Teleidoscopic Imaging System for Microscale 3D Shape Reconstruction Supplementary Material

Ryo Kawahara
Kyushu Institute of Technology
rkawahara@ai.kyutech.ac.jp

Meng-Yu Jennifer Kuo
University of Minnesota
kuo00052@umn.edu

Shohei Nobuhara
Kyoto University
nob@i.kyoto-u.ac.jp

This supplemental material provides details on the depth-of-field analysis of the monocentric lens introduced in Section 4, details of the linear calibration of the mirror distances in Section 5, details of the triangulation with the teleidoscopic system we used for estimating the 3D shape in Section 6, and also additional results of microscale 3D shape recovery.

1. Monocentric Lens (Section 4)

Monocentric lens is a spherical and homogeneous optical lens which often has a high refraction index. As shown in Fig. 1, rays from a single point (blue) towards a monocentric lens diverge at a wide angle on the other side. This indicates that a perspective camera located at the point can use the monocentric lens as a conversion lens to obtain a wider field-of-view. It approximately has a short focal length as a thick lens, however, it does not have a single focus strictly [1], as illustrated as the caustic by the red lines in Fig. 1. In order to model such rays efficiently, we first model the ray-pixel mapping through the monocentric lens.

1.1. Depth-of-Field of Thin Lens Camera

Ideally all the incident light rays from a subject point to the lens focus at a common point. As shown in Fig. 2, suppose that a point $\boldsymbol{o}_{s_c}$ is focused on the image plane of a camera C through its lens L. Then the following thin lens formula for paraxial ray holds:

$$\frac{1}{s_c} + \frac{1}{t_c} = \frac{1}{f_c}. \tag{1}$$

The depth-of-field is defined as the backprojection of the permissive circle-of-confusion centered at the focused point. If the subject distance s_c is not long enough, the near and the far depth-of-focus ϵ_N and ϵ_F are given by aperture size Φ and t_c as

$$\epsilon_N = \frac{\delta t_c}{\Phi - \delta} = \frac{\delta s_c f_c}{(\Phi - \delta)(s_c - f_c)}, \tag{2}$$
$$\epsilon_F = \frac{\delta t_c}{\Phi + \delta} = \frac{\delta s_c f_c}{(\Phi + \delta)(s_c - f_c)}.$$

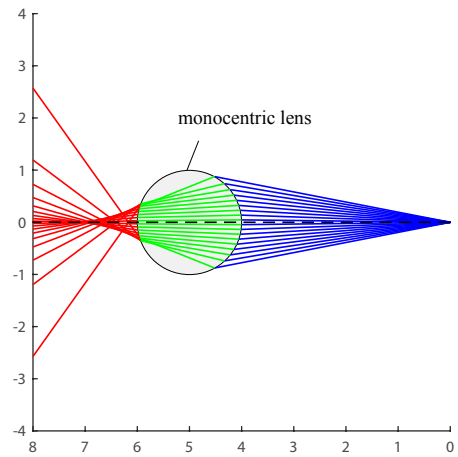

Figure 1. Refraction by Monocentric Lens. The blue, green, and red lines indicate incident, refracted, and emergent rays through a monocentric lens respectively. Notice that the emergent rays have a wider field-of-view than that of the incident rays, while they form a caustic.

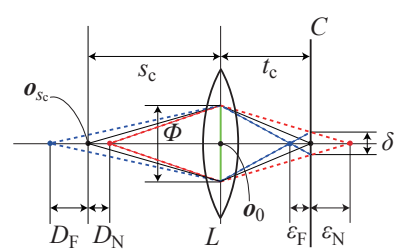

Figure 2. Depth-of-Field of Thin Lens Camera with Aperture. The red and blue lines show the backprojection of the permissible circle-of-confusion δ through the aperture Φ. D_F and D_N denote the near and the far depth-of-fields.

The near and the far depth-of-field D_N and D_F corre-

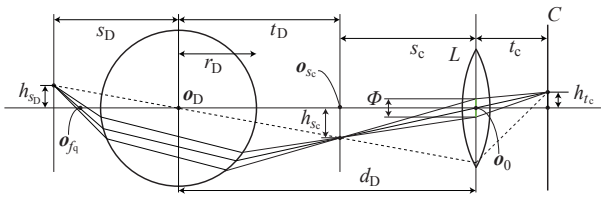

Figure 3. Light Paths through the Monocentric Lens.

sponding to ϵ_N and ϵ_F are then obtained by Eq. (1) as

$$D_N = \frac{\epsilon_N (s_c - f_c)^2}{f_c^2 + \epsilon_N (s_c - f_c)}, \quad D_F = \frac{\epsilon_F (s_c - f_c)^2}{f_c^2 - \epsilon_F (s_c - f_c)}. \tag{3}$$

As a result, D_N and D_F are described as follows:

$$D_N = \frac{\delta s_c f_c (s_c - f_c)}{(\Phi - \delta) f_c^2 + \delta s_c f_c}, \quad D_F = \frac{\delta s_c f_c (s_c - f_c)}{(\Phi + \delta) f_c^2 - \delta s_c f_c}. \tag{4}$$

Equation (4) indicates that the depth-of-field of camera C depends on the subject distance s_c, the aperture size Φ, and the permissible circle-of-confusion δ.

1.2. Light Paths through the Monocentric Lens

Fig. 3 illustrates the backprojection of the permissible circle-of-confusion through a thin lens and a monocentric camera. The light pass through the lens center $\boldsymbol{o}_0$. The key point is that an on-focus scene point at the distance s_D is also on-focus at the distance s_c between the two lenses.

2. Mirror Distance Calibration (Section 5)

Once the mirror normals are estimated, we can utilize the kaleidoscopic triangulation [2] to obtain linear constraints on the mirror distances d_i. That is,

$$\begin{aligned} (\boldsymbol{a}_v^{(0)} \times (H_i \boldsymbol{a}_v^{(i)}))^\top (\boldsymbol{o}_0 - \boldsymbol{t}_i) &= 0\,, \\ \Leftrightarrow (\boldsymbol{a}_v^{(0)} \times (H_i \boldsymbol{a}_v^{(i)}))^\top (\boldsymbol{0} - 2 d_i \boldsymbol{n}_i) &= 0\,. \end{aligned} \tag{5}$$

While this equation itself does not contribute to solve for d_i, the same constraint holds also for the first and the second reflections such as C_i and C_{ij} as

$$\begin{aligned} (\boldsymbol{a}_v^{(0)} \times (H_{ij} \boldsymbol{a}_v^{(ij)}))^\top (\boldsymbol{0} - \boldsymbol{t}_{ij}) &= 0 \\ \Leftrightarrow (\boldsymbol{a}_v^{(0)} \times (H_j H_i \boldsymbol{a}_v^{(ij)}))^\top (\boldsymbol{0} - 2 d_j \boldsymbol{n}_j - 2 H_j d_i \boldsymbol{n}_i) &= 0\,. \end{aligned} \tag{6}$$

By integrating Eq. (6) for $ij = \{12, 13, 21, 23, 31, 32\}$

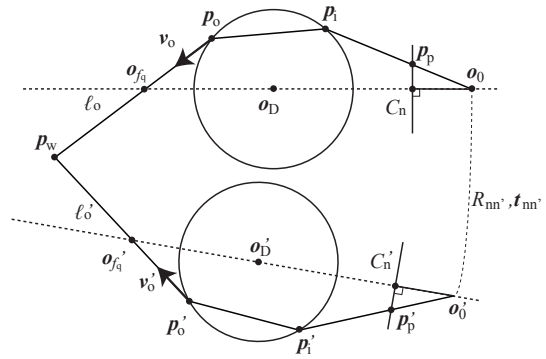

Figure 4. Teleidoscopic Triangulation. A scene point $\boldsymbol{p}_w$ is calculated as the intersection of the rays from $\boldsymbol{o}_{f_q}$ and $\boldsymbol{o}'_{f_q}$.

as a set of linear equations of d_i $(i = 1, 2, 3)$, we have:

$$\begin{pmatrix} 2h_{12}^\top H_2 \boldsymbol{n}_1 & 2h_{12}^\top \boldsymbol{n}_2 & 0 \\ 2h_{21}^\top \boldsymbol{n}_1 & 2h_{21}^\top H_1 \boldsymbol{n}_2 & 0 \\ 0 & 2h_{23}^\top H_3 \boldsymbol{n}_2 & 2h_{23}^\top \boldsymbol{n}_3 \\ 0 & 2h_{32}^\top \boldsymbol{n}_2 & 2h_{32}^\top H_2 \boldsymbol{n}_3 \\ 2h_{31}^\top \boldsymbol{n}_1 & 0 & 2h_{31}^\top H_1 \boldsymbol{n}_3 \\ 2h_{13}^\top H_3 \boldsymbol{n}_1 & 0 & 2h_{13}^\top \boldsymbol{n}_3 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = 0, \tag{7}$$

and we can obtain d_1, d_2, d_3 linearly up to scale.

3. Teleidoscopic Triangulation

Once the teleidoscopic imaging system parameters are calibrated, a scene point can be linearly triangulated in a DLT manner from its teleidoscopic projections as follows.

As shown in Fig. 4, suppose a 3D point $\boldsymbol{p}_w$ is projected to $\boldsymbol{p}_p$ in the real camera C_n and also to $\boldsymbol{p}'_p$ in a mirrored camera C'_n whose relative pose is $R_{nn'}, \boldsymbol{t}_{nn'}$. Since our ray-pixel camera model provides the direction $\boldsymbol{v}_o = (x_{v_o}, y_{v_o}, z_{v_o})^\top$ and the virtual focal length to $\boldsymbol{o}_{f_q} = (0, f_q)^\top$ directly, the collinearity constraint on ℓ_o and $\boldsymbol{p}_w = (x_{p_w}, y_{p_w}, z_{p_w})^\top$ can be expressed simply as

$$\begin{aligned} \boldsymbol{v}_o \times (\boldsymbol{p}_w - \boldsymbol{o}_{f_q}) &= 0, \\ \Leftrightarrow \begin{pmatrix} 1 & 0 & -\frac{x_{vo}}{z_{vo}} \\ 0 & 1 & -\frac{y_{vo}}{z_{vo}} \end{pmatrix} \begin{pmatrix} x_{pw} \\ y_{pw} \\ z_{pw} \end{pmatrix} &= -\frac{f_q}{z_{vo}} \begin{pmatrix} x_{vo} \\ y_{vo} \end{pmatrix}, \\ \Leftrightarrow A_0 \boldsymbol{p}_w &= \boldsymbol{b}_0. \end{aligned} \tag{8}$$

Similarly for C'_{v1}, we have

$$\begin{aligned} A_1 (R_1 \boldsymbol{p}_w + \boldsymbol{t}_1) &= \boldsymbol{b}_1, \\ \Leftrightarrow (A_1 R_1) \boldsymbol{p}_w &= \boldsymbol{b}_1 - A_1 \boldsymbol{t}_1. \end{aligned} \tag{9}$$

By combining these collinearity constraints, solving the linear system below estimates the position of $\boldsymbol{p}_w$:

$$\begin{pmatrix} A_0 \\ A_1 R_1 \end{pmatrix} \boldsymbol{p}_w = \begin{pmatrix} \boldsymbol{b}_0 \\ \boldsymbol{b}_1 - A_1 \boldsymbol{t}_1 \end{pmatrix}. \tag{10}$$

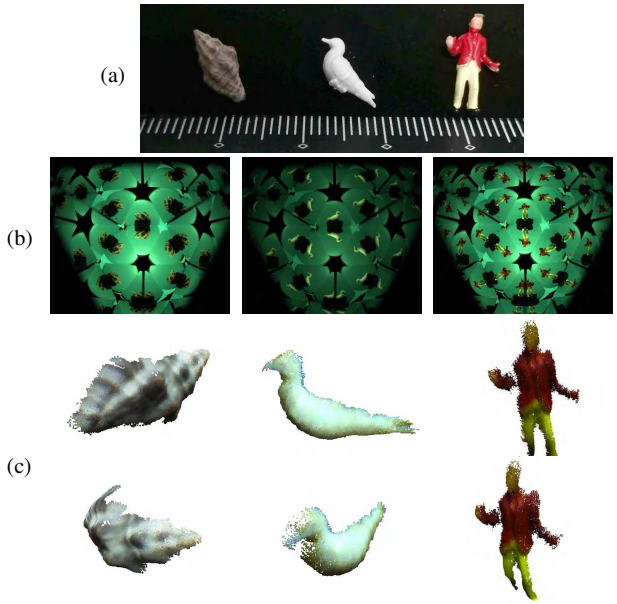

Figure 5. Additional results of 3D Shape Estimation. (a) Room-light appearance. (b) Teleidoscopic image. (c) Recovered 3D surface as point clouds with texture from two different viewpoints.

In the case of three or more viewpoints, depending on the number of the cameras, we can add equations in the same form into this system for triangulation.

4. Additional results

Fig. 5 shows the results of our teleidoscopic 3D reconstruction of various objects (a tiny seashell, and toys of a bird and a human). These results prove that our system realizes a close-up and surround-view capturing successfully.

References

[1] O. S. Cossairt, D. Miau, and S. K. Nayar. Gigapixel computational imaging. In *ICCP*, pages 1–8, 2011. 1

[2] Kosuke Takahashi, Akihiro Miyata, Shohei Nobuhara, and Takashi Matsuyama. A linear extrinsic calibration of kaleidoscopic imaging system from single 3d point. In *CVPR*, 2017. 2