Precise 3D Reconstruction from a Single Image

Changqing Zou[1,2,3], Jianbo Liu[1], and Jianzhuang Liu[1,4,5]

[1] Shenzhen Key Lab for CVPR, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[2] Graduate University of Chinese Academy of Sciences, Beijing, China
[3] Department of Physics and Electronic Information Science, Hengyang Normal University, Hengyang, China
[4] Department of Information Engineering, The Chinese University of Hong Kong
[5] Media Lab, Huawei Technologies Co. Ltd., China
{cq.zou,jb.liu}@siat.ac.cn, liu.jianzhuang@huawei.com

Abstract. 3D object reconstruction from single images has extensive applications in multimedia. Most of existing related methods only recover rough 3D objects and the objects are often required to be interconnected. In this paper, we propose a novel method which uses a set of auxiliary reference grids to precisely reconstruct 3D objects from a single uncalibrated image. In our system, the user first draws the line drawings of the objects. Then, the initial focal length f of the camera is computed with a calibration method, and then the initial focal length is refined by a reference grid. With the refined f, a 3D position measurement environment is constructed, and a world coordinate system is defined by the user. After that, a set of reference grids are used to find the precise 3D locations of the object points and the wireframes of the objects are recovered automatically. Finally, the system generates the surfaces and renders the complete 3D objects. Besides the precise 3D modeling, our reconstruction method does not require the objects in a scene are interconnected. A set of examples are provided to demonstrate the ability of handling complex polyhedral objects and curved surfaces within one framework.

1 Introduction

3D object reconstruction from a single 2D view of a scene has attracted considerable attention. It has many applications such as 3D environment modeling, 3D object design, and 3D object retrieval. This research is also a challenging problem because it is ill-posed with less information available from a single view.

Many methods have been proposed for 3D reconstruction from single images, such as [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. Debevec et al. [4] recover 3D buildings with parametric primitives from one or several photos taken by a calibrated camera. This system requires a large amount of user interaction involved [12]. Besides, it requires the camera to be pre-calibrated with known intrinsic parameters. Liebowitz et al. [6] create architectural models by geometric relationships from architectural scenes. Their method needs to

K.M. Lee et al. (Eds.): ACCV 2012, Part IV, LNCS 7727, pp. 271–282, 2013.

compute the vanishing lines of all object faces with auxiliary planes and the reconstruction errors are easy to accumulate with the face by face propagation. Sturm and Maybank's method [7] first does camera calibration, then recovers part of the points and planes, and finally obtains the 3D positions of other faces by propagation. The limitation of this method is that the parts of objects have to be sufficiently interconnected and the reconstruction errors may also be accumulated. Guillou et al. [5] carry out 3D reconstruction with a rectangular 3D box that fits at best with the potential objects within the scene. It can only recover simple planar objects. Jelinek and Taylor [8] propose a method of polyhedron reconstruction from single images using several camera models, which requires that the polyhedra have to be linearly parameterized. Zhang et al. [2] try to reconstruct free-form 3D models from single images. The method costs the user a lot of time to specify the constraints from an image.Shimodaira [11] uses the shading information, one horizontal or vertical face, and convex and concave edges to recover the shape of polyhedron in single images. This method handles simple polyhedral only. Li et al. [13] reconstruct objects from single images by obtaining a closed-form solution for the shape vector, using connectivity and perspective symmetry properties. This method only considers planar objects. Liu et al. [14], [15] try to reconstruct complex 3D planar objects from single images. These methods suppose the imaging is parallel or weak-perspective projection and they only obtain rough 3D models.

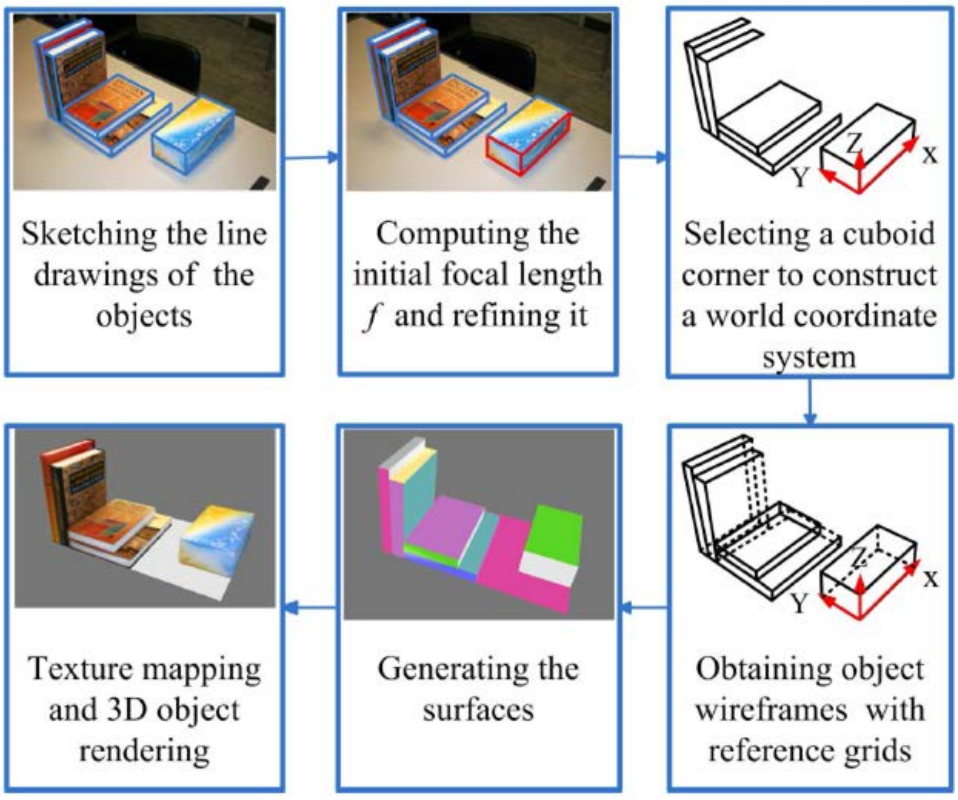

Fig. 1. Flow chart of our approach

In this paper, we propose a novel method which uses a set of auxiliary reference grids to precisely reconstruct both polyhedral objects and curved-surfaced objects from a single image with unknown camera parameters. In our system, the user first draws the edges and vertices of the objects. Then, the initial focal length f of the camera is computed with a calibration method. After that, we use a reference grid to obtain an accurate focal length. With the accurate f, we construct a 3D measurement environment and a world coordinate system. The 3D wireframes of the objects of interest are created by the reference grids. Finally, we generate the surfaces and render the complete 3D objects. The pipeline of our system is shown in Fig. 1.

Similar to previous works, we also focus on man-made objects. However, previous methods mainly consider plausible object reconstruction, while ours concentrates on precise reconstruction. Man-made objects usually have regular shapes such as rectangular faces, parallel lines, orthogonal corners,symmetrical structures, and lines perpendicular to the ground. These properties are easy to be identified by humans. Compared with previous works, ours has the following advantages: (1) Because of the refined camera focal length, the reference grids used to obtain accurate measurement, and a little user interaction, our method can precisely recover complex up-to-a-scale 3D objects. (2) Compared to other approaches it may handle a wider class of objects and demonstrates the ability of handling polyhedral objects and curved surfaces within one framework. (3) It may obtain accurate relative positions between objects in a scene which are not interconnected.

2 Calibration and Measurement Environment

In this work, we assume that there is at least one visible rectangular cuboid corner that can be perceived by the user. Actually, this is common in man-made objects. This corner is used to estimate the focal length and build a world coordinate system.

We consider the viewing camera coordinate system as an orthogonal Cartesian system with x, y and z axes, associated with a pin-hole camera. The origin of the system coincides with the view point, and the z direction is the optical axis orthogonal to the screen (image plane). The focal length determines the appearance of the object image in the camera. Therefore, to recover the objects, it is necessary to determine the focal length. We assume that the principle point (the intersection point between the z axis and the image plane) is located at the image center and a camera model is given with the following intrinsic matrix

$$\mathbf{k} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1}$$

where f is the focal length which needs to be found. The most widely used approach to find f is to use vanishing points and vanishing lines. However, obtaining the accurate f by the geometric constraints is not easy since computing

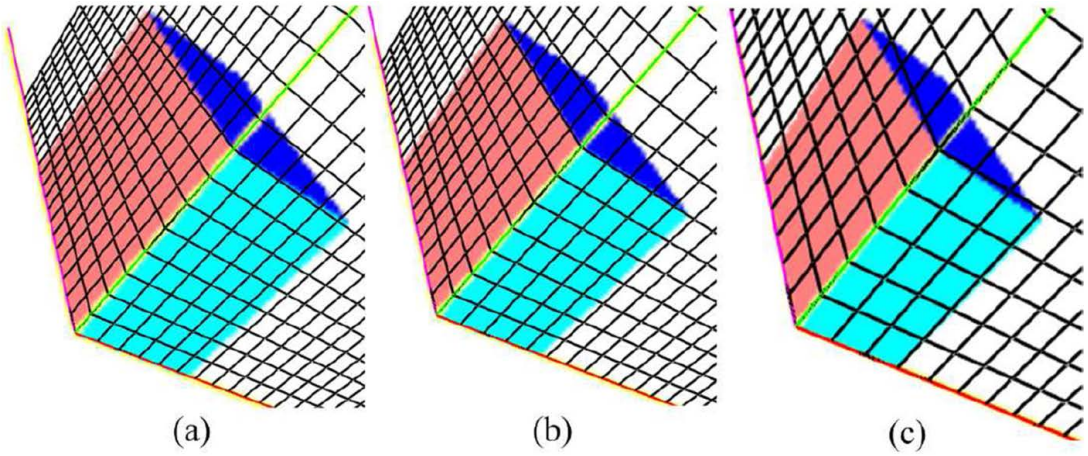

Fig. 2. Refining the focal length. Fig. 2a and 2c show the cases of $f_i \neq f_r$. Fig. 2b shows the case of $f_r = f_i$.

vanishing points and lines is numerically unstable. In our system, based on a known cuboid corner, we use the method in [16] to obtain an initial focal length.

After the initial focal length is obtained, the user employs an orthogonal reference grid (e.g., the one in Fig. 2) to refine the focal length, and then reconstructs the measurement environment. We take Fig. 2 as an example to illustrate this process, where the known cuboid corner is identified by the user.

The user first loads an image on the screen. OpenGL is used to render the image and the 3D reference grid. The initial focus length is used in the rendering . The reference grid, which can be rotated, translated, and zoomed by the user, is adjusted to match the lower cuboid corner of the block, as shown in Fig. 2 (In this process the cuboid corner point of reference grid is placed to a initialized position with a fixed Z coordinate). If the initial focal length f_i is not equal to the real focal length f_r, the upper edge of the block's faces will not overlap with two lines of the reference grid, as shown in Fig. 2a ($f_r > f_i$) and Fig. 2c ($f_i > f_r$). The accurate focal length can be found by adjusting it in an OpenGL rendering function such that the edges of the rectangular faces of the block coincide with the lines of the reference grid, as shown in Fig. 2b. Thanks to the initial focal length, this adjusting is done easily and quickly by the user.

Once the refined focal length has been obtained, the measurement environment is also constructed. It is the combination of the camera coordinate system and a world coordinate system. The origin of the world coordinate system is set at the cuboid corner and its axes coincide with the three edges of the cuboid (see Fig. 2b). In this environment, a set of 3D reference grids can be used to facilitate the 3D reconstruction of the objects.

3 Generating Object Wireframes

A 3D object wireframe consists of the 3D vertices and edges of the object. These 3D vertices and edges are derived with the help of a set of 3D reference grids put

on the image by the user. In this section, we first introduce the reference grids (Fig. 3), and then take the reference grids in Figs. 3a and b as an example to explain how to obtain 3D coordinates of an object's vertices in an image. Finally we discuss how to recover curves.

3.1 Reference Grids

Fig. 3 shows a set of 3D reference girds, which are used to conveniently find the 3D coordinates of the points of man-made objects in a scene. Fig. 3a is a rectangular reference plane used to measure coplanar object points. The reference grid in Fig. 3b has two orthogonal planes, which are used to refine the focal length and obtain the 3D point positions on a wedge with two orthogonal planes. Fig. 3c consists of three perpendicular planes used to recover cuboid objects. Sometimes the reference grid in Fig. 3c is more efficient than the combination of the reference grids in Fig. 3a and Fig. 3b. Fig. 3d shows a concentric disc, which can be moved along its axis. It is used to measure the height of a cylinder or cone and the radius of a circle. Fig. 3e is a combination of Fig. 3a and Fig. 3d, both on

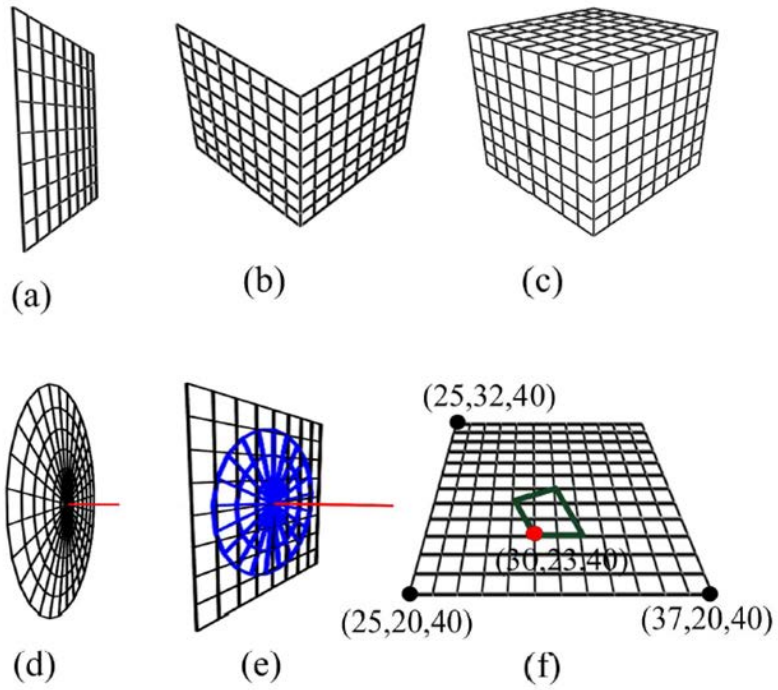

Fig. 3. 3D reference grids for measurement. (a) is a rectangular reference plane used to measure coplanar object points. (b) has two orthogonal planes, which are used to refine the focal length and obtain the 3D point positions on a wedge with two orthogonal planes. (c) is the combination of the reference grids in Fig. 3a and Fig. 3b. (d) is a concentric disc, which can be moved along its axis. It can be used to measure a surface of revolution (SOR). (e) is a combination of (a) and (d), both on the same plane. The concentric disc in (e) can be moved on the plane. It can be used to obtain the radii and center locations of circles. (f) shows the position of an arbitrary point on a plane can be obtained.

the same plane. The concentric disc can be moved on the plane. Fig. 3e can be used to obtain the radii and center locations of circles.

All these grids can be rotated, translated, and zoomed by the user, and the density and size of the grids can also be adjusted. To facilitate the measurement, we also develop an auxiliary measurement point which can be moved with a variable step size along the surface of a reference grid in 3D space by the user (as shown in Fig. 3f). The current 3D position of the point can be updated immediately when the user moves it, with the help of the known 3D location of the reference grid. Once the point is moved to coincide with an object vertex, the system obtains the precise 3D position of the vertex.

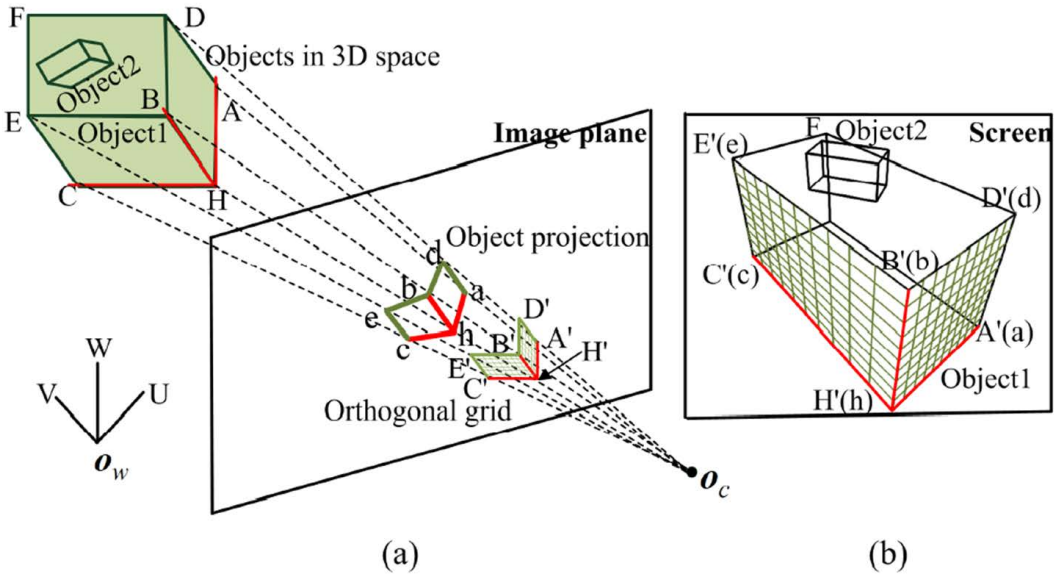

Fig. 4. 3D measurement with reference grids

3.2 3D Measurement

We take the reference grids in Figs. 3a and b as an example to explain how to obtain 3D coordinates of the object's vertices in Fig. 4. In Section 2, we have described how to establish the measurement environment and the 3D world coordinate system. Let the origin and three axes of this system be O_w, U, V, and W, respectively. Suppose that a rectangular block (Object1 in Fig. 4) is located on the plane spanned by U and V, i.e., $\overline{O_wU}$, $\overline{O_wV}$, $\overline{HA}$, and $\overline{HC}$ are on the same plane in 3D space. Then the 3D coordinates of H, A, and C can be obtained easily with the 3D reference grid in Fig. 3a and the auxiliary measurement point (as shown in Fig. 3f) when the grid is put on the same plane. (Actually, all the points on this plane can be measured.) It should be mentioned again that all the measurements are up to a scale.

Next we use the 3D reference grid in Fig. 3b to locate the 3D coordinates of D, B, and E of Object1. As shown in Fig. 4a, we first move the corner H' to

make it coincide with h (the projection of H), and rotate this grid such that its three axes $\overline{H'A'}$, $\overline{H'C'}$ and $\overline{H'B'}$ coincide with the three projected edges $\overline{ha}, \overline{hc}$ and $\overline{hb}$, respectively. Then, we can conveniently find the 3D coordinates of D, B and E by adjusting the density of the grid lines and the size of the grid. For example, the 3D coordinate of D is the same as that of D' when D' and d coincide.

If there is another object (Object2) located on Object1 as shown in Fig. 4, we can use the same method to obtain the 3D coordinates of Object2's vertices, because all the 3D points on the top of Object1 can be measured when the 3D locations of $\overline{BD}$ and $\overline{BE}$ are known.

3.3 Recovering Curves

Our system can also reconstruct objects with curved surfaces. As shown in Fig. 5a, $S1$ is a curved face, the wireframe of which consists of two curves l_1 and l_2, and two vertical lines. Using the reference grid in Fig. 3a and the auxiliary measurement point, we can obtain the 3D coordinates of several points on l_1 and l_2. Based on these 3D coordinates, we use Bezier curves to approximate l_1 and l_2 to obtain two smooth curves. Similarly, we can also find the 3D curves of the four circles in Fig. 5b.

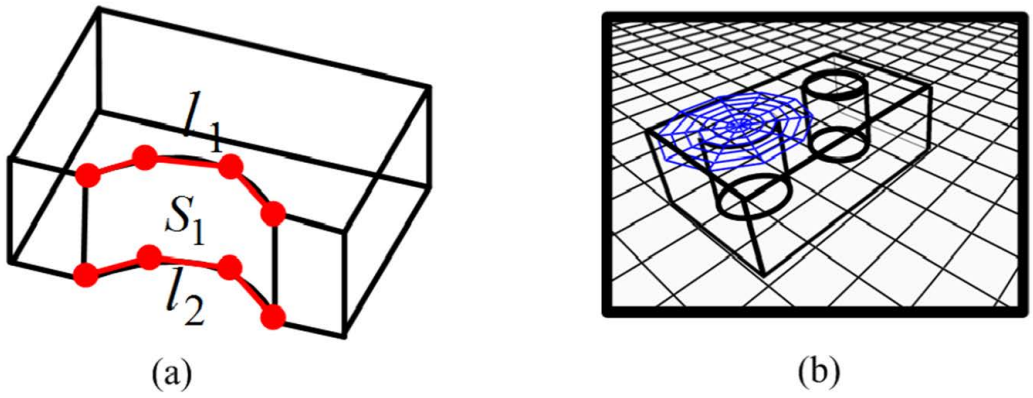

Fig. 5. 3D measurement with reference grids

4 Face Identification and Surface Generation

After the system obtains the wireframes of the 3D objects, the next step is to identify the circuits that represent the object faces and fill in them with surface patches. Face identification is not a trivial problem, and many methods have been proposed to find faces from a line drawing [17], [18], [19]. In this paper the algorithm proposed in [17] is used to detect the faces. In our interactive system, the user can also identify a face by selecting the edges of the face, which can correct possible wrongly detected faces by the algorithm in [17].

Generating a planar surface with its given 3D vertices is an easy task. For a curved face, after the system fits its curved edges with Bezier curves, the face is filled in with bilinearly blended Coons patches [20].

5 Experiments and Discussion

In this section, we show some experimental results to demonstrate the performance of our system. The most important step of our method is obtaining the 3D wireframes of objects, which depends on accurate 3D point localization.

Table 1 shows the precision of the reconstructed objects in Fig. 6a. We give the comparison between the measured result and the ground truth. The ground truth data are obtained manually when the scene is constructed. $L_g : W_g : H_g$ and $L_m : W_m : H_m$ are respectively the ratios of the length, width, and height of the ground truth and the measured result for an object. The error is obtained by setting

$$\alpha = L_g/L_m, \quad W_m^{'} = \alpha W_m, \quad H_m^{'} = \alpha H_m, \tag{2}$$

and computing

$$error = \frac{\sqrt{(W_m^{'} - W_g)^2 + (H_m^{'} - H_g)^2}}{\sqrt{W_g^2 + H_g^2}}. \tag{3}$$

From Table 1, we find that the measured ratios of the objects are very close to the ground truth. Even though there are small errors, they are still within an acceptable range. If the density of the grid lines is increased and the moving step size of the auxiliary measurement point is decreased, the errors can be reduced. It should be mentioned that most previous methods usually give only rough 3D reconstructed objects without precision shown.

We have conducted a number of experiments on real images to verify the effectiveness and precision of our system. Due to the space limitation, only some of them are given here. In Fig. 6, the first column shows the original images, and the other two columns show the reconstructed 3D objects with some with texture mapped, each in two views. Figs. 6a and b have objects with only planar faces and all the hidden edges are also drawn, the objects in Figs. 6c, d and e consist of both planar and curved faces and only the visible edges are drawn. These results show that, compared to previous methods like [1], [4], [5], [13],

Table 1. Comparison between the measured result and the ground truth for the objects in Fig. 6a

	$L_g : W_g : H_g$	$L_m : W_m : H_m$	Error
Book1	26.2 : 18.0 : 3.2	25 : 17 : 3	1.1%
Book2	24.2 : 18.0 : 4.7	24 : 18 : 5	0.8%
Book3	26.1 : 20.3 : 3.3	26 : 19 : 3	6.5%
Book4	23.5 : 16.3 : 3.2	24 : 17 : 3	2.5%
Box	23.0 : 10.5 : 7.8	22 : 10 : 7	3.7%

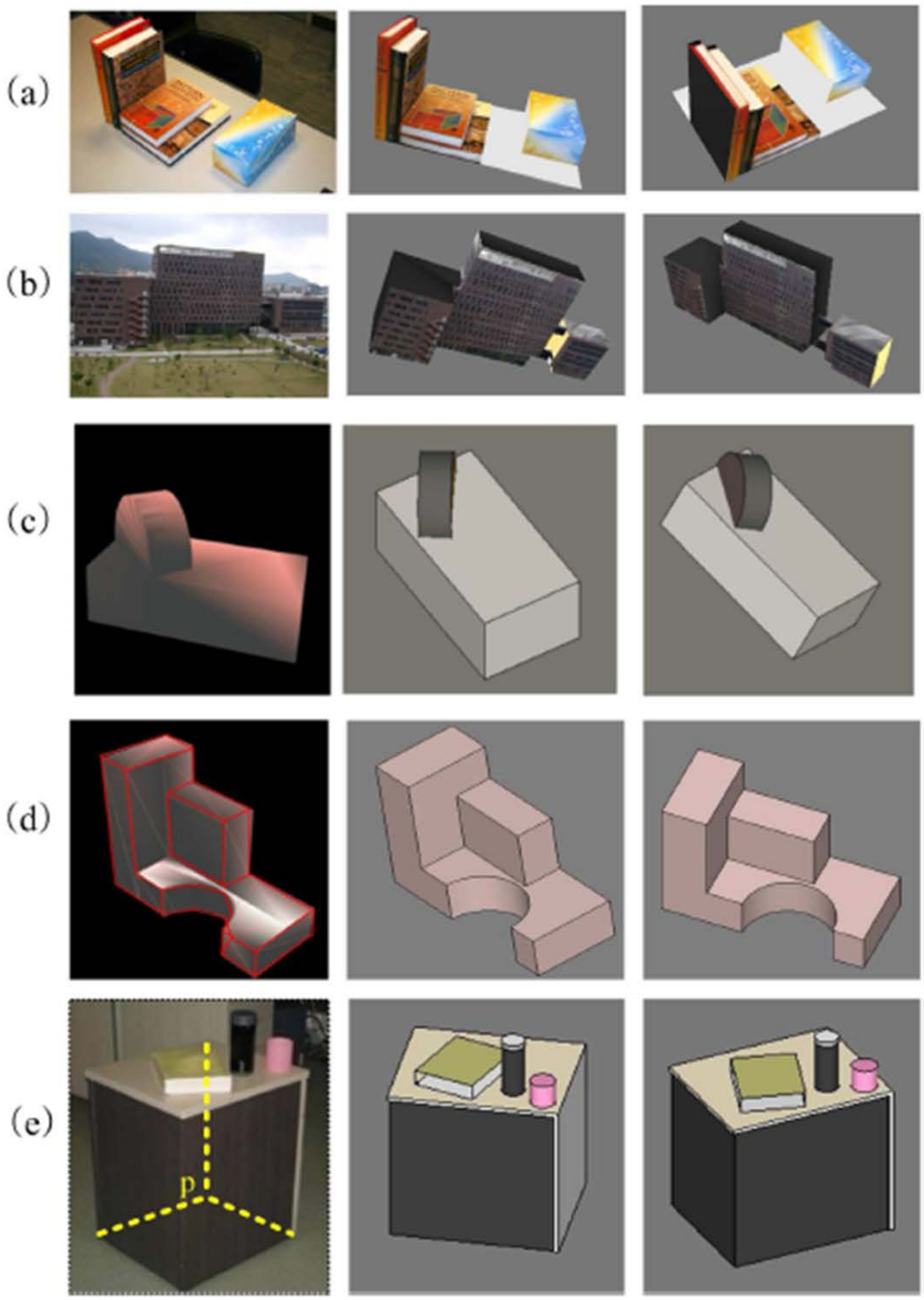

Fig. 6. Some experimental results

[6], [7], and [8], our system can precisely reconstruct complex objects with both planar and curved faces in one framework and the reconstruction time is within an acceptable range (e. g. Fig. 6c and d can be recovered within five minutes and the reconstruction time is less than the similar method in [4] and [5]).

Compared to previous approaches, which focuses on precise 3D reconstruction (like [13] which recovers planar objects with a closed form solution), our method can also recover the hidden parts precisely. As Fig. 6e shows, the

Table 2. Comparison between the measured results and the ground truth of internal diameters of the vases in Fig. 7

	Ground truth	Measured value	Error
vase1 (Fig. 7a)	$r_1 : r_2 : r_3 =$ $10.9 : 22.1 : 11.1$	$r_1' : r_2' : r_3' =$ $5 : 11 : 5$	7%
vase2 (Fig. 7b)	$L_1 : L_2 : L_3 : L_4 : L_5 =$ $10.2 : 4.4 : 10.2 : 9.1 : 11.1$	$L_1' : L_2' : L_3' : L_4' : L_5' =$ $20 : 9 : 20 : 18 : 22$	1.3%

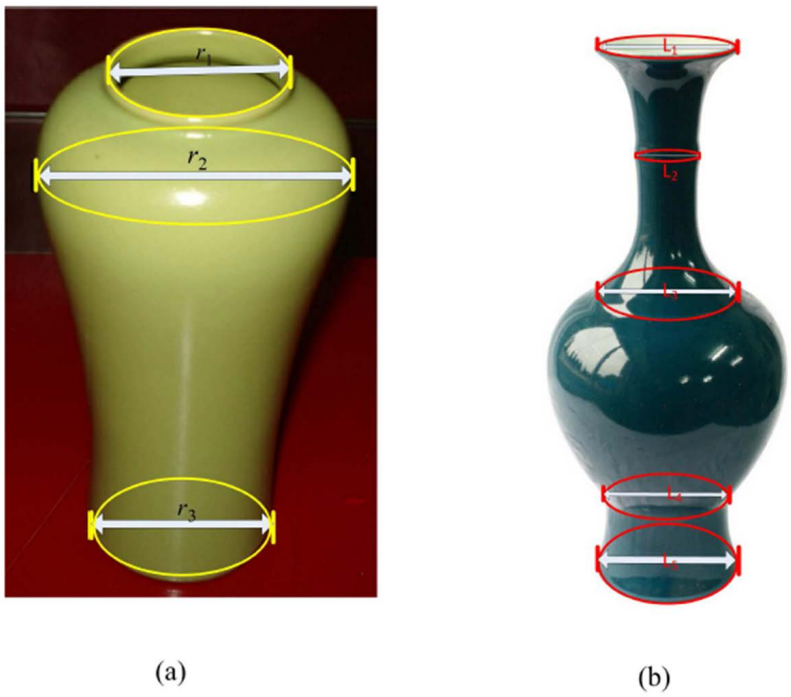

Fig. 7. Another application: measuring the geometric parameters with reference girds

hidden point P can be located accurately and sufficiently with the reference grid in Fig. 3a. Fig. 7a shows another example in which invisible parts of the cross section circles of a surface of revolution (SOR) can be recovered with the reference grid presented in this paper. Besides, Our approach overcomes the problem of accumulated reconstruction errors which exists in [6] and [7].

It should be mentioned that our method may not obtain the precise positions of the points on an irregular object. One example is the tip of a non-symmetric pyramid. Its position cannot be determined uniquely from a single view even if the bottom of the pyramid is located correctly. However, the user is still able to guess the location of the tip with a reference grid according to his/her perception (In Fig. 6e, the points of curved part are located by user's estimation).

Besides the 3D reconstruction, our method also provide another application: measuring the geometric parameters for perspective objects. Figs. 7a-b and Table 2 show a successful application in shape measurement with our method. From Table 2, we can see the internal diameters of typical parts of the two vases

can be measured precisely with the reference grid shown in Fig. 3d (Here we only demonstrate the measurement precision of our method and do not show the reconstruction results of Figs. 7a-b. Our method can also be used to recover perspective SORs since that the cross sections of a SOR can be obtained conveniently with reference grids in this paper and the focal length can be obtained by the algorithm in [21] and [22]).

6 Conclusion and Future Work

Most of existing methods of 3D reconstruction from a single image only recover rough 3D objects and the objects are often required to be interconnected. To address these problems, we have presented an efficient method using a set of reference grids to precisely locate both planar and curved 3D objects in a scene. Besides the precise 3D reconstruction, we also extended our approach to measuring the geometric parameters of perspective objects. The user interaction may be optimized and more complex 3D reference girds can be designed to facilitate the reconstruction in the future.

Acknowledgement. This work was supported by grants from Natural Science Foundation of China (60975029, 61070148), Science, Industry, Trade, Information Technology Commission of Shenzhen Municipality, China (JC200903180635A, JC201005270378A,ZYC201006130313A), Guangdong Innovative Research Team Program (No. 201001D0104648280), and the Construct Program of the Key Discipline in Hunan Province.

References

1. Criminisi, A., Reid, I., Zisserman, A.: Single view metrology. Int'l J. Computer Vision 40, 123–148 (2000)
2. Zhang, L., Dugas-Phocion, G., Samson, J., Seitz, S.: Single-view modelling of free-form scenes. The Journal of Visualization and Computer Animation 13, 225–235 (2002)
3. Prasad, M., Zisserman, A., Fitzgibbon, A.: Fast and controllable 3d modelling from silhouettes. In: Proc. Annual Conference of the European Association for Graphics (2005)
4. Debevec, P., Taylor, C., Malik, J.: Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach. In: Proc. SIGGRAPH (1996)
5. Guillou, E., Meneveaux, D., Maisel, E., Bouatouch, K.: Using vanishing points for camera calibration and coarse 3d reconstruction from a single image. The Visual Computer 16, 396–410 (2000)
6. Liebowitz, D., Criminisi, A., Zisserman, A.: Creating architectural models from images. Computer Graphics Forum 18, 39–50 (1999)
7. Sturm, P., Maybank, S., et al.: A method for interactive 3d reconstruction of piecewise planar objects from single images. In: Proc. BMVC (1999)

8. Jelinek, D., Taylor, C.: Reconstruction of linearly parameterized models from single images with a camera of unknown focal length. IEEE Trans. PAMI 23, 767–773 (2001)
9. Hong, W., Ma, Y., Yu, Y.: Reconstruction of 3-D Symmetric Curves from Perspective Images without Discrete Features. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3023, pp. 533–545. Springer, Heidelberg (2004)
10. Pavić, D., Schönefeld, V., Kobbelt, L.: Interactive image completion with perspective correction. The Visual Computer 22, 671–681 (2006)
11. Shimodaira, H.: A shape-from-shading method of polyhedral objects using prior information. IEEE Trans. PAMI 28, 612–624 (2006)
12. Jiang, N., Tan, P., Cheong, L.: Symmetric architecture modeling with a single image. ACM TOG 28, 113 (2009)
13. Li, Z., Liu, J., Tang, X.: A closed-form solution to 3d reconstruction of piecewise planar objects from single images. In: Proc. CVPR (2007)
14. Liu, J., Cao, L., Li, Z., Tang, X.: Plane-based optimization for 3d object reconstruction from single line drawings. IEEE Trans. PAMI 30, 315–327 (2008)
15. Liu, J., Chen, Y., Tang, X.: Decomposition of complex line drawings with hidden lines for 3d planar-faced manifold object reconstruction. IEEE Trans. PAMI 33, 3–15 (2011)
16. Svedberg, D., Carlsson, S.: Calibration, pose and novel views from single images of constrained scenes. Pattern Recognition Letters 21, 1125–1133 (2000)
17. Liu, J., Lee, Y., Cham, W.: Identifying faces in a 2d line drawing representing a manifold object. IEEE Trans. PAMI 24, 1579–1593 (2002)
18. Liu, J., Tang, X.: Evolutionary search for faces from line drawings. IEEE Trans. PAMI 27, 861–872 (2005)
19. Liu, J., Lee, Y.: Graph-based method for face identification from a single 2d line drawing. IEEE Trans. PAMI 23, 1106–1119 (2001)
20. Farin, G.: Curves and Surfaces for Computer-Aided Geometric Design: A Practical Code. Academic Press (1996)
21. Colombo, C., Del Bimbo, A., Pernici, F.: Metric 3d reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view. IEEE Trans. PAMI 27, 99–114 (2005)
22. Colombo, C., Comanducci, D., Del Bimbo, A.: Camera Calibration with Two Arbitrary Coaxial Circles. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 265–276. Springer, Heidelberg (2006)