Supplementary Document: Recovering Scene Geometry under Wavy Fluid via Distortion and Defocus Analysis

Mingjie Zhang[1], Xing Lin[1], Mohit Gupta[2], Jinli Suo[1] and Qionghai Dai[1]

[1]Department of Automation, Tsinghua University, [2]Columbia University

In this supplementary document, we provide detail derivations of the first-order Taylor expansions for Eq. 6 in the paper, and demonstrate the performance of our approach by using the reference images estimated by the method in [1].

1. Derivation of the Taylor Expansions

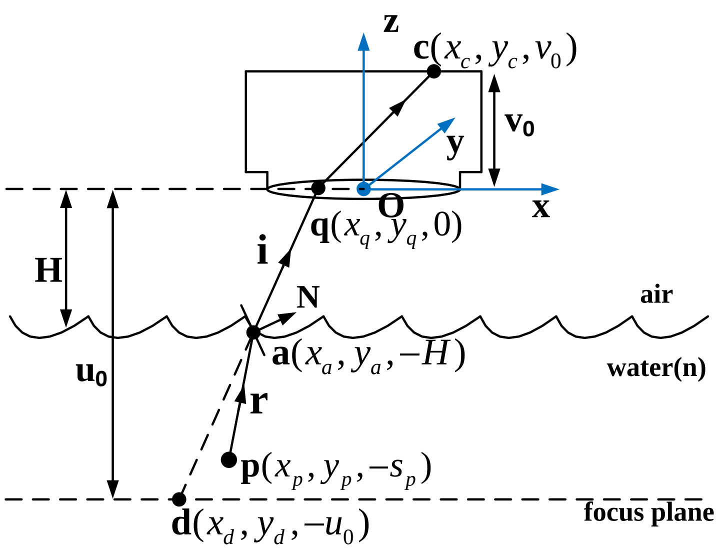


Fig. 1. The refraction blur and distortion geometry model

As illustrated in Fig. 1, the geometric function represented a ray, which emits from the immersed scene point $\mathbf{p}=(x_p, y_p, -s_p)$, passes through the water surface point $\mathbf{a}=(x_a, y_a, -H)$ and the aperture point $\mathbf{q}=(x_q, y_q, 0)$ and projects into the sensor point $\mathbf{c}=(x_c, y_c, v_0)$, can be formulated as (Eq. 6 in the paper)

$$\begin{cases} \dfrac{(x_q+\frac{u_0}{v_0}x_c)n_z(\mathbf{a})-u_0 n_x(\mathbf{a})}{\sqrt{\left(x_q+\frac{u_0}{v_0}x_c\right)^2+\left(y_q+\frac{u_0}{v_0}y_c\right)^2+u_0^2}} = n\dfrac{((1-\frac{H}{u_0})x_q-\frac{H}{v_0}x_c-x_p)n_z(\mathbf{a})-(s_p-H)n_x(\mathbf{a})}{\sqrt{\left((1-\frac{H}{u_0})x_q-\frac{H}{v_0}x_c-x_p\right)^2+\left((1-\frac{H}{u_0})y_q-\frac{H}{v_0}y_c-y_p\right)^2+(s_p-H)^2}} \\ \dfrac{(y_q+\frac{u_0}{v_0}y_c)n_z(\mathbf{a})-u_0 n_y(\mathbf{a})}{\sqrt{\left(x_q+\frac{u_0}{v_0}x_c\right)^2+\left(y_q+\frac{u_0}{v_0}y_c\right)^2+u_0^2}} = n\dfrac{((1-\frac{H}{u_0})y_q-\frac{H}{v_0}y_c-y_p)n_z(\mathbf{a})-(s_p-H)n_y(\mathbf{a})}{\sqrt{\left((1-\frac{H}{u_0})x_q-\frac{H}{v_0}x_c-x_p\right)^2+\left((1-\frac{H}{u_0})y_q-\frac{H}{v_0}y_c-y_p\right)^2+(s_p-H)^2}} \end{cases}, \quad (1)$$

where $\mathbf{N}=(n_x(\mathbf{a}), n_y(\mathbf{a}), n_z(\mathbf{a}))$ is the normal vector of the surface at the point **a**; x_* and x_* are the lateral coordinates. u_0 and v_0 are the object distance and image distance respectively; H is the distance from the camera to the flat water surface; and s_p is the depth of scene point **p**. However, Eq. 1 is too complex to analyze, thus we apply first order Taylor expansions for simplification.

Basically, the geometric function in Eq. 1 establishes the relationship between the immersed points **p** and its corresponding projected point on the sensor **c**, which can be formulated as

$$\begin{cases} g_1(x_c, y_c) = g_2(x_p, y_p, x_q, y_q, n_x(\mathbf{a}), n_z(\mathbf{a}), u_0, v_0, H, s_p, n) \\ g_3(x_c, y_c) = g_4(x_p, y_p, x_q, y_q, n_y(\mathbf{a}), n_z(\mathbf{a}), u_0, v_0, H, s_p, n) \end{cases} . \tag{2}$$

If we regard all the variables as constant except x_c, y_c, x_p *and* y_p, Eq. 2 can be expressed as

$$\begin{cases} g_1(x_c, y_c) = g_2(x_p, y_p) \\ g_3(x_c, y_c) = g_4(x_p, y_p) \end{cases} . \tag{3}$$

If the ray is parallel to the normal vector **N** of the surface at the point **a**, it will not be deflected by the water surface, and the immersed scene point and corresponding sensor point can be represented as

$$\begin{cases} \mathbf{p_0} = (x_q - \frac{n_x}{n_z} s(x_p, y_p), y_q - \frac{n_y}{n_z} s(x_p, y_p)) \\ \mathbf{c_0} = (\frac{v_0}{u_0}(\frac{n_x}{n_z} u_0 - x_q), \frac{v_0}{u_0}(\frac{n_y}{n_z} u_0 - y_q)) \end{cases} . \tag{4}$$

Then, the first-order derivative of Eq.3 at the point $\mathbf{p_0}$ is

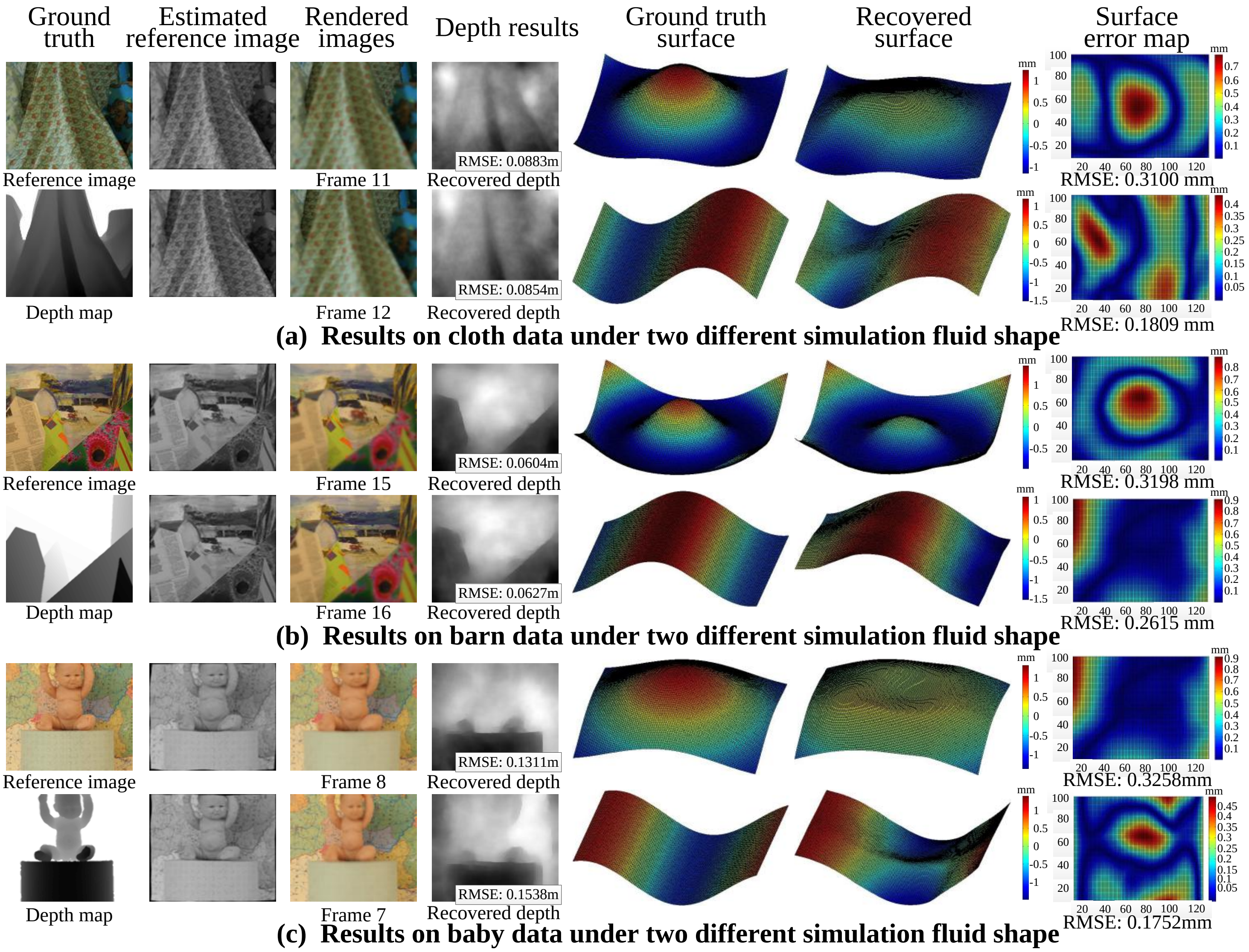
$$\begin{cases} \left.\frac{\partial x_c}{\partial y_p}\right|_{x_p = x_{p0}, y_p = y_{p0}} = \left.\frac{\partial y_c}{\partial x_p}\right|_{x_p = x_{p0}, y_p = y_{p0}} = 0 \\ \left.\frac{\partial x_c}{\partial x_p}\right|_{x_p = x_{p0}, y_p = y_{p0}} = \left.\frac{\partial y_c}{\partial y_p}\right|_{x_p = x_{p0}, y_p = y_{p0}} = -\frac{n v_0}{s(x_p, y_p) + (n-1)H} \end{cases} . \tag{5}$$

When the fluctuation of water is small or the immersed scene depth is within a certain range, the immersed scene point **p** will neighbor to the point $\mathbf{p_0}$. Thus we can apply the first-order Taylor expansion to Eq.1 at point $\mathbf{p_0}$:

$$\begin{cases} x_c \approx \frac{v_0}{u_0}(\frac{n_x}{n_z} u_0 - x_q) - (x_p - x_q + \frac{n_x}{n_z} s(x_p, y_p)) \frac{n v_0}{s(x_p, y_p) + (n-1)H} \\ y_c \approx \frac{v_0}{u_0}(\frac{n_y}{n_z} u_0 - y_q) - (y_p - y_q + \frac{n_y}{n_z} s(x_p, y_p)) \frac{n v_0}{s(x_p, y_p) + (n-1)H} \end{cases} , \tag{6}$$

which is Eq.7 in the paper.

2. Synthetic Results with the Estimated Reference Images



(a) Results on cloth data under two different simulation fluid shape

(b) Results on barn data under two different simulation fluid shape

(c) Results on baby data under two different simulation fluid shape

Fig. 2. The results of the synthetic experiments with the estimated reference image

We evaluate the performance of our approach with reference images reconstructed by [1], while keep the other experiment setttings the same to the paper. The experimental results are shown in Fig.2. In the experiments of wave 1 ($z(x,y,t) = -0.5 + 0.001\cos(\pi t\sqrt{x^2 + y^2} / 300)$ meters), the RMSE of the reconstructed scene depths are 0.0883m ("cloth"), 0.0604m ("barn") and 0.1311m ("baby"); and the average RMSE of the recovered water surfaces for three scenes are 0.2680mm ("cloth"), 0.3461mm ("barn") and 0.2980mm ("baby"), respectively.

In the experiments of wave 2 ($z(x,y,t) = -0.5 + 0.001\cos(\pi x / 60 + 9\pi t / 32)$ meters), the RMSE of the reconstructed scene depths are respectively 0.0854m ("cloth"), 0.0627m ("barn") and 0.1538m ("baby"); and the average RMSE of the recovered water surfaces for three scenes are 0.2191mm ("cloth"), 0.2128mm ("barn") and 0.3549mm ("baby"), respectively.

Since the reference image is estimated from only 17 distorted AIF frames under the same wavy water surface functions, the mean of all frames might largely deviate from the ground truth, which violates the assumption in [1]. In addition, the algorithm in [1] also has its inherent reconstruction errors. However, our method can provide the promissing water surface and scene depth reconstrution even when the reference images is not accurate.

References

[1] Oreifej, O., Shu, G., Pace, T., Shah, M.: A two-stage reconstruction approach for seeing through water. In: CVPR. pp. 1153-1160. IEEE (2011).