3D Shape Recovery by the Use of Single Image Plus Simple Pattern Illumination

Zhan Song and Ronald Chung*

Department of Mechanical & Automation Engineering
The Chinese University of Hong Kong, Hong Kong, China
{zsong,rchung}@mae.cuhk.edu.hk

Abstract. This paper presents a method of surface orientation and in turn shape recovery from a single image captured under projection of a simple checker-board pattern. The essences of the method include that only one image is required, that accurate correspondence establishment between the image and the projected pattern is not necessary, that the determination of 3D is much less sensitive to imaging noise and illumination condition than intensity-based methods like shape from shading. The method relies upon the fact that surface orientations at the grid points are only decided by image tangents in the image data. Experiments on planar, spherical, and ribbon-like surfaces show that, with accurate calibration of the projector-and-camera system through a mechanism we proposed earlier, 3D shape can be recovered with ease and precision both much better than before.

1 Introduction

Recovering 3D surface description of a scene or object has been one of the most important problems in computer vision. Many approaches have been proposed, such as stereo vision [1], structured light based approach [2], and laser strip scanning [3]. Correspondence problem is the key challenge in these approaches. To avoid tackling the correspondence problem, single image-based methods such as shape from texture [4], shape from shading [5], and shape from specularity [6] have also been suggested, which allow not absolute depth but relative depth information to be determined. Relative depth description has the scale ambiguity in comparison with the absolute depth description, but is a shape description sufficient for many applications including model registration, object recognition, pose estimation, and segmentation [7].

In this paper, we propose a method of recovering relative depth from a single image that is captured under projection of a simple pattern – the checker board pattern. The method does not require to establish correspondence between the image plane and the projector panel. Compared with methods like shape from shading and shape from specularity, the method does not depend upon the absolute values of image intensities and therefore is far less sensitive to imaging noise, illumination condition, and surface reflectance variance of the imaged scene. The method is based upon this principle. A known projected pattern, upon reflection from scene *S*, will have its image appearance

* Corresponding author.

G. Bebis et al. (Eds.): ISVC 2007, Part I, LNCS 4841, pp. 268–277, 2007.

modulated by the surface profile of *S*. With knowledge of the projected pattern which is under full control of the method, plus image observations of the reflected pattern, 3D information of the imaged scene *S* can be determined. In this paper, we shall show that surface orientation at any projected grid-point **P** on the target scene can be determined solely by the tangents to the imaged grid-lines that intersect to form the image position $\mathbf{p}_c$ of **P** on the image plane. Experimental results on a variety of surfaces – planar, spherical, ribbon-like ones – show that with accurate calibration of the projector-and-camera system using a system we proposed earlier [18], shape can be reconstructed with ease and precision far surpassing those of the previous work.

This paper is organized in the following way. In Section 2, previous work on surface orientation recovery is briefly reviewed. The principle of determining surface orientation from two image tangents is presented in Section 3. In Section 4, experimental results on a variety of surfaces are shown. Conclusions and future work are offered in Section 5.

2 Previous Work

Approaches for visual recovery of relative depth or surface orientation can be categorized into two classes. The first class makes use of the natural features on the target surface and assume certain ideal property (constant surface albedo, uniform textures etc.) of them. It is also named shape-from-X techniques, where X can be shading, texture, etc. The second class utilizes artificial textures or patterns that are projected onto the target surface. Since it does not depend upon the existence and observability of natural features, it makes less assumption on the target surface, and is generally more robust to imaging noise and illumination condition and variance in the reflectance property. Below we review in more details two reconstruction methods that are particularly related to this work.

Shape from shading is one representative method of the shape-from-X technique. The gradual variation of shading (i.e., the gray level intensity) in the image is utilized to recover the surface shape. The Lambertian reflectance model is usually assumed, in which the gray level at any position in the image is solely determined by the surface orientation and the incident angle of light there [9]. To a gray level image the method aims to recover the light source direction and the surface orientation in 3D at each image position. The shape from shading problem can be formulated as that of finding the solution of a nonlinear first-order partial differential equation called the brightness equation. It has been shown to be an ill-posed problem that generally has infinitely many solutions [10]. Additional heuristics on surface smoothness or intensity gradient are usually adopted to let surface orientations be determined in more precise form.

Structured light projection methods have also been proposed to recover local surface orientation. In the method proposed by Shrikhande and Stockman [11], a grid pattern is projected onto an object. Surface orientations at the grid-points (the intersections of the grid-lines) are induced from the change of the lengths of the grid edges in the image data. The result however indicates that there could be large errors at places where the grid-cell edges have large distortion, and such distortion is often induced by sharp edges in 3D. Woodham [12] proposed a photometric stereo method which determines surface orientation at each image position by changing the direction of light sources. Sugihara [13] proposed a method which infers surface orientation using the distortion of a known pattern projected on the surface from a structured light source. The projected textures on the object surface are detected and the distortion from the regular pattern is measured in order to estimate the surface orientation. Winkelbach and Wahl [14] used a uni-direction strip pattern to compute the surface normal at each edge point. A fringe pattern is projected onto the object twice from two different angles, each time with an image captured, so that two surface tangents are available for each image position. Surface normal at every image point can be interpolated from the data in the two images. A simplified method which determines surface normals from the slopes and intervals of the stripes in the image was also proposed [15] [16]. It is based on the assumption that the surface patch between two strip edges is planar or very smooth. Thus the tilt angle can be estimated from the deformed widths of the strip by comparing them with the strip width measured for a reference plane. In their system, pattern projection is assumed a parallel projection, and image capture 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 much simplified projection and imaging models are inevitable.

Our method is related to the method proposed in [14], but we use checker-board pattern not fringe pattern. We also assume not parallel projection nor parallel imaging, but perspective projection and imaging. With more accurate models for imaging (camera) and for illumination (pattern projection), more accurate 3D determination can be attained. We show that by the use of a calibration method we proposed [18] for precise calibration of the projector-camera system, surface orientation and relative depth or shape can be recovered precisely.

3 A Method Requiring Only One Image Under Illumination of Simple Pattern

3.1 The Underlying Principle

Suppose we program a pattern on the display panel of the projector, which consists of a 2D array of alternate dark and bright rectangular blocks, and project it to the object surface. The edge of the blocks forms the grid-lines, and the intersections of every two grid-lines form the grid-points. The grid-lines and grid-points are accessible in both the projector panel and the image data.

Consider any grid-point $\mathbf{p}_p$ in the pattern panel of the projector, and the two accompanying grid-lines that compose it, as illustrated by Fig. 1. Suppose the grid-point and grid-lines induce an imaged grid-point $\mathbf{p}_c$ and two imaged 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 at point $\mathbf{p}_p$ in the pattern projection panel are $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ respectively, and the tangents to the grid-lines at point $\mathbf{p}_c$ in the image plane are $\mathbf{t}_{c1}$ and $\mathbf{t}_{c2}$.

Tangent $\mathbf{t}_{p1}$ and the grid-point $\mathbf{p}_p$ in the pattern projection panel together form a plane $\prod(\mathbf{p}_p, \mathbf{t}_{p1})$ of illumination from the light source, which is reflected by the object surface at point **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 **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 **P** is merely the cross-product of the surface normals of the two light planes.

Similarly, another tangent $\mathbf{t}_2$ to the object surface at point **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 taking 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, the surface orientation $\mathbf{n}(x,y)$ of the object surface at the associated 3D point can be determined simply as $\mathbf{n}(x,y)=\mathbf{t}_1\times\mathbf{t}_2$ from image observations $\{\mathbf{p}_c, \mathbf{t}_{c1}, \mathbf{t}_{c2}\}$ and pattern data $\{\mathbf{p}_p, \mathbf{t}_{p1}, \mathbf{t}_{p2}\}$. This is the underlying principle of this work.

3.2 The Path from Image Tangents to 3D Orientation

Here we elaborate the calculations more precisely. Suppose $\mathbf{n}_{c1}$ and $\mathbf{n}_{c2}$ are the surface normals 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 $\mathbf{n}_{p2}$ the surface normals 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 is coming up out of it instead of going into it) have been calibrated, which are focal lengths f_c, f_p, and principal points $\boldsymbol{C}_c$ (u_{c0}, v_{c0}), $\boldsymbol{C}_p(u_{p0}, v_{p0})$. Then $\mathbf{n}_{c1}$, $\mathbf{n}_{c2}$, $\mathbf{n}_{p1}$, $\mathbf{n}_{p2}$ can be determined as:

$$\vec{\mathbf{n}}_{ci} = \left[(u_{ci} - u_{c0}) \quad (v_{ci} - v_{c0}) \quad -f_c\right] \times \left[\cos\theta_{ci} \quad \sin\theta_{ci} \quad 0\right] \quad i = 1,2 \tag{1}$$

$$\vec{\mathbf{n}}_{pi} = \left[(u_{pi} - u_{p0}) \quad (v_{pi} - v_{p0}) \quad -f_p\right] \times \left[\cos\theta_{pi} \quad \sin\theta_{pi} \quad 0\right] \quad i = 1,2 \tag{2}$$

where θ_* indicates the slant angle of the associated 2D tangent $\mathbf{t}_*$ with respect to the camera's image plane or projector's display panel as illustrated by Fig. 2a.

So the two 3D tangents $\mathbf{t}_1$, $\mathbf{t}_2$ to the object surface at point **P** that is associated with image point $\mathbf{p}_p=(x,y)$, if with reference to the camera coordinate system, can be obtained as:

$$\mathbf{t}_1 = \mathbf{n}_{c1} \times (\mathbf{R}\mathbf{n}_{p1}) \tag{3}$$

$$\mathbf{t}_2 = \mathbf{n}_{c2} \times (\mathbf{R}\mathbf{n}_{p2}) \tag{4}$$

where **R** represents the rotational relationship between the camera and projector coordinate frames.

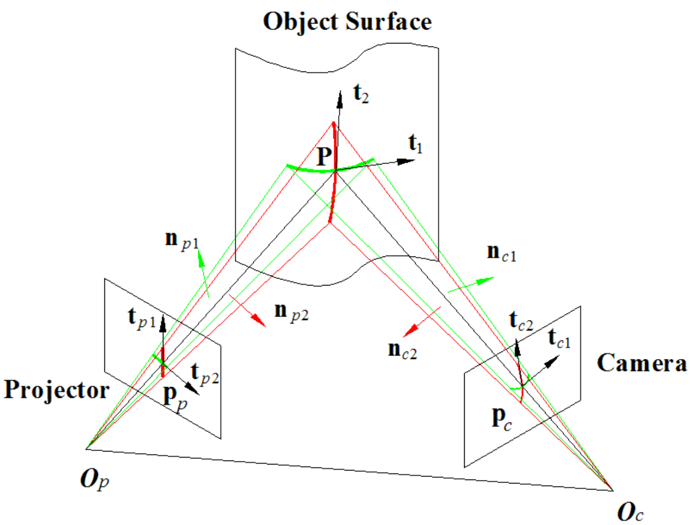
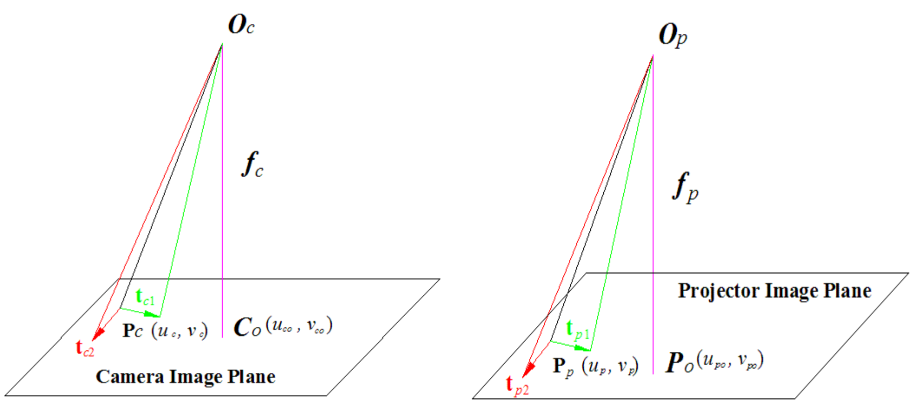


Fig. 1. Determining surface orientation from two image tangents, for object surface under pattern illumination



(a) (b)

Fig. 2. Definition of the slant angle of a 2D tangent as used in the calculation

Finally, the surface orientation to the object surface at point **P** that is associated with the image position (x,y) can be determined as:

$$\mathbf{n}(x, y) = \mathbf{t}_1 \times \mathbf{t}_2 \tag{5}$$

Notice that the determination of local surface orientation as expressed by Equation (5) is a deterministic process that requires only image information local to the specific point to operate. Unlike shading from shading and many other methods, 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.

3.3 The Needlessness of Correspondences

The previous analysis indicates that surface orientation in 3D can indeed be determined from image tangents, but it also points out that for each imaged grid-point $\mathbf{p}_p=(x,y)$ and the accompanying image tangents $\mathbf{t}_{c1}$ and $\mathbf{t}_{c2}$, it requires also the 2D tangents $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ and other local information in the projector's display panel to go with them. In other words, correspondence between grid-points in the image and grid-points in the projector panel seems necessary.

However, we shall show that with the suitable projection pattern and suitable light source, it is possible that the above correspondence be avoided. Specifically, we choose to use the checker-board pattern for projection, so that the 2D tangents $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ to the grid-lines at each grid-point on the projector panel are all the same. Under the assumption of parallel projection, the invariant $\mathbf{t}_{p1}$ and $\mathbf{t}_{p2}$ would lead to the determination of surface normals $\mathbf{n}_{p1}$ and $\mathbf{n}_{p2}$ (normals to the light planes from the projector) which are invariant with the grid-points. With this, the necessary local entities on the projector side can all be made global, and a once-and-for-all pre-estimation step will make correspondence between the image plane and the projector panel unnecessary. In previous work such as [14][15][16], the parallel projection is assumed not only the projector but also the camera.

However, the projector illumination process may not be exactly captured by the parallel projection model, and discrepancy from it will cause bigger error to the determination of $\mathbf{n}_{p1}$ and $\mathbf{n}_{p2}$ at positions farther away from the center (i.e.., the principal point) of the projector panel. We use a simple correction mechanism to fix the problem. For each grid-point $\mathbf{p}_c$ ($=(u_c,v_c)$) in the image plane, we adopt a linear mapping, as described below, to estimate more finely the associated grid-point $\mathbf{p}_p$ ($=(u_p,v_p)$) in the projector panel:

$$[u_p, v_p] = [\frac{u_c - u_{\min}}{u_{\max} - u_{\min}} W_p, \frac{v_c - v_{\min}}{v_{\max} - v_{\min}} H_p] \tag{6}$$

where $(u_{\min}, v_{\min})$, $(u_{\max}, v_{\max})$ represent the top-left and bottom-right grid point positions in the image plane, W_p and H_p are the width and height of the pattern on the projector panel counted in projector panel's pixels. This way, a projection model closer to perspective projection than parallel projection model is used to determine $\mathbf{n}_{p1}$ and $\mathbf{n}_{p2}$ at the estimated grid-point $\mathbf{p}_p$, and this is illustrated by Fig. 2b.

3.4 Calibration of Projector-and-Camera System

While in previous work like [14][15][16] illumination and imaging are modeled as parallel projections, here they are modeled as perspective projections. With this, the proposed method can expectedly reach more precision in 3D reconstruction.

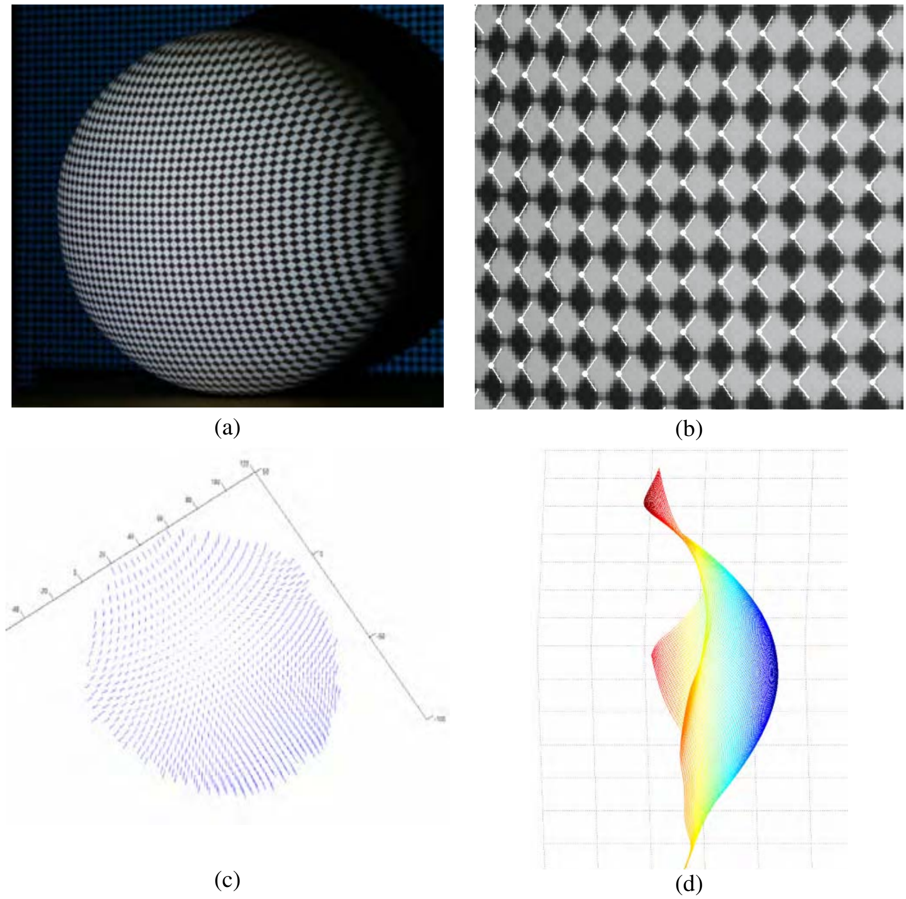


Fig. 3. Reconstruction of a spherical surface by the proposed method: (a) object surface under pattern illumination; (b) appearance of grid-points and grid-lines in the captured image; (c) reconstructed local surface orientations as viewed from one perspective; (d) interpolated surface as viewed from another perspective.

However, it is also obvious that the calibration of the intrinsic and extrinsic parameters of the projector-and-camera system has an important role to play. On this we use a calibration mechanism [18] that makes use of an LCD display panel (separate from the one of the projector) as the external reference object. The calibration mechanism has shown to be able to reach very high calibration accuracy, allowing more precise 3D reconstruction to take place.

3.5 Feature Extraction

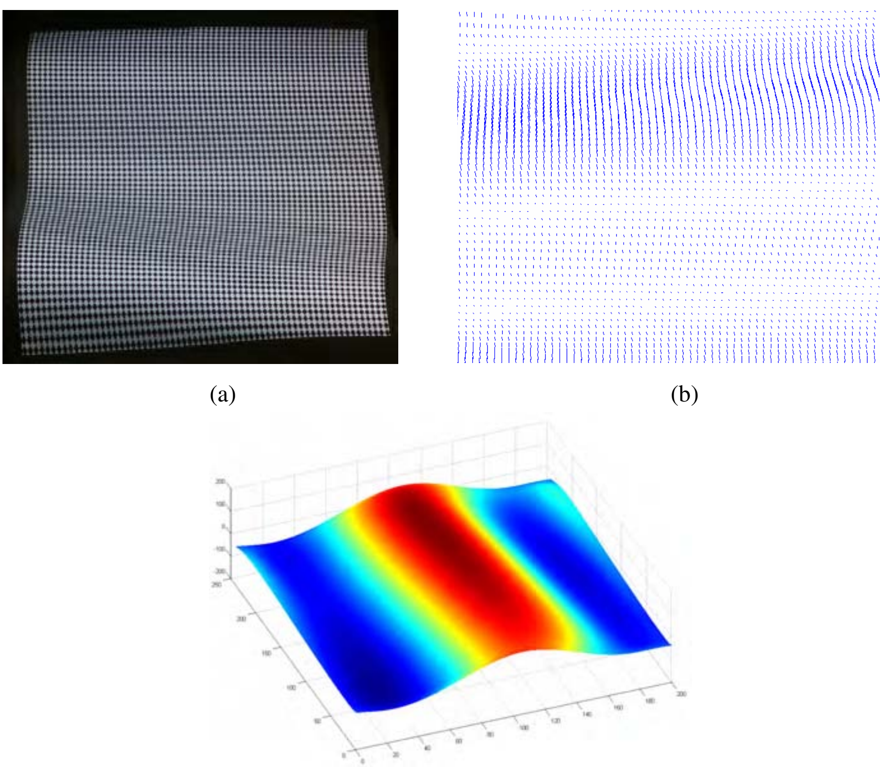
The reconstruction accuracy also depends upon how accurately the grid-lines and grid-points are located in the image, and how precisely 2D tangents to the grid-lines at the grid-points are extracted. There has been a rich body of works in the literature on the topic. In this work we use an algorithm that is of sub-pixel precision, whose description is beyond the page limit of this report and will be reported elsewhere.

3.6 Surface Interpolation

Once the local surface orientations at the grid-points are computed, a simple median filter is used to remove the outliers among them which mostly exist around the image boundary. Then the discrete data go through linear interpolation to form dense surface description, which embeds a regularization term related to smoothness. More precisely, we use the 2D integration method proposed by Frankot and Chellappa [17].

4 Experimental Results

A structured light system consisting of a DLP projector and digital camera was used for experiments. The system was first calibrated using the method described in [18] that made use of an LCD panel as the external reference object.



(a) (b)

(c)

Fig. 4. Reconstruction of a ribbon-like surface by the proposed method: (a) object surface under pattern illumination; (b) appearance of grid-points and grid-lines in the captured image; (c) reconstructed and interpolated surface as viewed from one perspective.

We used the LCD panel whose planarity was of industrial grade and which could be regarded as a perfectly planar surface to evaluate the quality of shape reconstruction. Shape reconstruction from the proposed method showed only very small deviation from planarity: of all the measured points, only an average discrepancy of 0.82° and standard deviation of 0.15° in the distribution of local surface orientations. Note that in methods like [14][15][16] and other earlier works a discrepancy in the range of 2° to 8° were reported.

We have also conducted experiment on a spherical object. The shape reconstruction result by the proposed method is as shown in Fig. 3: Fig. 3(a) shows the object under illumination of the checker-board pattern, Fig. 3(b) shows the appearance of grid-points and grid-lines in the image data, and Fig. 3(c) and (d) show the determined local surface orientations and the final interpolated shape as viewed from two viewpoints. The mean error of shape reconstruction was only 1.39°.

A ribbon-like surface was reconstructed and the result is shown in Fig. 4. Though ground truth of this surface was not available for evaluating in precise terms the reconstruction quality, visual check showed that the reconstruction was reasonable.

5 Conclusions and Future Work

Relative depth or shape description independent of scale is an important shape expression sufficient for many applications. We have described a method for its reconstruction that requires only off-the-shelf instruments which are not many more than a camera and a projector. The operation is also simple: requiring the projection of a checker-board pattern by the projector and just one image capture by the camera. The working mechanism is straightforward as well; it determines local surface orientations deterministically without the need of going through iterations, and it requires no feature correspondence between the projector panel and the image data either if relative depth is all that is desired, though feature correspondences are not excluded from the operation of the method and their presence could actually escalate the relative depth to absolute depth. Experiments show that with accurate calibration of the projector-and-camera system through system like [18], reconstruction could be of very promising precision.

As the method does not make use of raw intensities in the image data, it is more robust to imaging noise and illumination variations than intensity-based methods like shape from shading and other photometric methods, and it requires to make less assumption on the reflectance property of the object surface. Compared to other methods that are structured light based like [14][15][16], it is based upon more general models for projection and imaging, and it requires no scanning. Future work includes how additional images could boost the reconstruction precision, and how look-up table for all possible slant angles θ_{c*} can speed up the computations further.

References

1. Taylor, C.J.: Surface Reconstruction from Feature Based Stereo. In: 9th IEEE International Conference on Computer Vision, pp. 184–190 (2003)
2. Salvi, J., Pages, J., Batlle, J.: Pattern Codification Strategies in Structured Light Systems. Pattern Recognition 37(4), 827–849 (2004)

3. Kawasaki, H., Furukawa, R.: 3D Acquisition System using Uncalibrated Line-Laser Projector. In: 18th International Conference on Pattern Recognition, vol. 1, pp. 975–979 (2006)
4. Garding, J.: Direct Estimation of Shape from Texture. IEEE Trans. on PAMI 15, 1202–1208 (1993)
5. Zhang, R., Tsai, P., Cryer, J.E., Shah, M.: Shape from Shading: A Survey. IEEE Trans on. PAMI 21(8), 690–705 (1999)
6. Healey, G., Binford, T.O.: Local Shape from Specularity. In: 1st ICCV, pp. 151–160 (1987)
7. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Three-Dimensional Face Recognition. Computer Vision 64(1), 5–30 (2005)
8. Brooks, M.J., Chojnacki, W.: Direct Computation of Shape from Shading. In: 12th International Conference on Pattern Recognition, pp. 114–119 (1994)
9. Lange, H.: Advances in the Cooperation of Shape from Shading and Stereo Vision. Second International Conference on 3-D Imaging and Modeling, 46–58 (1999)
10. Woodham, R.J.: Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from Image Intensity. In: 22nd Int. Symp. SPIE, vol. 155, pp. 136–143 (1978)
11. Shrikhande, N., Stockman, G.: Surface Orientation from a Projected Grid. IEEE Trans. on PAMI 2(6) (1989)
12. Horn, B.K.P., Brooks, M.J.: The Variational Approach to Shape from Shading. Computer Vision Graphics Image Processing 33, 174–208 (1986)
13. Sugihara, K., Okazaki, K., Kaihua, F., Sugie, N.: Regular Pattern Projection for Surface Measurement. In: 2nd International Symp. Robotics Research, pp. 17–24 (1984)
14. Winkelbach, S., Wahl, F.M.: Shape from 2D Edge Gradients. In: Radig, B., Florczyk, S. (eds.) Pattern Recognition. LNCS, vol. 2191, pp. 377–384. Springer, Heidelberg (2001)
15. Winkelbach, S., Wahl, F.M.: Shape from Single Stripe Pattern Illumination. In: Van Gool, L. (ed.) Pattern Recognition. LNCS, vol. 2449, pp. 240–247. Springer, Heidelberg (2002)
16. Asada, M., Ichikawa, H., Tsuji, S.: Determining Surface Orientation by Projecting a Stripe Pattern. IEEE Trans. on PAMI 10(5), 749–754 (1988)
17. Robert, T., Frankot, R.: A Method for Enforcing Integrability in Shape from Shading Algorithms. IEEE Trans. on PAMI 10(4) (1988)
18. Song, Z., Chung, R., Yeung, S.W.-L.: An Accurate Camera-Projector System Calibration by the use of an LCD Panel. In: Proceedings of the 2007 International Conference on Image Processing, Computer Vision, and Pattern Recognition, pp. 164–170 (2007)