

Deskewing of Underwater Images

Karthik Seemakurthy and Ambasamudram Narayanan Rajagopalan, *Senior Member, IEEE*

Abstract—We address the problem of restoring a static planar scene degraded by skewing effect when imaged through a dynamic water surface. In particular, we investigate geometric distortions due to unidirectional cyclic waves and circular ripples, phenomena that are most prevalent in fluid flow. Although the camera and scene are stationary, light rays emanating from a scene undergo refraction at the fluid-air interface. This refraction effect is time varying for dynamic fluids and results in nonrigid distortions (skew) in the captured image. These distortions can be associated with motion blur depending on the exposure time of the camera. In the first part of this paper, we establish the condition under which the blur induced due to unidirectional cyclic waves can be treated as space invariant. We proceed to derive a mathematical model for blur formation and propose a restoration scheme using a single degraded observation. In the second part, we reveal how the blur induced by circular ripples (though space variant) can be modeled as uniform in the polar domain and develop a method for deskewing. The proposed methods are tested on synthetic as well as real examples.

Index Terms—Underwater imaging, skew, motion blur, unidirectional waves, circular ripples, space-invariant, space-variant, polar domain.

I. INTRODUCTION

UNDERWATER (UW) imaging has been receiving considerable attention in the last few decades. A driving force behind advances in UW technologies has been the need to improve image quality. Underwater images can suffer from different types of degradations including color loss, noise due to floating particles, low contrast, skewing and blurring [1]. In this paper, the primary aim is to restore scenes captured through a dynamic refractive medium. Even though visibility is limited to few tens of meters inside water, imaging through dynamic water surface is important for coral reef monitoring, examining the contamination of shallow waters, mapping the distribution of vegetation and seabed sediments etc. [2]. These capabilities find applications in commercial fishing zones as well as in boat safety. Thus, it is essential to develop image restoration schemes that can address the challenges that arise when imaging through flowing water.

Manuscript received February 17, 2014; revised October 13, 2014 and January 9, 2015; accepted January 13, 2015. Date of publication January 22, 2015; date of current version February 11, 2015. This work was supported by BRNS, Department of Atomic Energy, India, under Project 2012/21/14/BRNS/1681. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Aleksandra Pizurica.

The authors are with the Department of Electrical Engineering, IIT Madras, Chennai 600036, India (e-mail: karthikvjit@gmail.com; raju@ee.iitm.ac.in).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author. The file names starting with UCW_, CIRCULAR_ and OUTDOOR_ correspond to the videos captured for unidirectional cyclic waves, circular ripples and outdoor experiments, respectively. The total size of the videos is 7.1 MB. Contact karthikvjit@gmail.com for further questions about this work.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2015.2395814

We primarily focus on restoring a UW scene affected by skew and motion blur when imaged through a dynamic refractive medium. The phenomenon of skew is due to the wavy nature of water, as different light rays experience different amounts of refraction at the water surface. This space-varying refraction that occurs on the water surface results in non-rigid deformation of the scene when captured on the image plane. Several methods exist that primarily aim to restore skewed scenes. Shefer et al. [3] propose a method based on averaging that assumes that the sum of displacements that a particular scene point undergoes is zero as time tends to infinity. However, the method fails to recover high frequency contents. Murase et al. [4] propose a method based on optical flow. They initially recover the surface slope information and then use it to restore the distorted scene. Efros et al. [5] propose a ‘lucky’ approach in which they assume that when the water is locally flat, the scene points that a camera sees through that region would be undistorted. Hence, they take a video and use graph-embedding to collect the undistorted information across the frames and stitch them together to obtain the final restored image. Wen et al. [6] have also developed a similar method that uses a long video shot and collects the undistorted part of the scene from each frame by performing bispectral analysis [7]. Tian et al. [8] propose a model-based tracking approach for simultaneous recovery of image and water surface. Their method derives a model for image distortions using principal component analysis (PCA) and fits the model to each and every frame of the video. In a later work, they propose a data-driven approach [9] in which they assume that the original undistorted frame is available as a template and attempt to restore the distorted video by estimating the deformation parameters for each frame. A recent method by Turalev et al. [2] uses multiple illumination sources (red to illuminate the water surface and blue for the underwater object). The red and blue components of the captured color image are used to extract information about the slopes of the water surface in order to undo the geometrical distortion effects.

A basic requirement of the above-mentioned methods for restoring skewed scenes is that the exposure time should be short to avoid any incidental motion blur. While the causes of motion blur in underwater images can be many [10]–[16], we restrict ourselves in this work to motion blur resulting from a dynamic water surface. Incidentally, skewing of scenes can also be associated with motion blur when the imaging medium is dynamic in nature and camera exposure is long. In our model of image formation, we assume that the light ray suffers negligible attenuation and scattering. The only cause of distortion is attributed to dynamic refraction occurring at the surface of the water, which changes the direction

of the incoming light. The net effect is that a single pixel sees multiple scene points during the exposure time of the camera leading to motion blur. Recovering an underwater scene affected by skew as well as motion blur is an important preprocessing step for many computer vision algorithms.

Most of the existing methods restore the video by considering that the distortion is due only to skewness of the scene. The motion blur due to dynamic medium is treated to be almost negligible. However, it is well-known that the dynamic nature of the water surface does induce motion blur in the captured video. Very few works exist that deal with both skew and blur. Donate et al. [17] restore a distorted video by modeling motion blur and geometric distortion due to water waves as two different entities. They initially divide all the frames of a video into small overlapping sub-regions and then classify them as low and highly distorted patches by K-means clustering. By using the fact that the Fourier spectrum of a motion blurred frame consists of low energy in the high frequencies, they eliminate the patches with energy less than the mean energy of all frames. Finally, they use blending to mosaic all the patches in order to reduce tiling artifacts. Oreifej et al. [18] employ a two-stage approach. In the first stage, they attempt to register each and every frame of the video to the blurred mean by using a non-rigid registration technique developed by Rueckert et al. [19]. In the second stage, they reduce the effect of sparse noise by using convex rank-optimization [20].

We address the twin problems of deskewing and deblurring within a unified framework. While Oreifej et al. [18] claim that deblurring of the temporal mean does not restore the degraded video, we show that it is indeed possible to deblur the temporal mean of a skewed and blurred video, albeit under modest constraints. In contrast to existing methods, our method requires neither a template nor a non-rigid registration framework. Although the geometrical distortions that occur in the image plane will appear to be local when imaged through a refractive medium, we establish that it is possible to express these local distortions by global transformations for two very relevant and commonly observed scenarios of cyclic water flow, namely, unidirectional cyclic waves (UCW) and circular ripples. The interpretation of local distortions as global transformations on the image plane greatly reduces the computational complexity involved in restoring images degraded by cyclic waves.

In our approach, we elegantly relate the effects of skew to motion blur through a careful investigation and underpinning of the physics of the image formation process. In the first part of the paper, we describe the image formation model for UCW and a scaled orthographic camera, and derive an expression for the geometric warp at a pixel in the image. Subsequently, we show that a fixed set of warps can be obtained at every pixel in the image plane. This is followed by a methodology for restoring images degraded by UCW waves. In the second part of the paper, we discuss restoration of images degraded by circular ripples. Unlike the UCW, the set of transformations induced by these waves on the image plane is space-variant. We start with an analytical model for circular ripples and show that transforming the blurred observation into the polar domain effectively renders the problem space-invariant. The choice

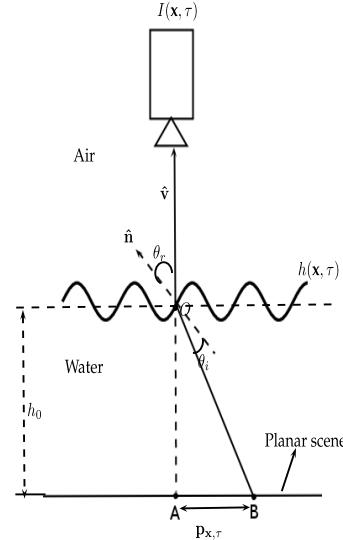


Fig. 1. Ray diagram for a scaled orthographic camera to demonstrate image formation.

of exposure time plays an important role in our approach. We show that a long exposure is necessary for our scheme and this is realized by a video of equivalent duration so as to avoid saturation effects. Both the proposed techniques are evaluated and also compared with state-of-the-art methods.

The organization of the paper is as follows. In section II, we describe the image formation model for a scaled orthographic camera where we analytically derive the relationship between the warping that a pixel undergoes and the gradient of the water surface that it encounters. In section III, we describe a mathematical model for blur formation in UCW. In section IV, we propose a maximum a posterior (MAP) formulation to restore scenes degraded by UCW followed by experiments in section V. In section VI, we investigate the case of circular ripples and propose an elegant method in the polar domain for restoration. The corresponding results are given in section VII. Finally, we conclude in section VIII with suggestions for future extensions.

II. IMAGE FORMATION IN FLOWING WATER

In this section, we describe image formation in the presence of a dynamic refractive medium. In Fig. 1, let $I_g(\mathbf{x})$ be the original image corresponding to the bottom planar scene of a waterpool which is imaged by a fronto-parallel camera. Since our method is primarily aimed at shallow water applications, we assume small water waves (low spatial frequency and small amplitude fluctuations). In practice, the wave will get attenuated as it progresses. Since we assume a limited field-of-view (FOV), the attenuation of water waves can be treated as negligible. Each video frame $I(\mathbf{x}, \tau)$ is a distorted version of the original scene $I_g(\mathbf{x})$ and the two images can be related [8] by the equation

$$I(\mathbf{x}, \tau) = I_g(\mathbf{x} + \mathbf{w}(\mathbf{x}, \tau)) \quad (1)$$

where $\mathbf{w}(\mathbf{x}, \tau)$ is the warping function acting on the pixel at location \mathbf{x} at time τ . Let n be the refractive index of water, and let θ_i and θ_r be the angle of incidence and refraction,

respectively, of the light ray reaching the camera from scene point B. Let us assume that water fluctuations with respect to the level h_0 are small. The amount of displacement $\mathbf{p}_{x,\tau}$ that a light ray experiences can be calculated as

$$\begin{aligned}\|\mathbf{p}_{x,\tau}\|_2 &= h_0 \tan(\theta_r - \theta_i) \\ &= h_0 \left[\frac{(\tan \theta_r - \tan \theta_i)}{1 + \tan \theta_r \tan \theta_i} \right]\end{aligned}\quad (2)$$

For small water waves, θ_r and θ_i will be small [8], which implies that $\cos \theta_r \approx 1$, $\cos \theta_i \approx 1$ and $1 + \tan \theta_r \tan \theta_i \approx 1$. Applying these approximations, Eq. 2 can be simplified as

$$\begin{aligned}\|\mathbf{p}_{x,\tau}\|_2 &\approx h_0 (\sin \theta_r - \sin \theta_i) \\ &= h_0 \sin \theta_r \left(1 - \frac{\sin \theta_i}{\sin \theta_r} \right)\end{aligned}\quad (3)$$

From Snell's law, $n \sin \theta_i = \sin \theta_r$ since the refractive index of air is 1. Therefore, Eq. 3 becomes

$$\begin{aligned}\|\mathbf{p}_{x,\tau}\|_2 &= h_0 \sin \theta_r \left(1 - \frac{1}{n} \right) \\ &= \alpha \sqrt{1 - \cos^2 \theta_r}\end{aligned}\quad (4)$$

where $\alpha = h_0(1 - \frac{1}{n})$. From Fig. 1, $\cos \theta_r = (\hat{\mathbf{v}} \cdot \hat{\mathbf{n}})$ where $\hat{\mathbf{v}} = (0 \ 0 \ 1)^T$ and $\hat{\mathbf{n}}$ can be calculated as

$$\hat{\mathbf{n}} = \frac{\mathbf{t}_x \times \mathbf{t}_y}{\|\mathbf{t}_x \times \mathbf{t}_y\|_2} \quad (5)$$

where ' \cdot ' and ' \times ' represent dot and cross product, respectively, while $\mathbf{t}_x = (1 \ 0 \ h_x)^T$ and $\mathbf{t}_y = (0 \ 1 \ h_y)^T$ are the tangential vectors defined on the water surface at point O , and h_x and h_y denote the partial derivatives of $h(\mathbf{x}, \tau)$ with respect to spatial coordinates. If $\nabla \mathbf{h}$ denotes the gradient of the water surface then $\nabla \mathbf{h} = [h_x \ h_y]^T$. Using these values of \mathbf{t}_x and \mathbf{t}_y in Eq. 5, we get

$$\hat{\mathbf{n}} = \frac{(-h_x, -h_y, 1)^T}{\sqrt{1 + h_x^2 + h_y^2}} \quad (6)$$

Substituting Eq. 6 in Eq. 4, we obtain

$$\begin{aligned}\|\mathbf{p}_{x,\tau}\|_2 &= \alpha \frac{\sqrt{h_x^2 + h_y^2}}{\sqrt{1 + h_x^2 + h_y^2}} \\ &= \alpha \frac{\|\nabla \mathbf{h}\|_2}{\sqrt{1 + \|\nabla \mathbf{h}\|_2^2}}\end{aligned}\quad (7)$$

For small water waves, $\|\nabla \mathbf{h}\|_2 \ll 1$ and Eq. 7 simplifies to

$$\|\mathbf{p}_{x,\tau}\|_2 \approx \alpha \|\nabla \mathbf{h}\|_2 \quad (8)$$

Thus, the displacement that a light ray undergoes corresponding to a pixel \mathbf{x} inside water is proportional to the gradient at the corresponding point on the water surface that the pixel sees. This can be mathematically expressed as

$$\mathbf{p}_{x,\tau} \approx \alpha \nabla \mathbf{h}(\mathbf{x}, \tau) \quad (9)$$

From Fig. 1, when the water surface is still, the camera will image scene point A. However, in flowing water, the light ray which reaches the camera above the water surface will

make a finite angle with the normal $\hat{\mathbf{n}}$ defined on the surface of the water at point O . From Snell's law, there will be refraction at the water surface due to which instead of A, the camera will actually see scene point B. A dynamic water surface is spatially-varying and depends on factors such as wind, gravitational force etc. The amount of translation that a pixel \mathbf{x} experiences on the image plane is given by

$$\mathbf{w}(\mathbf{x}, \tau) = \mathbf{P} \mathbf{p}_{x,\tau} \quad (10)$$

where \mathbf{P} is the projection matrix of the camera. From Eq. 10, it can be seen that the amount of translation that a pixel undergoes will vary spatially i.e., different pixels can have different amount of translations, which in turn leads to geometrical distortions (skew).

III. MOTION BLUR AND SKEW IN UCW

Motion blur is a widely occurring phenomenon and is typically attributed to relative motion between camera and scene. Motion blur due to incidental camera shake can be avoided by using tripods or hardware stabilizers. In the underwater situation that we are dealing with, both the camera and scene are stationary. However, because the medium is dynamic, the captured image is degraded by dynamic refraction that occurs on the surface of the medium during exposure, thus leading to blurring. This can be explained using the results derived in the previous section. When the water surface is dynamic, the normal $\hat{\mathbf{n}}$ (see Fig. 1) will be time-varying, because of which different scene points will be seen by a pixel during the exposure time. If the camera exposure is short (compared to the rate of flow of water), with the result that the scene is captured before the water moves on to the next state, then only the skewing effect exists; else, light rays from multiple scene points will reach the same pixel and the averaging effect during the exposure window will lead to motion blurring. The mathematical relationship between motion blur and skew in UCW follows.

Consider the example shown in Fig. 2. Without loss of generality, let the water surface be sinusoidal in nature (although our analysis is valid for any UCW). Assume that the camera is placed above the surface and looking vertically downwards at the scene through a dynamic water surface. At time $T = 0$, let A, B, C, and D be four pixels in the image plane corresponding to points E, F, G, and H, respectively, on the sinusoidal surface as shown in Fig. 2(a). At time $T = t_1$, the pixels A, B, C, and D will see the surface points F, G, H, and E, respectively (Fig. 2(b)). At a later time $T = t_2$, these pixels will see the surface points G, H, E, and F, respectively, and so on. Thus, at any instant of time, the pixels can have different gradients which in turn results in different translations. This leads to the skewing effect of the planar scene being imaged. As the water surface changes its state, the skewed images across different times are averaged during the time of exposure resulting in a motion blurred image. Interestingly, the information about skew is embedded in the blurred observation as discussed next.

Let \mathbf{f} and \mathbf{g} be the images captured under still and dynamic water, respectively. The image \mathbf{g} can then be interpreted as

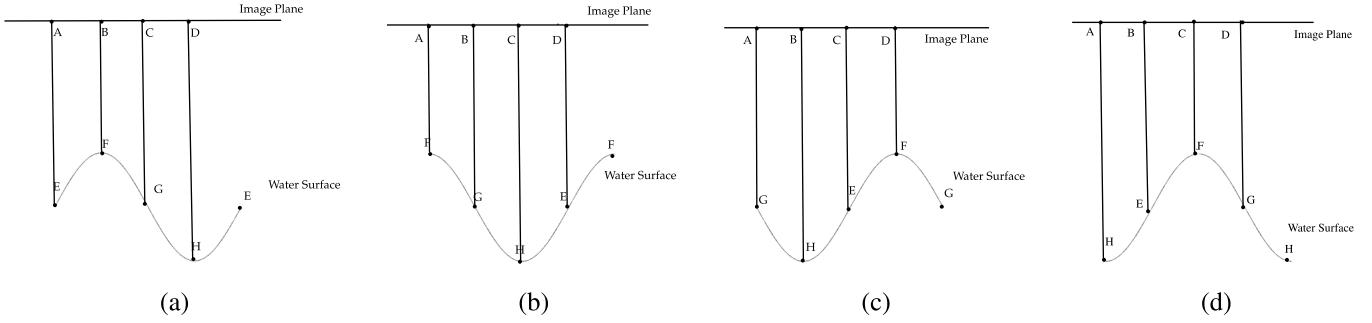


Fig. 2. Variation of gradients with time for a periodic water surface: Time (a) $T = 0$, (b) $T = t_1$, (c) $T = t_2$, and (d) $T = t_3$.

the average of all skewed versions of \mathbf{f} during the exposure time T_e . By using Eq. 1, the relation between \mathbf{f} and \mathbf{g} can be expressed as

$$g(\mathbf{x}) = \frac{1}{T_e} \int_0^{T_e} f(\mathbf{x} + \mathbf{w}(\mathbf{x}, \tau)) d\tau \quad (11)$$

It must be noted from Eq. 10 that, for our scenario, the transformation experienced by each pixel in the image plane depends only on the gradient of the water surface as we assume a homogeneous medium and planar scene. Hence, the set of transformations experienced by every pixel in the image plane will be the same under UCW. Provided the exposure time is long enough, the set of all possible warpings for each pixel will be the same, and consequently, we can conclude that the blur induced is space-invariant. This implies that the blurred image can be equivalently expressed as a weighted average of *global translations* of the undistorted image \mathbf{f} since the order in which the translations occur at each pixel is not relevant to the averaging process. Note that this elegant simplification was possible by leveraging both Eqs. 10 and 11 to succinctly depict the dependency between blur and skew. Strictly speaking, this argument will hold only for integer number of cycles (say c). Since in practice it may not be possible to enforce ' c ' to be an integer, we assume that the number of cycles is large enough to neglect any fractional (vestigial) effects due to $c - \lfloor c \rfloor$ which in turn implies long exposures. This enables us to address the problem of deskewing through a deblurring framework, since all the information about local distortions is embedded in the blurred observation.

A. Mathematical Model for Blur Formation

Let the origin be the camera center and $\mathbf{s} = [X \ Y \ Z]^T$ be a scene point on the planar surface that is imaged at pixel \mathbf{x} , where $\mathbf{x} = [x \ y]^T$ are image coordinates. The relation between \mathbf{x} and \mathbf{s} is given by $x = \frac{qX}{Z}$ and $y = \frac{qY}{Z}$ where q is the focal length of camera. In matrix notation, we can write $\tilde{\mathbf{x}} = \mathbf{K}[\mathbf{I}|\mathbf{0}]\tilde{\mathbf{s}}$ where

$$\mathbf{K} = \begin{bmatrix} q & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

\mathbf{I} is 3×3 identity matrix, $\mathbf{0}$ is 3×1 zero vector, and, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{s}}$ are homogeneous coordinates of \mathbf{x} and \mathbf{s} , respectively. Due to the dynamic nature of water surface during exposure, the scene point \mathbf{s} will be transformed to $\mathbf{s}_\tau = \mathbf{s} + \mathbf{p}_{\mathbf{x}, \tau}$ where

$\mathbf{p}_{\mathbf{x}, \tau}$ is the translation vector at time τ , and is of the form $\mathbf{p}_{\mathbf{x}, \tau} = [p_{x, \tau} \ p_{y, \tau} \ 0]^T$ since there is no movement in the z -direction.

Let \mathbf{x}_τ be the projection of \mathbf{s}_τ in the image plane. Then, we have $\tilde{\mathbf{x}}_\tau = \mathbf{K}[\mathbf{I}|\mathbf{0}]\tilde{\mathbf{s}}_\tau$. Assume that the point \mathbf{s} is on a plane with normal \mathbf{n} and at a distance d from the camera image plane. From the plane equation, we have $\mathbf{n}^T \mathbf{s} = d$. Therefore,

$$\begin{aligned} \mathbf{s}_\tau &= \mathbf{s} + \mathbf{p}_{\mathbf{x}, \tau} \\ &= \mathbf{s} + \mathbf{p}_{\mathbf{x}, \tau} \frac{\mathbf{n}^T \mathbf{s}}{d} \\ &= \left[\mathbf{I} + \frac{\mathbf{p}_{\mathbf{x}, \tau} \mathbf{n}^T}{d} \right] \mathbf{s} \end{aligned}$$

Thus, the relation between the transformed image pixel \mathbf{x}_τ and \mathbf{x} can be written as

$$\tilde{\mathbf{x}}_\tau = \mathbf{K} \left[\mathbf{I} + \frac{\mathbf{p}_{\mathbf{x}, \tau} \mathbf{n}^T}{d} \right] \mathbf{K}^{-1} \tilde{\mathbf{x}} \quad (12)$$

For a fronto-parallel plane, since all the scene points are at the same distance d from the camera, the transformation of image points can be expressed through matrix $\mathbf{H}_{\mathbf{x}, \tau}$ as $\tilde{\mathbf{x}}_\tau = \mathbf{H}_{\mathbf{x}, \tau} \tilde{\mathbf{x}}$, where $\mathbf{H}_{\mathbf{x}, \tau} = \mathbf{K} \left[\mathbf{I} + \frac{\mathbf{p}_{\mathbf{x}, \tau} \mathbf{n}^T}{d} \right] \mathbf{K}^{-1}$. Since we have shown that the blurred image \mathbf{g} (which is originally obtained by averaging the skewed frames) can also be equivalently obtained from the average of globally transformed versions of \mathbf{f} , we can relax the spatial dependency to yield $\mathbf{p}_{\mathbf{x}, \tau} = \mathbf{p}_\tau$. This implies that $\mathbf{H}_{\mathbf{x}, \tau} = \mathbf{H}_\tau$. Thus, Eq. 11 can be alternatively expressed as

$$g(\mathbf{x}) = \frac{1}{T_e} \int_0^{T_e} f(\mathbf{H}_\tau^{-1}(\mathbf{x})) d\tau$$

Let D be the set of all possible transformations that are induced on the image plane during the exposure time. The dimensionality of the space spanned by D is either 1 or 2, since the warpings at each pixel in our scenario are restricted to in-plane translations. Let us define $t_D: D \mapsto R^+$ to be the point spread function (PSF) such that each transformation $\lambda \in D$ is mapped to a non-negative real number which represents the fraction of time that the light ray spent in a particular position that caused the transformation λ . Since $t_D(\lambda)$ represents the fraction of the total time, the sum of all the elements of the range of t_D will be unity i.e., $\sum_{\lambda \in D} t_D(\lambda) = 1$. Note that the nature of blur in our scenario is space-invariant. The blurred image \mathbf{g} can then be expressed in terms of the PSF as

$$g(\mathbf{x}) = \int_{\lambda \in D} t_D(\lambda) f(\mathbf{H}_\lambda^{-1}(\mathbf{x})) d\lambda \quad (13)$$

where \mathbf{H}_λ is the homography matrix which transforms the pixel \mathbf{x} according to transformation λ . The discrete equivalent of Eq. 13 is given by

$$g(\mathbf{x}) = \sum_{\lambda \in D} t_D(\lambda) f(\mathbf{H}_\lambda^{-1}(\mathbf{x})). \quad (14)$$

IV. DEBLURRING AND DESKEWING OF IMAGES DEGRADED BY UCW

We estimate the latent image and the set of translational warps using maximum-a-posteriori (MAP) formulation. Let \mathbf{l} , \mathbf{t}_D and \mathbf{g} be the latent image, PSF to be estimated and the given blurred observation, respectively, all lexicographically ordered. By Bayes theorem

$$p(\mathbf{l}, \mathbf{t}_D | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{l}, \mathbf{t}_D) p(\mathbf{l}) p(\mathbf{t}_D) \quad (15)$$

where $p(\mathbf{l}, \mathbf{t}_D | \mathbf{g})$ represents the a posteriori probability, $p(\mathbf{g} | \mathbf{l}, \mathbf{t}_D)$ is the likelihood, while $p(\mathbf{l})$ and $p(\mathbf{t}_D)$ denote the priors on the latent image and PSF, respectively.

A. Likelihood

In the presence of noise, the blurred image is given by

$$\mathbf{g} = \sum_{\lambda \in D} t_D(\lambda) (\mathbf{H}_\lambda \mathbf{l}) + \boldsymbol{\eta} \quad (16)$$

where $\boldsymbol{\eta}$ is modeled as a set of i.i.d Gaussian random variables (since electronic sensor noise during image capture process is commonly modeled as Gaussian [21], [22]) with zero mean and standard deviation σ_0 :

$$p(\boldsymbol{\eta}) = \prod_i N(\eta_i | 0, \sigma_0) \quad (17)$$

where i indicates index of a pixel. Shan et al., [23] have shown that ringing artifacts can be considerably reduced if we work in the gradient domain along with intensity. In the light of this, Eq. 17 can be reformulated as

$$p(\boldsymbol{\eta}) = \prod_{\partial_* \in \bar{\partial}} \prod_i N(\partial_* \eta_i | 0, \sigma_k) \quad (18)$$

where $\bar{\partial} = \{\partial_0 \ \partial_x \ \partial_y \ \partial_{xx} \ \partial_{yy} \ \partial_{xy}\}$ denotes the set of partial derivative operators. If η_i takes a Gaussian distribution $N(\eta_i | 0, \sigma_0)$ then $\partial_k \eta_i$ also follows a Gaussian distribution as $N(\partial_k \eta_i | 0, \sigma_k)$ where $\sigma_k = \sqrt{2^{k+1}} \sigma_0$ and $k (> 0)$, denotes the order of the partial derivative. Hence, the likelihood can be expressed as

$$p(\mathbf{g} | \mathbf{l}, \mathbf{t}_D) = \prod_{\partial_* \in \bar{\partial}} \prod_i N(\partial_* g_i - \partial_* g'_i \mid \partial_* g'_i, \sigma_k) \quad (19)$$

where $\partial_* \mathbf{g}' = \sum_{\lambda \in D} t_D(\lambda) \partial_*(\mathbf{H}_\lambda \mathbf{l})$.

B. Prior on PSF

The PSF for motion blur was originally described in [24]–[26] for camera shake and a sparsity constraint was enforced. The PSF in our scenario represents the set of in-plane translations that any particular pixel experiences due to refraction. In section III, we have shown that the blur induced due to cyclic waves is space-invariant. Because waves

are directional in nature, the sparsity constraint on the PSF is equally applicable to our scenario. Mallat [27] has shown that the probability distribution of sparse images is heavy-tailed in nature and can be modeled by a Laplacian. We consider the PSF as an image and propose to use the prior

$$p(\mathbf{t}_D) = \prod_{\lambda \in D} \exp(-\beta t_D(\lambda)) \quad (20)$$

where β is the rate of decay and λ indexes over all elements of the PSF. Since the PSF weights are non-negative, we do not consider the modulus in the above prior.

C. Prior on Latent Image

It is well-known that the gradients of natural images are usually sparse. However, employing a sparse prior will render the latent image estimation problem non-convex [28]. Following [28], and to make our optimization step simpler, we use a Gaussian distribution for image gradients. The functional form for the latent image prior then becomes

$$\begin{aligned} p(\mathbf{l}) &\propto \prod_{i,j} \exp\left(-\frac{1}{2}\gamma((l_{i,j} - l_{i+1,j})^2 + (l_{i,j} - l_{i,j+1})^2)\right) \\ &\propto N(0, \Psi) \end{aligned} \quad (21)$$

where i, j are pixel indices, $\Psi^{-1} = \gamma(\mathbf{C}_{g_x}^T \mathbf{C}_{g_x} + \mathbf{C}_{g_y}^T \mathbf{C}_{g_y})$, where \mathbf{C}_{g_x} and \mathbf{C}_{g_y} are the convolution matrices corresponding to gradient vectors $\mathbf{g}_x = [1 \ -1]$, $\mathbf{g}_y = [1 \ -1]^T$, and γ indicates the rate of decay.

D. Energy Minimization

We estimate the latent image and the PSF by using an energy minimization approach which minimizes the negative logarithm of the a posteriori probability defined earlier in Eq. 15 i.e.,

$$\begin{aligned} E(\mathbf{l}, \mathbf{t}_D) &= -\log(p(\mathbf{l}, \mathbf{t}_D | \mathbf{g})) \\ &= -\log(p(\mathbf{g} | \mathbf{l}, \mathbf{t}_D)) - \log(p(\mathbf{l})) - \log(p(\mathbf{t}_D)) \\ &= \sum_{\partial_* \in \bar{\partial}} \sum_{\lambda \in D} t_D(\lambda) \partial_*(\mathbf{H}_\lambda \mathbf{l}) - \partial_* \mathbf{g}' \|_2^2 \\ &\quad + \beta \|\mathbf{t}_D\|_1 + \gamma \|\nabla \mathbf{l}\|_2^2 \end{aligned} \quad (22)$$

where $\bar{\alpha} = \{\alpha_0 \ \alpha_1 \ \alpha_2\}$ are the weights applied to the latent image and its first and second-order derivatives, respectively. The energy function in Eq. 22 is non-convex [29] and we propose an alternating minimization (AM) framework in which we fix one variable and minimize the energy function over the other variable and vice-versa. The user has to specify a rough estimate of the range of translations as an initial estimate for the PSF. We next describe the steps in the AM scheme.

• **Prediction:** The first step is prediction of the latent image estimate \mathbf{l} in which the given blurred observation is subject to bilateral filtering, shock filtering and gradient magnitude thresholding [30]. This step ensures that insignificant details and noise are eliminated. The output of this step is an estimate of the gradient of the latent image which is fed as input to the subsequent stages.

- *PSF Estimation:* In this step, we fix the latent image estimate (obtained from prediction) and minimize the energy function

$$E(\mathbf{t}_D) = \sum_{\substack{\partial_* \in \bar{\partial} \\ \alpha_* \in \bar{\alpha}}} \alpha_* \left\| \sum_{\lambda \in D} t_D(\lambda) \partial_*(\mathbf{H}_\lambda \mathbf{l}) - \partial_* \mathbf{g} \right\|_2^2 + \beta \left\| \mathbf{t}_D \right\|_1 \quad (23)$$

The optimization of Eq. 23 can be done by using nnLeastR of least absolute shrinkage and selection operator (LASSO) [31]. This algorithm combines the good features of both subset selection and ridge regression wherein it shrinks some of the PSF weights while setting others to zero [32]. The function nnLeastR imposes additional non-negativity constraint on the PSF weights. The equation that is solved is given by

$$\min_{\mathbf{x} \geq 0} \frac{1}{2} \left\| \mathbf{Ax} - \mathbf{y} \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{x} \right\|_2^2 + \lambda \left\| \mathbf{x} \right\|_1 \quad (24)$$

In order to express Eq. 23 in the form of Eq. 24, the parameter ρ is set to 0 so that Eq. 23 can be rewritten as

$$E(\mathbf{t}_D) = \left\| \mathbf{At}_D - \mathbf{g} \right\|_2^2 + \beta \left\| \mathbf{t}_D \right\|_1 \quad (25)$$

where $\mathbf{A} = [(\sqrt{\alpha_0} \partial_0 \mathbf{L})^T \ (\sqrt{\alpha_1} \partial_x \mathbf{L})^T \ (\sqrt{\alpha_1} \partial_y \mathbf{L})^T \ (\sqrt{\alpha_2} \partial_{xx} \mathbf{L})^T \ (\sqrt{\alpha_2} \partial_{yy} \mathbf{L})^T \ (\sqrt{\alpha_2} \partial_{xy} \mathbf{L})^T]^T$ and, \mathbf{L} is a block matrix whose columns are lexicographically-ordered and transformed versions of the latent image \mathbf{l} . If $\mathbf{l} \in \mathbb{R}^{m^2 \times 1}$ and $\mathbf{t}_D \in \mathbb{R}^{n \times 1}$ then $\mathbf{L} \in \mathbb{R}^{m^2 \times n}$.

- *Latent Image Estimation:* Following [10], we estimate the latent image using conjugate-gradient approach. In this step, we fix the PSF obtained from the PSF estimation step and minimize the energy function

$$E(\mathbf{l}) = \sum_{\substack{\partial_* \in \bar{\partial} \\ \alpha_* \in \bar{\alpha}}} \alpha_* \left\| \sum_{\lambda \in D} t_D(\lambda) \partial_*(\mathbf{H}_\lambda \mathbf{l}) - \partial_* \mathbf{g} \right\|_2^2 + \gamma \left\| \nabla \mathbf{l} \right\|_2^2 \quad (26)$$

The gradient of Eq. 26 is given by

$$\frac{dE(\mathbf{l})}{d\mathbf{l}} = \mathbf{X}\mathbf{l} - \mathbf{z} \quad (27)$$

where $\mathbf{X} = \left(\sum_{\substack{\mathbf{D}_* \in \bar{\mathbf{D}} \\ \alpha_* \in \bar{\alpha}}} 2\alpha_* \mathbf{D}_*^T \mathbf{H}^T \mathbf{H} \mathbf{D}_* \right) + 2\gamma \mathbf{D}_x^T \mathbf{D}_x + 2\alpha \mathbf{D}_y^T \mathbf{D}_y$, $\mathbf{z} = 2\mathbf{D}_*^T \mathbf{H}^T \mathbf{D}_* \mathbf{g}$, $\mathbf{H} = \sum_{\lambda \in D} t_D(\lambda) \mathbf{H}_\lambda$, and $\bar{\mathbf{D}} = \{\mathbf{D}_0 \ \mathbf{D}_x \ \mathbf{D}_y \ \mathbf{D}_{xx} \ \mathbf{D}_{yy} \ \mathbf{D}_{xy}\}$ is the set of matrices which perform partial derivative operation on the latent image.

- *Proof of Bi-Convexity:* Although the original problem is non-convex, we now prove that by adopting the AM framework, we solve a bi-convex problem. We first list the following two properties of convex functions [33].

P1: Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{b} \in \mathbb{R}^n$. Define $g : \mathbb{R}^m \rightarrow \mathbb{R}$ by $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$, with domain(g) = $\{\mathbf{x} | \mathbf{Ax} + \mathbf{b} \in \text{domain}(f)\}$. If f is convex so is g .

P2: Non-negative weighted sum of convex functions is convex.

From Eq. 25, we can infer that f is the square of 2-norm (which is convex [33]) and the term inside the norm in Eq. 25

Algorithm 1 Algorithm for Recovering a Planar Scene Distorted by UCW

Input: (a) Initial PSF estimate. (b) Single motion blurred image due to water waves.

Output: Deskewed and deblurred image.

- 1: **repeat**
 - 2: Obtain latent image estimate $\hat{\mathbf{l}}$ from prediction step.
 - 3: By using $\hat{\mathbf{l}}$ from prediction step and the initial PSF given by the user, Eq. 23 is minimized to get an estimate of PSF.
 - 4: By using the estimate of PSF obtained from the previous step, Eq. 26 is minimized to get an estimate of the latent image \mathbf{l} .
 - 5: **until** $\|\mathbf{l} - \hat{\mathbf{l}}\|_2 < 10^{-6}$ or maximum number of iterations is reached
-

is affine in nature. Hence, convexity follows from P1. The second term (L_1 norm) in Eq. 25 is also convex [33]. Therefore, from P2, we can conclude that the energy function given by Eq. 25 is convex. By a similar argument, Eq. 26 is also a convex function. A summary of the proposed method is given in Algorithm 1. Although, there is no guarantee that we will find the global minimum, our experiments confirm that a satisfactory solution is indeed obtained in about 10 iterations. For the purpose of visualization, we have given in Figs. 3 (a-c) the evolution of the image and PSF estimates as iterations progress. It can be clearly observed that our restored result becomes sharper while the PSF estimate gets refined as the algorithm progresses. The convergence plot is given in Fig. 4.

V. EXPERIMENTAL RESULTS FOR UCW

In this section, we provide synthetic as well as real results for deskewing and deblurring in the presence of UCW (please see the pdf). We evaluate the performance of our approach and also compare it with state-of-the-art methods [8], [18] as they have been recently developed for deskewing of videos. For each method, the corresponding parameters that yielded best results were employed. The values of the various constants best suited for our method are $\gamma = 8$, $\beta = 0.1$, $\alpha_0 = 50$, $\alpha_1 = 25$, and $\alpha_2 = 12.5$. The input given to the competing methods is in the form of video, whereas the input given to our algorithm is a single motion blurred image. As mentioned earlier in section III, a long exposure is necessary for our scheme to realize space-invariant blurring. But this can lead to saturation of effects. We overcome this challenge by providing the temporal average of the frames given to competing methods.

In the synthetic experiments, we need to realize the forward degradation model. The analytical model that we use to synthesize shallow water waves [34] is

$$h(\mathbf{x}, \tau) = h_0 + A \sin(\omega_x x + \omega_y y + \phi(\tau)) \quad (28)$$

where ω_x and ω_y are the spatial frequencies in X and Y directions, respectively, A is the amplitude, h_0 is the height of the water column and $\phi(\tau)$ is the phase of the sine wave

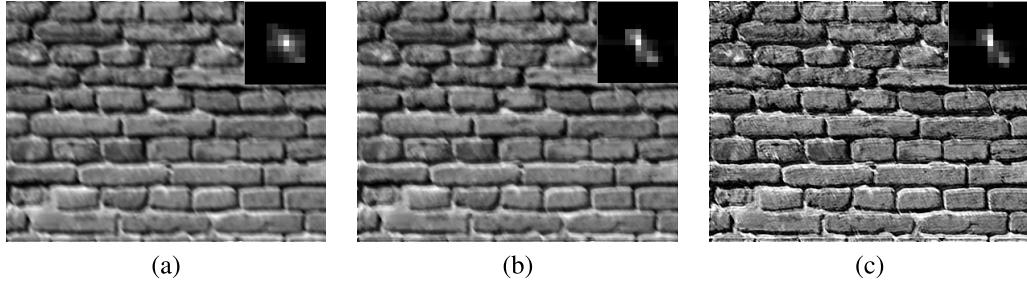


Fig. 3. Estimate of latent image and PSF with iterations. (a) 1st iteration, (b) 4th iteration, and (c) 10th iteration.

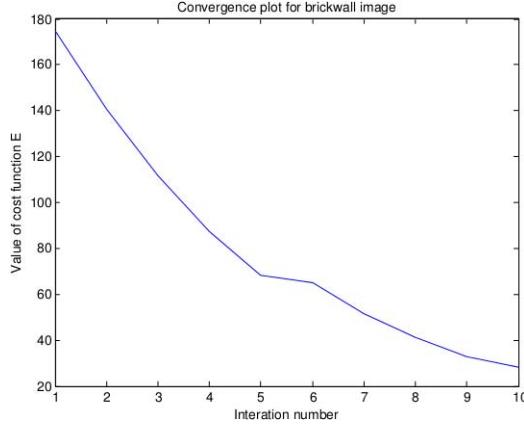


Fig. 4. Convergence plot with iterations.

which varies from 0 to 2π in steps of $\frac{2\pi}{k}$ to generate k number of frames, and $k = \max(\text{rows}, \text{cols})$, where rows and cols represent image dimensions. For our experiments, rows = 256 and cols = 256, and the variation in phase synthesizes the dynamic nature of water flow. All the k frames are averaged to synthesize the degraded observation. Although we use linear wave generation model for our synthetic experiments, there is no such restriction (except for cyclicity) on the real examples which can theoretically be of any shape. Also our method does not assume any functional form for the underlying wave.

Following other works, we also conduct our real experiments in an aquarium. We used a DSLR camera looking vertically downward at a planar scene immersed in an aquarium tank of dimensions 120cm × 60cm × 50cm. The waves were generated with a hair dryer which was kept away from the camera so as to generate 1D flow patterns within the aquarium. The approximate height of the ripples was between 1 and 2 cm with a spatial frequency of 4 cycles. The height of water column was about 20 cm. From the discussion in section III, the choice of c depends on the exposure time and the wind force generated from the hair dryer. For the operating conditions in our laboratory, the duration of video best suited for our experiments was found to be 2 seconds. The videos for the real experiments are included in the supplementary material.

In order to perform quantitative analysis, we chose three widely used metrics [18], [19], namely, peak-signal-to-noise ratio (PSNR), structure similarity index measure (SSIM), and normalized mutual information (NMI). They are defined as follows. Let MSE denote the mean square error between

two images (say I and K of size $m \times n$). The PSNR between them is calculated as

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right)$$

where MAX is the maximum possible gray level in an image (MAX = 255 for 8-bit images).

The structure similarity index measure (SSIM) between two image patches at locations \mathbf{x} and \mathbf{y} is calculated as [35]

$$\text{SSIM}_{\mathbf{x}, \mathbf{y}} = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (29)$$

where μ_x and μ_y are the means of the pixels in the window defined around \mathbf{x} and \mathbf{y} , respectively, σ_x^2 and σ_y^2 are the computed variances, and σ_{xy} is the covariance. The parameters $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ are used to stabilize the division, L is the dynamic range of the pixel values, $k_1 = 0.01$ and $k_2 = 0.03$ (by default). The SSIM value is the mean computed over all the locations using Eq. 29.

The normalized mutual information (NMI) [19] determines the accuracy of registration and is calculated as

$$\text{NMI}(I, K) = \frac{H(I) + H(K)}{H(I, K)}$$

where $H(I)$ and $H(K)$ are the entropies calculated for the images I and K , respectively, and $H(I, K)$ is the joint entropy. The expression for entropy of an image is given by $H(I) = -\sum_{i=0}^{255} p_i \log_2(p_i)$ where p_i is the probability of occurrence of intensity i at a pixel in image I . Similarly, the joint entropy between two images is defined as $H(I, K) = -\sum_{i=0}^{255} \sum_{k=0}^{255} p_{i,k} \log_2(p_{i,k})$, where $p_{i,k}$ is the joint probability of intensity i in image I and k in image K occurring together.

A. Synthetic Experiments

In this subsection, we present results on deblurring and deskewing of synthetically degraded images. The first column in Fig. 5 shows the ground truth images. Equation 28 was used to synthesize the dynamic water surface with $\omega_x = \frac{8\pi}{\text{rows}}$, $\omega_y = 0$, $A = 2$, $h_0 = 50$ (A and h_0 are in terms of pixels) while $\phi(\tau)$ is varied from 0 to 2π . The synthesized degraded observations shown in the second column are given as input to Algorithm 1. The third and fourth columns give results obtained from [8] and [18], respectively. The output of [8] is clearly blurry. Even though the result of [18] is sharper

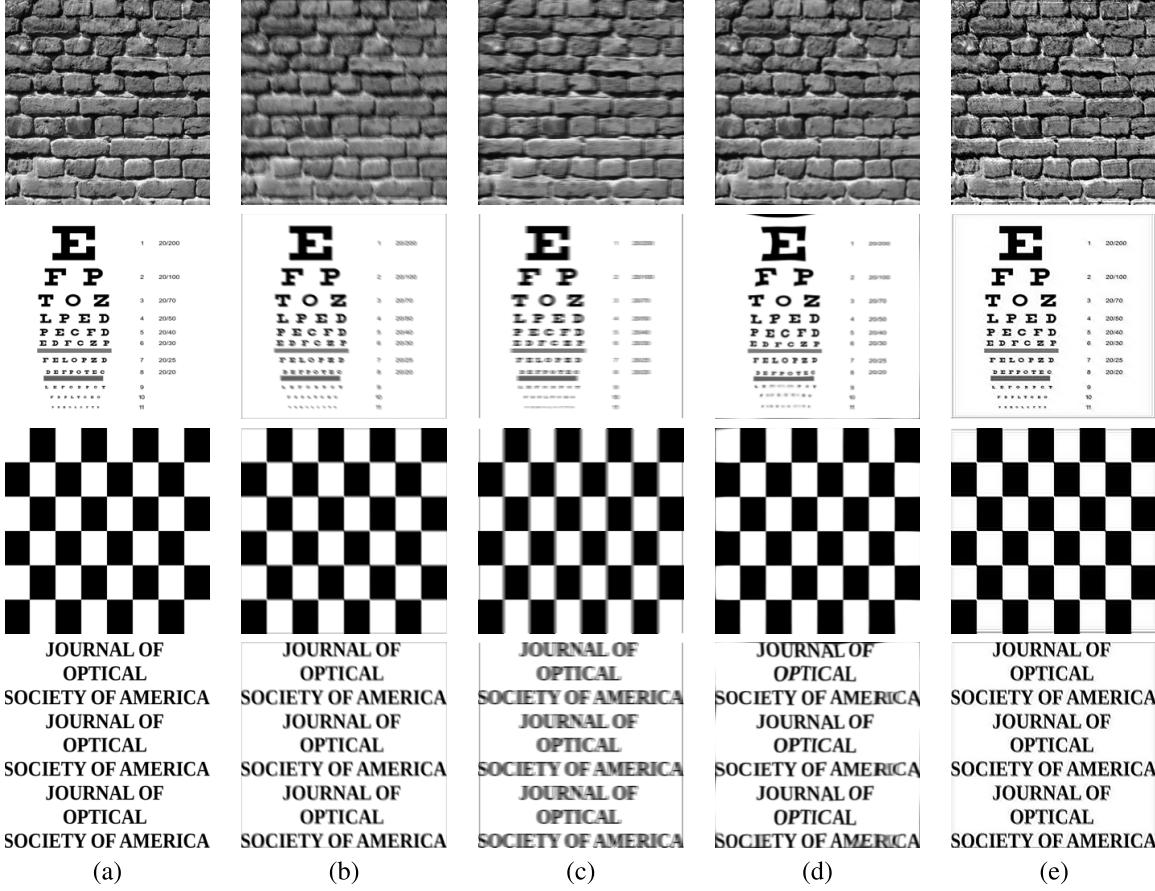


Fig. 5. Synthetic experiments to restore degraded images due to UCW: First row: brickwall. Second row: eyechart. Third row: checkerboard. Fourth row: text. (a) Ground truth, (b) blurred observation, (c) output of [8], (d) output of [18], and (e) output of our method.

TABLE I
COMPARISON WITH STATE-OF-THE-ART METHODS FOR SYNTHETIC UCW EXPERIMENTS

	PSNR (in dB)				SSIM			NMI				
	Blur	Tian [8]	Oreifej [18]	Algorithm 1	Blur	Tian [8]	Oreifej [18]	Algorithm 1	Blur	Tian [8]	Oreifej [18]	Algorithm 1
Brickwall	19.6513	22.4240	21.0315	23.7990	0.5572	0.7259	0.7197	0.8492	1.1208	1.1559	1.1452	1.1672
Eyechart	22.9896	21.4142	19.4327	23.5653	0.9190	0.8827	0.8563	0.9772	1.3892	1.4257	1.3575	1.4499
Checkerboard	23.7942	19.2969	22.0372	24.5278	0.9114	0.8215	0.9179	0.9280	1.4778	1.6368	1.5327	1.7404
Text	20.5040	15.0205	13.9449	22.9061	0.9232	0.7934	0.7796	0.9400	1.3384	1.2836	1.2359	1.3387

than [8], the skewness remains evident. This may be attributed to errors in non-rigid registration [19]. The last column gives the results obtained by our method. It is clearly evident that the shape of the bricks as well as sharpness are restored by our scheme. In the eyechart example, the skewness of the bigger alphabets as well as the blur in the smaller size letters stand corrected. Our restored result for the checkerboard is also quite sharp. In the 'JOSA' text example, it is again evident that the restored result is free from skew and blur. Table I shows the improvement in SSIM, PSNR (in dB) and NMI for our method along with comparisons. It is clearly evident that our method outperforms existing methods both qualitatively and quantitatively. The effect of curvature of the water waves on the performance of the proposed method is discussed in detail in the supplementary material [36].

B. Real Experiments

In this section, we give some representative results for real situations. The first column of Fig. 6 shows ground truth

images that were captured in still water in an aquarium. The second column shows degraded observations due to UCW generated by a hair dryer. Image degradation can be distinctly noted in the eyes of the birds, in the distorted shapes of the planes, deformed shapes of fishes, and in the letters in the last row. These observations are given as input to our algorithm. The outputs of [8] (third column) are blurred as they do not model blurring in their approach. Even though sharpness of the image is evident in the results obtained by [18] (fourth column), the skewing effect persists in the outputs. In contrast, the restored outputs for our method are close to the ground truth images both in terms of sharpness as well as degree of deskewing (although there are a few artifacts as we do not account for attenuation effects). In Table II, we provide quantitative comparisons. Note that the proposed method generally has higher PSNR, SSIM and NMI values as compared to the state-of-the-art. The failure of Tian's method [8] may be attributed to their not accounting for motion blur in the restoration process. Our method uses only one blurred observation which is the average of the frames

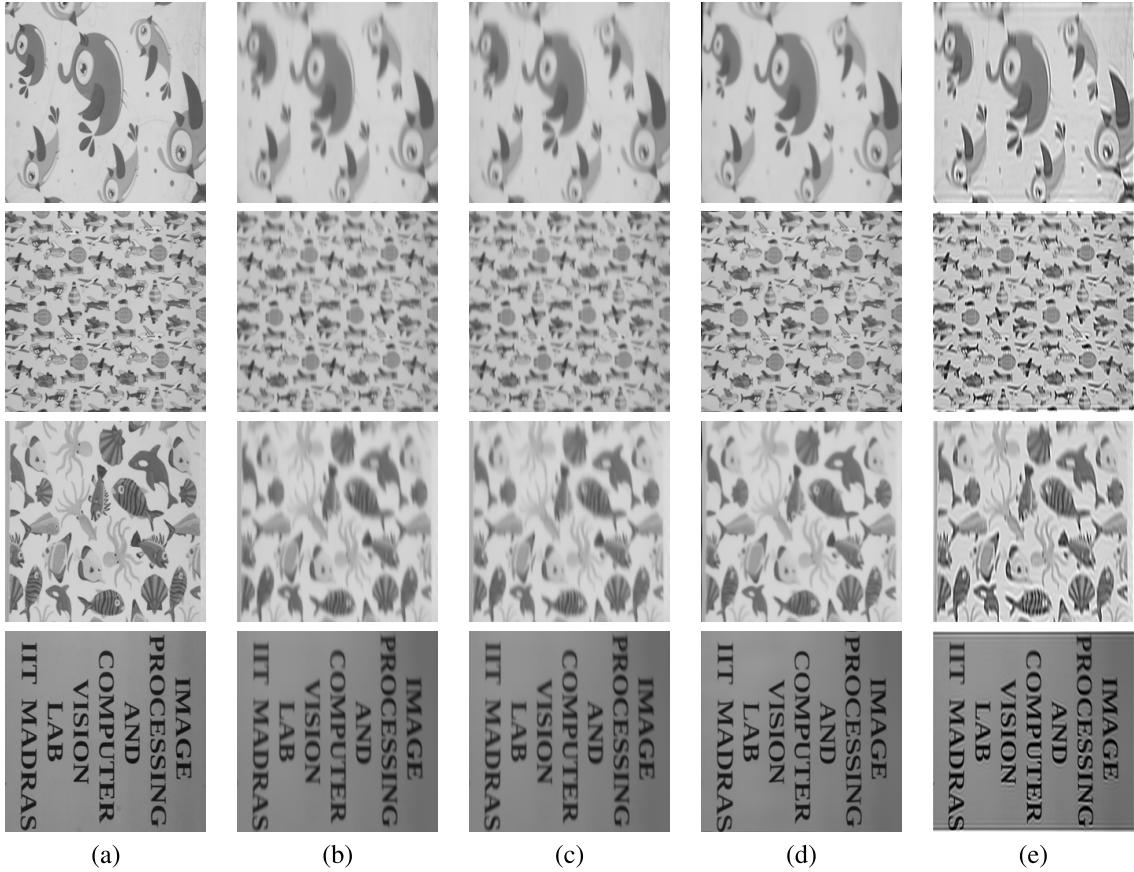


Fig. 6. UCW: Real indoor experiments for restoration of degraded images: First row: birds. Second row: planes. Third row: fishes. Fourth row: IPCV. (a) Ground truth, (b) skewed and motion blurred images, (c) restoration using [8], (d) restored output of [18], and (e) images restored using the proposed method.

TABLE II
COMPARISON OF OUR PROPOSED APPROACH IN THE PRESENCE OF UCW WITH OTHER METHODS ON REAL INDOOR DATA

	PSNR (in dB)				SSIM				NMI			
	Blur	Tian [8]	Oreifej [18]	Algorithm 1	Blur	Tian [8]	Oreifej [18]	Algorithm 1	Blur	Tian [8]	Oreifej [18]	Algorithm 1
Birds	27.2184	27.6205	27.9718	28.7190	0.8652	0.8734	0.8849	0.8913	1.2028	1.2028	1.2090	1.2829
Planes	22.0421	23.3494	22.0531	25.4122	0.6881	0.7590	0.7892	0.8506	1.2354	1.2463	1.2437	1.2617
Fishes	24.2461	23.8653	23.7464	25.6521	0.7819	0.7874	0.7650	0.8364	1.1538	1.1573	1.1450	1.1650
IPCV	23.5051	23.3494	23.3995	24.7142	0.8198	0.8766	0.8868	0.9140	1.1707	1.1488	1.1707	1.1802

(about 61 in number) provided to [8] and [18]. Because we use a single observation, the proposed method is (on an average) 5 times faster than [8] and 20 times faster than [18].

We also conducted outdoor experiments for the UCW scenario (as it is more natural) in a swimming pool in uncontrolled conditions. The planar scene was placed on the floor of the pool with a water depth of 1m. The source of waves was due to natural breeze. The results for three different planar scenes are given in Fig. 7. The output of [8] (first column of Fig. 7) is quite blurry, as it does not model blur. The second and third columns of Fig. 7 are the outputs of [18] and our method, respectively. The output of [18] is skewed at different places. This can be clearly seen in IPCV where the letters D and R in the word MADRAS were not restored properly. In contrast, our method uniformly delivers good results. Due to outdoor conditions, we could not capture observations corresponding to still water. To perform quantitative analysis, we adopted the cumulative probability for blur detection (CPBD) [37] which is a recently proposed no-reference metric and is based on the sensitivity of human blur perception at different sharpness levels. Table III gives

a comparison of CPBD values for the outputs obtained from our method vis-a-vis the state-of-the-art. Note that our method yet again performs the best both qualitatively and quantitatively. More details on the relevance of CPBD for our problem are given in the supplementary material [36].

Even though the main objective of this paper is not deblurring, yet for the sake of completeness, we also compare the performance of our algorithm with a well-known state-of-the-art space-invariant deblurring algorithm by Shan et al. [23]. Few representative results are given in Fig. 8. These results were obtained by using the executable file provided by the authors of [23]. Note that the results are inferior compared to our method (Figs. 5 and 6), as [23] is known to perform poorly in the presence of weak gradients.

VI. RESTORATION IN THE PRESENCE OF CIRCULAR RIPPLES

The previous sections dealt with UCW. However, circular wave pattern is also a commonly prevalent phenomenon in many situations such as when a stone is hurled into water

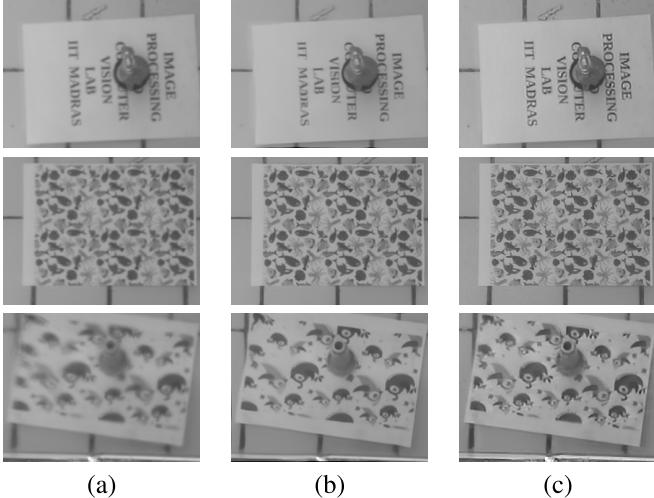


Fig. 7. Real outdoor experiments. First row: text. Second row: fishes. Third row: birds. (a) Output from [8], (b) result from [18], and (c) output of our method.

TABLE III
OUTDOOR UCW EXPERIMENTS: COMPARISON OF CPBD VALUES

	Tian et al. [8]	Oreifej et al. [18]	Proposed method
IPCV	0.0577	0.1739	0.6378
Fishes	0.0747	0.3545	0.6026
Birds	0.0521	0.1567	0.3360

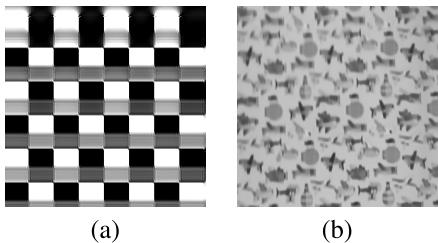


Fig. 8. Comparison with [23]. (a) Checkerboard, (b) planes.

or when rain drops land in a water pool or when water droplets descend from trees onto water puddles. Consider a situation when a scene of interest is inside a water body and a disturbance of the kind described above occurs on the water surface. The effect of refraction on the water surface will distort the underlying scene and it becomes a challenge to recover the original (undistorted) image. In this section, we will first objectively explain the generation of circular ripples. This will be followed by our approach for restoring planar scenes imaged in the presence of these waves.

A. Analytical Model for Circular Ripples

Circular ripples are widely used in physics for the study of interference and diffraction. The analytical model [38] for generating circular ripples is given by

$$h(\mathbf{x}, \tau) = h_0 + A \sin \left(\sqrt{(x' \omega_x)^2 + (y' \omega_y)^2} - \phi(\tau) \right) \quad (30)$$

where $x' = (x - x_0)$, $y' = (y - y_0)$, and (x_0, y_0) is the point on the water surface where the disturbance originates. As with

UCW, we assume negligible attenuation. We also assume that the ripple fluctuations are small, so that Eqs. 8 and 10 remain valid for circular ripples too. According to these equations, the warps that are to be applied on the image plane at time instant τ can be calculated by differentiating Eq. 30 to obtain

$$\mathbf{w}(\mathbf{x}, \tau) = \alpha \frac{A \omega \cos(\omega \sqrt{x'^2 + y'^2} - \phi(\tau))}{\sqrt{x'^2 + y'^2}} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (31)$$

where $\alpha = h_0(1 - \frac{1}{n})$. Parameterizing Eq. 31 by polar coordinates, we get

$$\mathbf{w}(\mathbf{x}, \tau) = \alpha A \omega \cos(\omega r - \phi(\tau)) \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (32)$$

where $r = \sqrt{x'^2 + y'^2}$, $x' = r \cos \theta$ and $y' = r \sin \theta$.

Claim 1: If A and ω are fixed, then the set of transformations experienced by pixels along the radial direction θ will be the same.

Proof: From Eq. 32, if we fix A , ω and θ , then the set of transformations induced at every pixel along a radial direction θ is determined by the factor $\cos(\omega r - \phi(\tau))$. Let \mathbf{w}_1 and \mathbf{w}_2 be the set of transformations induced at two pixels \mathbf{x}_1 and \mathbf{x}_2 which are at radial distances r and $r + \delta r$, respectively, along the direction θ . Thus, \mathbf{w}_1 is a function of $\cos(\omega r - \phi(\tau))$. While \mathbf{w}_2 is a function of $\cos(\omega(r + \delta r) - \phi(\tau))$, it can also be written as $\cos(\omega r + (\omega \delta r - \phi(\tau)))$ which is simply a phase-shifted version of $\cos(\omega r - \phi(\tau))$. This implies that if enough exposure is maintained such that a complete cycle of the wave passes through both the pixels, then the set of transformations in \mathbf{w}_1 and \mathbf{w}_2 will be the same. This argument can be extended to all the points along the radial direction θ . However, a major difference from UCW is that the blur induced due to circular ripples is invariant only along radial direction.

Claim 2: For fixed A and ω , the transformations at two different locations on a circular arc of radius r are related by the angular separation between those pixels.

Proof: Suppose we fix A , ω , and r . Eq. 32 then reduces to

$$\mathbf{w}(\mathbf{x}, \tau) = k(\tau) \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (33)$$

where $k(\tau) = \alpha A \omega \cos(\omega r - \phi(\tau))$. Let \mathbf{k} be the set of transformations induced at a particular radius r during exposure for $\theta = 0$. Then from Eq. 33, the set of transformations induced at any other point on the circular arc of radius r but at an angle θ will be given by $\mathbf{k} \angle \theta$ i.e., all the transformations in \mathbf{k} will be rotated by θ .

It is straightforward to further infer that claims 1 and 2, in fact, stand valid in general for circular ripples of any form (not necessarily sinusoidal). If \mathbf{f} and \mathbf{g} (both lexicographically-ordered) are the original and blurred observations, respectively, then they can be related by

$$\mathbf{g} = \mathbf{H}\mathbf{f} \quad (34)$$

where \mathbf{H} is the space-variant blurring matrix whose j^{th} column contains the blur kernel corresponding to the j^{th} entry in \mathbf{f} . There is an alternative interpretation for \mathbf{g} . From Eq. 31, it may be seen that the warps applied on the image plane at a particular time instant τ is dependent on the pixel location \mathbf{x} .

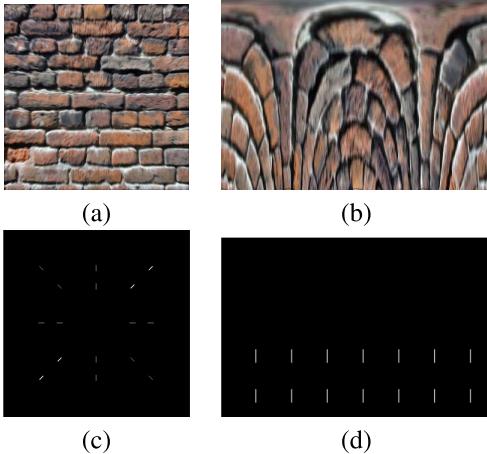


Fig. 9. (a) Brickwall in Cartesian coordinates, (b) image in polar domain, (c) estimated PSFs at various locations in (a), (d) estimated PSFs at different locations of (b).

This implies that every pixel will have its own translations based on its location, which induces geometric distortions on the image plane. Hence, the blurred image \mathbf{g} can also be treated as the average of different distorted or skewed versions of \mathbf{f} during the exposure time of the camera. Nevertheless, estimation of \mathbf{f} and \mathbf{H} given \mathbf{g} is a very ill-posed problem. If $\mathbf{f} \in \mathbb{R}^{M^2 \times 1}$ then $\mathbf{H} \in \mathbb{R}^{M^2 \times M^2}$. Hence, even for small M , the computational complexity will be very high. Interestingly, the claims 1 and 2 established in the earlier part of this section can be exploited judiciously as discussed next.

B. Deblurring in Polar Domain

In this subsection, we describe our method for restoring images degraded by circular ripples. In Fig. 9(a) we show a blurred image in rectangular coordinates. The same image when plotted in polar coordinates is as given in Fig. 9(b), where the horizontal and the vertical axis correspond to θ and r , respectively. The transformation from rectangular to polar domain was done by assuming the image center to be the origin. In Fig. 9(c), we show an image with PSFs estimated at different locations corresponding to the patches in Fig. 9(a). It can be clearly observed from this figure that the region of support for PSFs along a specific direction remains the same (which is in line with claim 1). Also, the PSFs on any circular contour are rotated versions of each other (in line with claim 2). In Fig. 9(d), we show the PSFs estimated at different locations in Fig. 9(b). All the pixels along a particular radial direction θ in Cartesian domain map to a vertical line in the polar domain (corresponding to θ). This implies that the PSF corresponding to a pixel along a radial direction θ will orient along a corresponding vertical line in the polar domain. Hence, the PSFs will map to different vertical lines in the polar domain based on the pixel to which they belong (as shown in Fig. 9(c)). This implies that it should be possible to invoke a space-invariant restoration formulation for this problem too, albeit in the polar domain. We would like to point to an interesting analogy to an earlier landmark work by Sawchuk [39] in which he had exploited the fact that the

Algorithm 2 Algorithm for Recovering a Distorted Scene in the Presence of Circular Ripples

Input: (a) Single motion blurred image due to circular ripples. (b) Initial PSF in polar domain.

Output: Deblurred and deskewed image.

- 1: Convert the blurred observation into polar domain.
 - 2: Use Algorithm 1 to obtain an estimate of the original image \mathbf{f} in the polar domain.
 - 3: Convert the restored image (which is in polar domain) back to rectangular coordinates.
-

space-variant nature of blur in rectangular coordinates due to camera rotation can be rendered as space-invariant in the polar domain. Our scenario is different from [39] in that the blur is space-invariant in the radial direction alone.

For restoration, we first convert the blurred observation into the polar domain. Since the blur in the transform domain is uniform (i.e. space-invariant), we invoke the restoration framework followed earlier for images degraded by UCW. The priors used in Algorithm 1 are equally applicable here too. The methodology adopted for deskewing and deblurring the images degraded by circular ripples is given succinctly in Algorithm 2.

VII. EXPERIMENTAL RESULTS FOR CIRCULAR RIPPLES

In this section, we demonstrate the performance of Algorithm 2 on synthetic and real data for restoring images degraded by skew and motion blur due to circular ripples. For the synthetic experiments alone, the analytical model described in Eq. 30 was used for generation of circular ripples. For the real experiments, we used a burette to generate circular ripples within our aquarium. The screw of the burette was adjusted to control the rate of flow of water droplets. The size of the planar scene was chosen such that the attenuation factor could be neglected. The mapping from Cartesian to polar domain varies with the choice of the reference point. For a scaled orthographic camera, back projection of the image center corresponds to the center of the FOV of the camera. Hence, for real experiments, we ensured that the water droplets from the burette fell approximately at the center of the camera's FOV so that the image center could be considered as the reference point for performing coordinate transformations.

A. Synthetic Experiments

To synthesize outward movement of circular ripples post-disturbance, we used Eq. 30 with $\omega = \frac{6\pi}{k}$, $A = 4$, $h_0 = 50$ (in terms of pixels) while $\phi(\tau)$ was varied from 0 to 2π . In Fig. 10, the first column contains the ground truth images. The second column shows some of the synthetically generated images which are distorted due to both blur and skew under the influence of circular ripples. Note that all the examples suffer from significant amounts of degradation. The third and fourth columns show the outputs of [8] and [18], respectively, while our results using Algorithm 2 are given in the last column. A comparison of the outputs reveals that our method yields

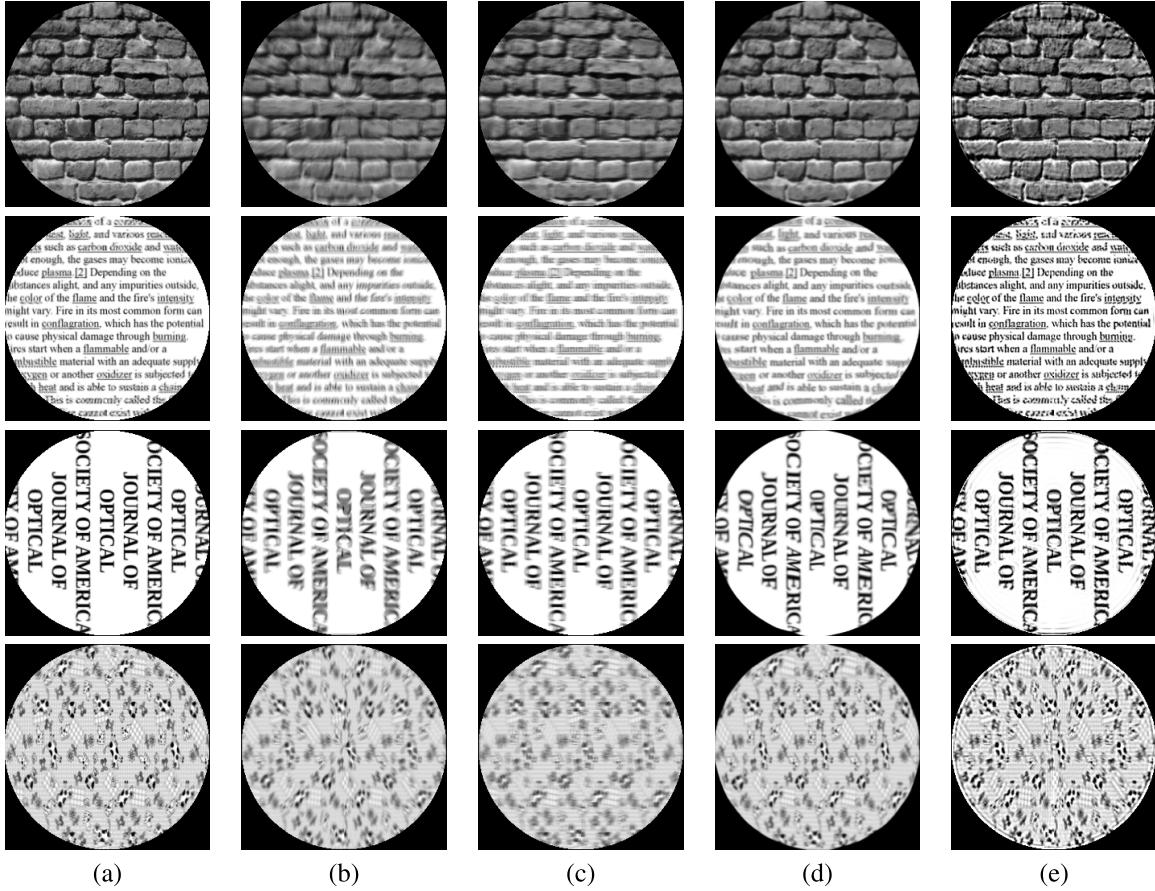


Fig. 10. Synthetic experiments to restore the degraded images in the presence of circular ripples: First row: brickwall. Second row: text. Third row: JOSA. Fourth row: texture. (a) Ground truth, (b) temporal average of distorted video, (c) result of [8], (d) result of [18], and (e) output of Algorithm 2.

TABLE IV

SYNTHETIC EXPERIMENTS: COMPARISON OF SCORES FOR DIFFERENT METHODS IN PRESENCE OF CIRCULAR RIPPLES

	PSNR (in dB)				SSIM				NMI			
	Blur	Tian [8]	Oreifej [18]	Algorithm 2	Blur	Tian [8]	Oreifej [18]	Algorithm 2	Blur	Tian [8]	Oreifej [18]	Algorithm 2
Brickwall	21.9779	22.3291	22.7864	23.2619	0.7801	0.8708	0.8437	0.8791	1.2157	1.2554	1.2296	1.2673
Text	18.4052	17.0964	17.4712	18.6798	0.7582	0.7000	0.7213	0.9066	1.2426	1.2451	1.2177	1.3036
JOSA	16.4220	18.9146	11.3665	20.7420	0.8124	0.8634	0.6825	0.9186	1.3226	1.3286	1.2997	1.3299
Texture	20.8349	19.3000	18.9889	21.2364	0.8031	0.7034	0.7297	0.8637	1.2074	1.1832	1.1848	1.3034

TABLE V

REAL EXPERIMENTS: RESULTS FOR RESTORATION IN THE PRESENCE OF CIRCULAR RIPPLES

	PSNR (in dB)				SSIM				NMI			
	Blur	Tian [8]	Oreifej [18]	Algorithm 2	Blur	Tian [8]	Oreifej [18]	Algorithm 2	Blur	Tian [8]	Oreifej [18]	Algorithm 2
Soldiers	23.2161	22.3897	22.0637	24.7994	0.9225	0.9060	0.8626	0.9405	1.2784	1.2695	1.2760	1.2877
Planes	26.4353	25.1718	24.6342	27.2596	0.9366	0.9390	0.9034	0.9401	1.3073	1.2947	1.2809	1.3804
Eyechart	23.1920	21.9767	20.1264	24.8690	0.9454	0.9325	0.8797	0.9520	1.3072	1.3081	1.3090	1.3133
Texture	23.8547	22.8205	21.6755	24.3010	0.9290	0.8928	0.8481	0.9319	1.2029	1.2780	1.2544	1.2847

the best results. From the PSNR, SSIM and NMI values given in Table IV, it is quite clear that our method scores even by quantitative metrics.

B. Real Experiments

In the first column of Fig. 11, we display images acquired in still water in the aquarium. These are treated as ground truth images. Once the burrette was actuated, videos were captured under the influence of circular ripples. The second column depicts the blurred image obtained by averaging

the frames of the video captured over 2 seconds. Results obtained using [8], [18] are given in columns three and four, respectively, while the output of our method (Algorithm 2) is shown in the last column. Observe that the restored results by our method are sharp and quite close to the images captured under still water conditions. The restored results of [8] are again blurred. Although the results of [18] are reasonably sharp, deskewing has not been achieved satisfactorily. This is quite apparent in the eyechart example where the letters as well as the straight line are visibly skewed. Table V gives PSNR, SSIM and NMI values for the restored images for

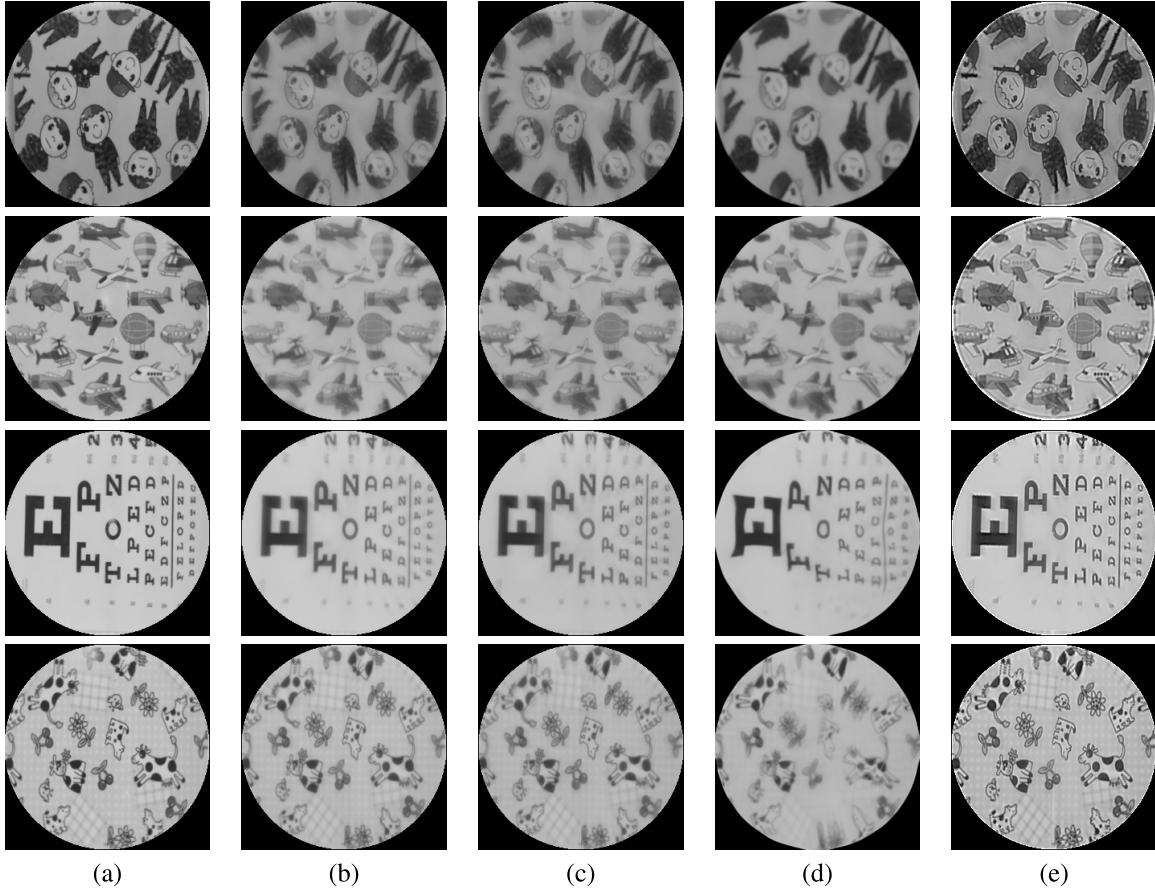


Fig. 11. Real indoor experiments (circular ripples): First row: soldiers. Second row: planes. Third row: eye chart. Fourth row: texture. (a) Ground truth, (b) degraded observations, (c) restoration by [8], (d) restoration by [18], and (e) restoration by Algorithm 2.

all the methods. The proposed Algorithm 2 yet again yields significant quantitative improvements. In comparison to the UCW case, the attenuation effects are barely visible since the circular ripples originate from the center of the image.

VIII. CONCLUSIONS

We proposed a method for deskewing and deblurring videos distorted by the dynamic nature of water surface. Existing methods typically need multiple observations to address this problem. In this work, we demonstrated that it is possible to perform deskewing and deblurring using a *single* blurred observation under certain modest constraints on the water flow. Initially, we established that the blur induced by UCW is space-invariant in nature and proposed a unified framework to deskew and deblur a distorted video. We then revealed that the nature of blur due to circular ripples is space-variant in the Cartesian plane and proved that a space-invariant restoration approach can still be employed but in the polar domain. The proposed algorithms advance the state-of-the-art as verified with several synthetic and real examples. We plan to extend the scope of these methods to include even attenuation effects.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor as well as the anonymous reviewers for their comments and suggestions.

REFERENCES

- [1] R. Schettini and S. Corchs, "Underwater image processing: State of the art of restoration and image enhancement methods," *EURASIP J. Adv. Signal Process.*, vol. 2010, Jan. 2010, Art. ID 14.
- [2] D. G. Turlaev and L. S. Dolin, "On observing underwater objects through a wavy water surface: A new algorithm for image correction and laboratory experiment," *Izvestiya, Atmosph. Ocean. Phys.*, vol. 49, no. 3, pp. 339–345, 2013.
- [3] R. Shefer, M. Malhi, and A. Shenhari. (2001). *Waves Distortion Correction Using Crosscorrelation*. [Online]. Available: <http://visl.technion.ac.il/projects/2000maor/>
- [4] H. Murase, "Surface shape reconstruction of a nonrigid transport object using refraction and motion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 10, pp. 1045–1052, Oct. 1992.
- [5] A. A. Efros, V. Isler, J. Shi, and M. Visontai, "Seeing through water," in *Proc. Adv. Neural Inf. Process. Syst.*, 2004, pp. 393–400.
- [6] Z. Wen, A. Lambert, D. Fraser, and H. Li, "Bispectral analysis and recovery of images distorted by a moving water surface," *Appl. Opt.*, vol. 49, no. 33, pp. 6376–6384, 2010.
- [7] Z. Wen, D. Fraser, and A. Lambert, "Bicoherence: A new lucky region technique in anisoplanatic image restoration," *Appl. Opt.*, vol. 48, no. 32, pp. 6111–6119, 2009.
- [8] Y. Tian and S. G. Narasimhan, "Seeing through water: Image restoration using model-based tracking," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2303–2310.
- [9] Y. Tian and S. G. Narasimhan, "A globally optimal data-driven approach for image distortion estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1277–1284.
- [10] Z. Hu and M.-H. Yang, "Fast non-uniform deblurring using constrained camera pose subspace," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 136.1–136.11.
- [11] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2010, pp. 157–170.

- [12] L. Xu, S. Zheng, and J. Jia, "Unnatural L_0 sparse representation for natural image deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1107–1114.
- [13] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf, "Fast removal of non-uniform camera shake," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 463–470.
- [14] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 607–614.
- [15] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186, 2012.
- [16] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2010, pp. 171–184.
- [17] A. Donate and E. Ribeiro, "Improved reconstruction of images distorted by water waves," in *Advances in Computer Graphics and Computer Vision*. Berlin, Germany: Springer-Verlag, 2007, pp. 264–277.
- [18] O. Oreifej, G. Shu, T. Pace, and M. Shah, "A two-stage reconstruction approach for seeing through water," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 1153–1160.
- [19] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
- [20] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 763–770.
- [21] P. Cattin, "Image restoration: Introduction to signal and image processing," 2013. [Online]. Available: <https://miac.unibas.ch/SIP/06Restoration.html>
- [22] A. C. Bovik, *Handbook of Image and Video Processing*. New York, NY, USA: Academic, 2010.
- [23] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. ID 73.
- [24] P. Chandramouli and A. Rajagopalan, "Inferring image transformation and structure from motion-blurred images," in *Proc. Brit. Mach. Vis. Conf.*, 2010, pp. 73–1–73–12.
- [25] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, 2006.
- [26] T. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. Philadelphia, PA, USA: SIAM, 2005.
- [27] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [28] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. ID 70.
- [29] F. Sroubek, G. Cristóbal, and J. Flusser, "A unified approach to super-resolution and multichannel blind deconvolution," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2322–2332, Sep. 2007.
- [30] S. Cho and S. Lee, "Fast motion deblurring," *ACM Trans. Graph.*, vol. 28, no. 5, 2009, Art. ID 145.
- [31] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, Mar. 2010.
- [32] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. B, Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [34] A. D. D. Craik, "The origins of water wave theory," *Annu. Rev. Fluid Mech.*, vol. 36, pp. 1–28, Jan. 2004.
- [35] 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.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [36] [Online]. Available: <http://www.ee.iitm.ac.in/pcvlab/research/deskew>
- [37] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678–2683, Sep. 2011.
- [38] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [39] A. A. Sawchuk, "Space-variant image motion degradation and restoration," *Proc. IEEE*, vol. 60, no. 7, pp. 854–861, Jul. 1972.



Karthik Seemakurthy received the B.Tech. degree in electronics and communications engineering from the Vidya Jyothi Institute of Technology, Hyderabad, India, in 2008, and the M.Tech. degree in visual information processing and embedded systems specialization from IIT Kharagpur, Kharagpur, India, in 2010. He is currently pursuing the Ph.D. degree with IIT Madras, Chennai, India. His research interests are in the areas of computer vision and image processing with specific focus on underwater imaging.



Ambasamudram Narayanan Rajagopalan received the Ph.D. degree in electrical engineering from IIT Bombay, Mumbai, India, in 1998. He was with the Center for Automation Research, University of Maryland, College Park, MD, USA, as a Research Faculty Member, from 1998 to 2000. He joined the Department of Electrical Engineering, IIT Madras, Chennai, India, in 2000, where he currently serves as a Professor. His research interests include 3D structure from blur, motion deblurring, registration, super-resolution, video microscopy, inpainting, matting, heritage resurrection, face recognition, image forensics, and underwater imaging. He has co-authored a book entitled *Depth from Defocus: A Real Aperture Imaging Approach* (New York, NY, USA: Springer-Verlag, 1999). He is a fellow of the Alexander von Humboldt Foundation in Germany, the Indian National Academy of Engineering, and the Institution of Electronics and Telecommunication Engineers in India (by invitation). He served as an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE from 2007 to 2011, and is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING. He was an Area Chair of CVPR and ICPR in 2012, and the Program Co-Chair of ICVGIP in 2010. He was a recipient of the DAE-SRC Outstanding Investigator Award in 2012 and the VASVIK Award in 2013. He received the Mid-Career Research and Development Award from IIT Madras in 2014. He co-edited a book entitled *Motion Deblurring: Algorithms and Systems* (Cambridge University Press, 2014) along with Prof. R. Chellappa of the University of Maryland.