Refraction Wiggles for Measuring Fluid Depth and Velocity from Video

Tianfan Xue[1], Michael Rubinstein[2,1], Neal Wadhwa[1], Anat Levin[3], Fredo Durand[1], and William T. Freeman[1]

[1]MIT CSAIL [2]Microsoft Research [3]Weizmann Institute

1 Refractive Stereo: Quantitative Evaluation

We evaluate our refractive stereo results quantitatively in two ways. First, for natural stereo sequences, we compare the recovered depths of the refractive fluid layers with that of the heat sources generating them, as computed using a standard stereo algorithm (since the depth of the actual heat sources, being solid objects, can be estimated well using existing stereo techniques). More specifically, we pick a region on the heat source and pick another region of hot air right above the heat source, and compare the average disparities in these two regions. Our experiments show that the recovered depth map of the (refractive) hot air matches well the recovered depth map of the (solid) heat source, with an average error of less than a few pixels.

Second, we evaluate the refractive stereo algorithm on simulated sequences with ground truth disparity. The simulation setup is similar to one we used for the refractive flow algorithm in the paper, except that we manually specify the ground truth disparity map as shown in Figure 1.

To evaluate the performance of the algorithm, we generated four different background patterns shown in Figure 1: 1) weakly textured in both x and y directions, 2) strongly textured in x direction, 3) strongly textured in y direction, 4) strongly textured in both directions. We generated all combinations of these backgrounds in the left and right views and get $4 \times 4 = 16$ stereo sequences. For comparison, we also implemented a refractive stereo algorithm that does not consider the uncertainty in the optical flow. Figure 1(a) and (b) show the disparities recovered by both refractive stereo (without uncertainty; the baseline) and probabilistic refractive stereo (with uncertainty), together with their corresponding root-mean-square error (RMSE). The simple refractive stereo only works well when the backgrounds of both the left and right views are strongly textured in both the x and y directions. When the background is weakly textured in one direction, the optical flow in that direction is noisy and thus the recovered disparity will be significantly less accurate.

The probabilistic refractive stereo algorithm is able to handle weaker textures. As long as one direction of the background is textured in both views, the algorithm is able to accurately recover the disparity map. For example, in the second row and second column in Figure 1(b), the optical flow in the x direction is noisy, while the optical flow in the y direction is clean, because the background is very smooth in x direction. In the probabilistic refractive stereo, the algorithm weights optical flow results by how accurately they are estimated when calculating the disparity. Therefore, the optical flow

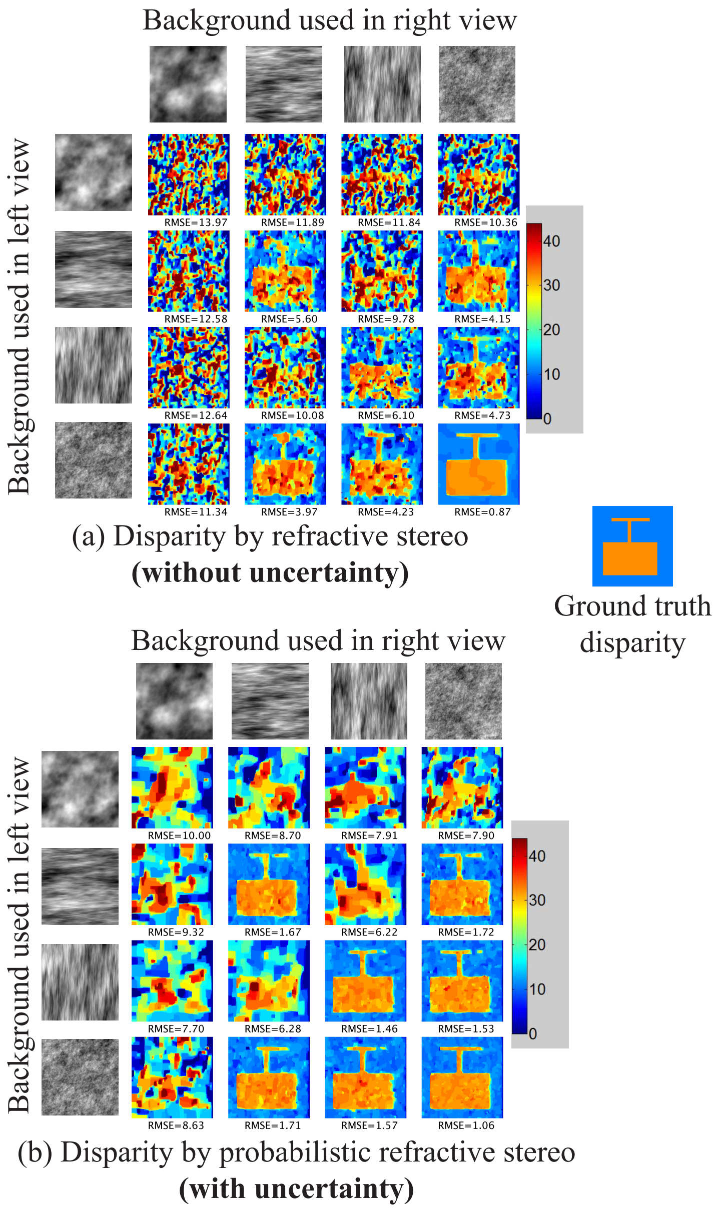

Fig. 1. Quantitative evaluation of refractive stereo using synthetic sequences.

in the x direction will have a smaller effect on the final result than the optical flow in the y direction, and the algorithm will infer the disparity map correctly. This demonstrates the robustness of our algorithm to partially textured backgrounds.

2 Calculating Fluid Flow Efficiently (Section 4)

Recall that in section 4 of the paper, the probabilistic refractive flow algorithm consists of two steps. First, we solve for the mean $\tilde{\mathbf{v}}$ and the variance Σ_v of the wiggle features $\mathbf{v}$ from the following Gaussian distribution:

$$P(\mathbf{v}|I) = \exp\left(-\sum_{\mathbf{x}} \alpha_1 \left\| \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} \right\|^2 + \alpha_2 \left\| \frac{\partial \mathbf{v}}{\partial x} \right\|^2 + \alpha_2 \left\| \frac{\partial \mathbf{v}}{\partial y} \right\|^2 \right). \tag{1}$$

To solve for the mean and variance of flow from (1), let the V be the vector formed by concatenating all the optical flow vectors in one frame. That is, $\mathbf{V} = (\cdots, \mathbf{v}(\mathbf{x}), \cdots)$. Also, let us represent (1) in information form $P(\mathbf{v}|I) = \exp(-\frac{1}{2}V^\intercal JV + \mathbf{h}^\intercal V)$, where $\mathbf{h}$ and J can be calculated from (1). Then the mean of V is $\tilde{\mathbf{V}} = J^{-1}\mathbf{h}$ and covariance of V is $\Sigma = J^{-1}$.

In the second step, the fluid flow is calculated by minimizing the following optimization problem based on the mean and variance of the wiggle features computed in the first step.

$$\tilde{\mathbf{u}} = \arg\min_{\mathbf{u}} \sum_{\mathbf{x}} \beta_1 \left\| \frac{\partial \tilde{\mathbf{v}}}{\partial x} u_x + \frac{\partial \tilde{\mathbf{v}}}{\partial y} u_y + \frac{\partial \tilde{\mathbf{v}}}{\partial t} \right\|^2_{\Sigma_v} + \beta_2 \left(\left\| \frac{\partial \mathbf{u}}{\partial x} \right\|^2 + \left\| \frac{\partial \mathbf{u}}{\partial y} \right\|^2 \right) + \beta_3 \|\mathbf{u}\|^2. \tag{2}$$

Calculating the covariance of each wiggle feature from (1) requires inverting the information matrix J. This step will be slow if the matrix is large. To avoid this time-consuming inversion, we make a slight change to the fluid flow objective function. Let $\tilde{\mathbf{V}}_x$, $\tilde{\mathbf{V}}_y$, and $\tilde{\mathbf{V}}_t$ be the vectors formed by concatenating all the partial derivatives of mean wiggle features in a frame, that is $\tilde{\mathbf{V}}_x = (\cdots, \frac{\partial \mathbf{v}}{\partial x}(\mathbf{x}), \cdots)$, $\tilde{\mathbf{V}}_y = (\cdots, \frac{\partial \mathbf{v}}{\partial y}(\mathbf{x}), \cdots)$, and $\tilde{\mathbf{V}}_t = (\cdots, \frac{\partial \mathbf{v}}{\partial t}(\mathbf{x}), \cdots)$. Similarly, let $\mathbf{U}_x$, $\mathbf{U}_y$ be the vectors formed by concatenating all the x-components and y-components of $\mathbf{u}$ in a frame respectively. Then we can calculate the refractive flow as follows:

$$\begin{aligned} \min_{\mathbf{U}} \beta_1 & \left(\tilde{\mathbf{V}}_x \cdot \mathbf{U}_x + \tilde{\mathbf{V}}_y \cdot \mathbf{U}_y + \tilde{\mathbf{V}}_t \right)^\intercal J \left(\tilde{\mathbf{V}}_x \cdot \mathbf{U}_x + \tilde{\mathbf{V}}_y \cdot \mathbf{U}_y + \tilde{\mathbf{V}}_t \right) \\ & + \beta_2 \left(\|D_x \mathbf{U}_x\|^2 + \|D_y \mathbf{U}_x\|^2 + \|D_x \mathbf{U}_y\|^2 + \|D_y \mathbf{U}_y\|^2 \right) + \beta_3 \|U\|^2 \end{aligned} \tag{3}$$

where D_x and D_y are the partial derivative matrices to x and y respectively. The smoothness term of (3) is exactly the same as that in (2), and the data term of (2) is

$$\begin{aligned} & (\mathbf{V}_x \cdot \mathbf{U}_x + \mathbf{V}_y \cdot \mathbf{U}_y + \mathbf{V}_t)^\intercal J (\mathbf{V}_x \cdot \mathbf{U}_x + \mathbf{V}_y \cdot \mathbf{U}_y + \mathbf{V}_t) \\ = & \|\mathbf{V}_x \cdot \mathbf{U}_x + \mathbf{V}_y \cdot \mathbf{U}_y + \mathbf{V}_t\|_{J^{-1}} = \|\mathbf{V}_x \cdot \mathbf{U}_x + \mathbf{V}_y \cdot \mathbf{U}_y + \mathbf{V}_t\|_{\Sigma}, \end{aligned} \tag{4}$$

which is also similar to the data term in (2) except that it jointly considers all the wiggle vectors in a frame. Therefore, this change will not affect the result too much, but the algorithm is more computationally efficient as we never need to compute J^{-1}. The term never appears in (3).

3 Probabilistic Interpretation in the Refractive Stereo (Section 5)

In this section, we will show that the data term defined in Section 5 for refractive stereo is equal to the negative log of the conditional marginal distribution. Let $\mathbf{v}_R(\mathbf{x}) \sim N(\overline{\mathbf{v}}_R(\mathbf{x}), \Sigma_R(\mathbf{x}))$ and $\mathbf{v}_L(\mathbf{x}+d(\mathbf{x})) \sim N(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})), \Sigma_L(\mathbf{x}+d(\mathbf{x})))$ be the optical flow from the left and right views, where $\overline{\mathbf{v}}_L$ and $\overline{\mathbf{v}}_R$ are the means of the optical flow and Σ_L and Σ_R are the variances of optical flow. Then the data term of the refractive stereo is

$$\begin{aligned}
&f(\mathbf{v}_R, \mathbf{v}_L)\\
=&-\log cov(\mathbf{v}_L, \mathbf{v}_R)\\
=&-\log \int_{\mathbf{v}} P_L(\mathbf{v}) P_R(\mathbf{v}) d\mathbf{v}\\
=&-\log \int_{\mathbf{v}} N(\mathbf{v}; \overline{\mathbf{v}}_L, \Sigma_L) N(\mathbf{v}; \overline{\mathbf{v}}_R, \Sigma_R) d\mathbf{v}\\
=&\frac{1}{2}\log|\Sigma_L + \Sigma_R| + \frac{1}{2}\|\overline{\mathbf{v}}_R - \overline{\mathbf{v}}_L\|^2_{\Sigma_L+\Sigma_R} + const
\end{aligned} \tag{5}$$

where $N(\mathbf{v}; \overline{\mathbf{v}}_L, \Sigma_L)$ is shorthand notation for the Gaussian distribution probability density function

$$N(\mathbf{v}; \overline{\mathbf{v}}, \Sigma) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\overline{\mathbf{v}}^\intercal \Sigma^{-1} \overline{\mathbf{v}}\right) \tag{6}$$

Recall that the optical flow calculated by the algorithm is degraded by noise. Specifically, let $\mathbf{v}(\mathbf{x})$ be the ground truth optical flow from the right view at $\mathbf{x}$. The mean optical flow from the right view (or left view) calculated by the algorithm equals the ground truth optical flow plus Gaussian noise with zero-mean and variance equal to Σ_R (or Σ_L), that is:

$$P(\overline{\mathbf{v}}_R(\mathbf{x})|\mathbf{v}(\mathbf{x})) = N(\overline{\mathbf{v}}_R(\mathbf{x}); \mathbf{v}(\mathbf{x}), \Sigma_R(\mathbf{x})), \tag{7}$$

$$P(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x}))|\mathbf{v}(\mathbf{x}), d(\mathbf{x})) = N(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})); \mathbf{v}(\mathbf{x}), \Sigma_L(\mathbf{x}+d(\mathbf{x}))) \tag{8}$$

To evaluate the probability of d, let us consider the marginal distribution

$$\begin{aligned}
&P(\overline{\mathbf{v}}_R(\mathbf{x}), \overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x}))|d(\mathbf{x}))\\
=&\int_{\mathbf{v}} P(\overline{\mathbf{v}}_R(\mathbf{x}), \overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})), \mathbf{v}(\mathbf{x})|d(\mathbf{x})) d\mathbf{v}(\mathbf{x})\\
=&\int_{\mathbf{v}} P(\overline{\mathbf{v}}_R(\mathbf{x})|\mathbf{v}(\mathbf{x})) P(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x}))|\mathbf{v}(\mathbf{x}), d(\mathbf{x})) P(\mathbf{v}) d\mathbf{v}(\mathbf{x})\\
=&\int_{\mathbf{v}} N(\overline{\mathbf{v}}_R(\mathbf{x}); \mathbf{v}(\mathbf{x}), \Sigma_R(\mathbf{x})) N(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})); \mathbf{v}(\mathbf{x}), \Sigma_L(\mathbf{x}+d(\mathbf{x}))) P(\mathbf{v}) d\mathbf{v}(\mathbf{x})
\end{aligned} \tag{9}$$

Assuming that $P(\mathbf{v})$ has an uniform prior, we have:

$$\log P(\overline{\mathbf{v}}_R(\mathbf{x}), \overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x}))|d(\mathbf{x})) \tag{10}$$

$$= \log \int_{\mathbf{v}} N(\overline{\mathbf{v}}_R(\mathbf{x}); \mathbf{v}, \Sigma_R(\mathbf{x})) N(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})); \mathbf{v}, \Sigma_L(\mathbf{x}+d(\mathbf{x}))) P(\mathbf{v}) d\mathbf{v} \tag{11}$$

$$= \log \int_{\mathbf{v}} N(\overline{\mathbf{v}}_R(\mathbf{x}); \mathbf{v}, \Sigma_R(\mathbf{x})) N(\overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})); \mathbf{v}, \Sigma_L(\mathbf{x}+d(\mathbf{x}))) d\mathbf{v} + const \tag{12}$$

$$= \log \int_{\mathbf{v}} N(\mathbf{v}; \overline{\mathbf{v}}_R(\mathbf{x}), \Sigma_R(\mathbf{x})) N(\mathbf{v}; \overline{\mathbf{v}}_L(\mathbf{x}+d(\mathbf{x})), \Sigma_L(\mathbf{x}+d(\mathbf{x}))) d\mathbf{v} + const \tag{13}$$

$$= -f(\mathbf{v}_R, \mathbf{v}_L) + const \tag{14}$$

Therefore, the data term is equal to the negative log of conditional marginal distribution (plus a constant).

4 Proofs for the Theorems in Section 3

Recall that in the paper, we define the refractive wiggle as:

Definition 1 (Refraction wiggle) *Let x_t be a point on the refractive fluid layer, $x_{t+\Delta t}$ be the intersection between the background and the light ray passing through x_t and the center of projection at time t, and Δt be a short time interval. Then the* ***wiggle*** *of x_t from time t to $t+\Delta t$ is the shift of the projection of $x_{t+\Delta t}$ on the image plane during this time.*

Then, let us prove the following lemma:

Lemma 1 (Refractive wiggle) *Let z, z', and z'' be the focal length of the camera, the depth of fluid, and the depth of background. Assuming that the fluid object is moving parallel to the camera plane, then the wiggle feature equals to:*

$$v = -\frac{(z''-z')z\sec^2\beta_t}{z''}(\alpha_{t+\Delta t} - \alpha_t) \tag{15}$$

Proof. Because wiggles are defined by shifts in the image plane, we first trace rays to determine which points in the image plane correspond to x'_t and $x'_{t+\Delta t}$ on the fluid object. At time t, an undistorted ray is emitted from the center of the projection o to point x'_t with angle β_t, with respect to vertical (solid lines in Fig. 2). This ray is bent by α_t due to the refraction when it passes through the fluid and finally hits the image plane at x''_t.

At successive time $t_i + \Delta t$, the fluid object moves to a new location (dashed gray blob in Fig. 2. The background point x''_t now correspond to a different point $x'_{t+\Delta t}$ on the fluid object and a different point $x_{t+\Delta t}$ on the image point. The ray from the background to the fluid object is now bent by $\alpha_{t+\Delta t}$ when passing through the fluid, and finally hits the image plane with angle of $\beta_{t+\Delta t}$ with respect to the vertical. The wiggle at time t is the distance between x_t and $x_{t+\Delta t}$, denoted as $\overrightarrow{x_t x_{t+\Delta t}}$ (blue or red arrows in Fig. 2. A simple geometric calculation shows:

$$x_{t+\Delta t} - x_t = z(\tan\beta_{t+\Delta t} - \tan\beta_t) \tag{16}$$

$$\begin{aligned} x''_t - o &= z'\tan\beta_{t+\Delta t} + (z''-z')\tan(\alpha_{t+\Delta t}+\beta_{t+\Delta t}) \\ &= z'\tan\beta_t + (z''-z')\tan(\alpha_t+\beta_t) \end{aligned} \tag{17}$$

From Eq. 17, we have:

$$\begin{aligned} 0 &= z'(\tan\beta_{t+\Delta t} - \tan\beta_t) + (z''-z')\Big(\tan(\alpha_{t+\Delta t}+\beta_{t+\Delta t}) - \tan(\alpha_t+\beta_t)\Big) \\ &\approx \sec^2\beta_t(\beta_{t+\Delta t}-\beta_t)z' + (z''-z')\sec^2(\beta_t+\alpha_t)(\beta_{t+\Delta t}-\beta_t+\alpha_{t+\Delta t}-\alpha_t) \\ &= \sec^2\beta_t(\beta_{t+\Delta t}-\beta_t)z' + (z''-z')\sec^2\beta_t(\beta_{t+\Delta t}-\beta_t+\alpha_{t+\Delta t}-\alpha_t), \end{aligned} \tag{18}$$

where the first line to the second line is based on the Taylor expansion, and second line to the third line is based on the assumption that α is small, because the index of refraction is close to 1. Then from Eq. 18, we have:

$$\beta_{t+\Delta t} - \beta_t = -\frac{z''-z'}{z''}(\alpha_{t+\Delta t}-\alpha_t). \tag{19}$$

Finally, combining Eq. 16 and Eq. 19, we have:

$$\begin{aligned} x_{t+\Delta t} - x_t &= z(\tan\beta_{t+\Delta t} - \tan\beta_t) \\ &\approx z(\beta_{t+\Delta t}-\beta_t)\sec^2\beta_t \\ &= -\frac{(z''-z')z\sec^2\beta_t}{z''}(\alpha_{t+\Delta t}-\alpha_t) \end{aligned} \tag{20}$$

This completes the proof.

Lemma 2

Based on this lemma, we can prove the refractive flow constancy and the refractive stereo constancy.

Theorem 1 (Refractive flow constancy). *Suppose the fluid object does not change its shape and index of refraction during a short time interval* $[t_1, t_2]$*. Then for any point on the fluid object, its wiggle* $\mathbf{v}(t_1)$ *at* t_1 *equals its wiggle* $\mathbf{v}(t_2)$ *at* t_2.

Proof. As shown in Fig. 2, let $o, x_{t_i}, x'_{t_i}, x''_{t_i}$ be the light ray that hits the background x''_{t_i} at time t, and $o, x_{t_i+\Delta t}, x'_{t_i+\Delta t}, x''_{t_i+\Delta t}$ be the light ray that hits the same background point x''_{t_i} at time $t_i+\Delta t$, $i = 1, 2$. According the definition of the refractive wiggle:

$$\mathbf{v}(t_1) = \overrightarrow{x_{t_1}x_{t_1+\Delta t}}, \mathbf{v}(t_2) = \overrightarrow{x_{t_2}x_{t_2+\Delta t}} \tag{21}$$

To prove the refractive flow constancy, first we will show that $x'_{t_1+\Delta t}$ and $x'_{t_2+\Delta t}$ are the same point on the fluid object, or equivalently, the shifts $\overrightarrow{x'_{t_1}x'_{t_1+\Delta t}}$ and $\overrightarrow{x'_{t_2}x'_{t_2+\Delta t}}$

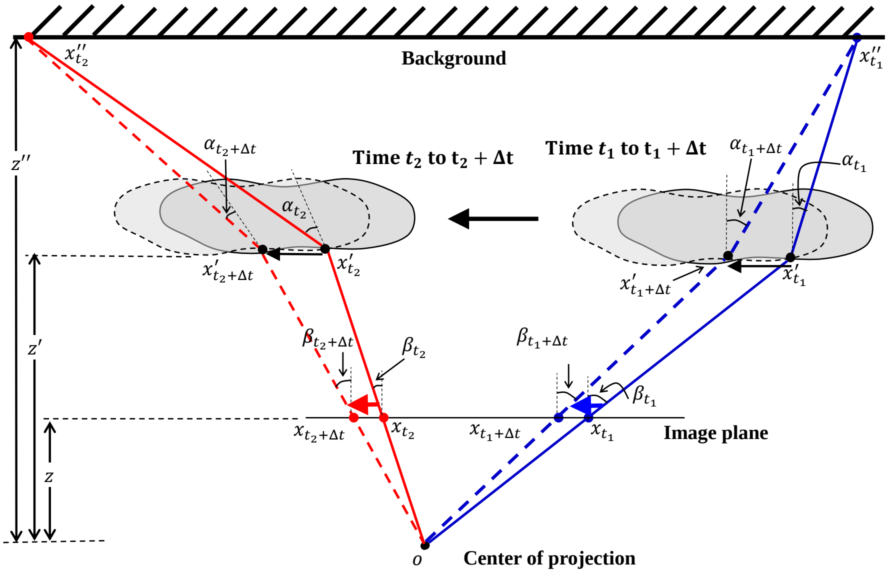

Fig. 2. Proof of refractive stereo constancy

are equal. By the definition of wiggle, we only know that x'_{t_1} and x'_{t_2} refers to the same point on the refractive object. We need to prove that $x'_{t_1+\Delta t}$ and $x'_{t_2+\Delta t}$ are equal.

Based on Lemma 1, we know that:

$$\overrightarrow{x_{t_1}x_{t_1+\Delta t}} = -\frac{(z''-z')z\sec^2\beta_{t_1}}{z''}(\alpha_{t_1+\Delta t}-\alpha_{t_1})$$
$$\overrightarrow{x_{t_2}x_{t_2+\Delta t}} = -\frac{(z''-z')z\sec^2\beta_{t_2}}{z''}(\alpha_{t_2+\Delta t}-\alpha_{t_2}) \quad (22)$$

By the similar triangle formula, we have:

$$\overrightarrow{x'_{t_1}x'_{t_1+\Delta t}} = \frac{z'}{z}\overrightarrow{x_{t_1}x_{t_1+\Delta t}} = -\frac{(z''-z')z'\sec^2\beta_{t_1}}{z''}(\alpha_{t_1+\Delta t}-\alpha_{t_1}), \quad (23)$$
$$\overrightarrow{x'_{t_2}x'_{t_2+\Delta t}} = \frac{z'}{z}\overrightarrow{x_{t_2}x_{t_2+\Delta t}} = -\frac{(z''-z')z'\sec^2\beta_{t_2}}{z''}(\alpha_{t_2+\Delta t}-\alpha_{t_2}). \quad (24)$$

As discussed in the paper, for a short duration, the refractive angle of a same point on the fluid object will remain constant, so $\alpha_{t_1} = \alpha_{t_2}$. Subtracting Eq. 23 by Eq. 24, we have:

$$\overrightarrow{x'_{t_1}x'_{t_1+\Delta t}} - \overrightarrow{x'_{t_2}x'_{t_2+\Delta t}} = -\frac{(z''-z')z'\sec^2\beta_{t_1}}{z''}(\alpha_{t_1+\Delta t}-\alpha_{t_2+\Delta t}) \quad (25)$$

Then we can solve $x_{t_2+\Delta t}$ from Eq. 25. If $x'_{t_1+\Delta t}$ and $x'_{t_2+\Delta t}$ are also the same point on the fluid object, then LHS of Eq. 25 is zero as $x'_{t_1+\Delta t} - x'_{t_1} = x'_{t_2+\Delta t} - x'_{t_2}$, and the RHS of Eq. 25 is also zero because the refractive angle of the same point on the fluid object are the same, that is $\alpha_{t_1+\Delta t} = \alpha_{t_2+\Delta t}$. This shows $x'_{t_2+\Delta t} = x'_{t_1+\Delta t}$ is one solution to Eq. 25. Assuming that there is only one light ray that emits from x''_{t_2} and hits at the center of projection, then $x'_{t_2+\Delta t} = x'_{t_1+\Delta t}$ is also the only solution to Eq. 25. This proves $x'_{t_2+\Delta t} = x'_{t_1+\Delta t}$ and consequently $\alpha_{t_1+\Delta t} = \alpha_{t_2+\Delta t}$.

Finally, we prove that wiggles at t_1 and t_2 are equal. Plugging

$$\alpha_{t_1+\Delta t} - \alpha_{t_1} = \alpha_{t_2+\Delta t} - \alpha_{t_2} \tag{26}$$

to Eq. 22, we have:

$$\begin{aligned} \overrightarrow{x_{t_1} x_{t_1+\Delta t}} &= -\frac{(z'' - z')z \sec^2 \beta_{t_1}}{z''}(\alpha_{t_1+\Delta t} - \alpha_{t_1}) \\ &= -\frac{(z'' - z')z \sec^2 \beta_{t_2}}{z''}(\alpha_{t_2+\Delta t} - \alpha_{t_2}) = \overrightarrow{x_{t_2} x_{t_2+\Delta t}} \end{aligned} \tag{27}$$