arXiv:2603.00908v1 [cs.CV] 1 Mar 2026

图文摘要

UD-SfPNet：一种用于三维法线重建的水下去散射偏振形状网络

Puyun Wang、Kaimin Yu、Huayang He、Feng Huang、Xianyu Wu、Yating Chen

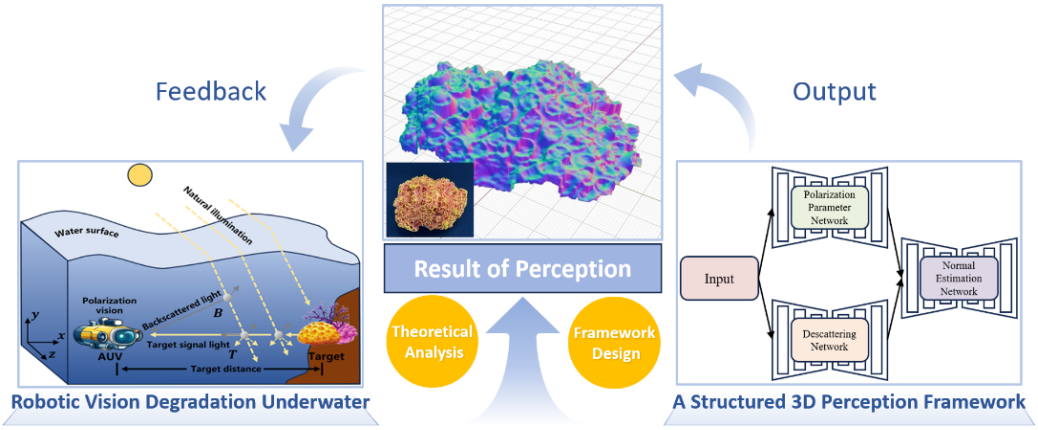

图1：所提出的水下机器人偏振三维视觉感知流程示意图。

亮点

UD-SfPNet：一种用于三维法线重建的水下去散射偏振形状网络

Puyun Wang、Kaimin Yu、Huayang He、Feng Huang、Xianyu Wu、Yating Chen

- UD-SfPNet 联合训练水下偏振去雾和 SfP 法线
- 端到端训练可避免级联流水线中的误差累积
- 颜色嵌入迁移通过 RGB编码法线确保几何一致性
- 细节增强卷积可改善两个阶段的高频几何
- 15.12° MuS-Polar3D 上的平均角误差，优于强基线方法

UD-SfPNet：一种用于三维法线重建的水下去散射偏振形状网络

Puyun Wang[a]、Kaimin Yu[a]、Huayang He[b]、Feng Huang[a]、Xianyu Wu[a,*]、Yating Chen[c,*]

[a]福州大学机械工程及自动化学院，福州，*35018*，福建，中国[b]交通运输部公路科学研究院，北京，*100088*，中国[c]精密测试技术及仪器国家重点实验室，清华大学精密仪器系，北京，*100084*，中国

摘要

水下光学成像受到散射的严重阻碍，但偏振成像兼具去散射和基于偏振的形状恢复（SfP）三维重建这两大独特优势。为充分发挥这些优势，本文提出 UD-SfPNet，这是一种水下去散射偏振形状恢复网络，利用偏振线索提升三维表面法线预测效果。该框架在统一流程中联合建模基于偏振的图像去散射与偏振形状恢复法法线估计，避免了串行处理导致的误差累积，并实现了两项任务的全局优化。UD-SfPNet 还引入了一种新颖的颜色嵌入模块，利用颜色编码与表面朝向之间的关系来增强几何一致性。此外，该网络还包含一个细节增强卷积模块，以更好地保留散射造成损失的高频几何细节。在 MuS-Polar3D 数据集上的实验表明，所提出的方法显著提高了重建精度，平均表面法线角误差达到 15.12°（在对比方法中最低）。这些结果证实了将去散射与基于偏振的形状推断相结合的有效性，并凸显了 UD-SfPNet 在具有挑战性的水下环境中用于光学三维成像的实用意义与潜在应用，

*通讯作者。电子邮箱：xwu@fzu.edu.cn（Xianyu Wu），chenyating@mail.tsinghua.edu.cn（Yating Chen）

用价值。代码可在 https://github.com/WangPuyun/UD-SfPNet 获取。

关键词： 偏振形状恢复，计算成像，去散射，三维重建，水下成像

---

1. 引言

水下机器人和载人潜水器依靠高性能水下成像与三维感知开展海洋探测。空气中的光学三维成像已发展得相当成熟。然而，在存在散射的水下环境中，由于水中的米氏散射，其性能会显著下降 [1]。这种性能退化表现为纹理细节模糊、有效探测范围大幅缩短以及噪声严重。受计算成像 [2], 启发，本文从成像流程的角度重新审视水下三维成像。完整的水下三维成像流程分为两个连续阶段：首先去散射，随后进行三维重建。在各种可用的光学先验中，光偏振在这两个阶段都具有独特优势。基于这一观察，本文提出一种全链路优化策略：在解耦的去散射与重建阶段之间编码信息，以突破传统光电子成像的局限，获得优于原始观测结果的成像效果。

大多数既有研究都直接使用深度学习模型处理水下数据 [3, 4]。Wang 等人 [5] 将水下三维成像明确构建为一个两阶段流程，包括去散射和后续的三维重建。然而，这两个模型是独立训练的，下游视觉任务则基于去散射结果执行，这可能导致级联处理流程中的误差累积。为解决这一问题，本文提出一种基于水下偏振的三维重建方法，将偏振物理纳入其中，并在统一框架内对去散射过程和三维重建过程进行联合建模。该方法能够对上游和下游视觉任务进行全局学习和联合优化，为机器人水下三维视觉感知提供了新的视角。

主要贡献如下：

1. 本文提出了 UD-SfPNet，这是一种面向水下偏振三维成像的统一结构化学习框架。该框架以端到端的方式联合学习基于偏振的去散射和 SfP 法线估计，从而实现跨模块的全局梯度优化。2. 本文为 SfP 法线估计引入了颜色嵌入模块。通过利用 RGB编码法线图与几何方向之间的对应关系，强化跨通道一致性，从而提升几何稳定性。3. 去散射和法线估计两个阶段均采用了细节增强卷积。这些卷积能够显式捕获方向／微分变化，从而改善高频几何细节的恢复。4. 在 MuS-Polar3D [5] 上的大量实验表明，UD-SfPNet 效果显著，平均角误差为 15.12° ，性能优于强基线方法。

2. 相关工作

2.1. 基于偏振的去散射

基于偏振的去散射，其核心思想是利用散射光与物体反射光之间偏振态的差异，通过偏振计算分离这些成分，从而恢复被散射掩盖的目标信息。在水下场景中，光子散射和吸收都比大气中更强。因此，结合衰减补偿的基于偏振的分离已成为主要研究方向。Schechner团队早期的工作将偏振分析引入水下能见度和结构恢复研究，并探讨了后向散射偏振特性与透射函数之间的关系 [6]。该团队随后提出了主动偏振去散射，通过主动照明增强水下偏振信号，减少对自然光和扫描硬件的依赖，成为该领域最具代表性的物理模型之一 [7]。随着深度学习的最新进展，这一领域迎来了新一轮发展。基于学习的散射去除方法通常分为两类：物理模型引导方法和端到端增强方法。Yang等人 [8]使用多个空洞卷积，提取偏振图像中目标像素周围的局部偏振特征，

并结合经典的 U-Net [9] 恢复高质量偏振信息，以用于下游任务。Liu 等人 [10] 更加关注端到端重建，以及动态或强散射条件下的实时性能；通过结合光学性质分析与基于图像的证据，该模型经过精心调优，实现了一种兼具实时速度和鲁棒性的去散射方法。

2.2 基于偏振的三维成像

基于偏振的三维成像，尤其是偏振形状恢复（SfP），传统上基于菲涅耳反射模型和斯托克斯参数。通过估计天顶角和方位角，从偏振角（AoP）和偏振度（DoP）推断表面法线，并使用额外先验进一步完善三维重建，如 [11, 12, 13] 所示。然而，此类方法对材料属性有严格假设，因此泛化能力有限。

2020年，Ba 等人 [14] 首次将深度学习引入 SfP。他们构建了一个包含复杂材料和多变照明条件的偏振数据集用于训练，证明深度网络能够适应多种条件进行表面法线估计，其潜力远高于经典偏振形状重建。此后，Lei 等人 [15] 将 SfP 拓展至户外复杂场景，Shao 等人 [16] 解决了透明物体中透射分量导致的三维建模难题，而 Lyu 等人 [17] 则研究了未知环境照明对 SfP 的影响。

专为水下散射设计的 SfP 仅在过去两年才开始受到研究。Wu 等人 [3] 首次将 SfP 应用于水下三维目标物重建。基于 U^2-Net [18], ，从原始偏振图像和偏振参数图中提取多尺度特征，从而能够在具有挑战性的水下环境中实现准确、稳健的三维重建。Li 等人 [4] 利用 Transformer 开发了性能更高的偏振三维形状恢复网络 SfP-underwater。然而，该方法依赖单一的编码器–解码器网络处理偏振输入，且未考虑去散射等上游任务对三维重建这一下游任务的影响。Wang 等人 [5] 提出对水下偏振形状重建中的去散射过程进行显式建模，但去散射和三维重建由两个独立的深度模型处理，去散射模型的输出则作为重建输入的一部分。这种级联设计也容易导致误差累积。

3. 方法

3.1. 偏振成像原理

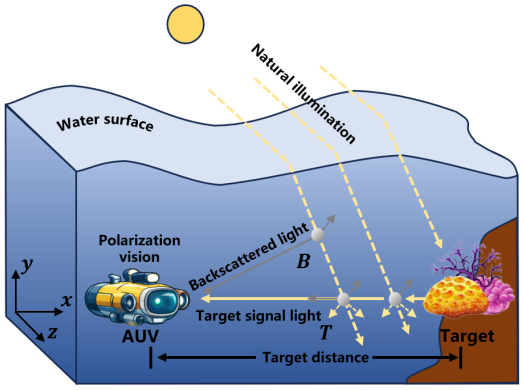

图2：水下偏振成像模型。

在偏振计算成像中，斯托克斯矢量被广泛用于表示和编码偏振特性 [19]。分焦平面（DoFP）偏振相机 [20] 可以同时采集偏振角不同的四幅强度图像，分别记为 $I_{0°}$、 $I_{45°}$、 $I_{90°}$ 和 $I_{135°}$。根据这四幅偏振图像，可以计算出斯托克斯参数：

$$\mathbf{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} I_{0°} + I_{90°} \\ I_{0°} - I_{90°} \\ I_{45°} - I_{135°} \end{bmatrix}. \tag{1}$$

此处， S_0 为总强度， S_1 为水平偏振分量与垂直偏振分量之差， S_2 为 $45°$分量与 $135°$ 分量之差。根据水下偏振成像模型 [8]，所采集的信号可进一步表示为

$$S_0(x, y) = T(x, y) + B(x, y), \tag{2}$$

其中， $T(x,\ y)$表示目标物信号， $B(x,\ y)$表示后向散射光，$(x,\ y)$为像素坐标。传统光电成像主要依赖光谱和强度线索。当目标物处于辐亮度相近的背景中时，其信号很容易被淹没。偏振成像利用后向散射光与目标反射光之间的偏振态差异，并通过

偏振测量分离这两个分量，从散射中恢复被遮蔽的目标信息。深度学习具有强大的非线性表征能力，能够学习偏振线索与散射干扰之间的内在关系，从而实现基于偏振的去散射的高效优化。

基于偏振的三维重建本质上是求解单位法向量$\mathbf{n}$的天顶角 θ和方位角 α。如图 3 所示，$\mathbf{n}$ 可由 θ 和 α 表示为

$$\mathbf{n} = \left[\sin\theta\cos\alpha, \sin\theta\sin\alpha, \cos\theta\right]^{\mathsf{T}}. \tag{3}$$

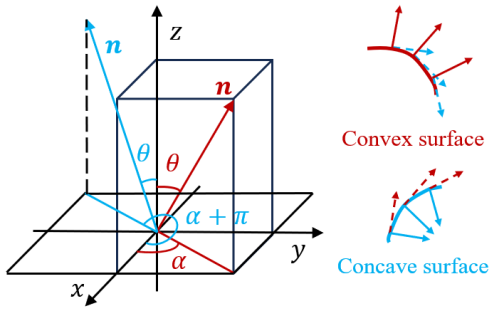

图 3：表面微面元模型中最小单元的示意图。

天顶角 θ 和方位角 α 都存在多解。对于 θ，歧义主要来自镜面反射条件下偏振度 ρ与 θ 之间的非线性关系。对于 α，歧义由正弦函数的周期性引起。这种周期性会产生两个相位相差 π的歧义解，进而可能使局部法向重建呈现为“凸”或“凹”表面（见图3）。

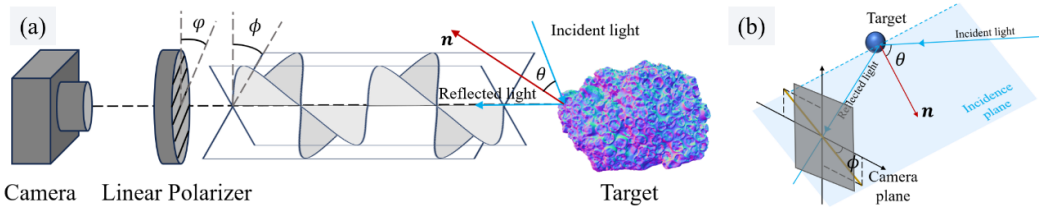

图 4：基于偏振的三维重建的物理模型：(a) 全局视图；(b) 相机视图。

传统的基于偏振的三维重建通常建立在图4所示的物理模型之上。其基本思想是计算偏振度 ρ和偏振角 ϕ，然后反演单位表面法线的天顶角 θ 和方位角 α 。偏振度 ρ和偏振角 ϕ可由斯托克斯参数得到，即

$$\rho = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \tag{4}$$

$$\phi = \frac{1}{2} \arctan\left(\frac{S_2}{S_1}\right). \tag{5}$$

3.1.1. 求解天顶角 θ

给定目标物表面的折射率 η，在镜面反射条件下，偏振度 ρ 与天顶角 θ 之间的关系为

$$\rho = \frac{2\sin^2\theta\cos\theta\sqrt{\eta^2 - \sin^2\theta}}{\eta^2 - \sin^2\theta - \eta^2\sin^2\theta + 2\sin^4\theta}. \tag{6}$$

对应的曲线如图 5(a) 所示。当且仅当 $\rho = 1$时，天顶角等于布儒斯特角，且 θ 唯一。否则，当$\rho \neq 1$时， θ 有两个不同的解。

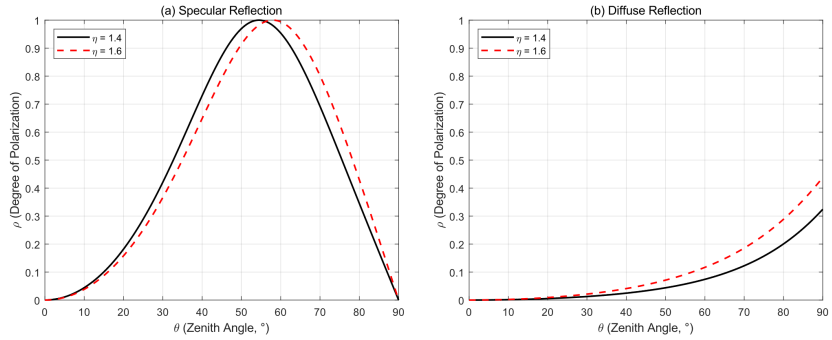

图 5：给定折射率 η时，偏振度 ρ与天顶角 θ 之间的关系：(a) 镜面反射；(b) 漫反射。

对于漫反射，偏振度 ρ与天顶角 θ 之间的关系为

$$\rho = \frac{\left(\eta - \frac{1}{\eta}\right)^2 \sin^2\theta}{2 + 2\eta^2 - \left(\eta + \frac{1}{\eta}\right)^2 \sin^2\theta + 4\cos\theta\sqrt{\eta^2 - \sin^2\theta}}, \tag{7}$$

相应的变化趋势如图5(b)所示。利用上述两个关系，可以根据测量得到的 ρ求出天顶角 θ 。

3.1.2 求解方位角 α

采集到的强度 I 随偏振片旋转角 φ变化，如下所示

$$I(\varphi) = \frac{I_{\max} + I_{\min}}{2} + \frac{I_{\max} - I_{\min}}{2} \cos\left(2\varphi - 2\phi\right). \tag{8}$$

相应的变化趋势如图6所示。当 $\varphi = \phi$时，上述方程给出 $I(\phi) = I_{\max}$。此时，入射面与相机成像平面坐标轴之间的夹角（即偏振角 ϕ，如图4(b)所示）等于偏振片旋转角 φ。对于漫反射，方位角满足 $\alpha = \phi$ 或 $\alpha = \phi + \pi$。对于镜面反射，$\alpha = \phi \pm \frac{\pi}{2}$。当 $\varphi = \phi \pm \frac{\pi}{2}$时，$I(\varphi) = I_{\min}$；这一特性也用于抑制玻璃等物体产生的镜面眩光。

利用 ϕ，即可求得方位角 α 。结合上一小节得到的天顶角 θ ，将 θ 和 α 代入方程3，即可得到表面微元的单位法向量。随后对表面上的法向量场进行积分，便可获得目标物的完整三维形状。

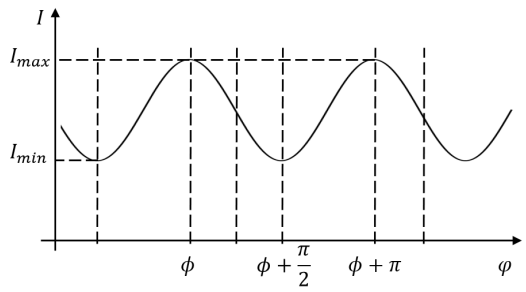

图6：强度 I 与偏振片旋转角 φ之间的关系。

传统的基于偏振的三维成像通常需要辅助传感器（例如深度相机），以提供额外的物理线索来消解天顶角和方位角的歧义。相比之下，深度学习可以通过使用真实标注进行监督训练，使模型隐式地学习偏振物理。尽管无法完全消除法向歧义，但基于学习的方法对这类歧义表现出很强的泛化能力，例如，能够在具有不同材料和光学性质的 MuS-Polar3D 样本上准确预测法向。

3.2. 网络架构

本节介绍所提出的结构化学习框架 UD-SfPNet（图 7）及其架构细节。该框架旨在融合偏振线索，以便在水下散射环境中进行三维重建。框架由三部分组成：（a）偏振参数网络、（b）去散射网络和（c）法线估计网络。

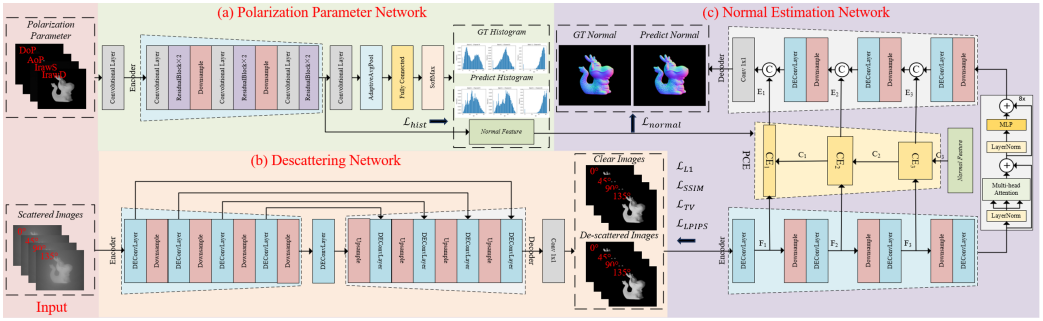

图 7：用于基于水下偏振的三维重建的结构化学习框架 UD-SfPNet 概览。

3.2.1. 偏振参数网络（*PPN*）

给定偏振输入后，该部分训练一个网络来学习从偏振特征到三维表面法线的映射。采用 [21]中所述的偏振输入组合，并将网络构建为“图像编码器 + 分类头”。分类头的输出由根据真实法线图计算得到的法线分布直方图进行监督。模型仅学习法线的全局占比（64 个区间），此阶段不考虑每个法线的空间分布。在训练过程中，保留编码器预测的高维法线特征（法线特征，NF），以供后续三维重建使用。该过程表示为

$$\mathrm{NF}_{\mathrm{pre}},\ \mathrm{Hist}_{\mathrm{pre}} = \mathrm{PPN}\left(\rho,\ \phi,\ I^{S}_{\mathrm{raw}},\ I^{D}_{\mathrm{raw}}\right), \tag{9}$$

其中，$\mathrm{NF}_{\mathrm{pre}}$ 是预测法线特征，$\mathrm{Hist}_{\mathrm{pre}}$ 是对 64 个法线分箱的预测直方图，PPN(·)表示偏振参数网络。这里，ρ 是偏振度，ϕ 是偏振角，I^{S}_{raw} 是镜面分量，I^{D}_{raw} 是漫反射分量，所有这些量

均由原始偏振图像计算得出。使用了一个 ℓ_1 约束：

$$\mathcal{L}_{\text{hist}} = \left\| \text{Hist}_{\text{pre}} - \text{Hist}_{\text{GT}} \right\|_1 , \tag{10}$$

其中， $\mathcal{L}_{\text{hist}}$ 是直方图损失，Hist_{GT} 是根据法线图计算得到的真实直方图。该损失约束分类头，并间接优化图像编码器，使其能够利用偏振线索进行学习，从而实现法线预测。

3.2.2. 去散射网络（*DN*）

去散射网络专注于低层视觉任务。其目标是在散射条件下增强偏振图像的对比度和清晰度，并恢复被散射淹没的目标信息。该网络采用经典的 U-Net [9] 设计，是一种具有四级跳跃连接的编码器–解码器。整体流程为

$$I_{\text{desc}} = \text{DN}(I_{\text{sc}}), \tag{11}$$

其中， I_{sc} 是输入的散射图像，DN(·) 表示去散射操作， I_{desc} 是去散射输出。监督训练使用四种损失：

$$\mathcal{L}_{\text{L1}} = \left\| I_{\text{desc}} - I_{\text{clean}} \right\|_1 , \tag{12}$$

$$\mathcal{L}_{\text{SSIM}} = 1 - \text{SSIM}(I_{\text{desc}}, I_{\text{clean}}) , \tag{13}$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)\,(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}, \tag{14}$$

$$\mathcal{L}_{\text{TV}} = \sum_{i,j} \left(|I_{i+1,j} - I_{i,j}| + |I_{i,j+1} - I_{i,j}| \right) , \tag{15}$$

$$\mathcal{L}_{\text{LPIPS}} = \text{LPIPS}\big(I_{\text{mean}}^{\text{desc}}, I_{\text{mean}}^{\text{clean}}\big) . \tag{16}$$

此处， I_{clean} 表示在清水中拍摄的偏振图像，并用作参考图像。 $\|\cdot\|_1$ 是 ℓ_1 范数，用于确保逐像素重建精度。SSIM 是结构相似性指数，用于约束去散射结果与参考图像在亮度、对比度和结构方面的一致性； μ和 σ 分别表示 SSIM 计算中的均值和方差。 $\mathcal{L}_{\text{TV}}$ 是总变分损失，有助于提升空间平滑性，并抑制散射引入的高频噪声和局部伪影。 $\mathcal{L}_{\text{LPIPS}}$ 是感知损失；先前的研究 [22] 表明，它有助于提升

去散射结果的感知自然度。由于 LPIPS 主要基于三通道自然图像进行训练，因此首先将四个偏振通道转换为平均强度图像，再将其复制为三个通道，作为 LPIPS 输入。 $I_{\text{mean}}^{\text{desc}}$ 和 $I_{\text{mean}}^{\text{clean}}$ 分别表示去散射输出和参考图像的平均强度图。

3.2.3. 法线估计网络（*NEN*）

法线估计网络执行高层三维重建任务。它融合去散射网络和偏振参数网络的输出，并联合恢复高精度表面法线。该架构包含一个编码器和两个解码器。在瓶颈层处使用多头注意力模块 [23] ，以提升高维特征聚合能力。值得注意的是，其中一个解码器采用了 [24]提出的金字塔颜色嵌入（PCE）模块。PCE 的颜色编码特性与 SfP 中的法线编码规则相匹配（见第 3.3 节）。本质上，它是对低级生物视觉线索进行特征融合与提升。因此，本文将其从二维低光增强扩展到三维机器人视觉感知，从而实现了较高的重建精度。法线估计表示为

$$N_{\text{pre}} = \text{NEN}(\text{NF}, I_{\text{desc}}) , \tag{17}$$

其中， N_{pre} 为预测的法线图，NEN(·) 表示整体法线估计过程。损失定义为

$$\mathcal{L}_{\text{normal}} = \frac{\sum_{i=1}^{W}\sum_{j=1}^{H}\left(1 - \left\langle N_{i,j}^{\text{pre}}, N_{i,j}^{\text{gt}} \right\rangle\right) - m}{W \times H - m}, \tag{18}$$

其中 $\langle \cdot, \cdot \rangle$ 表示余弦相似度， W 和 H 分别表示图像宽度和高度， m 表示背景像素的数量。该损失通过余弦相似度衡量角度偏差，并使用掩码仅对前景区域进行归一化，从而提高法线预测的稳定性和准确性。

3.3. 颜色嵌入模块

本小节介绍 DCC [24]中的颜色嵌入模块及其在 SfP 中的潜力。在 SfP 中，通常使用法线图来表示和监督表面法线。如图 8 所示，这种表示本质上是一种基于颜色的向量编码：对于每个像素，单位法向量$\mathbf{n} = (n_x, n_y, n_z)^{\mathsf{T}}$ 被映射到 RGB 通道，从而使三维几何能够紧凑地存储在二维图像平面上。更具体地说，标准

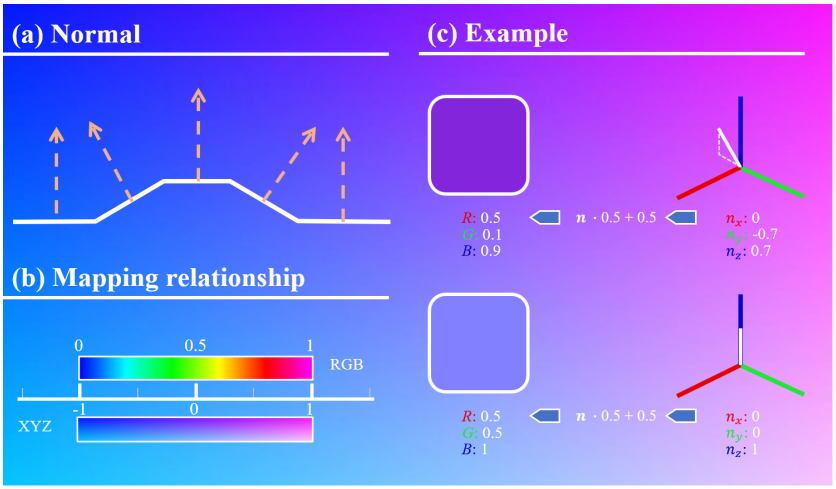

图 8：烘焙法线图的颜色编码机制。

法线图将 $[-1, 1]$ 到 $[0, 1]$, 进行线性归一化，从而得到颜色编码

$$\mathbf{c} = (R, G, B)^{\mathrm{T}} = \frac{\mathbf{n} + 1}{2}, \text{ where } R \leftrightarrow n_x,\ G \leftrightarrow n_y,\ B \leftrightarrow n_z. \tag{19}$$

这种编码广泛应用于图形学法线烘焙和基于法线的监督学习，使三维法线预测等同于在严格几何约束下恢复一幅“彩色图像”。

基于这一观察，偏振形状法线预测可以重新表述为一个由颜色一致性驱动的表征学习问题。如果网络能够在特征空间中学习到具有退化鲁棒性且一致的颜色嵌入，那么这种一致性便可通过法线图的 RGB 编码自然迁移为法线方向一致性，从而提升几何重建的稳定性和实用性。

DCC [24] 中提出的颜色嵌入模块旨在处理低照度下的色偏。其核心思想是在特征空间中对颜色分布进行结构化建模，使输出在空间和语义层面都能保持颜色一致性。尽管该模块最初是为低光照增强而开发的，但它所应对的核心挑战——在强退化下维持稳定且一致的颜色——与水下散射条件下 SfP 法线图的编码特性相契合。

基于上述分析，我们将颜色嵌入模块应用于水下散射环境中的偏振形状法线预测。这并非

一种简单的即插即用方案；它直接源于法线图的“颜色—几何”同构关系，其中颜色一致性意味着几何一致性。

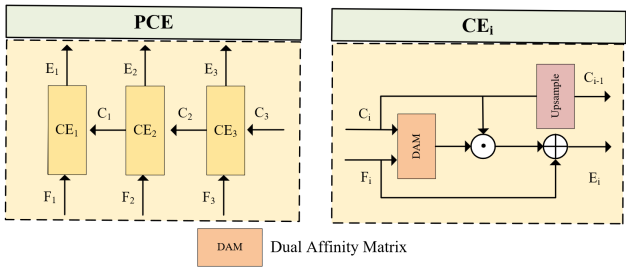

图9：改编自 DCC [24]的颜色嵌入模块：由多层颜色嵌入模块组成的金字塔颜色嵌入（PCE）（左），以及颜色嵌入（CE）模块的详细实现（右）。

3.4. 细节增强卷积模块

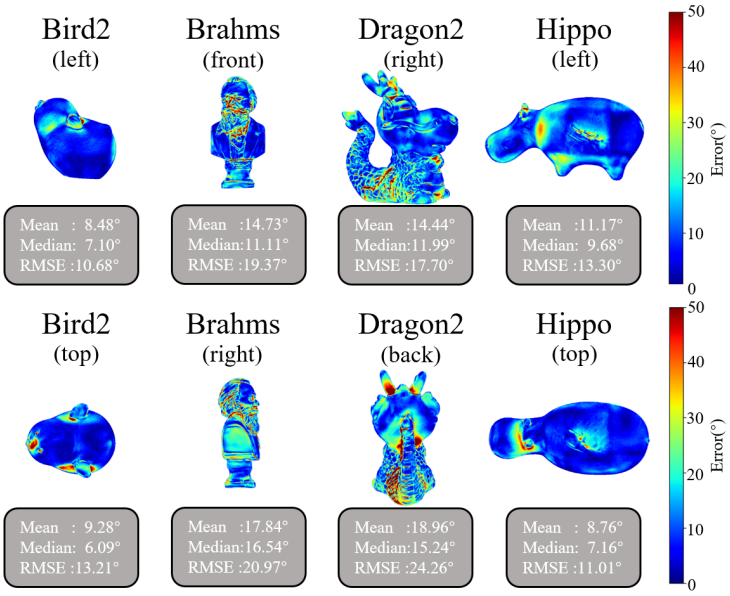

图10：不同样本的误差热图对比。

如图 10 所示，Wang 等人 [5]使用误差热图可视化表明，一旦去散射有效抑制了散射退化，三维重建的主要瓶颈便转移至目标物的内在几何

。具体而言，高频几何纹理更丰富的区域往往会出现更大的法线估计误差和重建误差。

这表明，法线估计的关键挑战在于准确建模高频几何细节。然而，标准卷积算子主要聚合局部邻域内的整体响应，对局部微分线索和方向变化的敏感度有限，因此可能会限制细节丰富区域的预测精度。

基于这一观察，本文将 DEA-Net [25] 中的细节增强卷积（DEConv）模块引入低层去散射网络和高层法线估计网络，作为二者的卷积骨干。该模块以标准卷积为基础，融入多个微分卷积算子，以显式建模局部像素差异和方向变化，从而增强对高频结构信息的表征能力。

在“去散射”和“3D重建”这两项关键任务中引入 DEConv，可提升网络整体稳定性以及复杂几何区域的重建精度。

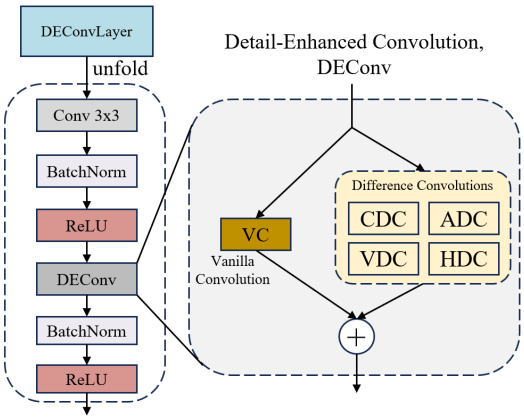

图11：DEConv 模块示意图。

3.5. 进一步讨论

所提出的用于水下偏振三维重建的结构化学习框架，对水下三维成像中的两个过程——“去散射”和“三维重建”进行了显式建模，从而能够在深度学习训练期间进行全局优化。此外，每个子网络都可以

替换为其他方法以进一步提升性能。例如，基于 Transformer 的模型已应用于散射去除 [26]和表面法线估计 [27]。Transformer 通常需要大规模训练数据。在这一结构化框架中，可以冻结其他组件，以便在大型公开去散射/法线估计数据集上预训练目标子网络，再在特定数据集上进行微调，从而获得最佳的水下三维成像性能。如果需要实时性能，则可以移除复杂模块，并用标准卷积块替换，以实现轻量化。

4. 实验与结果

4.1. 训练设置

从公开的 MuS-Polar3D 数据集中选取 726 个散射样本 [5] ，并按 8:1:1 的比例划分为训练集、验证集和测试集。所提出的 3D 成像框架在 PyTorch [28] 平台上使用 4 个 NVIDIA A100 GPU 训练 1000 个 epoch，学习率为 0.001。训练期间，256×256 随机裁剪图块，要求其中超过一半的像素有效，并采用基于图块的训练进行数据增强。推理也在 256×256 图块上进行；使用滑动窗口策略将图块拼接回全分辨率图像，并减少块状伪影。总损失为

$$\mathcal{L}_{\text{total}} = \lambda_1\mathcal{L}_{\text{hist}} + \lambda_2\mathcal{L}_{\text{L1}} + \lambda_3\mathcal{L}_{\text{SSIM}} + \lambda_4\mathcal{L}_{\text{TV}} + \lambda_5\mathcal{L}_{\text{LPIPS}} + \lambda_6\mathcal{L}_{\text{normal}}. \quad (20)$$

超参数根据经验设置为 $\lambda_1 = 1.0$、 $\lambda_2 = 10.0$、 $\lambda_3 = 1.0$、$\lambda_4 = 10.0$、$\lambda_5 = 2.0$ 和 $\lambda_6 = 30.0$。

4.2. 性能评估

所提出的 SfP 方法对水下环境中的去散射过程进行了显式建模。本小节首先评估去散射性能。如图 12 所示，采用 ORB [29] 配准算法的误匹配率，对代表性样本进行定性评估。经过去散射网络处理后，表示错误对应关系的红线数量显著减少，表明去散射网络效果显著。

表 1 进一步比较了原始图像、去散射结果和真实参考图像。报告的指标包括峰值信噪比（PSNR）、结构

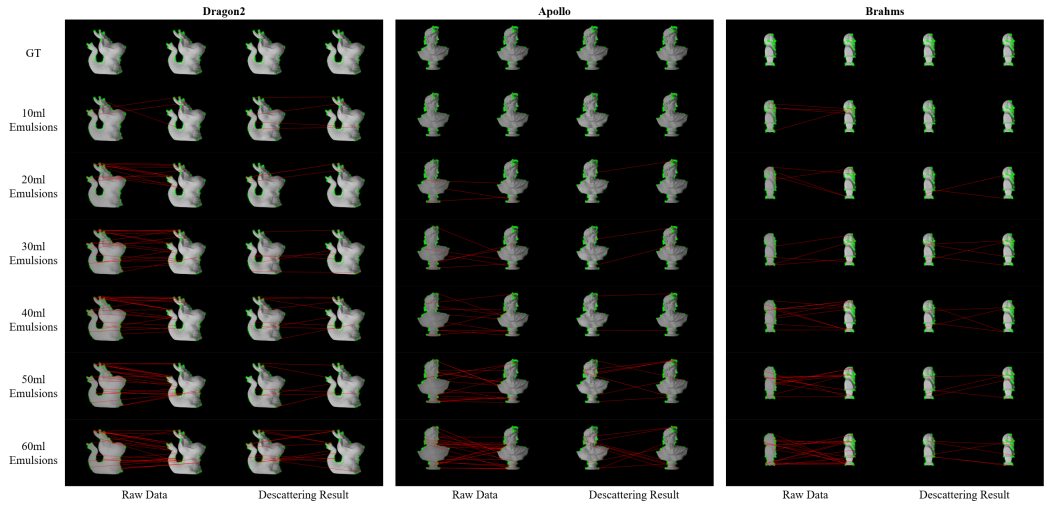

图12：使用基于 ORB 的配准误匹配率评估去散射效果。绿色圆圈表示检测到的关键点，红线表示错误匹配。

表1：去散射前后的定量比较。对真实值和待评估样本均应用掩膜，以聚焦于目标区域。

指标	PSNR↑	结构相似性↑	LPIPS↓
原始数据	30.80	0.9569	0.3830
去散射结果	36.87	0.9745	0.0356

相似性（SSIM）和学习感知图像块相似度（LPIPS）。为更好地反映目标区域的视觉增强效果，在计算指标时，对真实值和评估样本均应用掩膜。结果表明，去散射后，目标区域的所有指标均有所提升，这说明去散射能够改善视觉质量，并为下游任务提供更丰富的有效信息。

所提出的方法与五种法线估计基线方法进行了比较，包括 DeepSfP [14], SfP-wild [15], TransSfP [16], AttentionU^2- Net [3], 和 DSINE [30]。表 2 报告了 MuS-Polar3D测试集上的平均角误差（平均绝对误差）。所提出的方法取得了最低的平均绝对误差，证明了联合建模低层视觉（去散射）和高层视觉（三维重建）的有效性。图 13 进一步展示了代表性样本上的预测结果。DeepSfP、SfP-wild 和 TransSfP 更容易受到水下散射的影响，随着散射增强，其平均绝对误差也随之增大。AttentionU^2-Net、DSINE 和所提出的方法

表 2：不同基线方法在 MuS-Polar3D测试集上的性能比较。指标为平均角误差（平均绝对误差，°）；数值越低越好。

指标	DeepSfP	SfP-wild	TransSfP	注意力 U2-Net	DSINE	提出的方法
MAE↓	19.64°	21.64°	20.54°	15.72°	16.94°	15.12°

对水下环境具有更强的鲁棒性，其中所提出的方法取得了最低的平均绝对误差。

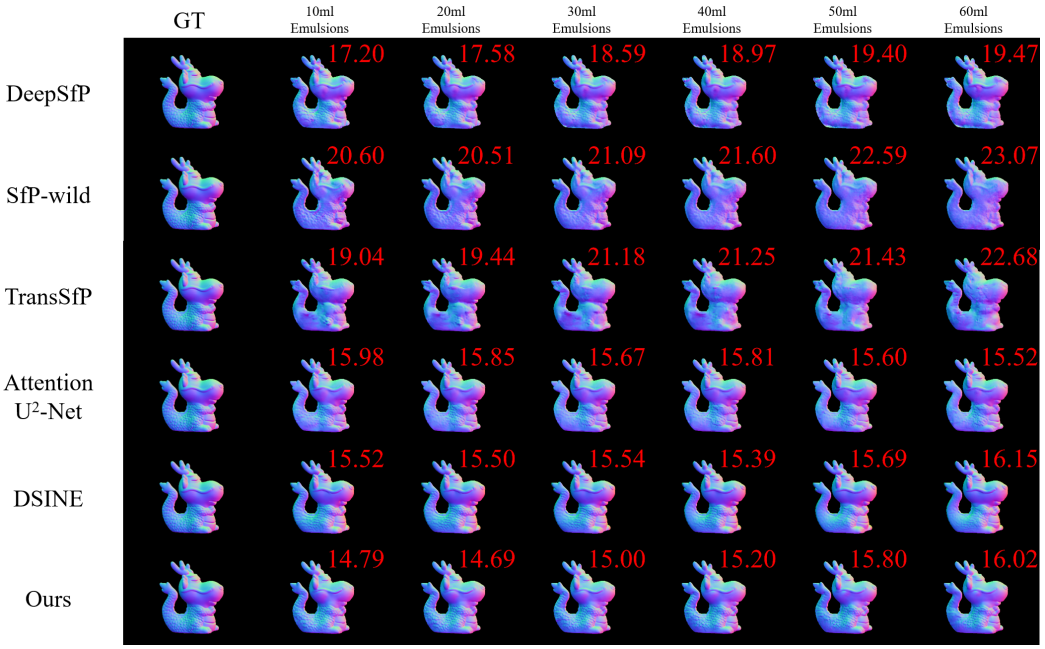

图13：在一个具有代表性的样本（Dragon2）上对比不同基线方法。每个子图右上角的红色数字表示平均绝对误差（°），数值越低表示性能越好。

为进一步直观比较水下散射条件下的三维重建性能，图14展示了MuS-Polar3D代表性测试样本的表面法线预测结果，以及放大的局部细节和相应的误差图。现有方法通常能够恢复整体形状轮廓和基本几何结构，但在曲率变化较大且高频细节丰富的区域（例如边缘、褶皱和局部纹理）仍会出现明显的过度平滑或几何失真。

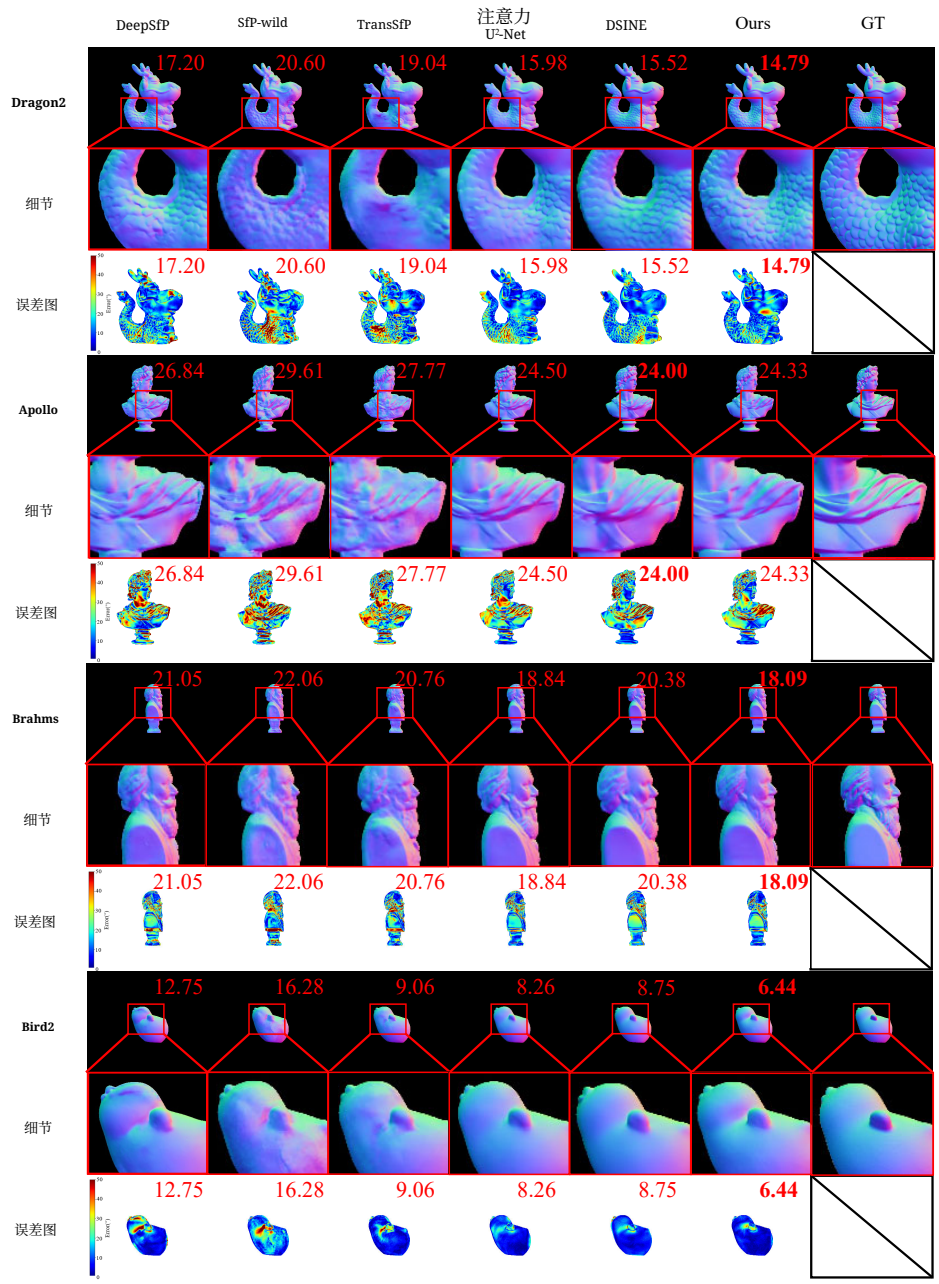

图14：在相同浊度条件（10 ml 乳浊液）下，多种法线估计方法在多个样本上的比较。每组结果依次展示预测法线图、局部放大细节和角度误差热力图。红色数字表示相应方法在各样本上的平均绝对误差（°）。

相比之下，所提出的方法在不同目标物和散射级别下表现出更稳定的几何重建能力。对于 Dragon2 和 Apollo 等细节丰富的样本，该方法能更好地保留局部表面起伏和连续性，避免法线方向发生突变。对于类肖像样本（Brahms）和小尺度目标物（Bird2），与对比方法相比，其预测结果也展现出更高的全局一致性和更精细的细节表现。

误差图进一步佐证了这些观察结果。基线方法的误差主要集中在曲率变化较大的区域，并且常呈现明显的结构化模式。相比之下，所提出的方法整体误差幅值更低，高误差区域显著减少，误差分布也更加均匀。这表明，在统一框架中联合建模去散射与表面法线估计，能够有效减轻水下散射导致的系统性几何推理误差，从而提高法线预测的稳定性和准确性。

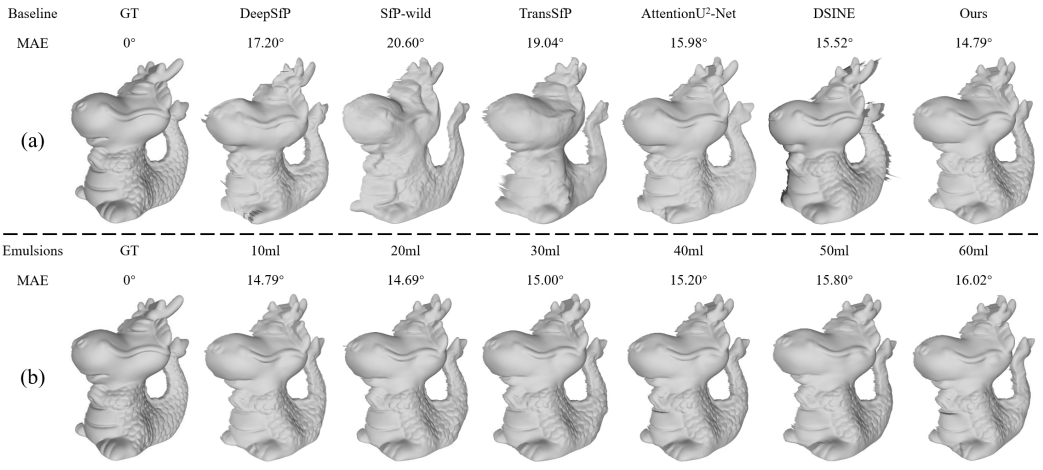

图15：使用法线积分方法，根据预测法线图得到的三维形状积分结果。(a) 10毫升乳液下不同基线方法的积分形状。(b) 不同散射条件下所提出方法的积分形状。

此外，使用 [31] 中的方法从预测法线图恢复三维形状，以展示重建质量。如图 15 所示，较小的平均绝对误差通常对应着更清晰、完整的集成三维形状。所提出的方法在不同散射条件下能够生成更生动的几何形状和更丰富的纹理细节（图 15(b)），表明其具有很强的鲁棒性。然而，角状区域

在该示例的单视图设置中存在深度不连续，这会破坏法向量场的可积性，并导致法线积分出现跳变，进而产生类似悬崖的重建伪影。未来的工作将解决由边界遮挡引起的此类深度不连续问题。

4.3. 消融研究

表3：关键子网络/模块的消融结果。指标为均值/中位数法线角误差（°）；越低越好。

模型	不含 PPN	不含 DN	不含 PPN & DN	不含 CE	不含 DEConv	所提出的方法
Mean	16.72°	15.37°	15.56°	15.46°	23.03°	15.12°
中位数	15.94°	15.38°	16.09°	15.73°	22.48°	15.21°

为验证所提出的结构化学习框架中关键子网络与模块的有效性，在MuS-Polar3D测试集上进行了系统的消融实验。定量结果见表3。实验分别移除偏振参数网络（PPN）、去散射网络（DN）、同时移除PPN与DN，以及颜色嵌入模块（CE）和细节增强卷积（DEConv），并与完整模型进行比较。由此可定量分析各组件对整体三维重建性能的贡献。

首先，移除PPN会使平均角误差从15.12° 增至16.72°，表明仅依据局部去散射线索预测法线，无法充分挖掘偏振信息中包含的全局几何先验。移除DN带来的性能下降较小但仍明显，说明显式的低层去散射建模能为后续散射条件下的几何推理提供更稳定的输入。当同时移除PPN与DN时，模型退化为直接从散射偏振图像进行端到端法线回归；误差仍高于完整模型，证实了联合建模“低层恢复+高层几何”的必要性。

对于建筑模块，移除 CE 会导致法线误差明显增加，表明基于颜色–几何同构的嵌入机制能够提升法线估计的一致性和稳定性。移除 DEConv 时性能下降最为严重：平均角误差急剧上升至 23.03°，这表明显式建模高频

几何细节和方向变化对于准确完成水下 SfP 任务至关重要。总体而言，这些组件在统一框架中相互补充，共同实现最佳性能。

5. 结论

水生环境中的严重散射仍是制约主流水下三维光学成像方法性能的根本挑战。尽管偏振成像在去散射和基于偏振形状恢复（SfP）的三维结构恢复方面展现出无可比拟的双重优势，但现有先进方法仍受限于顺序处理流程：去散射与三维重建被视为两个相互独立的任务，导致误差逐步累积，且无法在后续阶段得到纠正。为解决这一根本局限，本文从全链路成像流程优化的角度重新审视水下三维成像问题，并开发了一种端到端统一框架，实现偏振引导去散射与三维法线重建的联合学习和全局优化。该方法有效抑制了级联流程中的误差传播，并在实际水下场景中较传统方法取得了显著的性能提升。

本文提出 UD-SfPNet，将基于偏振的去散射与偏振驱动的法线估计紧密结合。该框架由偏振参数网络和去散射网络组成，二者共同为法线估计网络提供输入，从而实现低层恢复与高层几何推断的端到端联合建模。此外，本文还引入了两个关键模块以提升性能：CE 模块利用颜色与几何之间的映射关系来增强预测一致性；DEConv 模块则对高频几何特征和方向变化进行显式建模，以更好地保留目标区域中的细微细节。在物理模型的指导下，这些模块与偏振特征共同嵌入网络，使 UD-SfPNet 能够在水下散射条件下生成更清晰的图像和更准确的表面法线预测。

在 MuS-Polar3D 数据集上开展的大量实验表明，UD-SfPNet 具有良好的有效性。该方法在所有对比方法中取得了最低的平均法线角误差（15.12°），表明其重建精度得到显著提升。此外，该方法还表现出较强的鲁棒性和

在不同散射级别和目标属性下的泛化能力。总体而言，本研究将偏振物理与深度学习深度融合，为水下三维成像确立了新范式，并为复杂环境中的水下机器人视觉和海洋探测成像提供了实用解决方案。

关于生成式人工智能及人工智能辅助技术在稿件准备过程中使用的声明

在本研究工作的准备过程中，作者使用了ChatGPT（OpenAI）以提高稿件的可读性和语言质量。在使用该工具/服务后，作者根据需要对内容进行了审阅和编辑，并对所发表文章的内容承担全部责任。

参考文献

[1]E. Organelli, G. Dall'Olmo, R. J. Brewin, G. A. Tarran, E. Boss, A. Bricaud，开阔海域缺失的后向散射源于颗粒的结构复杂性，《自然·通讯》9 (1) (2018) 5439。[2] M. Xiang, F. Liu, J. Liu, X. Dong, Q. Liu, X. Shao，计算光学成像：挑战、机遇、新趋势与新兴应用，《成像前沿》3 (2024年) 1336829。[3] X. Wu, J. Chen, P. Li, X. Wang, J. Wu, F. Huang，基于深度学习的水下目标偏振三维成像方法，《光学快报》33 (2) (2025) 2068–2081。[4] K. Li, J. Liang, Z. Wan, Y. Liu, SFP-Underwater：面向水下散射环境的基于注意力的偏振形状恢复，《光学与激光技术》192 (2025) 113545。[5] P. Wang, K. Yu, H. He, X. Wu，MuS-Polar3D：多重散射条件下计算偏振三维成像的基准数据集，arXiv 预印本 arXiv:2512.21513 (2025)。[6] Y. Y. Schechner, N. Karpel，通过偏振分析恢复水下能见度与结构，《IEEE海洋工程期刊》30 (3) (2006) 570–587。

[7] T. Treibitz, Y. Y. Schechner，主动偏振去散射，IEEE 模式分析与机器智能汇刊31 (3) (2008) 385–399。[8] K. Yang, P. Han, R. Gong, M. Xiang, J. Liu, Z. Fan, T. Xi, F. Liu, B. Wang, X. Shao，通过深度偏振神经网络从散射场景中恢复高质量三维形状，《光学与激光工程》173 (2024年) 107934。[9] O. Ronneberger, P. Fischer, T. Brox，U-Net：用于生物医学图像分割的卷积网络，载于：医学图像计算与计算机辅助干预国际会议，施普林格，2015年，第234–241页。[10] H. Liu, F. Wang, Y. Jin, X. Ma, S. Li, Y. Bian, G. Situ，基于学习的动态散射介质实时成像，《光：科学与应用》13 (1) (2024年) 194。[11] A. H. Mahmoud, M. T. El-Melegy, A. A. Farag，基于偏振和明暗的形状恢复直接方法，载于：2012年第19届IEEE国际图像处理会议，IEEE，2012年，第1769–1772页。[12] D. Miyazaki, M. Kagesawa, K. Ikeuchi，基于一对偏振图像的透明表面建模，IEEE 模式分析与机器智能汇刊26 (1) (2004) 73–82。[13] W. A. Smith, R. Ramamoorthi, S. Tozza，基于未经标定的单目偏振图像进行线性深度估计，载于：欧洲计算机视觉会议，施普林格，2016年，第109–125页。[14] Y. Ba, A. Gilbert, F. Wang, J. Yang, R. Chen, Y. Wang, L. Yan, B. Shi, A. Kadambi，深度偏振形状恢复，载于：欧洲计算机视觉会议，施普林格，2020年，第554–571页。[15] C. Lei, C. Qi, J. Xie, N. Fan, V. Koltun, Q. Chen，面向真实复杂场景的偏振形状恢复，载于：IEEE/CVF 计算机视觉与模式识别会议论文集，2022年，第12632–12641页。

[16] M. Shao、C. Xia、Z. Yang、J. Huang、X. Wang，单视角偏振图像的透明形状，载于：IEEE/CVF国际计算机视觉会议论文集，2023年，第9277–9286页。[17] Y. Lyu、H. Guo、K. Zhang、S. Li、B. Shi，Sfpuel：未知环境光下的偏振形状恢复，神经信息处理系统进展 37（2024年）97184–97202。[18] X. Qin、Z. Zhang、C. Huang、M. Dehghan、O. R. Zaiane、M. Jagersand，U2-Net：采用嵌套U形结构实现更深层次的显著目标检测，模式识别 106（2020年）107404。[19] L. Li、M. Liao、Y. Zhang、Z. Zeng、S. Wang、W. Jia、J. Zhang、B. Zhang、Y. Dong、D. Zhang等，基于偏振复用液晶透镜引导的傅里叶神经网络的单目复振幅成像，《国家科学评论》13（2）（2026年）nwaf561。[20] T. Yamazaki、Y. Maruyama、Y. Uesaka、M. Nakamura、Y. Matoba、T. Terada、K. Komori、Y. Ohba、S. Arakawa、Y. Hirasawa等，采用2.5-μm背照式像素上空气隙线栅的四方向逐像素偏振CMOS图像传感器，载于：2016年IEEE国际电子器件会议（IEDM），IEEE，2016年，第8–7页。[21] X. Wu、P. Li、X. Zhang、J. Chen、F. Huang，基于偏振成像与深度学习的三维形状重建，《传感器》23（10）（2023年）4592。[22] R. Zhang、P. Isola、A. A. Efros、E. Shechtman、O. Wang，深度特征作为感知度量的不合理有效性，载于：IEEE计算机视觉与模式识别会议论文集，2018年，第586–595页。[23] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、Ł. Kaiser、I. Polosukhin，Attention Is All You Need，神经信息处理系统进展 30（2017年）。[24] Z. Zhang、H. Zheng、R. Hong、M. Xu、S. Yan、M. Wang，用于低照度图像增强的深度颜色一致性网络，载于：IEEE/CVF 计算机视觉与模式识别会议论文集，2022年，第1899–1908页。

[25] Z. Chen、Z. He、Z.-M. Lu，DEA-Net：基于细节增强卷积与内容引导注意力的单幅图像去雾，IEEE 图像处理汇刊 33（2024年）1002–1015。[26] Y. Song、Z. He、H. Qian、X. Du，用于单幅图像去雾的视觉变换器，IEEE 图像处理汇刊 32（2023年）1927–1941。[27] Z. Wan、K. Li、J. Liang、Z. Ouyang、Y. Liu，Attentivesfp：利用dualpool-former与注意力机制实现精确的偏振形状恢复，模式识别（2025）112714。[28]A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga等，PyTorch：一种命令式风格的高性能深度学习库，神经信息处理系统进展 32（2019）。[29] E. Rublee、V. Rabaud、K. Konolige、G. Bradski，ORB：SIFT或SURF的高效替代方案，载于：2011年国际计算机视觉会议，IEEE，2011年，第2564–2571页。[30] G. Bae、A. J. Davison，重新思考表面法线估计的归纳偏置，载于：IEEE/CVF 计算机视觉与模式识别会议论文集，2024年，第9535–9545页。[31] X. Cao、H. Santo、B. Shi、F. Okura、Y. Matsushita，双边法线积分，载于：欧洲计算机视觉会议，施普林格，2022年，第552–567页。