Optimised De Bruijn patterns for one-shot shape acquisition

Jacques Pagès, J. Salvi, J. Forest

▶ **To cite this version:**

Jacques Pagès, J. Salvi, J. Forest. Optimised De Bruijn patterns for one-shot shape acquisition. Image and Vision Computing, 2005, 23 (7), pp.707-720. hal-02596063

HAL Id: hal-02596063
https://hal.science/hal-02596063

Submitted on 30 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimised De Bruijn patterns for one-shot shape acquisition

Jordi Pagès, Joaquim Salvi and Josep Forest *

Computer Vision and Robotics Group, Institute of Informatics and Applications, University of Girona. Avda. Lluís Santaló s/n, E-17071 Girona, Spain

Abstract

Coded structured light is an optical technique based on active stereovision which allows shape acquisition. By projecting a suitable set of light patterns onto the surface of an object and capturing images with a camera, a large number of correspondences can be found and 3D points can be reconstructed by means of triangulation. One-shot techniques are based on projecting an unique pattern so that moving objects can be measured. A major group of techniques in this field define coloured multi-slit or stripe patterns in order to obtain dense reconstructions. The former type of patterns is suitable for locating intensity peaks in the image while the latter is aimed to locate edges. In this paper we present a new way to design coloured stripe patterns so that both intensity peaks and edges can be located without loss of accuracy and reducing the number of hue levels included in the pattern. The results obtained by the new pattern are quantitatively and qualitatively compared to similar techniques. These results also contribute to a comparison between the peak-based and edge-based reconstruction strategies.

Preprint submitted to Elsevier Science *25 April 2005*

Key words: Shape acquisition, Active Stereo, Coded structured light

1 Introduction

The shape acquisition problem has long been a challenge for the computer vision community [1,2]. The effort put into finding solutions has led to a large taxonomy of non-contact optical techniques [3]. In this field, two major groups of techniques can be distinguished: passive and active. Passive techniques are based on obtaining the shape of the object without introducing energy into the scene. Some of the cues used are shape from motion, texture and shadows which are extracted by using a single image or a set of images captured by a single camera or multiple cameras. One of the most well-known passive techniques is stereovision [4]. The main difficulty of stereovision consists of solving the correspondence problem between the different views of the same object. Even if the correspondence problem can be made easier thanks to geometrical constraints [5], its solution is still difficult and computationally expensive. Furthermore, stereovision systems cannot ensure a maximum resolution; in fact, the number of correspondences and the number of 3D points is quite reduced because it is highly dependent on the observed object. Moreover, when the measured object does not contain singular points, such as corners, or it is not highly textured, correspondences cannot be found. Despite these limitations, stereovision techniques still have an outstanding potential thanks to the huge amount of information that can be obtained by using cameras. Investigations into how to get the most from this potential led to a powerful solution: active stereovision. In general terms, active optical techniques project light onto the object and then the reflected light is measured; in this way, information on shape and range can be calculated. In the specific case of active stereovision, controlled illumination is projected onto the object and then a camera, or a set of cameras, is used to capture the images. The controlled illumination is used to project singular points onto the object surface.

* Corresponding author: Joaquim Salvi; tel.:+34 972 41 8483; fax: +34 972 41 8098.
Email address: {jpages,qsalvi,forest}@eia.udg.es (Jordi Pagès, Joaquim Salvi and Josep Forest).

Therefore, the correspondence problem can be solved easily.

In the development of active stereovision techniques, structured light was the first to appear. Structured light techniques basically involve projecting laser planes or laser dots and scanning the object surface. An image must be captured for every position of the scanning laser. For each image, all the pixels which are illuminated by the laser can be triangulated and reconstructed. The main advantages of structured light techniques include the easy image processing involved and the high accuracy that can be achieved in the 3D reconstruction. In some of these techniques, however, the mechanical operations required or the number of images that must be taken can become a drawback.

Coded structured light can be considered as a logical evolution of classical structured light techniques. The aim of this group of techniques is to increase the resolution achieved in every image. Such techniques are based on projecting bidimensional light patterns by using devices such as slide projectors and, more recently, LCD (liquid-crystal device) and DMD (digital mirror device) projectors. The systems based on these devices are more flexible since the projected patterns can be easily and inexpensively changed. The patterns are designed so that a set of points are encoded, i.e. its pattern position can be retrieved from the camera image(s) by gathering local information around this encoded point. A key element of coded structured light techniques is the coding strategy used to generate the light patterns. A coding strategy is needed to avoid ambiguities in the projected patterns, i.e. different pattern regions being equal and therefore, indistinguishable. The coding strategy used strongly determines the performance of the technique in terms of resolution, accuracy and the number of required patterns.

Two main groups of techniques based on coded structured exist [6]. The first group, called time-multiplexing, is based on projecting a sequence of patterns so that colour is not required but only static objects can be measured. The second group is based on projecting a unique pattern and it is referred as one-shot techniques. In this paper we focus on one-shot techniques based on a coloured pattern. Concretely, the paper presents a new way of encoding coloured one-shot patterns. Even if this

type of patterns are usually limited to colour neutral or pale objects, they can obtain very good results in terms of resolution and accuracy. The aim of the new coding strategy is to define a hybrid pattern so that it improves the performance of similar existing patterns in terms of resolution and accuracy. The structure of the paper is as follows. First, in Section 2, a brief overview of one-shot techniques based on coded structured light is presented, focusing on techniques using colour. Secondly, in Section 3, we describe two of the most usual coloured patterns in one-shot techniques and we present the new pattern which combines the advantages of both of them. Afterwards, details about the segmentation and decoding processes for the new pattern are exposed in Section 4. The system's calibration is briefly explained in Section 5. Then, the hybrid pattern is compared to similar patterns through experimental results from a quantitative and a qualitative point of view in Section 6. Finally, the conclusions of the paper are discussed in Section 7.

2 Overview of one-shot techniques

According to an early study of the state of the art on coded structured light [6], it can be stated that most part of one-shot techniques are based on spatial coding. In time-multiplexing techniques the labels of the encoded points are multiplexed along a sequence of patterns. In spatial coding (one-shot techniques) the label is encoded in a local neighborhood. Therefore, while time-multiplexing techniques assume that the object remains static, spatial techniques generally assume that the object surface is locally smooth.

Different coding strategies can be found in the bibliography leading to patterns where a set of columns, rows or both can be uniquely identified by decoding local spatial data. A large group of techniques exploit the *pseudo-random arrays theory* [7] in order to define patterns having the *window property* [8–11]. For example, Morano et al. proposed a pattern containing an array of coloured dots so that every window of $m \times n$ dots appears only once in the whole array. In this case, encoded points are provided since both pattern axis are encoded. Similarly, another major group

of techniques rely on coding strategies based on *pseudo-random sequences* [12] in order to colour multi-slit or stripe patterns. In the former, equidistant bands of pixels separated by black gaps are coloured so that consecutive slits can have the same colour. In the latter, the bands are adjacent so that consecutive stripes cannot share the same colour. By colouring such patterns with pseudo-random sequences each subset of n adjacent slits or stripes identified by its colour appears only once in each pattern line orthogonal to the stripes direction [13–15]. Point encoding can be fulfilled by using grid patterns where both sets of horizontal and vertical slits are encoded [16, 17]. These two groups of techniques are able to obtain large resolution with a unique image and high accuracy in the $3D$ measurements. However, the use of several colours makes them typically sensitive to the object albedo or texture so that they work properly only with colour neutral or pale objects.

Some efforts have been done in order to define patterns avoiding to exploit colour. For example, a binary pattern consisting of vertical slits containing random cuts which uniquely identifies each slit was proposed by Maruyama and Abe [18]. Even if the binary nature of the pattern makes it robust against the object colour, the segmentation of the segments can be difficult under certain textures and the decoding stage is quite complex and sensitive to surface discontinuities. In the work by Guan et al. [19] a method to combine multiple patterns into a single grey-level pattern is presented. The technique relies on the communications theory which is used to modulate several sinusoidal patterns (which are equidistantly phase-shifted) into a unique pattern. The perceived image can be decomposed by using demodulation and the phase of the sinusoidal signal can be recovered. The technique is pretty robust against the object appearance thanks that the albedo is cancelled during the phase calculation. However, the authors state that some problems are observed near edges or regions with sharp variations on albedo or depth. A real-time application for global scene reconstruction based on video sequences was implemented with this pattern. The resolution obtained was enough for allowing human-computer interaction. However, the resolution obtained by the modulated pattern could be insufficient in order to obtain dense and accurate reconstructions of objects with high level of detail. Another recent technique avoiding the use of multiple colours

was presented by Koninckx et al. [20]. In this case, a stripe pattern of alternating black and white stripes is projected. In order to distinguish the stripes several parallel coding-lines intersecting the stripes with a certain angle are also included in the pattern. All the coding-lines have the same colour, which is green by default even if it is on-line adapted taking into account the object predominant colour. A drawback of this technique is that no dense codification occurs, since only the intersection points between stripes and coding-lines are actually encoded. These intersection points can be decoded by using the epipolar geometry of the camera and the projector. Then, for every image scan-line along the orthogonal direction of the stripes a unique stripe is encoded while the rest must be relative labeled to this one. The authors proposed to solve the relative labeling of all the stripes by using cues from the graph theory [21].

In this paper we focus on classic one-shot techniques based on coloured stripe patterns. These techniques, even if they are typically limited to colour neutral or pale objects, are very powerful in terms of accuracy and resolution. Our goal is to propose a new coding strategy for improving the resolution of stripe patterns without loss of accuracy in the $3D$ measurements. Next section presents the new pattern and how it increases the resolution of similar existing patterns.

3 A proposal of a new hybrid pattern

In this section we present a brief review of two classical one-shot techniques and afterwards a new hybrid pattern combining the advantages of both of them.

As already mentioned in the previous section, there is an important group of one-shot techniques which are based on coloured multi-slit and stripe patterns. Multi-slit patterns introduce black gaps between the coloured bands of pixels so that two consecutive slits can have the same colour. The black gaps also allow intensity peaks to be detected in the images. Every intensity peak corresponds to the central position of a certain slit. Thus, the reconstruction step consists of triangulating the central point of the projected slits with the corresponding intensity peaks detected

in the image. We refer to this reconstruction strategy as *peak-based reconstruction*.

On the other hand, in stripe patterns no black gaps are introduced between the coloured bands of pixels. Therefore, adjacent stripes cannot share the same colour (hue level). In such patterns, edges between stripes are searched in the image. Every edge located in the image can be triangulated with the corresponding edge from the projected pattern so that an *edge-based reconstruction* is performed.

Note that in multi-slit patterns an edge-based reconstruction is not suitable since a certain amount of the intensity of a slit is integrated over the surrounding black regions. Therefore, the edges in the image do not really correspond to the edges in the pattern definition. Such problem already appears when multiplexing binary codes by projecting a sequence of binary stripe patterns of successively increasing frequency [22]. As stated by Trobina [23], the best solution in these cases consists of projecting every one of the binary stripe pattern and its negative version. Then, by intersecting the intensity profiles of every pattern pair the edges can be detected with high accuracy. Unfortunately, this solution is not applicable to one-shot techniques. Note also that in the case of coloured stripe patterns a peak-based reconstruction is neither feasible since the pattern is projected with a flat intensity profile.

Let us now briefly present two representative patterns belonging to this group of one-shot techniques. These patterns will be used to evaluate the performance of the new pattern presented at the end of the section.

3.1 An example of stripe pattern

A recent powerful technique based on a coloured stripe pattern was presented by Zhang et al. [15]. The pattern is composed of 125 vertical stripes coloured using 8 different hue levels and is shown in Figure 1a. The edges are located in each scan-line by searching local maxima (detected with sub-pixel precision) of the following function

$$f = dR^2 + dG^2 + dB^2 \tag{1}$$

where dR, dG and dB are the intensity gradients of the red, green and blue channels on the current scan-line. Every edge is labeled as $(dR(e_i), dG(e_i), dB(e_i))$ where e_i is the position of the edge on the scan-line. The pattern is generated so that every combination of three consecutive edge labels appears only once as maximum (window property equal to 3). The correspondence problem between the sequence of edges located in a scan-line and the sequence of emitted edges is solved by using dynamic programming. The score function to compare a projected edge with a received edge is a distance of their labels (see [15] for details).

Figure 1

3.2 An example of multi-slit pattern

The pattern chosen for representing the peak-based strategy consists of 64 coloured slits, separated by black gaps of the same width. Four colours have been used to generate the pattern shown in Figure 1b. The sequence of coloured stripes has window property equal to 3. Similar patterns can be found in the bibliography like, for example, the one by Chen et al. [24] or the one by Monks et al. [25]. In order to locate the slits in a camera scan-line, we define the *M channel* as function of the RGB channels as follows

$$\begin{aligned} M(i) &= \max(R(i), G(i), B(i)) \\ i &= 1..\text{width(scan-line)} \end{aligned} \tag{2}$$

Maxima in the M channel correspond approximately to the bounds of the coloured slit. A coloured slit is said to be between two consecutive maxima if the gradient of M in this region is predominantly descending. The accurate position of the central point of a coloured slit is located by using a normalised centroid peak detector [26]. Then, since only 4 colours are projected, the colour of the detected slits can be easily segmented. The matching of the detected slits on the current scan-line and the sequence of projected ones is also made through dynamic programming based on the RGB components of the slits.

3.3 The new coding strategy

We now propose a new type of coding strategy which defines patterns containing both edges and intensity peaks. The new hybrid pattern appears as a stripe-pattern in the RGB space while in the intensity channel it appears as a peak-based pattern. The intensity profile of the pattern is defined as a square function so that it alternates from a medium intensity stripe to a high intensity stripe.

The coding strategy requires only $n = 4$ different hue values in order to colour a pattern with 128 stripes so that a window property of $m = 3$ stripes is obtained. The formal definition of the pattern can be noted as follows. Let W_m and W_h be respectively the width in pixels of a medium-intensity stripe and a high-intensity stripe. Let $S = \{1, .., 2n^m\}$ be the set of stripe indices. Then, the mapping from the pattern abscissas to the stripe indices is defined by the following function

$$stripe : X \longrightarrow S$$

$$stripe(x) = 2\,((x\text{–}1) \div (W_m\text{+}W_h))\text{+}\begin{cases} 1 & ((x\text{–}1) \bmod (W_m\text{+}W_h))\text{+}1 \leq W_m \\ 2 & \text{otherwise} \end{cases} \quad (3)$$

where $\div$ denotes integer division. The following function determines the intensity level of a given stripe, which has two possible values $L = \{high, medium\}$

$$Int : S \longrightarrow L$$

$$Int(x) = \begin{cases} high & x \bmod 2 \;=\; 0 \\ medium & \text{otherwise} \end{cases} \quad (4)$$

The pattern is divided in n periods $P = \{1, .., n\}$ of n^m consecutive stripes each one. In order to determine to which period belongs to a given stripe the following

function is defined

$$\begin{aligned} &period : S \longrightarrow P \\ &period(s) = \left((s-1) \div 2n^{m-1}\right) + 1 \end{aligned} \tag{5}$$

Let $H = \{1, .., n\}$ the set of indices to the hue values chosen. Then, the hue value of a given stripe is set as follows

$$\begin{aligned} &Colour : S \longrightarrow H \\ &Colour(s) = \begin{cases} period(s) & Int(s) = medium \\ db\left[\frac{((s-1) \mod 2n^2)+1}{2}\right] & \text{otherwise} \end{cases} \end{aligned} \tag{6}$$

where db is an ordered vector containing a De Bruijn sequence with window property of 2 like the following one [6]:

$$db = [1, 1, 2, 1, 3, 1, 4, 2, 2, 3, 2, 4, 3, 3, 4, 4] \tag{7}$$

Every stripe with a certain hue value and a certain level of intensity can be labeled according to the set of labels $B = \{1, 2, 3, 4, 5, 6, 7, 8\}$ by using the following function

$$\begin{aligned} &Label : H \times L \longrightarrow B \\ &Label(h, l) = \begin{cases} h & l = high \\ h + n & l = medium \end{cases} \end{aligned} \tag{8}$$

The sequence of stripes generated by the proposed coding strategy and the De Bruijn sequence in equation 7 can be represented using the labeling function in equation 8 as follows

$$\begin{aligned} P_1 &= [5, 1, 5, 1, 5, 2, 5, 1, 5, 3, 5, 1, 5, 4, 5, 2, 5, 2, 5, 3, 5, 2, 5, 4, 5, 3, 5, 3, 5, 4, 5, 4] \\ P_2 &= [6, 1, 6, 1, 6, 2, 6, 1, 6, 3, 6, 1, 6, 4, 6, 2, 6, 2, 6, 3, 6, 2, 6, 4, 6, 3, 6, 3, 6, 4, 6, 4] \\ P_3 &= [7, 1, 7, 1, 7, 2, 7, 1, 7, 3, 7, 1, 7, 4, 7, 2, 7, 2, 7, 3, 7, 2, 7, 4, 7, 3, 7, 3, 7, 4, 7, 4] \end{aligned}$$

$$P_4 = [8, 1, 8, 1, 8, 2, 8, 1, 8, 3, 8, 1, 8, 4, 8, 2, 8, 2, 8, 3, 8, 2, 8, 4, 8, 3, 8, 3, 8, 4, 8, 4] \tag{9}$$

where P_i is the sequence of stripes corresponding to period i. Note that the elements $\{1, 2, 3, 4\}$ correspond to high intensity stripes and 1 indicates the first hue value, 2 the second, etc. On the other hand, the elements $\{5, 6, 7, 8\}$ correspond to medium intensity stripes so that, in this case, 5 indicates the first hue value, 6 the second, etc. Note that in every period, all the medium intensity stripes share the same hue value, while the hue value of the high intensity stripes is directly chosen according to the De Bruijn sequence in equation 7.

Figure 1c shows a pattern encoded according to the above strategy. The $n = 4$ hue values used are equally spaced in the Hue space so that $H = \{0°,\ 90°,\ 180°,\ 270°\}$. Finally, the two intensity levels used are $L = \{0.35,\ 0.5\}$. The intensity profile of the pattern is shown in Figure 1d.

Note that 128 stripes have been coloured by using only 4 levels of hue. Indeed, as less colours are projected, higher immunity against noise is obtained. The proposed pattern allows two adjacent stripes to share the same hue value but not the same intensity level.

The aim of this pattern is to allow both intensity peaks (corresponding to the stripes central point) and edges between stripes to be detected.

Next section presents the algorithmic details concerning the detection of both intensity peaks and edges and the decoding of the new hybrid pattern.

4 Pattern segmentation and decoding

Given the system configuration of our experimental setup, i.e. a camera and a projector positioned aside and the pattern consisting of vertical stripes, the decoding process is based on horizontal scan-lines (in case that the projector and the camera are not approximately aside, a stereo pair rectification algorithm can be used to

transform the images taking into account the geometry of the system [27]).

According to the aim of the proposed hybrid pattern, for every scan-line, the intensity peaks corresponding to the centre of every stripe and the edges between the stripes must be located. Afterwards, the detected stripes centres and the edges between adjacent stripes must be matched with the projected pattern in order to obtain correspondences which would be triangulated for reconstructing the illuminated object.

4.1 Scan-line segmentation

First of all, we take advantage of the pattern square intensity profile in order to segment regions in the scan-line corresponding to medium intensity and high intensity stripes of the pattern. Given the M channel defined in Section 3.2, its second derivative is calculated. The following numeric filtered linear derivative is used [28]

$$g(i) = \sum_{c=1}^{o/2} (f(i+c) - f(i-c)) \tag{10}$$

where f is the original function (in our case, the M channel of the current scan-line), g is the corresponding filtered linear derivative of order o, and i indicates the element index which is being filtered. This filtering operator is applied twice in order to obtain the filtered second derivative of the M channel. The order o of the filter must be chosen according to the apparent stripes width of the pattern in the images, which must be always greater than o (in our experiments we have used $o = 6$). The effects of applying this filtered derivative in a signal showing a maximum and a minimum is shown in Figure 2a. Note that the second derivative strongly enhances the intensity differences, producing more enhanced peaks. Then, the second derivative can be used to segment maximum and minimum intensity regions of the pattern by simply binarizing it. Concretely, regions where the second derivative is less than 0 are binarised to 1 and inversely. In Figure 2b a portion of a scan-line of the M channel of a real image is shown. The square signal is the binarised second derivative of the scan-line. The results of segmenting the stripes by

using this technique is also shown in Figure 2c-d. In this case, the M channel of an image with a human hand under the pattern illumination is shown (see Figure 2c). As can be seen in the resulting binarised image in Figure 2d, the second derivative is able to segment the stripes even in regions with low intensity (see the black background behind the hand). In order to avoid to process the image background a minimum intensity threshold applied to the M channel can be used in order to detect such regions. Note that the stripes corresponding to the human hand have all been correctly segmented even if there are contrast variations through the image.

Figure 2

The segmentation of the scan-line by using the second derivative allows us to distinguish between regions corresponding to medium and high intensity stripes. The centre of the stripe corresponding to every segmented region can be detected with sub-pixel accuracy using different peak detectors [26]. After some experiments, we have found that given the width of the imaged stripes the peak detector obtaining better accuracy is the one based on a normalised centroid. This peak detector first normalises the pixel intensities by dividing it by the maximum intensity in the region. Then, it calculates the sub-pixel centre of mass of the region taking into account only those pixels for which its normalised intensity is higher than a certain threshold (set to 0.9 during the experiments).

The edges between adjacent stripes can be located by using the following strategy. Given the stripe centre of a high intensity stripe, the sub-pixel position of the surrounding edges corresponds to the two closest maxima in the function

$$g = dR^2 + dG^2 + dB^2 + dM^2 \tag{11}$$

where dR, dG, dB and dM denotes the first derivative (calculated by the linear filter in equation 10) of the Red, Green, Blue and M channel of the scan-line. This strategy is very similar to the used in the edge-based stripe pattern by Zhang et al. [15] but in this case we take profit also of the alternating intensity profile of the pattern. The sub-pixel position of the maxima in g are calculated by using the peak detector by Blais and Rioux [28] when a zero-crossing is detected in the first

derivative of g.

4.2 Scan-line decoding

Once the stripes centres and the edges have been located, it is necessary to match them with the projected pattern, a process which is known as *pattern decoding*. The decoding strategy of our technique is based on the hue value of every segmented stripe in the scan-line.

The hue value of each segmented stripe is set to be the median hue value of all the pixels included in the corresponding region. Since only four hue values are projected, the hue of each region can be easily identified by using simple rules over the RGB colour like the following ones (taking into account the hue values used in the pattern proposed in Section 3)

$$\begin{cases} \max(R,G,B) = R \text{ and } G < 0.5R \text{ and } B < 0.5R \Rightarrow & H = 0^\circ \\ \max(R,G,B) = G \text{ and } B < 0.5G & \Rightarrow \; H = 90^\circ \\ \min(R,G,B) = R & \Rightarrow H = 180^\circ \\ \min(R,G,B) = G & \Rightarrow H = 270^\circ \end{cases} \tag{12}$$

Then, for every scan-line a sequence of stripes is recovered and a hue value and an intensity level (medium or high) is assigned to each one of them. By using the labeling function presented in equation 8 a numeric sequence representing the scan-line can be obtained. Ideally, if all the pattern stripes are segmented in the scan-line and their hue value and intensity level are correctly identified, the numeric sequence corresponding to the scan-line is equal to the encoded sequence in equation 9. In such an ideal case, the matching of correspondences between the projected sequence and the perceived one is straightforward. However, in most cases either the whole sequence of projected stripes is not visible in the image scan-line or some of them are incorrectly labeled or disorders may occur.

In order to robustly solve the correspondence problem that arises we also use dynamic programming as in the techniques introduced in Section 3. For details about the dynamic programming algorithm we refer to [15,29]. In our case, the score function measuring the similarity between a certain element of the projected sequence and an element of the perceived one returns only the maximum value when both elements are the same.

5 System's calibration

In this section we present the different modelling steps that must be considered in order to obtain the 3D reconstruction of the illuminated scene. First of all, we explain how the geometric calibration of the system is performed. Afterwards, some imperfections of the real system are modeled in order to increase the robustness and the accuracy.

5.1 Geometric calibration

Geometric calibration of a stereo vision system is the first step towards 3D reconstruction of the illuminated object. A calibration algorithm estimates a set of parameters of a mathematical model describing the physical behaviour of the sensor, making it possible to measure an object accurately by means of its projection on the image plane [4,30].

We refer to the comparative review presented by Salvi et al. [30] for the mathematical details of the calibration principle. This review concludes that non-linear models which take into account the radial distortion introduced by the lenses are enough to obtain high accuracy. In our case, an adaptation of the Faugeras calibration method which includes radial distortion has been used [4].

In order to calibrate the camera it is necessary to obtain a set of non-coplanar 3D points and its corresponding 2D projections. To obtain the sample points, we used

two orthogonal white panels each one containing 20 black squares, placed about 1 meter in front of the camera. The world frame is positioned at the intersection of both panels as shown in Figure 3. Then, the camera captures an image and the 2D corners are detected with sub-pixel accuracy by means of image processing. Finally, 2D and 3D points are used to iterate the calibration algorithm to convergence.

Figure 3

The LCD projector can be modelled as an inverse camera, so that the same camera model remains valid. In order to calibrate the projector, we assume that the camera has been previously calibrated and that the calibrating panels remain in the same position. The projector is placed aside the camera so that the angle between the optical axis of both devices is about 15 degrees. A white grid pattern is projected onto both panels. We chose the grid cross-points of the pattern as 2D points. Then, it is necessary to capture an image with the camera. The grid cross-points can be detected with sub-pixel accuracy. Afterwards, since the camera is already calibrated, the 3D points can be reconstructed, taking into account certain geometrical constraints. Remember that the grid is projected onto the calibration panels and the world frame is positioned on the bottom angle of both panels, as shown in Figure 6. Then, the left panel lies on the plane $Y = 0$, while the equation of the plane containing the right panel is $X = 0$. In this way, the 3D points, corresponding to the grid cross-points projected onto the panels, can be triangulated by intersecting the camera rays with the equation of these planes. Then, the set of 2D and 3D points obtained with this procedure are used to calibrate the projector's model.

5.2 RGB channel alignment

Ideally, colour cameras should perceive an intensity peak of white light at the same image coordinates in the three RGB channels. In practice there is an offset between the sub-pixel location of the intensity peaks in every RGB channel. This phenomenon is known as RGB channel misalignment. It is caused by spatial misalignments in the different CCD cells perceiving the red, green and blue light respectively. Al-

though the order of these misalignments is usually below or around one pixel, it can produce higher order errors in 3D reconstruction. Furthermore, it can be easily shown experimentally that it is not only cameras the ones suffering from RGB channel misalignment, but also LCD projectors.

Some authors propose to reduce the camera RGB channel misalignment by viewing an object providing reference points (like a checkerboard) and locating such points in the three channels separately. Afterwards, an homography can be calculated relating the position of the points in the red channel with respect to the ones in the green channel, and another homography doing the same between the points in the blue and the green channel. These homographies are then used to reduce the misalignment in the images [15]. Nevertheless, this method totally ignores the RGB misalignment in the LCD projector.

We propose to minimise the RGB misalignment observed in the camera images taking into account both the camera and the projector at the same time. Since the decoding process is made through horizontal scan-lines, we propose to minimise the RGB channel misalignment in every scan-line. The algorithm is simple. A flat white panel is set in front of the camera and the projector at the typical working distance (about 1 meter in our case). Three patterns consisting of a sequence of narrow stripes separated by black gaps are projected sequentially: each one having red stripes, green stripes and blue stripes, respectively. Images of every projected pattern are taken by the camera. For every scan-line the sub-pixel position of every intensity peak is located with the detector by Blais and Rioux [28]. In the image containing the red stripes, the red channel is used to locate the peaks. Similarly, the green channel is used in the image containing green stripes and the blue channel for the case of blue stripes. For every scan-line, the median of the sub-pixel offsets and the relative positions between the three channels are stored. We have observed that the relative positions of the channels coincide in all the scan-lines and that the relative offsets are very similar. That is why we finally store two unique offsets between the central channel and the other two. In our experimental setup we have found that in the images the central channel is the blue while the green channel is approximately at 1.26 pixels at its left, and the red channel is about 0.53 pixels at its

right. These offsets have been named ${}^{b}\mathbf{H}_{g}$ and ${}^{b}\mathbf{H}_{r}$, respectively. In order to reduce the global misalignment observed in an image it is necessary to apply the offset ${}^{b}\mathbf{H}_{g}$ to the green channel and the offset ${}^{g}\mathbf{H}_{r}$ to the red one and then combine the transformed channels with the original blue channel in order to obtained the rectified image. Note that the intensity of every transformed pixel in the new channels must be interpolated from the neighboring pixels in the corresponding source channel since the offsets are at sub-pixel precision.

5.3 Colour selection and calibration

The proposed hybrid pattern is based on the assumption that two different intensity levels will be distinguished in the image regions containing the projected pattern. This is necessary in order to be able to segment the stripes and to identify them as medium or high intensity stripes. Ideally, any discrete RGB instruction $\mathbf{c}$ with the same level of intensity i should produce the projection of light with the same intensity I so that the only varying parameter is the wavelength λ. Similarly, a perfect camera should be able to digitise any incident light of wavelength λ and a certain intensity I to a discrete RGB triplet $\mathbf{C}$ with intensity i. In real conditions, however, the mapping from the RGB projection instruction $\mathbf{c}$ to the imaged RGB triplet $\mathbf{C}$ is a strongly non-linear process. Several radiometric models of a structured light system composed of a LCD projector and a colour camera can be found in the bibliography, as for example, in [31,32]. Such models take also into account the reflective properties of the illuminated scene so that a model of the object albedo is also estimated. Indeed, the complete radiometric model identification of our experimental setup could be performed. However, taking into account that only 4 colours (hue values) and 2 different intensity levels are projected in the hybrid pattern, a simpler algorithm can be performed. Furthermore, since a one-shot technique is based on a unique pattern projection, the albedo of the object cannot be recovered so that this part of the radiometric model is unnecessary.

In Figure 4 the system response when projecting the four hue values with different intensities (from the point of view of the RGB instruction $\mathbf{c}$) is plotted. The response

is the value of the M channel obtained from the camera image. These curves have been obtained by projecting solid patterns with each one of the 4 hue values with different intensity levels. A neutral colour panel has been used in order to not excessively perturb the projected light. As can be seen, the system response for each one of the colours is different and it is clearly non-linear. Note that in order to obtain the same value in the M channel of the image, each one of the 4 colours must be projected with a different intensity level. Therefore, as shown in Figure 4, by choosing the two desired values of the M channel, the required projecting intensities for each one of the colours can be approximately obtained using the curves. With this simple calibration procedure, the hybrid pattern can be adapted in order to produce a suitable system response without need of performing a whole radiometric calibration.

Figure 4

Another typical problem of a structured light system setup is the projector-camera colour crosstalk [15, 31, 32]. In order to see the role played by this phenomenon let us remember the model presented by Caspi et al. [31]

$$\underbrace{\begin{pmatrix} R \\ G \\ B \end{pmatrix}}_{\mathbf{C}} = \underbrace{\begin{pmatrix} a_{rr} & a_{rg} & a_{rb} \\ a_{gr} & a_{gg} & a_{gb} \\ a_{br} & a_{bg} & a_{bb} \end{pmatrix}}_{\mathbf{A}} \mathbf{KP} \underbrace{\begin{Bmatrix} r \\ g \\ b \end{Bmatrix}}_{\mathbf{c}} + \underbrace{\begin{pmatrix} R_0 \\ G_0 \\ B_0 \end{pmatrix}}_{\mathbf{C}_0} \tag{13}$$

where $\mathbf{c}$ is the RGB projection instruction sent to the projector and $\mathbf{C}$ the corresponding RGB triplet digitised by the camera. The consign $\mathbf{c}$ is actually modified by the non-linear behaviour of the projector which actually projects a colour denoted by $\mathbf{P}$. $\mathbf{K}$ is a 3×3 matrix modelling the albedo of the illuminated object, and $\mathbf{A}$ is the colour crosstalk matrix, while $\mathbf{C}_0$ is the RGB tripled digitised by the camera when there is only ambient lighting. Therefore, $\mathbf{A}$ expresses how the RGB channels of the camera are affected by the RGB channels of the projector.

Experimentally it can be observed that usually the strongest crosstalk appears when projecting green, since it is not only detected by the green channel of the camera but also by the red one. In order to minimise the colour crosstalk, we perform two

calibration processes. First, a solid pattern is projected using each one of the 4 calibrated colours for medium intensity stripes. These colours are defined by the 4 hues selected in Section 3 and the 4 intensities selected in this section which aim to produce the same medium intensity in the M channel. Then, the following linear system is defined

$$\begin{pmatrix} R_1 & R_2 & R_3 & R_4 \\ G_1 & G_2 & G_3 & G_4 \\ B_1 & B_2 & B_3 & B_4 \end{pmatrix} = \mathbf{A}_m \begin{pmatrix} r_1 & r_2 & r_3 & r_4 \\ g_1 & g_2 & g_3 & g_4 \\ b_1 & b_2 & b_3 & b_4 \end{pmatrix} \tag{14}$$

where (R_i, G_i, B_i) is the average colour perceived by the camera when the instruction (r_i, g_i, b_i) is sent to the projector and using a colour neutral panel ($\mathbf{K} = \mathbb{I}_3$). Matrix $\mathbf{A}_m$ can be numerically calculated by using singular value decomposition. The inverse of this matrix can be used to partially remove the colour crosstalk of the stripes identified during the segmentation process as medium intensity stripes. Similarly, a matrix $\mathbf{A}_h$ is calculated by projecting the colours corresponding to the high intensity stripes.

6 Experimental results

In this section we show some experimental results validating the proposed hybrid pattern. The experimental setup consists of an Olympus Camedia digital camera and a Mitsubishi XL1U LCD projector which are positioned aside with a relative direction angle of about 15°. Both devices operate at 1024×768 pixels. The measuring volume is about 30 cm high, 40 *cm* wide and 20 cm deep.

The calibration of the system has been performed according to the steps presented in the previous section and summarised in Figure 5. The figure also shows the one-shot shape acquisition process and its interaction with the data obtained during the system calibration. Note that the image rectification step is not really necessary if the camera and the projector are positioned aside with similar tilt angle.

During the experiments three different patterns have been tested: the multi-slit

pattern (hereafter referred as peak-based pattern) shown in Figure 1a (the width of both the coloured slits and the black gaps have been set to 7 pixels), the stripe pattern (hereafter referred as edge-based pattern) by Zhang et al. shown in Figure 1b (the width of the stripes is also 7 pixels) and our hybrid pattern presented in Figure 1c (the stripes are also $W_m = W_h = 7$ pixels wide).

Figure 5

6.1 Quantitative results

The selected techniques have been first compared in terms of accuracy and resolution. A measure of the accuracy has been obtained by reconstructing a plane positioned in front of the camera and the projector at a distance of about 1.2 m. The accuracy has been characterised by the mean and the standard deviation of the distances between the reconstructed $3D$ points and the fitted plane. The figure for evaluating the resolution is the number of $3D$ reconstructed points.

When using the peak-based pattern, the $3D$ points are obtained by reconstructing the intensity peaks detected in the image. In the case of the edge-based pattern (Zhang et al.), the detected edges are used to triangulate $3D$ points. On the other hand, the hybrid pattern allows to obtain $3D$ points by reconstructing both intensity peaks and edges. Let us introduce the following notation to refer to the basic reconstruction strategies allowed in the hybrid pattern:

- M strategy: reconstruction of intensity peaks (maxima) corresponding to the centre of high intensity stripes.
- m strategy: reconstruction of intensity peaks (minima) corresponding to the centre of medium intensity stripes.
- E strategy: reconstruction of edges between adjacent stripes.

In fact, not only the individual strategies M, m and E have been used when testing the hybrid pattern, but also some combinations of them, namely $M + m$, $E + M$ and $E + M + m$.

Table 1 shows the numerical results obtained by every pattern and the different reconstruction strategies allowed by each one. In terms of accuracy the best result is obtained by the peak-based pattern. In second position we find the hybrid pattern when using the M strategy. In both cases, the average and the standard deviation of the error is much more lower that when using the edge-based pattern. Furthermore, we note that the accuracy of the hybrid pattern when reconstructing only edges (E strategy) is also better that the accuracy of the edge-based pattern. When using the m strategy with the hybrid pattern, the average error increases considerably while the standard deviation remains quite similar. This points out that the reconstruction of minima is much more affected by noise. Therefore, finding the sub-pixel position of an intensity minimum is a much more sensitive and inaccurate process than locating an intensity maximum. This is probably caused because the signal-to-noise ratio is lower in the medium intensity stripes of the hybrid pattern. In terms of resolution we note that strategies based on intensity peaks (M and m) obtain about half of the resolution that when reconstructing edges (E strategy). Therefore, from the numerical results it is not easy to decide which is the best individual strategy because there is a trade-off between accuracy and resolution.

Table 1

Let us now analyse the results when combining different reconstruction strategies thanks to the hybrid pattern. Using more than one strategy allows the resolution to be largely increased (see last column of Table 1). In general, the combination $M+m$ obtains the same order of resolution that the E strategy. Note however, that the accuracy of the hybrid pattern when using $M+m$ is slightly better than using the hybrid pattern and the E strategy. With respect to the edge-based pattern, the improvement on accuracy is much more significant. On the other hand, we can emphasise that the combination $E+M$ gets a resolution about 1.5 times greater than when using E or $M+m$ and obtains similar accuracy. Finally, the combination $E+M+m$ doubles the resolution of the E strategy. However, we observe a worse accuracy (the standard deviation of the error is higher). It seems therefore, that the inclusion of the minima intensity peaks degrades the performance of the hybrid pattern in terms of accuracy.

The results of the patterns are qualitative compared in the following section. These results will help us to decide which combination of reconstruction strategies give better results when using the proposed hybrid pattern.

6.2 Qualitative results

In this section we present and discuss the visual appearance of several reconstructions obtained with the selected patterns and the different reconstruction strategies. First of all we analyse the plane reconstruction explained in the previous section. In Figure 6a the reconstruction obtained using the peak-based pattern is shown. Note that the reconstructed surface is very smooth confirming the accuracy results presented before. Figure 6b shows the surface obtained when using the edge-based pattern. Note that some visual artifacts, concretely vertical ridges, appear along the surface. Figure 6c-d shows the surfaces obtained with the hybrid pattern when reconstructing only maxima intensity peaks and edges, respectively. Note that in this case a smooth surface is also obtained with the M strategy, while some ridges also appear when using the E strategy. Nevertheless, we remark that the ridges obtained with the hybrid pattern when reconstructing edges are much smoother that the ones appearing in the reconstruction provided by the edge-based pattern. This fact seems to confirm the accuracy results obtained by both techniques. The reconstruction obtained by the hybrid pattern when using the $M + m$ strategy is shown in Figure 6e. Note that a smooth plane is still obtained. The result given by the hybrid pattern and the $E + M$ strategy is depicted in Figure 6f. Even if a small loss of accuracy has been predicted in the previous section, we can observe that the visual appearance is still better that when using the edge-based pattern. Nevertheless, as expected, the inclusion of the edges in the reconstruction process make some smooth ridges to appear. Finally, Figure 6g presents the reconstruction when using the strategy $E+M+m$. The appearance is bit noisier than the obtained by the $E + M$ strategy.

Figure 6

Figure 7 presents three colour neutral test objects, namely a mask, a horse statue

and a sun statue. The images corresponding to the projected patterns (peak-based, edge-based and hybrid pattern) are also shown. The reconstruction results are plotted in Figure 8.

Figure 7

As expected, the peak-based pattern (first row) and the hybrid pattern using the $M + m$ strategy (second row) obtain both smooth surfaces and almost absence of artifacts. However, the reconstructions obtained by the peak-based pattern are too much smoothed due to the low resolution achieved, so that details like the mask mouth and the horse eye are not visible. Note that in the result corresponding to the hybrid pattern and the $M + m$ strategy such details begin to be appreciated. The gain in visual appearance obtained by the hybrid pattern and the $M + m$ strategy is confirmed in the sun reconstruction, where the eyes, nose and mouth are better represented than in the case of the peak-based pattern.

The edge-based pattern (third row) shows pretty performance producing highly detailed surfaces with absence of important holes. On the other hand, vertical ridges appear, degrading the smoothness of the surfaces. If we compare the edge-based pattern and the hybrid pattern using the $M+m$ strategy, the former seems to get a bit more level of detail. However, since the latter does not have the inconvenience of the vertical ridges, it is difficult to decide which results are better from a qualitative point of view.

Finally, the last two rows of Figure 8 present the results obtained by the hybrid pattern when using the $E + M$ and $E + M + m$ strategies. The $E + M$ strategy obtains a higher level of detail (see for example the horse mouth, which was not detected in the previous cases) than the previous techniques. Due to the inclusion of the edges in the triangulation process, the vertical ridges also appear. However, these artifacts are smoother than the ones observed in the edge-based pattern results. On the other hand, it seems that some additional small holes appear and that some contours near shadows are less well defined. Note also that the differences between the results of the $E + M$ strategy and the $E + M + m$ strategy can be hardly distinguished.

Figure 8

In general, we think that the results obtained by the hybrid pattern are globally better than the ones obtained by the edge-based pattern proposed by Zhang et al. [15]. We think that the increase in resolution achieved by the combination of edges and intensity peaks allows some small details of the objects to be better reconstructed. In order to show this, Figure 9 presents an ampliation of the sun reconstruction focusing on the zone corresponding to the nose, eyes and mouth. As can be seen in Figure 9b, the reconstruction obtained by the peak-based pattern is quite poor. Note that the quality of the reconstruction is already improved when using the hybrid pattern and the $M + m$ strategy as shown in Figure 9e. Even if the reconstruction obtained by the edge-based pattern (see Figure 9c) shows great level of detail, we can see that the results obtained by the hybrid pattern and the $E + M$ and $E + M + m$ strategies (Figure 9d and f) are even better. Note that the shape of the sun nose and mouth and the nose corresponding to the moon are much more clear in the hybrid pattern reconstructions. We can also point out that the fold appearing in the left cheek of the sun is better appreciated in the hybrid pattern results.

Figure 9

Finally, we present to reconstructions obtained with the hybrid pattern taking into account slightly coloured objects. As said in the beginning of the paper, one-shot techniques projecting colourful patterns have usually problems when dealing with non-colour neutral objects. We first show the results obtained when reconstructing a human hand. The skin usually introduces a strong gain on the red component of the pattern and at the same time attenuates the whole luminosity reaching the camera. Figure 10a shows the picture of the hand with the projected pattern. The stripe and colour segmentation obtained is shown in Figure 10b. The reconstruction results are plotted in form of points and surface in Figure 10c-d. As can be seen, the level of detail is as expected very high (see the veins passing through the hand) taking into account that a unique pattern is projected. The second example consists of three coloured sheets of papers as shown in Figure 10e. The obtained reconstruction is presented in Figure 10f.

Figure 10

7 Conclusions

This paper presents a new coloured pattern for one-shot shape acquisition. The new pattern aims to improve similar existing patterns in terms of resolution and accuracy. Among the one-shot colour-based techniques two of the most frequent patterns are based on multi-slits or stripes. The former is based on coloured bands of pixels (slits) separated by black gaps so that consecutive bands can have the same hue value. In the latter, the coloured bands are adjacent (stripes) so that two consecutive bands cannot have the same hue value. In multi-slit patterns the intensity profile has periodic peaks corresponding to the central point of the slits. These are used to find correspondences. In stripe patterns, since the intensity profile is flat, edges between adjacent stripes are used to find correspondences. Both types of patterns are usually coloured according to De Bruijn sequences. The advantage of stripe patterns over multi-slit patterns is that the resolution is higher since no black gaps must be inserted. However, the number of hue values required is also higher since adjacent stripes must be different.

In this paper we present a new hybrid pattern which combines the properties of a multi-slit pattern and the ones of an edge-based pattern. First of all, the new coding strategy allows a stripe pattern to be defined by using less hue values than usual. The new pattern contains $2n^m$ stripes, where the window length m is set to be 3, and n, the number of different hue values used, remains free in order to obtain the desired resolution. We remind that classical multi-slit patterns contain n^m slits, since consecutive slits can have the same hue. In the case of stripe patterns, the number of coloured bands is $n(n-1)^{m-1}$. Therefore, the resolution of the new hybrid pattern given a certain number of hue values is always larger than in classic multi-slit and stripe patterns.

The hybrid pattern has a multi-slit structure in the Intensity channel while it has a stripe appearance in the RGB space. Then, the odd stripes have a medium-intensity

value while the even stripes have high-intensity. Since all the stripes are coloured, both maxima and minima intensity peaks can be located in the images. Furthermore, edges between stripes can also be detected with accuracy. This allows the pattern resolution to be increased in a factor up to 2.

A hybrid pattern using the new coding strategy has been implemented with only $n = 4$ hue levels, obtaining 128 stripes. This pattern has been experimentally compared to a 64-slit pattern (and the corresponding 64 black gaps among them) using also $n = 4$ hue levels and the stripe pattern proposed by Zhang et al. [15] which contains 125 stripes based on 8 different hue levels. The three patterns have been analysed from quantitative and qualitative points of view. The accuracy and resolution of each technique has been characterised by reconstructing a plane and measuring the average and the standard deviation of the fitting error, and the number of $3D$ points, respectively. A first conclusion is that reconstructing intensity peaks always provides more accurate measurements than when reconstructing edges. Indeed, the sub-pixel estimation of an intensity peak is more stable than the sub-pixel location of an edge between coloured stripes. This numerical results are confirmed by the visual appearance of the reconstructions. When reconstructing intensity peaks, the obtained surfaces are smoother, while in the case of edges, visual artifacts like ridges appear.

The peak-based (multi-slit) pattern has obtained the better results in terms of accuracy. Nevertheless, its performance in terms of resolution is quite poor (half of the obtained by the edge-based pattern). The better results obtained with the hybrid pattern arise when reconstructing only the maxima intensity peaks. In this case, however, the resolution is also half of the obtained by the edge-based pattern (pure stripe pattern). The accuracy of the hybrid pattern when reconstructing only edges is still better than the one obtained by the edge-based pattern. This fact is visually confirmed by observing the plane reconstruction. The numerical results show that the accuracy diminishes when using the hybrid pattern and reconstructing only minima intensity peaks. This confirms that the medium-intensity stripes are more sensitive to noise. However, the most interesting results are obtained when using the hybrid pattern and combining edges and maxima intensity peaks. In this

case, the accuracy is not so good as for the peak-based pattern, but it is better than the edge-based pattern. In addition to this, the resolution is increased in a factor of 1.5. The resolution of the edge-based pattern is overcome by a factor of 2 when using the hybrid pattern and when reconstructing edges and maxima and minima intensity peaks. Nevertheless, the accuracy is slightly worse than when reconstructing only edges and maxima intensity peaks.

Several objects containing high level of details have been reconstructed by using the three selected patterns. The visual appearance of the reconstructions confirm that when using only intensity peaks the obtained surfaces are smoother. Otherwise, when using edges periodic ridges appear. On the other hand, the importance of increasing the resolution is visually confirmed. In fact, while with the peak-based pattern coarse object's reconstructions are obtained, with the edge-based and the hybrid pattern they are more detailed. Furthermore, it has been shown that when using the hybrid pattern and reconstructing edges and maxima intensity peaks, more object's details are distinguished in the reconstructions. Finally, two results showing the performance of the new hybrid pattern when reconstructing more difficult textures, like human skin, have been presented, obtaining good results even if the colours of the pattern are perturbed by the object's albedo.

References

[1] P. J. Besl, Advances in Machine Vision: Architectures and Applications, Springer-Verlag, 1988, Ch. Active Optical Range Imaging Sensors, pp. 1–53.

[2] R. Jarvis, Range sensing for computer vision, Advances in Image Communications. Elsevier Science Publishers, Amsterdam (1993) 17–56.

[3] F. Chen, G. Brown, M. Song, Overview of three-dimensional shape measurement using optical methods, Optical Engineering 1 (39) (2000) 10–22.

[4] O. Faugeras, Three-Dimensional Computer Vision, MIT Press, 1993.

[5] X. Armangué, J. Salvi, Overall view regarding fundamental matrix estimation, Image and Vision Computing 21 (2) (2003) 205–220.

[6] J. Salvi, J. Pagès, J. Batlle, Pattern codification strategies in structured light systems, Pattern Recognition 37 (4) (2004) 827–849.

[7] T. Etzion, Constructions for perfect maps and pseudorandom arrays, IEEE Transactions on information theory 34 (5) (1988) 1308–1316.

[8] P. Griffin, L. Narasimhan, S. Yee, Generation of uniquely encoded light patterns for range data acquisition, Pattern Recognition 25 (6) (1992) 609–616.

[9] C. J. Davies, M. S. Nixon, A hough transform for detecting the location and orientation of 3-dimensional surfaces via color encoded spots, IEEE Transactions on systems, man and cybernetics 28 (1) (1998) 90–95.

[10] H. J. W. Spoelder, F. M. Vos, E. M. Petriu, F. C. A. Groen, Some aspects of pseudo random binary array-based surface characterization, IEEE Transactions on instrumentation and measurement 49 (6) (2000) 1331–1336.

[11] E. M. Petriu, Z. Sakr, S. H. J. W., A. Moica, Object recognition using pseudo-random color encoded structured light, in: Proceedings of the 17th IEEE Instrumentation and Measurement technology Conference, Vol. 3, 2000, pp. 1237–1241.

[12] F. J. MacWilliams, N. J. A. Sloane, Pseudorandom sequences and arrays, Proceedings of the IEEE 64 (12) (1976) 1715–1729.

[13] H. Hügli, G. Maître, Generation and use of color pseudo random sequences for coding structured light in active ranging, in: Proceedings of Industrial Inspection, Vol. 1010, 1989, pp. 75–82.

[14] T. Monks, J. Carter, Improved stripe matching for colour encoded structured light, in: 5th International Conference on Computer Analysis of Images and Patterns, 1993, pp. 476–485.

[15] L. Zhang, B. Curless, S. M. Seitz, Rapid shape acquisition using color structured light and multi-pass dynamic programming, in: Int. Symposium on 3D Data Processing Visualization and Transmission, Padova, Italy, 2002, pp. 24–36.

[16] J. Salvi, J. Batlle, E. Mouaddib, A robust-coded pattern projection for dynamic 3d scene measurement, Pattern Recognition Letters (19) (1998) 1055–1065.

[17] P. Lavoie, D. Ionescu, E. Petriu, A high precision 3D object reconstruction method using a color coded grid and nurbs, in: Proceedings of the International Conference on Image Analysis and Processing, Venice, Italy, 1999, pp. 370–375.

[18] M. Maruyama, S. Abe, Range sensing by projecting multiple slits with random cuts, Pattern Analysis and Machine Intelligence 15 (6) (1993) 647–651.

[19] C. Guan, L. Hassebrook, D. Lau, Composite structured light pattern for three-dimensional video, Optics express 11 (5) (2003) 406–417.

[20] T. Koninckx, A. Griesser, L. Van Gool, Real-time range scanning of deformable surfaces by adaptively coded structured light, in: Int. Conf. on 3-D Digitial Imaging and Modelling, Banff Canada, 2003, pp. 293–300.

[21] T. Koninckx, I. Geys, T. Jaeggli, L. Van Gool, A graph cut based adaptive structured light approach for real-time range acquisition, in: Int. Symposium on 3D Data Processing, Visualization and Transmission, Thessaloniki, Greece, 2004, pp. 413–421.

[22] J. L. Posdamer, M. D. Altschuler, Surface measurement by space-encoded projected beam systems, Computer Vision, Graphics and Image Processing 18 (1) (1982) 1–17.

[23] M. Trobina, Error model of a coded-light range sensor, Technical report, Communication Technology Laboratory, ETH Zentrum, Zurich (1995).

[24] C. Chen, Y. Hung, C. Chiang, J. Wu, Range data acquisition using color structured lighting and stereo vision, Image and Vision Computing 15 (1997) 445–456.

[25] T. P. Monks, J. N. Carter, C. H. Shadle, Colour-encoded structured light for digitisation of real-time 3D data, in: IEE Int. Conf. on Image Proc., 1992, pp. 327–30.

[26] E. Trucco, R. B. Fisher, A. W. Fitzgibbon, D. K. Naidu, Calibration, data consistency and model acquisition with laser stripers, Int. Journal Computer Integrated Manufacturing 11 (4) (1998) 293–310.

[27] A. Fusiello, E. Trucco, A. Verri, A compact algorithm for rectification of stereo pairs, Machine Vision and Applications 12 (1) (2000) 16–22.

[28] F. Blais, M. Rioux, Real-time numerical peak detector, in: Signal Processing, Vol. 11, 1986, pp. 145–155.

[29] I. J. Cox, S. L. Hingorani, S. B. Rao, B. M. Maggs, A maximum likelihood stereo algorithm, Computer Vision and Image Understanding 63 (3) (1996) 542–567.

[30] J. Salvi, X. Armangué, J. Batlle, A comparative review of camera calibrating methods with accuracy evaluation, Pattern Recognition 35 (7) (2002) 1617–1635.

[31] D. Caspi, N. Kiryati, J. Shamir, Range imaging with adaptive color structured light, Pattern Analysis and Machine Intelligence 20 (5) (1998) 470–480.

[32] S. K. Nayar, H. Peri, M. Grossber, P. Belhumeur, A projector with automatic radiometric screen compensation, Tech. rep., Department of Computer Science, Columbia University, New York (2003).

Figure legends

Figure 1: One-shot patterns compared in this paper. a) Stripe pattern by Zhang et al. b) Multi-slit pattern similar to the one proposed by Monks et al. c) Hybrid pattern in the RGB space. d) Hybrid pattern intensity profile.

Figure 2: Stripes segmentation through a filtered second derivative. a) Behaviour of the 1st and 2nd derivative on a synthetic signal. b) Segmentation of maxima and minima using the 2nd derivative. c) M channel of an image of a human hand under the pattern illumination. d) Stripes segmented by using the 2nd Derivative.

Figure 3: Geometric calibration setup.

Figure 4: Non-linear response of the system when projecting the four selected hue values with different intensity levels.

Figure 5: Schema of the system calibration and the one-shot acquisition steps.

Figure 6: Plane reconstruction. a) Peak-based pattern. b) Edge-based pattern. c) Hybrid pattern: M strategy. d) Hybrid pattern: E strategy. e) Hybrid pattern: M+m strategy. f) Hybrid pattern: E+M strategy. g) E+M+m strategy.

Figure 7: Test objects under the patterns illumination.

Figure 8: Reconstruction results of the test objects. The first row corresponds to the peak-based pattern; second row: hybrid pattern and the $M + m$ strategy; third row: edge-based; fourth row: hybrid pattern and the $E + M$ strategy; sixth row: hybrid pattern and $E + M + m$ strategy.

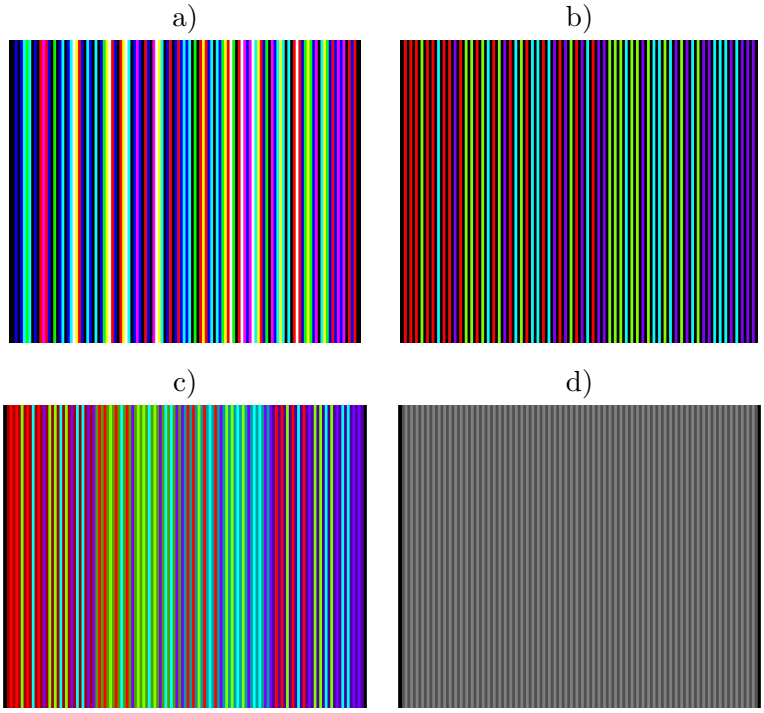
Figure 9: Zone enlargement of the sun reconstruction. a) Picture of the real object. b) Peak-based pattern. c) Edge-based pattern. d) Hybrid pattern: M+E strategy. e) Hybrid pattern: M+m strategy. f) Hybrid pattern: E+M+m strategy.

Figure 10: Reconstruction example of non-colour neutral objects. Human hand: a) Skin under the pattern illumination. b) Colour segmentation of the received pattern. c) Cloud of reconstructed points. d) Surface fitting the reconstructed hand. Coloured papers: e) the coloured sheets used in the experiment. f) The reconstructed surface.

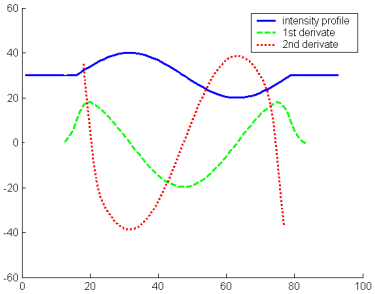
Table legends

Table 1: Accuracy results of plane fitting obtained from the three strategies compared in the paper. E denotes edges, M maxima intensity peaks and m minima intensity peaks.

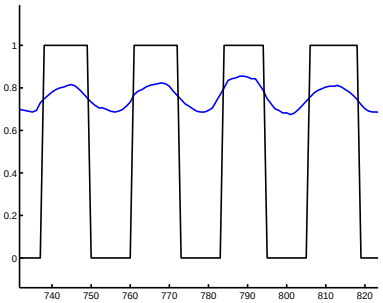
Figures



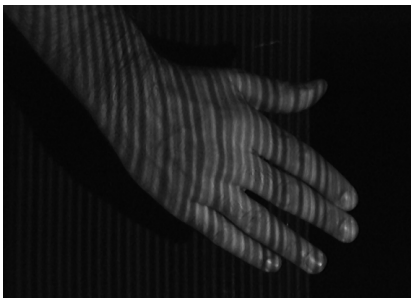
a)


intensity profile
1st derivate
2nd derivate
60
40
20
0
-20
-40
-60
0
20
40
60
80
100

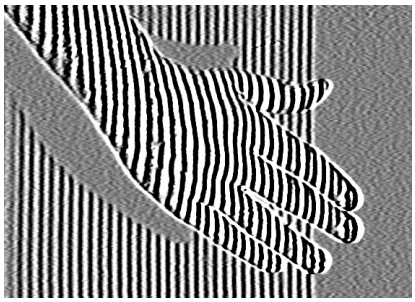
b)

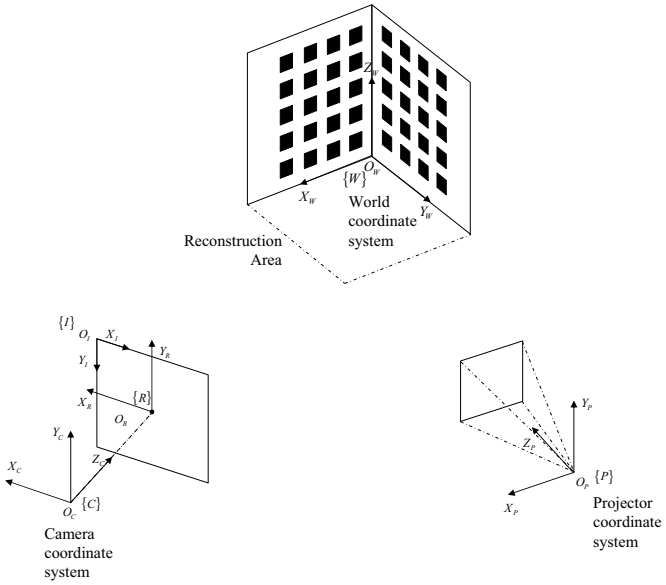

1
0.8
0.6
0.4
0.2
0
740
750
760
770
780
790
800
810
820

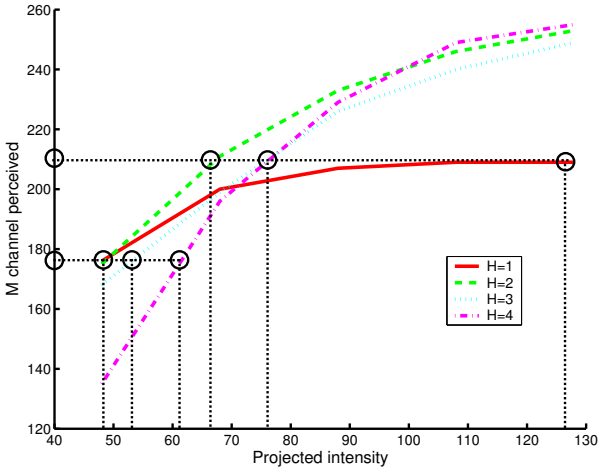
c)

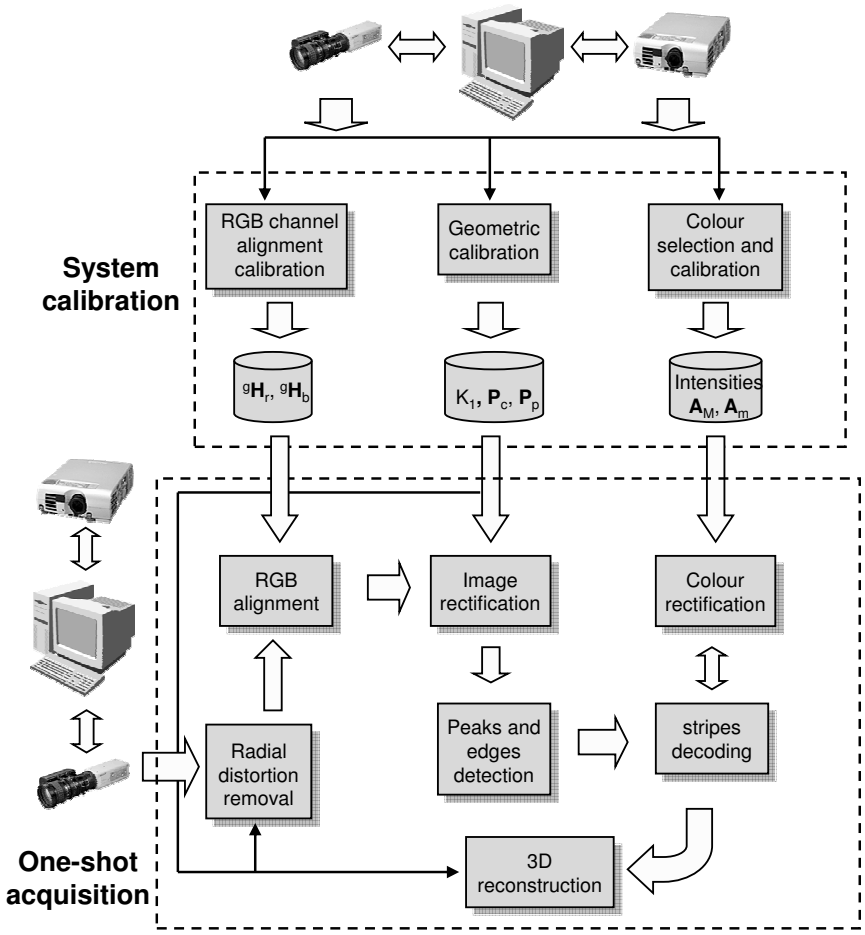


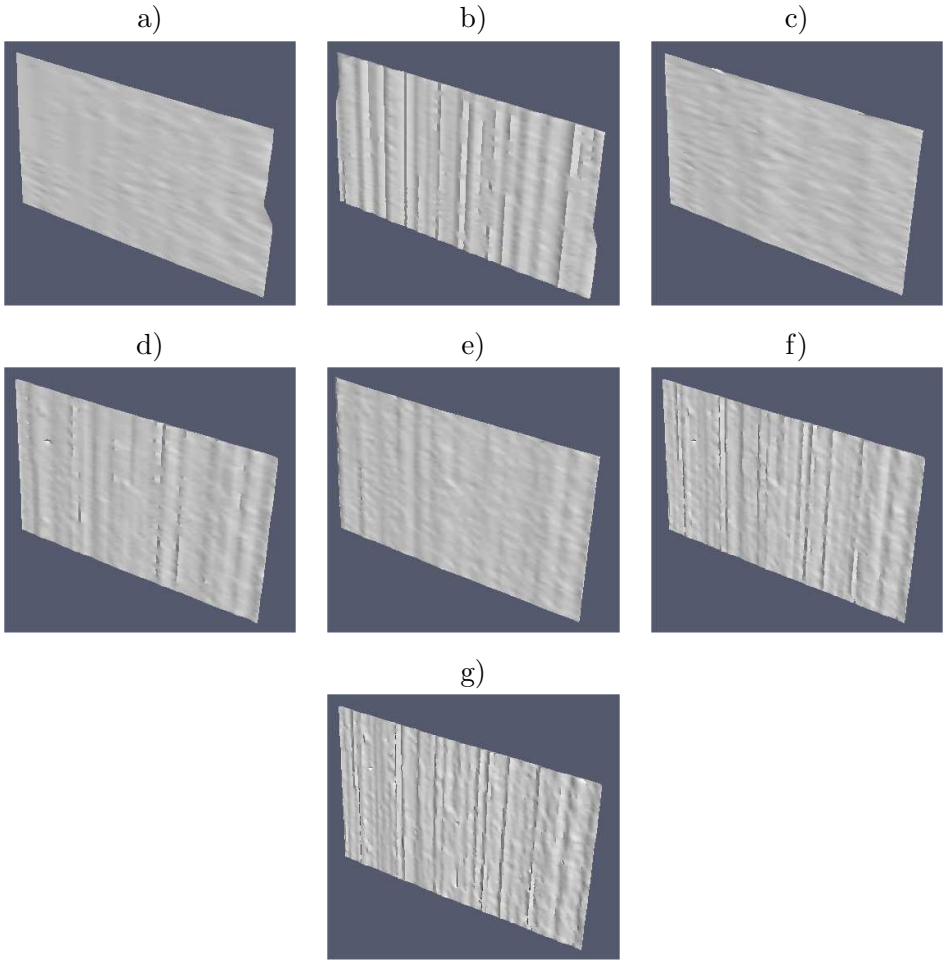
d)

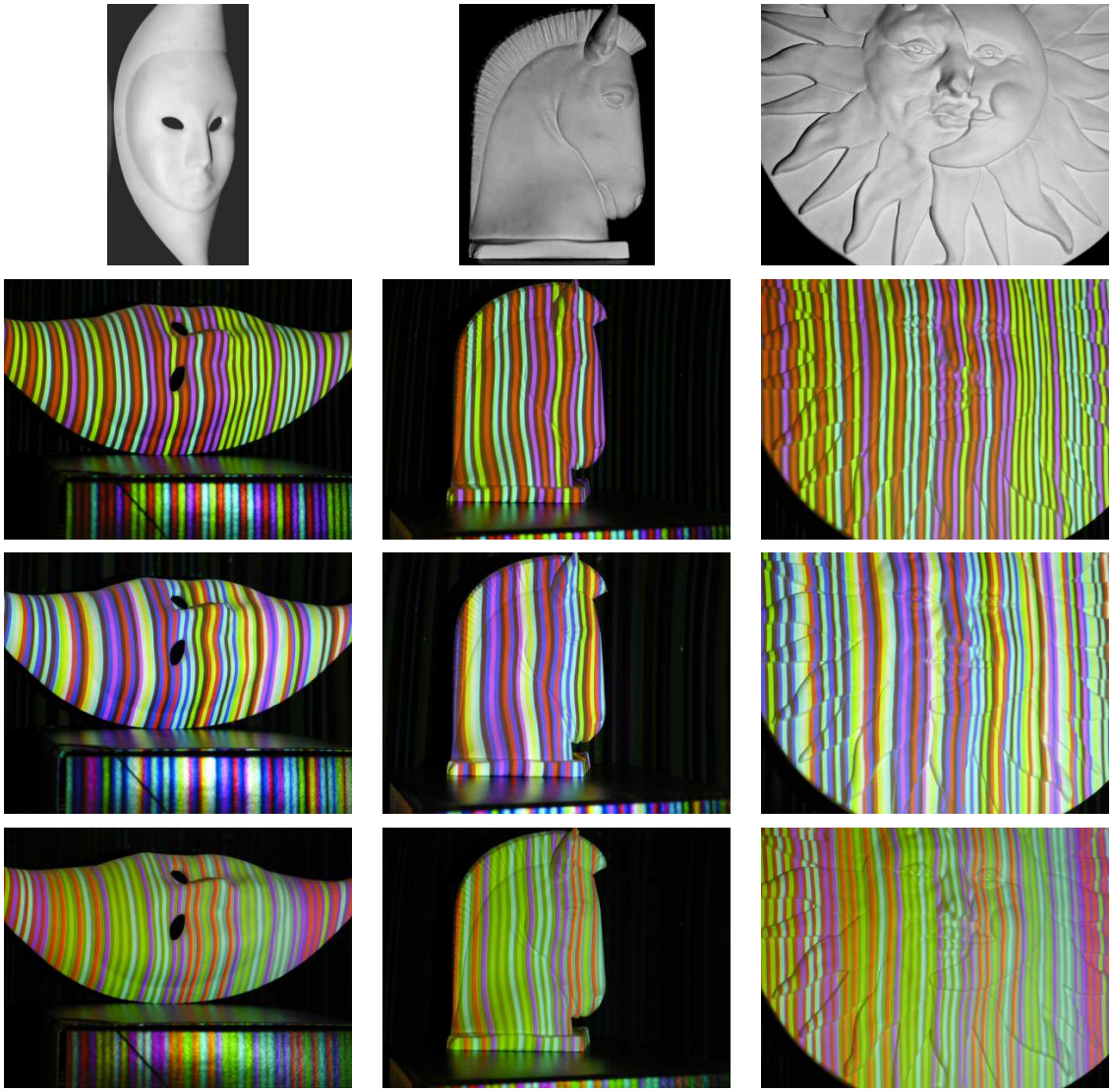


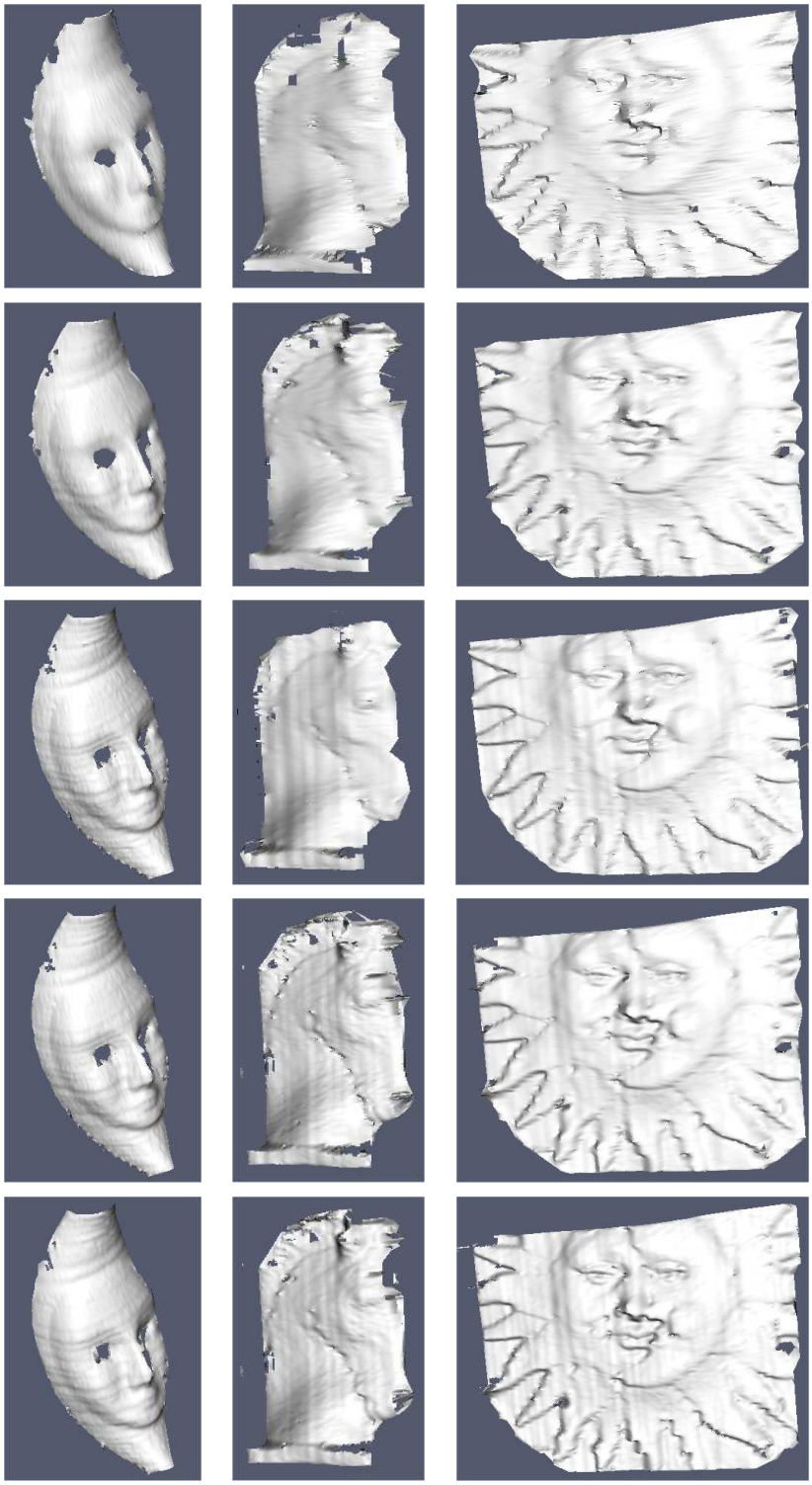

Z_W
O_W
$\{W\}$
X_W
Y_W
World coordinate system
Reconstruction Area
$\{I\}$
O_I
X_I
Y_I
Y_R
X_R
$\{R\}$
O_R
Y_C
X_C
Z_C
O_C
$\{C\}$
Camera coordinate system
Y_P
Z_P
O_P
$\{P\}$
X_P
Projector coordinate system


260
240
220
200
180
160
140
120
40
50
60
70
80
90
100
110
120
130
M channel perceived
Projected intensity
H=1
H=2
H=3
H=4


System calibration
RGB channel alignment calibration
Geometric calibration
Colour selection and calibration
$^{g}\mathbf{H}_r$, $^{g}\mathbf{H}_b$
K_1, $\mathbf{P}_c$, $\mathbf{P}_p$
Intensities $\mathbf{A}_M$, $\mathbf{A}_m$
One-shot acquisition
RGB alignment
Image rectification
Colour rectification
Radial distortion removal
Peaks and edges detection
stripes decoding
3D reconstruction


a)
b)
c)
d)
e)
f)
g)

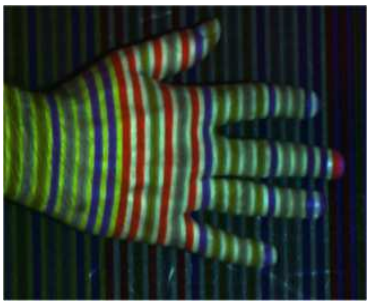




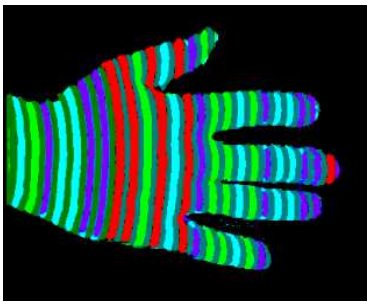
a) b) c)

d) e) f)

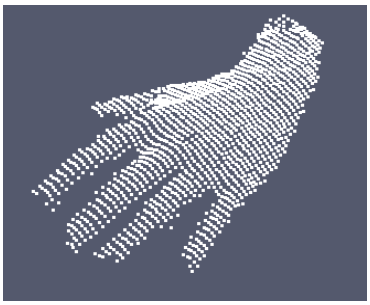
a)



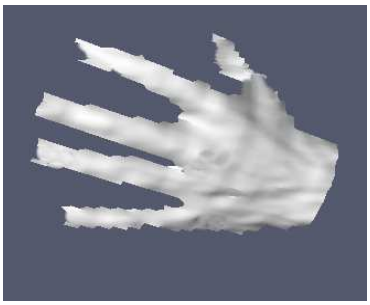
b)



c)

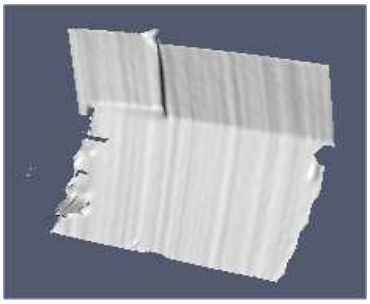


d)



e)

f)



Tables

Pattern type	Strategy	Mean (mm)	StDev (mm)	Reconstructed points
peak-based	M	0.30	0.22	1964
edge-based	E	0.63	0.37	3938
hybrid	E	0.51	0.34	3886
hybrid	M	0.43	0.33	1943
hybrid	m	0.52	0.41	1943
hybrid	M+m	0.43	0.34	3886
hybrid	E+M	0.48	0.34	5829
hybrid	E+M+m	0.46	0.44	7973