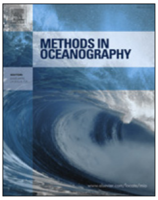
Methods in Oceanography 17 (2016) 118–137

Contents lists available at ScienceDirect

Methods in Oceanography

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



Full length article

Dynamic 3D capture of swimming fish by underwater active stereo

Ryo Kawahara[*], Shohei Nobuhara, Takashi Matsuyama

Graduate School of Informatics, Kyoto University, Japan

HIGHLIGHTS

- A practical underwater active stereo system for dynamic 3D capture of object in water with explicit refraction modeling.
- A practical projector–camera calibration in water.
- An efficient projection computation of 3D points in water to 2D camera images.

ARTICLE INFO

Article history:
Received 30 June 2015
Received in revised form
4 July 2016
Accepted 5 August 2016

Keywords:
Underwater active stereo
3D shape reconstruction
Refraction
Calibration

ABSTRACT

This paper presents an underwater active stereo system that realizes 3D capture of dynamic objects in water such as swimming fish. The key idea on realizing a practical underwater 3D sensing is to model the refraction process by our pixel-wise varifocal camera model that provides efficient forward (3D to 2D) projections as well as an underwater projector–camera calibration. Evaluations demonstrate that our method achieves reasonable calibration accuracy using off-the-shelf cameras and projectors, and provides a 3D capture of real swimming fish in water.

1. Introduction

Image-based 3D shape modeling has been a fundamental research goal in computer vision for decades, and recent advances realized practical solutions for capturing dynamic objects such as human (Starck et al., 2006; Furukawa and Ponce, 2007; Matsuyama et al., 2012). Today it provides

---

[*] Corresponding author.
E-mail addresses: kawahara@vision.kuee.kyoto-u.ac.jp (R. Kawahara), nob@i.kyoto-u.ac.jp (S. Nobuhara), tm@i.kyoto-u.ac.jp (T. Matsuyama).

http://dx.doi.org/10.1016/j.mio.2016.08.002

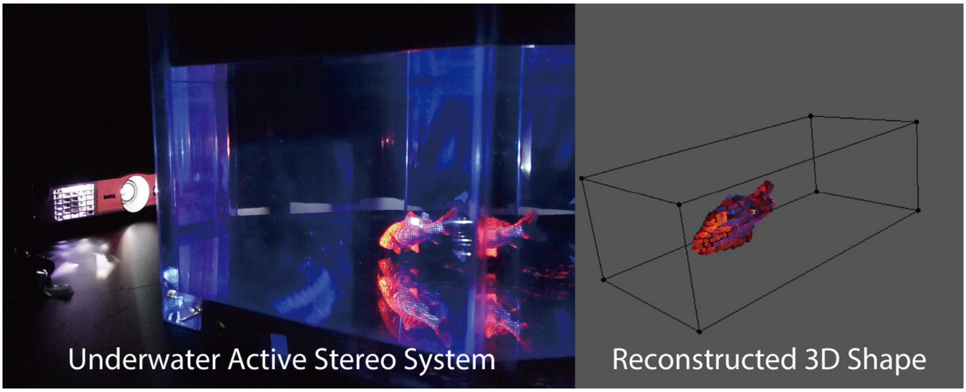


Fig. 1. Dynamic 3D capture of swimming fish by underwater active stereo. Left: our capture system consisting of cameras and projectors observing the fish via flat surfaces. Right: reconstructed 3D shape.

nonconstrained and noninvasive measurements of its 3D shape and motion, and widely used as a marker-less motion capture (Moeslund et al., 2006; Ballan and Cortelazzo, 2008; Corazza et al., 2010; Liu et al., 2013) for movie production, medical analysis, and so forth. The goal of this paper is to realize such an image-based 3D acquisition for underwater objects as shown in Fig. 1. We believe that providing an automated quantitative 3D sensing of underwater objects will help advancing studies in oceanography, marine biology, aquaculture, etc.

Realizing an image-based underwater 3D capture is not a trivial problem even though many algorithms have been developed for objects in the air. The difficulty can be found in the environment and the object. Capturing images in water inevitably involves refractive image distortions by housings (Treibitz et al., 2008, 2012), and image noise and intensity attenuations by unclear media (Narasimhan et al., 2005). Also the objects to be captured in water are likely to show poorly-textured surfaces with transparency and specularity. These properties are not dominant in 3D capture in the air, and hence should be handled explicitly for underwater 3D capture.

The key problem addressed in this paper is the refraction by water-proof housings that lead to incorrect 3D reasoning unless modeled correctly. That is, if the refraction is ignored and the ray is wrongly modeled as a straight line passing through the camera center as in the air, 3D measurement by stereo cannot return the correct triangulation point obviously since the ray passing through the camera center and the pixel should be refracted by the housing.

The refraction is well described by Snell's law. The motivation of this paper is not to introduce another law, but to introduce a new practical representation of the refraction described by Snell's law. We show that our representation provides a faster computation of 3D-to-2D forward projection as well as a linear calibration of projector–camera systems in water.

The key idea of our approach is to introduce a virtual camera model that defines a focal length on a per-pixel basis, and is to exploit a radial structure of such pixel-wise focal lengths to realize a compact and efficient representation of the refraction. As is well known, the refraction by the housing result in a caustic structure of rays as shown in Fig. 2(a), and this structure deforms depending on the relative pose of the housing and the camera. We show that our model can model cameras and projectors behind flat housings regardless of their poses, and that each of them can be converted to a virtual camera or a projector having radially-symmetric pixel-wise focal lengths (Fig. 2(b)).

Based on this new camera model, this paper shows that we can realize a linear calibration of projectors and cameras in water, and can realize a practical 3D capture of underwater objects such as swimming fish.

Notice that earlier versions of this paper are partially reported in Kawahara et al. (2013, 2014). The contribution of this paper is twofold. We evaluate the accuracy of our underwater projector–camera system using a real dataset, and also provide a practical system that realizes underwater 3D capture by active stereo.