Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

Correction of geometrically distorted underwater images using shift map analysis

Kalyan Kumar Halder,[1,*] Manoranjan Paul,[2] Murat Tahtali,[1] Sreenatha G. Anavatti,[1] and Manzur Murshed[3]

[1]School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT 2600, Australia
[2]School of Computing and Mathematics, Charles Sturt University, Bathurst, NSW 2795, Australia
[3]School of Information Technology, Federation University, VIC 3842, Australia
*Corresponding author: k.halder@unsw.edu.au

Received 3 October 2016; revised 11 December 2016; accepted 12 December 2016; posted 13 December 2016 (Doc. ID 278065); published 31 March 2017

In underwater imaging, water waves cause severe geometric distortions and blurring of the acquired short-exposure images. Corrections for these distortions have been tackled reasonably well by previous efforts but still need improvement in the estimation of pixel shift maps to increase restoration accuracy. This paper presents a new algorithm that efficiently estimates the shift maps from geometrically distorted video sequences and uses those maps to restore the sequences. A nonrigid image registration method is employed to estimate the shift maps of the distorted frames against a reference frame. The sharpest frame of the sequence, determined using a sharpness metric, is chosen as the reference frame. A *k*-means clustering technique is employed to discard too-blurry frames that could result in inaccuracy in the shift maps' estimation. The estimated pixel shift maps are processed to generate the accurate shift map that is used to dewarp the input frames into their nondistorted forms. The proposed method is applied on several synthetic and real-world video sequences, and the obtained results exhibit significant improvements over the state-of-the-art methods.

OCIS codes: (010.7295) Visibility and imaging; (100.3020) Image reconstruction-restoration; (110.3000) Image quality assessment.

https://doi.org/10.1364/JOSAA.34.000666

1. INTRODUCTION

When imaging through an air–water surface, i.e., the camera is in the air above the water surface and the object of interest is immersed, the captured images suffer from random geometric distortions and nonuniform blurring [1–6]. These effects come from attenuation due to scattering and absorption throughout the imaging path [7–11]. Scattering is due to dissolved organic matter or small suspended biological organisms, which causes a change in the direction of the light beam. In addition, the absorption process causes a reduction in the beam energy. Even if the water is clear, i.e., free of the small observable suspended particles, it is still inhomogeneous due to random temperature and salinity fluctuations [12,13]. Furthermore, light rays bend when they travel from air to water or vice versa; the amount of bending is determined by the refractive indices of these media. This significantly contributes to the geometric distortions of the acquired images [2]. As a consequence, underwater imaging poses significant challenges at extended ranges when compared with similar problems in the air.

Researchers from different fields, including computer vision, computer graphics, and optical engineering, have attempted to address the general problems of recovering distorted images due to water waves. In [14], an algorithm to solve the image recovery problem was proposed by modeling a 3D structure of the water waves. The 3D shape of the water surface is reconstructed using the optical flow fields of the distorted frames. The algorithm assumes that the distance between the object and the water surface is known, and the surface waviness is mild enough so that the short-exposure images pose similar features. An adaptive optics (AO) technology was proposed in [13] for a possible solution of the underwater imaging problem. Although AO systems, which are comparatively popular in astronomical imaging, provide real-time corrections of images, the systems are complex, and the cost of full correction in real-time is high. Taking these factors into account, a simplified model was proposed, and a low-cost low-order AO system was developed in [13], which could be controlled by a standard personal computer. The complete system is developed based on the assumption that the degradations of the underwater images are predominantly low-frequency low-order aberrations and image motion.

A simple algorithm was proposed in [15], where it is assumed that each point on an object viewed through a wavy water surface oscillates around its true position, i.e., when the water surface is in a steady-state condition. Also, the average of the displacements

of each point adds up to zero when time tends to infinity. Based on this assumption, the algorithm restores the image of a submerged object by taking the time average of a large number of individual frames captured over a certain period of time. This approach can provide a reasonable result when the waviness of the water is low, but with higher energy waves it fails to restore all features of the image, providing a blurry output. Based on the similar assumption, a graph-embedding method was proposed in [16] to reconstruct through-the-water images from a sequence of submerged objects. A shortest-path algorithm, assuming the distribution of surface normals of the wavy water to be Gaussian, is employed to select the image having the shortest overall path to all the remaining images. A similar approach was presented in [17,18] where the entire image is considered as a group of fixed small image patches. Unlike [16], the method compensates for the leakage problem, which is caused by false shortest distances among similar but motion-blurred patches, by separately analyzing the motion blur and the refraction effects. The authors in [19] first emphasized the estimation of the shape of the distorted water surface and then reconstructing the scenes. Their algorithm assumes that the water surface follows the linear wave dynamics. The radial surface slope, which is used to estimate the water surface, is deduced along extinction boundaries considering the transmission region and internal reflection region separately, and it is recursively improved by minimizing the quadratic measurement error. The slopes of the estimated surface are then employed to reconstruct the target image by an inverse ray tracing.

A lucky region technique was proposed in [20] for reconstruction of underwater images. The algorithm divides the whole image frame into an ensemble of smaller image patches similar to [16–18], but in this case instead of the whole image only the image patches of good quality are selected. The good image patches are then further processed to produce a high-quality image. The algorithm was tested for both through-the-water imaging and water-surface-reflection imaging. In [6], an underwater image reconstruction approach was presented based on motion compensation, lucky patch selection, and denoising. The motion compensation is performed by an image registration technique, whereas the lucky patches are selected by calculating the nonlinear gain coefficient for each pixel of the prototype current frame pair by means of an image quality metric. The final reconstructed frame is obtained by fusing the lucky patches. Tian and Narasimhan [3,4] introduced an underwater image reconstruction approach using model-based tracking. In the approach, a simulator of fluctuations of the water surface was first built using the forward Euler method, and then the original images were recovered using the estimated water surface. In [5], Oreifej *et al.* proposed a two-stage image reconstruction approach, where in the first stage an iterative nonrigid image registration is employed to register all the input frames, after being pre-processed, against their temporal mean, which, in turn, is used to produce a better-structured sequence. The second stage involves extracting the unstructured sparse errors from the sequence through rank minimization. In [21], Oreifej *et al.* applied the image reconstruction approach for background estimation in the cases where the videos contain both turbulence effects and moving objects. Although this method works well for turbulence mitigation, it is computationally expensive.

In this paper, a new image restoration approach is proposed to recover a nondistorted image from a distorted video sequence of it based on a motion compensation technique. This approach introduces a new way of choosing the reference frame for motion estimation by employing a blind image quality (BIQ) metric and discarding the severely degraded frames from further processing by means of a k-means algorithm to avoid inaccuracies in the output. An efficient image registration technique is employed to accurately estimate the motion vector maps, in other words, the pixel shift maps of the distorted frames against the reference frame. The centroid of the estimated shift maps is then calculated and further processed to generate correct shift maps to obtain a high-quality frame. The results presented are of both theoretical and practical interest and offer a new efficient tool for substantial improvement of infrastructure of machine vision-based systems in general and of intelligent surveillance systems in particular.

The rest of this paper is organized as follows. Section 2 presents the details of the proposed image restoration method. Section 3 provides the experimental results on several synthetic and real-world data sets and a comparison of them against those from the state-of-the-art methods [4,21]. Finally, Section 4 contains comments and conclusions on the results.

2. PROPOSED IMAGE RESTORATION METHOD

The proposed geometric distortion correction method is illustrated as a block diagram in Fig. 1. In the first step of the method, an image sharpness assessment metric is used to calculate the sharpness values of the input frames, which are used to select the reference frame as well as a subset of fairly sharp frames from the input frames using k-means clustering. Next, a pixel registration technique is employed to calculate the pixel shift maps of the selected frames against the reference frame. After that, the shift maps are used to calculate the correct shift maps for the input frames through 2D interpolations. Finally, those shift maps are utilized to restore the distorted frames back to their true ones. The complete approach is discussed in detail below.

A. Assumptions

The proposed method is based on the Cox–Munk law [22], which states that, in the case when the water surface is sufficiently large and the waves are stationary, the normals of the water surface, and therefore the displacements of the points on the object, are approximately Gaussian distributed centered around their true locations. Considering this law, it can be assumed that the wander of each pixel of an image over a certain period of time oscillates around its true position, i.e., the mean wander of the pixel would be zero. In a different viewpoint, if the pixel's wander is measured with respect to any fixed position, the mean of its wander will represent its shift from the true position. This shift can be then used to restore the pixel to its true position. The goal of the proposed method is to accurately estimate the center of each pixel's wander; hence it is called a centroid-based method.

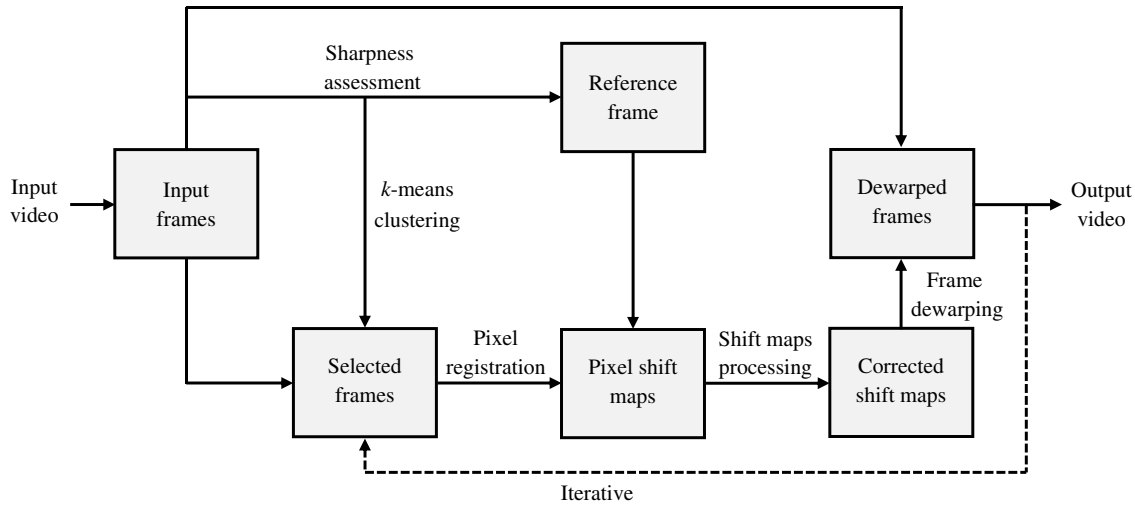

Fig. 1. Simple block diagram of the proposed image restoration method.

B. Sharpness Assessment

Based on the above-mentioned assumptions, instead of a single pixel, if a whole frame is considered, there needs to be a reference frame to calculate the pixel shift maps of all the input frames against it. The simplest way could be calculating the mean of the warped frames and use it as the reference frame. Although a mean frame is somewhat geometrically stable, it contains motion blur; thus, it cannot provide high-quality output [23,24]. In this method, one of the warped frames from the whole video sequence is considered as the reference frame. There are three options to choose a reference frame in this approach: (1) first frame, (2) any random frame, and (3) the sharpest frame. Although the first two options do not involve too much processing time for the reference frame selection, the output video may not be of high quality if the frame suffers from blur or noise [2]. Therefore, in this study, a BIQ metric is employed, which calculates the sharpness of all input frames and uses these values to choose the reference frame.

Assuming P as the 1D Pseudo–Wigner distribution, the Rényi entropy R can be calculated as [25]

$$R(n) = -\frac{1}{2}\log_2\left(\sum_{\kappa=1}^{N}\check{P}_n^3(\kappa)\right), \tag{1}$$

where n and κ are the time and frequency discrete variables, respectively, N is the number of pixels, and $\check{P}(\kappa) = P(\kappa)\cdot P^*(\kappa)$. Here $P^*(\kappa)$ indicates the complex conjugate of $P(\kappa)$.

Using Eq. (1), the expected entropy for an image can be computed as [25]

$$\bar{R}(t,\theta_s) = \sum_n R(n,\theta_s)/M, \tag{2}$$

where $\bar{R}(t,\theta_s)$ is the expected value of entropy for image $t \in [1,2,\ldots,M]$, measured in directions $\theta_s \in [\theta_1,\theta_2,\ldots,\theta_S]$. Here M and S represent the image size and the number of directions, respectively. The metric employed to select the sharpest frame is [25]

$$Q = \sqrt{\sum_{s=1}^{S}(\mu_t - \bar{R}(t,\theta_s))^2/S}, \tag{3}$$

where μ_t is the mean of the values $\bar{R}(t,\theta_s)$. The frame that poses a higher value of Q is deemed to be the sharpest frame.

C. K-Means Clustering

Because the distortion condition changes nonuniformly and rapidly, some frames in the video sequence may have poor sharpness, and use of those frames to calculate pixel shift maps may add additional errors in the restoration process. K-means clustering is an efficient tool to select a set of subframes due to its high convergence rate and invariant property to data ordering [26–28]. Therefore, a k-means algorithm is applied on the sharpness values of the input frames, calculated using Eq. (3), to create two clusters of frames: poor sharpness and fair sharpness. The frames belonging to the fair sharpness cluster are passed to the next pixel registration process, whereas the frames belonging to the poor sharpness cluster are discarded.

D. Pixel Registration

A correlation coefficient-based backward mapping image registration technique [29] is employed to estimate the pixel shift maps (s_x and s_y for x and y directions, respectively) of the selected input frames with respect to the reference frame. For the purpose of restoration of geometric deformations, it is advantageous to have the shift maps in the form of backward mapping rather than forward mapping. The latter will tell where a pixel went in the next frame and can lead to gaps if two or more pixels from the source frame collapse into one pixel in the target frame, whereas the former will tell where a pixel came from the previous frame. Backward mapping is easily implementable using a 2D interpolation without any noticeable artifacts [30].

E. Shift Maps Processing

After estimating the shift maps for all selected frames through the backward mapping, their centroid is calculated as

$$p_x(x,y) = \frac{\sum_{k=1}^{K} s_x(x,y,k)}{K},$$
$$p_y(x,y) = \frac{\sum_{k=1}^{K} s_y(x,y,k)}{K}, \quad (4)$$

where p_x and p_y are the centroid shift maps, k is the frame index, and K is the total number of selected frames.

Due to the use of backward mapping, inversion of the centroid shift maps is required before they are warped with the individual shift maps in order to generate corrected shift maps, which transforms each individual frame into their dewarped selves. However, simply calculating the matrix inverse or changing the sign of the centroid shift maps may not provide a desired inversion. A possible solution of this inversion is described in [31], where the inverse of a shift map is approximated by warping of that shift map with its negative matrix. Therefore, the inverses of p_x and p_y are approximated as

$$p_x^{-1}(x,y) \approx -p_x(x - p_x(x,y), y - p_y(x,y)),$$
$$p_y^{-1}(x,y) \approx -p_y(x - p_x(x,y), y - p_y(x,y)), \quad (5)$$

where p_x^{-1} and p_y^{-1} represent the inverses of p_x and p_y, respectively. According to the initial assumption, this shift map represents the actual shift of the reference frame from its ground truth; therefore, it can be used to warp the frame back into its unwarped position through a 2D interpolation technique. For all other frames, the actual shift maps are calculated using the inverse centroid shift map and individual shift maps as

$$s_x^*(x,y,k) = p_x^{-1}(x,y) + s_x(x + p_x^{-1}(x,y), y + p_y^{-1}(x,y), k),$$
$$s_y^*(x,y,k) = p_y^{-1}(x,y) + s_y(x + p_x^{-1}(x,y), y + p_y^{-1}(x,y), k), \quad (6)$$

where s_x^* and s_y^* are the corrected shift maps for all the selected frames.

F. Frame Dewarping

Using the corrected shift maps in Eq. (6), the input frames can be dewarped as

$$I^*(x,y,k) = I(x + s_x^*(x,y,k), y + s_y^*(x,y,k), k), \quad (7)$$

where I^* is the dewarped version of the warped frame I.

The dewarped sequence, after the first iteration, is more stabilized than the original one. However, a few more iterations may provide higher accuracy and better stabilization of the output video. In that case, the dewarped version of the reference frame after an iteration is used as the reference frame in the next iteration.

3. SIMULATION EXPERIMENTS

The proposed method was implemented in MATLAB and tested on an Intel Core i7-2600 CPU 3.40 GHz machine with 8 GB RAM. Experimental analyses were carried out to verify its performance against two state-of-the-art methods, namely, the Tian method [4] and the Oreifej method [21]. As discussed earlier, the Tian method is quite different from the Oreifej method and the proposed method, which recovered the distorted frames by estimating the water surface fluctuations based on a distortion model. However, the proposed method and the Oreifej method are somewhat similar, where an image registration is employed to calculate the pixel shift maps that are later used to dewarp the frames. The main difference between these two methods is in the way of calculating the pixel shift maps. The Oreifej method uses the mean of the input frames as the reference frame, blurs the input frames with the same blur presented in the reference frame, and registers them against the reference frame to get the pixel shift maps. On the other hand, the proposed method uses the sharpest warped frame as the reference frame, registers the input frames directly against it, and uses the mean of the individual pixel shift maps to calculate the accurate ones. In short, both methods rely on averaging: the Oreifej method does the averaging of the frames, whereas the proposed method does the averaging of the pixel shift maps avoiding the blurring operation and thus provides a more accurate estimation of the pixel shift maps.

All three methods were first applied on four synthetically distorted underwater video sequences, namely, the brick-wall, small-fonts, medium-fonts, and large-fonts sequences. Each sequence consists of 60 frames, and the underwater distortion of different strengths was added to them using a spatial distortion model, which is developed based on Snell's law of refraction and the wave equation [3]. Later on, the methods were applied on a real degraded video sequence, namely, the water–Colgate sequence [20]. This sequence contains 60 frames that were captured in the laboratory by a video camera fixed above the water surface. A still object was laid on the planner ground under the clean water. Waves were generated by wind produced by a fan, while additional waves were created by dropping an object into the water. Any other attenuation of a turbid medium may have not been present during the video recording, the distortions were only due to the wavy water surface. This process generated distortions with a maximum amplitude of 15 pixels, with an average amplitude of 3.2 pixels and a standard deviation of 2.8 pixels. The test sequences contain both texts and textures in order to provide a good means of clearly demonstrating underwater distortion effects. Along with visual comparison, various image quality metrics are used to quantitatively evaluate the performance of the methods.

A. Quality Metrics

Several relevant quality metrics are used to analyze the performance of restoration by the proposed method. These metrics require the existence of a distortion-free copy of an image, called the ground-truth image, which can be used for comparison with the image whose quality is to be measured.

1. Mean Square Error

The mean square error (MSE) is a frequently used measure of how close the restored frame is to the ground-truth frame. It can be defined as

$$\mathrm{MSE} = \frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} (R(x,y) - G(x,y))^2, \quad (8)$$

where R and G are any restored frame and the ground-truth frame, respectively. $m \times n$ is the size of each frame.

2. Peak Point Spread Function

The point spread function (PSF) describes, as the name implies, the response of an imaging system to a point source or point

object. It is basically the image of a bright spot. The sharper and narrower the PSF of an imaging system the better it is at resolving finer detail. In order to use PSF as a quality measure for the restored frames, it is calculated as

$$\mathrm{PSF}(x,y) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(R(x,y))}{\mathcal{F}(G(x,y))}\right), \tag{9}$$

where $\mathcal{F}$ denotes the Fourier transform and $\mathcal{F}^{-1}$ its inverse. The peak point spread function (PPSF), therefore, can be calculated as

$$\mathrm{PPSF} = \mathrm{MAX}(\mathrm{PSF}(x,y)), \tag{10}$$

where $\mathrm{MAX}(\cdot)$ represents the maximum possible value.

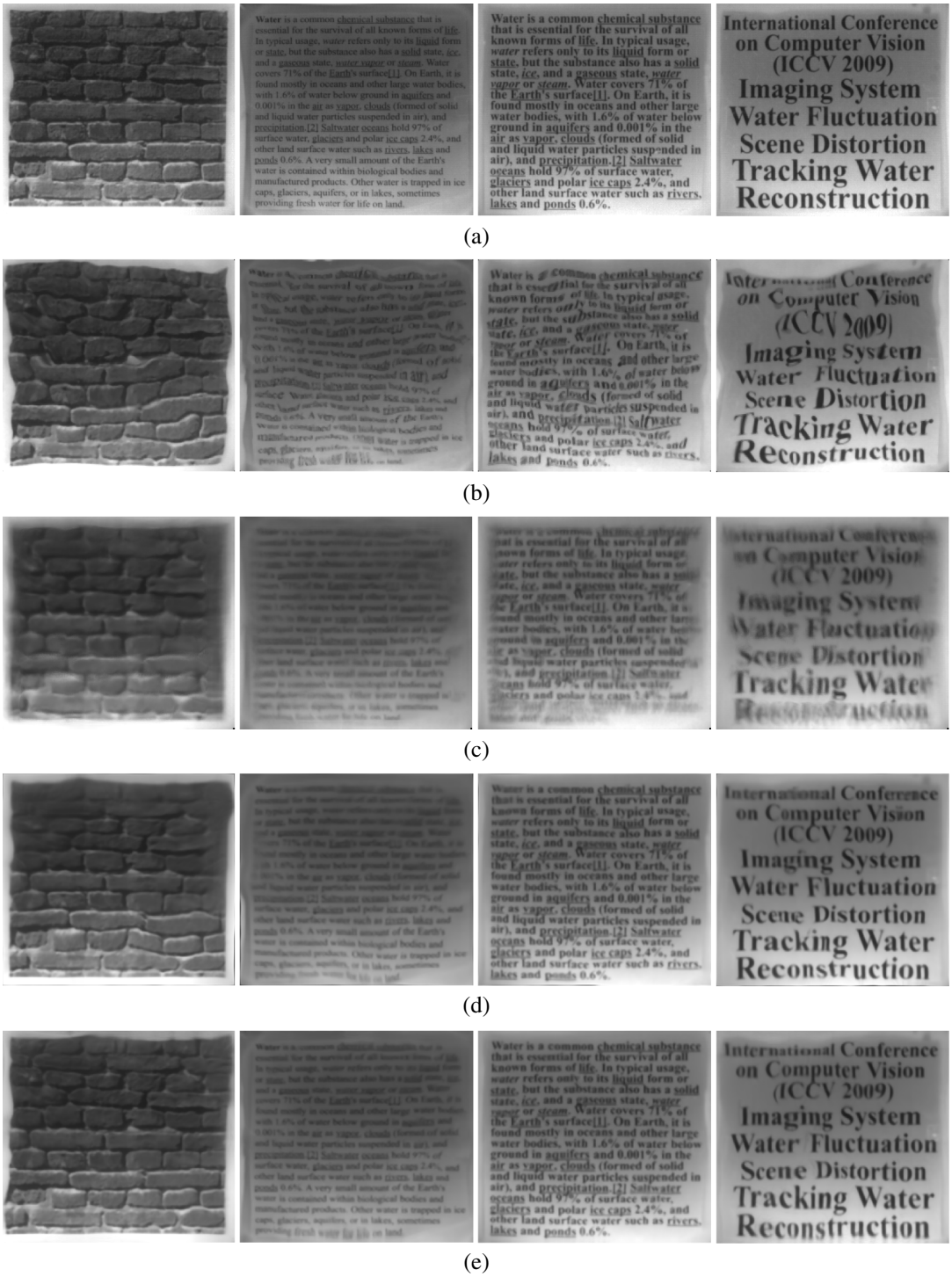

Fig. 2. Image restoration results. Simulation tests were carried out on the synthetically generated brick-wall, small-fonts, medium-fonts, and large-fonts sequences [3]. (a) Ground truth frames. (b) Sample distorted frames. (c) Restored frames using Tian method. (d) Restored frames using Oreifej method. (e) Restored frames using proposed method.

3. Peak Signal-to-Noise Ratio

The term peak signal-to-noise ratio (PSNR) is an expression for the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. The mathematical expression for the PSNR is as follows:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{(\mathrm{MAX}(G))^2}{\mathrm{MSE}}\right). \tag{11}$$

4. Structural Similarity

The structural similarity (SSIM) is an image quality metric that estimates the visual impact of changes in image luminance, contrast, and structure. It can be expressed as [32]

$$\mathrm{SSIM} = \frac{(2\mu_R\mu_G + C_1)(2\sigma_{RG} + C_2)}{(\mu_R^2 + \mu_G^2 + C_1)(\sigma_R^2 + \sigma_G^2 + C_2)}, \tag{12}$$

where μ_R, μ_G, σ_R, σ_G, and σ_{RG} are the local means, standard deviations, and cross-covariance for the restored frame R and the ground-truth frame G. C_1 and C_2 are constants, which can be defined as

$$C_1 = (0.01 \times L)^2, \qquad C_2 = (0.03 \times L)^2,$$

where L is the specified dynamic range value.

B. Performance Analysis and Comparison

The computational times for the Tian method and the Oreifej method, which required us to process a single frame of size 268×292, are obtained as 48s and 37s, respectively. For the proposed method, the processing time for the same frame is 23s, which is 52% shorter than the Tian method and 38% shorter than the Oreifej method. The simulation results of the Tian, Oreifej, and proposed methods on the four synthetic sequences are shown in Fig. 2. The ground-truth frames, which were used to synthetically generate the test sequences, are shown in Fig. 2(a). Samples of the distorted frames from each sequence are shown in Fig. 2(b). The means of the restored frames using the Tian, Oreifej, and proposed methods are shown in Figs. 2(c)–2(e), respectively. A mean frame is used here to show the image restoration performance, as it is a good measure of how well the frames are close to each other. The sharper the mean frame, the more the frames are geometrically aligned. It is evident that all three methods compensate for the geometric distortions to some extent, especially when compared with the warped frames separately. However, the Tian method is much less effective in doing so, which is notable from the uncompensated regions and intensity distortions of the restored frames in Fig. 2(c). The Oreifej method performs comparatively well in both geometric and intensity corrections, as can be seen in Fig. 2(d). For the proposed method, the restored frames are more geometrically aligned and, therefore, provide a mean frame closer to the ground truth. It can be argued that there are still some distortions in the restored frames in Fig. 2(e), but these came from the pixel registration and interpolation technique that introduce some processing errors.

The quantitative analysis is conducted for all the synthetic sequences by calculating the MSE, PSNR, PPSF, and SSIM of the mean of restored frames with respect to their ground truths. A comparison of the calculated metrics is shown in Table 1. It can be noted from the numerical comparison that the proposed method outperforms the other two methods in all the quality metrics for each of the test sequences.

Furthermore, Fig. 3 shows the PSF of the sample frame in comparison with the mean of restored frames by the three methods for the randomly selected brick-wall sequence. As expected, the PSF of the warped frame in Fig. 3(a) is indistinct without any notable peak at its center. In the case of the Tian method, although its PSF could provide a reasonable restoration of the warped frames, it is not distinguishable, as it only sports a weak raise at its center, as shown in Fig. 3(b). The Oreifej method's PSF is slightly more pronounced, as shown in Fig. 3(c), while also exhibiting increased PSF activity away from the center. On the other hand, the proposed method's PSF shown in Fig. 3(d) has a distinct peak at the center; although not as sharp as to indicate perfect correlation, the peak is significantly higher with less activity away from the center.

The performances of the Tian, Oreifej, and proposed methods are also compared by their geometric correction capabilities under noisy conditions. Figure 4 shows the SSIM index versus SNR plots for the brick-wall sequence for the three methods. As expected, the restoration performance of all the methods degrades with the increase of noise level. In the case of the

Table 1. Comparison of Quality Metrics among the Tian, Oreifej, and Proposed Methods for the Frames in Fig. 2

		Video Sequences			
Quality Metrics	**Methods**	**Brick Wall**	**Small Fonts**	**Medium Fonts**	**Large Fonts**
MSE (Low values—Good)	Tian method	0.0108	0.0197	0.0150	0.0153
	Oreifej method	0.0051	0.0029	0.0038	0.0078
	Proposed method	0.0031	0.0015	0.0022	0.0072
PSNR (High values—Good)	Tian method	19.68	17.05	18.24	18.14
	Oreifej method	22.92	25.38	24.20	21.08
	Proposed method	25.09	28.24	26.58	21.43
PPSF (High values—Good)	Tian method	0.0403	0.0429	0.0651	0.0372
	Oreifej method	0.0526	0.0435	0.1082	0.0788
	Proposed method	0.1017	0.0856	0.1778	0.0856
SSIM (High values—Good)	Tian method	0.6017	0.5465	0.5778	0.6199
	Oreifej method	0.6331	0.6190	0.7107	0.6317
	Proposed method	0.7726	0.7326	0.8100	0.7412

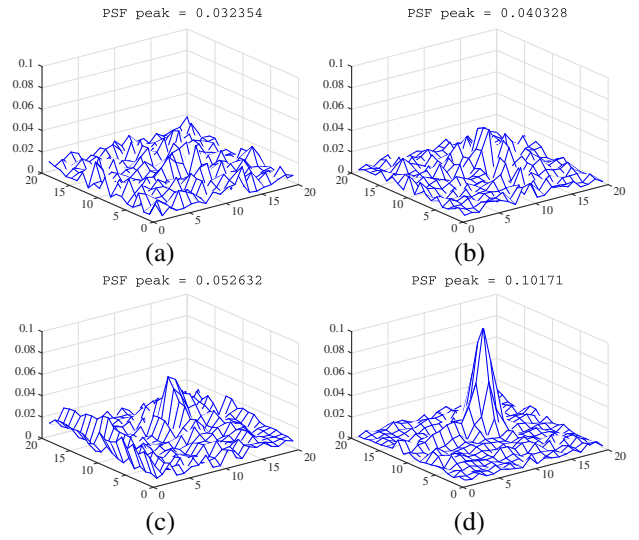

Fig. 3. PSF plot for the randomly selected brick-wall sequence. (a) Sample frame. (b) Using Tian method. (c) Using Oreifej method. (d) Using proposed method.

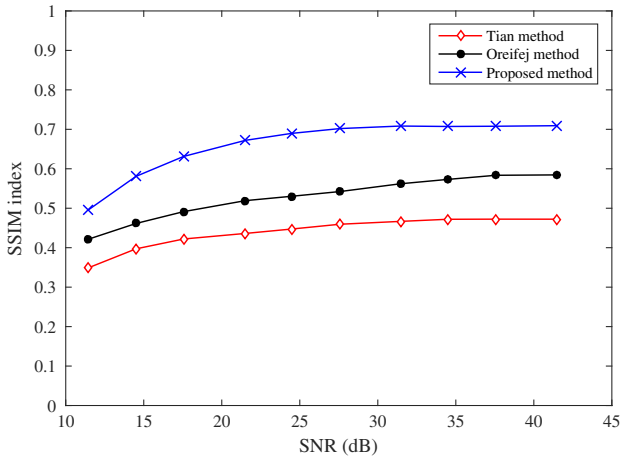

Fig. 4. SSIM index of the restored mean frame versus SNR for the brick-wall sequence.

proposed method, the SSIM index is consistently higher than the other two methods for all the values of SNR.

After verifying the performance of the three methods on the synthetic sequences, they were applied on the real-world water–Colgate sequence [20]. A sample warped frame is shown in Fig. 5(a) for visual comparison. The means of the restored frames using the Tian, Oreifej, and proposed methods are shown in Figs. 5(b)–5(d), respectively. The Tian method improves the geometric alignments at the cost of considerable blurring in the restored frame. The Oreifej method provides better visual quality of the frame than the Tian method, although the restored frame still has some distortions. However, the proposed method outperforms both state-of-the art methods by providing more details of the frame, which is visually evident from the better legibility of the word "stationery" compared with the other frames in Figs. 5(a)–5(c).

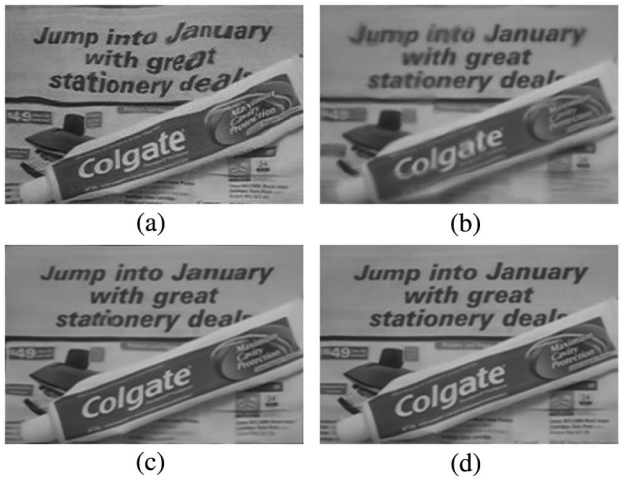

Fig. 5. Restoration results on real-world water–Colgate sequence [20]. (a) Sample distorted frame. (b) Restored frame using Tian method. (c) Restored frame using Oreifej method. (d) Restored frame using proposed method. Colgate images used with permission.

Table 2. Comparison of Quality Metric among the Tian, Oreifej, and Proposed Methods for the Real-World Water–Colgate Sequence

		Video Sequence
Quality Metric	**Methods**	**Water–Colgate**
BIQ (High values—Good)	Tian method	0.0056
	Oreifej method	0.0063
	Proposed method	0.0115

Because the ground-truth frame is not available in this case, the quality metrics presented in Table 1 are not applicable for the quantitative assessment. Still, an object comparison is made for these results using the BIQ metric, defined in Eq. (3), and the calculated values are presented in Table 2. For the proposed method, the value of the metric is higher than that of the other two methods, confirming its better geometric correction capability.

4. DISCUSSION AND CONCLUSION

An efficient algorithm to recover the nondistorted image of an underwater scene from a distorted video sequence is successfully implemented in this paper. The proposed approach is based on estimating the centroid of the random shifting of the pixels of short-exposure frames using a motion estimation technique to obtain a single high-quality image. The performance of the proposed method is highly dependent on the accuracy of the pixel shift map estimation, which is challenging due to large random fluctuations of refractive index. A highly accurate pixel registration technique is employed for this purpose. Also, a sharpness assessment strategy provides the sharpest frame to be used as the reference frame, while a *k*-means clustering algorithm allows only reasonably good quality frames for shift maps computation. The proposed scheme therefore demonstrates a convincing way of calculating the nearly accurate

pixel shift maps of the distorted frames that is further processed and used to warp the reference frame back into its truth position.

The proposed method outperforms the state-of-the-art methods on both synthetic and real-world video sequences, especially, in terms of visual quality and accuracy. Although an iterative scheme is considered in the proposed method, its performance with a single iteration also provides substantial improvement. The computational time of the proposed method can be easily controlled by various means, such as by introducing some stopping criteria in the pixel registration algorithm and using the fastest interpolation technique. The implementation of the proposed method on graphical processing units (GPUs) or field programmable gate arrays (FPGAs) could allow its real-time application to optical communications, mine detection, and surveillance. In its current form, only geometrical correction of the frames is of concern, while the consideration of other imaging problems, such as localized blur and CCD noise, is a future aim. The proposed method also can be extended to imaging through the water surface reflection and imaging from underwater to observe objects above the water surface.

REFERENCES

1. J. Gilles and S. Osher, "Wavelet burst accumulation for turbulence mitigation," J. Electron. Imaging **25**, 033003 (2016).
2. K. K. Halder, M. Tahtali, and S. G. Anavatti, "Simple algorithm for correction of geometrically warped underwater images," Electron. Lett. **50**, 1687–1689 (2014).
3. Y. Tian and S. G. Narasimhan, "Seeing through water: image restoration using model-based tracking," in *Proceedings of IEEE International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, 2009), pp. 2303–2310.
4. Y. Tian and S. G. Narasimhan, "Globally optimal estimation of nonrigid image distortion," Int. J. Comput. Vis. **98**, 279–302 (2012).
5. O. Oreifej, S. Guang, T. Pace, and M. Shah, "A two-stage reconstruction approach for seeing through water," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, 2011), pp. 1153–1160.
6. A. V. Kanaev, W. Hou, S. Woods, and L. N. Smith, "Restoration of turbulence degraded underwater images," Opt. Eng. **51**, 057007 (2012).
7. A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," J. Visual Commun. Image Represent **26**, 132–145 (2015).
8. K. Seemakurthy and A. N. Rajagopalan, "Deskewing of underwater images," IEEE Trans. Image Process. **24**, 1046–1059 (2015).
9. W. Hou, "A simple underwater imaging model," Opt. Lett. **34**, 2688–2690 (2009).
10. M. Boffety, F. Galland, and A.-G. Allais, "Influence of polarization filtering on image registration precision in underwater conditions," Opt. Lett. **37**, 3273–3275 (2012).
11. G. Wang, B. Zheng, and F. F. Sun, "Estimation-based approach for underwater image restoration," Opt. Lett. **36**, 2384–2386 (2011).
12. R. Schettini and S. Corchs, "Underwater image processing: state of the art of restoration and image enhancement methods," EURASIP J. Adv. Signal Process. **2010**, 1–15 (2010).
13. M. L. Holohan and J. C. Dainty, "Low-order adaptive optics: a possible use in underwater imaging?" Opt. Laser Technol. **29**, 51–55 (1997).
14. H. Murase, "Surface shape reconstruction of a nonrigid transport object using refraction and motion," IEEE Trans. Pattern Anal. Mach. Intell. **14**, 1045–1052 (1992).
15. R. Shefer, M. Malhi, and A. Shenhar, Waves distortion correction using crosscorrelation, 2001, http://visl.technion.ac.il/projects/2000maor/.
16. A. Efros, V. Isler, J. Shi, and M. Visontai, "Seeing through water," in *Proceedings of Neural Information Processing Systems* (Neural Information Processing Systems Foundation, 2004).
17. A. Donate, G. Dahme, and E. Ribeiro, "Classification of textures distorted by water waves," in *Proceedings of International Conference on Pattern Recognition* (Institute of Electrical and Electronics Engineers, 2006), pp. 421–424.
18. A. Donate and E. Ribeiro, "Improved reconstruction of images distorted by water waves," in *Advances in Computer Graphics and Computer Vision*, J. Braz, A. Ranchordas, H. Araújo, and J. Jorge, eds., Communications in Computer and Information Science (Springer, 2007), Vol. **4**, pp. 264–277.
19. D. M. Milder, P. W. Carter, N. L. Flacco, B. E. Hubbard, N. M. Jones, K. R. Panici, B. D. Platt, R. E. Potter, K. W. Tong, and D. J. Twisselmann, "Reconstruction of through-surface underwater imagery," Waves Random Complex Media **16**, 521–530 (2006).
20. Z. Wen, A. Lambert, D. Fraser, and H. Li, "Bispectral analysis and recovery of images distorted by a moving water surface," Appl. Opt. **49**, 6376–6384 (2010).
21. O. Oreifej, X. Li, and M. Shah, "Simultaneous video stabilization and moving object detection in turbulence," IEEE Trans. Pattern Anal. Mach. Intell. **35**, 450–462 (2013).
22. C. Cox and W. Munk, "Slopes of the sea surface deduced from photographs of sun glitter," Bull. Scripps Inst. Oceanogr. **6**, 401–479 (1956).
23. M. Tahtali, D. Fraser, and A. J. Lambert, "Restoration of non-uniformly warped images using a typical frame as prototype," in *Proceedings of IEEE Region 10 Conference* (Institute of Electrical and Electronics Engineers, 2005), pp. 1–6.
24. K. K. Halder, M. Tahtali, and S. G. Anavatti, "Geometric correction of atmospheric turbulence-degraded video containing moving objects," Opt. Express **23**, 5091–5101 (2015).
25. S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," J. Opt. Soc. Am. A **24**, B42–B51 (2007).
26. A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," ACM Comput. Surv. **31**, 264–323 (1999).
27. S. Yu, L.-C. Tranchevent, X. Liu, W. Glänzel, J. A. K. Suykens, B. De Moor, and Y. Moreau, "Optimized data fusion for Kernel *k*-means clustering," IEEE Trans. Pattern Anal. Mach. Intell. **34**, 158–173 (2012).
28. M. H. Furhad, M. Tahtali, and A. Lambert, "Restoring atmospheric-turbulence-degraded images," Appl. Opt. **55**, 5082–5090 (2016).
29. A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," IEEE Trans. Med. Imaging **29**, 1882–1891 (2010).
30. M. Tahtali, A. J. Lambert, and D. Fraser, "Graphics processing unit restoration of non-uniformly warped images using a typical frame as prototype," Proc. SPIE **7800**, 78000H (2010).
31. M. Tahtali, "Imaging techniques through the atmosphere," Ph.D. dissertation (The University of New South Wales, 2007).
32. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).