

Article

Reconstruction of the Instantaneous Images Distorted by Surface Waves via Helmholtz–Hodge Decomposition

Bijian Jian ^{1,2}, Chunbo Ma ¹, Yixiao Sun ¹, Dejian Zhu ¹, Xu Tian ¹ and Jun Ao ^{1,*}¹ School of Information and Communication, Guilin University of Electronic Technology, Guilin 541000, China² School of Artificial Intelligence, Hezhou University, Hezhou 542800, China

* Correspondence: aojun@guet.edu.cn; Tel.: +86-135-5723-6811

Abstract: Imaging through water waves will cause complex geometric distortions and motion blur, which seriously affect the correct identification of an airborne scene. The current methods mainly rely on high-resolution video streams or a template image, which limits their applicability in real-time observation scenarios. In this paper, a novel recovery method for the instantaneous images distorted by surface waves is proposed. The method first actively projects an adaptive and adjustable structured light pattern onto the water surface for which random fluctuation will cause the image to degrade. Then, the displacement field of the feature points in the structured light image is used to estimate the motion vector field of the corresponding sampling points in the scene image. Finally, from the perspective of fluid mechanics, the distortion-free scene image is reconstructed based on the Helmholtz–Hodge Decomposition (HHD) theory. Experimental results show that our method not only effectively reduces the distortion to the image, but also significantly outperforms state-of-the-art methods in terms of computational efficiency. Moreover, we tested the real-scene sequences of a certain length to verify the stability of the algorithm.

Keywords: water-air imaging; image restoration; structured light; Helmholtz–Hodge Decomposition



Citation: Jian, B.; Ma, C.; Sun, Y.; Zhu, D.; Tian, X.; Ao, J. Reconstruction of the Instantaneous Images Distorted by Surface Waves via Helmholtz–Hodge Decomposition. *J. Mar. Sci. Eng.* **2023**, *11*, 164. <https://doi.org/10.3390/jmse11010164>

Academic Editor: Fausto Pedro García Márquez

Received: 17 November 2022

Revised: 20 December 2022

Accepted: 2 January 2023

Published: 9 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Viewing the airborne scene through a water surface using a submerged camera constitutes a virtual periscope. Unlike a true periscope, the system is entirely submerged and does not draw attention. Images acquired in this way suffer from serious refractive distortions due to the random fluctuations of the water–air interface (WAI). Such effects seriously affect the acquisition of real-scene information and limit the development of underwater vehicles in practical applications, such as path planning, air target monitoring, and marine biological research [1–6]. However, removing such distortions from a single image is an inherently ill-posed problem. It is similar to the process of blind deconvolution; however, the kernel changes randomly in space, which is much more complicated than the process of image deblurring. As such, most previous studies [7–14] have mainly relied on high-frame rate video frames or templates, which limits their applicability in practice.

The studies on this challenging problem have been carried out for decades. One way to approach the problem is to create a spatial distortion model using a long video sequence and then fit it to each frame to recover the undistorted image synchronously. Tian et al. [15] creatively proposed a model-tracking method to recover real underwater scene images. In later work [16], they proposed a data-driven iterative method to reduce the distortions of images, which introduced the notion of a “pull-back” operation. K. Seemakurthy et al. [17] proposed the responding methods to remove distortions caused by unidirectional cyclic waves and circular ripples. However, it is difficult for this simple spatial model, such as that reported in [15–17], to represent various complex fluid motion phenomena in natural scenes. More recently, Li et al. [18] first attempted to remove the dynamic refractive effects using a single image via a deep learning method. Later, James et al. [19] put forward

to a novel method combining compressive sensing (CS) and a local polynomial image representation. Simron Thapa et al. [20] proposed a distortion-guided network (DG-Net) to remove the dynamic refractive effects. They first introduced a physically constrained convolutional network to predict the distortion map from the warped image and then used it to guide the generation of a distortion-free image. For the single image undistortion problem, a deep learning framework can indeed achieve end-to-end processing without a prior image and active excitation. However, a large amount of training data and time is often required in the network-training phase. Obviously, it is very difficult to construct a widely applicable dataset for complex and changeable fluid motion.

Another typical approach is by estimating the instantaneous shape of surface waves, which tries to first estimate the slope distribution of the instantaneous water surface by studying the optical properties of light rays at the interface of the medium and then restores the distortion-free image synchronously. Milder et al. [21,22] first presented a scheme of recovering the real scene by estimating the shape of the water surface. Assuming the sky was uniform in brightness without sunlight and clouds, the total reflection extinction area outside the Snell window was completely dark. On this basis, they first estimated the shape of the wavy water surface based on a map among sky brightness, the angle of incidence, and the slope of the water surface. Then, the distortion-free image was recovered through reverse ray tracing. Levin et al. [23] designed an experimental setup to remove the geometric distortion effect of underwater target images. At first, a color camera was used to simultaneously acquire an image of the object and a glitter pattern on the water surface using multiple illumination sources and then process the glitter pattern to obtain the values of surface slopes at a limited number of points and to use these slopes for the retrieval of image fragments [24]. Finally, a complete distortion-free image was generated through multiple accumulations of short-exposure images [25–28]. Alterman et al. [29] estimated the shape of the wavy water interface by adding an additional refractive imaging sensor. The system uses the sun as the guide star. The wavefront sensor includes a pinhole array, a diffuser plate, and a submerged camera. The position of each spot that goes through a pinhole is used to figure out the sampled normal vector of the water surface. This is then used to estimate the water interface. Y. Schechner et al. [30] utilized multiple viewpoints along a wide baseline to observe the scene under uncorrelated distortion conditions and recover sparse point clouds. R.H. Gardashov et al. [31] proposed a method for recovering the instantaneous images distorted by surface waves. They first used the signature of the sun's flickering to determine the instantaneous shape of the wavy water surface and then corrected the warped image based on back-projection. However, this is only suitable for scenarios, such as monitoring from an airplane, because there is no sun glint features if viewing an airborne scene from a submerged camera [21,29,30]. The above methods either rely on special natural illumination conditions [21,23,29,31] or require a multi-viewpoint device [30], and the imaging accuracy is not satisfactory, so it is difficult to apply it to practical application scenarios.

The three-dimensional measurement technology based on structured light has the advantages of high precision, fast speed, and strong adaptability. In recent years, it has been widely used in various fields [32–34]. This paper proposed a novel method for recovering the instantaneous images distorted by surface waves. In contrast to previous work, our method does not need a long video sequence [7–17], a large amount of training data [18–20], special illumination conditions [21,23,29,31], multiple-viewpoints [30], or a special illumination source [23]. We only require a simple projection setup and single distorted scene image. The main contributions of this paper are as follows: (1) in a virtual periscope, propose a based active excitation water–air imaging model; (2) from the perspective of fluid mechanics, an underwater image-correction algorithm based on Helmholtz–Hodge Decomposition is proposed; (3) demonstrate the approach and analyze its effectiveness.

2. Optical Analysis of Imaging through WAI

2.1. Snell's Window

2.1.1. Snell's Window in Flat WAI

Imaging through a water surface, light ray transmission follows Snell's Law [35,36]. When the water surface is flat, the field of view of a submerged camera only needs 97.2° , which can observe the whole sky scene regardless of the observer's depth. As shown in Figure 1, a submerged camera observes from an underwater location of **O**. OM and ON are the total reflection boundaries, where $\angle MON = 97.2^\circ$. The conical area formed by the rotation of OM and ON around the optical axis is called the Snell cone, and the circular area on the water surface between M and N is the Snell window (SW), the boundary of which is the extinction boundary. SW is surrounded by a dark region that represents light that is entirely internally reflected from the sea and back to the observer from the water below.

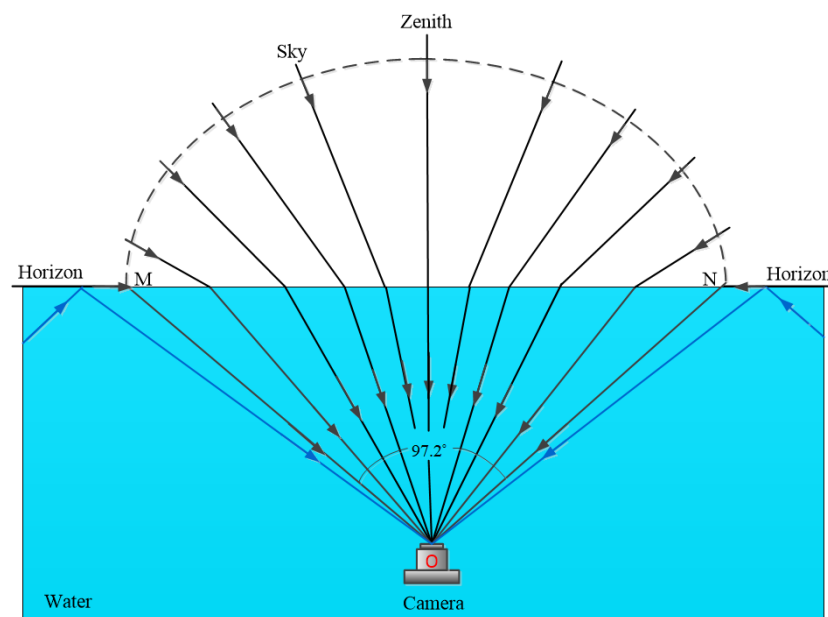


Figure 1. Example of Snell's window as the WAI is calm.

2.1.2. Snell's Window in Wavy WAI

Assuming that the brightness distribution of airborne scenes is uniform, only the effect of the water–air interface for the illumination distribution on the image plane is considered, ignoring the absorption and scattering of the water body and the receiving loss of a camera. Letting the initial value of the sky brightness be 1, according to the Fresnel formula [37], the illuminance on the image plane can be expressed as

$$E(u_x, v_y) = E_{\text{air}}(x', y') * \left(1 - 0.5 * \left[\left(\frac{\sin(\theta_i - \theta_t)^2}{\sin(\theta_i + \theta_t)^2} \right) + \left(\frac{\tan(\theta_i - \theta_t)^2}{\tan(\theta_i + \theta_t)^2} \right) \right] \right), \quad (1)$$

where $E_{\text{air}}(x', y')$ represents the radiant emittance of the objective scene located at (x', y') , $E_{\text{air}}(x', y') = 1$. $E(u_x, v_y)$ is the brightness of the corresponding pixel. θ_i and θ_t are the angle of incidence and refraction, respectively. Figure 2 shows the contour distribution of irradiance on the image plane. Among them, WAI is numerically simulated based on the wave spectrum theory [38,39] by a wind speed of 5.0 m/s, and the water depth of the camera of 8 m.

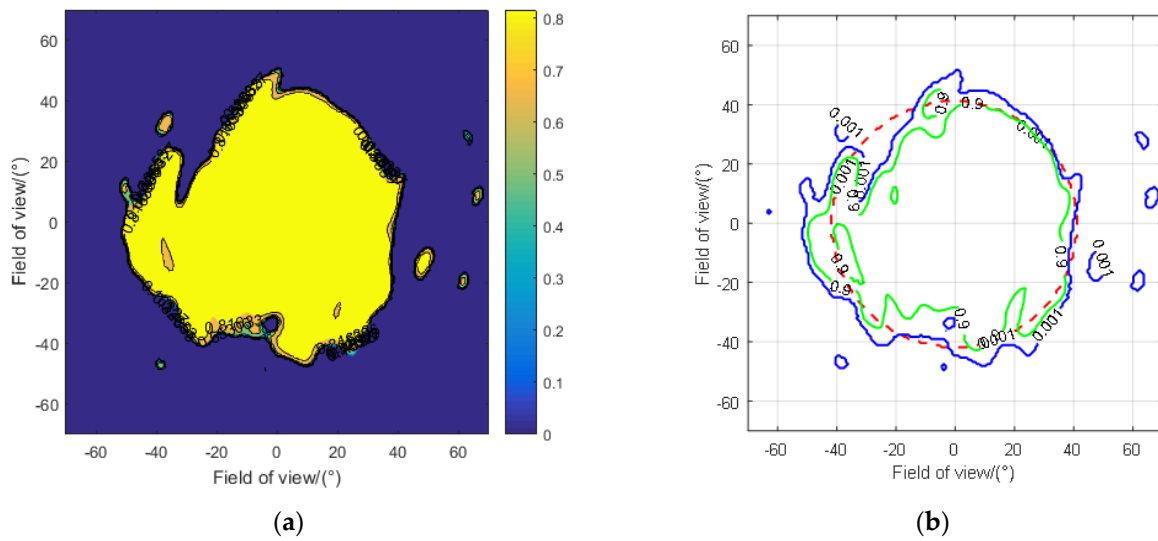


Figure 2. Contour distribution of irradiance on the image plane. **(a)** Contour distribution of normalized illuminance for downwelling radiation from the sky on the image plane. **(b)** Contour distribution of normalized illuminance (the red dashed line is the boundary of Snell's window in flat water; the blue solid line is the boundary of Snell's window in wavy water).

The research results show that there is always a Snell window on the WAI when observing the airborne scenes through water surface. When the water is flat, the Snell window is a regular, smooth circle area. However, in the presence of surface waves, the edge of the Snell window will change irregularly and extend outwards. In addition, there might be a dark extinction region within the window due to the total reflection of light (Figure 2a), which can cause the loss of detail of the edge field of view.

2.2. Imaging through Wavy WAI

Imaging through the wavy water–air interface using a submerged camera, the light rays of the airborne scene will be bent by the dynamic refractive media, resulting in complex geometric distortions. Therefore, the image degradation model of random distortions can be expressed as

$$I(\mathbf{x}) = I_g(\mathbf{x} + \tilde{\mathbf{w}}(\mathbf{x})), \quad (2)$$

where $\mathbf{x} \in \mathbb{R}^2$ represents the pixel coordinates on the image plane, I_g is the distortion-free image, and I is the distorted image. $\tilde{\mathbf{w}}(\mathbf{x})$ is a 2-D warping field that represents the pixel displacement between I and I_g caused by the refraction of the water–air interface.

As Figure 3 shows, the pixel coordinate $\mathbf{x}_i$ will image the scene point **A** when the water surface is calm. However, in the presence of surface waves, the slope of the intersection $\mathbf{q}_i$, which is the crossover point between WAI and the back-projected ray from $\mathbf{x}_i$ through $\mathbf{o}_{lab}$ (optical center), will change across the fluctuation of water. At this moment, the camera will actually observe scene point **B** rather than **A**. Thus, the amount of translation of pixel $\mathbf{x}_i$ on the image plane can be given by

$$\mathbf{w}(\mathbf{x}_i) = -\mathbf{P}\mathbf{d}_s(\mathbf{x}_i), \quad i \in M \times N, \quad (3)$$

where $\mathbf{w}(\mathbf{x}_i)$ and $\mathbf{d}_s(\mathbf{x}_i)$ are both representations of homogeneous coordinates. $\mathbf{d}_s(\mathbf{x}_i) = [d_x, d_y, 0]^T$ represents the amount of displacement that the back-projected ray of the pixel $\mathbf{x}_i$ experiences above water, where d_x , d_y and 0 are the axial components of the vector $\mathbf{d}_s(\mathbf{x}_i)$.

in the camera coordinate frame, respectively. M and N are the pixel size of the image. $\mathbf{P}$ is the projection matrix of the camera [34,35].

$$\mathbf{P} = \frac{1}{z_c} \begin{bmatrix} 1/d_x & 0 & 0 \\ 0 & 1/d_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

where f represents the focus length of the camera, and z_c is the distance from the optical center $\mathbf{o}_{lab}$ to the planar scene above water (ignoring the 3-D properties of the object). d_x and d_y represent the physical sizes of each pixel on the image plane at the x -axis and y -axis, respectively. It can be seen from Equation (3) that since the fluctuations of WAI are spatially variable, different pixel points may have different pixel displacements, resulting in complex geometric distortions of the image.

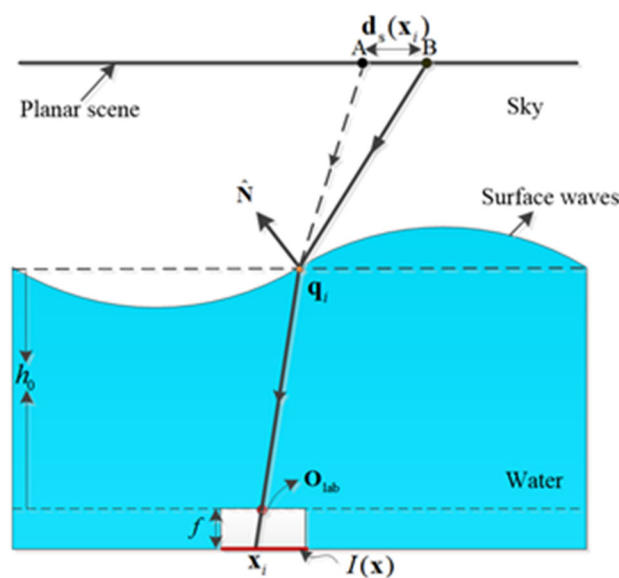


Figure 3. Illustration of refractive distortion.

3. Methods

3.1. Overview

This paper proposed an image-recovery scheme based on the Helmholtz–Hodge Decomposition to recover the instantaneous images distorted by surface waves. The key idea is to leverage the feature information of the preset structural light pattern to realize the water–air interface(WAI) sampling and simultaneously acquire the scene image in the air and the structured light image on the diffuser plate. Then, the deformation vector field of the control points in the scene image were estimated using the displacement relationship of the corresponding points between the structural light image and the scene image. Finally, the distortion-free scene image is recovered based on the Helmholtz–Hodge Decomposition theory.

As Figure 4 shows, similar to our previous work [6], we first introduced a water–air imaging model with structured light projection to simultaneously acquire the distorted structured image and the warped scene image. In which the camera s obtains the warped structured light image from the diffuser plane, while the camera v acquires the scene image through the same sampled region. Next, we will describe in detail how the algorithm works.

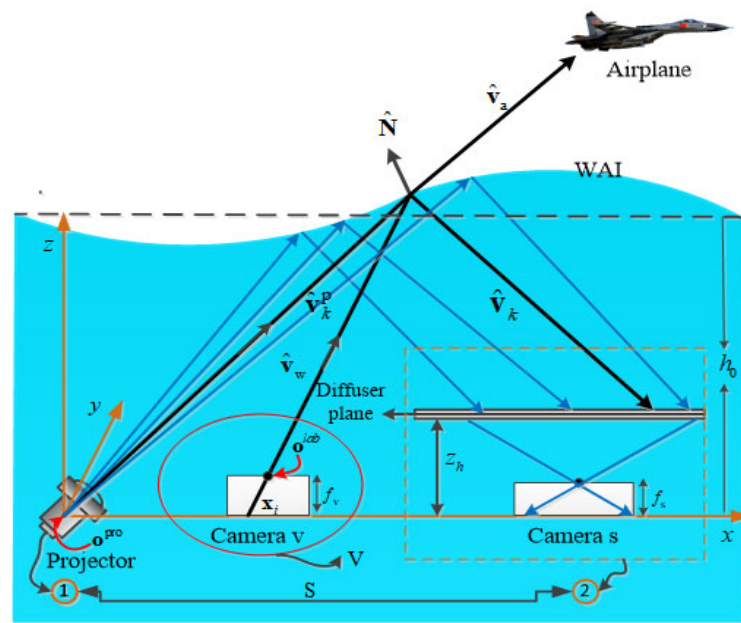


Figure 4. Geometry of the image-restoration model via structured light projection, comprising a structured light projection system S and a viewing camera V. Component S consists of a projector, diffuser plane, and a camera s.

3.2. Sampling of WAI

First, an adaptive and adjustable structured light pattern is projected onto the water surface to accomplish the sampling of the instantaneous wave surface, as shown in Figure 4. Once the system parameters are given, the distribution of WAI sampling points and the corresponding control point distribution on the scene image are easy to acquire through perspective projection transformation [40,41]. A simulation example for the WAI sampling is shown in Figure 5. Figure 5a is a preset structural pattern, and Figure 5b is the virtual image formed on the calm water surface (note: whether it is daytime or a moonlit night, no projected image can be formed on the water. For ease of analysis, we introduce a virtual image).

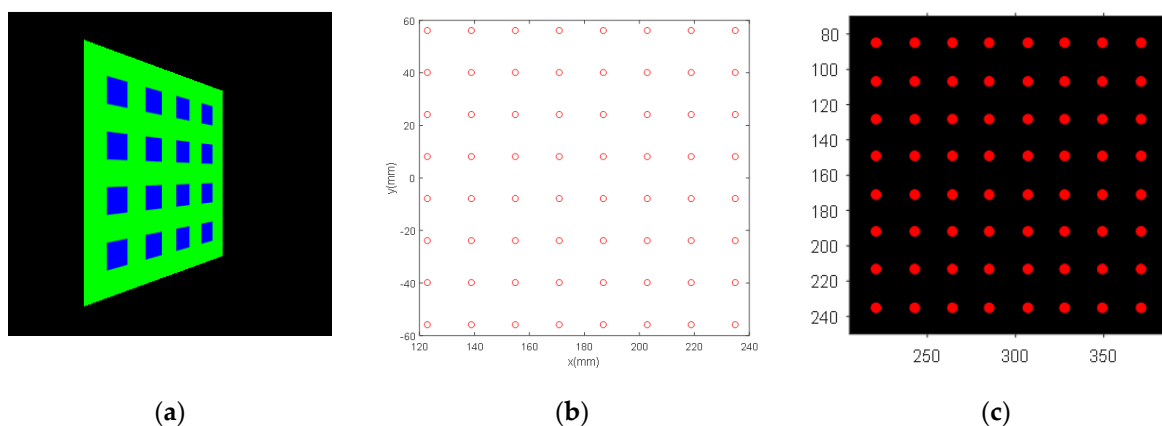


Figure 5. Examples of WAI sampling via structured light projection. (a) A preset structured light pattern in the projector. (b) Distribution of WAI sampling points. (c) Control point distribution on the image plane, where the points correspond to WAI sampled points.

3.3. Feature Extraction and Matching

A distorted structural light image will be formed in the diffusion plate plane after reflection through the water surface, as shown in Figure 6. Then, feature point extraction

and matching were performed on the reference and the twisted structured light images using a new corner detection algorithm (described in Section 3.3.1), respectively. Thereby, the displacement field of the feature points in the structured light image can be expressed as

$$\mathbf{w}_k^{\text{str}} = \mathbf{x}_k - \mathbf{x}'_k = [\omega_{x,k}, \omega_{y,k}, 0]^T, k \in m \times n, \quad (5)$$

where $m \times n$ represents the number of feature points. $\mathbf{x}_k$ and $\mathbf{x}'_k$ are the pixel coordinates of the corresponding feature points on the distorted structure light image $J(\mathbf{x})$ and the reference one $J_g(\mathbf{x})$, respectively.

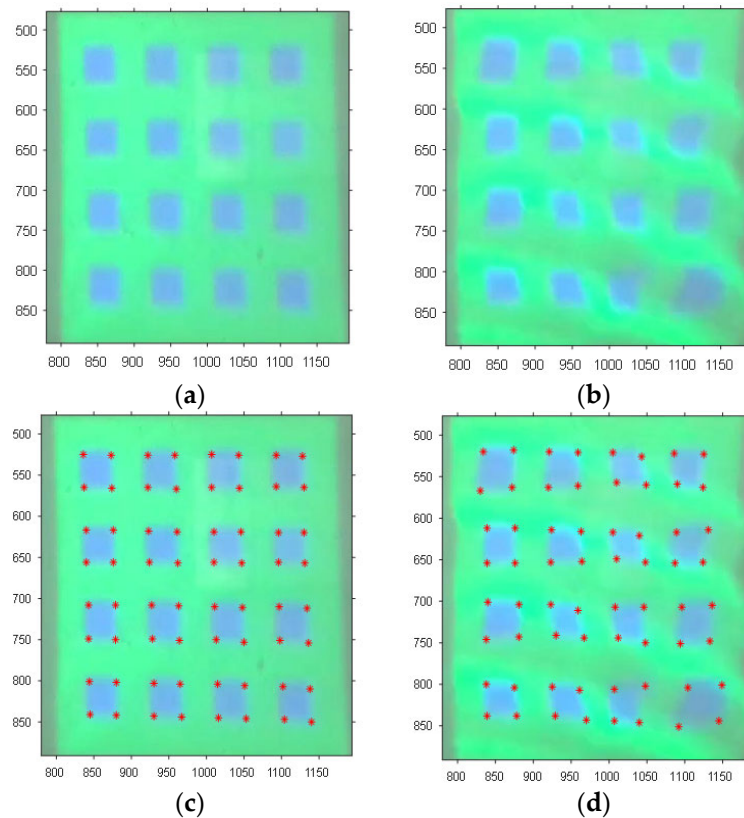


Figure 6. Processing of structured light images in the experiment. (a) Reference structured light image for flat water surface. (b) Distorted structured light image in the presence of surface waves. (c) Feature extraction for the reference frame. (d) Feature extraction for the distorted frame.

3.3.1. A Slope Search-Based Corner Detection Algorithm

Corner detection is a method used to extract features in computer vision systems, which is widely used in motion detection, image registration, 3D reconstruction, and object recognition. The edge profile of the structured light image will no longer be smooth due to the influence of random fluctuations of the water surface, absorption, and scattering based on the water body, which makes it difficult for the current corner detection methods [42–44] to accurately extract feature information. Hence, a slope search-based corner detection method is proposed in this paper.

The framework of the feature-extraction method is shown in Figure 7. Firstly, distorted structured light images are converted to an HSI color model from an RGB color model, where H is the hue, representing the frequency of the color; S represents the depth of the

color, called saturation; and I represents the intensity or lightness. The HSI model can be given as

$$\begin{cases} H = \begin{cases} \theta, & \text{if } b \leq g \\ 2\pi - \theta, & \text{if } b > g \end{cases}, \quad \theta = \cos^{-1} \left\{ \frac{0.5 * [(r-g) + (r-b)]}{[(r-g)^2 + (r-b)(g-b)]^{1/2}} \right\}, \\ S = 1 - \frac{3}{(r+g+b)} [\min(r, g, b)], \\ I = \frac{1}{3}(r + g + b), \end{cases} \quad (6)$$

where H component represents the frequency of the color, which is related to the wavelength. Random fluctuations in the water surface and absorption and scattering by water bodies affect the brightness of the image, but not the wavelengths of the colors. As such, we choose the H channel as processing-frame $J_H(\mathbf{x})$ to extract the feature information for distorted structured light images. Moreover, we further enhanced the edge information of $J_H(\mathbf{x})$ using the guider filtering [45] and transformed it to the binary image $J_H^b(\mathbf{x})$ using Otsu's method [46].

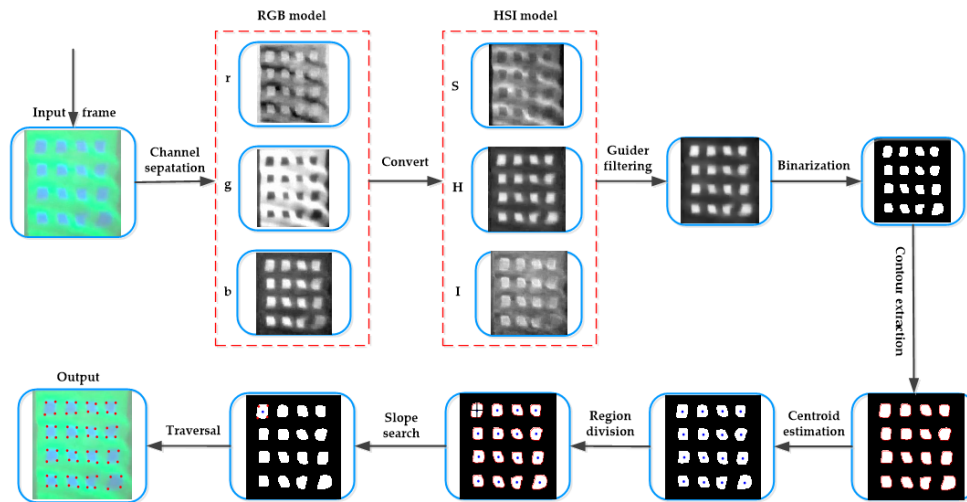


Figure 7. Framework of the feature extraction method in this paper.

Then, the boundary pixel set Θ_l of the arbitrary connected domain Ω_l is obtained based on the connected domain operator, and the geometric center of Ω_l is further calculated based on the spatial moment function, which represents the centroid ε_l . The spatial moment function is given as

$$\psi_{pq} = \sum_{\varphi_x} \sum_{\varphi_y} \varphi_x^p \varphi_y^q \Theta_l(\varphi_x, \varphi_y), \quad (7)$$

where φ_x and φ_y represent the column and row coordinates of an arbitrary pixel in the contour set Θ_l , respectively. From Equation (7), the centroid ε_l of Ω_l can be calculated as

$$\begin{cases} \bar{\varphi}_x = \frac{\psi_{10}}{\psi_{00}}, \\ \bar{\varphi}_y = \frac{\psi_{01}}{\psi_{00}}, \end{cases} \quad (8)$$

where $\bar{\varphi}_x$ and $\bar{\varphi}_y$ are the column and row of the corresponding centroid ε_l .

Subsequently, the contour set Θ_l can be divided into k_s different subsets according to the shape of the structured light and the centroid ε_l . As our structured light pattern is rectangle, set k_s to 4. The specific classification standards are as follows:

$$\begin{cases} K_1, & \text{if } \varphi_x \leq \bar{\varphi}_x, \varphi_y \leq \bar{\varphi}_y \\ K_2, & \text{if } \varphi_x > \bar{\varphi}_x, \varphi_y \leq \bar{\varphi}_y \\ K_3, & \text{if } \varphi_x \leq \bar{\varphi}_x, \varphi_y > \bar{\varphi}_y \\ K_4, & \text{if } \varphi_x > \bar{\varphi}_x, \varphi_y > \bar{\varphi}_y \end{cases}. \quad (9)$$

The corner points of the image are generally the points where the local curvature suddenly changes. As shown in Figure 7, the grid will exhibit varying degrees of compression or expansion as a result of water surface changes. Hence, we can define a corner point of the structured light image as the intersection of the sub-contour K_τ , $\tau = 1, 2, 3, 4$ and the diagonal of the smallest external rectangle, where the external rectangle is defined by the sub-contour K_τ and the centroid ε_l and the diagonal passes through the centroid ε_l . The diagonal slope λ_τ of each subset K_τ of connected domain Θ_l can be calculated using Equation (10), specifically

$$R_k^p \equiv \left\{ \mathbf{o}^{\text{pro}} + \hat{\mathbf{v}}_k^p l_p \right\}_{\forall l_p > 0} \begin{cases} \lambda_1 = \frac{\min(\varphi_y | \varphi_y \in K_{1,y}) - \bar{\varphi}_y}{\min(\varphi_x | \varphi_x \in K_{1,x}) - \bar{\varphi}_x}, \text{ if } (\varphi_x, \varphi_y) \in K_1 \\ \lambda_2 = \frac{\min(\varphi_y | \varphi_y \in K_{2,y}) - \bar{\varphi}_y}{\max(\varphi_x | \varphi_x \in K_{2,x}) - \bar{\varphi}_x}, \text{ if } (\varphi_x, \varphi_y) \in K_2 \\ \lambda_3 = \frac{\max(\varphi_y | \varphi_y \in K_{3,y}) - \bar{\varphi}_y}{\min(\varphi_x | \varphi_x \in K_{3,x}) - \bar{\varphi}_x}, \text{ if } (\varphi_x, \varphi_y) \in K_3 \\ \lambda_4 = \frac{\max(\varphi_y | \varphi_y \in K_{4,y}) - \bar{\varphi}_y}{\max(\varphi_x | \varphi_x \in K_{4,x}) - \bar{\varphi}_x}, \text{ if } (\varphi_x, \varphi_y) \in K_4 \end{cases} \quad (10)$$

where $K_{\tau,x}, K_{\tau,y}$, $\tau = 1, 2, 3, 4$ represent all column and row coordinates of K_τ , respectively. Then, the corner coordinates of Θ_l can be given as

$$[\varphi_{x,\tau}, \varphi_{y,\tau}] = \left\{ (\varphi_x, \varphi_y) \left| \frac{\varphi_y}{\varphi_x} = \lambda_\tau, (\varphi_x, \varphi_y) \in K_\tau \right. \right\}, \tau = 1, \dots, 4. \quad (11)$$

Finally, all connected domains are traversed to obtain all corner points of the input-structured light image.

After generating all corner points, the corner points in the reference image and the twisted image are respectively sorted using the bubble sorting method. The key idea is to update the corner matrix by the column. The specific process is as follows: assuming that the corner point matrix is C and the size is $m \times n$, firstly, the candidate set Cr is constructed from the corner points, which are located at the current column and the next column. Then, sort them in ascending order according to the row values to get a vector Cm with a size of $2m \times 1$. Next, divide the vector Cm into m subsets Cs_i , $i = 1, \dots, m$ from smallest to largest, $i = 1, \dots, m$. Finally, select the corner point with the smallest row value from each subset as the sorting result of the current column and update the matrix C . The entire process for corner point detection is described in Algorithm 1.

Algorithm 1: A Slope Search-based Corner Detection Algorithm

Input: Distorted structured light image $J \in \mathbb{R}^{M \times N}$,

Output: Corner points matrix $C \in \mathbb{R}^{m \times n}$

$[H, S, I] \leftarrow \text{HSI}(\text{rgb}(J));$

$J_H \leftarrow \text{GuiderFilter}(H);$

$J_H^b(\mathbf{x}) \leftarrow \text{Binarization}(J_H);$

$\Theta \leftarrow \text{Boundaries}(J_H^b);$

While $\text{ite} \leq \text{lenth}(\Theta)$ do

$\text{Temp} = \text{zeros}(M, N);$

 for each boundary pixel set $\Theta_l, l = 1, 2, \dots, \text{length}(\Theta)$

$\varepsilon_l \leftarrow \text{CentroidsCalculate}(\Theta_l);$

$K_\tau \leftarrow \text{DivideSubsets}(\varepsilon_l), \tau = 1, 2, 3, 4;$

 for $\tau = 1 : 4$

$[\varphi_{x,\tau}, \varphi_{y,\tau}] = \text{SlopeSearch}(K_\tau);$

$\text{Temp}(\varphi_{y,\tau}, \varphi_{x,\tau}) = 1;$

 end

 end

End

$C \leftarrow \text{Traverse}(\text{Temp} \equiv 1), C \in \mathbb{R}^{m \times n};$

Feature matching

$C = \text{BubbleSort}(C), C \in \mathbb{R}^{m \times n}$

3.4. Distortion Estimation of the Control Points

Water fluctuations will cause the radiated rays passing through the wave surface to be angularly offset, which means that the offsets of any two light rays passing through the same point on the water surface are correlated. Inspired by this, we attempted to establish the displacement relationship of the corresponding points between the structural light image and the scene image from the perspective of geometric optics.

As Figure 8 shows, the global coordinate system is established with the projection center $\mathbf{o}^{\text{pro}}$ as the origin, and the z-axis is vertically upward. Let $\mathbf{p}_k$ represent the 3-D position of the arbitrary feature point of the structured light pattern to be projected, and the corresponding projected ray is

$$R_k^{\text{p}} \equiv \left\{ \mathbf{o}^{\text{pro}} + \hat{\mathbf{v}}_k^{\text{p}} l_p \right\}_{\forall l_p > 0}, \quad (12)$$

where $\hat{\mathbf{v}}_k^{\text{p}}$ is the projected ray direction vector and l_p represents the propagation length along the ray. The projected ray intersects the WAI at $\mathbf{q}_k$, namely the WAI sampling point,

$$\mathbf{q}_k = \text{WAI} \cap R_k^{\text{p}}, \quad (13)$$

and the corresponding reflected ray is

$$R_k \equiv \left\{ \mathbf{q}_k + \hat{\mathbf{v}}_k l_k \right\}_{\forall l_k > 0}, \quad (14)$$

where $\hat{\mathbf{v}}_k$ denotes the reflected ray direction vector and l_k is the propagation length along the ray $\hat{\mathbf{v}}_k$. The reflected ray irradiates a spot on the diffuser plane Π^{diffuser} , at 3D location $\mathbf{s}_k$. The WAI normal is $\hat{\mathbf{N}}_k$, and $\mathbf{s}'_k$ is the corresponding spot location as the water is flat. Thus, the displacement vector $\delta(\mathbf{p}_k)$ of the feature points in the diffusion plane Π^{diffuser} is equal to

$$\delta(\mathbf{p}_k) = [\delta_x, \delta_y, 0]^T = \mathbf{s}_k - \mathbf{s}'_k = \mathbf{P}_s' \mathbf{w}_k^{\text{str}}, \quad (15)$$

where $\mathbf{w}_k^{\text{str}}$ denotes the displacement amount of corresponding feature points in the structured light image. $\mathbf{P}_s'$ is the inverse projection matrix of the camera s

$$\mathbf{P}_s' = z_s \begin{bmatrix} 1/f_s & 0 & 0 \\ 0 & 1/f_s & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dx' & 0 & 0 \\ 0 & dy' & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (16)$$

where f_s is the focal length of camera s. z_s is the distance from the diffuser plane Π^{diffuser} to camera s. dx' and dy' represent the physical size of a single pixel. Relative to the global coordinate system, the pose of camera s can be defined by a rotation matrix and a translation vector, and the 3D position of the origin of camera s is $\mathbf{o}_s^{\text{lab}} = -\mathbf{R}_s^T \mathbf{t}_s$. Similarly, the origin of camera v is $\mathbf{o}_v^{\text{lab}} = -\mathbf{R}_v^T \mathbf{t}_v$.

Camera v observes an airborne scene through wavy WAI, and the back-projected ray forms $\mathbf{x}_i^k$ through $\mathbf{o}_v^{\text{lab}}$ in water as

$$R_w^k(\mathbf{x}_i^k) \equiv \left\{ \mathbf{o}_v^{\text{lab}} + \hat{\mathbf{v}}_w^k l_w \right\}_{\forall l_w > 0}, \quad (17)$$

where $\hat{\mathbf{v}}_w^k$ is the back-projected ray direction vector of the pixel $\mathbf{x}_i^k$ and l_w represents the propagation length along the ray $\hat{\mathbf{v}}_w^k$. The back-projected ray intersects the WAI at

$$\mathbf{q}_k = \text{WAI} \cap R_w^k(\mathbf{x}_i^k). \quad (18)$$

where $h(\mathbf{q}_k)$ represents the altitude of intersection point $\mathbf{q}_k$ and z_h indicates the height of the water depth of the diffuser plate Π^{diffuser} .

The projected ray R_k^p and the reverse projected ray $R_w^k(\mathbf{x}_i^k)$ of the pixel point $\mathbf{x}_i^k$ intersect the same point $\mathbf{q}_k$ on the WAI; hence, in the presence of surface waves, the displacement amounts $\delta(\mathbf{p}_k)$ and $\mathbf{d}_s(\mathbf{x}_i^k)$ are only related to the slope and height of the intersection point $\mathbf{q}_k$ and the height of the diffuser plane z_h .

As Figure 5c shows, the set $\{\mathbf{q}_k\}$ (denoted as $\{\mathbf{q}'_k\}$) is periodic when the water is flat. However, in the presence of surface waves, $\{\mathbf{q}_k\}$ is quasi-periodic, which has a perturbation of periodicity. Considering that the variations in the height of the water surface are smaller than the work depth of the system ($\Delta h \ll h_0$), then

$$\mathbf{q}_k \approx (q_x, q_y, h_0) = \mathbf{p}_k + \hat{\mathbf{v}}_k^p (h_0/c_k), \quad (24)$$

where h_0 is the system height, which is the time-averaged underwater depth of the projector and can be determined in the field using a pressure-based depth gauge. c_k is the z-axis component of the vector $\hat{\mathbf{v}}_k$. Hence, the displacement vectors $\delta(\mathbf{p}_k)$ and $\mathbf{d}_s(\mathbf{x}_i^k)$ can be expressed as

$$\begin{cases} \delta(\mathbf{p}_k) = \mathbf{s}_k - \mathbf{s}'_k = \hat{\mathbf{v}}_k(h_0 - z_h)/c_k - (\hat{\mathbf{v}}'_k(h_0 - z_h)/c'_k), \\ \mathbf{d}(\mathbf{x}_i^k) = \mathbf{s}_b^k - \mathbf{s}_a^k = \hat{\mathbf{v}}_a^k(h_0 - z_h)/c_a^k - (\hat{\mathbf{v}}'_a(h_0 - z_h)/c'^k_a) \end{cases}, \quad (25)$$

where c_k , c'_k , c_a^k , and c'^k_a are the z-axis components of the vectors $\hat{\mathbf{v}}_k$, $\hat{\mathbf{v}}'_k$, $\hat{\mathbf{v}}_a^k$, and $\hat{\mathbf{v}}'^k_a$, respectively; $\hat{\mathbf{v}}'_k$ and $\hat{\mathbf{v}}'^k_a$ represent the reflected ray direction vector of feature point $\mathbf{q}_k$ and the airborne viewing ray direction vector of the corresponding pixel $\mathbf{x}_i^k$ as the water is flat, respectively. According to the Snell' law, the vectors $\hat{\mathbf{v}}_k$, $\hat{\mathbf{v}}'_k$, $\hat{\mathbf{v}}_a^k$, and $\hat{\mathbf{v}}'^k_a$ can be further expressed as

$$\begin{cases} \hat{\mathbf{v}}_k = \hat{\mathbf{v}}_k^p + 2\hat{\mathbf{N}}_k(-\hat{\mathbf{N}}_k \cdot \hat{\mathbf{v}}_k^p), \\ \hat{\mathbf{v}}'_k = \hat{\mathbf{v}}_k^p + 2\hat{\mathbf{N}}_0(-\hat{\mathbf{N}}_0 \cdot \hat{\mathbf{v}}_k^p), \\ \hat{\mathbf{v}}_a^k = \hat{\mathbf{v}}_w^k + \hat{\mathbf{N}}_k \left[\sqrt{1 - n_w^2 + n_w^2(\hat{\mathbf{v}}_w^k \cdot \hat{\mathbf{N}}_k)^2} - n_w \hat{\mathbf{v}}_w^k \cdot \hat{\mathbf{N}}_k \right], \\ \hat{\mathbf{v}}'^k_a = \hat{\mathbf{v}}_w^k + \hat{\mathbf{N}}_0 \left[\sqrt{1 - n_w^2 + n_w^2(\hat{\mathbf{v}}_w^k \cdot \hat{\mathbf{N}}_0)^2} - n_w \hat{\mathbf{v}}_w^k \cdot \hat{\mathbf{N}}_0 \right], \end{cases} \quad (26)$$

where $\hat{\mathbf{N}}_0$ is the WAI normal as the water surface is flat, $\hat{\mathbf{N}}_0 = [0, 0, 1]^T$; $\hat{\mathbf{N}}_k$ is the normal of the WAI sampling point $\mathbf{q}_k$ when the water fluctuates; n_w is the refractive index of water, $n_w = 4/3$.

From Equations (25) and (26), it can be seen that the displacement vectors $\delta(\mathbf{p}_k)$ and $\mathbf{d}_s(\mathbf{x}_i^k)$ are only related to the WAI normal $\hat{\mathbf{N}}_k$, specifically only to the instantaneous slope of $\mathbf{q}_k$, once the system parameters are determined.

Give the normal of the WAI sampling point $\hat{\mathbf{N}}_k$, the mapping relationship between $\delta(\mathbf{p}_k)$ and $\mathbf{d}_s(\mathbf{x}_i^k)$ is easily obtained by ray tracing,

$$\mathbf{d}_s(\mathbf{x}_i^k) = F[\delta(\mathbf{p}_k)], \quad (27)$$

where $F(\cdot)$ is the transformation function between $\delta(\mathbf{p}_k)$ and $\mathbf{d}_s(\mathbf{x}_i^k)$. According to the displacement vector model derived from Equations (25) and (26), given the normal vector

$\hat{\mathbf{N}}_k$ of WAI, this paper estimates the mapping relationship $F(\cdot)$ between the two based on the polynomial representation method.

Moreover, the WAI normal vectors fluctuate randomly around the z -axis; hence, this perturbation can be divided into two components. One is the xoz plane component, at which the normal vector is

$$\hat{\mathbf{N}}_k^{xoz} = (\sin \theta_k, 0, -\cos \theta_k)^T, \quad (28)$$

where θ_k is the inclination angle of the WAI. The other one is the yoz plane component, and the corresponding normal is

$$\hat{\mathbf{N}}_k^{yoz} = (0, \sin \theta_k, -\cos \theta_k)^T, \quad (29)$$

According to Equations (15), (25), (26), (28), and (29), using the least squares polynomial representation, Equation (27) can be further expressed as

$$\mathbf{d}_s(\mathbf{x}_i^k) = \left[\sum_{i=0}^{\eta_1} a_i \delta_x^{\eta_1}, \sum_{j=0}^{\eta_2} b_j \delta_y^{\eta_2}, 0 \right]^T, \quad (30)$$

where η_1 and η_2 denote the orders of the fitted polynomial, respectively; a_i and b_j represent the corresponding coefficients, respectively.

In summary, based on Equations (5), (15), (21), and (30), this paper realizes the estimation of the distortion vector field $\mathbf{w}(\mathbf{x}_i^k)$ of the pixel points in the scene image using the deformation vector field $\mathbf{w}_k^{\text{str}}$ of the corresponding feature points in the structured light image.

3.5. Image Restoration Algorithm Based on Helmholtz–Hodge Decomposition

The Helmholtz–Hodge Decomposition of vector fields is one of the fundamental theorems in fluid dynamics, which has been widely used in various fields [47–50]. Viewing the scene above water with a submerged camera, ignoring the absorption and scattering of water, the essential reasons for image distortion can be attributed to the following three aspects: (1) the height variation of the water surface; (2) the whirlpool of water; (3) the relative position of the camera and the WAI. In 1858, Hermann von Helmholtz et al. [51] first explained the connection between fluid motion and vector field decomposition. From the perspective of fluid mechanics, this paper proposes an image restoration algorithm based on the Helmholtz–Hodge Decomposition. The algorithm flow is shown in Figure 9.

3.5.1. Image Recovery Algorithm Based on HHD

In a bounded space domain, given the divergence ($\nabla \cdot \xi$), curl ($\nabla \times \xi$), and boundary conditions of a smooth vector field ξ , the vector field ξ has the following unique Helmholtz–Hodge decomposition:

$$\begin{aligned} \xi &= \mathbf{d} + \mathbf{r} + \mathbf{h} \\ &= \nabla D + J(\nabla R) + \mathbf{h} \end{aligned} \quad (31)$$

where $\mathbf{d}$ is the divergence component, which is normal to the boundary; $\mathbf{r}$ represents the curl component, which is parallel to the boundary; $\mathbf{h}$ denotes the harmonic component. In $\mathbb{R}^2$, the components $\mathbf{d}$ and $\mathbf{r}$ can be calculated as the gradient of the scalar potential D and the curl of the scalar potential R , respectively. $J(\cdot)$ is an operator [47] that rotates a 2D vector counterclockwise by $\pi/2$; specifically, if $\vec{v} = (v_1, v_2)$, then $J\vec{v} = (-v_2, v_1)$.

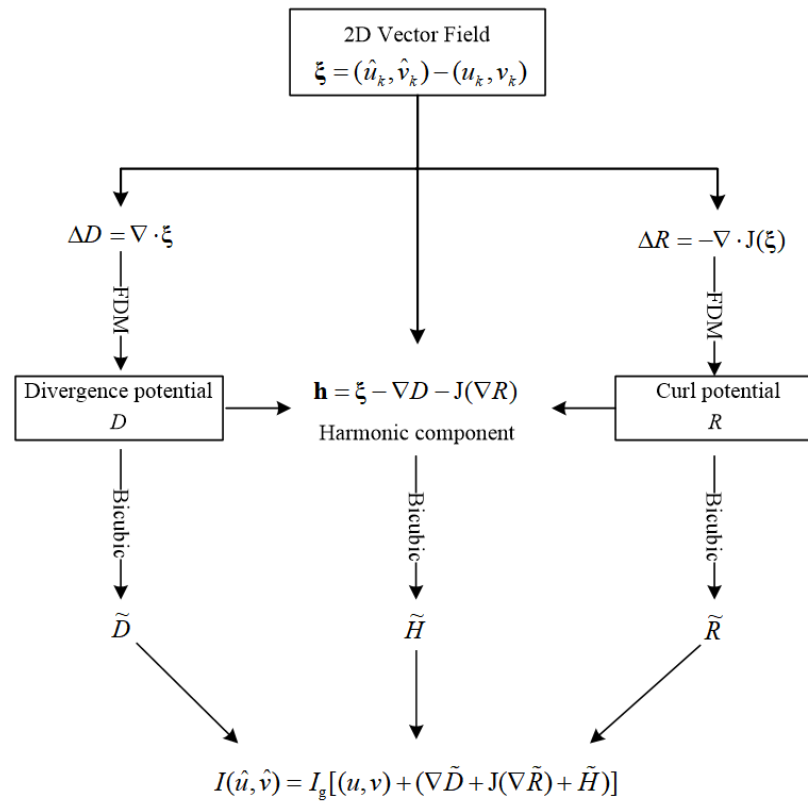


Figure 9. Flow chart of image recovery algorithm based on HHD.

Let (u_k, v_k) and $(\hat{u}_k, \hat{v}_k)$ be the corresponding control point coordinates in the ground-truth scene image $I_g(\mathbf{x})$ and the distorted scene image $I(\mathbf{x})$, respectively, then the 2D distortion vector field ξ is defined as

$$\xi = (\hat{u}_k, \hat{v}_k) - (u_k, v_k), \quad (32)$$

where the homogeneous coordinate form of ξ is $\mathbf{w}(\mathbf{x}_i^k) = [(\hat{u}_k - u_k), (\hat{v}_k - v_k), 0]^T$, which is obtained based on Section 3.4. Moreover, the divergence $(\nabla \cdot \xi)$ and the curl $(\nabla \times \xi)$ of ξ satisfy the following Poisson equations, respectively

$$\begin{aligned} \nabla \cdot \xi &= \nabla \cdot (\nabla D) = \Delta D, \\ \nabla \times \xi &= \nabla \times (J(\nabla R)) = \Delta R, \end{aligned} \quad (33)$$

where $\nabla \cdot$ and $\nabla \times$ are the divergence operator and the curl operator of a vector field, respectively; $\Delta(\cdot)$ represents the Laplace operator. The divergence potential D and curl potential R can be calculated by solving the Poisson Equation (33). Then, the harmonic components can be acquired based on $\mathbf{h} = \xi - \nabla D - J(\nabla R)$, where the harmonic components are regarded as residuals.

Due to the WAI sampling $\{\mathbf{q}_k\}$ is sparse, and the corresponding 2D vector field ξ is also discrete and sparse. Therefore, the divergence component $\tilde{D}$, the curl component $\tilde{R}$, and the harmonic components $\tilde{H}$ of the 2-D distortion vector field $\tilde{\mathbf{w}}$ of the scene image can be further calculated using the bicubic interpolation algorithm [52]. Then, the 2D warped vector field $\tilde{\mathbf{w}}$ of the scene image can be obtained by the inverse HHD, specifically

$$\tilde{\mathbf{w}} = (\nabla \tilde{D} + J(\nabla \tilde{R}) + \tilde{H}). \quad (34)$$

Finally, from Equations (2) and (34), the distortion-free scene image can be acquired via bi-linear interpolation.

3.5.2. Finite Difference Method for HHD

According to the analysis in Section 3.5.1, it can be seen that the key to the decomposition of a two-dimensional vector field lies in the solution of the Poisson equation:

$$\begin{cases} \Delta D = \nabla \cdot \xi \\ \Delta R = \nabla \times \xi \end{cases}, \quad (35)$$

where the divergence ($\nabla \cdot \xi$) and the curl ($\nabla \times \xi$) are known, and $\Delta(\cdot)$ represents the Laplace operator. Considering that the wave surface is sampled periodically when the water is calm, we solve the Poisson equation using the finite difference method [53].

Firstly, the problem of solving Poisson's equations is transformed into a system of linear equations for solving grid nodes. In a two-dimensional regular grid, the finite difference method uses the difference quotient to approximate the derivative; that is, the second derivative of the Laplace operator is approximated using the central difference formula, so the Poisson equation can be expressed as

$$\begin{aligned} \Delta D &= \frac{\partial^2 D}{\partial x^2} + \frac{\partial^2 D}{\partial y^2}, \\ \Delta R &= \frac{\partial^2 R}{\partial x^2} + \frac{\partial^2 R}{\partial y^2}, \end{aligned} \quad (36)$$

Specifically, for each sampling point (x, y) , Equation (34) can be further expressed as

$$\begin{aligned} \frac{D(x+1,y)+D(x-1,y)-2D(x,y)}{(\Delta x)^2} + \frac{D(x,y+1)+D(x,y-1)-2D(x,y)}{(\Delta y)^2} &= F(x,y), \\ \frac{R(x+1,y)+R(x-1,y)-2R(x,y)}{(\Delta x)^2} + \frac{R(x,y+1)+R(x,y-1)-2R(x,y)}{(\Delta y)^2} &= G(x,y), \end{aligned} \quad (37)$$

where $x = 1, 2, \dots, n$, $y = 1, 2, \dots, m$; $\Delta x, \Delta y$ represent the horizontal spacing and vertical spacing of the control points, respectively; $F(x, y)$ and $G(x, y)$ are the divergence and curl of the sampling point (x, y) , respectively. Assuming that the number of feature points of a structured light pattern is $m \times n$, then mn sets of linear equations can be constructed using Equation (33). Then, the Poisson equation can be further expressed in the form of a matrix:

$$\begin{cases} \Phi * D = F \\ \Phi * R = G \end{cases}, \quad (38)$$

where Φ are a rectangular sparse matrix of size $mn \times mn$; D, R, F , and G are the vector forms of the divergence potential, curl potential, divergence, and curl of the two-dimensional vector field ξ , respectively, with a size of $mn \times 1$. Finally, D and R can be calculated using the generalized inverse of the matrix, specifically

$$\begin{cases} D = \Phi^{-1} F \\ R = \Phi^{-1} G \end{cases}, \quad (39)$$

where Φ^{-1} is the generalized inverse matrix of Φ .

4. Results and Discussion

In the experiment, we first set up a simplified water–air imaging platform with structured light projection and then tested the proposed method on MATLAB (Math Work Co., Natick, MA, USA) with real through-water scene images. The datasets can be acquired from [54]. Moreover, we tested the real-scene sequences of a certain length to verify the stability of the algorithm.

4.1. Experiment Setup

As Figure 10 shows, we first built an experimental platform for imaging through water in the laboratory. The experimental setup included: an acrylic water tank with a size of $100 \times 70 \times 60$ cm, for which the water-level was 42 cm, a projector (model: LOGO 35w) that was placed at the left side of the water tank, with the vertical height from the bottom of the water tank being zero, a diffuser plate, and a camera (model: MV-EM200C). Component S and V need simultaneous imaging of the diffuser plane and airborne scene as Figure 4 shows; for ease of operation, we used a single camera to image both regions simultaneously as Figure 10 shows. System calibration utilizes Zhang's calibration method [55]. Other system parameters include: $\theta^{\text{pro}} = 40^\circ$, $h_0 = 42$ cm, and $z_h = 30$ cm; a target scene was placed at $z_a = 160$ cm above water.

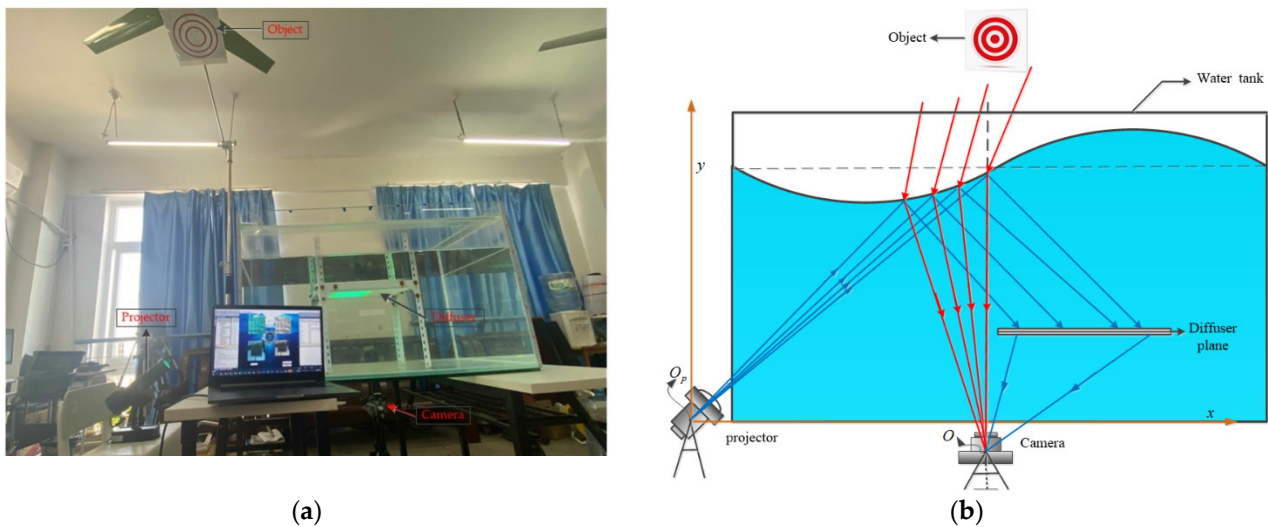


Figure 10. Experiment setup. (a) Our real laboratory scene. (b) Scheme of the experiment, comprising a projector, diffuser plane, a viewing camera, and a water tank.

In the experiment, the water surface ripples in the real environment were simulated using the artificial wave-making method, in which the agitation causes the water surface to generate natural oscillations close to surface waves. A sample image acquired by the camera in the presence of surface waves is shown in Figure 11, which is composed of two parts. One part is the distorted structured light image, and the other is the corresponding scene image above water. Both pass through the same area on the water surface.

4.2. Image Restoration

The processing procedure of the proposed method is shown in Figure 12. In order to simplify the system operation, the camera firstly collected the structured light projection on the WAI (a diffusion plate was placed on WAI) and the structured light image on the diffusion plane to obtain the WAI sampling information and the reference structured light image when the water surface was calm. Of course, we can also obtain the above two images using a full computer system simulation, once the system parameters are known. The results are shown in Figure 12a,b. Then, feature extraction and matching were performed on the reference structured light image and the distorted structured light image (Figure 13c), respectively, and the deformation vector field of the control points in the scene image were estimated. Finally, a distortion-free image was generated using the method described in Section 3.5 (Figure 13f). Intuitively, it can be seen that the distortion of the details of the target scene has been corrected, which indicates that our method can significantly reduce the distortion.

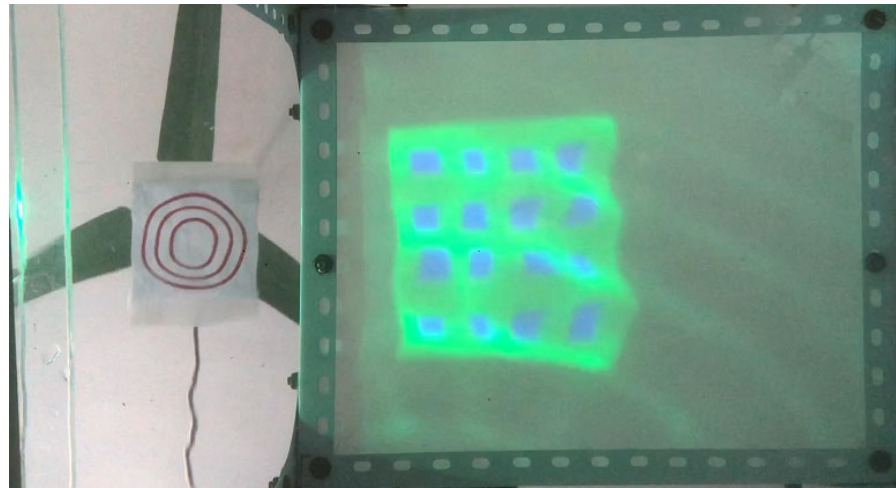


Figure 11. A real scene image captured by the camera.

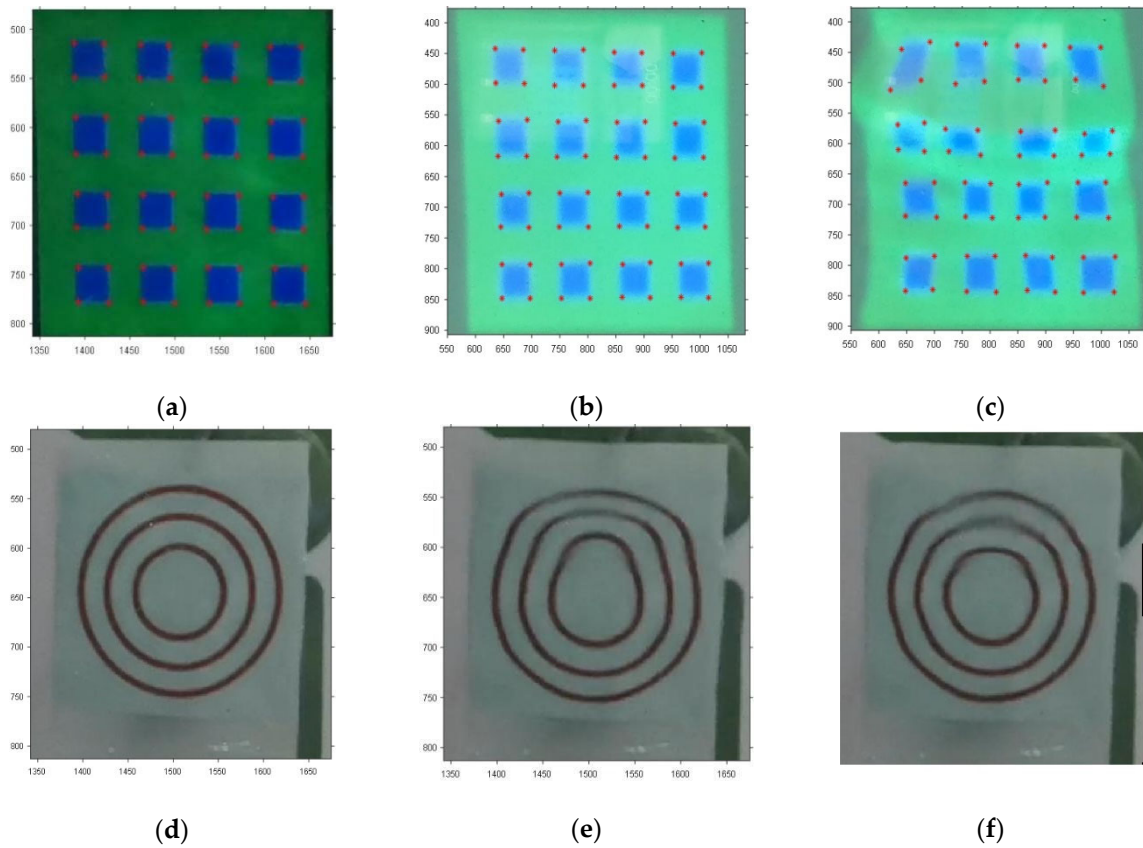


Figure 12. Processing image. (a) WAI sampling. (b) Reference structured light image as the WAI was flat. (c) Distorted structured light image. (d) Ground-truth image. (e) Distorted image. (f) Recovered image using our method.

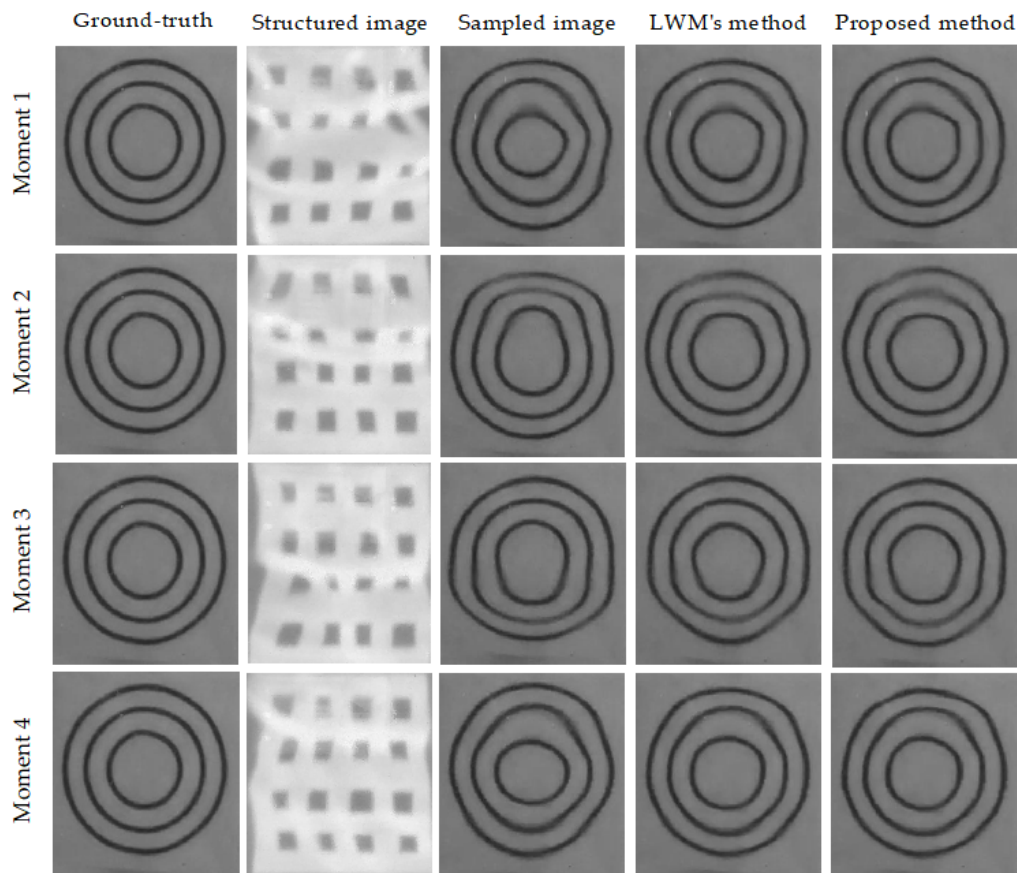


Figure 13. Recovery results of different methods. Left to right: the ground-truth image, the distorted structured light image, the sampled image at the corresponding moment, the results of the LWM method [51], and the results of our method. Up to down: recovery results of sampled images at different times using the above two methods.

4.3. Image Quality Evaluation

To verify the performance of the proposed method, we first made a comparison with the same type of method, namely the LWM method [56,57], which also requires only a single distorted image. They used the same datasets from the real through-water scenes [54].

Moreover, we utilized the three standard image quality metrics, namely GMSD [10], PSNR [11], and SSIM [14], to evaluate the recovery results of different methods. This paper randomly selected some sampling images from the real scene sequences [54] for testing, in which the scene image size of interest was 256×256 . Figure 13 shows the restoration results of the two algorithms, and Table 1 presents the results of the numerical comparison in Figure 13. It was obvious that both methods can significantly reduce the distortions of the target scene images as Figure 13 shows. However, the performance of the proposed method outperforms the LWM method significantly [57], as shown in Table 1.

Table 1. Numerical comparison among sampled images, the LWM method, and our method.

	SSIM(H)			GMSD (L)			PSNR (H)		
	Samples	LWM	Ours	Samples	LWM	Ours	Samples	LWM	Ours
Moment 1	0.7702	0.8647	0.9088	0.2171	0.1635	0.1232	22.8897	26.3590	28.7916
Moment 2	0.6978	0.8860	0.9037	0.2535	0.1508	0.1322	21.2091	27.2867	28.6392
Moment 3	0.7890	0.8506	0.8760	0.2087	0.1752	0.1684	23.2404	25.5079	26.7243
Moment 4	0.7606	0.8780	0.8905	0.2283	0.1565	0.1445	22.6351	27.0957	27.8371

In order to more accurately evaluate the performance of our method and its scope of application, we further compared our experiment results with other state-of-the-art methods, such as Oreifej's Method [7], Tao Sun's method [11], and James's methods [19]. It should be noted that these methods [7,11,19] require that the input parameter is not a single image but an image sequence, while our method is only used to recover the single instantaneous scene image. Considering the recovery result of the same sampled frame, we took the sampled frame at a specific moment as the center, selected a certain length of image sequence forward or backward as the input sequence of the previous methods [7,11,19], and then compared the restoration result of the corresponding frame with our method. Moreover, we also considered the image processing time to evaluate the computing efficiency, in which all of the above methods were performed in the same operating environment, which contains a CPU (i7-6700HQ) and RAM (8GB).

All of the above methods were tested using the same source data form [54], for which the size of the target scene image was 256×256 . The maximum number of iterations for Oreifej's [7] and T. Sun's method [11] was set to 3 (after 3 iterations for our dataset, the results have been stabilized), and each patch in the related sequence was set as 74×74 in [11].

The recovery results of different methods are shown in Figure 14. It can be seen from the figure that all methods are able to effectively reduce the distortions to the image; however, due to the respective limitations of the algorithms, the restoration results still have more or fewer residual distortions. The numerical comparison results of different methods are displayed in the Table 2. Here, we focus on analyzing the experimental results from the two aspects of restoration accuracy and computational efficiency.

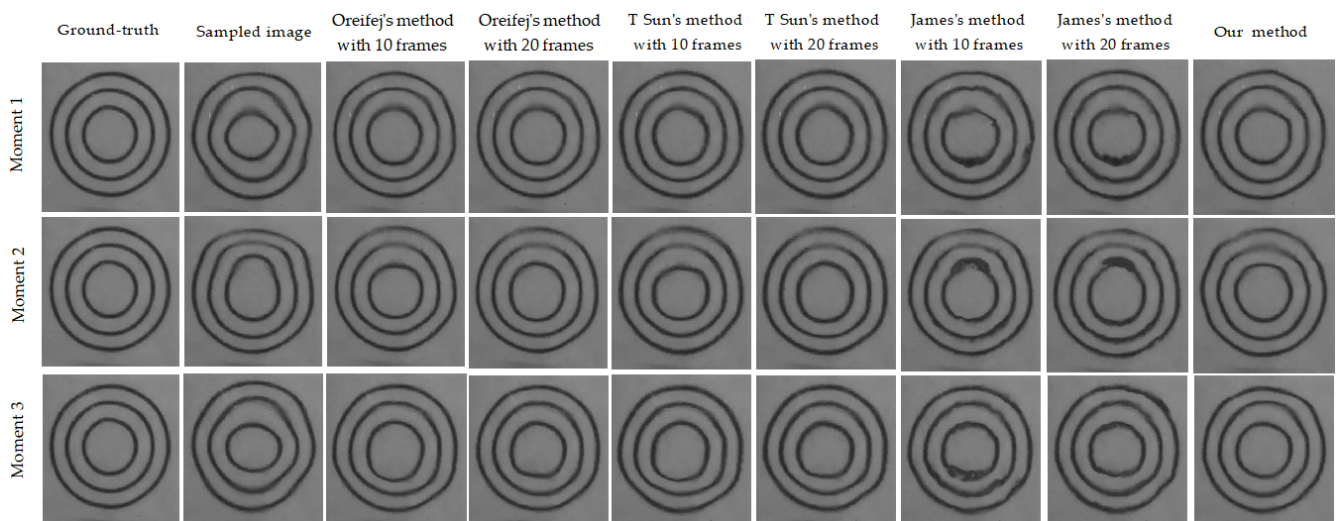


Figure 14. Recovery results of different methods. Left to right: the ground-truth image, the sampled image, the results of Oreifej's method using 10 frames, the results of Oreifej's method using 20 frames, the results of T. Sun's method using 10 frames, the results of T. Sun's method using 20 frames, the results of James's method using 10 frames, the results of James's method using 20 frames, and the results of our method using a single image.

Table 2. Numerical comparison based on the real scene datasets [51] (bold values are the best).

	Methods	Sequence Length	GMSD(L)	SSIM(H)	PSNR(H)	Time(L)
Moment 1	Oreifej [7]	10	0.1419	0.8890	27.5162	141.1623 s
	T. Sun [11]		0.1607	0.8677	26.1908	167.5526 s
	James [19]		0.1674	0.8596	25.9702	9.0383 s
	Oreifej [7]	20	0.1274	0.9028	28.3086	349.6122 s
	T. Sun [11]		0.1397	0.8898	27.2052	341.5579 s
	James [19]		0.1545	0.8739	26.5241	11.0007 s
	Our method	1	0.1232	0.9088	28.7916	0.9686 s
Moment 2	Oreifej [7]	10	0.1430	0.8883	27.3552	177.2077 s
	T. Sun [11]		0.1514	0.8779	26.5544	168.7399 s
	James [19]		0.1711	0.8591	25.5864	3.4897 s
	Oreifej [7]	20	0.1658	0.8690	26.1154	350.2627 s
	T. Sun [11]		0.1694	0.8582	25.5472	339.0449 s
	James [19]		0.2017	0.8271	24.2426	6.4821 s
	Our method	1	0.1322	0.9037	28.6392	1.0178 s
Moment 3	Oreifej [7]	10	0.1710	0.8549	25.4536	177.7958 s
	T. Sun [11]		0.1911	0.8262	24.3011	182.5569 s
	James [19]		0.1761	0.8543	26.2393	3.5057 s
	Oreifej [7]	20	0.1395	0.8904	27.9088	351.2541 s
	T. Sun [11]		0.1661	0.8615	26.1695	362.3698 s
	James [19]		0.1620	0.8738	26.8630	4.9811 s
	Our method	1	0.1445	0.8905	27.8371	1.0399 s

First, we discuss the restoration accuracy of the methods. Compared with other methods, the results of Oreifej's method and T. Sun's method seem to perform better in terms of visual effects; as shown in Figure 14, the edge details of the target scene are smoother. However, surprisingly, their restoration accuracy is not the best, as shown in Table 2. The main reasons for this result can be attributed to the following two aspects: (1) the length of the image sequence is not long enough, resulting in fluctuations in the restoration accuracy; (2) the difference between frames is large, resulting in serious degradation of the mean image or fusion image, reducing the restoration accuracy. In recent years, deep learning theory has been used as a brand-new technology to solve such problems. As a model based on deep learning, James's method also requires a long video sequence of a static target scene for accurate reconstruction. Obviously, as shown in Figure 14, it is difficult for James's method to generate a high-quality result using a shorter sequence, such as 10 or 20 frames. In contrast, we attempt to solve a single image distortion-removing problem through the theory of a vector field decomposition and a distorted structured image. Our method cleverly utilizes the characteristics of structured-light images to achieve periodic sampling of the WAI, which can effectively estimate the displacement field of the corresponding target scene image. It can be seen from Table 2 that compared with other state-of-the-art methods, in terms of restoration accuracy, our method performs better.

Moreover, we also compared our experimental results from the perspective of computational efficiency. The results show that our method is significantly better than other methods. Such as the Moment 1, to generate the distortion-free image of the same time, Oreifej's method spends 141.1623 s with 10 frames, T. Sun's method spends 167.5526 s, James's method spends 9.0383 s, and our method only takes 0.9686 s, which means that our method can be better applied to real-time observation scenarios.

In summary, experimental results show that compared with other state-of-the-art methods, the proposed method can not only significantly reduce the distortion, but also significantly outperform other methods in terms of computational efficiency. Therefore, this means that the proposed method can be applied to scenarios of dynamic object detection,

such as marine biological research, airborne target monitoring, obstacle avoidance of submersibles, etc.

4.4. Robustness Analysis of Our Method

In order to further verify the reliability of the proposed method, 50 distorted frames were randomly selected from the real scene sequence [54] for processing, and the image quality metrics (including SSIM, MSE, PSNR, and Times) were statistically averaged. As described above, the algorithm was also run on the same laptop computer. Table 3 shows the processing results of the proposed method and LWM method [57]. The experimental results show that the algorithm in this paper shows good reliability, and the algorithm is obviously better than the LWM method in terms of restoration accuracy and computational efficiency.

Table 3. Comparison results of robustness of LWM method and our method.

	SSIM(H)	GMSD(L)	PSNR(H)	Times (s)
Samples	0.7733	0.2238	22.8823	-
LWM [57]	0.8887	0.1443	27.7107	0.8946
Our method	0.8962	0.1384	28.0632	0.2319

4.5. Analysis of Corner-Detection Algorithm

As we all know, the accuracy of feature point extraction will affect the quality of the restoration results of an image. In [58], a corner detection algorithm based on Euclidean distance was used to solve the feature extraction problem of distorted structured light images, and compared with other classic algorithms, such as SURF, Harris, and SUSAN, it was proven that the method performs better than the classic algorithms. However, as shown in Figure 15, we found that there are still misdetection problems (purple marks) in the detection results, which directly reduce the restoration accuracy and visualization effect of the corresponding scene image. On the contrary, the corner detection scheme proposed in this paper can effectively correct this problem (red mark), and compared with the previous work, the image restoration results are significantly better. Moreover, compared with the ground-truth image, there are still several residual distortions in our results, which means that we can further improve our corner-detection scheme to optimize our restoration result.

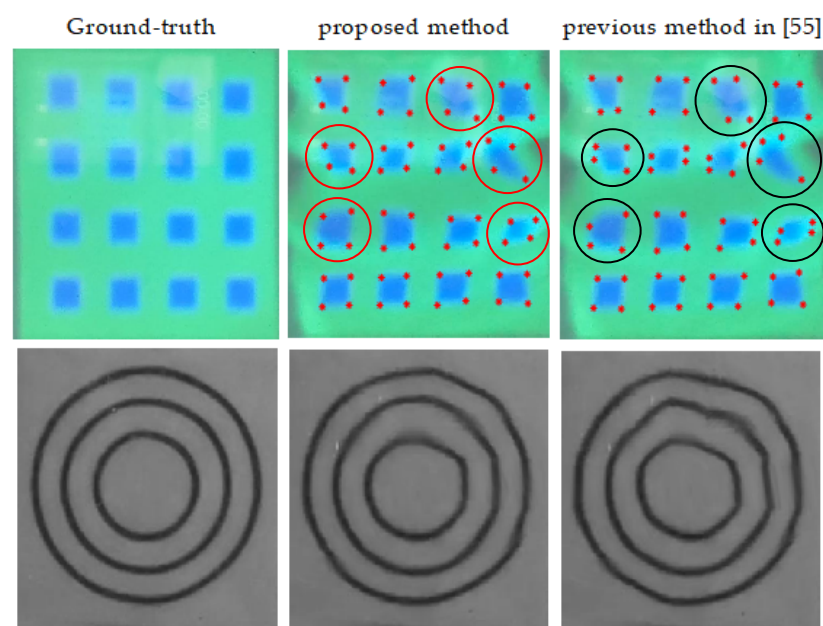


Figure 15. An example of recovery results using the different corner-detection algorithms.

4.6. Analysis of the Limitations of the Algorithm

Figure 16 demonstrates an example of the recovery result of our method. It is obvious that our method can effectively remove the geometric distortion of the image, as shown in Figure 16; however, the removal results for motion blur seem to be not as satisfactory (red marks). In future work, we need to further study additional image deblurring methods to improve and compensate for the shortcomings of our algorithm.

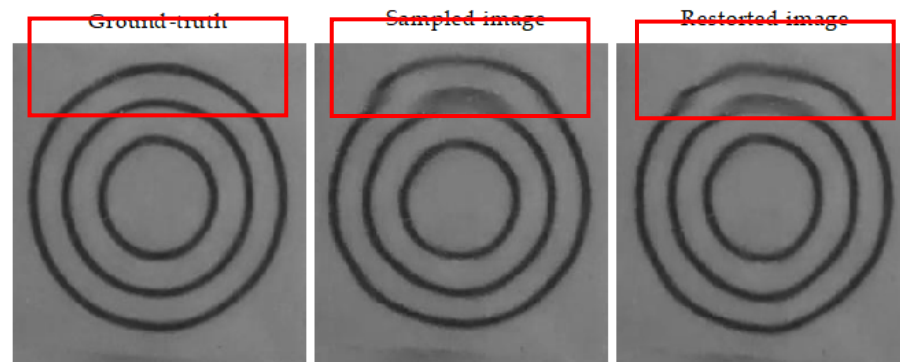


Figure 16. Recovery result of a sampled frame using the proposed method.

5. Conclusions

This paper proposed a based Helmholtz–Hodge Decomposition recovery method for the instantaneous images distorted by surface waves. Compared with previous studies, we do not need long video sequences of static scenes [7–17], large training datasets [18–20], special illuminate conditions [21,23,29,31], or multiple viewpoints [30]; we only need a simple projection device and a single scene image.

We conducted extensive tests on our method and other state-of-the-art methods [7,11,19,57] using the real-scene dataset from [54]. Experimental results show that our method not only effectively reduces the distortion to the image, but also significantly outperforms state-of-the-art methods in terms of computational efficiency. This means that our method can be applied to a wider range of practical scenarios, especially for dynamic target detection.

Moreover, we found that the accuracy of the Helmholtz–Hodge Decomposition of the vector field is directly related to the sampling interval of the WAI (described in Section 3.5). Once the complexity of image distortion increases, with the fixed interval structured light projection method, it is obviously difficult to guarantee the accuracy of image restoration. It is necessary to further study the method of adaptive structured light projection in the next work.

Author Contributions: All authors have made significant contributions to this paper. B.J. designed the research, conducted experiment analysis, and wrote the manuscript; C.M. provided guidance; Y.S. conducted data curation; D.Z. conducted validation; X.T. investigation; J.A. provided supervision and funding. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Guangxi National Science Foundation, grant number 2018GXNSFAA294056 and the Guangxi Young and Middle-Aged Teachers’ Basic Research Ability Improvement Project, grant number 2022KY0703.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used to support the findings of this study are available from Ref. [54].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alterman, M.; Schechner, Y.Y.; Perona, P.; Shamir, J. Detecting motion through dynamic refraction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 245–251. [\[CrossRef\]](#)
2. Molkov, A.A.; Dolin, L.S. The Snell's window image for remote sensing of the upper sea layer: Results of practical application. *J. Mar. Sci. Eng.* **2019**, *7*, 70. [\[CrossRef\]](#)
3. Li, T.; Yang, Q.; Rong, S.; Chen, L.; He, B. Distorted underwater image reconstruction for an autonomous underwater vehicle based on a self-attention generative adversarial network. *Appl. Opt.* **2020**, *59*, 10049–10060. [\[CrossRef\]](#)
4. Cortese, A.R.B.; Reimchen, T.E. View from below: Experimental manipulation of water surfaces to evaluate visual integrity of heron-like models through Snell's window. *Biol. J. Linn. Soc.* **2022**, *138*, 132–140. [\[CrossRef\]](#)
5. Zhang, R. Modeling and simulation of imaging through the water-air Snell's window from underwater // Eighth Symposium on Novel Photoelectronic Detection Technology and Applications. *SPIE* **2022**, *12169*, 2799–2808.
6. Jian, B.; Ma, C.; Zhu, D.; Sun, Y.; Ao, J. Seeing through Wavy Water–Air Interface: A Restoration Model for Instantaneous Images Distorted by Surface Waves. *Future Internet* **2022**, *14*, 236. [\[CrossRef\]](#)
7. Oreifej, O.; Shu, G.; Pace, T.; Shah, M. A two-stage reconstruction approach for seeing through water. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; pp. 1153–1160.
8. Kanaev, A.V.; Hou, W.; Restaino, S.R.; Matt, S.; Gładysz, S. Restoration of images degraded by underwater turbulence using structure tensor oriented image quality (STOIQ) metric. *Opt. Express* **2015**, *23*, 17077–17090. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Zhang, R.; He, D.; Li, Y.; Huang, L.; Bao, X. Synthetic imaging through wavy water surface with centroid evolution. *Opt. Express* **2018**, *26*, 26009–26019. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Zhang, Z.; Yang, X. Reconstruction of distorted underwater images using robust registration. *Opt. Express* **2019**, *27*, 9996–10008. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Sun, T.; Tang, Y.; Zhang, Z. Structural information reconstruction of distorted underwater images using image registration. *Appl. Sci.* **2020**, *10*, 5670. [\[CrossRef\]](#)
12. Cai, C.; Meng, H.; Qiao, R.; Wang, F. Water–air imaging: Distorted image reconstruction based on a twice registration algorithm. *Mach. Vis. Appl.* **2021**, *32*, 1–11. [\[CrossRef\]](#)
13. Zhang, Z.; Tang, Y.G.; Yang, K. A two-stage restoration of distorted underwater images using compressive sensing and image registration. *Adv. Manuf.* **2021**, *9*, 273–285. [\[CrossRef\]](#)
14. Jian, B.; Ma, C.; Zhu, D.; Huang, Q.; Ao, J. Water–Air Interface Imaging: Recovering the Images Distorted by Surface Waves via an Efficient Registration Algorithm. *Entropy* **2022**, *24*, 1765. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Tian, Y.; Narasimhan, S.G. Seeing through water: Image restoration using model-based tracking. In Proceedings of the IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 27 September–4 October 2009.
16. Tian, Y.; Narasimhan, S.G. Globally optimal estimation of nonrigid image distortion. *Int. J. Comput. Vis.* **2012**, *98*, 279–302. [\[CrossRef\]](#)
17. Seemakurthy, K.; Rajagopalan, A.N. Deskewing of underwater images. *IEEE Trans. Image Process.* **2015**, *24*, 1046–1059. [\[CrossRef\]](#)
18. Li, Z.; Murez, Z.; Kriegman, D.; Ramamoorthi, R.; Chandraker, M. Learning to see through turbulent water. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 512–520.
19. James, J.G.; Agrawal, P.; Rajwade, A. Restoration of non-rigidly distorted underwater images using a combination of compressive sensing and local polynomial image representations. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 7839–7848.
20. Thapa, S.; Li, N.; Ye, J. Learning to Remove Refractive Distortions from Underwater Images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 5007–5016.
21. Milder, D.M.; Carter, P.W.; Flacco, N.L.; Hubbard, B.E.; Jones, N.M.; Panici, K.R.; Platt, B.D.; Potter, R.E.; Tong, K.W.; Twisselmann, D.J. Reconstruction of through-surface underwater imagery. *Waves Random Complex Media* **2006**, *16*, 521–530. [\[CrossRef\]](#)
22. Schultz, H.; Corrada-Emmanuel, A. System and Method for Imaging through an Irregular Water Surface. U.S. Patent No. 7,630,077, 8 December 2009.
23. Levin, I.M.; Savchenko, V.V.; Osadchy, V.J. Correction of an image distorted by a wavy water surface: Laboratory experiment. *Appl. Opt.* **2008**, *47*, 6650–6655. [\[CrossRef\]](#)
24. Weber, W.L. Observation of underwater objects through glitter parts of the sea surface. *Radiophys. Quantum Electron.* **2005**, *48*, 34–47. [\[CrossRef\]](#)
25. Dolin, L.S.; Luchinin, A.G.; Turlaev, D.G. Algorithm of reconstructing underwater object images distorted by surface waving. *Izvestiya. Atmos. Ocean. Phys.* **2004**, *40*, 756–764.
26. Luchinin, A.G.; Dolin, L.S.; Turlaev, D.G. Correction of images of submerged objects on the basis of incomplete information about surface roughness. *Izvestiya. Atmos. Ocean. Phys.* **2005**, *41*, 247–252.
27. Dolin, L.; Gilbert, G.; Levin, I.; Luchinin, A. *Theory of Imaging through Wavy Sea Surface*; IAP RAS: Nizhny Novgorod, Russia, 2006.
28. Dolin, L.S.; Luchinin, A.G.; Titov, V.I. Correcting images of underwater objects distorted by sea surface roughness[C] // Current Research on Remote Sensing, Laser Probing, and Imagery in Natural Waters. *SPIE* **2007**, *6615*, 181–192.
29. Alterman, M.; Swirski, Y.; Schechner, Y.Y. STELLA MARIS: Stellar marine refractive imaging sensor. In Proceedings of the IEEE International Conference on Computational Photography (ICCP), Santa Clara, CA, USA, 2–4 May 2014.

30. Alterman, M.; Schechner, Y.Y. 3D in natural random refractive distortions. *Int. Soc. Opt. Photonics* **2016**, 9867, 64–75.
31. Gardashov, R.H.; Gardashov, E.R.; Gardashova, T.H. Recovering the instantaneous images of underwater objects distorted by surface waves. *J. Mod. Opt.* **2021**, 68, 19–28. [\[CrossRef\]](#)
32. Gong, B.; Wang, G. Underwater 2D Image Acquisition Using Sequential Striping Illumination. *Appl. Sci.* **2019**, 9, 2179. [\[CrossRef\]](#)
33. Hurtado-Aviles, E.A.; Vila, M.; Vilatela, J.J.; Martines-Arano, H.; Bornacelli, J.; García-Merino, J.A.; Cervantes-Sodi, F.; Torres-Torres, C. Structured light using carbon nanostructures driven by Kerr nonlinearities and a magnetic field. *Phys. Chem. Chem. Phys.* **2022**, 24, 1081–1090. [\[CrossRef\]](#)
34. Zhang, B.; Lin, S.; Lin, J.; Jiang, K. Single-shot high-precision 3D reconstruction with color fringe projection profilometry based BP neural network. *Opt. Commun.* **2022**, 517, 128323. [\[CrossRef\]](#)
35. Lynch, D.K. Snell's window in wavy water. *Appl. Opt.* **2015**, 54, B8–B11. [\[CrossRef\]](#)
36. Hu, P.; Yang, J.; Guo, L.; Yu, X.; Li, W. Solar-tracking methodology based on refraction-polarization in Snell's window for underwater navigation. *Chin. J. Aeronaut.* **2022**, 35, 380–389. [\[CrossRef\]](#)
37. Born, M.; Emil, W. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*; Elsevier: Amsterdam, The Netherlands, 2013.
38. Holthuijsen, L.H. *Waves in Oceanic and Coastal Waters*; Cambridge University Press: Cambridge, UK, 2010.
39. Masselink, G.; Hughes, M.; Knight, J. *Introduction to Coastal Processes and Geomorphology*; Routledge: London, UK, 2014.
40. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis, and Machine Vision*; Cengage Learning: Boston, MA, USA, 2014.
41. Partheepan, R.; Raja Paul Perinbam, J.; Krishnamurthy, M.; Shanker, N.R. Visualization of dentate nucleus, pontine tegmentum, pontine nuclei from CT image via nonlinear perspective projection. *Signal Image Video Process.* **2022**, 16, 137–145. [\[CrossRef\]](#)
42. Ryu, J.B.; Park, H.H.; Park, J. Corner classification using Harris algorithm. *Electron. Lett.* **2011**, 47, 536–538. [\[CrossRef\]](#)
43. Jian, C.; Xiang, X.; Zhang, M. Mobile terminal gesture recognition based on improved FAST corner detection. *IET Image Process.* **2019**, 13, 991–997. [\[CrossRef\]](#)
44. Grycuk, R.; Knop, M.; Mandal, S. *Video Key Frame Detection Based on SURF Algorithm*; Springer: Cham, Switzerland, 2015.
45. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, 35, 1397–1409. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, 9, 62–66. [\[CrossRef\]](#)
47. Bhatia, H.; Norgard, G.; Pascucci, V.; Bremer, P.T. The Helmholtz-Hodge decomposition—A survey. *IEEE Trans. Vis. Comput. Graph.* **2012**, 19, 1386–1404. [\[CrossRef\]](#)
48. Suda, T. Application of Helmholtz–Hodge Decomposition to the study of certain vector fields. *J. Phys. A Math. Theor.* **2020**, 53, 375703. [\[CrossRef\]](#)
49. Caltagirone, J.P. On a reformulation of Navier–Stokes equations based on Helmholtz–Hodge decomposition. *Phys. Fluids* **2021**, 33, 063605. [\[CrossRef\]](#)
50. Vieira, M.B.; Giralaldi, G.A.; Ribeiro, A.C.A.; Renhe, M.C.; Esperança, C. Anisotropic Helmholtz decomposition for controlled fluid simulation. *Appl. Math. Comput.* **2021**, 411, 126501. [\[CrossRef\]](#)
51. Helmholtz, H. Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen. *J. Die Reine Angew. Math.* **2009**, 1858, 25–55.
52. Gao, S.; Gruev, V. Bilinear and bicubic interpolation methods for division of focal plane polarimeters. *Opt. Express* **2011**, 19, 26161–26173. [\[CrossRef\]](#)
53. Oberman, A.M. Finite difference methods for the infinity Laplace and p-Laplace equations. *J. Comput. Appl. Math.* **2013**, 254, 65–80. [\[CrossRef\]](#)
54. Jian, B. Imaging through wavy water. *Figshare* **2022**. [\[CrossRef\]](#)
55. Zhang, Z. Flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, 22, 1330–1334. [\[CrossRef\]](#)
56. Goshtasby, A. Image registration by local approximation methods. *Image Vis. Comput.* **1988**, 6, 255–261. [\[CrossRef\]](#)
57. Ma, C.; Zhu, D.; Ao, J.; Jian, B.; Sun, Y. Jian B.; Sun, Y. An Image Restoration Method Based on Local Surface Fitting. Patent CN114022385A, 8 February 2022.
58. Mai, C.; Jian, B.; Ling, Y. A Corner Detection Algorithm Based on Regional Center of Mass in Imaging through Water Surface. *J. Phys. Conf. Series.* **2021**, 2083, 032090. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.