

水下光学图像中目标探测关键技术研究综述

林森^{1,3,4}, 赵颖^{1,2*}

¹ 辽宁工程技术大学电子与信息工程学院, 辽宁 葫芦岛 125105;

² 辽宁工程技术大学研究生学院, 辽宁 葫芦岛 125105;

³ 中国科学院沈阳自动化研究所机器人学国家重点实验室, 辽宁 沈阳 110016;

⁴ 中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110016

摘要 对水下图像预处理和目标检测识别、跟踪关键技术进行了详细的归纳总结。根据是否需要构建模型,将水下图像预处理分为图像增强和图像复原,论述了增强和复原方法的基本思想及方法特点。对水下光学图像的目标检测识别与跟踪原理和方法的研究进展进行全面阐述。通过对以上各个过程研究成果的总结分析,梳理了水下目标探测各关键技术中存在的重点解决问题与相关难点问题,讨论了其解决思路和进一步发展方向。

关键词 图像处理; 数字图像处理; 水下图像预处理; 水下目标检测识别; 水下目标跟踪; 深度学习

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP57.060002

Review on Key Technologies of Target Exploration in Underwater Optical Images

Lin Sen^{1,3,4}, Zhao Ying^{1,2*}

¹ School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China;

² Institute of Graduate, Liaoning Technical University, Huludao, Liaoning 125105, China;

³ State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

⁴ Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China

Abstract We summarize key technologies of underwater target exploration in more details, including underwater image preprocessing, target detection and recognition, and target tracking algorithms. First, underwater image preprocessing algorithms are divided into image enhancement and restoration by judging whether they need to build a model, and then we discuss the basic ideas and characteristics of enhancement and restoration. Besides, the principles and methods of target detection and recognition, and target tracking for underwater optical images are reviewed across the board. After summarizing and analyzing the research results of the above processes, we untangle the core issues to be addressed and related difficulties existing in key technologies of underwater target exploration, and discuss the solutions and further development directions as well.

Key words image processing; digital image processing; underwater image preprocessing; underwater target detection and recognition; underwater target tracking; deep learning

OCIS codes 100.2000; 100.2980; 100.3008; 010.7340

收稿日期: 2019-07-09; 修回日期: 2019-08-07; 录用日期: 2019-08-28

基金项目: 国家自然科学基金(61473280,61773367)、国家自然科学基金重大研究计划(91648118)、辽宁省教育厅科学研究一般项目(L2014132)、辽宁省自然科学基金面上项目(2015020100)

* E-mail: 1460419171@qq.com

1 引言

近年来,海洋信息处理技术蓬勃发展,水下目标探测技术的应用也日益广泛,涉及海底光缆的铺设、水下石油平台的建立与维修、海底沉船的打捞、海洋生态系统的研究等^[1]领域。水下光学图像分辨率较高,信息量较为丰富,在短距离的水下目标探测任务中具有突出优势。然而,由于受水下特殊成像环境的限制,水下图像往往存在噪声干扰多、纹理特征模糊、对比度低及颜色失真等诸多问题。因此,水下目标探测任务面临诸多挑战,如何在图像可视性较差的情况下,精确、快速、稳定地检测识别和跟踪水下目标物体是亟待解决的问题。

根据水下目标探测任务的执行步骤,将基于光学图像的水下目标探测关键技术分为图像预处理和目标探测两部分。其中,水下目标探测特指水下目标检测、识别与跟踪。近年来,国内外研究人员对基于光学图像的水下目标探测关键技术进行了大量研究,水下目标探测技术取得了迅速发展,一些研究人员总结了关键技术的发展现状。Sahu 等^[2]总结了一系列水下图像增强算法,Han 等^[3]对水下图像智能去雾和色彩还原算法进行了综述,Kaeli 等^[4]概述了一组用于水下图像颜色校正改进的算法,郭继昌等^[5]对水下图像增强和复原算法进行了系统归纳并通过实验对比了不同算法,Moniruzzaman 等^[6]梳理了近年来深度学习在水下图像分析中的应用。然而,这些综述仅总结了水下目标探测某一关键技术的研究成果,目前仍缺少对水下目标探测关键技术的系统概述。本文从水下图像预处理和水下

目标检测、识别、跟踪技术入手,详细归纳了水下目标探测关键技术的研究现状。根据是否需要构建模型,将水下图像预处理分为图像增强和图像复原,并重点分析了水下图像增强的各类方法(基于直方图处理、基于 Retinex 理论、基于图像融合和基于深度学习的方法)的优缺点。由于水下目标跟踪技术的相关研究论文较少,本文主要从传统方法和深度学习两个角度讨论了水下目标检测与识别相关算法,并简要介绍了常用的水下图像数据集。在上述基础上指出了水下光学图像中的目标探测技术亟待解决的问题,讨论了解决思路和进一步发展方向。

2 水下图像预处理

与大气光学成像技术相比,水下光学成像技术深受水体光吸收和散射的影响,可见光在水体中传播的波长依赖性使得水下图像呈现蓝绿色调,水体中的杂质微粒对光的散射导致图像细节模糊以及表面雾化。为解决上述问题,研究人员提出了大量水下图像预处理算法,分为基于非物理模型的图像增强方法和基于物理模型的复原方法。

2.1 水下图像增强

图像增强不需要构建模型,通过主观处理消除水下图像的畸变和噪声,加强感兴趣的物体特征,削弱不相关的背景特征。通过对现有水下图像增强方法的调研得出常用的水下图像增强方法有基于直方图处理、基于 Retinex 理论、基于图像融合和基于深度学习的水下图像增强,如表 1 所示。

表 1 水下图像增强方法对比

Table 1 Comparison of underwater image enhancement methods

Method	Principle	Characteristic	Limitation
Histogram-based	Improving image contrast through nonlinear stretching of pixel values	Bright color, real-time	Lack of texture details, increasing noise
Retinex-based	Removing the influence of illuminating light from the image to obtain the reflective properties of the object	Bright color, high contrast, more detail information	Over-enhancement for the brighter patch, unsaturation for darker region
Fusion-based	Combining relevant information from two or more images into a single image, which is more informative than any of the inputs	Effective color correction, high contrast	Less robust, affected by artificial illumination
Deep learning-based	Imitating the working of the human brain in processing data, and improving the quality of underwater image through network training	Effective color correction, robust contrast stretching	Hard to get ground-truth images, time-consuming for training

2.1.1 基于直方图处理的水下图像增强

高对比度的图像通常呈现灰度细节丰富且动态范围较大的特点,基于直方图处理的水下图像增强方法通过改善图像像素值的分布范围提高对比度,如直方图均衡化方法和直方图拉伸方法。直方图均衡化方法利用概率密度函数和分布函数拉伸和均衡直方图,对比度受限的自适应直方图均衡化(CLAHE)算法^[7]是较为成熟的直方图均衡化算法,能够较好地改善水下图像的对比度,但会引入噪声。Qiao等^[8]将CLAHE算法与小波域滤波相结合,最大限度地消除了图像处理过程中产生的噪声和图像自身带有的噪声。直方图拉伸方法利用变换函数将像素值从一个区间变换到另一个区间,与直方图均衡化方法相比,计算复杂度较低。考虑到水下图像存在色偏现象,许多研究人员在不同的颜色空间(RGB、HSV、Lab)中对水下图像进行直方图拉伸^[9-11],有效地提高了水下图像的对比度,但盲目地使用全局直方图拉伸容易使图像丢失细节信息,且出现过度增强现象。

2.1.2 基于Retinex理论的水下图像增强

人眼接收到的图像由照射光图像和物体的反射属性组成^[12]。Retinex图像增强是从图像中去除照射光的影响获得物体的反射属性,通常采用中心环绕函数与原图像卷积的方式估计照射光图像,高斯函数作为中心环绕函数具有明显的优势。但单个标准差的选取无法同时保证图像的动态范围压缩和亮度一致性,多尺度Retinex(MSR)选取多个标准差,并通过加权叠加实现动态范围压缩和亮度一致性之间的权衡。有学者将双边滤波、三边滤波与高斯核相结合缓解边缘过度平滑^[13],或利用多尺度同态滤波代替高斯滤波估计照射光^[14],但卷积运算的计算成本随着卷积次数的增加而显著增加,频繁的卷积操作会严重影响处理速度,改进的MSR颜色保护(IMSRCP)^[15]算法采用多次下采样和无限脉冲响应高斯滤波提高了处理速度。

2.1.3 基于融合的水下图像增强

多尺度融合水下图像增强,即利用不同的融合权重提取输入图像的对比度、颜色、细节纹理等信息,并将其融合到输出图像中。针对水下图像存在色偏和对比度低的问题,输入图像常为水下退化图像颜色校正和对比度增强后的图像。然而深水环境中水体对光的波长选择性衰减使得直接应用颜色校正算法易出现红色伪影^[16],有学者根据水下可见光的衰减特性采用光补偿的方式校正色偏,提升图像

的对比度并补偿图像细节信息^[17-18],有效地解决了水下图像颜色失真、对比度低的问题,但处理后的图像存在局部亮度偏大的问题。多尺度融合过程中融合权重的设计也至关重要,具有高权重值的像素在输出图像中更具代表性。常用的权重如对比度权重、饱和度权重、显著性权重、曝光度权重及亮度权重等,考虑了如何提取纹理和边缘信息,区分平滑区域和变化快的区域,每个像素及其邻域像素之间的关系,如何在突出显著性较强区域的同时保证与非显著性区域间差值不能过大等。权重的设计并不是对所有水下图像都有较好的结果,这取决于水下图像的退化类型、时间复杂度和期望的结果等因素。

2.1.4 基于深度学习的水下图像增强

随着深度学习网络在多个领域的成功应用,基于深度学习的水下图像增强方法也逐渐取得了国内外研究人员的关注。目前,学者们主要将有监督的深度学习应用于水下图像增强,这需要大量的训练集,包括水下退化图像及其对应的真值图像。由于受特殊成像环境的限制,水下真值图像很难获得,若采用传统算法处理后的清晰图像作为真值图像,难以突破传统算法处理效果的限制。有学者从多种传统算法处理结果中选取最佳图像作为真值图像^[19-20],或通过估计水下成像的随机参数合成水下图像^[21-22],但合成的水下图像存在假设不准确、场景单一等问题,与真实水下图像相差甚远^[23],由合成水下图像训练的深度网络在真实水下环境中的图像增强能力也远不如人意。

用于水下图像增强的深度网络主要有生成对抗网络(GAN)(UGAN^[22]、FGAN等)和卷积神经网络(CNN)(UIENet、DUIENet等)^[23]。GAN的核心思想是通过使用对抗训练过程,生成一个足以以假乱真的生成图像替代真实图像。CNN方法能够更好地提高图像的峰值信噪比,GAN方法在图像颜色校正方面具有更好的效果。但这些方法的鲁棒性和泛化能力与传统的最先进的算法相比仍有一定差距。

2.2 水下图像复原

基于增强的方法是在特定意义上提高图像质量,可有效地提高图像的视觉效果,但由于没有考虑物理退化原理,忽略了退化程度与场景深度的关系,增强结果不能正确反映场景的真实色彩特征。相反,基于复原的方法使用构建的水下成像模型逆转退化过程,需估计未知参数,其难点在于如何保证在降低计算模型复杂度的前提下准确估计参数,使恢

复后的图像更接近真实场景。

有学者采用大气散射模型描述水下图像的退化过程,但未考虑水下光衰减的波长依赖性,其研究仅适用于清澈的浅水域。简化的水下成像模型中融入了水体吸收光的波长选择特性,受多数学者的青睐,但并未考虑到衰减系数随环境因素的变化;修正的水下成像模型在简化模型的基础上进一步完善,但其复杂度较高,不能得到广泛应用^[23]。

常用的场景系数估计方法有基于改进暗通道先验理论(DCP)的方法和基于深度学习的方法,如表2所示。基于改进DCP的方法多数基于特定的

假设,在特定环境中能很好地校正色偏和去雾,但一些假设忽略了前向散射的影响,并不适用于所有的水下环境,鲁棒性不强,且易受非均匀光照的影响。基于深度学习的方法,如用于学习透射率图的水下残卷积神经网络(URCNN)、水下图像复原网络(UIRNet)及水下图像实时颜色校正的无监督生成对抗网络(WaterGAN)等^[24],通过有效的网络训练在一定程度上降低了基于DCP假设方法的估计误差,但在水质较为浑浊的情况下,URCNN出现曝光过度的现象。此外,网络越深,参数越多越复杂,给网络训练带来一定的影响。

表2 水下场景系数估计方法对比分析

Table 2 Comparison of underwater scene coefficient estimation methods

Method	Improvement	Characteristic
DCP-based	Tang <i>et al.</i> ^[25] : the depth-of-field map is obtained by disparity between bright and dark channels, so as to estimate the background color of water body and transmittance map more accurately	High real-time performance, more authentic colors, influence of artificial light sources, and inadequate robustness and adaptability
	Xu <i>et al.</i> ^[26] : estimation of transmittance map of images by red channel algorithms	
	Xie <i>et al.</i> ^[27] : using the relationship between the wavelength of visible light and the scattering coefficient to obtain transmittance maps of each color channel	
Deep learning-based ^[23]	URCNN: a convolutional layer and ReLU are used to generate feature maps, then the batch normalization is added between convolutional layer and ReLU to speed up training process.	More authentic colors, robustness, and being time-consuming for training
	This pattern is repeated until the transmission map is output	
	UIRNet: the transmission map and the background light are estimated and computed by BL-Net and TM-Net, respectively	
	WaterGAN ^[24] : taking in-air images and depth maps as input and generating corresponding synthetic underwater images as output	

3 水下目标探测

无论是对水下坝基裂缝的检测、水下管道电缆的铺设与检修等工程应用领域,还是对鱼类等水生生物种群的海洋生态监测领域,水下目标探测都具有重要的研究意义。本文的水下目标探测特指水下目标检测、识别与跟踪。从概念上说,水下目标检测是在给定目标物体的前提下,根据其特征和先验知识确认水下图像中是否存在该目标物体;水下目标识别是在给定一幅图像的情况下,通过提取图像中的特征信息,并将其进行分类,进而识别图像中的目标物体;水下目标跟踪是在目标检测的基础上对目标的状态进行连续估计的过程。检测和识别间并没

有非常分明的界限,多数文献也将检测和识别的概念混在一起,因此,本文对水下目标检测和识别并不加以严格区分。

3.1 水下目标检测与识别

近年来,国内外研究机构及学者对水下目标检测与识别算法进行了大量研究,水下目标检测与识别技术发展迅速,根据目标类型可将其分为基于传统特征的水下目标检测识别和基于深度学习网络的水下目标检测识别。

3.1.1 基于传统特征的水下目标检测与识别

水下机器人在执行工程任务时需检测识别的目标物体多数为管道、电缆等具有规则形状的人造目标物体,特征明显。基于传统特征的水下目标检测

与识别方法通过特征描述符刻画水下目标物体,具有方法简单、实时性好的优点,但也受目标物体尺寸变化、旋转、被遮挡和拍摄角度变化、物种类别变化

的影响。常用的水下图像特征有颜色特征、形状特征、纹理特征等,图 1 归纳了传统的水下图像特征提取的主要方法。

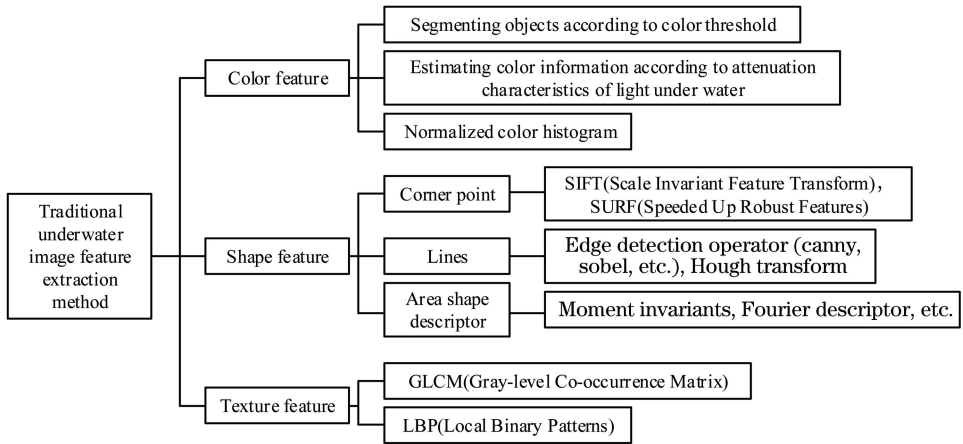


图 1 传统的水下图像特征提取的主要方法

Fig. 1 Traditional methods of underwater image feature extraction

1) 颜色特征

虽然水下光学图像存在色偏现象,使得颜色信息在水下目标探测中应用时,存在一定的局限性,但颜色本身固有的对尺度变化、目标位置变化及局部遮挡的鲁棒性,使得颜色成为一个简单且非常可靠的信息特征。由于 HSV 颜色空间可分离图像的颜色信息和光强度,颜色特征的提取通常在 HSV 颜色空间中进行,通过设置特定的颜色阈值分割目标和背景环境^[28-29],但易受相同颜色的背景物体的干扰。有学者结合水下颜色恒常性问题,检测给定物体颜色的相容色^[30],直接对原始数据进行处理,所需计算机资源少,实时性好,且对局部遮挡、目标位置和尺度变化具有很强的鲁棒性,但未能考虑水下较暗的颜色和光学成像模型中的后向散射分量。此外,颜色直方图也是采用颜色特征进行目标检测的一种常用方法,但直方图基于统计特性,缺少空间分布信息。

2) 形状特征

形状特征具有较好的鲁棒性,在水下环境中抗干扰能力强,且典型的水下人造目标物体,如管道、电缆等,均具有规则的形状,因此常通过形状特征提取检测识别水下目标。形状特征的提取方法主要有边缘检测算子、霍夫(Hough)变换、形状特征描述符法等。Hough 变换是检测直线的最好工具之一, Lin 等^[31]将 Hough 变换用于识别物体的线性或中心特征, Fatan 等^[32]采用 Hough 变换检测水下电缆和管道。虽然 Hough 变换可检测直线,但不能提供线段长度等完整性信息,且计算量大,耗费存储空

间。尺度不变特征转换(SIFT)是较为可靠的特征描述符,包括尺度空间的极值检测、关键点定位、方向分配、生成关键点描述符四个步骤。加速稳健特征(SURF)对 SIFT 进行了优化,采用 box 滤波器替代计算高斯拉普拉斯函数的过程来检测感兴趣点,并增加了两个新的概念减少运算时间,描述符向量的长度也减少了一半。

3) 纹理特征

纹理特征是一种全局特征,包含信息多且复杂,描述图像或其中小块区域的空间颜色分布和光强分布,通过像素及其周围空间邻域的灰度分布来表现,抗噪能力强,具有旋转不变性,不会因局部偏差导致无法成功匹配。灰度共生矩阵(GLCM)^[33]常用来描述图像的纹理特征, Li 等^[34]同时利用 GLCM 和分形维数描述图像视觉纹理特征, Cheng 等^[35]采用 GLCM 提取水下模糊图像的纹理细节信息捕捉图像的模糊度。虽然 GLCM 能够很好地提取图像的纹理特征,但也存在一些缺点,如计算量大,耗时较长,且有时会对不同结构模式产生相同的二进制码。局部二进制模式(LBP)根据局部结构特征对纹理进行结构描述和统计描述,具有计算简单、效率高、纹理识别性能好等特点,但 LBP 算子的不足在于当邻域数增加时,特征向量的维数变得很高, $Z \oplus T Z L B P$ (Z with Tilted Z LBP)^[36]通过减少特征向量的维数降低了 LBP 的计算复杂度,成功用于珊瑚图像的分类中。

4) 多特征结合

单个特征难以完全描述目标物体的特性,一些

研究人员融合多个特征检测识别水下目标物体,如 Huang 等^[37]从鱼的不同部位提取 66 种特征,并采用正向序列特征选择方法降低特征维数;Tharwat 等^[38]采用韦伯局部描述符和颜色矩分别提取了鱼类图像的纹理和颜色特征;Li 等^[39]依据水下图像的颜色和轮廓特性构建块特征描述符;Rizzini 等^[40]根据水下图像的颜色均匀性和轮廓的锐度搜索目标物体。多特征结合的方法能够更加准确地描述水下目标物体,检测成功率普遍有了提升,但也存在计算复杂度增加的不足。

传统的水下目标检测与识别方法预先指定了目标物体的特征,这些特征采用常用的低层图像特征描述符,能够很好地识别简单目标物体。然而在监测鱼类等水生生物种群信息时,由于鱼类的巨大多样性,为特定物种设计的一组特征并不能保证对其他物种也具有有良好的适应性。此外,人工选择的特征也可能导致物体的识别性能不佳。

3.1.2 基于深度学习的水下目标检测与识别

在 2012 年的 ImageNet 大尺度视觉识别挑战

表 3 深度神经网络用于水下目标检测与识别的对比分析

Table 3 Comparison of depth neural network for underwater target detection and recognition

Author	Method	Advantage and limitation
Salman <i>et al.</i> ^[43]	A hybrid approach involving GMM, optical flow, and deep R-CNN to fine-tune the categorization of fish	Higher classification accuracy; requirement on relatively more computational resources
Siddiqui <i>et al.</i> ^[44]	A cross-layer pooling algorithm using a pre-trained convolutional neural network as a generalized feature detector	No need for a large amount of training data; high classification accuracy; requirement of extensive computations
Sun <i>et al.</i> ^[45]	A CNN knowledge transfer framework for underwater object recognition and extracting discriminative features from relatively low contrast images	High real-time performance; no need for a large amount of training data; lower robustness
Chuang <i>et al.</i> ^[46]	An underwater fish recognition framework that consists of a fully unsupervised feature learning technique and an error-resilient classifier	Successfully handled data uncertainty and class imbalance in practical classification applications; lower robustness
Cao <i>et al.</i> ^[42]	A method to combine CNN and hand-designed features to improve classification performance	Being robust, reducing feature dimensionality without decreasing classification performance; contour extraction needing original video frames

随后,CNN 出现了很多变体,如 R-CNN、Fast R-CNN、YOLO (You Only Look Once) CNN、YOLOv2 CNN、YOLOv3 CNN 等。Li 等^[47]将高精度的 Fast R-CNN 应用于复杂的水下环境中进行鱼类的检测与识别,与可变形部件模型(DPM)相比,Fast R-CNN 的平均精度(mAP)提高了 11.2%,

赛中,CNN 方法显示出更高的图像分类精度。随着 GPU 等硬件系统的发展日趋成熟,深度学习面临的数据量异常庞大,训练时间冗长的问题得到了有效解决,基于深度学习的目标检测与识别技术发展迅速。

1) 水下目标检测识别算法

许多机器视觉任务中,深度学习网络从像素级原始数据到抽象的语义概念逐层提取信息,使其在提取图像的全局特征和上下文信息方面具有突出的优势,与传统方法相比,表现更好。研究人员通过实验对比了深度学习网络 and 传统分类算法的性能^[41],发现传统方法分类的准确率明显更低。为了进一步提高深度网络在水下特殊环境中的目标检测识别精度,许多学者将深度神经网络与其他算法相结合,如表 3 所示,充分发挥了各算法的优势,在水下目标检测时降低了大量训练数据的需求,提高了检测精度及适应性。Cao 等^[42]将人工提取的特征与 CNN 提取的特征相结合,也达到了比单独应用 CNN 更高的精度。

达到了 81.4%,并且在单个鱼图像上检测速度比 R-CNN 快 80 倍。Mandal 等^[48]将 Faster R-CNN 与三种分类网络相结合,用于鱼类物种的检测和识别。Lu 等^[49]所提出的深度卷积滤波网络(FDCNet)的分类精度比一些经典的分类方法,如 AlexNet、GoogLeNet、ResNet50 和 ResNet100,分别提高了

1.8%, 2.9%, 2.0%和 1.0%。Christensen 等^[50]提出了一种基于光学的鱼类检测网络 OFDNet(Optical Fish Detection Network), 利用水下摄像机获取的视觉数据进行鱼类检测、定位和物种分类。针对水下视频鱼类检测的实时性问题, Sung 等^[51]提出了一种基于 YOLO 的卷积神经网络, 使用真实环境下的鱼类视频图像对该方法的可靠性和准确性进行评价, 分类准确率为 93%。Pedersen 等^[52]对 YOLOv2 和 YOLOv3 CNN 进行了调整, 并用于水下目标检测。通过端到端训练的学习任务, 深度学习可实现多层非线性变换, 将底层特征有效地组合成一个更加抽象的高层表示属性, 当存在目标遮挡, 目标尺寸较小时也能成功检测, 但由于深度神经网络结构复杂, 这使其应用到特定环境中时需调整大量参数。

2) 用于水下目标检测与识别的水下图像数据集

有效的训练集数据是提高基于深度学习的水下目标检测与识别精度的关键。通常, 用于水下目标检测与识别的水下图像数据集应包含以下特点: 水下场景环境的多样性, 不同程度和不同类型的图像

质量退化, 以及目标种类、目标物体姿态的多样性等。fish4knowledge 数据集^[53]以复杂的珊瑚礁环境为背景, 包含丰富的运动生物, 如游动的鱼类、海藻、生长的藻类、漂动的水生植物等, 是应用最为广泛的水下图像数据集。Brackish 数据集^[52]包含了不同能见度的咸水中的鱼、蟹和海星图像序列及其对应带注释的图像序列。同时, 文献^[52]中也详细总结了用于鱼类、底栖生物检测和海洋环境监测的 J-EDI、Croatian Fish Dataset、QUT Fish Dataset、HabCam、Benthos-15、Tasmania Coral Point Count、The Moorea Labeled Corals、Eilat Fluorescence、MOUSS、AFSC、MBARI 和 NWFSC 数据集。此外, 还有用于场景识别和目标检测的 SUN Dataset^[54]中的水下图像, 用于海洋自主机器人的 MARIS Dataset^[55]等。

3.2 水下目标跟踪

水下目标跟踪过程中会存在各种问题, 如目标形变、目标遮挡、目标与背景中相似的物体特征、阴影问题、光照的变化、背景景物的运动和跟踪实时性与准确性的平衡问题等。常用的水下目标跟踪方法如表 4 所示。

表 4 水下目标跟踪方法对比

Table 4 Comparison of underwater target tracking methods

Method	Improvement	Principle	Characteristic
Optical flow-based	Modified Lucas-Kanade optical flow method based on pyramid hierarchy and affine transformation	Estimating the position by calculating the velocity in two consecutive frames	Need of a large number of precise image feature points
Mean shift-based	1) Adaptive mean shift algorithm through color histogram based on background and target region; 2) adaptive mean shift algorithm combined with edge information; 3) combined with color, texture, HOG features, and deformable multi-core algorithm	Using the color histogram of the target as the search feature, and iterating the mean shift vector continuously	Being robust to range variations; being influenced by multi-objects
CNT-based	1) Fast-CNT: improving the computing performance by selecting adaptive K value and omitting background filter; 2) improved Fast-CNT: extracting each region containing a moving target through Gaussian mixture model	Tracking target through two-layer forward convolution networks learning image features	No need for off-line training which is time-consuming and requires lots of training data

光流法和均值偏移法常被用于基于视觉的跟踪过程中。光流法通过计算两个连续帧中的流速估计

图像点的位置, 需要大量精确的图像特征点, 当图像存在模糊问题时, 图像特征点提取困难, 进而影响跟

踪效果。均值偏移法能够非常有效地跟踪非刚性物体,对距离变化具有鲁棒性,但图像中存在两个互相重叠的物体时,无法跟踪到目标。Wang 等^[56]采用自适应均值偏移算法跟踪鱼类,通过对尺度的估计,利用背景和目标区域的颜色直方图,处理鱼姿态的变化和光照不均匀性,很好地克服了上述不足。Zheng 等^[57]提出了一种自适应的基于边缘的均值偏移(MS)跟踪算法,该算法可有效地跟踪灰度变化目标,但严重的漂移会导致目标区域的扩展和目标位置的不精确。Chuang 等^[58]采用基于颜色、纹理特征和 HOG 特征的均值漂移算法有效地估计核运动,然后与 DPM 相结合,提出了基于可变形多核的跟踪算法,不仅利用了基于内核的跟踪的低计算成本,而且利用了 DPM 检测目标定位的鲁棒性。

近年来,深度学习的兴起为视觉跟踪领域提供了新的方向。然而,超长的离线训练时间和深度学习所需的大量图像极大地阻碍了学习的进展,CNT(Convolutional Network Based Tracker)^[59]不需要大量辅助数据进行离线训练,简单的两层前向卷积网络就足以学习图像特征来进行鲁棒的视觉跟踪。CNT 的第一层卷积网络由一组滤波器定义的简单单元特征图构成,其中每个滤波器用 K-means 算法从第一帧中提取归一化图像块;第二层将简单单元特征图叠加成复杂单元特征图作为目标表示,对目标的局部结构和几何布局信息进行编码,采用软收缩策略对目标表示进行降噪。此外,CNT 还采用一种有效的在线方案更新目标表示,以适应跟踪过程中目标外观的变化,但 CNT 的不足在于跟踪目标时存在漂移现象。Fast-CNT 算法采用自适应 K 值代替 K-means 算法中的固定 K 值,省略了 CNT 中使用的背景滤波器,在不影响性能的前提下节省了计算时间,平均计算性能提高了 2~10 倍,SURF 特征点与粒子滤波的结合解决了 CNT 中存在的漂移问题^[60]。Huang 等^[61]在跟踪前采用 GMM(Gaussian Mixture Model)提取每个包含运动目标(鱼)的区域,并识别包含鱼的相应区域,再采用 Fast-CNT 网络进行水下多目标跟踪,即使鱼的姿态有很大的变化,该方法仍然能够成功跟踪多条鱼,但一条鱼被另一条完全遮挡时,会出现跟踪失败的情况。

4 结束语

水下目标探测关键技术海洋信息处理中具有很大的应用前景和发展潜力,受到越来越多学者的

关注,经过几十年的发展取得了许多成果,但目前对水下目标探测关键技术的研究现状仍缺乏统一归纳。对近年来国内外水下图像预处理、水下目标检测识别与跟踪的研究进展进行了调研总结,为科研人员提供了参考和依据。虽然基于光学图像的水下目标探测关键技术取得了许多研究成果,但仍存在一些问题需要进一步探讨,可以从以下几个方面入手。

1) 提高水下图像预处理方法的有效性、实时性以及鲁棒性、自适应性。

目前针对水下图像增强、复原技术的研究较多,且大部分算法是解决色偏、对比度低、非均匀光照的问题,与此同时还要兼顾算法的有效性、实时性以及鲁棒性、自适应性,在这些方面还有许多工作要做。目前的预处理算法几乎都仅作用于单幅图像,而水下目标探测任务中视频的作用较大,针对水下视频预处理算法,帧间一致性的问题仍需考虑^[5]。

2) 深度学习在水下目标探测各关键技术中的应用。

深度学习的动机在于建立、模拟人脑的机制进行分析学习并处理数据。基于深度学习的水下目标探测关键技术可以显著提高算法的性能,但该算法存在可解释性差的问题,应用到目标探测各关键技术中时也有一些问题。首先,现有的水下图像预处理方法很少采用深度学习算法,原因在于采用深度学习对水下图像预处理时需要大量的真值图像,由于受特殊成像环境的限制,水下真值图像很难获得,若采用传统算法处理后的清晰图像作为真值图像,又难以突破传统算法处理效果的限制。其次,将深度学习应用于水下目标探测算法中时,也存在训练样本繁多、训练过程漫长、网络结构复杂、在普通设备上的应用受到限制等问题。如何解决这些问题对以后的研究提出了很高的要求。此外,由于非监督深度网络无需真值图像,可以完美解决水下真值图像难以获得的问题,但目前还没有基于非监督的深度神经网络应用到水下目标探测关键技术中的相关研究。

3) 仿照水下生物视觉的信息处理模式进行水下目标探测。

在水下图像中检测、识别、跟踪目标物体时都会进行图像预处理以提高成功率,这势必会增加处理时间,在实时性要求较高的应用中很难满足要求。水下生物能够在复杂多变的水下环境生存,它们拥有各自特殊的视觉系统,仿照水下生物视觉的信息

处理模式进行水下目标检测、识别与跟踪也是一个可以考虑的目标探测研究方向。

4) 将陆上目标探测算法迁移到水下环境中。

由于受水下特殊光学环境的限制,陆上目标探测算法的研究遥遥领先于水下目标探测算法。目前,研究较为先进的陆上目标检测算法有 RetinaNet、Mask R-CNN、RefineDet、M2Det 等^[62],克服水下环境的特殊限制,将这些算法迁移到水下环境中,对推动水下目标探测技术的发展具有重要意义。

参 考 文 献

- [1] Guo G, Wang X K, Xu H P. Review on underwater target detection, recognition and tracking based on sonar image[J]. Control and Decision, 2018, 33(5): 906-922.
郭戈, 王兴凯, 徐慧朴. 基于声呐图像的水下目标检测、识别与跟踪研究综述[J]. 控制与决策, 2018, 33(5): 906-922.
- [2] Sahu P, Gupta N, Sharma N. A survey on underwater image enhancement techniques [J]. International Journal of Computer Applications, 2014, 87(13): 19-23.
- [3] Han M, Lyu Z Y, Qiu T, et al. A review on intelligence dehazing and color restoration for underwater images [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018: 1-13.
- [4] Kaeli J W, Singh H, Murphy C, et al. Improving color correction for underwater image surveys[C] // OCEANS'11 MTS/IEEE KONA, September 19-22, 2011, Waikoloa, HI, USA. New York: IEEE, 2011: 12470441.
- [5] Guo J C, Li C Y, Guo C L, et al. Research progress of underwater image enhancement and restoration methods[J]. Journal of Image and Graphics, 2017, 22(3): 273-287.
郭继昌, 李重仪, 郭春乐, 等. 水下图像增强和复原方法研究进展[J]. 中国图象图形学报, 2017, 22(3): 273-287.
- [6] Moniruzzaman M, Islam S M S, Bennamoun M, et al. Deep learning on underwater marine object detection: a survey[M] // Blanc-Talon J, Penne R, Philips W, et al. Advanced concepts for intelligent vision systems. Lecture notes in computer science. Cham: Springer, 2017, 10617: 150-160.
- [7] Reza A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement [J]. The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology, 2004, 38(1): 35-44.
- [8] Qiao X, Bao J H, Zhang H, et al. Underwater image quality enhancement of sea cucumbers based on improved histogram equalization and wavelet transform[J]. Information Processing in Agriculture, 2017, 4(3): 206-213.
- [9] Ghani A S A, Isa N A M. Automatic system for improving underwater image contrast and color through recursive adaptive histogram modification [J]. Computers and Electronics in Agriculture, 2017, 141: 181-195.
- [10] Ao J, Ma C B. Adaptive stretching method for underwater image color correction[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2018, 32(2): 1854001.
- [11] Huang D M, Wang Y, Song W, et al. Underwater image enhancement method using adaptive histogram stretching in different color models [J]. Journal of Image and Graphics, 2018, 23(5): 640-651.
黄冬梅, 王龔, 宋巍, 等. 不同颜色模型下自适应直方图拉伸的水下图像增强[J]. 中国图象图形学报, 2018, 23(5): 640-651.
- [12] Fu Q. Underwater image processing of low level vision[D]. Harbin: Harbin Engineering University, 2015: 1-112.
付强. 水下图像低层视觉问题研究[D]. 哈尔滨: 哈尔滨工程大学, 2015: 1-112.
- [13] Zhang S, Wang T, Dong J Y, et al. Underwater image enhancement via extended multi-scale Retinex [J]. Neurocomputing, 2017, 245: 1-9.
- [14] Li Q W, Zhu W Q, Cao M. Underwater color image enhancement using improved multi-scale Retinex and histogram linear quantification [C] // Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing - ICC '17, March 22-23, 2017, Cambridge, United Kingdom. New York: ACM, 2017: 189.
- [15] Tang C, von Lukas U F, Vahl M, et al. Efficient underwater image and video enhancement based on Retinex[J]. Signal, Image and Video Processing, 2019, 13(5): 1011-1018.
- [16] Ancuti C, Ancuti C O, Haber T, et al. Enhancing underwater images and videos by fusion [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI,

- USA. New York: IEEE, 2012: 81-88.
- [17] Ancuti C O, Ancuti C, de Vleeschouwer C, et al. Color balance and fusion for underwater image enhancement [J]. *IEEE Transactions on Image Processing*, 2018, 27(1): 379-393.
- [18] Dai C G, Lin M X, Wang Z, et al. Color compensation based on bright channel and fusion for underwater image enhancement [J]. *Acta Optica Sinica*, 2018, 38(11): 1110003.
代成刚, 林明星, 王震, 等. 基于亮通道色彩补偿与融合的水下图像增强[J]. *光学学报*, 2018, 38(11): 1110003.
- [19] Li C Y, Guo C L, Ren W Q, et al. An underwater image enhancement benchmark dataset and beyond [EB/OL]. (2019-01-11) [2019-07-08]. <https://arxiv.org/abs/1901.05495>.
- [20] Chen X Y, Yu J Z, Kong S H, et al. Towards quality advancement of underwater machine vision with generative adversarial networks [EB/OL]. (2018-01-16) [2019-07-08]. <https://arxiv.org/abs/1712.00736>.
- [21] Hou M J, Liu R S, Fan X, et al. Joint residual learning for underwater image enhancement [C] // 2018 25th IEEE International Conference on Image Processing (ICIP), October 7-10, 2018, Athens, Greece. New York: IEEE, 2018: 4043-4047.
- [22] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks [EB/OL]. (2018-01-11) [2019-07-08]. <https://arxiv.org/abs/1801.04011>.
- [23] Anwar S, Li C. Diving deeper into underwater image enhancement: a survey [EB/OL]. (2019-07-17) [2019-08-30]. <https://arxiv.org/abs/1907.07863>.
- [24] Li J, Skinner K A, Eustice R M, et al. WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images [EB/OL]. (2017-10-26) [2019-07-08]. <https://arxiv.org/abs/1702.07392>.
- [25] Tang Z Q, Zhou B, Dai X Z, et al. Underwater robot visual enhancements based on the improved DCP algorithm [J]. *Robot*, 2018, 40(2): 222-230.
汤忠强, 周波, 戴先中, 等. 基于改进 DCP 算法的水下机器人视觉增强 [J]. *机器人*, 2018, 40(2): 222-230.
- [26] Xu Y, Zeng X B. Underwater image restoration based on red-dark channel prior and inverse filtering [J]. *Laser & Optoelectronics Progress*, 2018, 55(2): 021009.
徐岩, 曾祥波. 基于红色暗通道先验和逆滤波的水下图像复原 [J]. *激光与光电子学进展*, 2018, 55(2): 021009.
- [27] Xie H L, Peng G H, Wang F, et al. Underwater image restoration based on background light estimation and dark channel prior [J]. *Acta Optica Sinica*, 2018, 38(1): 0101002.
谢昊伶, 彭国华, 王凡, 等. 基于背景光估计与暗通道先验的水下图像复原 [J]. *光学学报*, 2018, 38(1): 0101002.
- [28] Yahya M F, Arshad M R. Tracking of multiple light sources using computer vision for underwater docking [J]. *Procedia Computer Science*, 2015, 76: 192-197.
- [29] Susanto T, Mardiyanto R, Purwanto D. Development of underwater object detection method base on color feature [C] // 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM), November 26-27, 2018, Surabaya, Indonesia. New York: IEEE, 2018: 254-259.
- [30] Bazeille S, Quidu I, Jaulin L. Color-based underwater object recognition using water light attenuation [J]. *Intelligent Service Robotics*, 2012, 5(2): 109-118.
- [31] Lin Y H, Chen S Y, Tsou C H. Development of an image processing module for autonomous underwater vehicles through integration of visual recognition with stereoscopic image reconstruction [J]. *Journal of Marine Science and Engineering*, 2019, 7(4): 107.
- [32] Fatan M, Daliri M R, Shahri A M. Underwater cable detection in the images using edge classification based on texture information [J]. *Measurement*, 2016, 91: 309-317.
- [33] Cheng K C, Cheng X M, Hao Q. A review of feature extraction technologies for plankton images [C] // Proceedings of the 2018 International Conference on Information Hiding and Image Processing-IHIP 2018, September 22-24, 2018, Manchester, United Kingdom. New York: ACM, 2018: 48-56.
- [34] Li Y, Liu S J, Zhu P Q, et al. Extraction of visual texture features of seabed sediments using an SVDD approach [J]. *Ocean Engineering*, 2017, 142: 501-506.
- [35] Cheng E, Lin X Z, Chen Y, et al. GLCM based no-reference perceptual blur metric for underwater blur image [J]. *International Journal of Circuits, Systems and Signal Processing*, 2016, 10: 291-296.
- [36] Ani Brown Mary N, Dejeu D. Classification of coral

- reef submarine images and videos using a novel Z with tilted Z local binary pattern ($Z \oplus$ TZLBP) [J]. *Wireless Personal Communications*, 2018, 98(3): 2427-2459.
- [37] Huang P X, Boom B J, Fisher R B. Underwater live fish recognition using a balance-guaranteed optimized tree[M] // Lee K M, Matsushita Y, Reh J M, et al. *Computer vision-ACCV 2012. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2013, 7724: 422-433.
- [38] Tharwat A, Hemedan A A, Hassanien A E, et al. A biometric-based model for fish species classification [J]. *Fisheries Research*, 2018, 204: 324-336.
- [39] Li J, Eustice R M, Johnson-Roberson M. High-level visual features for underwater place recognition[C] // 2015 IEEE International Conference on Robotics and Automation (ICRA), May 26-30, 2015, Seattle, WA, USA. New York: IEEE, 2015: 3652-3659.
- [40] Rizzini D L, Kallasi F, Oleari F, et al. Investigation of vision-based underwater object detection with multiple datasets [J]. *International Journal of Advanced Robotic Systems*, 2015, 12(6): 77.
- [41] Villon S, Chaumont M, Subsol G, et al. Coral reef fish detection and recognition in underwater videos by supervised machine learning: comparison between deep learning and HOG+SVM methods[M] // Blanc-Talon J, Distant C, Philips W, et al. *Advanced concepts for intelligent vision systems. Lecture notes in computer science*. Cham: Springer, 2016, 10016: 160-171.
- [42] Cao Z, Principe J C, Ouyang B, et al. Marine animal classification using combined CNN and hand-designed image features [C] // OCEANS 2015 - MTS/IEEE Washington, October 19-22, 2015, Washington, DC, USA. New York: IEEE, 2015: 15798739.
- [43] Salman A, Siddiqui S A, Shafait F, et al. Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system [J]. *ICES Journal of Marine Science*, 2019: fsz025.
- [44] Siddiqui S A, Salman A, Malik M I, et al. Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data [J]. *ICES Journal of Marine Science*, 2018, 75(1): 374-389.
- [45] Sun X, Shi J Y, Liu L P, et al. Transferring deep knowledge for object recognition in low-quality underwater videos[J]. *Neurocomputing*, 2018, 275: 897-908.
- [46] Chuang M C, Hwang J N, Williams K. A feature learning and object recognition framework for underwater fish images [J]. *IEEE Transactions on Image Processing*, 2016, 25(4): 1862-1872.
- [47] Li X, Shang M, Qin H W, et al. Fast accurate fish detection and recognition of underwater images with Fast R-CNN [C] // OCEANS 2015 - MTS/IEEE Washington, October 19-22, 2015, Washington, DC, USA. New York: IEEE, 2015: 15798821.
- [48] Mandal R, Connolly R M, Schlacherz T A, et al. Assessing fish abundance from underwater video using deep neural networks[EB/OL]. (2018-07-16) [2019-07-08]. <https://arxiv.org/abs/1807.05838>.
- [49] Lu H M, Li Y J, Uemura T, et al. FDCNet: filtering deep convolutional network for marine organism classification [J]. *Multimedia Tools and Applications*, 2018, 77(17): 21847-21860.
- [50] Christensen J H, Mogensen L V, Galeazzi R, et al. Detection, localization and classification of fish and fish species in poor conditions using convolutional neural networks[C] // 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), November 6-9, 2018, Porto, Portugal. New York: IEEE, 2018: 18740228.
- [51] Sung M, Yu S C, Girdhar Y. Vision based real-time fish detection using convolutional neural network[C] // OCEANS 2017 - Aberdeen, June 19-22, 2017, Aberdeen, United Kingdom. New York: IEEE, 2017: 17290648.
- [52] Pedersen M, Haurum J B, Gade R, et al. Detection of marine animals in a new underwater dataset with varying visibility [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 16-20, 2019, Long Beach, California, USA. New York: IEEE, 2019: 18-26.
- [53] Hsiao Y H, Chen C C, Lin S I, et al. Real-world underwater fish recognition and identification, using sparse representation [J]. *Ecological Informatics*, 2014, 23: 13-21.
- [54] Xiao J X, Hays J, Ehinger K A, et al. SUN database: large-scale scene recognition from abbey to zoo[C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 3485-3492.
- [55] Oleari F, Kallasi F, Rizzini D L, et al. An underwater stereo vision system: from design to

- deployment and dataset acquisition [C] // OCEANS 2015 - Genova, May 18-21, 2015, Genova, Italy. New York: IEEE, 2015: 15473467.
- [56] Wang J H, Zhao M, Zou L, et al. Fish tracking based on improved TLD algorithm in real-world underwater environment [J]. Marine Technology Society Journal, 2019, 53(3): 80-89.
- [57] Zheng H C, Mao X, Chen L J, et al. Adaptive edge-based mean shift for drastic change gray target tracking[J]. Optik, 2015, 126(23): 3859-3867.
- [58] Chuang M C, Hwang J N, Ye J H, et al. Underwater fish tracking for moving cameras based on deformable multiple kernels [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(9): 2467-2477.
- [59] Zhang K H, Liu Q S, Wu Y, et al. Robust visual tracking via convolutional networks without training [J]. IEEE Transactions on Image Processing, 2016, 25(4): 1779-1792.
- [60] Huang R J, Tsao C Y, Kuo Y P, et al. Fast visual tracking based on convolutional networks [J]. Sensors, 2018, 18(8): 2405.
- [61] Huang R J, Lai Y C, Tsao C Y, et al. Applying convolutional networks to underwater tracking without training [C] // 2018 IEEE International Conference on Applied System Invention (ICASI), April 13-17, 2018, Chiba, Japan. New York: IEEE, 2018: 342-345.
- [62] Liu L, Ouyang W, Wang X, et al. Deep learning for generic object detection: a survey [EB/OL]. (2019-08-22) [2019-08-30]. <https://arxiv.org/abs/1809.02165>.