

Article

Structural Information Reconstruction of Distorted Underwater Images Using Image Registration

Tao Sun, Yugui Tang and Zhen Zhang * 

School of Mechatronic Engineering and Automation, Shanghai University, Nanchen Road 333, Shanghai 200444, China; suntao@shu.edu.cn (T.S.); tangyugui@shu.edu.cn (Y.T.)

* Correspondence: zhangzhen_ta@shu.edu.cn

Received: 18 July 2020; Accepted: 13 August 2020; Published: 15 August 2020



Featured Application: The proposed method can be applied to the restoration for turbulence degraded video in nuclear power plant reactors, underwater target or aerial target monitoring, and surveillance of shallow riverbeds to observe vegetation.

Abstract: Imaging through wavy air-water surface suffers from uneven geometric distortions and motion blur due to surface fluctuations. Structural information of distorted underwater images is needed to correct this in some cases, such as submarine cable inspecting. This paper presents a new structural information restoration method for underwater image sequences using an image registration algorithm. At first, to give higher priority to structural edge information, a reference frame is reconstructed from the sequence frames by a combination of lucky patches chosen and the guided filter. Then an iterative robust registration algorithm is applied to remove the severe distortions by registering frames against the reference frame, and the registration is guided towards the sharper boundary to ensure the integrity of edges. The experiment results show that the method exhibits improvement in sharpness and contrast, especially in some structural information such as text. Furthermore, the proposed edge-first registration strategy has faster iteration velocity and convergence speed compared with other registration strategies.

Keywords: image restoration; guided filter; underwater images; image registration

1. Introduction

When imaging through a fluctuating air-water surface using a fixed camera, the images acquired by observing the fixed objects that are completely immersed in the water, often suffer from severe distortions due to the water surface fluctuation. Furthermore, refraction occurs when light passes through the air-water interface, especially on the wavy water surface [1,2]. The refraction angle is affected by the refractive index of the medium and angle of incidence. In addition, the attenuation derived from the scattering and absorption of the imaging path is also an important consideration that results in observation distortion. The suspended particles in the medium, such as organic matter, mineral salt, microorganisms, etc., are constantly consuming the beam energy and changing the propagation path of light [3,4]. Hence, underwater image restoration is more challenging than similar problems in other environments.

Research on this challenging problem has been carried out for decades. Some researchers consider that the distorted wavefront due to the fluctuated surface is the important factor of image degradation. Therefore, the adaptive optics method, which is popular in astronomical imaging, was used for underwater image correction at first. Holohan and Dainty [5] assumed that the distortions are mainly low-frequency aberration and phase shift, and proposed a simplified model based on adaptive optics.

There are several methods using ripple estimation technology for underwater reconstruction. A simple algorithm was proposed in [6], where it reconstructs the 3D model of water surface ripple by optical flow estimation and statistical motion features. In [7], an algorithm based on the dynamic nature of the water surface was presented, which uses cyclic waves and circular ripples to express local aberration. Tian et al. [8] proposed a model-based tracking algorithm. The distortion model, based on the wave equation, is established, then the model is fitted to frames to estimate the shape of the surface and restore the underwater scene.

According to Cox–Munk law [9], if the water surface is sufficiently large and still, the normal water surface will roughly assume a Gaussian distribution. Inspired by the law, some approaches focus on finding the center of the distribution of patches from the image sequence as the orthoscopic patch [10]. Many strategies related to lucky regions have been proposed to deal with the problem. In [11], the graph-embedding technology was applied to block frames and calculate the local distance between them, and the shortest path algorithm was used to select patches to form the undistorted image. Donate et al. [12,13] came up with a similar method, where the motion blur and geometric distortion were modeled respectively, and the K-means algorithm was applied to replace the shortest path algorithm. Wen et al. [14] combined bispectrum analysis with lucky region selection, and smooth edge transition by using patches fusion. A restoration method based on optical flow and lucky region was proposed by Kanaev et al. [15]. The selection of a lucky patch is realized by image measurement metric and calculating the nonlinear gain coefficient of the current frame to each point. Later in [16,17], Kanaev et al. improved the resolution of their algorithm by developing structure tensor oriented image quality metrics. Recently in [18], on the basis of lucky region fusion, Zhang et al. put forward to a method that the input of distorted frame and output of restored frame are carried out at the same time. The quality of reconstruction is ameliorated by successive updating of the subsequent distorted frame.

The registration technology [19–22], which is originally used to recover the atmospheric turbulence image, is applied to address this issue. Oreifej et al. [20] presented a two-stage nonrigid registration approach to overcome the structural turbulence of waves. In the first stage, a gaussian blur is added to sequence frames to improve the registration effect, and in the second stage, a rank minimization is used to dislodge sparse noise. In our previous research [21], an iterative robust registration algorithm was employed to overcome the structural turbulence of the waves by registering each frame to a reference frame. The high-quality reference frame is reconstructed by the patches selected from the sequence frames and a blind deconvolution algorithm is performed to improve the reference frame. Halder et al. [22] proposed a registration approach using pixel shift maps, where registers image sequence against the sharpest frame to obtain pixel shift maps.

Motivated by deep learning technology, Li et al. [23] introduced the trained convolution neural network to dewarp dynamic refraction, which has a larger requirement of training samples and training cycles.

Recently, James et al. [24] assumed that water fluctuation possesses spatiotemporal smoothness and periodicity. Based on this hypothesis, compressive sensing technology was combined with local polynomial image representation. Later, based on the periodicity, a Fourier based pre-processing was proposed to correct the apparent distortion [25].

Most of the above studies were performed to restore the whole image from an underwater sequence. However, some structural information of the image is expected to be obtained in some cases, such as underwater cable number observing. In this paper, a new image restoration approach for underwater image sequences using an image registration algorithm is proposed. In our approach, the lucky patches fusion was employed to discard the patches with severer warping, then a guided filter algorithm was used to enhance object boundary in the fused image. An iterative registration algorithm was applied to remove most of the distortions in the frames, which registers frames against the enhanced image. After the registration process, the unstructured sparse noises were eliminated by principal component analysis (PCA) and patches fusion technology to produce undistorted image

sequences and frames. Experiments present that the proposed method has better performance on structural information reconstruction [20,21].

The remaining part of this paper proceeds as follows: Section 2 presents an overview of the proposed method and introduces it in detail, Section 3 shows the results of the experiment, Section 4 analyzes and discusses the experimental results, and Section 5 concludes the paper.

2. Methods

2.1. Overview

In this section, we propose a new image restoration method, which is illustrated as the flow chart in Figure 1. The primary object of the proposed method is to recover an undistorted sequence $V_f = \{I_{f1}, I_{f2}, \dots, I_{fn}\}$ and a dewarped frame F from a distorted underwater sequence $V_d = \{I_{d1}, I_{d2}, \dots, I_{dn}\}$, where $V \in \mathbb{R}^{h \times w \times n}$. The restoration algorithm is composed of three major parts:

1. Lucky patches fusion: the severely distorted image patches within each frame and across the entire image sequence are clustered and discarded, and the patches with less distortion are selected to fuse a reference image.
2. Guided filter: a guided filter algorithm is employed to preserve the edge and smooth the non-edge of the fused image.
3. Image registration: to dewarp the frames sequence, a nonrigid registration is applied to register frames against the reference R .

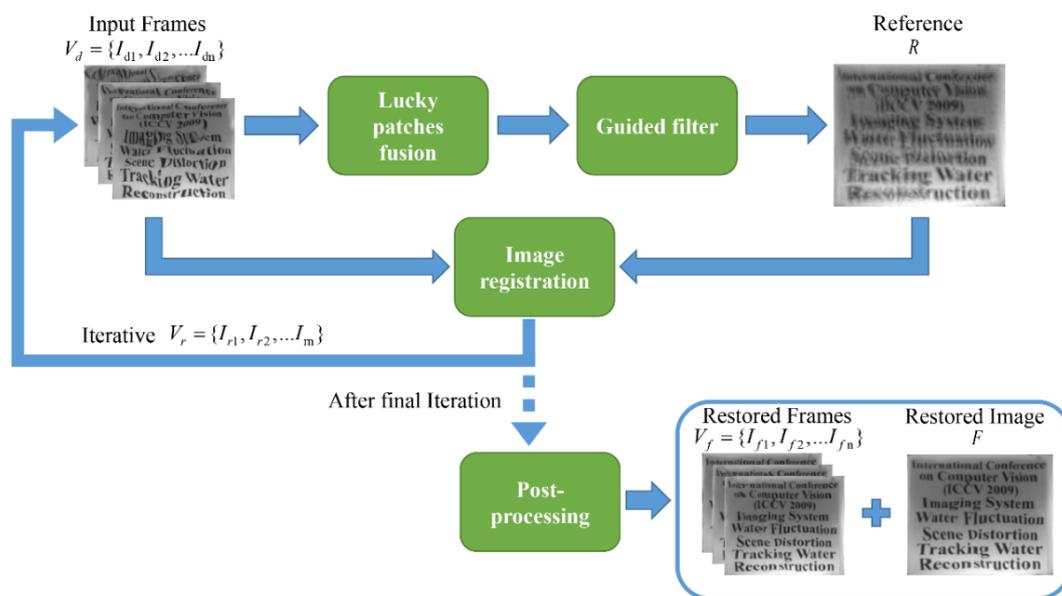


Figure 1. Simple flow diagram of the proposed method.

After applying the registration algorithm several times, the frames obtained still contain several misalignments and residual noises. The robust PCA technique is employed as the post-processing to refine the frames, and patches fusion is used to obtain a single dewarped image.

2.2. Principles

On the basis of Cox–Munk law [9], for the image sequence observed over a period of time, each pixel of the mean should be its real position. However, when the water surface is severely disturbed, some points in the simple mean deviate from correct positions because of the seriously uneven shift. On account of this, we intend to add patches selection before getting the mean image, only the less

distorted parts of the sequence are chosen to form a new mean frame with better quality. Then, in view of the clear structural edge is more contributing to improving image quality and visual feeling, we intend to enhance the boundary information of the new mean so that it guides the registration toward the more interesting place.

2.3. Lucky Patches Fusion

At first, the input frames are subdivided into a group of smaller patches with the same size. The overlap area of each pair of neighboring patches is chosen to 50% [21]. Then the K-means algorithm is applied to cluster patches into two groups in accordance with warp intensity. Finally, the less warped patches within each patch sequence, are picked out and fused into a single frame, while the rest patches are discarded.

The warp intensity is quantified by the structure similarity index (SSIM) [26], which represents the deviation of the patch I from the mean M of the current patch sequence. The SSIM can be described as

$$\text{SSIM}(I, M) = \frac{(2\mu_I\mu_M + c_1)(2\sigma_{IM} + c_2)}{(\mu_I^2 + \mu_M^2 + c_1)(\sigma_I^2 + \sigma_M^2 + c_2)} \quad (1)$$

Here $\mu_I, \mu_M, \sigma_{IM}, \sigma_I, \sigma_M$ represents local means, standard deviations, and cross-covariance for I, M , respectively. c_1, c_2 are constants. In the fusion step of each patch sequence, the patch group with higher SSIM value is preserved and averaged to a patch. All the fused patches are stitched together to form an entire frame, and a two-dimensional weighted Hanning window [14] is used to mitigate the boundary effects of every patch.

2.4. Guided Filter

The edges of the fused image, which plays a critical role in the reconstruction of the text, are closer to their true positions. In order to guide the registration to focus on the more important region, a guided filter proposed in [27] is employed to enhance the text boundary. In general, a general linear translation-variant filtering process involves a guidance image G , an input image p , and an output image q . Both G and p are given beforehand according to the application, and they can be identical. The filtering output q at the position i is presented as a weighted average:

$$q_i = \sum_j W_{ij}(G)p_j \quad (2)$$

where i, j are position indexes, and W_{ij} is filter kernel related to the guidance image G , reflects the linear relationship between the output and the input image.

In the definition of the guided filter, the output q can be modeled as a linear transform of guidance G in a window ω_k centered at pixel k :

$$q_i = a_k G_i + b_k, \forall i \in \omega_k \quad (3)$$

where a_k, b_k are linear coefficients, which are constant in the window with the radius r . For the purpose of determining the linear coefficients, the cost function is set as follow:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k G_i + b_k - p_i) + \varepsilon a_k^2) \quad (4)$$

where ε is the regularization parameter that is used to distinguish the boundary. The solution to Equation (4) can be obtained by linear regression [27]. It should be noted that the guidance G and input p are identical in our algorithm, then the solution to (4) can be expressed as

$$a_k = \frac{\sigma_k^2}{\sigma_k^2 + \varepsilon} \tag{5}$$

$$b_k = (1 - a_k)\overline{p_k} \tag{6}$$

where σ_k^2 is the variance of G in ω_k , $\overline{p_k}$ is the mean of p in ω_k . Based on the Equation (5) and Equation (6), the windows can be divided into “flat area” and “high variance area”. The former means that the guidance G is approximately constant in ω_k , and the solution is $a_k = 0, b_k = \overline{p_k}$, whereas the latter implies that the intensity of guidance G changes a lot within ω_k , then the solution becomes $a_k \approx 1, b_k \approx 0$. More specifically, the ε is selected as the discrimination threshold of “flat area” and “high variance area”. Therefore, the edge patches with σ^2 much larger than ε are preserved, while non-edge patches with smaller variance are smoothed.

After computing coefficients for all windows ω_k in image, the entire output image can be given by:

$$q_i = \overline{a_i}G_i + \overline{b_i} \tag{7}$$

2.5. Image Registration

Similar to [28], a nonrigid B-spline based registration is applied to yield fine local alignment and deskew the distortion. On the basis of this algorithm, an underlying mesh of control points is driven to deform the input frames into reference. For each point (x, y) in the image, the free-form deformation can be expressed by a mesh of control points $\phi_{i,j}$ of size $n_x \times n_y$:

$$T(x, y) = \sum_{m=0}^3 \sum_{l=0}^3 B_m(v)B_l(u)\phi_{i+l,j+m} \tag{8}$$

where $i = \lfloor x/n_x \rfloor - 1, j = \lfloor y/n_y \rfloor - 1, u = x/n_x - \lfloor x/n_x \rfloor, v = y/n_y - \lfloor y/n_y \rfloor$ and B is the standard cubic B-spline based functions, are defined as $B_0(t) = (1 - t)^3/6, B_1(t) = (3t^3 - 6t^2 + 4)/6, B_2(t) = (-3t^3 + 3t^2 + 3t + 1)/6$ and $B_3(t) = t^3/6$, where $0 \leq t < 1$. According to the distance to the point (x, y) , every control point is assigned a weight to free-form deformation by these B-spline based functions. Besides, the coarse-to-fine method, which determines the resolution of controlled mesh, is employed to achieve the optimal balance between the model flexibility and computational cost [29,30]. The optimal solution of transformation can be calculated by the formulas given in the previous publication [21].

2.6. Post-Processing

The results after several time registrations are less distorted but with several misalignments and remaining sparse noises. For producing an undistorted sequence $V_f = \{I_{f1}, I_{f2}, \dots, I_{fn}\}$, the PCA algorithm is applied to dispel random noises. Furthermore, to reconstruct a single dewarped image, the lucky patches fusion is employed on V_f again.

2.7. Summary of Algorithm

The complete restoration strategy is summarized in Algorithm 1.

3. Results

In the experiment, the proposed method was implemented on MATLAB (MathWorks Co., Natick, MA, USA). The basic data sets used were the same as Tian’s in [8]. The source codes are opened online for the public and available in the following link: <https://github.com/tangyugui1998/Reconstruction-of-distorted-underwater-images>. To verify the performance of the proposed method, we made a comparison with our previous study [21] and Oreifej et al. [20], whose source codes are also available online. The same image sequence was processed by three methods individually, and for all three methods, the maximum number of iterations was set as five.

The tested data sets contain “checkboard”, “Large fonts”, “Middle fonts”, and “Small fonts”. Each data set was composed of 61 frames. The size of the frame and patch used in patches fusion was manually preset before running the algorithm, which is shown in Table 1.

Algorithm 1: Structural Information Reconstruction

Input: Distorted image sequence

$$V_d = \{I_{d1}, I_{d2}, \dots, I_{dn}\}, I_{dk} \in \mathbb{R}^{h \times w} (k = 1, 2, \dots, n)$$

Output: Undistorted image sequence and dewarped image

$$V_f = \{I_{f1}, I_{f2}, \dots, I_{fn}\}, I_{fk} \in \mathbb{R}^{h \times w} (k = 1, 2, \dots, n), F \in \mathbb{R}^{h \times w}$$

While $s \leq \text{MaximalIteration}$ **do**

Step 1: Lucky patches fusion

for each patch sequence $B^i = \{B_1, \dots, B_n\} (i = 1, \dots, N)$

$$Q = \{q_1, \dots, q_n\} \leftarrow \text{ComputeQualityMetric}(B^i);$$

$$B^* = \{B_1^*, \dots, B_n^*\} \leftarrow \text{KmeansClustering}(B^i, Q);$$

$$B_m^i \leftarrow \text{Mean}(B^*);$$

End

$$R_m \in \mathbb{R}^{h \times w} \leftarrow \text{SpliceTogether}(B_m^i);$$

Step 2: Guided filter

$$R_g \in \mathbb{R}^{h \times w} \leftarrow \text{GuidedFilter}(R_m);$$

Step 3: Image registration

$$T = \{T_1, \dots, T_n\} \leftarrow \text{ComputerDeformation}(V, T);$$

$$V_r = \{I_{r1}, I_{r2}, \dots, I_{rn}\} \leftarrow \text{Dewarp}(V, T);$$

$$s = s + 1;$$

End

Step 4: Post-processing

End

Table 1. The Frame Size and Patch Size of Different Data Sets.

	Checkboard	Large Fonts	Middle Fonts	Small Fonts
Frame size	238 × 285	238 × 324	238 × 285	238 × 285
Patch size	68 × 114	68 × 108	68 × 114	68 × 114

The registration results are shown in Figure 2. The results indicate that all three methods can compensate for the distortion, defocusing, and double image of an underwater image sequence to some extent, while our method has better performance in some details, especially on three text data sets. The adjacent letters are blurred together by the Oreifej method and the previous research, which has been marked in red rectangles. Meanwhile, there are some ghosting and misalignment in marked red regions, such as the Oreifej’s results of “Large fonts” and “checkboard” and the result of our previous method in “Large fonts”. On the contrary, the proposed method successfully avoids the above problems, and the letters in results are clearer and easier to be observed. Moreover, for some data sets such as “checkboard”, the proposed method is inferior to our previous method in regional restoration (marked with green rectangles), and some letters in the result of “Large fonts” suffered from slight aberration. Although some areas marked in green rectangles are more distorted compared with other methods, the letter observation is not affected dramatically.

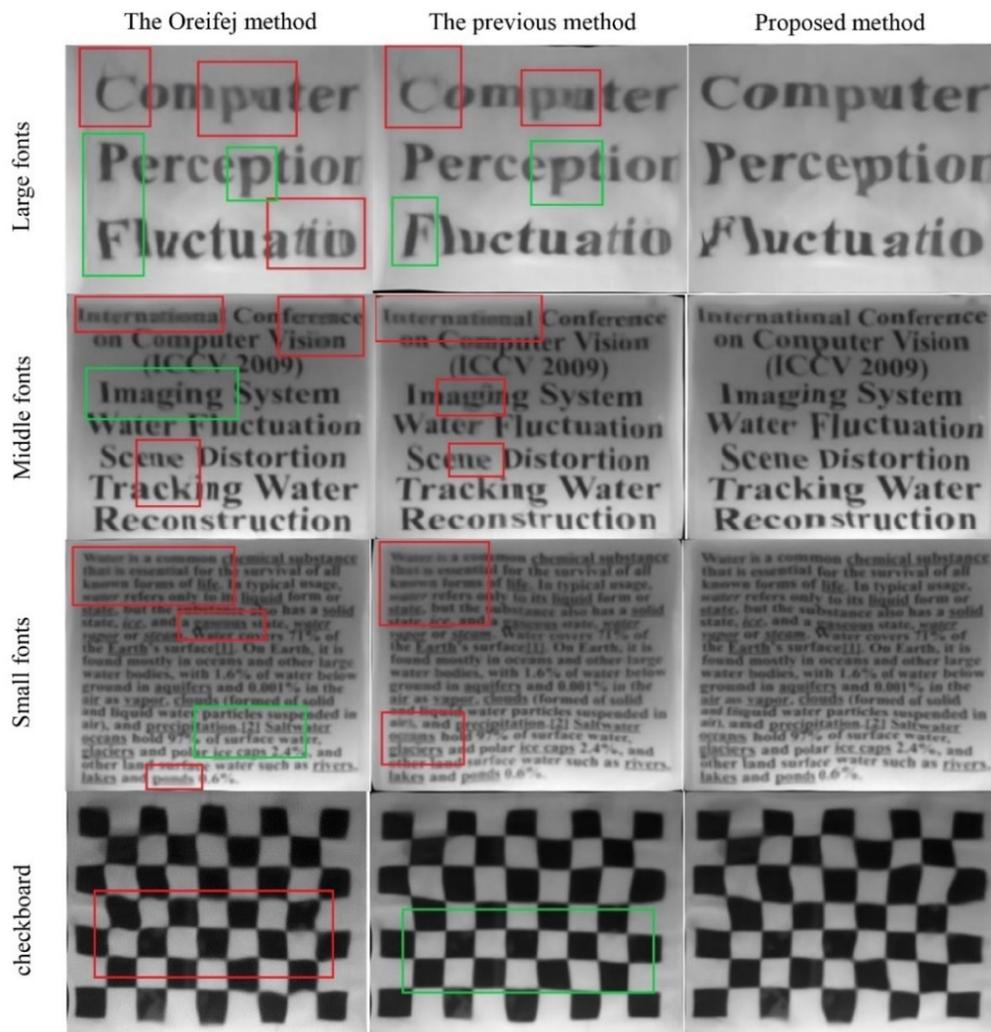


Figure 2. The recovery results of different methods after five iterations. The regions marked with red rectangles are inferior to our method. The green rectangles mark the regions with less distortion than our method.

In addition, to compare the running time of three methods, we used a laptop computer (Intel Core i5-9300H, 8 GB RAM) to measure the processing time. The “checkboard” data set was used and the running time of each step in the proposed method as follows: The patches fusion step cost 4.609 s. The guided filter step cost 0.016 s. The comparison of each registration step of the three methods is shown in Figure 3. Each iteration time of the Oreifej method fluctuated around 260 s, whereas the iteration time of the proposed method decreases with increasing of iterations, and the time was less than the Oreifej method except for in the first iteration. Furthermore, the average running time of the proposed method was 249.859 s, less than the Oreifej method, which was 263.060 s. Compared with the previous method, obviously, the proposed method had a decreasing running time as iterations went on, whereas uptime of the previous method rebounded evidently with an average of 255.522 s. Overall, the proposed method kept a clear pattern of deceleration in the iteration time, which is very useful with the improvement of processor hardware acceleration.

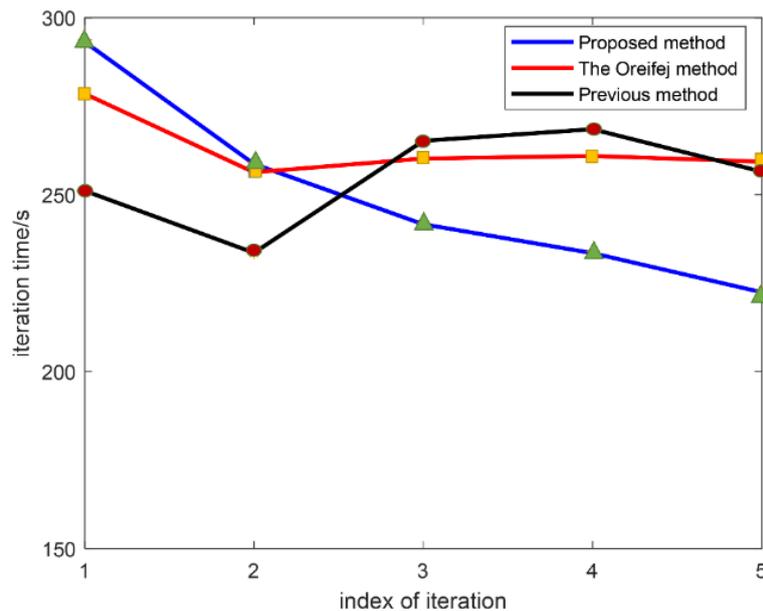


Figure 3. The running time at each iteration of different methods.

4. Discussion

4.1. Better Quality Reference Image

Usually, the real image as a reference image used in image registration can not be obtained in a real video system. So, the most important step in image registration is the selection of the reference image, as mentioned in the introduction, the mean or the sharpest frame of the image sequence is usually used as the reference image in the past. However, if the mean frame is so blurred that it causes feature loss in the image, while the sharpest frame has severe geometric distortion that cannot be ignored, the registration is seriously restricted. In order to solve the problem, Oreifej selected the mean of an image sequence as the reference and blurred the sequence frames using a blur kernel estimated from the sequence. When the reference image and sequence are at the same blur level, the whole registration process is guided to the sharper region. However, in the Oreifej method, the blurred frames tend to introduce unexpected local image quality deterioration. Some edges in sequence are blurred and shifted because of joined frame blurring, and the same regions in the mean are also ambiguous and warped. When the registration process of these blurred regions is directed to other regions, and the registration of the boundary is sacrificed. It should be noted that the loss and confusion of edge information are probably irreversible, and the distortion and misalignment may be aggravated with the increase of iterations.

In the proposed method, reconstruction of the higher quality reference frame is performed at first. The reference frame is obtained through discarding severely distorted parts and enhancing the edge. Compared with the mean frame, the reconstructed reference frame can greatly avoid introducing artificial misalignment, and its pixel points are closer to their real positions. Then, the registration is guided to the sharper boundaries with the preserve and enhancement of edge. The comparison between the reconstructed reference image with the mean reference image used in other methods described above, is shown in Figure 4. It is clear that the reference processed by the proposed method has more distinct letters edges.



Figure 4. Reference selected by different methods.

4.2. Analysis of Restoration Results

The results shown in Figure 2 suggest that the proposed method performed better than our previous study and the Oreifej method in the adjacent letter reconstruction. The mean frame of frame sequence was selected as the reference frame of the Oreifej method, and the originally sharp letter boundaries in the image sequence were destroyed by bringing blur to the frames. It is difficult in the registration process to act precisely on the edge of text, even causing some mismatches. The fuzzy and double image caused by mismatches gets worse as the iteration step goes on. Our previous study adds a deblur step to the reference image before registration, which makes the registration pay too much attention to the nonboundary part and neglects the restoration of the border. In other words, Oreifej adds blurring to the original sequence and registers the blurred image sequence against the mean frame of the original sequence. Our previous method registers the original sequence against the reconstructed image using lucky patches fusion and blind deconvolution. However, Figure 2 shows that the above methods fail to deal with the fuzzy of structural information, especially the outline of adjacent letters. We think the reason is that the regions registered preferentially, which are relatively sharper, are random and irregular because of the uncertainty of sharper region generation and lack of explicit information priority. Therefore, considering the integrity of structural edge information is useful to improve the visual effect, we decided to intentionally raise the status of edge registration and guarantee the priority of boundary reconstruction to recover the edge information to a greater extent. We spliced the patches with less distortion into an image. Then the guided filter was used to keep the boundary information with great gradient change and ensure the edges of the character were clearer than non-edges. The registration process of edges took precedence over other positions, and with increasing iterations, the reference quality was constantly improved. Nevertheless, because the nonuniform geometric degradation in the first iteration was unable to be completely compensated by lucky patches fusion, and guided filter directed the registration to focus on the structural details instead of regional distortion, some of the migrated boundaries were mistaken for correct ones and partial distortion was retained, leading to some letters suffering from deformations shown in the result of the “Large fonts”. Although there are still some local distortions in the results, slight distortions are more conducive to be observed than blurring.

To quantify the results of all methods, gradient magnitude similarity deviation (GMSD) [31], and feature similarity (FSIM) [32], were chosen as relevant quality metrics. Above metrics belong to full-reference, which reflect the difference between the evaluated image and the ideal image, were defined as follows:

GMSD is a metric that estimates image quality score by gradient magnitude, which presents structural information. The formal can be shown as:

$$\text{GMS}(i) = \frac{2m_r(i)m_d(i) + c}{m_r^2(i) + m_d^2(i) + c} \quad (9)$$

$$\text{GMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{GMS}(i) - \text{GMSM})^2} \quad (10)$$

where m_r and m_d denote gradient magnitude of the reference image and the distorted image, respectively, c represents constant, and GMSM denotes the mean of gradient magnitude similarity.

FSIM is an expression for emphasizing the structural features of visual interest, to estimate the frame quality. It can be expressed as:

$$PC_m = \max(PC_1(x), PC_2(x)) \quad (11)$$

$$\text{FSIM} = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (12)$$

where PC represents phase congruency, Ω denotes the entire image, and S_L denotes similarity that is determined by phase congruency similarity and GMS.

Furthermore, the underwater image quality measure (UIQM) [33] is employed as a no-reference quality metric, considering that the true images of the “Large fonts” and “checkboard” are unavailable. The metric is attributed by colorfulness measure (UICM), sharpness measure (UISM), and contrast measure (UIConM). In our experiment, the data sets are grayscale images so that the UICM is identically equal to zero.

The comparison results of image quality are presented in Table 2. It shows that the proposed method outperforms our previous method and is close to the Oreifej method in full-reference metrics. For most of the no-reference metrics, the proposed method performs the best numerical value, which indicates the sharpness and contrast of the whole image are significantly improved. The higher value of UISM shows the image contains richer structural edge information and details, and the greater value of UIConM reflects the figure is more suitable for the human visual system. For all metrics, the UIQM, which reflects the underwater image quality from different aspects, is presented based on human visual system and is a special image evaluation for the underwater environment. However, GMSD and FSIM are not as relevant to people’s intuitive feelings as UIQM. Therefore, we prefer UIQM as the main indicator. Although the residual distortions of the proposed method are indicated by GMSD and FSIM, compared with the removal of all distortions, the sharp letter distinction and details, indicated by higher contrast and sharpness, contribute more to the subsequent observation.

Furthermore, from Table 2, we found that the proposed method performs better in UISM, UIConM, and UIQM, which means that it is suitable for the applications that emphasize the human visual effects of restored images and emphasizes structural boundary information, such as underwater text image.

Table 2. Comparison of image quality metrics among the Oreifej method [20], the proposed method, and the previous method [21] ¹.

Data sets	Methods	GMSD (L)	FSIM (H)	UISM (H)	UIConM (H)	UIQM (H)
Middle fonts	Oreifej [20]	0.1530	0.8239	6.7602	0.6896	4.4619
	Our previous method [21]	0.2100	0.7291	6.6876	0.6774	4.3968
	Proposed method	0.1938	0.7613	6.8826	0.7002	4.5358
Small fonts	Oreifej [20]	0.1349	0.8068	6.6793	0.8195	4.9022
	Our previous method [21]	0.2033	0.7073	6.5596	0.7959	4.7825
	Proposed method	0.1850	0.7389	6.7231	0.8225	4.9259
Tiny fonts	Oreifej [20]	0.1524	0.7537	6.8084	0.5654	4.0321
	Our previous method [21]	0.1798	0.7168	6.7185	0.5508	3.9533
	Proposed method	0.1679	0.7439	7.1807	0.6256	4.3571
Large fonts	Oreifej [20]	X	X	6.4618	0.4136	3.3870
	Our previous method [21]	X	X	6.6355	0.4003	3.3906
	Proposed method	X	X	6.9208	0.4395	3.6151
checkboard	Oreifej [20]	X	X	6.8788	0.6291	4.2804
	Our previous method [21]	X	X	7.0863	0.6161	4.2953
	Proposed method	X	X	7.0246	0.6308	4.3295

¹ Because the true image of “Large fonts” is unavailable, the PSNR, GMSD, FSIM are presented with “X”. The “L” means that the low value is good, whereas the “H” means that the high value is good.

4.3. Number of Iterations and Running Time

For the purpose of raising efficiency, the \Downarrow , which presents the difference between the frames and the mean of the current sequence V , is applied to determine the end of the registration process. The value can be obtained from Equation (14). The threshold is set to 0.025, the same as Oreifej [20]. As long as the \Downarrow is less than the threshold, it means that the registration tends to be stable, then the post-processing is carried out. The comparison of convergence results is listed in Table 3.

$$\Downarrow(V) = \frac{\sum_k \sum_{x \in \Omega} |I_k(x) - M(x)|}{h \times w \times n} \quad (13)$$

Table 3. The number of iterations when termination conditions are met.

	Checkboard	Large Fonts	Middle Fonts	Small Fonts	Tiny Fonts
Oreifej method [20]	2	>5	>5	2	4
Our previous method [21]	4	5	5	2	3
Proposed method	2	4	4	2	2

From Table 3, we can see that the iteration numbers of the proposed method are obviously less than other methods, and neither of the data sets has more iterations than the preset maximum iteration times.

We think the reason that, in the Oreifej method, the numbers of iterations keep a high level is because frame blurring is used. In our previous study, using the higher quality reference frame could speed up the registration. However, due to the continuous improvement of the overall definition of the reference frame, the nature of registration being guided to the sharper region is weakened, resulting in a decrease of later iteration velocity. However, a processed reference frame, which has preserved edges and smoothed non-edges simultaneously, is employed to the registration of our proposed method. The edges of reference are always sharper than other parts as the iterations go on,

thus further shortening the registration time. It is also shown by the decreasing trend of iteration time, as shown in Figure 3.

Therefore, the proposed method has better performance on computing speed and greater development potential when the numbers of iterations are the same or unknown.

4.4. Analysis of Patch Fusion

Figure 5 shows the registration results of using the mean frame or the fused frame as the reference frame after the guided filter. It can be seen that there are some distortions and double image (marked by the red rectangle) that affect the recognition and even the overlapping of letters in the mean-to-frames registration. However, the results of our proposed method can express more accurate structural information. The results show that the patch fusion step must be carried out to eliminate the seriously distorted parts before the guided filtering, so as to enhance the correct boundaries of the objects as much as possible, which could decrease the errors that occurred in the registration step.

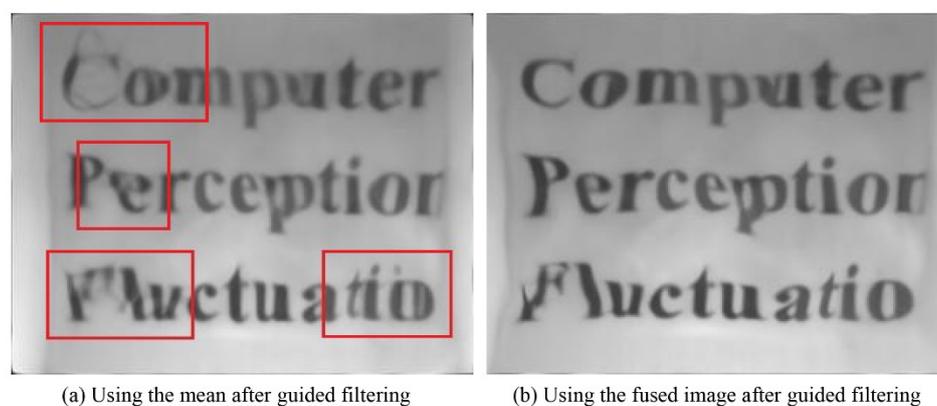


Figure 5. The results after five iterations of our method. (a) Use the mean as the reference after guided filtering. (b) Use the fused image as the reference after guided filtering.

4.5. Analysis of Guided Filter

Compared with other popular edge-preserving algorithms such as the bilateral filter, the guided filter has better behavior near the edges, a fast and non-approximate linear-time characteristic. Furthermore, the guided filter step, which enhances the structure information's boundaries, is the key factor in determining the quality of restoration, especially for "Small fonts" and "Tiny fonts". After the fusion of the sequences, the letters are closely arranged, and their edges are mildly shifted and blurred, resulting in the small variance of the window centered on the boundary points. If the selected regularization parameter ε is too large, the guided filter will mistake the boundaries between letters as the "flat area" to be smoothed, then the loss of boundary information makes it impossible to distinguish letters, whereas if the parameter is too small, then the guided filter will mistakenly regard the ghosting as the gradient edge to preserve, which may cause registration error.

We tested different combinations of regularization parameter ε and window size r , the reconstruction results are compared, as shown in Figure 6. If the ε is not selected properly, the boundaries will be deleted and blurred, as marked by the red rectangle, while the selection of r has little effect on the results. Therefore, we could find that ε plays a greater role in edge restoration than r . In the case, ε and r are set to 0.022 and 4, respectively, to achieve better reconstruction result.

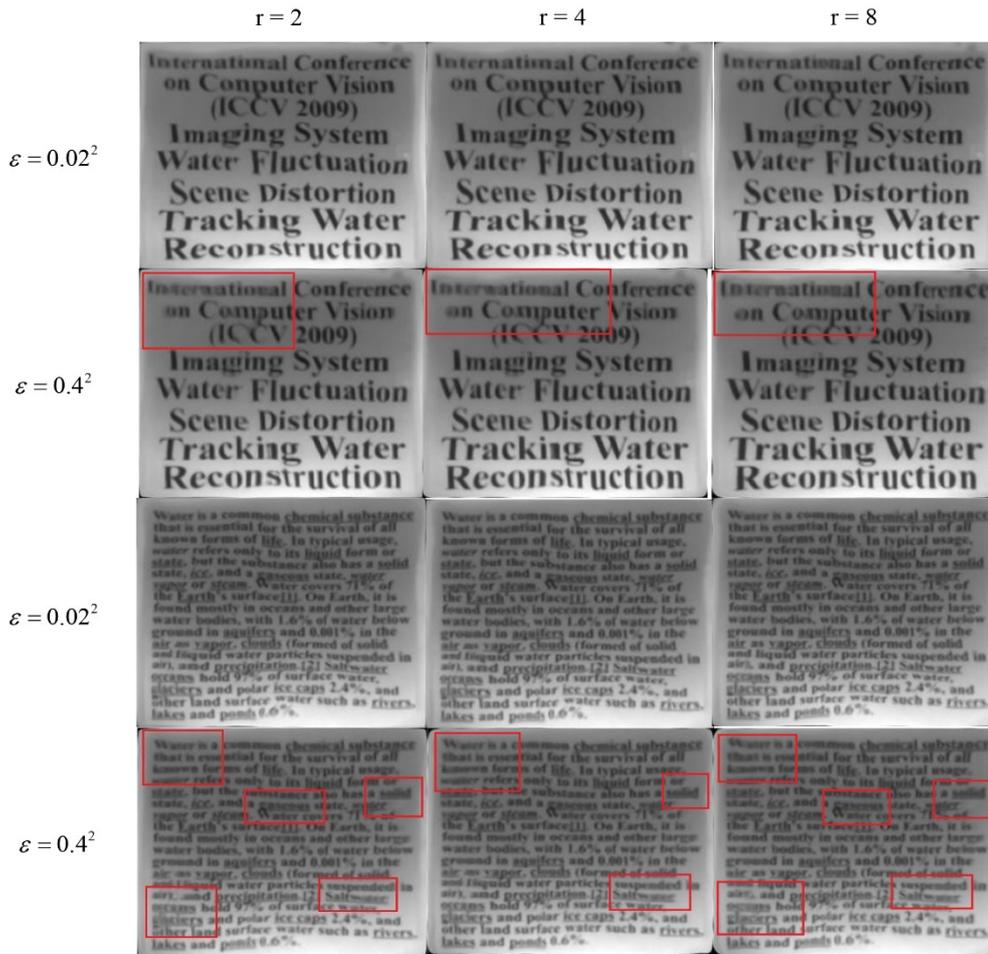


Figure 6. The results of the proposed method with different combinations of ϵ and r . The blurred regions are marked with red rectangles.

4.6. Other Underwater Text Data Set

To further verify the robustness of the proposed method, we dealt with a real text data set in underwater scenes [24]. The data set contains 101 frames, and the size of each frame is 512×512 . The result is also compared with other registration algorithms [20,21]. As shown in Figure 7, the proposed method still outperforms other registration strategies, and the boundary between letters is successfully restored without blurring.

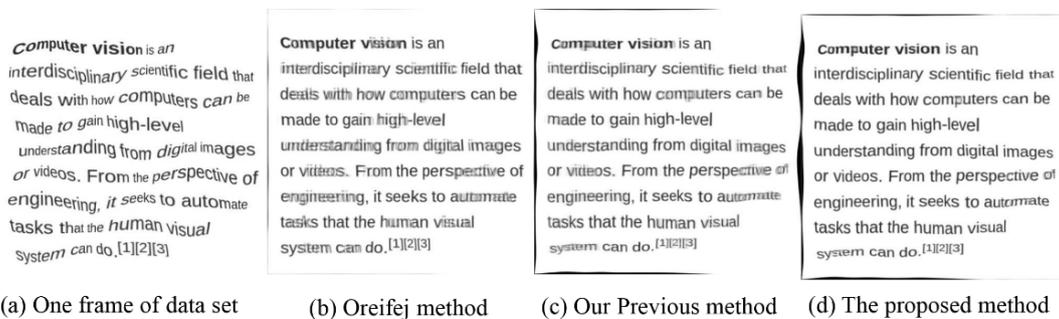


Figure 7. The result comparison between the proposed method and other registration strategies. (a) One frame of the original data set. (b) The result recovered by the Oreifej method. (c) The result recovered by our previous method. (d) The result recovered by the proposed method.

4.7. Compared with Nonregistration Method

The above discussions have proven that the proposed method has faster convergence speed and better behavior than other registration strategies. In order to verify the performance more objectively, the proposed method was also compared with state-of-the-art technology without registration. Considering the reproducibility, the James method [24], whose resource code is available online, it was selected as a reference method. The same data set “Large fonts” with 61 frames were applied to the proposed method and the James method simultaneously. Figure 8 describes one frame of data set and the results of two methods. We found that the James method failed to deal with details under strong disturbance, while the proposed method was able to reconstruct better boundaries and more details. The comparison illustrates that the periodicity hypothesis may fail in severely disturbed environments. Instead, we intend to treat all distortions as local deformation that can be corrected through registration, and achieve better recovery in a strongly disturbed situation.

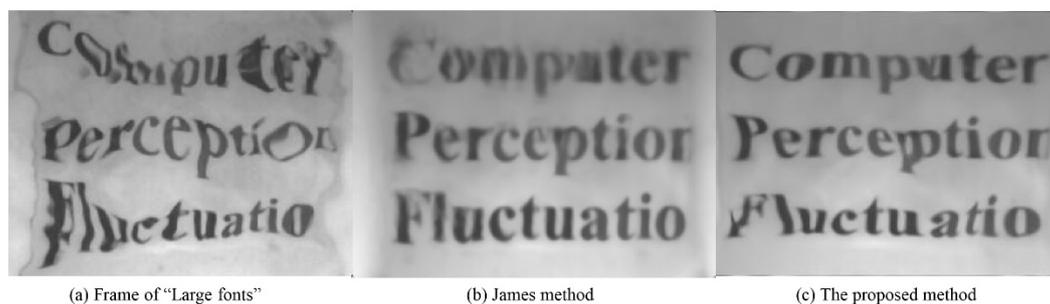


Figure 8. The comparison between the proposed method and the James method [24]. (a) One frame of the original data set. (b) The result recovered by the James method. (c) The result recovered by the proposed method.

5. Conclusions

This article proposes a new structural information restoration method for underwater image sequences using an image registration algorithm. Different from previous studies [20,21], the proposed method follows a new registration strategy, that is, it gives higher priority to edge information before registration, and intentionally guides the registration to focus on the boundary area of interest. At first, the regions with less misalignment across the input warped sequence, which are picked by using SSIM as a metric, are fused into a single frame. Then the guided filter is employed to recognize and hold back the gradient border of the fused image while blurring other areas. During the iterative registration, using the output of the guided filter as a reference, most unexpected fringe anamorphoses can be corrected, and the structural edge information is restored to a greater extent. As iterations go on, the quality of the boundary improves progressively and tends to be stable with the increasing of the iteration. Finally, to dispel the random noise and produce an undistorted frame, the PCA algorithm and patches fusion technology are applied to the registered frames.

From the experiment, we have tested and compared our method with other registration methods [20,21]. In comparison with other registration strategies [20,21], these methods fail to deal with the contour relationship between adjacent letters, bringing about fuzzy blocks that restrict character recognition. Instead, our method can effectively restore the more prominent letter boundary. Meanwhile, as the iteration of registration increases, the running time of our method is constantly shortened, which makes our method more advantageous in the real scene with unknown iterations. In comparison with a nonregistered method [24], our method can also cope better with the highly disturbed underwater scenes. In the future, the application of edge priority in underwater observing will be further studied.

Author Contributions: Software, T.S.; validation, T.S.; conceptualization, Z.Z.; methodology, Z.Z.; project administration, Z.Z.; formal analysis, Y.T.; investigation, Y.T.; writing—original draft preparation, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China under Grant No. 2018YFB1309200.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Gilles, J.; Osher, S. Wavelet burst accumulation for turbulence mitigation. *J. Electron. Imaging* **2016**, *25*, 33003. [CrossRef]
2. Haleder, K.K.; Tahtali, M.; Anavatti, S.G. Simple algorithm for correction of geometrically warped underwater images. *Electron. Lett.* **2014**, *50*, 1687–1689. [CrossRef]
3. Galdran, A.; Pardo, D.; Picón, A.; Alvarez-Gila, A. Automatic Red-Channel underwater image restoration. *J. Vis. Commun. Image Represent.* **2015**, *26*, 132–145. [CrossRef]
4. Boffety, M.; Galland, F.; Allais, A.G. Influence of polarization filtering on image registration precision in underwater conditions. *Opt. Lett.* **2012**, *37*, 3273–3275. [CrossRef] [PubMed]
5. Holohan, M.L.; Dainty, J.C. Low-order adaptive optics: A possible use in underwater imaging? *Opt. Laser Technol.* **1997**, *29*, 51–55. [CrossRef]
6. Murase, H. Surface shape reconstruction of a nonrigid transparent object using refraction and motion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 1045–1052. [CrossRef]
7. Seemakurthy, K.; Rajagopalan, A.N. Deskewing of Underwater Images. *IEEE Trans. Image Process.* **2015**, *24*, 1046–1059. [CrossRef]
8. 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.
9. Cox, C.; Munk, W. Slopes of the sea surface deduced from photographs of sun glitter. *Bull. Scripps Inst. Oceanogr.* **1956**, *6*, 401–479.
10. Shefer, R.; Malhi, M.; Shenhar, A. Waves Distortion Correction Using Cross Correlation: Technical Report, 2011. Israel Institute of Technology. Available online: <http://visl.technion.ac.il/projects/> (accessed on 4 March 2020).
11. Efros, A.; Isler, V.; Shi, J.; Visontai, M. Seeing through water. In Proceedings of the Conference and Workshop on Neural Information Processing Systems, Vancouver, BC, Canada, 13–18 December 2004.
12. Donate, A.; Dahme, G.; Ribeiro, E. Classification of textures distorted by water waves. In Proceedings of the International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006.
13. Donate, A.; Ribeiro, E. Improved reconstruction of images distorted by water waves. In Proceedings of the International Conferences VISAPP and GRAPP, Setúbal, Portugal, 25–28 February 2006.
14. Wen, Z.; Lambert, A.; Fraser, D.; Li, H. Bispectral analysis and recovery of images distorted by a moving water surface. *Appl. Opt.* **2010**, *49*, 6376–6384. [CrossRef]
15. Kanaev, A.V.; Ackerman, J.; Fleet, E.; Scribner, D. Imaging Through the Air-Water Interface. In Proceedings of the Computational Optical Sensing and Imaging, San Jose, CA, USA, 13–15 October 2009.
16. Kanaev, A.V.; Hou, W.; Restaino, S.R.; Matt, S.; Gladysz, S. Restoration of images degraded by underwater turbulence using structure tensor oriented image quality (STOIQ) metric. *Opt. Express* **2015**, *23*, 17077–17090. [CrossRef]
17. Kanaev, A.V.; Hou, W.; Restaino, S.R.; Matt, S.; Gladysz, S. Correction methods for underwater turbulence degraded imaging. In Proceedings of the Remote Sensing of Clouds and the Atmosphere XIX and Optics in Atmospheric Propagation and Adaptive Systems XVII, Amsterdam, The Netherlands, 24–25 September 2014.
18. 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]
19. Wu, J.; Su, X. Method of Image Quality Improvement for Atmospheric Turbulence Degradation Sequence Based on Graph Laplacian Filter and Nonrigid Registration. *Math. Probl. Eng.* **2018**, *2018*, 1–15. [CrossRef]

20. Oreifej, O.; Guang, S.; Pace, T.; Shah, M. A two-stage reconstruction approach for seeing through water. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011.
21. Zhang, Z.; Yang, X. Reconstruction of distorted underwater images using robust registration. *Opt. Express* **2019**, *27*, 9996–10008. [[CrossRef](#)] [[PubMed](#)]
22. Halder, K.K.; Paul, M.; Thatali, M.; Anavatti, S.G.; Murshed, M. Correction of geometrically distorted underwater images using shift map analysis. *J. Opt. Soc. Am. A* **2017**, *34*, 666–673. [[CrossRef](#)]
23. 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, Lake Tahoe, NV, USA, 12–15 March 2018.
24. 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 International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019.
25. James, J.G.; Rajwade, A. Fourier Based Pre-Processing for Seeing Through Water. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Snowmass, CO, USA, 1–5 March 2020.
26. Zhou, W.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.
27. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1397–1409. [[CrossRef](#)]
28. Rueckert, D.; Sonoda, L.I.; Hayes, C.; Hill, D.L.G.; Leach, M.O.; Hawkes, D.J. Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Trans. Med. Imaging* **1999**, *18*, 712–721. [[CrossRef](#)]
29. Xiao, S.; Liu, S.; Wang, H.; Lin, Y.; Song, M.; Zhang, H. Nonlinear dynamics of coupling rub-impact of double translational joints with subsidence considering the flexibility of piston rod. *Nonlinear Dyn.* **2020**, *100*, 1203–1229. [[CrossRef](#)]
30. Berdinsky, A.; Kim, T.W.; Cho, D.; Bracco, C.; Kiatpanichgij, S. Based of T-meshes and the refinement of hierarchical B-splines. *Comput. Methods Appl. Mech. Eng.* **2015**, *283*, 841–855. [[CrossRef](#)]
31. Xue, W.; Zhang, L.; Mou, X.Q.; Bovik, A.C. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Trans. Image Process.* **2013**, *23*, 684–695. [[CrossRef](#)]
32. Zhang, L.; Zhang, L.; Mou, X.Q.; Zhang, D. FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)] [[PubMed](#)]
33. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).