

文章编号 1004-924X(2023)21-3212-09

基于无监督学习和注意力机制的 水下偏振图像融合

巩文哲, 褚金奎, 成昊远, 张 然*

(大连理工大学 辽宁省微纳米系统重点实验室, 辽宁 大连 116024)

摘要:针对光线在水中传播会受到吸收和散射影响,仅使用传统的强度相机获取水下图像存在成像结果亮度偏低、图像模糊、细节丢失等问题,将深度融合网络应用于水下偏振图像,即用深度学习的方法将水下偏振度图像与光强图像融合。分析水下主动偏振成像模型,搭建实验装置获取水下偏振图像构建训练数据集,并进行适当变换以扩充数据集。构建基于无监督学习和注意力机制引导的用于融合偏振度和光强图像的端到端学习框架,并对损失函数及权重参数进行阐述。使用制作的数据集在不同的损失权重参数下进行网络训练,基于不同的图像评价指标对融合后的图像进行评估。实验结果表明,融合后的水下图像细节更为丰富,相比于光强图像信息熵提升 24.48%,标准差提升 139%。相比于其他传统融合算法,该方法不需要人工调节权重参数,运算速度较快,具有较强的稳定性和自适应性,对于海洋探测、水下目标识别等领域的应用研究具有重要意义。

关键词:偏振成像;水下图像增强;图像融合;深度学习

中图分类号: TN29; TP391.4 **文献标识码:** A **doi:** 10.37188/OPE.20233121.3212

Underwater polarization image fusion based on unsupervised learning and attention mechanisms

GONG Wenzhe, CHU Jinkui, CHENG Haoyuan, ZHANG Ran*

(Key Laboratory for Micro/Nano Technology and System of Liaoning Province,
Dalian University of Technology, Dalian 116024, China)

* Corresponding author, E-mail: zhangr@dlut.edu.cn

Abstract: As light propagation in water is subject to absorption and scattering effects, acquiring underwater images using conventional intensity cameras can result in low brightness of imaging results, blurred images, and loss of details. In this study, a deep fusion network was applied to underwater polarimetric images; the underwater polarimetric images were fused with light-intensity images using deep learning. First, the underwater active polarization imaging model was analyzed, an experimental setup was built to obtain underwater polarization images to construct a training dataset, and appropriate transformations were performed to expand the dataset. Second, an end-to-end learning framework was constructed based on unsupervised learning and guided by attention mechanism for fusing polarimetric and light intensity images and the loss function and weight parameters were elaborated. Finally, the produced dataset was used to train

收稿日期: 2023-02-23; 修订日期: 2023-04-13.

基金项目: 国家自然科学基金资助项目 (No. 52175265, No. 52275281); 国家自然科学基金创新研究群体项目 (No. 51621064); 中央高校基本科研业务费资助项目 (No. DUT21ZD101, No. DUT21GF308)

the network under different loss weight parameters and the fused images were evaluated based on different image evaluation metrics. The experimental results show that the fused underwater images are more detailed, with 24.48% higher information entropy and 139% higher standard deviation than light-intensity images. Unlike other traditional fusion algorithms, the method does not require manual weight parameter adjustment, has faster operation speed, strong robustness, and self-adaptability, which is important for ocean detection and underwater target recognition.

Key words: polarization imaging; underwater image enhancement; image fusion; deep learning

1 引言

海洋面积占地球总面积的 70% 以上,海洋生态系统是地球上生产力最高、最具活力的生态系统之一,人们在海洋资源勘探、生物考察和水下航行器导航等领域展开了研究^[1-5]。水下光学图像是目前探索海洋的重要媒介之一,但是由于光线在水中受到大量漂浮微粒的影响,实际水下获得的图像严重退化,存在背景噪声大、对比度低和细节丢失等问题^[1]。因此,水下图像增强技术的研究对于海洋探测和获取水下的有效信息具有重要的意义和价值。

在众多水下成像技术中,水下偏振光学成像技术作为新型光学成像技术备受关注意^[6-9]。偏振是光的波动性的表现,不同物体反射光的偏振态不同,在一定程度上能表征目标物体的特征信息。研究表明,水下偏振成像技术利用散射光的偏振特性可以在一定程度上减小后向散射光对水下成像的影响^[8]。偏振探测中,通常用斯托克斯矢量表征光波的强度和偏振态。其中,总光强包含反射光和透射光并倾向于描述物体的反射率和透射率;线性偏振度(Degree of Linear Polarization, DoLP)图像用于表征偏振特性并提供详细的特征。为了获得同一场景的互补信息,需要融合强度和 DoLP 图像。近年来,图像处理领域出现了许多图像融合方法。图像融合方法根据融合过程可以分为像素级、特征级和决策级 3 种不同层级的融合。其中,像素级图像融合的研究和应用最多,分为基于变换域的融合方法(图像金字塔^[10]、小波变换^[11]、Ridgelets 变换^[12]、Curvelet 变换^[13])和基于空间域的融合方法(HIS 变换融合方法^[14]、主成分分析逆变换融合方法^[15]、像素值加权法、统计模型融合方法)两类,可从像素

层面实现多源图像的融合,融合后的图像信息更加全面、精确、可靠。

近年来,深度学习技术发展飞快,计算机视觉和图像处理领域的许多问题取得了突破性进展。目前,基于深度学习的图像融合技术研究日益增多,在数字成像(多焦点图像融合^[16]、多曝光图像融合^[17]等)、多模态成像(可见光与红外光图像融合^[18-21]、医学多模态图像融合^[22]、偏振图像融合^[23]等),以及遥感成像^[24](多光谱与全色图像融合、多光谱与高光谱图像融合等)等多个方向。与传统方法表征输入图像和目标图像复杂映射关系的能力有限相比,深度学习模型在表征不同信号之间复杂关系方面具有强大的能力,并且可从数据中自动提取有效的特征,解决人工设计的困难。本文将深度学习应用于水下偏振图像融合,提出了一种基于无监督学习和注意力机制引导的网络模型,主要分为特征提取、特征融合和图像重构 3 个模块。其中,特征提取模块融合了注意力机制,其次构建了相关损失函数及权重参数。该方法能够有效融合水下光强和偏振度图像信息,提高视觉成像质量,具有较强的稳定性和自适应性。

2 成像原理

2.1 水下成像模型

Jaffe-McGlamery 模型^[25-26]是常用的水下成像模型之一,诸多水下图像复原算法都是基于此模型提出的。如图 1 所示, Jaffe-McGlamery 模型中探测器最终接收到的图像是 3 个分量的线性组合,这 3 个分量分别是探测器接收到的目标反射光 $S(x, y)$ 、光源在到达目标物前经水体散射的后向散射光 $B(x, y)$ 和部分目标反射光经水体散射到达探测器的前向散射光 $F(x, y)$, 则该图像可

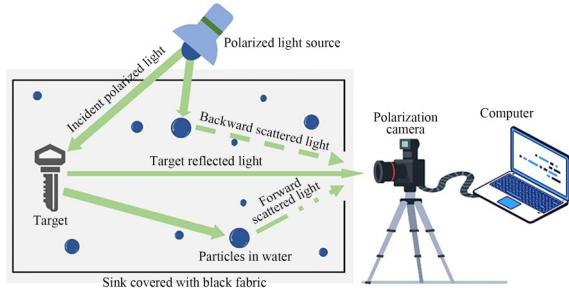


图 1 Jaffe-McGlamery 模型

Fig. 1 Jaffe-McGlamery model

表示为:

$$I(x, y) = S(x, y) + B(x, y) + F(x, y). \quad (1)$$

假定目标最初的辐照度为 $J(x, y)$, 由于散射和吸收的作用, 光从目标传播到探测器时损失部分能量, 目标反射光 $S(x, y)$ 可表示为:

$$S(x, y) = J(x, y) \cdot t(x, y), \quad (2)$$

式中 $t(x, y)$ 为介质透射率, 其表达式为:

$$t(x, y) = e^{-\beta(x, y)\rho(x, y)}. \quad (3)$$

在单一均匀介质中, 介质透射率由衰减系数 $\beta(x, y)$ 和传播距离 $\rho(x, y)$ 决定, 衰减系数 $\beta(x, y)$ 在空间上是恒定的, 即 $\beta(x, y) = \beta_0$, 传播距离 $\rho(x, y)$ 是指物体与相机之间光路的水下部分。

后向散射光 $B(x, y)$ 是背景光经水中颗粒反射到探测器处, 可以表示为:

$$B(x, y) = B_{\infty}(1 - t(x, y)), \quad (4)$$

其中 B_{∞} 表示无限远处的水下环境光光强。由于前向散射对成像质量的影响极小, 通常可以忽略, 则式(1)可简化为:

$$I(x, y) = J(x, y) \cdot t(x, y) + B_{\infty}(1 - t(x, y)). \quad (5)$$

物体最初的辐照度 $J(x, y)$ 可表示为:

$$J(x, y) = \frac{I(x, y) - B_{\infty}(1 - t(x, y))}{t(x, y)}. \quad (6)$$

2.2 偏振成像原理

在偏振探测领域最常用的偏振特性表示方法之一是斯托克斯矢量法, 该方法能够全面地表征入射光波以及物质作用后光波的偏振特性。该矢量由 4 个参量组成, 如下:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_0 + I_{90^\circ} \\ I_0 - I_{90^\circ} \\ I_{45^\circ} - I_{135^\circ} \\ I_r - I_l \end{pmatrix}, \quad (7)$$

式中: S_0 代表探测器接收到的总光强, S_1 代表 0° 和 90° 偏振分量 I_0 和 I_{90° 的光强度差值, S_2 代表 45° 和 135° 偏振分量 I_{45° 和 I_{135° 的光强度差值, S_3 代表光波中左旋和右旋圆偏振分量 I_r 和 I_l 的光强差。不同方向的光强可以通过穆勒矩阵得出, 即:

$$S' = \begin{pmatrix} S_0' \\ S_1' \\ S_2' \\ S_3' \end{pmatrix} =$$

$$\frac{1}{2} \begin{pmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \cos 2\theta \sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta \sin 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}, \quad (8)$$

式中: θ 是主光轴与参考线之间的夹角, S' 表示角度为 θ 的出射光。根据式(8)可得角度为 θ 的出射光光强为:

$$I'(\theta) = \frac{1}{2}(S_0 + S_1 \cos 2\theta + S_2 \sin 2\theta). \quad (9)$$

偏振相机能够获得偏振方向为 $0^\circ, 45^\circ, 90^\circ$ 和 135° 的光强图, 是因为 CMOS 传感器的每个像元上都有 4 个不同角度的偏振片 ($0^\circ, 45^\circ, 90^\circ, 135^\circ$)。如图 2 所示, 每 4 个像元一组作为一个计算单元, 进而获得 θ 为 $0^\circ, 45^\circ, 90^\circ$ 和 135° 的光强, 联立后可以得到式(7)。利用斯托克斯矢量能够进一步计算入射光的线偏振度 DoLP 以及偏振角 φ :

$$DoLP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \quad (10)$$

$$\varphi = \frac{1}{2} \arctan\left(\frac{S_2}{S_1}\right), \quad (11)$$

其中: DoLP 代表光波中线偏振分量占总光强的比例, φ 是入射光的优势偏振方向。

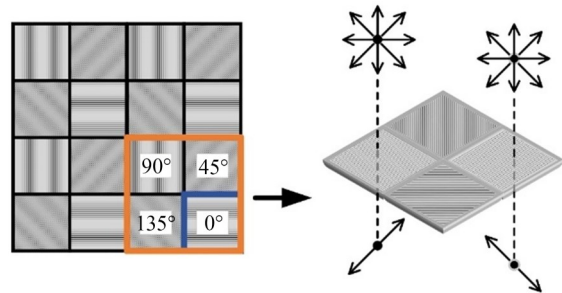


图 2 偏振传感器原理

Fig. 2 Schematic diagram of polarization sensor

3 本文算法

3.1 网络结构

本文采用的网络结构如图 3 所示,它主要包含:特征提取模块,特征融合模块和图像重构模块 3 个模块。在特征提取模块,光强图像(S_0 image)和偏振度图像(DoLP image)采用双通道输入,第一层为包含 3×3 卷积核的卷积层和激活函数 ReLU,用来提取低级特征;之后是包含 3 个卷积层的 DenseBlock 模块,用来提取高级特征,每个卷积层也采用 3×3 的卷积核,卷积核的运算步长为 1。在卷积运算之前还有 BN 层和 ReLU 激活函数,这种排序方式能够加快网络的训练速度,DenseBlock 结构能够保留较多的深度特征,增强特征传播,确保之后的融合过程有足够的深度特征,也能缓解梯度消失的问题。光强图像和偏振度图像输入的两个通道共享相同的权重,减少了参数量,有助于降低网络的计算复杂度。之后是注意力单元(见 3.2 节),注意力单元将前一层特征图作为输入,能够捕获数据中的全局关

系,引导网络学习特征图的分布。在特征融合模块,将特征提取模块输出的特征图叠加,两个特征图的通道大小均为 128,叠加后融合特征图的通道大小为 256。最后特征融合模块的输出作为图像重构模块的输入,图像重构模块包括 5 个转置卷积层,每个转置卷积层的卷积核尺寸为 3×3 ,通过重构模块从融合特征中重建出融合结果。详细的网络结构配置如表 1 所示。

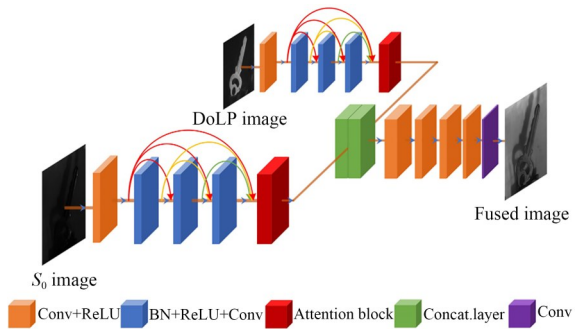


图 3 水下偏振图像融合算法的网络结构
Fig. 3 Network architecture of under-water polarized image fusion algorithm

表 1 水下偏振图像融合算法的网络结构配置

Tab. 1 Network structure configuration of under-water polarized image fusion algorithm

Modules	Inputs	Layers	Convolution kernel size	Input channel	Output channel	Activation functions	Outputs
Feature extraction	S_0	Conv	3×3	1	16	ReLU	F_{S_0}
		Conv	3×3	16	16	ReLU	
		Conv	3×3	32	16	ReLU	
		Conv	3×3	48	16	ReLU	
	DoLP	AttentionBlock	—	64	128	ReLU	
		Conv	3×3	1	16	ReLU	
		Conv	3×3	16	16	ReLU	
		Conv	3×3	32	16	ReLU	
		Conv	3×3	48	16	ReLU	
		AttentionBlock	—	64	128	ReLU	
Feature fusion	F_{S_0}, F_{DoLP}	Concat	—	128×2	256	—	F_{fused}
		Conv	3×3	256	128	ReLU	
		Conv	3×3	128	64	ReLU	
Image reconfiguration	F_{fused}	Conv	3×3	64	32	ReLU	I_f
		Conv	3×3	32	16	ReLU	
		Conv	3×3	16	1	—	

3.2 注意力单元

注意力机制是一种捕获全局依赖的机制。

注意力单元结合了通道注意力和空间注意力,通道注意力能够使网络学习通道域的特征重要性,

给与特征图不同的权重,从而在通道域上实现光强图像和偏振图像的选择性结合;而空间注意力聚焦于学习每一层特征图的有效信息分布,提高显著性特征的传递。注意力单元包括全局平均池化层、卷积层、激活层和拼接层等结构,如图4所示。

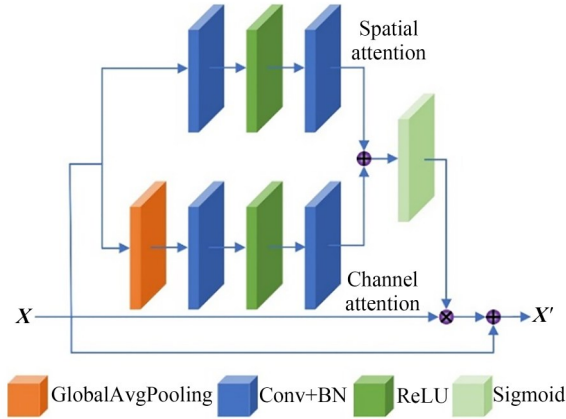


图4 注意力单元结构

Fig. 4 Structure of attention unit

给定 $X \in \mathbb{R}^{H \times W \times C}$ 和 $X' \in \mathbb{R}^{H \times W \times C}$ 为注意力单元的输入和输出,则注意力单元的计算过程如下:

$$X' = \sigma(F_c(X) \oplus F_s(X)) \otimes X + X, \quad (12)$$

式中: σ 为 Sigmoid 激活函数, F_c 为通道注意力分支, F_s 为空间注意力分支, $\oplus$ 为广播加法运算, $\otimes$

$$Loss^{MSW-SSIM} = 1 - \frac{1}{5} \cdot \sum_{\omega \in [3, 5, 7, 9, 11]} (\gamma_\omega \cdot Loss^{SSIM}(I_{S_0}, I_f; \omega) + (1 - \gamma_\omega) \cdot Loss^{SSIM}(I_{DoLP}, I_f; \omega)), \quad (15)$$

式中: $Loss^{SSIM}(x, y; \omega)$ 为基于 SSIM 的损失函数,代表在窗口 ω 上图像 x 和 y 的结构相似性,其表达式为:

$$Loss^{SSIM}(x, y; \omega) = \frac{(2\varpi_x \varpi_y + C_1)(2\omega_x \omega_y + C_2)}{(\varpi_x^2 + \varpi_y^2 + C_1)(\sigma_{\omega_x}^2 + \sigma_{\omega_y}^2 + C_2)}. \quad (16)$$

SSIM 是衡量两幅图像相似度的指标,包括亮度、对比度和结构 3 个不同因素。SSIM 的取值为 $-1 \sim 1$, SSIM 越接近 1, 两个图像越相似。

该损失函数基于多窗口的 SSIM, 解决了不同尺度的图像细节问题。使用的窗口尺寸包括 3, 5, 7, 9, 11, 不同窗口可以提取不同尺度的特征。此外, 在 $Loss^{SSIM}(I_{S_0}, I_f; \omega)$ 和 $Loss^{SSIM}(I_{DoLP}, I_f; \omega)$ 之间使用权重系数, 权重系

为逐元素乘法运算。

作为输入的特征图 $X \in \mathbb{R}^{H \times W \times C}$ 经过通道注意力分支时, 首先经过全局平均池化层得到通道特征 $X_c \in \mathbb{R}^{1 \times 1 \times C}$, 然后用逐点卷积 $PWConv_1$, BN 层和 ReLU 激活函数得到尺寸为 $1 \times 1 \times r$ 的通道特征, 再经过逐点卷积 $PWConv_2$ 和 BN 层得到尺寸为 $1 \times 1 \times C$ 的通道注意力特征图 X_c 。 F_c 表示为:

$$F_c(X) = BN(PWConv_2(\delta(BN(PWConv_1(GAP(X)))))), \quad (13)$$

其中: δ 为 ReLU 激活函数, GAP 为全局平均池化。与通道注意力分支相似, 在经过空间注意力分支时, 先用 3×3 卷积 $Conv_1$, BN 层和 ReLU 激活函数得到尺寸为 $H \times W \times r$ 的特征图, 再经过 1×1 卷积 $PWConv_2$ 和 BN 层得到尺寸为 $H \times W \times C$ 的空间注意力特征图 X_s 。 F_s 表示为:

$$F_s(X) = BN(PWConv_2(\delta(BN(Conv_1(X))))). \quad (14)$$

3.3 损失函数

本文中的损失函数包括全局加权结构相似性 (Structural Similarity Index Measure, SSIM) 损失函数 (MSW-SSIM)^[23], 如下:

数基于 $\sigma_{\omega_{S_0}}^2$ 和 $\sigma_{\omega_{DoLP}}^2$ 确定, 定义如下:

$$\gamma_\omega = \frac{g(\sigma_{\omega_{S_0}}^2)}{g(\sigma_{\omega_{S_0}}^2) + g(\sigma_{\omega_{DoLP}}^2)}, \quad (17)$$

其中: $g(x)$ 为防止异常情况设置的截断函数, $g(x) = \max(x, 0.0001)$ 。当窗口内 S_0 图像的方差大于对应 DoLP 图像的方差时, 说明 S_0 图像的局部区域具有更多的图像细节, S_0 图像的权重系数 γ_ω 应该更大。

此外, MSW-SSIM 可以保留高频信息, 但对均匀的偏差不敏感容易导致亮度的改变, 所以将它与 L_1 范数损失函数整合, 可以保证融合结果的亮度。 L_1 范数损失函数表示为:

$$Loss^{L_1} = \frac{1}{M \cdot N} \|I_{avg} - I_f\|_1, \quad (18)$$

式中: M 和 N 分别是图像的高度和宽度, I_{avg} 为 I_{S_0} 和 I_{DoLP} 的平均值。最终损失函数可以表示为:

$$L^{\text{Mix}} = \alpha \cdot \text{Loss}^{\text{MSW-SSIM}} + (1 - \alpha) \cdot G_\sigma \cdot \text{Loss}^{L_1}, \quad (19)$$

式中: α 为控制 MSW-SSIM 损失函数和高斯加权 L_1 范数损失函数之间权重的超参数,设置为 0.84。

4 实验与结果分析

4.1 实验配置

为了获取数据集,进行水下成像实验,实验装置及布局如图 1 所示,主要包括偏振相机、玻璃水箱、偏振光源和目标物体。偏振相机型号为 PHX050S-P,目标分辨率为 $2\,448 \times 2\,048$,采用焦平面偏光成像,镜头焦距为 10.5 mm,可以一次拍摄 4 张偏振角度分别为 $0^\circ, 45^\circ, 90^\circ$ 和 135° 的线偏振光强度图像。采用透明玻璃水箱作为容器,并在水箱内壁覆盖黑色绒布,以避免环境光和内壁反射的干扰。将目标放置在装满水的玻璃水缸中,获取光强图和线偏振度图像,最终构建了包含 150 组图像的数据集。每组图像由对应的光强图和线偏振度图构成,图像尺寸为 $1\,224 \text{ pixel} \times 1\,024 \text{ pixel}$,其中 100 组作为训练集,剩余 50 组为验证集和测试集。对数据集图像进行翻转,并裁剪为 80×80 大小用作网络训练的输入,训练过程在显卡为 NVIDIA GeForce RTX 2080 Ti 的服务器上进行。在进行权值初始化后,使用 mini-batch 大小为 128 的 Adam 优化器进行优化。学习率最初设置为 0.000 1,并以 0.99 的速率呈指数衰减,最大 Epoch 设置为 200。

4.2 结果分析

基于无监督学习和注意力机制对网络进行训练,并对水下图像融合性能进行测试。为了客观衡量融合图像的质量,采用信息熵(EN)、标准差(SD)、互信息(MI)和结构相似性(SSIM)4 种图像评价指标。信息熵表示图像的平均信息量,其计算公式如下:

$$H(X) = E \left[\log \frac{1}{p(a_i)} \right] = - \sum_{i=1}^n p(a_i) \log p(a_i). \quad (20)$$

信息量越大,信息熵越大。图像融合后会造造成图像信息量的增加,而信息熵可以反映出变化的程度。

标准差指图像像素灰度值相对于均值的离散程度。标准差越大,表明图像中灰度级分布越分散,图像质量也就越好,其计算公式为:

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (F(i, j) - \mu)^2} \quad (21)$$

$$\mu = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N F(i, j)$$

互信息可度量两幅图像之间的相似程度,即融合图像获取了多少原图像信息量。互信息越大,表示融合图像保留更多源图像信息,质量越好。互信息由图像的信息熵 $H(A)$ 和联合信息熵 $H(A, B)$ 决定,即:

$$MI(A, B) = H(A) + H(B) - H(A, B). \quad (22)$$

这里 MI 的计算公式为:

$$MI = \frac{1}{2} MI(I_{S_0}, I_f) + \frac{1}{2} MI(I_{\text{DoLP}}, I_f). \quad (23)$$

SSIM 是一个广泛使用的图像质量评价指标,它是基于人眼观看图像时会提取其中的结构化信息的假设。SSIM 越接近 1,相似度越高,融合质量越好。本文采用的计算公式为:

$$SSIM = \frac{1}{2} \cdot SSIM(I_{S_0}, I_f) + \frac{1}{2} \cdot SSIM(I_{\text{DoLP}}, I_f). \quad (24)$$

图像融合结果如图 5 所示,可以发现,光强图像质量较差,场景细节退化严重,在加入偏振图像进行融合后,目标物变得更清晰,钥匙的纹理轮廓能够清楚地识别。根据采用的图像指标进行评价,融合后的信息熵提升 24.48%,标准差提升 139%,表明本方法可提升水下图像质量。

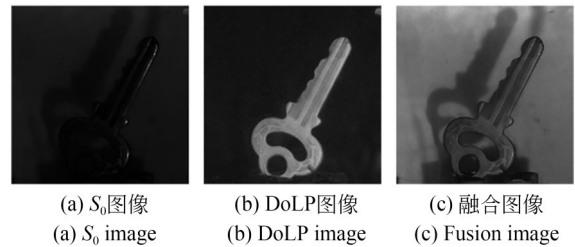


图 5 融合结果

Fig. 5 Fusion result

此外,进行了消融实验,并将本方法得到的融合图像与其他几种图像融合方法进行了比较,包括曲波变换(Curvelet Transform, CVT)^[13]、梯度转移(Gradient Transfer, GTF)^[27]、多分辨率奇

异值分解 (Multi-resolution Singular Value Decomposition, MSVD)^[28]、低通率金字塔变换(Ratio of low-pass Pyramid, RP)^[10]、离散小波变换 (Discrete Wavelet Transform, DWT)^[11] 以及移除注意力单元后的网络。如图 6 所示,RP 的结果具有较差的视觉效果,钥匙边缘和阴影部分均产生了一定程度的伪影,且噪点较多;CVT,DWT 和 MSVD 的结果有一定程度的颗粒感,且对比度偏低,对于阴影的处理较差;GTF 的结果具有较

高的对比度,但钥匙的纹理细节不够清晰;而本文方法具有较为真实的视觉效果,没有明显的伪影和失真且对阴影的处理效果较好,相对于移除注意力单元的网络在细节方面更为精细。

为了客观地评估方法的性能,使用之前介绍的 4 种图像评价指标对测试集中的图像进行评估,最终结果取平均值,结果如表 2 所示。本方法在信息熵、标准差、MI 和 SSIM 3 种图形评价指标中均具有较好的表现,进一步说明了方法的有效性。

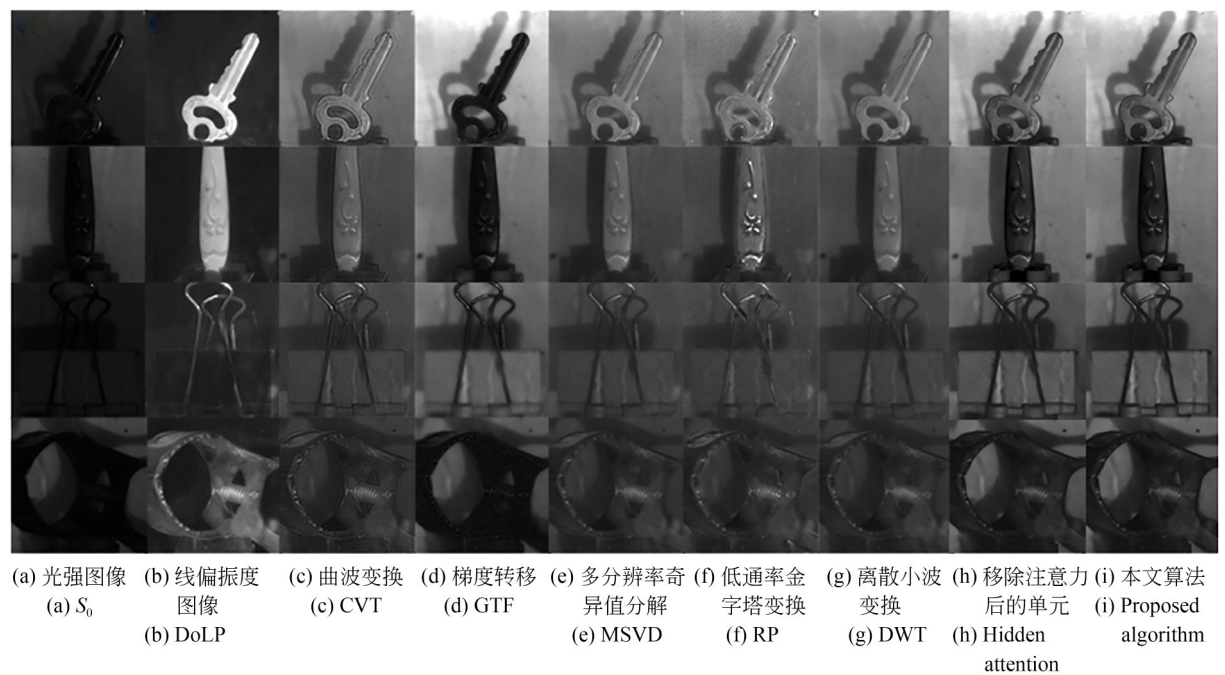


图 6 部分测试集图像不同方法融合结果对比

Fig. 6 Comparison of fusion results of different methods for partial test set images

表 2 不同方法对偏振图像融合的定量结果

Tab. 2 Quantitative results of polarization image fusion by different methods

Method	EN	SD	MI	SSIM
CVT	6.112 3	17.87	0.737 4	0.495 3
GTF	6.981 8	42.62	1.272 5	0.354 2
MSVD	6.059 1	17.11	0.757 4	0.495 9
RP	6.481 5	21.93	0.360 0	0.240 9
Wavelet	6.083 4	18.21	1.471 7	0.504 9
Hidden Attention	7.284 1	37.96	1.617 3	0.503 3
Proposed method	7.307 3	40.46	1.680 8	0.518 6

最后,在配置为 NVIDIA GeForce RTX 2080 Ti, 3.1GHz Intel Core i9-9600X, 128G RAM 的服务器上进行运行时间评估,结果如表 3

所示。所有方法均在 Python 语言下实现,采用多组取平均值,本文方法的处理速度为 0.051 s,优于其他方法。

表 3 不同方法的运算时间

Tab. 3 Operating time of different methods

Method	Operating time/s
CVT	1.042 0
GTF	0.068 0
MSVD	0.356 8
RP	0.082 0
Wavelet	0.508 6
Ours	0.051 0

5 结 论

本文针对水下光学成像质量差的问题,通过分析水下主动偏振成像模型,提出了将深度融合网络应用于水下偏振图像的方法,搭建实验装置获取水下偏振图像构建训练数据集,构造了基于

无监督学习和注意力机制引导的端到端网络模型并设计损失函数。实验结果表明,本文方法提高了图像视觉质量,融合后的水下图像细节更为丰富,相比于光强图像信息熵提升了24.48%,标准差提升了139%。传统的图像融合方法通常需要设定权重参数来控制不同输入图像之间的融合程度,而权重参数的设定需要经过反复实验和调整才能得到最佳效果。相比之下,该方法不需要手动设定权重参数,可以从数据中学习到的融合策略。下一步研究包括构建更为全面的数据集、改进损失函数和网络模块以进一步提升融合图像质量,以满足实际应用的要求;同时还要提高算法的处理效率,减小运算时间,以满足实时处理水下视频的需求,实现对水下目标的实时检测。

参考文献:

- [1] LI C Y, GUO C L, REN W Q, *et al.* An underwater image enhancement benchmark dataset and beyond[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4376-4389.
- [2] GONZÁLEZ-GARCÍA J, GÓMEZ-ESPINOSA A, CUAN-URQUIZO E, *et al.* Autonomous underwater vehicles: localization, navigation, and communication for collaborative missions[J]. *Applied Sciences*, 2020, 10(4): 1256.
- [3] 褚金奎,林木音,王寅龙,等. 偏振光传感器的无人船导航与编队应用[J]. *光学精密工程*, 2020, 28(8):1661-1669.
CHU J K, LIN M Y, WANG Y L, *et al.* Application of polarization sensor to unmanned surface vehicle navigation and formation[J]. *Opt. Precision Eng.*, 2020, 28(8): 1661-1669. (in Chinese)
- [4] 张然,桂心远,成昊远,等. 基于偏振成像的低光照强背景噪声下的目标位姿估计[J]. *光学精密工程*, 2021, 29(4):647-655.
ZHANG R, GUI X Y, CHENG H Y, *et al.* Target pose estimation based on polarization imaging in low light and strong background noise[J]. *Opt. Precision Eng.*, 2021, 29(4):647-655. (in Chinese)
- [5] CHENG H Y, CHU J K, ZHANG R, *et al.* Real-time position and attitude estimation for homing and docking of an autonomous underwater vehicle based on bionic polarized optical guidance[J]. *Journal of Ocean University of China*, 2020, 19(5): 1042-1050.
- [6] 韩平丽. 水下目标偏振成像探测技术研究[D]. 西安: 西安电子科技大学, 2018.
HAN P L. *Research on Polarization Imaging Detection Technology of Underwater Target*[D]. Xi'an: Xidian University, 2018. (in Chinese)
- [7] 褚金奎,张培奇,成昊远,等. 基于特定偏振态成像的水下图像去散射方法[J]. *光学精密工程*, 2021, 29(5):1207-1215.
CHU J K, ZHANG P Q, CHENG H Y, *et al.* Descattering method of underwater image based on imaging of specific polarization state[J]. *Opt. Precision Eng.*, 2021, 29(5):1207-1215. (in Chinese)
- [8] HUANG B J, LIU T G, HU H F, *et al.* Underwater image recovery considering polarization effects of objects[J]. *Optics Express*, 2016, 24(9): 9826-9838.
- [9] TREIBITZ T, SCHECHNER Y Y. Active polarization descattering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(3): 385-399.
- [10] TOET A. Image fusion by a ratio of low-pass pyramid[J]. *Pattern Recognition Letters*, 1989, 9(4): 245-253.
- [11] PAJARES G, MANUEL DE LA CRUZ J. A wavelet-based image fusion tutorial[J]. *Pattern Recognition*, 2004, 37(9): 1855-1872.

- [12] CANDÈS E J, LEIGH DONOHO D. *Ridgelets: Theory and Applications* [M]. Stanford University, 1998.
- [13] NENCINI F, GARZELLI A, BARONTI S, *et al.* Remote sensing image fusion using the curvelet transform [J]. *Information Fusion*, 2007, 8(2): 143-156.
- [14] RAHMANI S, STRAIT M, MERKURJEV D, *et al.* An adaptive IHS pan-sharpening method [J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 7(4): 746-750.
- [15] XIE H Q, LIU Y F, THUERING T, *et al.* Photon-counting spectral CT with de-noised principal component analysis (PCA) [C]. *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine. June 2-6, 2019. Philadelphia, USA. SPIE*, 2019, 11072: 177-181.
- [16] LIU Y, CHEN X, PENG H, *et al.* Multi-focus image fusion with a deep convolutional neural network [J]. *Information Fusion*, 2017, 36: 191-207.
- [17] KALANTARI N K, RAMAMOORTHY R. Deep high dynamic range imaging of dynamic scenes [J]. *ACM Transactions on Graphics*, 2017, 36(4): 1-12.
- [18] MA J, MA Y, LI C. Infrared and visible image fusion methods and applications: a survey [J]. *Information Fusion*, 2019, 45: 153-178.
- [19] 刘志豪, 金伟其, 李力, 等. 四波段共光轴成像实验平台及其图像融合 [J]. *光学精密工程*, 2022, 30(1): 1-11.
- LIU ZH H, JIN W Q, LI L, *et al.* Four-band coaxial imaging experimental platform and its image fusion methods [J]. *Optics and Precision Engineering*, 2022, 30(1): 1-11. (in Chinese)
- [20] 丁贵鹏, 陶钢, 李英超, 等. 基于非下采样轮廓波变换与引导滤波器的红外及可见光图像融合 [J]. *兵工学报*, 2021, 42(9): 1911-1922.
- DING G P, TAO G, LI Y CH, *et al.* Infrared and visible images fusion based on non-subsampled contourlet transform and guided filter [J]. *Acta Armamentarii*, 2021, 42(9): 1911-1922. (in Chinese)
- [21] 郭全民, 王言, 李翰山. 改进 IHS-Curvelet 变换融合可见光与红外图像抗晕光方法 [J]. *红外与激光工程*, 2018, 47(11): 432-440.
- GUO Q M, WANG Y, LI H SH. Anti-halation method of visible and infrared image fusion based on improved IHS-Curvelet transform [J]. *Infrared and Laser Engineering*, 2018, 47(11): 432-440. (in Chinese)
- [22] ELZEKI O M, ABD ELFATTAH M, SALEM H, *et al.* A novel perceptual two layer image fusion using deep learning for imbalanced COVID-19 dataset [J]. *PeerJ Computer Science*, 2021, 7: e364.
- [23] ZHANG J, SHAO J, CHEN J, *et al.* Polarization image fusion with self-learned fusion strategy [J]. *Pattern Recognition*, 2021, 118: 108045.
- [24] LIU X, LIU Q, WANG Y. Remote sensing image fusion based on two-stream fusion network [J]. *Information Fusion*, 2020, 55: 1-15.
- [25] MCGLAMERY B L. A computer model for underwater camera systems [C]. *Proc SPIE 0208, Ocean Optics VI*, 1980, 0208: 221-231.
- [26] JAFFE J S. Computer modeling and the design of optimal underwater imaging systems [J]. *IEEE Journal of Oceanic Engineering*, 1990, 15(2): 101-111.
- [27] MA J, CHEN C, LI C, *et al.* Infrared and visible image fusion via gradient transfer and total variation minimization [J]. *Information Fusion*, 2016, 31: 100-109.
- [28] NAIDU V P S. Image fusion technique using multi-resolution singular value decomposition [J]. *Defence Science Journal*, 2011, 61(5): 479.

作者简介:



巩文哲(1999—),男,河北邢台人,硕士研究生,2019年于燕山大学获得学士学位,主要从事仿生视觉认知技术方面的研究。E-mail: wenzhe_gong@mail.dlut.edu.cn

通讯作者:



张然(1984—),男,辽宁锦州人,博士,副教授,硕士生导师,2007年、2013年于大连理工大学分别获得学士和博士学位,主要从事仿生视觉认知技术、亚波长光学结构设计和微纳传感器与执行器等方面的研究。E-mail: zhangr@dlut.edu.cn