CONJUGATE TRAJECTORY AND ITS UNCERTAINTY OF FLAT-REFRACTIVE STEREO-PAIR IN MULTIMEDIA ENVIRONMENT

Szu-Han Wang[1] and Jen-Jer Jaw[2]
[1] National Taiwan University, Civil Engineering, No. 1, Sec. 4, Roosevelt Rd., Da'an Dist., Taipei, Taiwan, Email: r10521808@ntu.edu.tw
[2] National Taiwan University, Civil Engineering, No. 1, Sec. 4, Roosevelt Rd., Da'an Dist., Taipei, Taiwan, Email: jejaw@ntu.edu.tw

KEY WORDS: underwater photogrammetry, dense image matching, axial camera, flat-port

ABSTRACT: In a single-media environment, dense image matching can be effectively conducted along one-dimensional epipolar line or by row-to-row corresponding epipolar imagery. In contrast, when the light passes through multi-media in an underwater flat-port camera, incident ray is refracted, and conjugate trajectory becomes a quartic curve, complicating the image matching task. Although the quartic curve can be realized by ray tracing or refractive fundamental matrix, the former is lack of explicit equation of the trajectory, while the latter is mathematically complex. Furthermore, not only the curve itself is crucial, but also the uncertainty, resulting from the errors of relevant parameters, of the trajectory should be taken into consideration during the image matching process, which has not been clearly addressed in the literature yet. In this research, first, the result of conjugate trajectory is provided based on a coplanar condition in axial camera configuration. Second, one of the influential factors, the distance from perspective center it the interface, of the trajectory is analyzed by error propagation with the analytical function. Thus, it is the aim of this study to provide a more efficient way to locate the quartic trajectory and define the error range of the trajectory for the mission of underwater dense image matching as employing flat-refractive imaging system.

1. INTRODUCTION

Photogrammetry has been applied to several 3D reconstruction tasks under the water, such as the reconstruction of corals, sea beds, river habitats and so on. In underwater environment, cameras are usually equipped with water-proof flat-ports. When the light bounces off object and passes through the lens system, the ray travels through three mediums, water, glass and air. According to Snell's law, the light is refracted when the media changes which results in that the object point, perspective center and the image point are not on the same line. When this phenomenon occurs on a stereo-pair, the conjugate trajectory becomes a quartic-curve, in contrast to epipolar line in a single-media situation. In a single-media environment, the conjugate rays and baseline vector are coplanar. Based on this geometry, dense image matching is conducted along one-dimensional epipolar line or row-to-row corresponding epipolar imagery, which are effective and powerful. However, these kinds of operations cannot apply to non-coplanar quartic curve trajectory in multi-media system, complicating the image matching task.

To locate the quartic conjugate trajectories, researchers have proposed two main approaches. First, the ray tracing method, given any image point on the left image and the possible depth of the object point, it can be traced back to the conjugate points on the right image. Gedge et al. (2011) simulated the conjugate trajectories under water-air environment, and they solved a degree of four equations for each point on the trajectories, point-by-point, thus computationally expensive. Lee et al. (2017) further reduced the searching area and computational resources by confining the possible range of depth of the object points. Although ray tracing method can be used as a tool for finding conjugate trajectories, it does not provide any clear function between stereos; besides, the depth of object points in the computational process is needed.

In contrast to ray tracing method, the other approach tried to build the relative geometry relationship between a stereo-pair. Chari and Sturm (2009) proposed "Refractive fundamental matrix". This equation is the first one which describes a clear relationship between the left and right conjugate rays; moreover, there is no requirement of the depth of object points. Based on the refractive fundamental matrix and the concept of "axial camera system (Agrawal et al., 2012)", Elnashef and Filin (2022) developed a new equation that improved the refractive fundamental matrix, they well applied this equation to calibrate the exterior orientation parameters (EOPs), interior orientation parameters (IOPs) and interface parameters.

With a clear formula, the trajectories can be calculated without the information of object point. However, the uncertainty, resulting from the errors of relevant parameters, of the trajectory has not been clearly addressed in the literature yet. Thus, in this research, we develop the equation first, and utilize it to calculate and analyze the uncertainty of trajectory. The method is described in chapter 2, the results and discussion are shown in chapter 3, and

finally the chapter 4 is the conclusion.

2. METHODOLOGY

2.1 Coplanar Geometry Between Stereo-Pair Under Axial Camera System

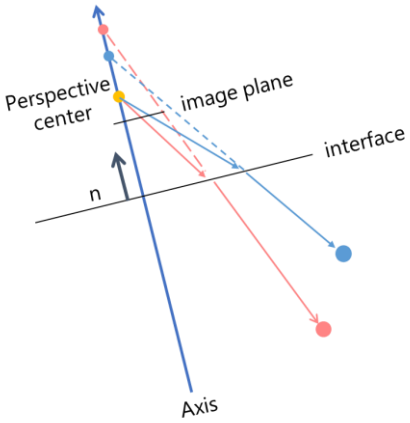


Figure 1. Axial camera system

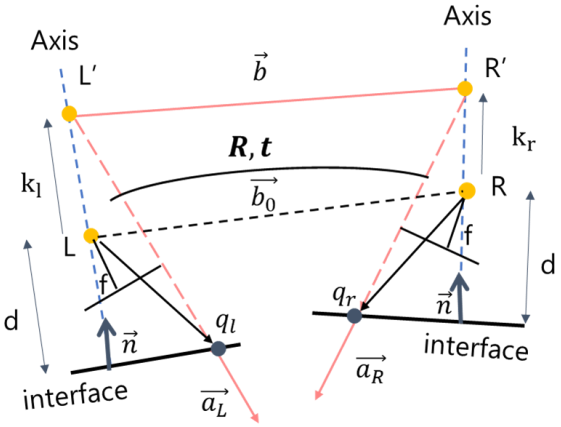


Figure 2. Coplanar geometry between stereo under axial camera system (revised from Elnashef and Filin (2022))

2.1.1 Axial Camera System: In a multi-media environment, the light rays refract when passing through the interface of two media, so the object point, perspective center and the image point are not on the same line which violates the properties of pinhole camera system. Nevertheless, if the lights go straight without bending while passing through the interface, as the dashed line in figure 1, the lights will intersect with the axis, which is parallel to the normal vector of interface and penetrates the perspective center. The intersection point varies with different object points. Therefore, this system is called axial camera system that the lights would intersect at different position along the axis (Agrawal et al., 2012). Gaining more insights in this system, we can regard the intersection points on axis as the new perspective centers, and the interface as a big image plane so that the points on interface is the new image points in this camera system.

2.1.2 Axial Camera System in Stereo Pair: It's shown in figure 2 that the conjugate rays $\overrightarrow{a_L}, \overrightarrow{a_R}$ are not coplanar with original baseline vector $\overrightarrow{b_0}$, but with the new baseline vector $\vec{b}$ formed by new perspective centers L′ and R′ under axial camera system, the coplanar geometry is reconstructed. Equation (1) expresses the coplanar condition within three vectors, and the matrix form is expressed as equation (2).

$$\vec{b} \cdot [\overrightarrow{a_L} \times (\overrightarrow{a_R})] = \overrightarrow{a_L} \cdot [\vec{b} \times (\overrightarrow{a_R})] = \overrightarrow{a_L}^{\mathrm{T}} \mathrm{K_b} (\overrightarrow{a_R}) = 0 \tag{1}$$

$$[\mathrm{x_{a_L}} \quad \mathrm{y_{a_L}} \quad \mathrm{z_{a_L}}] \begin{bmatrix} 0 & -\mathrm{b_z} & \mathrm{b_y} \\ \mathrm{b_z} & 0 & -\mathrm{b_x} \\ -\mathrm{b_y} & \mathrm{b_x} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{x_{a_R}} \\ \mathrm{y_{a_R}} \\ \mathrm{z_{a_R}} \end{bmatrix} = 0 \tag{2}$$

where

$$\overrightarrow{a_L} = [\mathrm{x_{a_L}} \quad \mathrm{y_{a_L}} \quad \mathrm{z_{a_L}}]^{\mathrm{T}}, \overrightarrow{a_R} = [\mathrm{x_{a_R}} \quad \mathrm{y_{a_R}} \quad \mathrm{z_{a_R}}]^{\mathrm{T}}, \vec{b} = [\mathrm{b_x} \quad \mathrm{b_y} \quad \mathrm{b_z}]^{\mathrm{T}}$$

2.2 Coplanar Geometry Between Stereo-Pair in Left Interface Coordinate System

The coplanar condition above is based on the refracted vectors $\overrightarrow{a_L}$ and $\overrightarrow{a_R}$ and the baseline vector $\vec{b}$. However, in order to build the relationship between the stereo-pair, we need to put the exact image points into equation (2), and rewrite $\overrightarrow{a_L}, \overrightarrow{a_R}$ and $\vec{b}$ as the function of left and right image points. Here we introduce two main coordinate system, image coordinate system and interface coordinate system. The former is defined by an image plane and the optical axis, and the origin of the coordinate is at the perspective center; the latter is defined by the normal vector of the interface and the interface plane, and its origin is at the perspective center too, as shown in figure 3. If the optical axis is parallel to the normal vector of the interface, then two coordinate system are the same. Besides, left and right images have their own interface coordinate systems.

2.2.1 Transformation Between Image and Interface Coordinate System: Considering the small angle of installation error between the flat-port and the image plane, image coordinate system and the interface coordinate system are not always identical. This inclination could be understood from the calibrated normal vector of the interface in the image coordinate system $\overrightarrow{n_c} = [n_x \quad n_y \quad n_z]^T$. The normal can further define the rotation matrix R_η, rotating from the interface coordinate system to the image coordinate system, as shown in equation (3).

$$R_\eta = \left[\frac{\overrightarrow{n_c} \times \vec{z}}{\|\overrightarrow{n_c} \times \vec{z}\|} \quad \frac{\overrightarrow{n_c} \times (\overrightarrow{n_c} \times \vec{z})}{\|\overrightarrow{n_c} \times (\overrightarrow{n_c} \times \vec{z})\|} \quad \frac{\overrightarrow{n_c}}{\|\overrightarrow{n_c}\|} \right], \qquad \text{where } \vec{z} = [0 \quad 0 \quad 1]^T \tag{3}$$

Given an image point (x, y), the vector from the perspective center to the point is $\vec{m} = [x \quad y \quad -f]^T$. Same vector in the interface coordinate system is $\overrightarrow{a_{i0}} = R_\eta{}^T \vec{m} = [x_{a_{io}} \quad y_{a_{io}} \quad z_{a_{io}}]^T$, where i = *l* or *r*, which means that the vector is in the left interface coordinate system and the right interface coordinate system, respectively.

2.2.2 Refraction Under Interface Coordinate System: Based on Snell's law (refer to equation (4) and figure 3), the refracted vector $\overrightarrow{a_i}$ can be expressed as the function of the incident ray $\overrightarrow{a_{i0}}$ and the ratio of refractive coefficients $\lambda = n_a/n_w$, where n_a is the refractive coefficient in the air, and n_w is the refractive coefficient in the water, as shown in equation (5) (Chari and Sturm 2009; Elnashef and Filin ,2022). It is noteworthy that $\overrightarrow{a_l}$ and $\overrightarrow{a_r}$ here are not in the same coordinate system, but in the left and right interface coordinate system, respectively.

$$n_a \sin \alpha_i = n_w \sin \beta_i \tag{4}$$

$$\overrightarrow{a_i} = \begin{bmatrix} x_{a_i} \\ y_{a_i} \\ z_{a_i} \end{bmatrix} = \begin{bmatrix} \lambda x_{a_{io}} \\ \lambda y_{a_{io}} \\ -\sqrt{\lambda^2 z_{a_{io}}{}^2 + (1-\lambda^2)\|\overrightarrow{a_{i0}}\|^2} \end{bmatrix} = \lambda \overrightarrow{a_{i0}} + \delta_i \vec{n}$$

$$= \lambda \overrightarrow{a_{i0}} + \left(-\lambda \left(\overrightarrow{a_{i0}}{}^T \vec{n} \right) - \sqrt{\lambda^2 (\overrightarrow{a_{i0}}{}^T \vec{n})^2 + (1-\lambda^2)\|\overrightarrow{a_{i0}}\|^2} \right) \vec{n} \tag{5}$$

$$= \lambda R_\eta{}^T \vec{m} + \left(-\lambda \left(R_\eta^{13} x + R_\eta^{23} y + R_\eta^{33} (-f) \right) - \sqrt{\lambda^2 \left(R_\eta^{13} x + R_\eta^{23} y + R_\eta^{33} (-f) \right)^2 + (1-\lambda^2)(x^2 + y^2 + f^2)} \right) \vec{n}$$

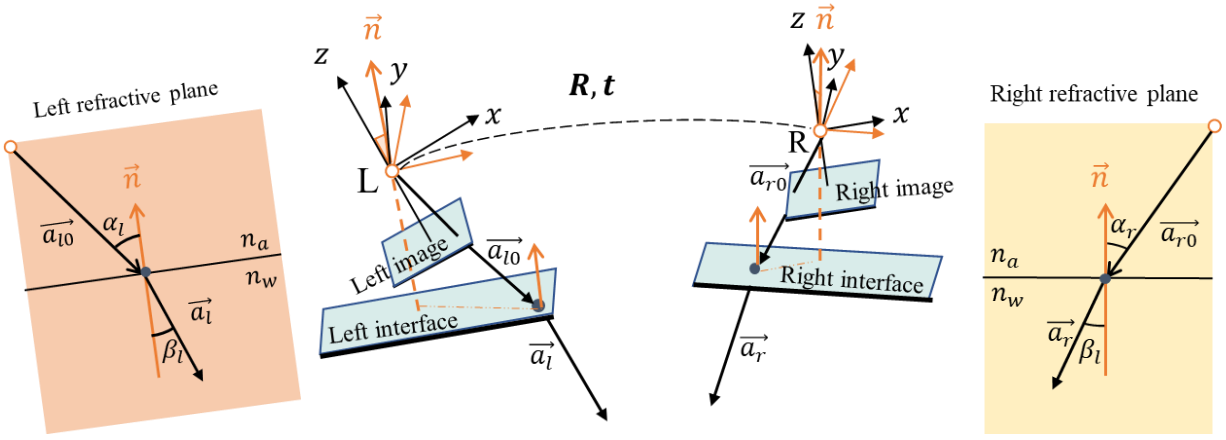


Figure 3. In the middle part of the figure is two coordinate system: Image coordinate system is in black; interface coordinate system is in orange. On the left and right side of the figure is the refractive plane of left and right images.

2.2.3 New Perspective Center in Interface Coordinate System: Elongating the refracted vector until it encounters the axis, the intersection point is the new perspective center. In the interface coordinate system, the points on the axis is always on the z-axis that x and y coordinates are always zero (equation (6)). Thus, the offset of the perspective center k_i, as shown in figure 2, is expressed as equation (7).

$$q_i + \frac{\mathrm{d} + \mathrm{k}_i}{\overrightarrow{a_l}^{\mathrm{T}}\vec{\mathrm{n}}}\overrightarrow{a_l} = \frac{-\mathrm{d}}{\overrightarrow{a_{l0}}^{\mathrm{T}}\vec{\mathrm{n}}}\overrightarrow{a_{l0}} + \frac{\mathrm{d} + \mathrm{k}_i}{\overrightarrow{a_l}^{\mathrm{T}}\vec{\mathrm{n}}}\overrightarrow{a_l} = \mathrm{k}_i\vec{n} \tag{6}$$

$$\begin{aligned}\mathrm{k}_i &= -\mathrm{d} - \frac{\mathrm{d}\sqrt{\lambda^2(\overrightarrow{a_{l0}}^{\mathrm{T}}\vec{\mathrm{n}})^2 + (1-\lambda^2)\|\overrightarrow{a_{l0}}\|^2}}{\lambda\overrightarrow{a_{l0}}^{\mathrm{T}}\vec{\mathrm{n}}} \\ &= -\mathrm{d} - \frac{\mathrm{d}\sqrt{\lambda^2\left(\mathrm{R}_\eta^{13}\mathrm{x} + \mathrm{R}_\eta^{23}y + \mathrm{R}_\eta^{33}(-\mathrm{f})\right)^2 + (1-\lambda^2)(\mathrm{x}^2 + \mathrm{y}^2 + \mathrm{f}^2)}}{\lambda\left(\mathrm{R}_\eta^{13}\mathrm{x} + \mathrm{R}_\eta^{23}\mathrm{y} + \mathrm{R}_\eta^{33}(-\mathrm{f})\right)}\end{aligned} \tag{7}$$

2.2.4 Integrate All Vectors into The Model Coordinate System: Setting the left interface coordinate system as the model coordinate system, the vectors in the right interface coordinate system need to be further transformed to the same space as the model coordinate system. As the relative orientation parameters, ***R*** rotates the vectors in the left image coordinate system to the right image coordinate system, and ***t*** is defined by the coordinate of right perspective center in the left image coordinate system.

The vectors already in left interface coordinate system remains unchanged, including $\overrightarrow{a_l}$ and $\mathrm{k}_l\vec{n}$, such that $\overrightarrow{a_L} = \overrightarrow{a_l}$. In contrast, the vectors in right interface coordinate system should be first transformed to the right image coordinate system, next the left image coordinate system, and end at the left interface coordinate system, which is the model coordinate system, such that $\overrightarrow{a_R} = {R_\eta}^T \boldsymbol{R}^T R_\eta \overrightarrow{a_r}$, and $\mathrm{R}' = {R_\eta}^T\left(\boldsymbol{R}^T R_\eta \mathrm{k}_r\vec{n} + \boldsymbol{t}\right)$. The new baseline vector $\vec{b}$ is shown in equation (8). Finally, the coplanar equation can be represented as equation (9). Through this equation, given any image points on the left image, it could solve the conjugate trajectory on the right image; moreover, an analytical solution of the y coordinate on right image can be provided for further error propagation.

$$\vec{b} = \mathrm{R}' - \mathrm{L}' = {R_\eta}^T\left(\boldsymbol{R}^T R_\eta \mathrm{k}_r\vec{n} + \boldsymbol{t}\right) - \mathrm{k}_l\vec{n} \tag{8}$$

$$\left(\lambda {R_\eta}^{\mathrm{T}}\overrightarrow{m_l} + \delta_l\vec{n}\right)\cdot\left\{\left[{R_\eta}^T\left(\boldsymbol{R}^T R_\eta \mathrm{k}_r\vec{n} + \boldsymbol{t}\right) - \mathrm{k}_l\vec{n}\right] \times \left({R_\eta}^T \boldsymbol{R}^T R_\eta\left(\lambda {R_\eta}^{\mathrm{T}}\overrightarrow{m_r} + \delta_r\vec{n}\right)\right)\right\} = 0 \tag{9}$$

3. EXPERIMENT RESULT AND DISCUSSION

Based on the equation (9), we investigate 2 questions with conjugate trajectories. The first one is the conjugate trajectories on the right image and how the trajectories behave at different positions on the image; the other one is how the elements in the equation influence the trajectories, we specify the error on y-axis for each point. The setting of the experiment is described in chapter 3.1.

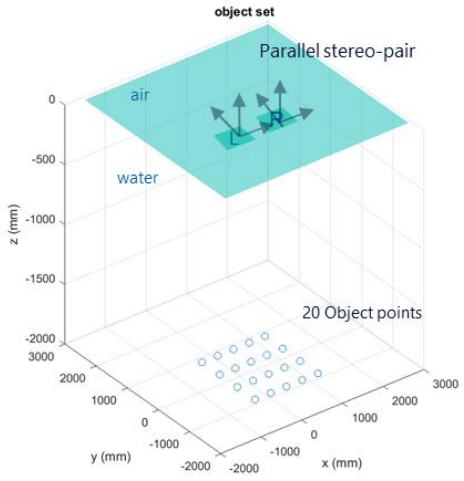


Figure 4. Setting up

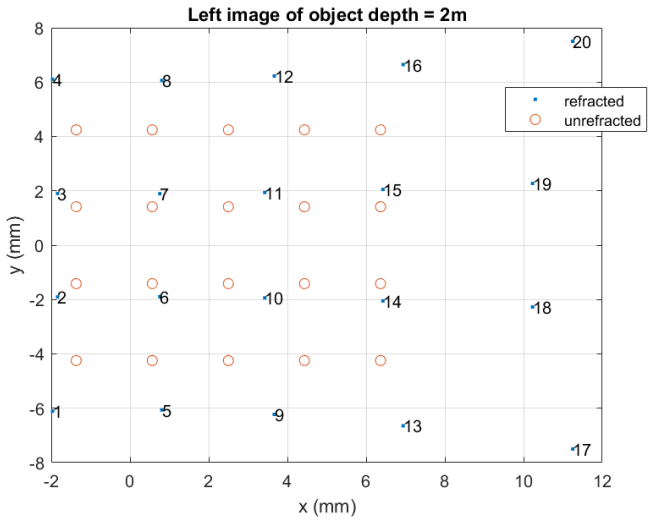


Figure 5. Image points on left image

3.1 Experiment Setting

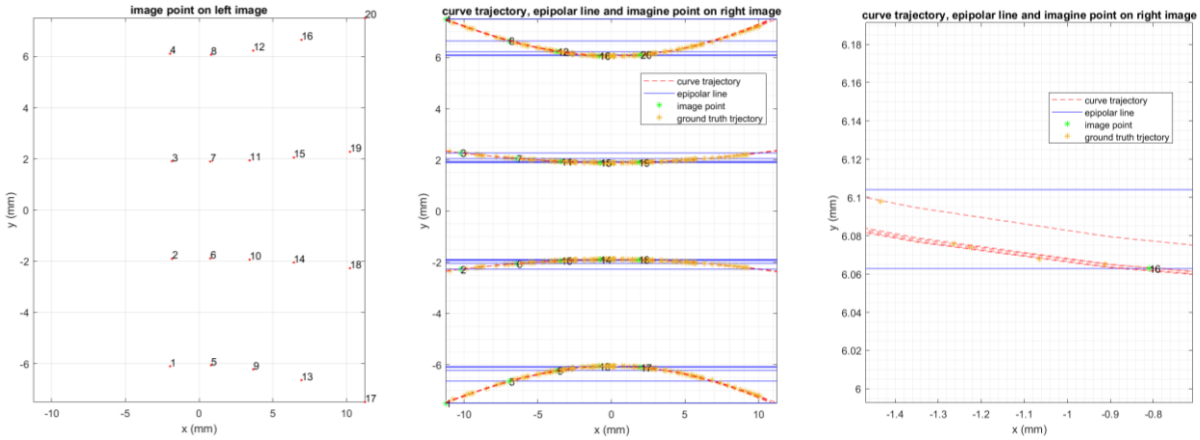
We simulated a stereo-pair that are paralleled to each other. The left perspective center is located at the origin and the right perspective center is located 1m away along the x direction, so $\boldsymbol{R}$ is a 3-by-3 identity matrix and $\boldsymbol{t} = [1000 \quad 0 \quad 0]^T$ (mm). The principal distance is 10 mm, and the perpendicular distance from the interface to the perspective center is 30 mm. Here, we assume the thickness of the glass is negligible, so this configuration is close to a two-media environment. In addition, the normal vector of the interface is parallel to the z-axis, such that $n_c = [0 \quad 0 \quad 1]^T = n$, which means R_η is a 3-by-3 identity matrix. Apart from the setting of the camera and the interface, the object points are 2-meter-deep along the z-axis from the perspective centers, as shown in figure 4.

Following this configuration, the imaging points of those 20 object points on the left image are shown in figure 5. In addition to the points considered the refraction effect, we also calculated the points without the refraction as a reference.

3.2 Computing the Conjugate Trajectory

Based on equation (9), given an image point on the left image $\overrightarrow{m_l}$ (figure 6(a)), relative orientation $(\boldsymbol{R}, \boldsymbol{t})$, and the parameters of interior orientation (f), interface (d), refractive coefficient (λ), we can solve the conjugate trajectory point-by-point, as shown in figure 6(b). Besides, we also plot the epipolar lines for each point, which are the conjugate trajectories in the single-media environment. To take a closer look into the area near the point 16 (figure 6 (c)), we can observe that due to the parallel setting of the stereo-pair and the object points; though, 5 curves are close to each other, they are not identical.

Furthermore, in order to understand how big the error is, if we search for conjugate points on epipolar lines instead of curve trajectories, we calculated the y differences between the two trajectories and the average value is summarized in figure 7. In addition, the incident angle of each point for each curve is shown in figure 8.



(a) Points on left image as the input (b) The conjugate trajectories (curves) and the epipolar line. (c) The zoom in view

Figure 6. Input points on the left side and the conjugate trajectories of right image and epipolar lines on the right side

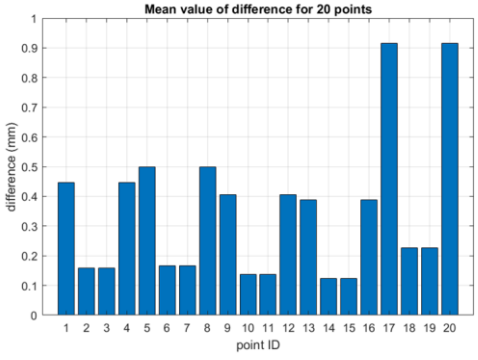


Figure 7. The mean value of the differences

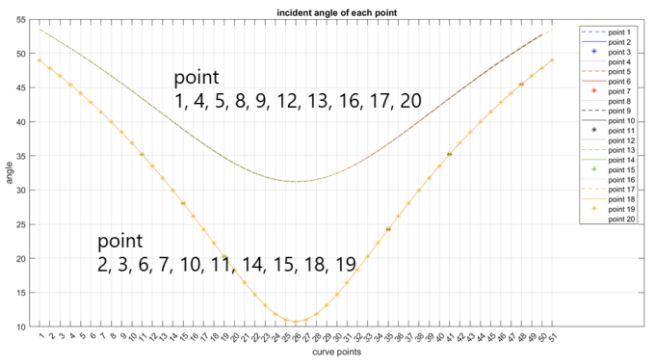


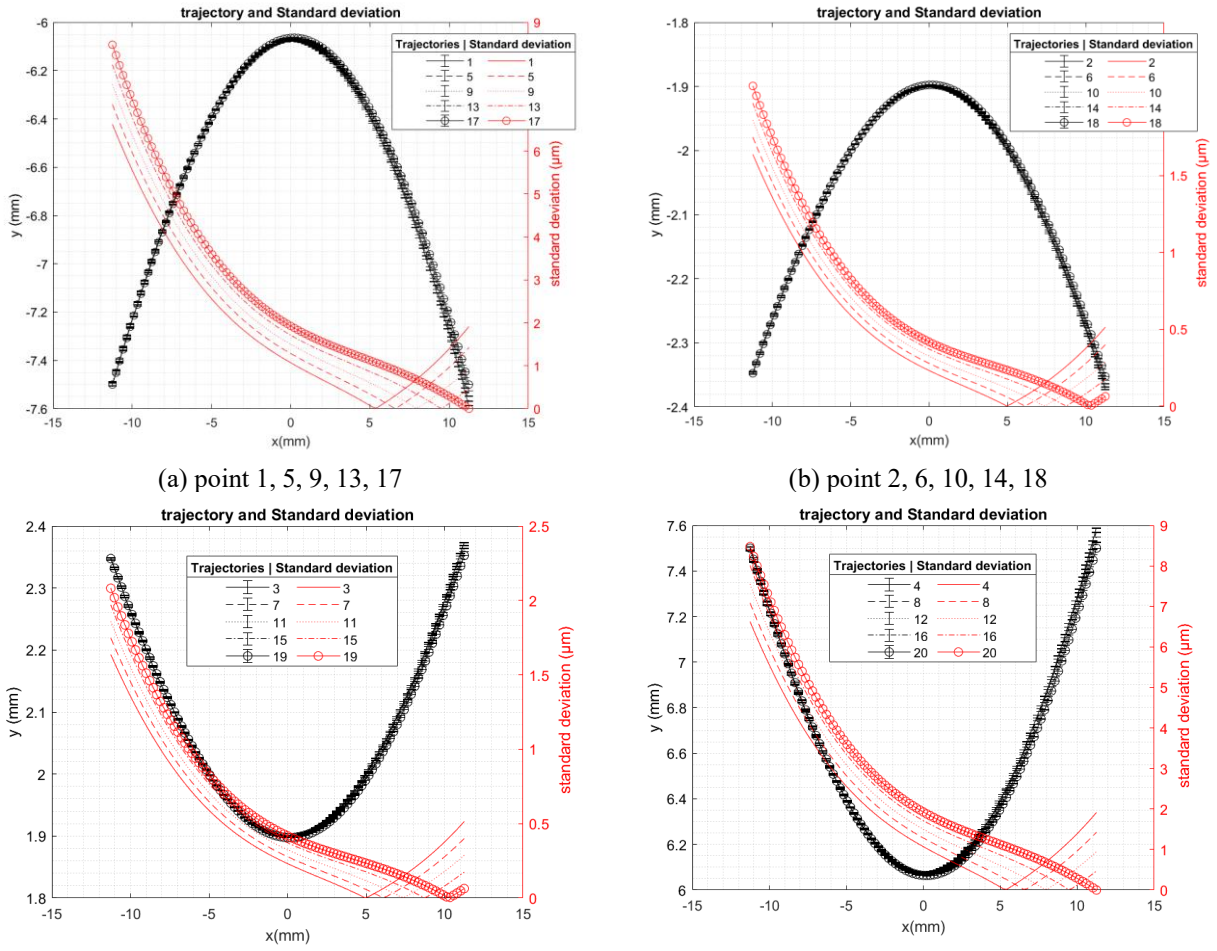
Figure 8. The incident angle (°) of each point for each curve

3.3 Error Propagation: The Distance of the Interface to the Perspective Center (d)

Under the parallel setting, the rotation matrix, such as $\boldsymbol{R}$ and R_η can be simplified to the identity matrix. According to this simplification, the analytical solution of y coordinate can be provided, and this analytical solution is the function of all the elements in the coplanar equation except y. Thus, it could be utilized for further error propagation, and analysis of factors affecting the behavior of y.

Here we take the distance from the perspective center to the interface (d) as the first example to be propagated. The value and the standard deviation of d is 30mm $\pm$ 1mm. The analytical function F= $solve(\overrightarrow{a_L}^{\mathrm{T}} \mathrm{K_b} (\overrightarrow{a_R}) = 0, y_r)$, and the error propagation equation is shown in equation (10). The propagated results of the 20 trajectories are shown in figure 9. To express the details in each curve, we categorized all of them into 4 groups based on their location, and showed the value of standard deviation independently on the second y-axis (in red). Overall, the error decreases from the left side of trajectories to the right side, but there is some "bounce off" at the end of the trajectories.

$$\sigma_y{}^2 = \frac{\partial F}{\partial d} \sigma_d{}^2 \left(\frac{\partial F}{\partial d}\right)^T \tag{10}$$



(a) point 1, 5, 9, 13, 17

(b) point 2, 6, 10, 14, 18

(c) point 3, 7, 11, 15, 19

(d) point 4, 8, 12, 16, 20

Figure 9. Trajectories with the standard deviation on the right image plane

4 DISCUSSION AND CONCLUSION

Observing Figure 7, the average differences are varying among different image points, and the image points near the center of image usually have the smaller differences. Comparing the average differences with the incident angles (figure 8), it is shown that the amount of the differences might be related to the value of incident angles. However, this correlation remains unclear now.

In the second part of the experiment, we propagated the error of the distance from the perspective center to the interface under the parallel stereo-pair setting. In figure 9 (a-d), we could observe that the left side of the images have

larger standard deviation than the right side of the image, but the standard deviation suddenly bounces off at a point, and starts to increase. This trend is same on four of the figures; however, the values of the standard deviation are not the same. In figure 9 (a)(d), the largest value is approximately 8.5μm; in contrast, the largest value in figure 9 (b)(c) is only 2μm, which might be within a sub-pixel level. When we conduct the image dense matching, the sub-pixel value of error is neglectable.

To sum up, based on the coplanar condition under the axial camera system, equation (9) can be well applied to compute the curve trajectory, and also analyze the error of trajectories. However, the results in chapter 3 are not enough to summarize the characteristic of the curve trajectory and the behavior of the influential factors. Therefore, we will conduct more experiment including different camera setting, different influential factors to meet the needs of the real-life underwater task.

REFERENCE

Agrawal, A., Ramalingam, S., Taguchi, Y., & Chari, V., 2012. A theory of multi-layer flat refractive geometry. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3346-3353.

Chari, V., & Sturm, P., 2009. Multi-View geometry of the refractive plane. In BMVC 2009-20th British Machine Vision Conference, pp. 1-11. The British Machine Vision Association (BMVA).

Elnashef, B, & Filin, S., 2022. Target-free Calibration of Flat Refractive Imaging Systems Using Two-view Geometry. Optics and Lasers in Engineering, 150 (2022) 106856.

Gedge, J., Gong, M., & Yang, Y.-H., 2011. Refractive epipolar geometry for underwater stereo matching. In 2011 Canadian Conference on Computer and Robot Vision, pp. 146-152.

Lee, Y.C., Chang, Y.S., Lee, H.B., Chang, W.M., Chia. L., 2017. Underwater Dense Stereo Matching Based on Depth Constraint, ACTA PHOTONICA SINICA, 46(7), pp. 198-207 (Chinese)