Calibration-free Approach to 3D Reconstruction
Using Light Stripe Projections on a Cube Frame

Chang Woo Chu, Sungjoo Hwang and Soon Ki Jung
1Virtual Reality Laboratory, Department of Computer Engineering
1Department of Dermatology, School of Medicine
Kyungpook National University, Daegu, Korea
cwchu@vr.knu.ac.kr, doctorhair@naver.com, skjung@knu.ac.kr

Abstract

This paper presents a new approach based on light striping for reconstructing a 3D model from a real object. The proposed system consists of a light plane projector, camera and cube frame with LEDs attached. As in other light striping systems, the correspondence problem is solved by projecting light plane onto an object inside a frame. However, we use only cross-ratios and vanishing points to set up the world coordinates of the object, while the intrinsic and extrinsic parameters of the camera and the position of light source are not required. As such, the proposed system does not calibrate the camera and the light source. Furthermore, the computed 3D point data does not require any registration process because the data is directly measured based on unified world coordinates. Experimental results proved the accuracy of the measurements and consistency of the outcomes without any knowledge of the camera and light source parameters.

1. Introduction

3D reconstruction and modeling of objects are the first steps in building many applications, such as manufacturing, virtual simulation, human-computer interaction, consumer marketing, medicine, and scientific exploration [2]. However, 3D modeling using CAD-like tools is not only tedious and time-consuming but also has limitations in expressing realism. As a result, various other methods have been proposed to gain models from real world objects. These methods can be basically classified into passive and active techniques. The former type senses visible radiation that is already present in the scene, whereas the latter controls the illumination of the scene.

There are many passive techniques including stereo, shape from shading, and structure from motion [2]. They use visual cues in images, such as parallax, shading, focus, and reflection. In particular, 3D photogrammetry deals with the extraction of highly accurate measurements from images. Most of these techniques require very precise calibration and there is almost no automation. Recently, many researchers have tried to reduce the calibration requirements and automate the acquisition process. Tomasi and Kanade proposed affine factorization method based on the assumption of orthographic projection [14]. Pollefeys [12, 11] and Liebowitz [7] proposed stratified methods using perspective projection, through which they can reconstruct objects up to an isometric scale factor. Pollefeys establishes the correspondences of the corner points from multi-views and then uses them to create a dense depth map, whereas Liebowitz uses the vanishing points, priori known angles, and length ratios from a single image. However, their approaches are only applicable to the scenes with parallel edges and corner points, such as an architectural environment. Plus the results are somewhat inaccurate.

The active methods are more accurate than the passive methods. There are also many active techniques, including radars, Moire interferometry, focusing/defocusing, and optical triangulation. Among these, light striping based on optical triangulation is commonly used in most commercial 3D scanner, such as Cyberware’s products [16]. Recently, several large-scale scanning projects have been accomplished using these systems, like Stanford’s The Digital Michelangelo Project [6], IBM’s The Pieta Project [17], and Tokyo University’s The Great Buddha Project [9]. Most of the commercial systems are accurate yet expensive because of actuators required for the precise control of the camera and light projector. Moreover, the camera and light projector must also be accurately calibrated. Many researchers have been studying calibration problems and proposed various effective methods, for example, Jokinen [5] and M’Ivor [8].

In the other hand, low-cost light striping systems are
proposed by Bouguet [1], Fisher [4] and Takatsuka [13]. Bouguet uses a camera, desk-lamp, pencil, and checkerboard. The user moves a pencil in front of the calibrated desk-lamp to cast a moving shadow on the object. The 3D shape of the object is then extracted from the spatial and temporal location of the observed shadow. The camera is calibrated using Tsai’s method [15], and the light source is calibrated using the pencil. In Fisher’s system, the position of light source is determined based on visual feedback from a specially designed wand rather than calibration. Takatsuka proposed a low-cost interactive active monocular range finder consisting of a camera and laser pointer with three LEDs attached. When the user scans the laser along the surfaces of the objects, the camera captures an image comprised of spots, thereafter triangulation is carried out using the camera’s viewing direction and the optical axis of the laser. Fisher’s and Takatsuka’s methods also calibrate camera parameters in advance.

Most systems based on light striping can only digitize the viewing surface and use a calibrated camera. As a result, in order to scan all the surfaces of an object, the camera should be controlled to suitable position and it must be calibrated again. Finally, each range data must be merged into one after an alignment or registration process.

This paper proposes a calibration-free system that includes several properties. First, the calibration of the camera and light plane projector is not required, because the light stripe is analyzed with respect to the cube frame that is a reference coordinates. Second, the range data does not need to be merged because the position of the object surface is calculated relative to the cube frame. As such, only collecting is involved. Finally, the proposed method applies several simple linear equations that use vanishing points and cross-ratio to compute the 3D coordinates for the points on a light stripe.

The rest of this paper is organized as follows. Section 2 gives an overview of the system configurations. The background notions of the projective geometry are presented in section 3. The proposed method is described in section 4 along with experimental results in section 5. The final conclusions are presented in section 6.

2. System configurations

Figure 1 shows our proposed system configurations which consist of a light plane projector, hand-held camera and cube frame with ten LEDs attached. Eight of the LEDs are attached to the corners of the cube frame in order to facilitate the detection of the cube frame. The remaining two LEDs are used to identify the origin and z-axis. They are attached to the inner and outer sides of an edge so that at least one of them is captured irrespective of the position of the camera. The light plane projector emits a line beam to the cube frame. This line beam is projected onto the objects as well as the edges of the frame. The points on the edge are called LDPs (Light plane Definition Points). The light should always be projected to meet three edges of the frame so that the resulting three non-collinear points can identify the 3D light plane uniquely without any position information on the light source. As in any other structured light range finder, the camera is positioned at a certain distance from the light plane projector, however, this does not need to be strictly controlled.

3. Backgrounds

Concepts in projective geometry [3, 10, 11], such as a vanishing point and cross-ratio, are used to calculate the 3D coordinates of the LDPs and the points on the light stripe. This section gives an overview of these concepts.

General imaging process using pin-hole camera model can be viewed as the transformation from Euclidean geometry to projective geometry. It’s the same as vision system of the human. However, in this case, the object geometry is seriously distorted by perspective image projection. Especially, under perspective projection, parallel lines do not remain parallel but instead meet at a point called the vanishing point. When a cubic object is transformed by perspective projection, there are three cases of distortion in view of the number of the vanishing points, as shown in Figure 2.

The difference in the number of the vanishing points is caused by the position and direction of the camera or eyes. In fact, the number of vanishing points is equal to the number of parallel line groups. Yet in some cases this is infinite. In the first example in Figure 2, the two vanishing points of front face are imaged to infinity. This occurs when the image plane is physically parallel with the front face of the cube. Similarly, in the second case, three vertical lines are parallel with the image plane.

Vanishing line is the set of the vanishing points, which the parallel line pairs determining the vanishing points are
prior step. The procedure is independently applied to each image, however, the recovery of the cube frame is skipped if the camera is not moved. Finally, all the 3D points of each image are collected, so that they are point clouds representing the object.

Step 1. Recover cube frame.
(1) extract the bright pixels from the input image
(2) group them into connected component
(3) classify the components into lines and points
(4) identify corner points from convex hull and vanishing points
(5) identify z-axis and origin
(6) assign 3D coordinates to each corner point

Step 2. Calculate the 3D coordinates of the LDPs.

Step 3. Calculate the 3D coordinates of the points on line stripe.

Step 4. Collect all 3D coordinates of points on line stripe.

Figure 3. Procedure of 3D coordinates calculation.

4. Calibration-free approach

An input image is converted into a binary image based on pixel intensity. The bright pixels of the image may be the corner points, LDPs, or points on the light stripe. This process may be more robust if IR (infra-red) LEDs, an IR light source, and a camera with an IR filter lens are used; however, in this study, common lens and a chip laser pointer was used and the light conditions were controlled. The bright pixels are grouped into connected components. Every component is classified into lines and points based on its size and shape. The convex hull of the point set defines the corner points of the cube frame, as shown in Figure 4. In this case, three vanishing points are defined by three pairs of lines, e.g., (p0p3, p0p4), (p0p1, p0p2) and (p0p5, p0p6) even if infinite. An inner corner point p0 can be identified by the intersection of lines, which are defined by the three vanishing points and in-between points like p6, p4, and p7.

Auxiliary points can be identified based on prior knowledge of the distance from the upper corners of the frame. As such, the z-axis and origin can be determined and any other axes can be decided using the right-handed coordinates system. Normalized 3D coordinates are assigned to each corner point of the cube frame, which will be unique with respect to the extrinsic parameters of the camera.

4.2. LDP coordinates calculation

To calculate the 3D coordinates of the LDPs, two corner points are interpolated. For example, to calculate \( P_{u} \).
position, $P_a$ and $P_f$ are used, seen in Figure 4, where $P_a$ represents the 3D coordinates for the corresponding image point $p_a$. As shown in equation (2), it is required to know the ratio of length in Euclidean geometry. But, the ratio of $p_x$ to $P_aP_f$ does not represent the accurate proportion, because the ratio of length in the image is not deserved in perspective projection.

$$P_a = P_a + \frac{T_x P_u}{T_y P_f}(P_f - P_a)$$  \hspace{1cm} (2)

The cross-ratio and vanishing points are used to know the ratio of length in Euclidean geometry. The vanishing point is an intersection of parallel lines in space when they are viewed in perspective. The cross-ratio is the ratio of length ratios. The cross-ratio has an important property in that the cross-ratio of four points on a line is preserved under projective transformations. It also remains invariant after inverse projective transformation. In Figure 5, the 3D coordinates of three LDPs are calculated from four points including vanishing points.

![Figure 5. LDP calculation.](image)

For example, to calculate $P_u$, the cross-ratios $Cr\{p_{u1}, p_{u2}; p_u, p_f\}$ and $Cr\{P_u, P_a; P_u, P_f\}$ are used, which are defined as equation (3).

$$Cr\{p_{u1}, p_{u2}; p_u, p_f\} = \frac{[p_{u1} - p_u]}{[p_{u1} - p_f]} ; \frac{[p_u - p_a]}{[p_u - p_f]}$$
$$Cr\{P_u, P_a; P_u, P_f\} = \frac{[P_{u1} - P_u]}{[P_{u1} - P_f]} ; \frac{[P_u - P_a]}{[P_u - P_f]}$$  \hspace{1cm} (3)

Because $P_{u1}$ is an infinite point, the accurate position cannot be known. However, the distances from $P_{u1}$ to $P_u$ and from $P_{u1}$ to $P_f$ are approximately the same. This allows the ratio between $P_uP_a$ and $P_uP_f$ to be calculated as equation (4).

$$\frac{P_uP_a}{P_uP_f} = \frac{[p_u - p_a]}{[p_u - p_f]} ; \frac{[p_{u1} - p_f]}{[p_{u1} - p_u]}$$  \hspace{1cm} (4)

$P_u$ can be calculated using equation (2) and (4) and the other LDPs can be done similarly. If any edge of the cube frame is parallel with the image plane, the number of vanishing points will be one or two, as the first two cases in Figure 2. In these cases, even though the number of vanishing points is not three, the same methods are applicable to calculate the coordinates of LDPs. For example, assume that $P_uP_f$ and $P_uP_f'$ are parallel. Then, $|p_{u1} - p_u|$ is approximately equal to $|p_{v1} - p_f|$ because vanishing point, $p_{u1}$, is imaged to an infinite. As such, the ratio in equation (2) can be calculated directly from the ratio in the image.

### 4.3. Light stripe calculation

Using the known corner points, vanishing points and LDPs, the coordinates of the points on the light stripe can then be calculated incrementally. To do so, additional vanishing points are necessary. In Figure 6, $p_{v2}$ is the infinite point that all lines parallel with $P_uP_v$ meet under perspective projection. It is the crossing point of all parallel lines with vanishing line $l_1$ in the image plane. Similarly, $P_{v2}$ is a vanishing point that all parallel lines with $P_uP_v$ meet by the imaging processes, which is also the intersection of $p_{v2}P_{v2}'$ and $l_2$.

In order to compute the 3D position of the points on the light stripe, certain unknowns need to be determined incrementally. These are presented in Figure 7. The most significant unknown to be aware of is the coordinate of point $P_v$. This can be calculated using the cross-ratios $Cr\{p_{v2}', p_{v1}; p_v, p_{v2}\}$ and $Cr\{P_{v2'}, P_{v1}; P_v, P_{v2}\}$ as described in section 3.2. However, $P_{v1}$ and $P_{v2}$ are still unknown. The position of $P_{v2}$ can be solved directly from the cross-ratios $Cr\{p_{v2}', p_v; p_{v2}, p_v\}$ and $Cr\{P_{v2}', P_{v1}; P_v, P_{v2}\}$. Yet, $P_{v2}$ includes more unknowns that need to be calculated beforehand, that is, $P_{v2}'$ and $P_{v1}$. The corner point occluded by the object in the cube frame, $p_b$, is estimated as the crossing point of $P_{v2}P_b$ and...
5. Experimental results

The accuracy of the proposed approach was tested from two perspectives. The first, was how similar the measured object was to the real thing. To do this, the planes of a CD box with known lengths were measured and the box was located in a known position. The second, was the consistency of the observations of one point with respect to the position of the camera. Figure 8 shows our experimental environment. Every edge of cube frame has length with 26cm and the height of the box was 13.5cm. The box was located to the distance of 8.6cm from x-axis.

5.1. Planarity of the measured plane

The accuracy of modeling was evaluated by comparing the original thing with the measured one. We tested it with the planes of the CD box. This was done by fixing camera on a tripod and then sweeping the light plane projector over the cube frame. Figure 9 shows some of the views of the rendered gauge points.

As shown in Figure 9, the measured planes seem to be planar. These results are presented numerically in Table 1. We fit a plane with planar patch of the faces of the box and estimated a standard variance of the distances from the points to the fitted plane. This represents a roughness of the plane. The results were steady with relatively small errors and a standard variance. In Table 1, we presented only the faces of the box, excluding the bottom.
5.2. Consistency of one point

One of the most important properties of our approach is that it is calibration-free. As such, the measurement of the same point must be consistent irrespective of the camera parameters, especially extrinsic ones. Therefore, one point of the input image sequences was checked while moving camera manually. The point concerned was the corner point where the front face met the upper one, as shown in Figure 8. The XYZ-time line of the point is shown in Figure 10 and the numerical values are presented in Table 2. The point remained its place with only a slight deviation.

The surface of a mannequin was also measured. When the camera is fixed in one place, only the viewed surface can be gauged, therefore, several surveys are required. However, in the proposed approach, registering the measurements for every camera position is not required. Instead, the measurements only need to be collected because the proposed method produces coordinates relative to the cube frame, the corner points of which have unique coordinates, as described in the previous section. Figure 11 shows the original mannequin and its reconstructed model.

<table>
<thead>
<tr>
<th>Table 2. Consistency of one point.</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>z</td>
</tr>
</tbody>
</table>

6. Conclusions

This paper presented a new calibration-free 3D reconstruction system composed of simple apparatuses. The proposed system includes three basic advantages. First, the camera and light plane projector do not need to be calibrated. Second, the merging of range data is not required. Finally, its calculation mechanism involves simple linear interpolation equations. These properties enable an easy scanning process. A human operator selects the camera position so that all surfaces can be scanned. Although some parts of the objects may be occluded by the edge of the cube frame, they can be digitized by changing the position of the camera. Experimental results showed that the proposed system is applicable to the 3D model reconstruction of real object.

Acknowledgements

This work was supported by Korea Research Foundation Grant(KRF-99-041-E00294) and the Korea Science and Engineering Foundation (KOSEF) through the Virtual Reality Research Center at KAIST.

References


