Pattern Recognition 69 (2017) 251–269

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Two-view underwater 3D reconstruction for cameras with unknown poses under flat refractive interfaces

Lai Kang[a,*], Lingda Wu[b,c], Yingmei Wei[b], Songyang Lao[b], Yee-Hong Yang[d]

[a] Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha, 410073 China
[b] College of Information System and Management, National University of Defense Technology, Changsha, 410073 China
[c] The Key Lab, The Academy of Equipment Command and Technology, Beijing, 101416 China
[d] Department of Computing Science, University of Alberta, Edmonton, T6G 2E8 Canada

ARTICLE INFO

Article history:
Received 28 June 2016
Revised 16 March 2017
Accepted 4 April 2017
Available online 21 April 2017

Keywords:
Structure and Motion (SaM)
Refractive distortion
Underwater calibration
Underwater 3D reconstruction
Hybrid optimization

ABSTRACT

In an underwater imaging system, a refractive interface is introduced when a camera looks into the water-based environment, resulting in distorted images due to the refraction of light. Simply ignoring the refraction effect or using the lens radial distortion model causes erroneous 3D reconstruction. This paper deals with a general underwater imaging setup using two cameras, of which each camera is placed in a separate waterproof housing with a flat glass window. In order to handle refraction properly, a simplified refractive camera model is used in this paper. Based on two new concepts, namely the Ellipse of Refrax (EoR) and the Refractive Depth (RD) of a scene point, we derive two new formulations of the underwater known rotation structure and motion (SaM) problem. One gives a globally optimal solution and the other is robust to outliers. The constraint of known rotation is further relaxed by incorporating the robust known rotation SaM into a new hybrid optimization framework. Our method is able to simultaneously perform underwater camera calibration and 3D reconstruction automatically without using any calibration object or additional calibration device. In order to evaluate the performance and practical applicability of our method, extensive experiments using synthetic data, synthetically rendered images and real underwater images were carried out. The experimental results demonstrate that the proposed method can significantly improve the accuracy of the reconstructed 3D structure (within 0.78 mm for an object of dimension over 200 mm compared with the ground truth model captured by a land-based system) and of the system parameters for underwater applications. Compared with bundle adjustment using the refractive camera model initialized with traditional 3D reconstruction methods, our proposed optimization method has significantly better completeness and accuracy and lower 3D errors in the reconstructed models.

1. Introduction

Underwater imaging and 3D reconstruction has a wide range of potential applications, which include biological observation, robot navigation, seafloor digitization, etc. While remarkable success has been achieved for land-based systems [1], accurate 3D reconstruction from images captured by underwater cameras, however, has not attracted much attention from the computer vision community only until recently [2–5]. In a typical underwater imaging system, a perspective camera is often placed outside a water-filled tank or in a waterproof housing with a flat glass window. Such a system faces challenges posed by the underwater environment and image distortion caused by the refraction of light. In particular, the quality of image degrades due to the effects of light attenuation and scattering. A more severe issue in underwater vision is the change of light direction due to the refraction of light, which occurs when a light ray passes through the water-glass and air-glass interface, rendering the conventional single viewpoint (SVP) camera model invalid in theory. The refractive distortion is known to be highly nonlinear and depends not only on the distance of a scene point from the optical axis but also on its depth [2].

In order to compensate for refractive distortion, a physically correct refractive camera model has been introduced to explicitly model the effect of refraction [6]. Although exciting improvements in both camera calibration and 3D reconstruction have been achieved recently [2–5,7,8], dense underwater 3D reconstruction based on the refractive camera model remains an active research topic. In fact, solutions for many key components (e.g., fundamental matrix estimation, triangulation, stereo matching, etc.) of

* Corresponding author.
E-mail address: lkang.vr@gmail.com (L. Kang).

http://dx.doi.org/10.1016/j.patcog.2017.04.006

image-based underwater 3D reconstruction have not been well-established till now.

In this paper, we focus on automatic 3D reconstruction solely based on images captured by a general underwater imaging setup consisting of two cameras, which are placed in separated waterproof housings each with a flat glass window, without using any additional calibration object or additional device. In particular, two new concepts, namely the Ellipse of Refrax (EoR) and the Refractive Depth (RD) of a scene point for the refractive camera model are presented. These two concepts facilitate the derivation of two new algorithms for the estimation of relative camera translation, refractive interface distances and 3D structure under L_∞-norm, given known rotation (camera rotation and interface normal). Furthermore, a new hybrid optimization framework is proposed to tackle the two-view underwater structure and motion (SaM) problem. Within this framework, the constraint of known rotation is relaxed and the geometrically meaningful reprojection error in the image space is minimized. In order to evaluate our proposed method qualitatively and quantitatively, extensive experiments using synthetic data, synthetically rendered underwater images and real underwater images were carried out. Additionally, we also show in our experiments that our results are better than that using bundle adjustment.

2. Related work

In early works, the effects of refraction in underwater 3D reconstruction are simply ignored [9] or modeled by approximate methods, such as focal length adjustment [10], lens radial distortion [11] and a combination of the two [12]. Unfortunately, these methods are insufficient since the effect of refraction is known to be highly nonlinear and depends on the 3D location of a scene. As shown by Treibitz et al. [2], assuming a single viewpoint (SVP) model can be erroneous for camera calibration in underwater applications. A more desirable method to compensate for refraction is to use a physically correct refractive camera model. Chari and Sturm [6] analyze using theoretical analysis the underlying multi-view relationships between two cameras when the scene has a single refractive planar surface separating two different media. The authors demonstrate the existence of geometric entities such as the refractive fundamental matrix and the refractive homography matrix. However, no practical application of these theoretical results has been demonstrated so far [6]. A limitation for most existing underwater photography works is that a calibration target with known dimensions or an additional calibration device is required to perform system calibration [2,4,5,13–15].

Recently, Agrawal et al. [4] recognize that the underlying refractive geometry corresponds to an axial camera and develop a general theory for calibrating such systems using a planar checkerboard. Chang and Chen [3] study a similar configuration involving multiple views of a scene through a single interface. In their work, refractive distortion is explicitly modeled as a function of depth, and an additional piece of hardware called the inertial measurement unit (IMU) is required to provide the roll and pitch angles of the camera. Also, the normal of the refractive interface is assumed to be known. Based on the additional information, the authors derive a linear solution to the relative pose problem and a closed-form solution to the absolute pose problem. More recently, Yau et al. [5] propose a novel underwater calibration method using a specially designed calibration device consisting of a watertight acrylic enclosure containing 122 LEDs arranged in a grid pattern. The method takes advantage of the dispersion of light and achieves improved calibration accuracy as compared with that of Agrawal et al. [4].

Methods that do not require a calibration object or additional calibration device also exist, aiming at minimizing virtual camera errors rather than the commonly used reprojection errors in the image space. Jordt-Sedlazeck and Koch [16] propose a method for the calibration of housing parameters in underwater stereo camera rigs. The error on the outer interface plane is minimized by deriving the virtual perspective projection [14] for each 3D point. One issue of this method is, as reported in [16], that the optimization process is time consuming (in the order of 3 hours). More recently, the idea of minimizing virtual camera errors [16] is incorporated into bundle adjustment [17] and a more efficient method for underwater structure-from-motion (SfM) is developed in [7].

In order to analyze the distortions in non-SVP imaging systems, Swaminathan et al. [18] introduce a taxonomy of distortions based on the geometry of imaging systems and derive a metric to quantify caustic distortions. The authors also present an algorithm to compute minimally distorted images using simple priors of a scene structure. Treibitz et al. [2] also analyze the refractive distortion and show the error in SVP approximation of a refractive camera using simulations. For multi-view (more than two views) underwater 3D reconstruction, the influence of refractive distortion on the accuracy of 3D reconstruction is evaluated quantitatively and systematically in [19].

Unlike most existing recent works mainly focusing on underwater camera calibration using calibration object, this paper addresses the problems of system calibration and 3D scene reconstruction without using any calibration object or special devices. A preliminary version of this paper appeared in [20], compared with which several major extensions are given in this paper:

- First, we discuss in more detail and accompany with simulations the effects of refractive distortion and epipolar errors. Such errors creates erroneous results in camera calibration and in stereo matching for an underwater imaging system.
- Second, we develop a new hybrid optimization framework for two-view underwater calibration, in which a new local sliding (LS) procedure based on bundle adjustment is used to guide a local search. Such a seemingly simple strategy improves the convergence of our proposed hybrid optimization framework significantly.
- Third, by incorporating the robust underwater known rotation SaM algorithm into our proposed hybrid optimization framework, we realize a new fully automatic method for two-view underwater structure and motion estimation. In contrast, in our previous work [20], a subset of outlier free image correspondences is selected manually.
- Fourth, in addition to performance evaluation by visual inspection, the accuracy of 3D reconstruction using our proposed method is evaluated quantitatively by comparing the results to ground truth 3D models reconstructed using a land-based system.
- Finally, the influence of non-zero thickness and non-parallel interface on underwater 3D reconstruction is studied experimentally. The rationale of the assumption of zero thickness and parallel refractive interface in the simplified refractive camera model is discussed and verified in our experiments.

The remainder of this paper is organized as follows: Section 4.1 presents preliminaries on the refractive camera model and an analysis of refraction induced distortion in images. In Section 4, we first introduce two new concepts, namely the Ellipse of Refrax (EoR) and the Refractive Depth (RD) under the refractive camera model. Then we propose two new formulations of the underwater structure and motion (SaM) with known rotation problem. In Section 5, we propose a new hybrid optimization framework for two-view underwater calibration and 3D reconstruction and introduce a way to perform dense underwater 3D reconstruction. Section 6 presents qualitative and quantitative evaluation and comparison using synthetic data, synthetically