Geometrical constraint based 3D reconstruction using implicit coplanarities

Ryo Furukawa

Faculty of information sciences, Hiroshima City University, Japan
ryo-f@cs.hiroshima-cu.ac.jp

Hiroshi Kawasaki

Faculty of engineering, Saitama University, Japan
kawasaki@cgv.ics.saitama-u.ac.jp

Abstract

Coplanarity is a relationship of a set of points that exist on a single plane. Coplanarities can be easily observed in a scene with planer surfaces, and these types of coplanarities have been widely used for 3D reconstructions based on geometrical constraints. Other types of coplanarities that can be observed from images are those observed as cross sections of planes and scenes; for example, points lit by a line laser, or boundary points of a shadow of a straight edge. Although these types of coplanarities have been implicitly used in variations of light sectioning methods, they have not been used in an unified manner with the former types. In this paper, we describe a new 3D reconstruction method based on coplanarities and other geometrical constraints. In particular, we make use of the above two types of coplanarities in an unified manner. This enables us to reconstruct 3D scenes scanned using line lasers or shadows of straight edges observed by a partially-calibrated single camera utilizing geometrical relationships between the planes in the scenes and the planes of line lasers or the planes of shadow boundaries.

1 Introduction

If a set of points exist on a plane, they are said to be coplanar. For example, if a scene includes a planer surface, points on the surface are coplanar. A scene composed of plane structures has many coplanarities.

On the other hand, there are other types of coplanarities. In a 3D space, there exist an infinite number of coplanarities that are not explicitly observed in ordinary situations, but could be observed under specific conditions. For example, points lit by a line laser are coplanar points. Another example is a set of points on a boundary of a cast-shadow of a straight edge. These types of coplanarities are not visible until the lasers or the shadows are cast on the scene. Let us call the former types of coplanarities as *explicit coplanarities* since they can be observed as visible surfaces of the scene, and let us call the latter types as *implicit coplanarities*.

Explicit coplanarities can be observed in scenes composed of planer surfaces, and have been widely used as geometrical constraints for 3D reconstructions [11, 9, 1, 10, 7, 6].

In vision researches, implicit coplanarities have been used in variations of light sectioning methods. In most of these researches, the light planes are first calibrated using some kind of calibration objects [3, 4, 5] and then the points on the laser planes are reconstructed using triangulation. In these researches, implicit coplanarities were not treated as geometrical constraints that can be solved by themselves to reconstruct 3D structures.

In this paper, we describe a new 3D reconstruction method based on coplanarities and other geometrical constraints. In our method, we use both two types of coplanarities in an unified way to reconstruct projective 3D information. This enables us to reconstruct projective 3D scenes with curved surfaces by using implicit coplanarities obtained by scanning the scenes with line lasers, or shadows of straight edges, observed by a partially-calibrated single camera.

Although coplanarities play an important role for shape reconstruction, it is known that the the geometrical constraints other than coplanarities (such as orthogonalities or parallelisms) are needed to achieve Euclidean reconstructions of the scenes[11]. Because of the unified treatment of both the implicit and explicit coplanarities, we can use geometrical constraints between both types of planes without discriminating them to achieve Euclidean reconstructions. This widens the applicabilities of our method. For example, a scene with curved surfaces can be densely reconstructed either by scanning the scene with a projector composed of two line lasers and utilizing geometrical constraints between the line lasers, or by scanning the scene with a single line-laser projector and utilizing geometrical constraints found in the scene (such as orthogonalities of the surfaces of the objects).

2   Related studies

Explicit coplanarities have been used in analysis of line drawings or 3D reconstructions based on geometrical constraints [11, 9, 1, 10, 7, 6]. In those studies, only scenes with planer surfaces are targeted, because they use only visible coplanarities and geometrical constraints that exist for those planer surfaces.

In computer vision researches, implicit coplanarities have been used, although unconsciously, in light sectioning methods. Recently, several researchers developed handheld 3D scanners based on light sectioning methods [3, 4, 5]. In these methods, the laser planes are calibrated by using calibration objects such as fixed frames, markers, or known planes, then the points on the laser (shadow) planes are reconstructed using triangulations. Bouguet *et al.* proposed a method in which the scene is scanned by shadows of a straight edge to reconstruct the scene [2]. Their technique requires calibration of camera parameters, a light source position, and a reference plane. Implicit coplanarities in these works are only planes for triangulations, and they should be calibrated first by using some calibration objects(known frames, markers, or planes). In contrast to these methods, our method does not require any special calibration objects.

3   Shape reconstruction from coplanarities

Reconstruction in the proposed method is realized by solving the simultaneous equations constructed from both the coplanarities and the metric constraints. As described later, metric constraints are formulated with nonlinear equations, whereas coplanarity
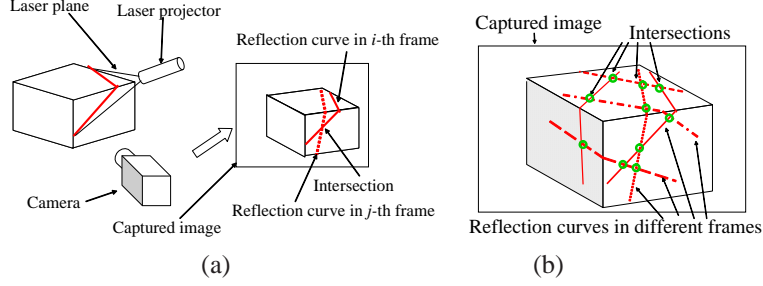
Figure 1: (a)An example configuration of the system. (b)Points of intersections.

constraints can be described by linear equations. Therefore, our method first solves linear simultaneous equations achieving projective reconstruction, and upgrades the solution to the Euclidean space.

3.1 Projective reconstruction

An example of our system consists of a camera and a line laser projector, as shown in Figure 1(a). The focal length of the camera may be unknown. The line laser beam from the projector is reflected at the surfaces of the scene and detected by the camera. These points are implicit-coplanar. A scanning process is performed by capturing a sequence of images with the camera while moving the projector back and forth. Scanning can also be performed by moving a cast shadow of a straight edge over the scene. Multiple reflection curves are obtained from the image sequence since they move in the image with the motion of the projector. The problem to be solved is the estimation of the positions of the projected laser planes from the observed implicit coplanarities. By drawing all the reflections in different frames in a image, those curves have intersections (Figure 1(b)). We can obtain geometrical constraints of coplanarities from these intersections since each of those points exists on multiple planes.

Suppose a set of N planes including both implicit and explicit planes. Let j-th plane of the set be π_j. We express the plane π_j by the form

$$a_j x + b_j y + c_j z + 1 = 0 \tag{1}$$

in the camera coordinates system.

Suppose a set of points such that each point of the set exists on intersections of multiple planes. Let the i-th element of the set be represented as ξ_i and exist on the intersection of π_j and π_k. Let the coordinates (u_i, v_i) be the location of the projection of ξ_i onto the image plane. We represent the camera intrinsic parameter by $\alpha = p/f$, where f is the focal length and p is the size of the pixel. We define $a_j^* = \alpha a_j$ and $b_j^* = \alpha b_j$. The direction vector of the line of sight from the camera to the point ξ_i is $(\alpha u_i, \alpha v_i, -1)$. Thus,

$$a_j(-\alpha u_i z_i) + b_j(-\alpha v_i z_i) + c_j(z_i) + 1 = 0, \tag{2}$$

where z_i is the z-coordinate of ξ_i. By dividing the form by z_i and using the substitutions of $t_i = 1/z_i$, $a_j^* = \alpha a_j$, and $b_j^* = \alpha b_j$, we get

$$-(\alpha u_i)a_j^* - (\alpha v_i)b_j^* + c_j + t_i = 0. \tag{3}$$

Since ξ_i is also on π_k,

$$-(\alpha u_i)a_k^* - (\alpha v_i)b_k^* + c_k + t_i = 0. \tag{4}$$

From the forms (3) and (4), the following simultaneous equations with variables $a_j^*, b_j^*, c_j, a_k^*, b_k^*$ and c_k can be obtained:

$$-u_i a_j^* + u_i a_k^* - v_i b_j^* + v_i b_k^* + c_j - c_k = 0 \tag{5}$$

We define $\mathbf{L}$ as the $M \times 3N$ coefficient matrix of the above simultaneous equations, and $\mathbf{x} = (a_0^*, b_0^*, c_0, a_1^*, b_1^*, c_1, \cdots, a_{N-1}^*, b_{N-1}^*, c_{N-1})^\top$ as the solution vector for all the M intersections and the N planes. Then, the equations can be described by a matrix form as

$$\mathbf{L}\mathbf{x} = \mathbf{0}. \tag{6}$$

Simultaneous equations of forms (5) have trivial equations that satisfy

$$a_j^* = a_k^*, b_j^* = b_k^*, c_j = c_k, (i \neq j). \tag{7}$$

Let $\mathbf{x}_1$ be the solution of $a_i^* = 1, b_i^* = 0, c_i = 0 (i = 1, 2, \ldots)$, $\mathbf{x}_2$ be the solution of $a_i^* = 0, b_i^* = 1, c_i = 0$, and $\mathbf{x}_3$ be the solution of $a_i^* = 0, b_i^* = 0, c_i = 1$. Then, the above trivial solutions form a linear space spanned by the bases of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, which we represent as T.

We describe a numerical solution of the simultaneous equations assuming the observed coordinates (u_i, v_i) on the image plane include errors. Since the equation (6) is over-constrained, the equation generally cannot be fulfilled completely. First, we consider the n-dimensional linear space S_n spanned by the n eigenvectors of $\mathbf{L}^\top \mathbf{L}$ associated with the n minimum eigenvalues. Then, S_n becomes the solution space of $\mathbf{x}$ such that $\max_{\mathbf{x} \in S_n} |\mathbf{L}\mathbf{x}|/|\mathbf{x}|$ is the minimum with respect to all possible n-dimensional linear spaces.

Even if coordinates of u_i, v_i are perturbed by additive errors, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ remain trivial solutions that completely satisfies equations(5) within the precision of floating point calculations. Thus, normally, the 3D space S_3 becomes equivalent with the space of trivial solutions T. For non-trivial solution, we can define a unit solution $\mathbf{x}_s = \operatorname{argmin}_{\mathbf{x} \in T^\perp} (|\mathbf{L}\mathbf{x}|/|\mathbf{x}|)^2$, where $T^\perp$ is the orthogonal complement space of T. $\mathbf{x}_s$ is the solution that minimizes $|\mathbf{L}\mathbf{x}|/|\mathbf{x}|$ and is orthogonal to $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_3$. Since T and S_3 are normally equal, $\mathbf{x}_s$ can be calculated as the eigenvector of $\mathbf{L}^\top \mathbf{L}$ associated with the 4-th minimum eigenvalue.

Thus, the general form of the non-trivial solutions are represented as

$$\mathbf{x} = f_1 \mathbf{x}_1 + f_2 \mathbf{x}_2 + f_3 \mathbf{x}_3 + f_4 \mathbf{x}_s = \mathbf{M}\mathbf{f}, \tag{8}$$

where f_1, f_2, f_3, f_4 are free variables, $\mathbf{f}$ is a vector of $(f_1 \; f_2 \; f_3 \; f_4)^\top$, and $\mathbf{M}$ is a matrix of $(\mathbf{x}_1 \; \mathbf{x}_2 \; \mathbf{x}_3 \; \mathbf{x}_s)$. The four DOFs of the general solution basically correspond to the DOFs of generalized projective bas-relief (GPBR) transformations described in the work of Kriegman *et al.* [8].

As far as we know, there are no previous studies that reconstruct 3D scenes by using the linear equations from the 3-DOF implicit and explicit planes. Advantages of this formulation are that the solution can be obtained stably, and the wide range of geometrical constraints can be used as metric constraints.

3.2 Euclidean reconstruction using metric constraints

The solution obtained in the previous section has four DOFs from $\mathbf{f}$. In addition, if camera parameters are unknown, additional DOFs should be resolved to achieve metric reconstruction. To achieve this, constraints other than coplanarities should be used.

For many scenes, we can find geometrical constraints among explicit and implicit planes. Examples of such information are explained here.
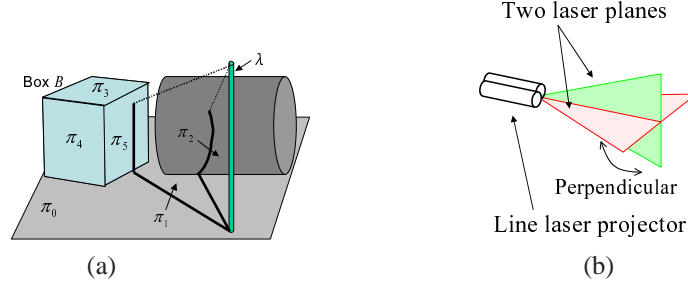
Figure 2: Metric constraints of coplanarities in a scene:(a) Rectangular box with shadow of straight bar. $\pi_0 \perp \pi_1$ and $\pi_0 \perp \pi_2$ if $\lambda \perp \pi_0$. $\pi_3 \perp \pi_4$, $\pi_4 \perp \pi_5$, $\pi_3 \perp \pi_5$, and $\pi_3 \parallel \pi_0$ if box B is rectangular and on π_0. (b) A laser projector with two line lasers.

1. In figure 2(a), the ground is plane π_0, and linear object λ is standing vertically on the ground. If the planes corresponding shadows of λ are π_1 and π_2, $\pi_0 \perp \pi_1, \pi_0 \perp \pi_2$ can be derived from $\lambda \perp \pi_0$.

2. In the same figure, the sides of box$B$ are π_3, π_4, and π_5. If box$B$ is rectangular, π_3, π_4, and π_5 are orthogonal with each other. If box$B$ is on the ground, π_3 is parallel to π_0.

3. Figure 2(b) shows a line projector with two line lasers that are aligned by the right angle. By scanning the scene with this type of projector, orthogonalities between the implicit planes are automatically obtained.

Normally, metric constraints can be represented as nonlinear equations using the free variable vector $\mathbf{f}$ and the unknown intrinsic parameters. To solve these nonlinear equations we use nonlinear optimization. The advantage of nonlinear optimization is that because of the freedom in the definition of the objective function, we can easily deal with many kinds of metric constraints.

To implement a stable nonlinear optimization, we propose a two step optimization. The first step involves optimizing the objective function with respect to the free variable vector $\mathbf{f}$ by using constant intrinsic parameters. The unknown intrinsic parameters are fixed to appropriate initial values in this step. The second step involves optimizing the objective function with respect to both $\mathbf{f}$ and the unknown intrinsic parameters. In many cases, the given information only allows us to reconstruct the scene up to scale. In this case, we fix one of the elements of $\mathbf{f}$, and the optimization is conducted for the rest of the variables.

The determination of the initial value of $\mathbf{f}$ may be a problem. In the experiments described in this study, the initial vector $\mathbf{f}_I$ is calculated from the initial plane parameter $\mathbf{x}_I$ by $\mathbf{f}_I = \mathbf{M}^\top \mathbf{x}_I$. Since $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and $\mathbf{x}_s$ (column vectors of $\mathbf{M}$) are unit and orthogonal with each other, $\mathbf{M}\mathbf{f}_I = \mathbf{M}\mathbf{M}^\top \mathbf{x}_I$ can be considered as the projection of the $\mathbf{x}_I$ (the vector of initial plane parameters) onto the solution space of the projective reconstruction (8) such that the Euclidean distance between $\mathbf{x}_I$ and $\mathbf{M}\mathbf{f}_I$ is minimum. Using this process, we can obtain a set of plane parameters which fulfills the coplanarity conditions for an arbitrary set of plane parameters.

For example, suppose that the orthogonality between the planes π_s and π_t is assumed. We denote the unit normal vector of plane π_s as a vector function $\mathbf{n}_s(\mathbf{f}, \alpha) =$
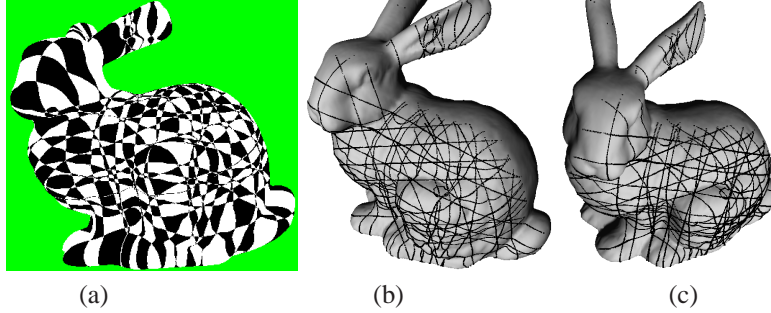
Figure 3: Reconstruction of a CG-synthesized scene: (a) the input image, and (b)(c) the reconstructed scene (the curves) with ground truth (the shaded surface).

$N((a_s(\mathbf{f}, \alpha)\ b_s(\mathbf{f}, \alpha)\ c_s(\mathbf{f}, \alpha))^\top)$ whose parameters are $\mathbf{f}$ and the camera parameter α, where $N()$ means an operation of normalization. Then, the orthogonality between π_s and π_t can be expressed as

$$\{(\mathbf{n}_s(\mathbf{f}, \alpha)\}^\top \{\mathbf{n}_t(\mathbf{f}, \alpha)\} = 0. \tag{9}$$

Another example of metric constraints is parallelism. Suppose that the planes π_s and π_t are parallel. The parallelism can be expressed as

$$\{(\mathbf{n}_s(\mathbf{f}, \alpha)\} \times \{\mathbf{n}_t(\mathbf{f}, \alpha)\} = 0. \tag{10}$$

Other than the above objective functions, we can use any functions that are described by the parameters of the points and planes and become minimum for the correct Euclidean reconstruction.

3.3 Dense reconstruction from video

After Euclidean reconstruction of sparse points, a dense 3D shape can be reconstructed by using all the captured frames. The actual process is as follows. First, we detect the intersections between a reflected curve of an unknown implicit laser plane and the curves of already reconstructed laser planes. Since the 3D positions of such intersections are known, we can estimate the parameters of the unknown plane by fitting it to the intersection 3D points using principal components analysis (PCA). We iterate the process for all frames and finally a dense 3D shape can be reconstructed.

4 Experiments

4.1 CG synthesized scene scanned by line lasers

We performed experiments on the reconstruction of 3D scenes with curved surfaces based on the implicit coplanarities. In the experiments, the nonlinear equations obtained from the metric constraints are solved using optimizations based on the Levenberg-Marquardt method.

For the first experiment, we synthesized a test data by CG as shown in 3(a), assuming a laser projector composed of two line lasers, whose laser planes are configured to be
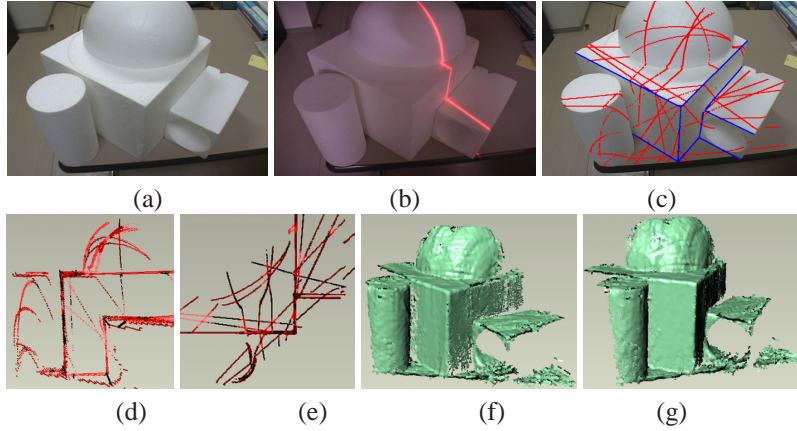
Figure 4: Reconstruction of the real scene from implicit and explicit coplanarities: (a) the target scene, (b) images used to extract the reflections of the line lasers, (c) extracted reflections(red curves) and explicit coplanarities(blue lines), (d)(e) reconstructed scene of line-lasers, and (f)(g) result of dense reconstruction .

perpendicular as shown in figure 2(b). By using the orthogonalities between the laser planes, the scene can be reconstructed without any metric constraints from the scene itself. For this scene, the cross sections of the laser planes and the model were calculated for various positions of the laser projector. The borders of the black and white patterns on the scene represent the cross sections. The images are taken 20 times, and 40 laser planes exist in the scene. The metric constraints are 20 orthogonalities between the planes. The Euclidean reconstruction was performed assuming $\alpha_u = \alpha_v, u_c = v_c = 0$. Since the scaling factor cannot be solved, we represented the solution using the average distance from the camera to the points of the model as the unit length. Using this scale, the bounding box of the ground truth points was $-0.29 \leq x \leq 0.25, -0.23 \leq y \leq 0.31, -1.34 \leq z \leq -0.93$. Figure 3(b),(c) show the solution (the curves) and the shaded ground truth model. α_u was estimated to be 7.467×10^2, whereas its true value is 7.464×10^2. The RMS of the error was 4.822×10^{-5}; therefore, the reconstruction was very accurate.

4.2 Real scene scanned by line lasers

To conduct experiments for a real object, we use a system consisting of a line laser projector and a video camera. A scanning process is performed by capturing a sequence of images with a fixed camera and moving the line laser back and forth manually. The reflections on the scene are observed as curves, and multiple curves are obtained from the image sequence. Then, we select a few images and detect the cross sections of the reflection curves. By using the points, we can reconstruct projective 3D shapes.

In the first experiment, we used a single line laser. We selected 20 images from a captured image sequence and reconstructed the 3D shape. From the scene, orthogonalities of the faces of the boxes are used as the metric constraints. Figures 4(a)–(e) show the inputs and results. We can clearly observe that the orthogonalities of the rectangular box and the parallelisms of the edges are successfully reconstructed. Then, we conducted a dense 3D reconstruction by using all the captured frames. Figures 4(f) and (g) show the
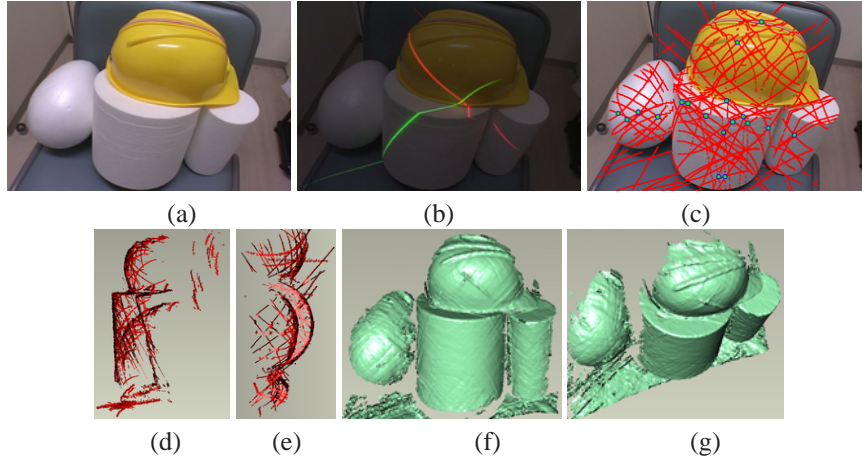
Figure 5: Reconstruction of the real scene from implicit coplanarities: (a) the target scene, (b) images used to extract the reflections of the line lasers, (c) extracted reflections (red curves) and cross sections where metric constraints are imposed(green points), (d)(e) reconstructed line-lasers, and (f)(g) result of dense reconstruction .

recovered dense 3D points. We can confirm that a dense reconstruction with an arbitrary shaped object was achieved.

Next, we built a special laser projecting device consisting of two line lasers that were aligned precisely at 90° as shown in Figure 2(b). In this case, no metric constraints were required from the scene. We selected 23 images and reconstructed the 3D shape. We also conducted a dense reconstruction. Figures 5(a)–(g) show all the inputs and results. We can see that an arbitrary shape is successfully reconstructed.

4.3 Real scene reconstruction from shadows of static objects

We conducted a shape reconstruction from images acquired by an outdoor fixed uncalibrated camera. Images from the camera were captured periodically and a shape and the focal length of the camera was reconstructed by the proposed technique from shadows in the scene. Since the scene also contained many shadows generated by non-straight edges, the automatic extraction of shadows based on background subtraction technique was difficult, and thus these noises were eliminated by human interactions. The figure 6 (a) shows the input frame, (b) shows the detected coplanar shadow curves, (c) shows all the coplanar curves and their intersections, and (d) to (f) show the reconstruction result. The proposed technique could correctly reconstruct the scene by using images from a fixed camera.

4.4 Real scene reconstruction from active scan by cast shadows

Next, we conducted an indoor experiment on an actual scene by using a point light source. A video camera was directed toward a target scene of an object of a ceramic jug shaped like a cock and multiple boxes. The target scene was captured to obtain a series of images
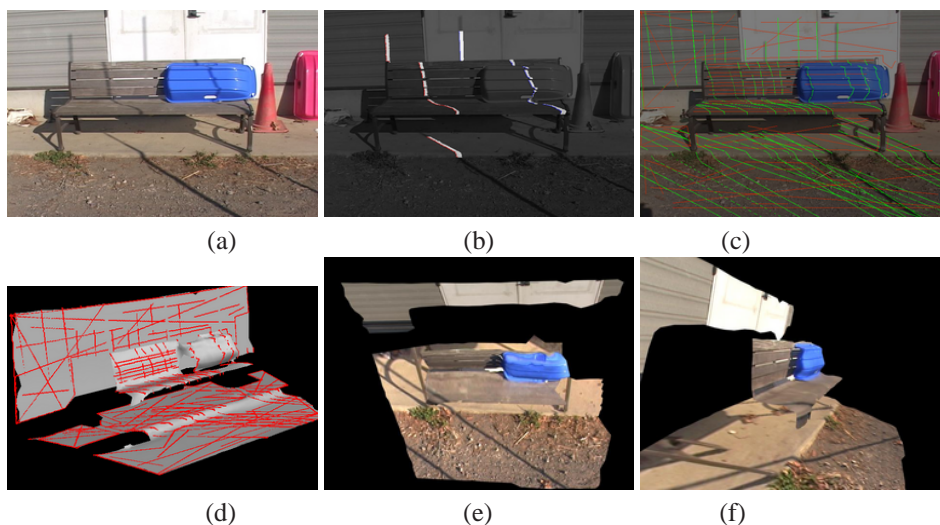
Figure 6: Reconstruction of outdoor scene: (a) input image, (b) a frame of the 3D segmentation result, (c) implicit (green) and explicit (red) coplanar curves, (d) reconstructed result of coplanar curves(red) and dense 3D points(shaded), and (e)(f) the textured reconstructed scene.

while the light source and the bar for shadowing were being moved freely. From the series of images, several images were selected and curves created by the shadow were detected from the images. By using detected coplanar shadow curves, we performed the 3D reconstruction up to 4 DOFs. For the metric reconstruction, orthogonalities of faces of the boxes were used.

Figures 7 (a)-(f) show the capturing scenes and the reconstruction result. In this case, since there were only small noises extracted because of indoor environment, shadow detection based on background subtraction technique worked well and no human interaction was required. The side orthogonalities of the rectangular box and the coplanarities of points on each plane are well reproduced. Unlike 3D photography, the proposed technique realizes reconstruction even if both the light source and the bar are moved freely.

5 Conclusion

In this paper, we propose a novel 3D reconstruction method that utilizes both the coplanarities of points lit by line lasers or those on the boundaries of shadows of straight edges. For obtaining a solution, we first obtain a projective reconstruction by solving the linear equations that are derived from the coplanarity constraints. Then, to upgrade the projective solution to the Euclidean space, we solve the nonlinear equations formulated from the metric constraints, using a nonlinear optimization method. We can use geometrical constraints such as orthogonalities and parallelisms among both the real surfaces and the laser (shadow) planes. By implementing the technique and conducting an experiment using simulated and real images, correct and dense shape reconstruction could be achieved.
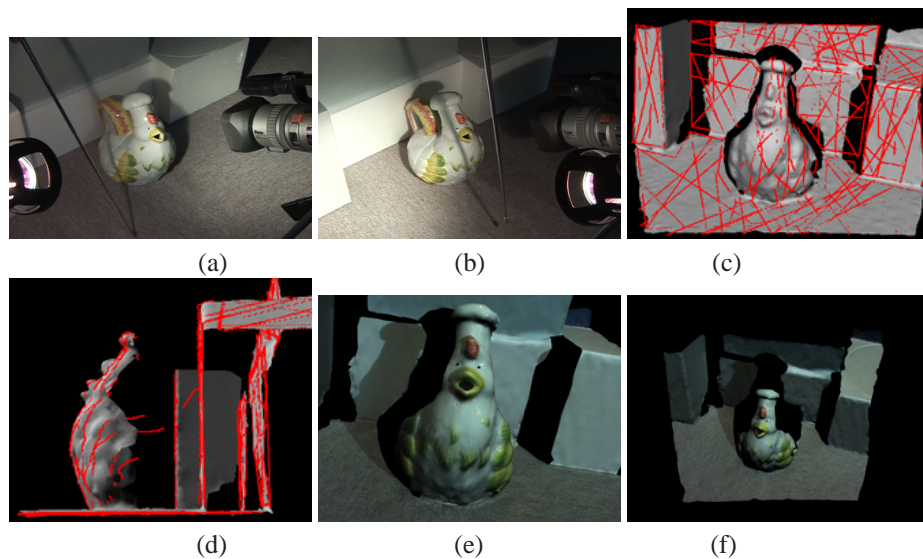
Figure 7: Reconstruction of an indoor real scene: (a)(b) the capturing scenes, (c)(d) the reconstructed coplanar shadow curves (red) with dense reconstructed model(shaded), and (e)(f) the textured reconstructed model.

References

[1] Didier Bondyfalat, Bernard Mourrain, and Theodore Papadopoulo. An application of automatic theorem proving in computer vision. In *Automated Deduction in Geometry*, pages 207–231, 1998. 1, 2

[2] J. Y. Bouguet and P. Perona. 3D photography on your desk. In *ICCV*, pages 129–149, 1998. 2

[3] Chang Woo Chu, Sungjoo Hwang, and Soon Ki Jung. Calibration-free approach to 3D reconstruction using light stripe projections on a cube frame. In *Third Int. Conf. on 3DIM*, pages 13–19, 2001. 2

[4] R. B. Fisher, A. P. Ashbrook, C. Robertson, and N. Werghi. A low-cost range finder using a visually located, structured light source. In *Second Int. Conf. on 3DIM*, pages 24–33, 1999. 2

[5] Ryo Furukawa and Hiroshi Kawasaki. Interactive shape acquisition using marker attached laser projector. In *Int. Conf. on 3DIM2003*, pages 491–498, 2003. 2

[6] E. Grossmann, D. Ortin, and J. Santos-Victor. Single and multi-view reconstruction of structured scenes. In *ACCV*, pages 93–104, 2002. 1, 2

[7] Etienne Grossmann, D. Ortin, and José Santos-Victor. Algebraic aspects of reconstruction of 3d scenes from one or more views. In *BMVC*, 2001. 1, 2

[8] David J. Kriegman and Peter N. Belhumeur. What shadows reveal about object structure. *Journal of the Optical Society of America*, 18(8):1804–1813, 2001. 4

[9] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *CVPR*, page 482, 1998. 1, 2

[10] Peter Sturm and Steve Maybank. A method for interactive 3d reconstruction of piecewise planar objects from single images. In *BMVC*, pages 265–274, Sep 1999. 1, 2

[11] Kokichi Sugihara. An algebraic approach to shape-from-image problems. *Artificial Intelligence*, 23:59–95, 1984. 1, 2