A novel robotic visual perception framework for underwater operation[*][#]

Yue LU[1], Xingyu CHEN[2], Zhengxing WU[1], Junzhi YU[†‡1,3], Li WEN[4]

[1]State Key Laboratory of Management and Control for Complex Systems, Institute of Automation,

Chinese Academy of Sciences, Beijing 100190, China

[2]Ytech, Kuaishou Technology, Beijing 100085, China

[3]State Key Laboratory for Turbulence and Complex Systems, Department of Advanced Manufacturing and Robotics,

College of Engineering, Peking University, Beijing 100871, China

[4]School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China

[†]E-mail: junzhi.yu@ia.ac.cn

Abstract: Underwater robotic operation usually requires visual perception (e.g., object detection and tracking), but underwater scenes have poor visual quality and represent a special domain which can affect the accuracy of visual perception. In addition, detection continuity and stability are important for robotic perception, but the commonly used static accuracy based evaluation (i.e., average precision) is insufficient to reflect detector performance across time. In response to these two problems, we present a design for a novel robotic visual perception framework. First, we generally investigate the relationship between a quality-diverse data domain and visual restoration in detection performance. As a result, although domain quality has an ignorable effect on within-domain detection accuracy, visual restoration is beneficial to detection in real sea scenarios by reducing the domain shift. Moreover, non-reference assessments are proposed for detection continuity and stability based on object tracklets. Further, online tracklet refinement is developed to improve the temporal performance of detectors. Finally, combined with visual restoration, an accurate and stable underwater robotic visual perception framework is established. Small-overlap suppression is proposed to extend video object detection (VID) methods to a single-object tracking task, leading to the flexibility to switch between detection and tracking. Extensive experiments were conducted on the ImageNet VID dataset and real-world robotic tasks to verify the correctness of our analysis and the superiority of our proposed approaches. The codes are available at https://github.com/yrqs/VisPerception.

1 Introduction

Within the last few years, great efforts have been made in underwater robotics (Gong et al., 2018; Li B et al., 2018; Zhu DQ et al., 2019; Cai et al., 2020). For example, Gong et al. (2018) designed a soft robotic arm for underwater operation. Cai et al. (2020) developed a hybrid-driven underwater vehicle-manipulator system for collecting marine products. Concerning intelligent autonomous

---