

Fig 1. Experimental setting. The camera and projector situation are presented; and also the back and front planes.

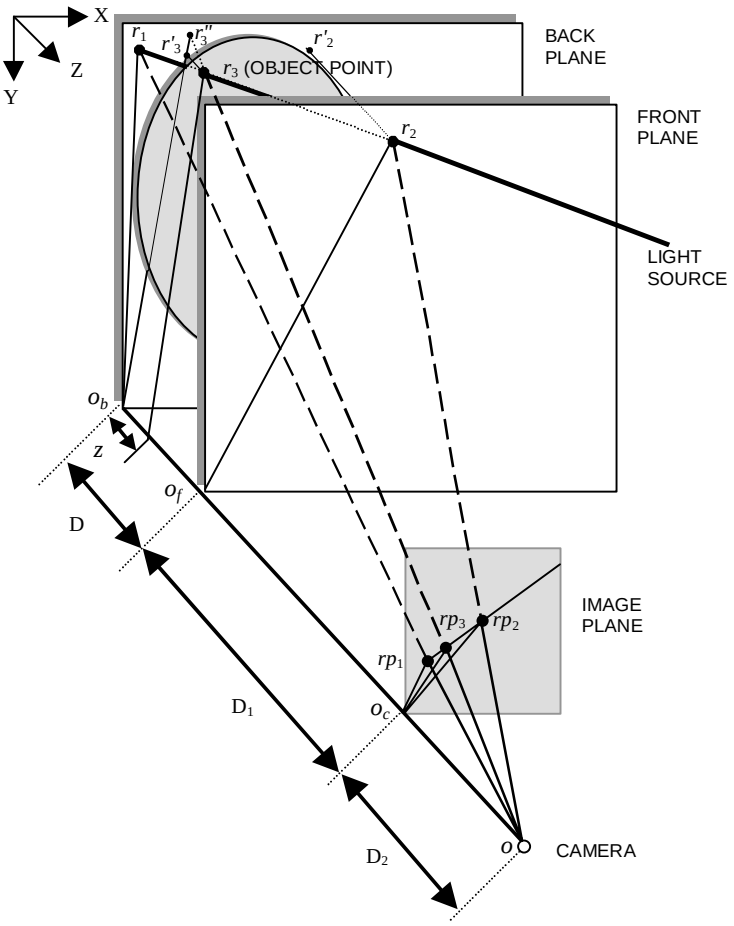

Fig. 2. Representation of the intersection points between a light beam and the front plane, the object surface, and the back plane. In addition, the positions where those three points are imaged are also displayed.

As indicated, the back and front nets establish the light propagation lines for the pattern grid nodes. The object net contains in its nodes the information related to the distance between the back plane and the object net nodes. To obtain these distances the following scheme must be followed: Let a light beam projected from the light source (as displayed in Fig. 2). This line intersects with the front plane, object surface, and back plane at r_2, r_3, and r_1, respectively, and the camera *sees* those points aligned in the image plane (at rp_2, rp_3, and rp_1). Let r'_2 and r'_3 the projections of r_2 and r_3 on the back plane, and r''_3 the intersection of the line Or_3 in that plane, and O_b the co-ordinate origin for the back plane, O_f the one for the front plane, O_c for the image plane, and O the focal point (see Fig. 2).

Using triangle similarities, it is possible to establish the following relations: for the triangles OO_crp_3 and $OO_br''_3$:

$$r''_3 = \frac{rp_3\left(D + D_1 + D_2\right)}{D_2} \tag{1}$$

and for the triangles OO_crp_1 and OO_br_1:

$$r_1 = \frac{rp_1\left(D + D_1 + D_2\right)}{D_2} \tag{2}$$

for the triangles OO_brp_2 and OO_fr_2:

$$r_2 = \frac{rp_2\left(D_1 + D_2\right)}{D_2} \tag{3}$$

In addition, the triangles $r_2r'_2r_1$ and $r_3\,r'_3r_1$ are also similar, so:

$$r'_3 = \frac{z\left(r'_2 - r_1\right)}{D} + r_1 \tag{4}$$

And, finally, the triangle $r_3r'_3r''_3$ is similar to OO_crp_3:

$$r'_3 = r''_3 - \frac{rp_3 z}{D_2} \tag{5}$$

Making (4)=(5) and solving for z:

$$z = \frac{(r'_3 - r_1)\, D_2 D}{(r'_2 - r_1)D_2 + rp_3 D} \tag{6}$$

If the r's are replaced by their values according to Eqs. (1), (2), and (3), and taking into account that r_2 and r'_2 can be interchanged, then

$$z = \frac{(rp_3 - rp_1)(D + D_1 + D_2)D}{(rp_2 - rp_1)(D_1 + D_2) + (rp_3 - rp_1)D} \tag{7}$$

Dividing the numerator and denominator of (7) by (D_1+D_2) and letting $k = (D+D_1+D_2)/(D_1+D_2)$, the value of z for the object point (r_3) is expressed as a function of the positions in the image plane and distances between the reference planes (all known):

$$z = \frac{D \cdot k \cdot (rp_3 - rp_1)}{(k-1)\cdot(rp_2 - rp_1) + (rp_3 - rp_1)} \tag{8}$$

This way, the z co-ordinate of each point can be calculated without the necessity of knowing the whole geometry of the setting, only from the positions of the nodes in pixels. Only k needs to be calculated for the system to be calibrated.

In certain conditions, the distance between the camera focal point and the image plane, D_2, is not needed to be known. Let's consider two beams A, B traversing the calibrated zone. The following relation for the distances between the beam intersections in the reference planes and their images in the image plane is obtained, using Eqs. (1) and (2):

$$\frac{rp_1^A - rp_1^B}{rp_2^A - rp_2^B} = \frac{r_1^A - r_1^B}{r_2^A - r_2^B} \cdot \frac{D_1 + D_2}{D + D_1 + D_2} = \frac{1}{k} \cdot \frac{r_1^A - r_1^B}{r_2^A - r_2^B} \tag{9}$$

where $(r^A{}_1 - r^B{}_1)$ and $(r^A{}_2 - r^B{}_2)$ are the distances in the back and front planes between the light dots produced by both beams. If the light source is far enough, the beams can be considered parallel, and those distances will be the same, so k will be a function of the distances in the camera image plane:

$$k = \frac{rp_2^A - rp_2^B}{rp_1^A - rp_1^B} \tag{10}$$

This way, once the images of the front and back planes have been digitized, the co-ordinates of every node are calculated, and the correspondence among the nodes in both nets establish the path followed by the beams between both planes, and the system is calibrated. Then, given an image of the same net projected on an object located in the calibrated space, the co-ordinate z relative to the back plane for each node is calculated using the position of the node in the image and the calibration information through Eq. (8).

3 METHOD UNCERTAINTY ESTIMATION

Some efforts have been carried out [10,14] trying to estimate the inaccuracy in range measurement using structured light. Two main error sources are reported: due to the setting and due to discretization. For the latter, [14] the use of subpixel operators has been proposed to improve the precision. Here, we will focus on the first problem.

The inaccuracy in measuring z can be estimated from Eq. (8), making $d_p = (rp_3 - rp_1)$ and $d_0 = (rp_2 - rp_1)$ for simplicity:

$$z = \frac{D \cdot k \cdot d_p}{(k-1)\cdot d_p + d_0} \tag{11}$$

Thus, $z = f(D,k,d_p,d_0)$, and to investigate how the uncertainties of each variable, v_i, influences the relative error of z, $\varepsilon_r(z)$ the following expression will be applied to Eq. (11):

$$\varepsilon_r(z) = \frac{\varepsilon(z)}{z} = \frac{1}{z}\sum_i \left|\frac{\partial z}{\partial v_i}\right| \varepsilon_r(v_i) \tag{12}$$

Considering the same errors when measuring distances in the image plane, $\varepsilon_r(d_r) = \varepsilon_r(d_0)$,

$$\varepsilon_r(z) = \varepsilon_r(D) + \left|\frac{(d_0 - d_p)}{d_0 + (k-1)d_p}\right| \varepsilon_r(k) + \left(\frac{d_0}{d_p} + 1\right) \left|\frac{1}{d_0 + (k-1)d_p}\right| \varepsilon(d_0) \tag{13}$$

Since $D << D_1 + D_2$ always, then $k \,/\, 1$ and this equation can be simplified obtaining this approximation:

$$\varepsilon(z) = \varepsilon_r(D) + \frac{d_0 - d_p}{d_0 + d_p}\varepsilon_r(k) + \varepsilon_r(d_0) \tag{14}$$

Let's now consider each of the terms of this expression by separate, since each one is related to a different item of the method. $\varepsilon_r(D)$ is related to the calibration of the system, $\varepsilon_r(k)$ depends on the positions of some parts of the setting, and $\varepsilon_r(d_0)$ depends on the method design. This way, four error types can be considered: calibration errors, errors due to the experimental setting, those from the image processing techniques, and other possible not considered errors.

Regarding the first, with the typical values of the setting (with a screen of $1m^2$ and $D = 500 \pm 1$ mm), $\varepsilon_r(D) = 0.2\%$. The setting errors are related to the term $\varepsilon_r(k)$ of Eq. (14). For the setting used for analyzing large objects, the ratio k is determined with an error of 2%. The errors associated to the image processing techniques are reflected in the term $\varepsilon_r(d_0)$ of Eq. (14). In this regard, the key point is a precise location of the nodes. The procedure for fringe location and skeletonization (see below in ∋4.1 and ∋4.2) has proved to be very accurate, and we can consider that, in the worst case, the positioning errors of lines and nodes could be of ±1 pixel. For a spatial resolution of 256×256 pixels, this error contributes in average with $1/128 \cong 0.8\%$.

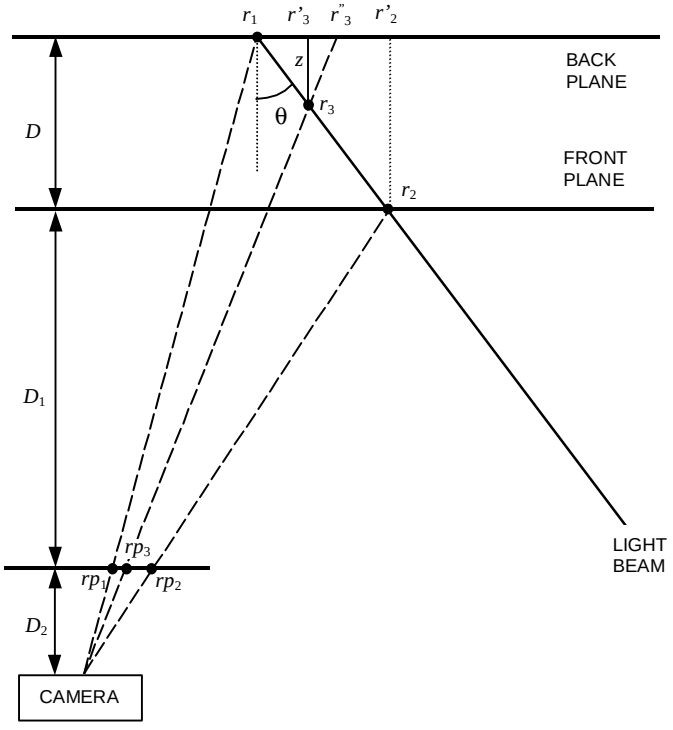

Fig. 3. (an upper view of Fig.2) Relations between the projection angle, θ, and the distances to the image plane.

Applying these considerations, and adding their contributions, an $\varepsilon_r(z) \approx 3\%$ is estimated. In addition, other possible sources of error may be the lens distortion and others unknown. For optical distortion in the projector, a correction has been made on the lines in the reference images. Those lines are parameterized, in such a way that the net nodes are not the light dots in the image but the mathematical intersections of parametric straight lines. Also, the errors due to the experimental setting may vary with the projection angle, θ. The points r_2, r_3, and r_1 are related to the points in the image plane where they are imaged rp_2, rp_3, and rp_1. This relationship has not the same associated error for all the angles.

From Fig. 3, an expression relating z to the angle θ can be obtained. Thus, the following relations can be established:

$$r_3'' - r_1 = \frac{z \cdot rp_3}{D_2} + z \tan\theta \tag{15}$$

Making Eq. (1) – Eq. (2) and replacing the result in (15), it turns that:

$$z = \frac{d_p \left(D + D_1 + D_2\right)}{rp_3 + D_2 \tan\theta} \tag{16}$$

Applying again the Eq. (12), the relative error $\varepsilon_r(z)$ and its angular dependence can be observed:

$$\varepsilon_r(z) = \frac{D_2 \varepsilon(\theta)}{D_2 \cos\theta \sin\theta + rp_3 \cos^2\theta} + \frac{\varepsilon\left(rp_3\right)}{D_2 \tan\theta + rp_3} \tag{17}$$

Typically, $\varepsilon(\theta) = 1^o$ and $\varepsilon(rp_3) = 1$ pixel. In addition, the dominant factor of the error is $1/\cos^2\theta$ for high values of $\theta = 0^o$ and the other influent term, $1/\tan\theta$, only contributes if the angles are very small. It has been proven experimentally that the error increases dramatically for $\theta > 45^o$ and if θ is small, local fluctuations appear.

An experiment was designed to study these variations and quantify the hypothetical unknown error. Some known shapes were digitized from different projection angles, and the retrieved dimensions were compared to the original ones to calculate the relative errors in the measurement. The data show an average error of $\varepsilon_r(z) \cong 3.5\%$ for a zone of best behavior of $\theta \in [25^o, 35^o]$.

4 APPLICATION ISSUES

The structured light techniques have been commonly used in robotics and artificial vision laboratories for locating and characterising three-dimensional objects. One of the advantages of the proposed method is that the calibration of a large zone of the space permits its application to the characterisation of large objects. Our goal is to show its suitability not only for small objects (typical of the formerly mentioned applications) but also for large ones, like the human body, with medical applications.

The key problem for calculating the z values of the nodes is to establish the correspondences among the nodes in the three nets. For this, a mark is introduced in the pattern, consisting of a white square that must be visible on the problem surface, which indicates the position of a labelled node, from which the reconstruction of the net is made.

There are a number of factors to keep in mind to design the kind of net to be projected: the line spacing and their thickness. The first one depends on the typical size of the studied objects, and the thickness of the lines depends on

Fig. 4. Projection of the net on an object with steep slopes and with uneven textures in those zones.

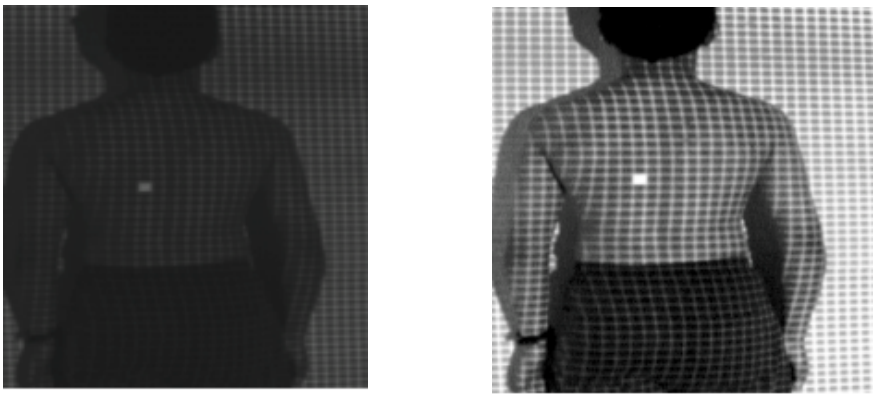

Fig. 5.a (left) Image of the back with the projected net as it is digitized; 5.b (right) The same image but equalized in order to make the pattern lines evident.

the typical smoothness of the object surface [6]. This factor and the inclination of the shape condition the continuity of the lines and the possible overlapping of the nodes, as can be observed in Fig. 4. The determination of all these parameters can be found elsewhere [8].

4.1 Image Processing and Pattern Reconstruction

When projecting a pattern onto large objects like the human body, the beam is dispersed, so the image may appear obscure if the light source has not enough power (Fig.5.a). To overcome this difficulty, the original image is equalised (Fig. 5.b). The mark detection and location is made in this enhanced image, and then it is deleted.

The next step is the object segmentation. Since the back net and the object net are only different in the zone occupied by the object, a logic difference between both images is useful in order to segment the object and thus restrict the image analysis just to the object zone and shadows. But in these latter zones no fringes are found, so we can focus on the object net using this difference image.

The imaged net provides spike-shaped grey profiles, where each peak corresponds to a line of the projected pattern. Once segmented the object, a maxima detection in the profile lines inside the segmented object zone is performed. For that maxima detection, a scanning for vertical and horizontal lines is executed. Being the light source and the camera at similar heights, the horizontal lines are always close to straight lines, making their detection and tracking very easy. The vertical ones are usually curved. Trying to avoid irregularities in the maxima detection the scanning is made by pixel fringes. Considering a width of 7 pixels has proven to be an efficient decision in all the cases tested. Even if vertical lines appear very curved, the method is able to detect the maxima and therefore segment the pattern lines.

Nevertheless, still some discontinuities may appear in this line location process that could condition the performance of the method. To solve this, a semiautomatic mechanism has been devised for reconstruction of discontinuities (if they appear) that is able to work correctly either for smooth surfaces or uneven ones, like the one in Fig. 4. This algorithm provides segment connection to join the discontinuous lines. It is based on the location of both segment extremes and then, from the end of one segment to seek the beginning of another that might be the prolongation of the former with a high likelihood.

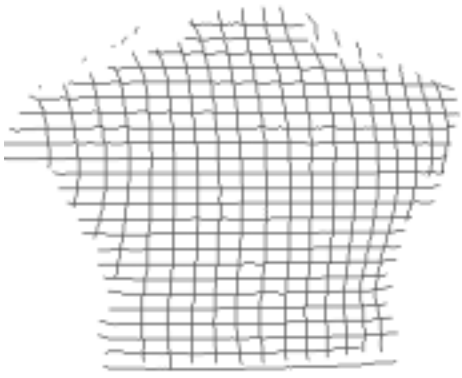

Fig. 6. Image obtained after line segmentation and pattern reconstruction.

Once found, a line is drawn from both connected extremes. This way, two binary images are constructed: one contains the lines with an average slope higher than one (vertical) and the other containing the lines with a slope lower than one (horizontal). The logic sum of both images is the image where the lines appear skeletonised and without discontinuities, i.e. the reconstruction of the distorted projected pattern, like in Fig. 6.

4.2 Detection of the Object Net Nodes and Landmark Mapping

The node-seeking algorithm performs a line tracking in the four cardinal directions, looking for line crossing points. These points must hold a crossing condition for at least three lines. Then, since more than one possible crossing point may locally appear, the final crossing point will be the pixel closest to the geometric centroid of them (Fig. 7). This part of the process is very important since the existence of false or missing nodes affects the method's reliability, and their correspondences in the three nets.

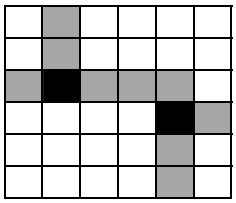

Fig. 7. The "double node" problem.

Once this process has been carried out, we have the positions of all the nodes for the object and reference nets, and we establish that the first node will be the one at the upper left corner of the mark in the images.

For the back and front nets, the lines are stored in an array as slopes and zero crossings, setting the co-ordinate origin at the first node. For the object net, the pattern is traversed using a neighbour protocol, starting from the first node, storing for each one its co-ordinates, the identification number according to a predetermined line ordering criterion, and the identification numbers of its neighbour nodes.

Once the correspondence among the nodes in the three nets has been established, the Eq. (8) is applied to each of the triads of nodes for calculating the *z* values of the correspondent node in the object net. For this, the node list from the object net must be traversed and the co-ordinates of the intersections for the correspondent lines in the back and front nets must also be located.

5 RESULTS

We present here the performance of the proposed method and its ability to retrieve the surface of large objects, like the tested application of the human back.

With the *z* values of the nodes, we get a mesh of landmarks that sketches the object surface (see Fig. 8.) That information can be enough for many applications, but for the medical application of this technique, accuracy is an important issue, so in order to retrieve

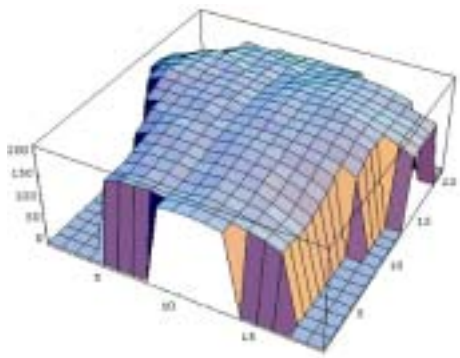

Fig. 8. 3D projection of the *z* values of the landmarks for the back of Fig. 5. As observed, the fact of assuming a suitable grid is very important to get a suitable model of the surface.

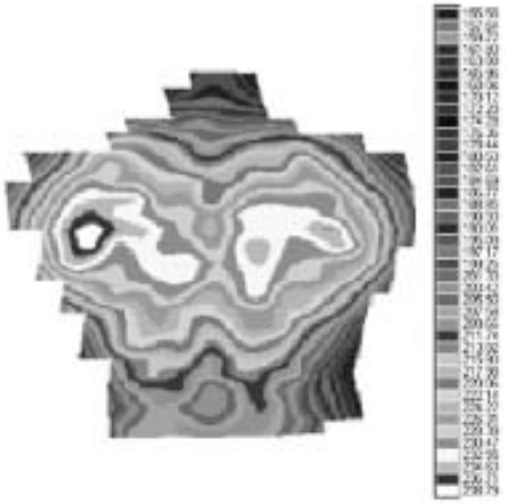

Fig. 9. Topographic representation of the surface of the back, showing the values of *z* calibrated in millimeters. At the right side, the scale is presented.

more information, a two-dimensional interpolation over those landmarks is applied to calculate the *z* values for the rest of points in the problem surface.

This method consists of a segment bicubic interpolation. This has the advantage of assuring the continuity of the first and second derivatives, and to satisfy a continuity condition in the extremes of each interval (at the pattern nodes). We impose that the second derivatives $x''(t)$ are zero for all those points. The interpolating functions will be polynomials of the kind $x(t) = a_1t^3 + a_2t^2 + a_3t + a_4$, with $t \in [0,1]$, where the surface parameterisation will be provided by the tensorial product of two polynomials of this kind.

The final result obtained is a surface topography (Fig. 9) that permits us to recover the shape of the imaged object (a human back in Fig. 9). As stated in ∋3, the inaccuracy of the method can be estimated in less than 4%.

6 CONCLUSIONS AND FUTURE WORK

A new method has been proposed to make it easier the calibration process in a structured light environment. Considering only the information from two calibration images: one for the illumination pattern projected onto a back plane and one projected onto a front plane, it permits to know the direction of the light beams produced by the pattern. This fact leads to set the space between both planes ready to retrieve, by means of triangulation, the range of the nodes of the imaged pattern when it is projected onto the problem surface placed in that space, measured with respect to the back plane.

The indexation problem is solved introducing a mark in the pattern that labels a given node of the square net. This mark must be visible on the problem surface. This approach restricts the applicability of the method to scenes with a single object whose position is a priori known or, at least, expected. If the scene presents multiple objects or highly changing positions, the mark should be replaced by any kind of encoded pattern (colour, binary-coded, etc.) without changing the calibration strategy.

The method is designed to permit the topography of large surfaces, like the human back. In this case, the absolute error for the measurements is 3 mm, which is in accordance with the experts that recommend 2 or 3 mm as upper bound to that inaccuracy for anatomic measurements of diagnostic validity.

From the application of the method on the human back, interesting studies can be derived related to pathological conditions of the patient. For this, high level reasoning must be carried out from the range level map. The first step to do this is to characterise the spine line in the range data. We are currently working on it through the use of 3D active shape models [5] to get a model of the back for measuring lateral raquidic deviations, curvatures and torsion. These values permit to obtain patterns for scoliosis diagnosis, where vertebrae lateral deviation and rotation occurs [9]. In addition, the study of the curvature of the spine leads to the measurement of other back abnormalities like kyphosis and lordosis.

Other research directions point to the analysis of local and global mirror symmetries, for characterisation and quantification of symmetric deficiencies [15] in 3D surfaces.

Acknowledgements

This work has been funded by the Fundación Bancaja Research Project no. P1A96-18.

References

[1] J. Batlle et al., Recent progress in coded structured light as a technique to solve the correspondence problem. Pattern Recognition 31 (1998) 963-982.

[2] K.L. Boyer and A.C. Kak, Color-Encoded Structured Light for Rapid Active Ranging. IEEE Trans. Patt. Anal. Mach. Intell. 9 (1987) 14-28.

[3] O.D. Faugeras and G. Toscani, The calibration problem for stereo, Proc. of Computer Vision and Pattern Recognition (1986) 15-20.

[4] K. G. Harding and K. Goodson, Hibrid, high accuracy structured light profiled. Prod. Soc. Photo-Opt. Instrum. Engineers. Vol. 728 (1986) 132-145.

[5] A. Hill et al., Model-Based Interpretation of 3D Medical Images. Proc. of 4th British machine Vision Conference, Guilford, England, 1993, pp. 339-348.

[6] J. Le Moigne and A.M. Waxman, Structured Light Patterns for Robot Mobility. IEEE Journal of Robotics and Automation. 4 (1988) 541-548.

[7] A.C. Sanderson et al., Structured Highlight Inspection of Specular Surfaces. IEEE Trans. Patt. Anal. Mach. Intell. 10 (1988) 44-55.

[8] J.M. Sotoca, Un nuevo método de calibración en entornos de luz estructurada para la obtención de topografía de superficies. MSc Thesis, Dept. Informática, Univ. Valencia, Spain (in spanish), 1999.

[9] A.R. Turner-Smith and J.R. Jefferson, Analysis and Presentation of Human Back Shape in Scoliosis. In D.C. Handscomb (ed.), The mathematics of surfaces III. Clarendon Press. Oxford, 1989.

[10] R.J. Valkenburg and A.M. McIvor, Accurate 3D measurement using a structured light system. Image and Vision Computing 16 (1998) 99-110.

[11] P. Vuylsteke and A. Oosterlinck, Range Image Acquisition with a Single Binary-Encoded Light Pattern. IEEE Trans. Patt. Anal. Mach. Intell. 12 (1990) 148-163.

[12] Y.F. Wang et al., Computation of Surface Orientation and Structure of Objects Using Grid Coding. IEEE Trans. Patt. Anal. Mach. Intell. 9 (1987) 129-136.

[13] P.M. Will and K.S. Pennington, Grid coding: a preprocessing technique for robot and machine vision. Proc. of Int. Joint. Conf. on Artif. Intell., 1971, pp.66-70.

[14] Z. Yang and Y.F. Wang, Error analysis of 3D shape construction from structure lighting. Pattern Recognition 29 (1996) 189-206.

[15] H. Zabrobsky et al., Symmetry as a Continuous Feature. IEEE Trans. Patt. Anal. Mach. Intell. 12 (1995) 1154-1165.