Determining Both Surface Position and Orientation in Structured-Light-Based Sensing

Zhan Song and Chi-Kit Ronald Chung, *Senior Member*, *IEEE*

Abstract—Position and orientation profiles are two principal descriptions of shape in space. We describe how a structured light system, coupled with the illumination of a pseudorandom pattern and a suitable choice of feature points, can allow not only the position but also the orientation of individual surface elements to be determined independently. Unlike traditional designs which use the centroids of the illuminated pattern elements as the feature points, the proposed design uses the grid points between the pattern elements instead. The grid points have the essences that their positions in the image data are inert to the effect of perspective distortion, their individual extractions are not directly dependent on one another, and the grid points possess strong symmetry that can be exploited for their precise localization in the image data. Most importantly, the grid lines of the illuminated pattern that form the grid points can aid in determining surface normals. In this paper, we describe how each of the grid points can be labeled with a unique color code, what symmetry they possess and how the symmetry can be exploited for their precise localization at subpixel accuracy in the image data, and how 3D orientation in addition to 3D position can be determined at each of them. Both the position and orientation profiles can be determined with only a single pattern illumination and a single image capture.

Index Terms—Structured light system, 3D reconstruction, surface orientation, feature detection.

1 Introduction

Orientation information is complementary to position information in describing shape or interpolating surface; while position describes the location of a surface point, orientation—often expressed as surface normal—describes how location changes locally on the associated surface, or equivalently speaking, how the neighborhood of the surface point is oriented in space. It has been shown that orientation information can aid in improving the precision of describing 3D geometry [1], in reconstructing object shape [2], and even in object recognition [3].

Indeed, certain orientation description can be inferred from position information by relating adjacent position data. However, such description is merely a certain abstraction of the original position data, not independent information; the result (of such orientation inference) is dependent on how dense the position data are and what method (or, more precisely, what shape prior, say a planar one or a quadric one) is used in the inference. Different density of the position data, or a different method used (upon even exactly the same position data), could generate different orientation descriptions.

In the literature, various vision cues, like shading [4] and texture [5], have been proposed to determine surface orientation. However, their application requires certain properties of the object to be present, like constant albedo or uniform distribution of surface texture, etc., and that limits the set of objects they can operate upon. This work addresses how orientation can be determined in structured light sensing, *independently* of position.

Structured Light System (SLS) [6] is an effective and low-cost means of measuring 3D information. By illuminating artificial pattern on the target object, it avoids the demand of natural texture on the object surface. By embedding unique coding in the illuminated pattern, it can make each pattern position distinct so that correspondences can be established readily between the display panel of the projector and the image plane of the camera, and 3D reconstruction be achieved.

On the illumination's coding, there are at least the temporal and spatial schemes. The temporal coding scheme (with gray-coded illumination as an example) makes use of a sequence of structured light projections and image captures to generate for each illuminated position a distinct code that is accumulated over time. In contrast, the spatial coding scheme (with pseudorandom color grid projection as an example) encodes each illuminated position by the illumination profile of the position's neighborhood. Each scheme has its own pros and cons. While the temporal scheme involves no use of spatial neighborhood and is thus less affected by the presence of occlusions in the scene, the spatial scheme could achieve 3D determination with a single image, and thus can tackle dynamic scenes better. In any case, in either scheme, a certain compromise needs to be struck on the use of colors in the illumination. While the use of monochrome light generally incurs better 3D determination quality, the use of more colors in the illumination can reduce the number of required images (in the case of temporal coding) or the size of the coding neighborhood (in the case of spatial coding), at the expense of possibly more disturbance from the measured object's original color profile.

- *Z. Song is with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Rm 301, CIE Lab, Research Building I, No. 1068, Xueyuan Avenue, University Town, Xili, Nanshan District, ShenZhen, Guangdong 518055, China. E-mail: zhan.song@sub.siat.ac.cn.*
- *C.-K.R. Chung is with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Rm 213, Shatin, Hong Kong. E-mail: rchung@cuhk.edu.hk.*

Manuscript received 1 Oct. 2008; revised 23 Apr. 2009; accepted 29 July 2009; published online 18 Nov. 2009.
Recommended for acceptance by R. Ambasamudram.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2008-10-0660.
Digital Object Identifier no. 10.1109/TPAMI.2009.192.

The above issues of SLS have all been addressed in the literature, and this work is not to study them. Existing SLSs are, however, largely restricted to determining position information only. This work seeks to explore how SLS in the case where the illumination is spatially coded can also determine orientation information.

To determine new information, we need a new set of features to encode, localize, and process. Previous works [7], [8], [9], [10], [11], [12], [13], [14], [15] on SLS generally use the centroids of the illuminated pattern elements as the feature points. The use of centroid features, however, has the following drawbacks: Being features related to the region spread of a pattern element, centroids generally have their positions in the image sensitive to surface albedo of the object surface, uneven ambient lighting, image noise, and other disturbances on image intensity distribution. Perspective distortion could also distort the centroid positions. In addition, the computation of centroids requires that the boundary of each pattern element be distinguished in the first place. As the segmentations of neighboring pattern elements are not independent, neither are the localizations of individual centroids. All of these limit how precisely the centroid feature points can be located.

More importantly, purely positional features like centroids could not supply orientation information by each of them alone, irrespective of how accurately they are located. The reason is, each single position datum itself does not offer constraint on the surface normal at the position. To also determine orientation, we need features that not only reveal where the surface points are in space, but also constrain how the neighborhood of the surface points are oriented.

In this work, we propose using *grid points* between the pattern elements as the features. Grid points are points formed by the intersection of the grid lines of the illuminated pattern. Irrespective of the perspective distortion, their locations are always where the grid lines intersect. In addition, compared to neighboring pattern element centroids, neighboring intersections of grid lines are not as intimately related. Grid points are thus features generally more localizable than pattern element centroids.

In addition, grid points possess a certain symmetry, termed *twofold rotation symmetry*, that is quasi-invariant with image noise, image blur, and perspective projection. Such a symmetry can be exploited for their precise localization in the image domain.

More importantly, each grid point carries not only information related to a position, but also the local curvatures of the grid lines that compose the position. In this paper, we present a mechanism of deducing from such local curvatures in the illumination and in the image data the local surface normal of the imaged surface in space.

Notice that the orientation information we determine is not meant to replace position information, but as supplementary data in describing or interpolating shapes. Notice also that this work does not address how independently obtained position and orientation data should be combined for surface description or interpolation. The subject warrants separate study on its own; in fact, there is already a rich collection of works in the literature on it. This work focuses on the possibility of, in SLS, obtaining orientation information that is not inferred from position.

This paper is organized as follows: In Section 2, previous work on position and orientation recovery is briefly reviewed. The illumination pattern and the grid point features we use are described in Section 3. Like the centroid features, the grid point features must each be coded uniquely in the illumination, or else the essence of SLS is destroyed. In Section 3, we also describe how each grid point can be labeled uniquely. In Section 4, we present the symmetry that the grid points possess and how the symmetry can be used for their localization in the image domain in subpixel accuracy. The mechanism of determining surface orientation at each grid point is presented in Section 5. In Section 6, experimental results on parametric surfaces and free-form surfaces, including a human face in natural appearance, are shown. Conclusion and possible future work are offered in Section 7.

2 Previous Work

SLS has been extensively researched and increasingly used in industry. Below, we offer a brief review of some of the key works related to SLS.

2.1 Position Determination via Structured Light

As discussed earlier, the key to the success of SLS is how unique code can be attributed to each position of the illuminated pattern. On this, there are the temporal [16], [17], [18], [19], [20], [21], [22], [23], [24] and spatial [7], [8], [9], [10], [11], [12], [13], [14], [15], [25], [26], [27], [28], [29], [30] coding schemes. The temporal coding scheme can achieve denser data samples and higher accuracy in the measurement, but at the expense of requiring multiple illuminations and image captures over time. Examples of temporal coding methods are Gray code [18], Gray code combined with sine wave pattern [21], phase shifting [22], line shifting [23], and strip shifting methods [24]. In contrast, the spatial coding scheme labels each pattern position by the appearance profile of the neighboring positions. The appearance profile can be about various gray levels, colors [26], or geometric primitives [27]. De-Bruijn sequences [29], [30], pseudorandom arrays, and M-arrays [7], [8], [9], [10], [11], [12], [13], [14], [15] are the coding methods often employed. The spatial coding scheme has the advantage that 3D determination can be achieved with a single illumination and a single image capture. It is therefore particularly suitable for use in dynamic applications.

As this work assumes the use of a spatial coding scheme in the SLS, below, we offer more details of previous works related to it.

Earlier works that structure the illumination as stripes and use spatial coding to encode each of them include [28], [29], [30]. Their contributions are on either of the following: the use of colors in the stripes, careful selection of the colors on neighboring stripes to reduce color confusion in the image data, and the integrated use of region segmentation and edge detection to segment the stripes. In particular, in [28], a window size smaller than what is necessary for unique spatial coding is used, and dynamic programming is employed to resolve the correspondence ambiguity arisen from the nonunique codes of the stripes. As for the accuracy of position determination, Zhang et al. [28] report 0.18 mm,

Chen et al. [29] report the range of 0.15-0.32 mm, and Pages et al. [30] report 0.3-0.6 mm.

There were also works like [9], [13], [26] that structure the illumination as an array of pattern elements not stripes, and assign colors to the pattern elements with the objective that each pattern element has a neighborhood of colors that are unique. In all of these works, the centroids of the pattern elements (being either circular or rectangular) are used as the feature points. In [26], a spatial resolution of around 5 mm for a surface 1 m away is attained. In [9], the position determination error is within 2 mm for an object located about 1.5 m from the sensor. In [13], a relative error of 2 percent (of the working distance) is reported.

Because of possible color confusion in the image data, it is desired that not only is each pattern element's color window unique, adjacent color windows are also distant. There were works like [14] that address how color assignment to the illumination can ensure that the hamming distance of adjacent color windows can be kept far. This is usually achieved at the expense of using a window size bigger than necessary.

2.2 Orientation Determination

As discussed, orientation information independently acquired is an important supplement to position information. A number of vision cues have been proposed to determine orientation. Established ones include shape from shading [4], texture [5], specularity [31], polarization [32], and photometric stereo [33]. Such methods, however, require certain restrictive conditions, like constant surface albedo or uniform texture distribution, etc.

There have been a few attempts in determining orientation from structured illumination. In the work of Shrikhande and Stockman [34], a grid pattern is projected onto the target object. Surface orientations at the grid points (the intersections of the grid lines) are inferred from the change of lengths of the grid edges in the image data. The presented experimental results, however, indicate that there could be large errors at some places of the image, specifically those where the surface orientation is quite different from that of the reference plane, or where the surface curvature is nonzero. Winkelbach and Wahl [35] used a strip pattern to compute the surface normal at each edge point. A fringe pattern is projected onto the object surface twice from two different angles, each time with a separate imaging step, so that two surface tangents are available for some image positions. In [36], a simplified method of determining surface normals from the slopes and intervals of the stripes in the image is also proposed. The method is based upon the assumption that the surface patch between two strip edges is planar or very smooth. With that, the tilt angle can be estimated from the deformed width of the strip by comparing it with the strip width measured for a reference plane. In the system, pattern illumination is assumed to be a parallel projection, and imaging a parallel imaging as well. In other words, the intrinsic parameters of both the camera and projector are not considered, and errors due to the oversimplified projection and imaging models are inevitable.

In the work by Davies and Nixon [15], a hexagonally tessellated array of circular spots is used as the illumination pattern. Coding is based upon the neighborhood color profile of each circular spot using Griffin's method [10]. It is assumed that each circular pattern spot, upon illumination of the object surface, reflection from it, and imaging by the camera, will appear as an ellipse in the image. With the center points of such ellipses regarded as the feature points, a method is proposed to deduce the surface normal in space associated with each of such elliptical centers, which is by examining the shear and scaling factors of the ellipse in the direction of the epipolar line.

However, the elliptical model for the pattern element in the image implicitly assumes that the object surface is planar not only in the immediate vicinity of the center point of the ellipse, but in the entire pattern element. Such a model is hardly justified for an object surface with substantial curvature. In addition, ambiguity could arise if an epipolar line happens to pass through the centers of multiple pattern elements of the same color. Also, very few experimental results were reported; only the result on a human face painted white using theatrical makeup is shown. There was no evaluation of how accurately the surface normals and positions are determined either.

In contrast with the previous works, this work does not assume planarity of the object surface over the entirety of each pattern element; we only determine surface normals at points—the grid points—from the local curvatures of the grid lines in the image that form the grid points. There is also no oversimplified approximation of the illumination and imaging processes; we use the perspective model to describe both processes.

3 Design of Our System

Our SLS has an illuminated pattern designed after the principle of pseudorandom array [37]. Below, we first describe how colors are assigned to each pattern position so that each position can have a neighborhood window of unique color profile. If centroids of the pattern elements are used as the feature points, the coding task is achieved. However, the feature points we use—the grid points—are each attached not to a single pattern element, but to a number of them. In fact, each grid point is right at the intersection of a number of pattern elements. We therefore need a separate mechanism of encoding each grid point uniquely. In this section, we also offer the description of the labeling scheme.

3.1 Illumination Design

A pseudorandom color pattern made of rhombic elements, as shown in Fig. 1, is used in our SLS. It is constructed in the following way. A primitive polynomial $h(x)$, which is defined over *Galois Field* with four elements ($GF(4)$), is first used to generate the pseudorandom sequence:

$$h(x) = 2x^6 + 2x^5 + x^4 + 3x^3 + 2x^2 + 2x + 1. \quad (1)$$

From the above primitive polynomial, a pseudorandom sequence of length 4,095 is generated. Then, a pseudorandom array of size 65×63 can be produced by folding the sequence [37]. By the window property of the pattern, every window of size 2×3 in the pattern is unique upon the elements the window is composed of. It means that every window of size 2×3 appears only once in the whole

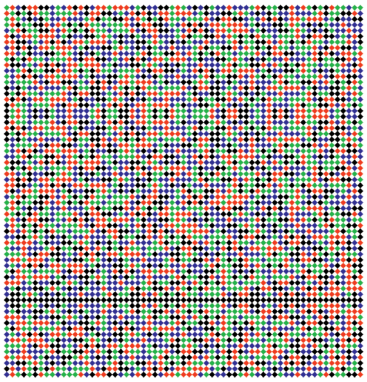

Fig. 1. A 2D pseudorandom color pattern of 65×63 elements that have red, green, blue, or black colors for the pattern elements (foreground), and white color for the background.

pattern. Since the pseudorandom array is constructed over $GF(4)$, there have to be, in total, four different elements in the pattern and, on this, we use four different colors (red, green, blue, and black) which have sharp contrast between them. The background of the pattern is set to white.

The use of rhombic elements in our design has the advantage that adjacent pattern elements (i.e., elements in the foreground) do not share a boundary, and thus color confusion in the image between adjacent pattern elements is avoided. However, the advantage comes with a price. The number of pattern elements is halved for the same pattern element size. If the centroids of pattern elements are used as the feature points, as practiced classically, the density of feature points would be compromised. One might argue that the issue can be resolved by also adding the centroids of the background elements to the use, but then that would mean treating background elements like pattern elements, and the original purpose of rhombic element design is defeated. There would be the issue of color confusion in the image data, and the extraction of the centroid of a foreground element is not independent of that of the nearby background.

3.2 Feature Choice and Coding

While, in traditional designs, the centroids of the pattern elements are used as the feature points, here we use the grid points between neighboring rhombic elements as the feature points.

As is apparent from Fig. 1, the use of grid points has the immediate advantage that the density of feature points is retained despite the use of rhombic pattern elements.

However, like any other features used in SLS, each grid point has to be attributed with a unique code in the first place. Below, we show how that can be achieved.

As illustrated by Fig. 2, the grid points can be classified into two types: P_1 and P_2, according to whether the pattern elements on the immediate left and right of the point are white (background) or colored (foreground) pattern elements. A grid point of P_1 type can be encoded by the color profile of the 2×3 rhombic elements surrounding it, in a fixed order, say the order of $c1$-$c2$-$c3$-$c4$-$c5$-$c6$ in Fig. 2. As for a grid point of P_2 type, we search for the grid point of P_1 type nearest it in the bottom-right direction and the code word of that grid point is adopted as its code word. This is illustrated in Fig. 3.

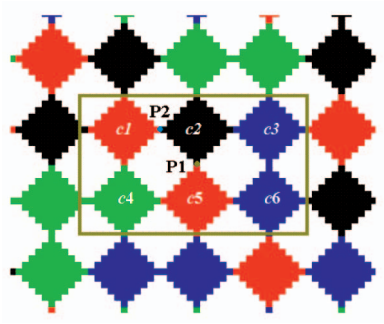

Fig. 2. The code word window of a grid point which can be of either P1 type or P2 type.

For convenience, the projector-camera system in our design is configured so that the x-axes of the two observation planes (i.e., the projector's display panel and the camera's image plane) and the baseline between the camera and projector are approximately parallel. This allows the descriptions "left" and "right" to be readily defined for any position (x, y) on the two observation planes, "left" referring to those positions with the x-coordinate being slightly smaller than x and "right" those positions with the x-coordinate being slightly larger than x. In our system, such "left" and "right" descriptions are used only for identifying the immediate neighboring pattern elements of a grid point, for the purpose of classifying the grid point type. They are not for identifying pixel-level features but features as large as the pattern elements, and thus, small errors in the above configuration do not matter.

For features as large as the rhombic elements, the topology relation as being on the "left" or "right" of a grid point in the projector's display panel is quasi-preserved in the image data. With this, the grid point type ($P1$ or $P2$) and the code word together uniquely distinguish a grid point and allow easy correspondence of the grid points between the projector side and the camera side.

By this new feature definition, for an illumination of 65×63 resolution, the number of encoded feature points can reach 7,808 as opposed to 3,904 when centroid features are used. However, we must emphasize that the use of grid points and the use of centroid features are not mutually exclusive; if the centroid features are also included and encoded, the feature points' density can be further improved.

4 Feature Extraction

The above shows that, like centroid features, grid points can also be attributed with unique labels if a code word

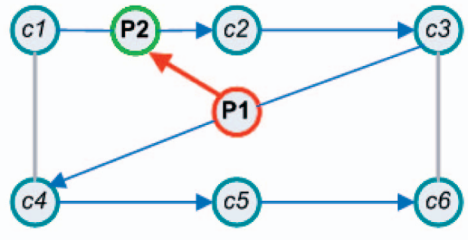

Fig. 3. Code word construction for grid points of type P1 and type P2, respectively.

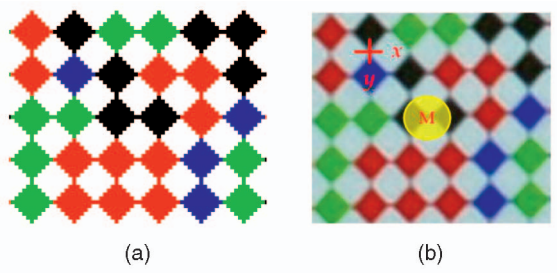

Fig. 4. An illumination pattern that is made of rhombic elements: (a) the original illumination in which adjacent pattern elements touch upon each another over their corners; (b) the image of the projection in which the corner point connection can be compromised by noise and distortion from the sensor system and the surface property of the imaged object.

from the pseudorandom pattern is augmented with grid point type description. In addition, they have the advantage that they are invariant with illumination variations and image blur and perspective distortion (that could impact the localization of region-based features like centroids), for the simple reason that they are not global features of surface patches but points at the intersections of grid lines. In principle, they can be extracted with much higher precision than the centroid features. Below, we describe their extraction in our system.

Due to influence of the various optical characteristics (Modulation Transfer Function (MTF), depth of field, sensor noise, etc.) of the projector-camera system, and of the target object's surface property (texture, reflectance, curvature, etc.), adjacent rhombic elements of the illuminated pattern can be disconnected in the image. An example is shown in Fig. 4b. It means it is not trivial to pinpoint where the grid points are in the image data. Traditional feature detection algorithms such as LoG [38], Harris [39], or SUSAN [40] cannot deal with this issue. Some corner extractors have been proposed in the literature for use in camera system calibration that involves the use of a reference planar grid pattern in 3D. In our case, since the true object surface under consideration can be a curved surface, such corner extractors, which often assume ideal planarity of the pattern, are not applicable either. To locate the grid point or corner features precisely in the image data, a new mechanism is needed.

Fig. 4 also illustrates a certain feature of grid points that could be positive for their localization. This is that they display a certain local symmetry. Suppose the color pattern is abstracted to be a monochrome one, with the color elements viewed as the foreground elements and the rest the background. With this, a small circular disc of intensities centered at any grid point resembles itself when rotated by 180 degrees. In the study of plane symmetry, this is referred to as the *twofold rotation* symmetry or *cmm* symmetry. Such a symmetry on the illumination side is quasi-invariant with image noise, image blur, and projective distortion, and thus, largely preserved in the image data for the reason that the colinearity of the segments that divide the circular disc into two symmetrical halves is preserved under the processes. If the circular disc is sufficiently small, even nonzero curvature of the illuminated surface in 3D has little effect to the symmetry. This quasi-invariance of the symmetry is what we exploit for precise grid point detection from the image.

4.1 Hypothesis of Grid Point Positions

A grid point in the rhombic pattern we use is characterized by the presence of neighboring regions of drastically different intensities in the vertical (y) and horizontal (x) directions. More precisely, in the vertical directions, there are the colored pattern elements, and in the horizontal directions, there are the background (white) regions. We use this property to hypothesize candidate positions of grid points in the image. We first convert the original image to a monochrome one $I(x, y)$, as described above. A mask in the shape of a cross, as shown in Fig. 4b, is then used for convolution with all of the same-size windows of $I(x, y)$. The absolute difference d, as expressed below, is defined as the response value of the mask at any image position (x, y):

$$d = \left| \sum_{i=-\varepsilon}^{\varepsilon} I(x+i, y) - \sum_{j=-\varepsilon}^{\varepsilon} I(x, y+j) \right|, \tag{2}$$

where ε indicates the size of the mask. The measures of d indicate how different the image intensity accumulations along the x and y directions, respectively, are at the position. The image positions with high response values are picked as candidates of grid points. Such a template-based scheme is simple and efficient. However, it merely considers differentials of raw intensities and that compromises its accuracy and robustness against image noise. It is used here only as a preliminary step to hypothesize possible grid point positions for speeding up the process.

4.2 Confirmation and Position Refinement

Twofold rotation symmetry is displayed at the positions of true grid points. This can be used for confirmation of the grid point features as well as for their precise localization in subpixel accuracy. To measure the strength of the twofold symmetry at any candidate positions, we use the correlation coefficient between a small circular window of the position and its 180-degree rotation. Since image intensities are, in general, normally distributed, the Pearson's Product-Moment Correlation Coefficient (or PMCC) [41] is adopted in our method.

Suppose M_C is a circular mask centered at any candidate image position C and $M_C{}'$ the mask created by rotating M_C by 180 degrees around C. By the definition of PMCC, we can write the correlation coefficient ρ_C as:

$$\rho_C = \frac{\sum_{i=1}^{n}(M_{Ci} - \overline{M}_C)(M_{Ci}{}' - \overline{M}_C{}')}{\sqrt{\sum_{i=1}^{n}(M_{Ci} - \overline{M}_C)^2 \sum_{i=1}^{n}(M_{Ci}{}' - \overline{M}_C{}')^2}}, \tag{3}$$

where i refers to an individual element of the mask M_C or $M_C{}'$, n is the size of the masks, M_{Ci}, $M_{Ci}{}'$ indicate image intensities of the ith elements of the M_C and $M_C{}'$ masks, respectively, and $\bar{M}_C, \bar{M}_C{}'$ are the mean values of intensity distributions of the two masks. The above equation can be replaced by an equivalent formula which avoids the use of the mean values of the masks, and is therefore faster:

$$\rho_C = \frac{n\sum_{i=1}^{n} M_{Ci}M_{Ci}{}' - \sum_{i=1}^{n} M_{Ci}\sum_{i=1}^{n} M_{Ci}{}'}{\sqrt{n\sum_{i=1}^{n} M_{Ci}^2 - \left(\sum_{i=1}^{n} M_{Ci}\right)^2}\sqrt{n\sum_{i=1}^{n} M_{Ci}'^2 - \left(\sum_{i=1}^{n} M_{Ci}{}'\right)^2}}. \tag{4}$$

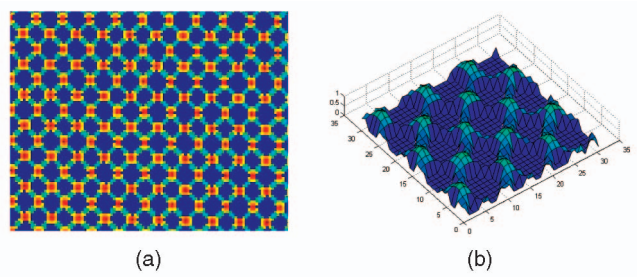

Fig. 5. Response values of the proposed twofold symmetry detector at various positions of an image.

The operator is only applied to image positions C that pass the above hypothesis step. The response values at these points are calculated, and a map named the response image ρ can be generated from the values. An example response image is shown in Fig. 5.

A threshold t is used on the ρ-map to extract image positions, where the ρ value is a sufficiently strong local maximum. A small region of size, say 3×3 pixels, is then selected around every such position. The final grid position (x_g, y_g) in the vicinity of every such point C is then computed in subpixel accuracy as the weighted average of all positions in this 3×3 region, with the weight being the ρ value ρ_i of each ith position (x_{Ci}, y_{Ci}) in the region:

$$(x_g, y_g) = \sum_{i=1}^{9} \rho_i \cdot (x_{Ci}, y_{Ci}) \Big/ \sum_{i=1}^{9} \rho_i. \tag{5}$$

Since structure rather than raw image intensity at the grid point is used, the described method has higher robustness than intensity-based methods against image noise, blur, surface texture, curvature, and projective distortions. More details about the grid-point detector, such as the robustness test, comparison with traditional operators, can be found in a separate paper [42]. As the extractor is based upon local symmetry of each grid point, the extractions of grid points are independent of one another. This is an additional advantage over the use of centroid points as the feature points.

5 Surface Orientation Determination

Once the feature points are located in the image data, correspondences over them between the image plane (of the camera) and the display panel (of the projector) can be readily made by the unique labeling of the points, and hence, 3D reconstruction can be achieved. On position determination, the use of grid points is no different from that of other feature points.

The question is, how can we go beyond the traditional goal of structured light system, determining not only position but also orientation. As said, orientation [1], [2], [3], [4], [5], [15], [32], [33], [34], [35], [36] is an indispensable piece of information that can aid much in describing 3D shape. Below, we describe a mechanism of deriving the surface normal at each grid point, using the grid lines that form it.

5.1 The Underlying Principle

Consider any grid point $\mathbf{p}_p$ in the display panel of the projector, and the two accompanying grid lines that compose it, as illustrated by Fig. 6. Suppose the grid point and grid lines induce an image grid point $\mathbf{p}_c$ and two image grid lines (or more correctly, grid curves, as the original grid lines are generally modulated by the curvature of the surface in 3D and appear not as lines in the image data) on the image plane via a 3D point $\mathbf{P}$ on the object surface in space. Suppose the tangents to the grid lines right at the point $\mathbf{p}_p$ in the display panel are $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$, respectively, and the corresponding tangents to the grid lines right at the point $\mathbf{p}_c$ in the image plane are $\mathbf{t}_{c1}$ and $\mathbf{t}_{c2}$, respectively.

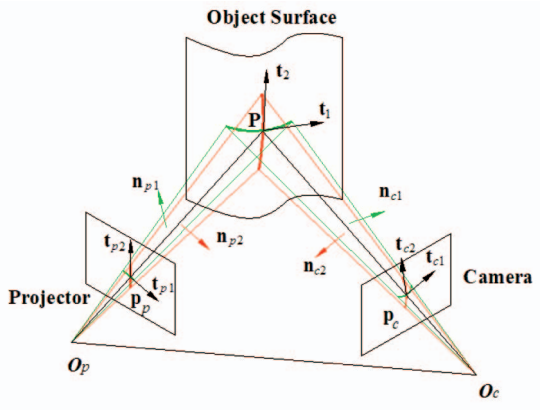

Fig. 6. The geometry behind the determination of surface orientation at a grid point from the component grid lines.

Tangent $\mathbf{t}_{p1}$ and grid point $\mathbf{p}_p$ (and the projection center of the projector) in the display panel together form a plane $\Pi(\mathbf{p}_p, \mathbf{p}_{p1})$ of illumination from the light source. Such a plane of light is reflected by the object surface at point $\mathbf{P}$ and becomes the plane of projection $\Pi(\mathbf{p}_c, \mathbf{t}_{c1})$ to the image plane. The intersection of the two light planes $\Pi(\mathbf{p}_p, \mathbf{t}_{p1})$ and $\Pi(\mathbf{p}_c, \mathbf{t}_{c1})$ actually defines a tangent $\mathbf{t}_1$ in 3D to the object surface at point $\mathbf{P}$. $\mathbf{p}_p$ and $\mathbf{t}_{p1}$ are fully accessible as they are entities under system design, and so are $\mathbf{p}_c$ and $\mathbf{t}_{c1}$ as they are entities observable from the image data. Thus, the two light planes $\Pi(\mathbf{p}_p, \mathbf{t}_{p1})$ and $\Pi(\mathbf{p}_c, \mathbf{t}_{c1})$ are both constructible, and their intersection $\mathbf{t}_1$ can be determined. In fact, the tangent $\mathbf{t}_1$ to the object surface at point $\mathbf{P}$ is merely the cross-product of the surface normals of the two light planes. Similarly, another tangent $\mathbf{t}_2$ at surface point $\mathbf{P}$ can be determined as the cross-product of the surface normals of two other light planes: $\Pi(\mathbf{p}_p, \mathbf{t}_{p2})$ and $\Pi(\mathbf{p}_c, \mathbf{t}_{c2})$, which are both accessible from design and image observations.

In other words, by simply grabbing one image of the object surface that is under projection of a proper pattern, for any imaged grid point $\mathbf{p}_p$ at position (x, y) on the image plane, it is possible to determine the surface orientation $\mathbf{n}(x, y)$ of the object surface at the associated 3D point $\mathbf{P}$. If the image observations are $\{\mathbf{p}_c, \mathbf{t}_{c1}, \mathbf{t}_{c2}\}$ and the pattern data are $\{\mathbf{p}_p, \mathbf{t}_{p1}, \mathbf{t}_{p2}\}$, the 3D tangents $\{\mathbf{t}_1, \mathbf{t}_2\}$ can be determined, and the surface normal is simply $\mathbf{n}(x, y) = \mathbf{t}_1 \times \mathbf{t}_2$.

Note that if the projected pattern is a regular grid, image tangents $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ on the projector side are the same at all grid points, and known.

5.2 Image Tangent Detection

Here, we describe how we extract the image tangents to the observed grid lines at a grid point. Suppose we are to determine image tangent $\mathbf{t}_{c1}$ at a grid point A. Fig. 7 displays $\mathbf{t}_{c1}$ as two short line segments. Two approaches can be

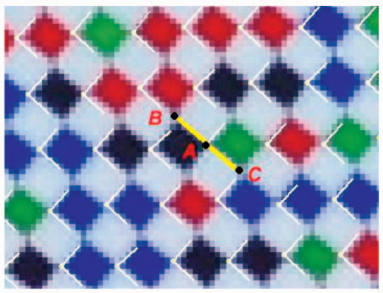

Fig. 7. An image tangent at any grid point A of one grid point type (type P_1 in the figure) can be determined from two adjacent grid points B and C of the other grid point type (type P_2 in the figure).

adopted. In one approach, the grid points adjacent to A are employed. As shown in Fig. 7, along the associated grid line through A, there are two adjacent grid points B and C. The three grid points are distinguished by the fact that while point A is of one grid point type (type $P1$ in Fig. 7), the adjacent grid points are of the other type (type $P2$ in Fig. 7).

The image tangent $\mathbf{t}_{c1}$ at point $\mathbf{A}$ can be determined from the three grid points by simple finite differencing and averaging as a good approximation. The second approach uses the image intensities instead. A small window around point A is selected, and the window of intensities is used to determine two directions about A such that the intensity change orthogonal to the directions is locally maximum. The two directions are allocated to $\mathbf{t}_{c1}$ and $\mathbf{t}_{c2}$, respectively. The first approach is less sensitive to image noise and surface texture, but the second approach requires no assumption of smoothness of the 3D shape. In our implementation, both approaches are used. For grid points where the local density of detected grid points is high, we use the first approach; otherwise, we use the second approach.

5.3 From Image Tangents to Surface Orientation

Suppose $\mathbf{n}_{c1}$ and $\mathbf{n}_{c2}$ are the normal vectors of the light planes $\Pi(\mathbf{p}_c, \mathbf{t}_{c1})$ and $\Pi(\mathbf{p}_c, \mathbf{t}_{c2})$ on the camera side, and $\mathbf{n}_{p1}$ and $\mathbf{n}_{p2}$ those of the light planes $\Pi(\mathbf{p}_p, \mathbf{t}_{p1})$ and $\Pi(\mathbf{p}_p, \mathbf{t}_{p2})$ on the projector side. Suppose the intrinsic parameters of the camera and projector (which is modeled as another perspective camera, except that light comes out of it instead of going into it) have been calibrated, which are focal lengths f_c, f_p and principal points $\mathbf{C}_c(x_{c0}, y_{c0}), \mathbf{C}_p(x_{p0}, y_{p0})$, respectively. Then, the four plane normals $\mathbf{n}_{c1}$, $\mathbf{n}_{c2}$, $\mathbf{t}_{p2}$, $\mathbf{n}_{p2}$ can be determined as:

$$\vec{\mathbf{n}}_{ci} \cong \mathbf{p}_c \times \begin{bmatrix} \mathbf{t}_{ci} \\ 0 \end{bmatrix} \quad \vec{\mathbf{n}}_{pi} \cong \mathbf{p}_p \times \begin{bmatrix} \mathbf{t}_{pi} \\ 0 \end{bmatrix}, \tag{6}$$

where $\cong$ represents equality up to arbitrary nonzero scaling,

$$i = 1, 2, \mathbf{p}_c = [(x_c - x_{c0}), (y_c - y_{c0}), f_c]^{\mathrm{T}},$$
$$\mathbf{p}_p = [(x_p - x_{p0}), (y_p - y_{p0}), f_p]^{\mathrm{T}},$$

with (x_c, y_c), (x_p, y_p) being the image positions of the grid points on the camera side and on the projector side, respectively.

If the planar orientations of the image tangents $\mathbf{t}_{ci}$ and $\mathbf{t}_{pi}$ are θ_{ci} and θ_{pi}, respectively $(i = 1, 2)$, the above plane normals of the incident light planes can be expressed as:

$$\mathbf{n}_{ci} \cong \begin{bmatrix} x_c - x_{c0} \\ y_c - y_{c0} \\ f_c \end{bmatrix} \times \begin{bmatrix} \cos\theta_{ci} \\ \sin\theta_{ci} \\ 0 \end{bmatrix},$$
$$\mathbf{n}_{pi} \cong \begin{bmatrix} x_p - x_{p0} \\ y_p - y_{p0} \\ f_p \end{bmatrix} \times \begin{bmatrix} \cos\theta_{pi} \\ \sin\theta_{pi} \\ 0 \end{bmatrix}, \tag{7}$$

which can be simplified to

$$\begin{aligned} \mathbf{n}_{pi} \cong [&-f_c \sin\theta_{ci}, f_c \cos\theta_{ci}, (x_c - x_{c0}) \sin\theta_{ci} \\ &- (y_c - y_{c0}) \cos\theta_{ci}]^{\mathrm{T}}, \mathbf{n}_{pi} \cong [-f_p \sin\theta_{pi}, f_p \cos\theta_{pi}, \\ &(x_p - x_{p0}) \sin\theta_{pi} - (y_p - y_{p0}) \cos\theta_{pi}]^{\mathrm{T}}. \end{aligned} \tag{8}$$

The two 3D tangents $\mathbf{t}_1, \mathbf{t}_2$ to the target surface at point $\mathbf{P}$ with reference to the camera coordinate system are then:

$$\mathbf{t}_1 = \mathbf{n}_{c1} \times (\mathbf{R}\mathbf{n}_{p1}), \quad \mathbf{t}_2 = \mathbf{n}_{c2} \times (\mathbf{R}\mathbf{n}_{p2}), \tag{9}$$

where $\mathbf{R}$ represents the rotational relationship between the camera and projector coordinate frames.

Finally, the surface orientation at point $\mathbf{P}$ can be determined as:

$$\mathbf{n}(x, y) = \mathbf{t}_1 \times \mathbf{t}_2. \tag{10}$$

Notice that the determination of local surface orientation as expressed by (10) is a deterministic process that requires only image information local to the specific point to operate. Unlike shape from shading and shape from texture, it does not require assumption about how local surface orientations at neighboring points are related. More specifically, it requires no process of iterations to determine local surface orientations.

The above analysis shows that surface orientation in 3D can indeed be determined from image tangents, but it also points out that for each imaged grid point $\mathbf{p}_c = (x_c, y_c)$ and the accompanying image tangents $\mathbf{t}_{c1}$ and $\mathbf{t}_{c2}$, it also requires the 2D tangents $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ and other local information in the projector's display panel for precise determination of orientation. However, in previous works such as [34], [35], [36], parallel projection is assumed on both the projector and the camera. The rather crude approximations of the illumination and imaging models inevitably bring error to the determined orientation. The error increases as the considered image position is farther away from the principal point of the observation plane.

6 Experimental Results

Experiments have been conducted to use the described system to reconstruct a variety of shapes. We determined position and orientation in isolation, and examined what quality they were of on their own; we leave to future work how they can be best integrated for improved 3D reconstruction. The experimental platform consisted of a DLP projector of XGA resolution (i.e., $1{,}024 \times 768$ pixels) and a camera of $1{,}500 \times 1{,}000$ resolution. The working

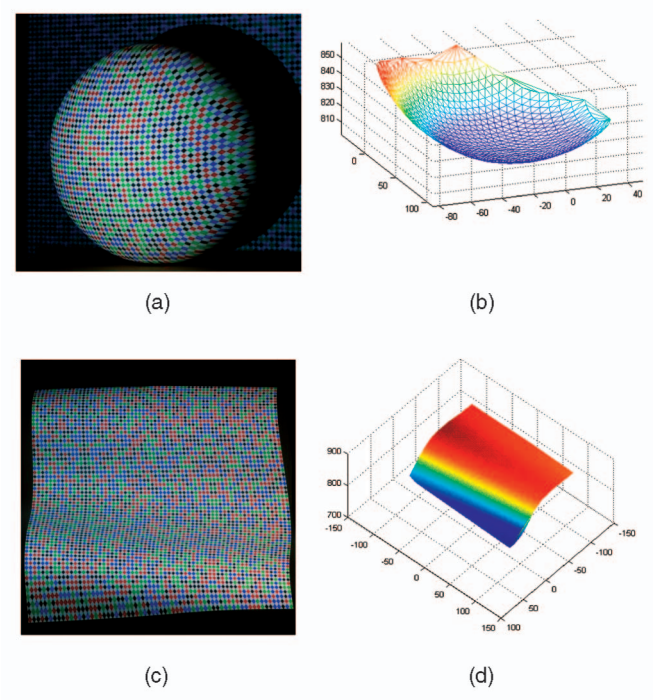

Fig. 8. Reconstruction of a sphere and a ribbon-like surface using the described system: (a) a sphere under pattern projection; (b) depth reconstruction of the spherical surface; (c) a ribbon-like surface under pattern projection; (d) depth reconstruction of the ribbon-like surface.

distance of the system was about 850 mm. The projector-camera system was first calibrated using a method we developed ourselves [43]. Once the intergeometry of the projector and camera and the intrinsic parameters of the two instruments were made known, we proceeded with the reconstruction. Below, we present a few results on reconstructing parametric shapes, like planar and spherical surfaces, and free-form shapes, like a bust object and a real human face.

6.1 Position Determination of Parametric Shapes

Experiments on parametric shapes allowed the reconstruction quality to be explicitly evaluated because the shapes were known parametrically.

In one experiment, we had the structured light projected onto a planar board to determine positions. The planarity of the board was then measured by fitting a plane in the least-square error sense to the discrete 3D points recovered. The plane fitting process had an absolute mean error of 0.131 mm (or 0.015 percent of the working distance 850 mm), and a standard deviation of fitting error of only 0.1 mm or 0.011 percent.

We had also conducted a similar experiment on a sphere of radius about 97 mm. The sphere under pattern projection is shown in Fig. 8a. The position determination result is shown in Fig. 8b. A sphere was fitted in the least-square error sense to the reconstructed data points. The fitting process had an absolute mean residue of 0.202 mm, with a standard deviation of only 0.067 mm. Shown in Figs. 8c and 8d are reconstruction results of a ribbon-like surface. The reconstruction results were of quality usable for many applications.

Table 1 lists out the 3D determination results of a number of other methods reported in the literature. Since different

TABLE 1
Performances of Various Structured-Light Systems on Plane Reconstruction

Method	*Basis of Method*	*Plane-fitting Error (as % of working distance)*
Zhang 2002, [28]	Color Strip	Std. 0.18mm (NA)
Chen 1997, [29]	Color Strip	0.15-0.32 mm (NA)
Pages 2005, [30]	Color Strip	0.3-0.6 mm (NA)
Kiyasu 1995, [7]	2D Pattern	1 mm (0.2% of 500 mm)
Fong 2005, [8]	2D Pattern	0.7 mm (0.11% of 650 mm)
Desjardins 2007, [9]	2D Pattern	2 mm (0.11% of 1.5 m)
Chen 2008, [13]	2D Pattern	about 2%
Sadlo 2005, [22]	Gray code & Phase Shifting	about 0.1- 0.5 mm (NA)
Gühring 2000, [23]	Gray code & Line Shifting	Std. 0.17-0.28 mm (NA)
Song 2008, [24]	Gray code & Strip Shifting	0.11 mm, std. 0.08mm (850mm)
Proposed Method	2D Pattern	0.11-0.13 mm or 0.015%, Std. 0.1 mm or 0.011%

experimental targets and error evaluation criteria were used in the methods, the listing is meant not for direct comparison but only for coarse reference. Nonetheless, there is indication that the proposed system is of a performance comparable to, if not better than, those of the other SLSs that are 2D-pattern-based. The result of the proposed system, which is a spatial coding one and requires only one image capture to operate, was even comparable to those of some temporal methods whose operations demand multiple image captures. The comparison was, of course, inconclusive, as a fair comparison should involve building all of these systems with the same pattern size, using the same color set, adopting the same coding window size, and having all of them calibrated with the same projector-camera calibration system. That would require a separate study. Nonetheless, the above experiments at least indicate that the proposed system has promising performance, and we attribute it to the precise localization of the feature points and the accurate calibration of the projector-and-camera setup in our system.

6.2 Orientation Determination of Parametric Shapes

We also reconstructed the surface normals of the above objects separately using the method described in Section 4. On the planar shape, the surface normals were determined over an area of a size about 100×100 mm. The recovered distribution of local surface orientations only had a small difference from those of a perfect plane—a mean discrepancy of 0.69 degree and a standard deviation of 0.13 degree. Had the parallel projection model been used instead, a much larger difference—a mean discrepancy of 1.65 degrees and a standard deviation of 0.48 degree—would have been obtained. The experiment shows that the adoption of a more accurate projection model, as is the case in our system, could incur a substantial difference. For reference, in previous works like [34], [35], [36], orientation discrepancies in the range of 2 to 8 degrees were reported.

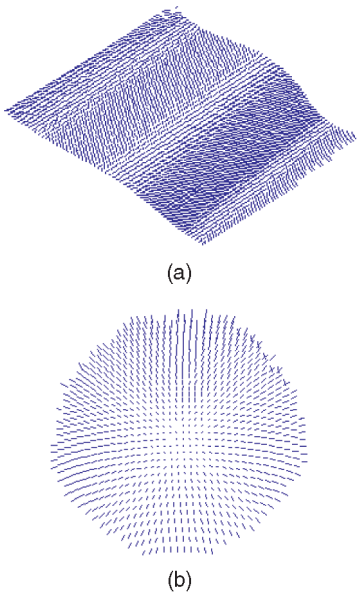

Fig. 9. Surface orientation maps recovered for (a) sphere and (b) ribbon-like surface.

The orientation maps of the above spherical object and ribbon-like surface were also determined. Fig. 9 shows the recovered orientation maps of the two shapes.

Though ground truth about the surface normals of the objects was not available, we had used some inferred references to evaluate the quality of our orientation determination. On the spherical object, we assumed that the object was an almost-perfect sphere, and identified the ideal sphere that best fitted the above determined position data in the least-squares error sense. The surface normals of such a sphere were then used as the reference. The orientation data we determined directly from the image data, not via the position data, had a difference of only 1.13 degrees, on average, from the above reference.

For comparison, we also inferred the orientation data from the aforementioned position data using a classical method, and compared such orientation abstraction of the position data with the above reference. We first generated a triangular mesh from the position data that had the reconstructed 3D positions as the corners of the triangular patches. The mesh for the spherical object is shown in Fig. 8b. At each of the patches, the local surface normal could be determined from the three corner positions of the patch. The surface normal at each grid point was then calculated from those of the surrounding triangular patches as a weighted average, with the weight for each triangular patch being in proportion to the area of the patch. Such orientation information that was abstracted from the position data however showed larger difference with the above reference—a difference of 5.41 degrees, on average.

Our comparison experiments indicated that the orientation information determined directly from the grid lines is different from and generally more accurate than that inferred from the position data.

6.3 Reconstruction of Free-Form Shapes

We also use the proposed system to determine position and orientation of some free-form shapes, like a bust object and a human face in natural appearance. The inspected surfaces under patterned illumination are shown in Figs. 10a and 11b.

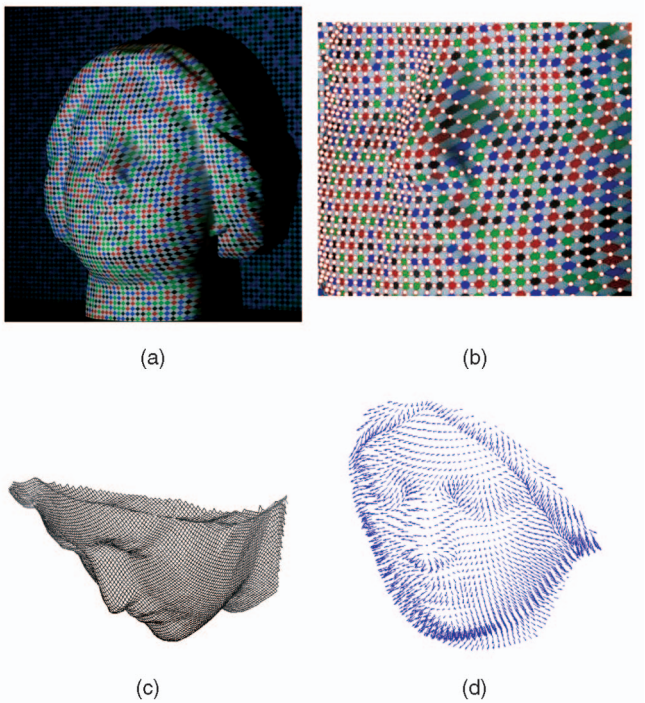

Fig. 10. 3D reconstruction of a bust object. (a) A bust object under patterned illumination. (b) Grid-point detection result. (c) Interpolated depth reconstruction using only the recovered 3D points. (d) Interpolated surface orientation map using only the recovered surface normals at the grid points.

Grid points detection was the first step to both position and orientation determinations. Fig. 10b shows the grid point detection result on the image of the bust object. Most of the grid points could be detected correctly despite the relatively large distortion of the pattern in the image projection due to varying curvature of the shape. The region near the nose was of curvature value varying quite drastically. A few grid points there were missed, yet most of the grid points in the region were detected correctly by our grid point detector.

Grid point detection on the image of the human face was more challenging because of the larger color variance and presence of texture. Nevertheless, our grid point extractor was effective in retrieving most of the grid points, as displayed in Fig. 11c. The extractor failed in a few regions, specifically those near the nose and the eyebrows where there was heavy occlusion. The occasional misses of feature detection could be compensated by the interpolation of the object surface, and the final reconstruction was still of adequate quality. We had tried traditional feature detectors like LoG, Harris's, SUSAN, and the template methods to detect the centroid features, but the results were much less favorable. Details of the comparison on the feature detection results can be found in another report [42]. In summary, the use of grid points as the feature points, together with the described grid point detector, allow more feature points to be detected in higher accuracy than what has been demonstrated on the traditional centroid features.

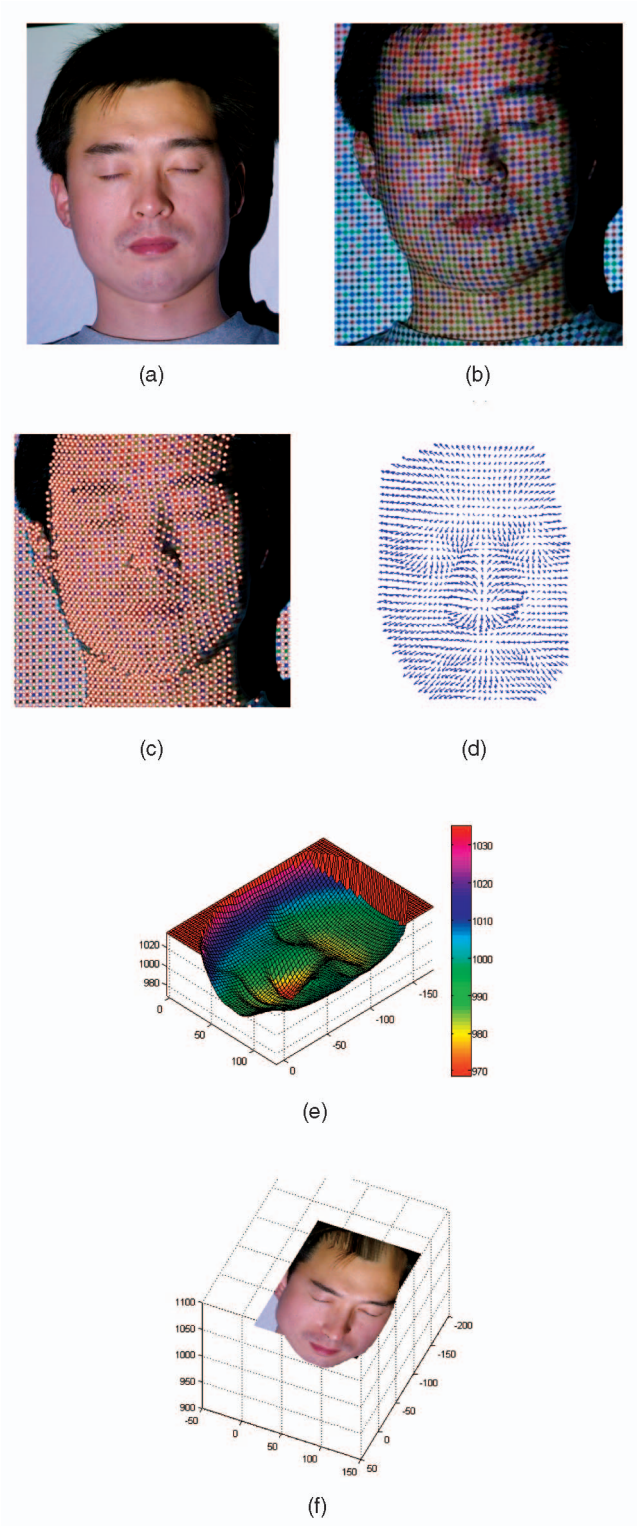

Fig. 11. 3D reconstruction of a human face. (a) The human face from the frontal perspective. (b) The face under patterned illumination. (c) Grid point detection result. (d) Surface orientation map interpolated from only the recovered surface normals at the grid points. (e) Reconstructed depth reconstruction using only the recovered 3D points. (f) 3D model of human face (as interpolated from the position data) with texture mapping.

Upon position determination at the extracted grid points, depth maps of the two objects could be constructed and they are shown in Figs. 10c and 11e, respectively. In particular, in the experiment on a human face, as mentioned above, there were grid points missing or encoded incorrectly. We used median filtering to remove outliers of the recovered 3D points, and used simple interpolation over the local neighborhood to fill in the occasional omissions of position data. With these, our final result had more than 2,200 grid points reconstructed successfully. The interpolated 3D model is shown in Fig. 11e. Fig. 11f shows the determined 3D human face with the original texture mapped on top.

We also determined the orientation maps separately, and the results are shown in Figs. 10d and 11d. By projecting the recovered 3D information to different viewpoints, we could examine the quality of the reconstruction visually. Such examinations showed that the reconstructed depth maps and orientation maps were of reasonable quality.

7 Conclusion and Future Work

We have described how a spatially coded structured light system can be used to determine not only 3D position but also 3D orientation independently. This is made possible by having a number of specific design features in the system. They include the adoption of grid points as opposed to centroid points as the feature points, and the inclusion of a mechanism that can extract orientation information at the grid points from the local curvatures of the grid lines that compose them. A few issues of the system are important, namely, how the grid points can be uniquely labeled and how they can be precisely extracted. To these we have presented solutions.

Empirical experiments show that not only can the system determine orientation that traditional SLSs cannot, it also offers position information in higher density and accuracy due to the use of grid point features and the better localizability of such features.

Future work will be about how the position and orientation information can be integrated in a complementary fashion for higher quality shape reconstruction.

Acknowledgments

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4195/04E).

References

[1] D. Nehab et al., "Efficiently Combining Positions and Normals for Precise 3D Geometry," *ACM Trans. Graphics,* vol. 24, no. 3, pp. 536-543, 2005.

[2] K. Ikeuchi and B.K.P. Horn, "Numerical Shape from Shading and Occluding Boundaries," *Artificial Intelligence,* vol. 17, no. 3, pp. 141-184, 1981.

[3] R.J. Campbell and P.J. Flyn, "A Survey of Free-Form Object Representation and Recognition Techniques," *Computer Vision and Image Understanding,* vol. 81, pp. 166-210, 2001.

[4] R. Zhang, P. Tsai, J.E. Cryer, and M. Shah, "Shape from Shading: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 8, pp. 690-705, Aug. 1999.

[5] J. Garding, "Direct Estimation of Shape from Texture," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp. 1202-1208, Nov. 1993.

[6] J. Salvi, J. Pagès, and J. Batlle, "Pattern Codification Strategies in Structured Light Systems," *Pattern Recognition,* vol. 37, no. 4, pp. 827-849, 2004.

[7] S. Kiyasu et al., "Measurement of the 3-D Shape of Specular Polyhedrons Using an M-Array Coded Light Source," *IEEE Trans. Instrumentation and Measurement,* vol. 44, no. 3, pp. 775-778, June 1995.
[8] P. Fong and F. Buron, "High-Resolution Three Dimensional Sensing of Fast Deforming Objects," *Proc. Int'l. Conf. Intelligent Robotics and Systems,* pp. 1606-1611, 2005.
[9] D. Desjardins and P. Payeur, "Dense Stereo Range Sensing with Marching Pseudorandom Patterns," *Proc. Canadian Conf. Computer and Robot Vision,* pp. 216-226, 2007.
[10] P. Griffin, L. Narasimhan, and S. Yee, "Generation of Uniquely Encoded Light Patterns for Range Data Acquisition," *Pattern Recognition,* vol. 25, no. 6, pp. 609-616, 1992.
[11] J. Salvi, J. Batlle, and E. Mouaddib, "A Robust-Coded Pattern Projection for Dynamic 3D Scene Measurement," *Pattern Recognition Letters,* vol. 19, pp. 1055-1065, 1998.
[12] Y.C. Hsieh, "Decoding Structured Light Patterns for Three-Dimensional Imaging Systems," *Pattern Recognition,* vol. 34, no. 2, pp. 343-349, 2001.
[13] S.Y. Chen and Y.F. Li, "Vision Processing for Realtime 3-D Data Acquisition Based on Coded Structured Light," *IEEE Trans. Image Processing,* vol. 17, no. 2, pp. 167-176, Feb. 2008.
[14] R.A. Morano et al., "Structured Light Using Pseudorandom Codes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 3, pp. 322-327, Mar. 1998.
[15] C.J. Davies and M.S. Nixon, "A Hough Transform for Detecting the Location and Orientation of Three-Dimensional Surfaces via Color Encoded Spots," *IEEE Trans. Systems, Man, and Cybernetics, Part B,* vol. 28, no. 1, pp. 90-95, Feb. 1998.
[16] I. Ishii et al., "High-Speed 3D Image Acquisition Using Coded Structured Light Projection," *Proc. Int'l. Conf. Intelligent Robots and Systems,* pp. 925-930, 2007.
[17] F. Gray, "Pulse Code Communication," US Patent 2632058, Mar. 1953.
[18] R.J. Valkenburg and A.M. McIvor, "Accurate 3D Measurement Using a Structured Light System," *Image and Vision Computing,* vol. 16, no. 2, pp. 99-110, 1998.
[19] W. Krattenthaler, K.J. Mayer, and H.P. Duwe, "3D-Surface Measurement with Coded Light Approach," *Proc. 17th Meeting Austrian Assoc. for Pattern Recognition on Image Analysis and Synthesis,* pp. 103-114, 1994.
[20] S. Zhang and P.S. Huang, "High-Resolution, Real-Time Three-Dimensional Shape Measurement," *Optical Eng.,* vol. 45, no. 12, p. 123601, 2006.
[21] D. Scharstein and R. Szeliski, "High-Accuracy Stereo Depth Maps Using Structured Light," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 195-202, 2003.
[22] F. Sadlo and T. Weyrich, "A Practical Structured Light Acquisition System for Point-Based Geometry and Texture," *Proc. Eurographics Symp. Point-Based Graphics,* pp. 89-98, 2005.
[23] J. Gühring, "Dense 3D Surface Acquisition by Structured Light Using Off-the-Shelf Components," *Proc. Int'l Soc. Optical Eng.,* pp. 220-231, 2000.
[24] Z. Song and R. Chung, "Off-the-Shelf Structured Light-Based System for Accurate 3D Reconstruction," *HKIE Trans.,* vol. 15, no. 4, pp. 44-51, 2008.
[25] P. Vuylsteke and A. Oosterlinck, "Range Image Acquisition with a Single Binary-Encoded Light Pattern," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 2, pp. 148-163, Feb. 1990.
[26] A. Adan et al., "3D Feature Tracking Using a Dynamic Structured Light System," *Proc. Second Canadian Conf. Computer and Robot Vision,* pp. 168-175, 2005.
[27] I.C. Albitar, P. Graebling, and C. Doignon, "Robust Structured Light Coding for 3D Reconstruction Computer Vision," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 1-6, 2007.
[28] L. Zhang, B. Curless, and S. Seitz, "Rapid Shape Acquisition Using Color Structured Light and Multi-Pass Dynamic Programming," *Proc. First Int'l Symp. 3D Data Processing, Visualization, and Transmission,* pp. 22-36, 2002.
[29] C.S. Chen, Y.P. Hung, C.C. Chiang, and J.L. Wu, "Range Data Acquisition Using Color Structured Lighting and Stereo Vision," *Image and Vision Computing,* vol. 15, no. 6, pp. 445-456, 1997.
[30] J. Pages, J. Salvi, and J. Forest, "Optimized De Bruijn Patterns for One-Shot Shape Acquisition," *Image and Vision Computing,* vol. 23, no. 8, pp. 707-720, 2005.
[31] G. Healey and T.O. Binford, "Local Shape from Specularity," *Proc. IEEE Int'l Conf. Computer Vision,* vol. 1, pp. 151-160, 1987.
[32] D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, "Determining Surface Orientations of Transparent Objects Based on Polarization Degrees in Visible and Infrared Wavelengths," *J. Optical Soc. Am. A,* vol. 19, pp. 687-694, 2002.
[33] R.J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images," *Optical Eng.,* vol. 19, no. 1, pp. 139-144, 1980.
[34] N. Shrikhande and G. Stockman, "Surface Orientation from a Projected Grid," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 11, no. 6, pp. 650-655, June 1989.
[35] S. Winkelbach and F.M. Wahl, "Shape from 2D Edge Gradients," *Proc. 23rd DAGM Symp. Pattern Recognition,* pp. 377-384, 2001.
[36] S. Winkelbach and F.M. Wahl, "Shape from Single Stripe Pattern Illumination," *Proc. 24th DAGM Symp. Pattern Recognition,* pp. 240-247, 2002.
[37] F.J. MacWilliams and N.J.A. Sloane, "Pseudo-Random Sequences and Arrays," *Proc. IEEE,* vol. 64, no. 12, pp. 1715-1729, Dec. 1976.
[38] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. Royal Soc. London,* vol. 207, pp. 187-217, 1980.
[39] C. Harris and M.J. Stephens, "A Combined Corner and Edge Detector," *Proc. Fourth Alvey Vision Conf.,* pp. 147-152, 1998.
[40] S.M. Smith and M. Brady, "SUSAN—A New Approach to Low Level Image Processing," *Int'l. J. Computer Vision,* vol. 23, no. 1, pp. 45-78, 1997.
[41] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences.* Lawrence Erlbaum Assoc., 1988.
[42] Z. Song and R. Chung, "Grid-Point Extraction Exploiting Point Symmetry in a Pseudorandom Color Pattern," *Proc. Int'l. Conf. Image Processing,* pp. 1956-1959, 2008.
[43] Z. Song and R. Chung, "Use of LCD Panel for Calibrating Structured Light-Based Range Sensing System," *IEEE Trans. Instrumentation and Measurement,* vol. 57, no. 11, pp. 2623-2630, Nov. 2008.

Zhan Song received the PhD degree in mechanical and automation engineering from the Chinese University of Hong Kong in 2008. He is currently with the Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Sciences (CAS), as an assistant researcher. His current research interests include structured light-based sensing, image processing, 3D face recognition, and human-computer interaction.

Chi-Kit Ronald Chung received the BSEE degree from the University of Hong Kong and the PhD degree in computer engineering from the University of Southern California, Los Angeles. He is currently with the Chinese University of Hong Kong as director of the Computer Vision Laboratory and a professor in the Department of Mechanical and Automation Engineering. His research interests include computer vision and robotics. He is a senior member of the IEEE and a member of MENSA. He was the chairman of the IEEE Hong Kong Section, Joint Chapter on Robotics and Automation Society and Control Systems Society from 2001 to 2003.