

• 四川大学建校 130 周年特刊约稿 •

DOI:10.12454/j.jsuese.202600541



本刊网刊

基于深度学习的低噪声全息图生成方法

李赵松¹, 樊宇博¹, 宋德江¹, 李移隆², 王 迪^{1,2}, 王琼华^{1,2*}

(1. 杭州市北京航空航天大学国际创新研究院(北京航空航天大学 国际创新学院), 浙江 杭州 311115;
2. 北京航空航天大学 仪器科学与光电工程学院, 北京 100191)

摘 要:全息 3D 显示能重建 3D 物体的完整波前信息, 因此, 被视为下一代裸眼 3D 显示的重要发展方向。然而, 传统的全息图生成方法难以兼顾全息重建质量和全息图计算效率, 尤其在散斑噪声抑制方面存在瓶颈。近年来, 深度学习为实现低噪声全息 3D 显示提供了新的解决思路。为此, 本文综述了基于深度学习的低噪声全息图生成方法, 旨在梳理该领域的研究进展、存在挑战与未来发展方向。首先, 简要回顾了传统全息图生成方法的类型及其局限性。然后, 重点介绍了低噪声全息图生成方法的基本原理, 阐明了数据监督方法、模型监督方法和混合监督方法的基本思想。接着, 分别围绕数据监督方法、模型监督方法和混合监督方法, 整理了其最新研究进展。围绕数据监督方法, 研究者验证了深度神经网络在全息图生成任务中的可能性, 引入了复杂的网络结构来生成低噪声全息图, 并尝试了从 2D 目标图像直接生成低噪声全息图的方案; 围绕模型监督方法, 研究者基于物理模型约束, 通过硬件反馈或软件反馈实现了低噪声全息图的生成, 并在彩色全息显示、遮挡关系处理和轻量化网络设计方面取得了突破; 围绕混合监督方法, 研究者通过加权组合数据损失与物理损失, 从多阶段训练策略到非配对学习方法, 再到光学器件与算法的协同设计, 逐步突破了数据监督方法和模型监督方法的性能瓶颈。最后, 针对数据监督方法、模型监督方法和混合监督方法面临的数据依赖性强、计算速度慢、网络性能增益有限等难题, 展望了未来可以采取的解决思路和研究路线。

关键词:全息 3D 显示; 全息图; 深度学习; 低噪声

中图分类号: O438.1

文献标志码: A

文章编号: 2096-3246(2026)05-0252-10

裸眼 3D 显示技术已成为光显示领域产业升级和革新的重要推力, 有望带来显著的经济增益。其中, 全息 3D 显示技术能记录和重建 3D 物体的完整波前信息, 提供观看者能感知的所有深度线索, 因此, 已成为裸眼 3D 显示技术中具有重要应用前景的新型显示技术之一^[1-3]。

全息 3D 显示技术由光学全息术发展而来, 光学全息术最初由 Gabor 发明, 旨在提高电子显微镜的分辨率, 这一突破使其荣获 1971 年诺贝尔物理学奖。基于光的干涉理论, 光学全息术将 3D 物体的波前信息记录到存储介质中, 当存储介质被相干光照射时, 即可获得 3D 物体的全息再现像。目前, 全息 3D 显示技术中广泛使用的是计算全息术。不同于光学全息术, 计算全息术通过计算机模拟光的衍射传播过程来替代

光学记录过程, 从而无须搭建复杂的全息记录系统。经过多年的发展, 计算全息术中的全息图生成方法大体可分为两类: 迭代法与非迭代法。

迭代法通过高迭代次数可逼近高精度解, 但每张全息图需经历数十到数百次的正逆向衍射计算, 耗时较长, 且易陷入局部最优^[4-6]。非迭代法可以直接合成全息图, 速度较快, 但散斑噪声难以得到有效抑制^[7]。针对 3D 物体的不同描述形式, 现有的非迭代法可以分为点元法、面元法和分层法。在计算 3D 物体的全息图时, 点元法的计算量较大, 面元法的衍射计算过程复杂, 而分层法则难以处理遮挡问题。目前, 传统全息图生成方法仍面临全息图计算效率与全息重建质量难以权衡的困境, 无法满足低噪声实时全息 3D 显示的需求^[8]。

收稿日期: 2026-06-20 修回日期: 2026-08-09 网络出版日期: 2026-08-20

基金项目: 国家自然科学基金项目(62535001); 北京市自然科学基金项目(4262080); 中国博士后科学基金项目(BX20250478)

作者简介: 李赵松(1998—), 男, 博士后。研究方向: 全息 3D 显示。E-mail: lzs1998@buaa.edu.cn

* 通信作者: 王琼华, 教授, E-mail: qionghua@buaa.edu.cn

近年来,计算机算力的提升推动了深度学习技术在人工智能各领域的应用,也促进了计算全息术与深度学习技术的融合。根据在深度神经网络(DNN)训练过程中约束条件的不同,基于深度学习的低噪声全息图生成方法一般可分为3类:数据监督方法、模型监督方法和混合监督方法。这些方法以大量全息实验数据为基础,通过定义全息图与3D物体之间的损失函数关系,完成对DNN的训练。训练后的DNN能快速生成高质量全息图,在解决全息显示质量与全息图计算速度的权衡问题方面已取得初步成效,是实现低噪声实时全息3D显示的潜在途径^[9]。然而,数据监督方法、模型监督方法和混合监督方法具有不同的适用条件和局限性。为推动全息3D显示技术的进一步发展,亟需梳理基于深度学习的低噪声全息图生成方法的最新进展、存在挑战及未来发展趋势。

针对上述问题,首先,本文介绍了数据监督方法、模型监督方法和混合监督方法的基本原理。然后,比较了3类方法在抑制散斑噪声、提升全息再现像质量等方面的优势与不足。最后,展望了3类方法在全息3D显示应用中面临的难题及未来的发展方向。

1 低噪声全息图生成方法的基本原理

在全息图的重建过程中,由于相干光的广泛使用,通常会产生显著影响成像质量的散斑噪声。特别是当3D物体的全息图复振幅分布中包含了随机相位时,散斑噪声将变得难以约束。抑制散斑噪声的关键在于平滑全息图的相位分布。由于DNN在特定的网络结构和训练方法下会表现出低通滤波器的性质,因此,通过DNN生成的全息图通常能较好地抑制散斑噪声。

DNN在输入层与输出层之间设有多个隐藏层,通过可学习的权重参数和非线性激活函数来完成从输入到输出的映射^[10-11]。本质上,DNN就是一种由多个自适应函数层组成的能够处理复杂任务的非线性模型。训练DNN的根本目标在于寻找能使预测的输出值与目标值偏差最小的网络参数。寻找网络参数是为了解决高维优化问题,即给定数据集 (X, Y) ,DNN将输入 X 映射到预测值 $\hat{Y}$,随后通过优化DNN的参数使输出值与目标值 Y 相匹配,该过程中预测值 $\hat{Y}$ 、优化后的输出值 Y_{opt} 可分别描述为:

$$\hat{Y} = f_{\text{DNN}}(X) \quad (1)$$

$$Y_{\text{opt}} = \underset{Y}{\text{minimize}} \mathcal{L}(\hat{Y}, Y) \quad (2)$$

式(1)、(2)中, f_{DNN} 为DNN的参数, $\text{minimize } \mathcal{L}(\cdot)$ 表示求损失函数 $\mathcal{L}$ 的最小值。

数据监督方法的基本思想是利用DNN从大量的“目标图像-目标全息图”的配对数据中学习3D物体和低噪声全息图间的映射关系。相比于传统的全息图迭代生成方法,数据监督方法能显著缩短全息图的推理时间。基于数据监督方法生成全息图的流程图如图1所示。

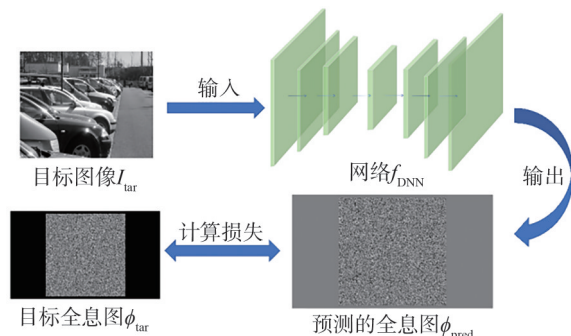


图1 基于数据监督方法生成全息图的流程图

Fig. 1 Flowchart of the data-supervised-based method for hologram generation

在训练阶段,DNN以目标全息图像 I_{tar} 作为输入,预测对应的全息图,并通过损失函数衡量预测结果与目标值之间的差异,进而借助反向传播算法优化DNN的参数。上述过程得到的DNN预测全息图 ϕ_{pred} 、训练完成后DNN输出的全息图 ϕ_{out} 可分别描述为:

$$\phi_{\text{pred}} = f_{\text{DNN}}(I_{\text{tar}}) \quad (3)$$

$$\phi_{\text{out}} = \underset{\phi}{\text{minimize}} \mathcal{L}(\phi_{\text{pred}}, \phi_{\text{tar}}) \quad (4)$$

训练后的DNN会偏向于生成包含训练数据集中广泛存在的低散斑噪声且平滑的相位/强度分布,从而实现低噪声全息图的快速预测,并有效避免传统全息图生成方法的高计算开销。

数据监督方法普遍采用卷积神经网络(CNN)进行训练,如U-Net。值得注意的是,数据监督方法的散斑噪声抑制能力依赖于训练数据集对低噪声样本的覆盖;当训练数据充足时,该方法能实现对散斑噪声的良好抑制。因此,该方法的泛化能力受限于训练数据的分布范围,而生成高质量训练数据集通常需要耗费大量时间。

基于模型监督方法生成全息图的流程图如图2所示。与数据监督方法不同的是,模型监督方法不依赖预先生成的训练数据集,其基本思想是在DNN的训练过程中嵌入可微的衍射传播方程等物理模型,在DNN输出全息图后,基于物理模型计算全息再现像与目标图像之间的损失函数,进而通过反向传播算法优化DNN参数。上述过程中,通过训练完成后DNN输出的全息图 ϕ_{out} 得到的全息再现像 I_{pred} 、优化后的全息图 ϕ_{opt} 可分别描述为:

$$I_{\text{pred}} = f_{\text{prop}}(\phi_{\text{out}}) = f_{\text{prop}}(f_{\text{DNN}}(I_{\text{tar}})) \quad (5)$$

$$\phi_{\text{opt}} = \underset{\phi}{\text{minimize}} \mathcal{L}(I_{\text{pred}}, I_{\text{tar}}) \quad (6)$$

式(5)中, f_{prop} 为使用的衍射传播模型。

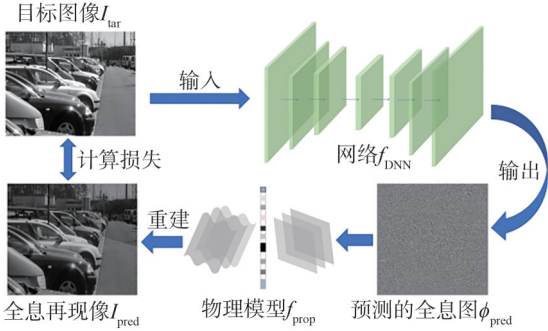


图2 基于模型监督方法生成全息图的流程图

Fig. 2 Flowchart of the model-supervised-based method for hologram generation

由此可以发现,模型监督方法的训练过程由物理模型提供显式约束,在无需或仅需少量配对数据的情况下,就能完成从目标图像到全息图的无监督或自监督学习,有效避免了数据监督方法对高质量训练数据集的依赖。同时,相比于传统的衍射传播方法,模型监督方法将全息图的衍射传播过程可学习化,实现了对散斑噪声的自动优化。

由于通过全息再现像与目标图像来计算损失函数,模型监督方法能够在物理模型的约束下产生具备物理一致性的低噪声全息图,并有效抑制在光学传播过程中产生的散斑噪声。然而,模型监督方法受限于物理模型的精度需求,并且计算损失函数时需反复执行衍射传播计算。因此,当3D物体的数据量较大时,模型监督方法的训练效率相对较低。数据监督方法和模型监督方法的特性对比如表1所示。

表1 数据监督方法和模型监督方法的特性对比

Tab. 1 Comparison of the characteristics of data-supervised and model-supervised methods

特性	数据监督方法	模型监督方法
约束条件	全息图数据集	衍射传播模型
对数据集依赖程度	高	低
计算效率	可实时	可实时
重建质量	高	高
泛化能力	弱	强
可解释性	低	高

混合监督方法的基本思想是将数据监督方法与模型监督方法相结合,从而兼具二者的优势。基于混合监督方法生成全息图的流程图如图3所示。在DNN的训练过程中,混合监督方法同时引入了两类约束:一方面,利用DNN生成的全息图与预先准备的数据集

构建数据损失;另一方面,将DNN预测的全息再现像与目标图像之间的差异作为物理损失;两者加权构成总损失函数。通过这种加权组合方式,既能利用数据先验避免过度平滑,又可通过显式约束保证物理一致性,从而兼顾散斑噪声的抑制与全息重建细节的保真,这是传统的全息图迭代生成方法和衍射传播方法所无法实现的。通过混合监督方法,既能使DNN从数据集中学习映射关系,加快DNN的收敛速度,又能借助物理模型约束抑制散斑噪声与伪影,提升全息再现像质量和DNN的泛化能力。在推理阶段,混合监督方法对散斑噪声的抑制效果一般优于数据监督方法和模型监督方法。

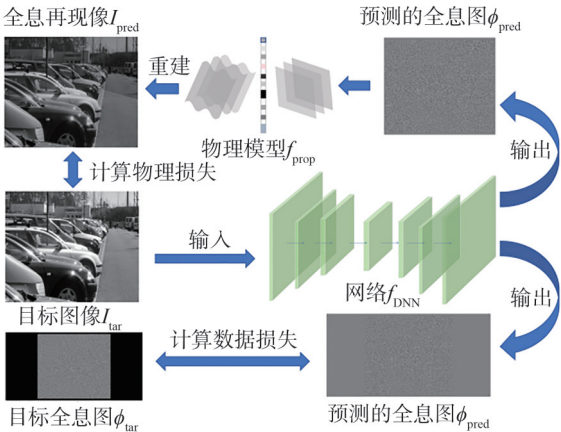


图3 基于混合监督方法生成全息图的流程图

Fig. 3 Flowchart of the hybrid-supervised-based method for hologram generation

2 低噪声全息图生成方法

近年,研究者使用数据监督方法、模型监督方法和混合监督方法在抑制散斑噪声方面均取得诸多进展,下面对3类方法逐一展开综述,并比较、分析各自的特点,以更清晰地呈现其研究思路。

2.1 数据监督方法

早期的数据监督方法主要是验证DNN生成全息图的可行性。例如:在二值振幅型全息图生成任务中,Goi等^[12]提出了基于深度学习的全息图生成方法,通过优化DNN直接输出二值振幅型全息图,有效规避了全息图的质量退化问题。为生成训练DNN的可靠数据集,Lee等^[13]提出了控制随机点密度和组合基本图形来构建训练数据集的策略,并在多深度全息图的CNN训练任务中对该策略进行了验证,其显著提升了全息图的重建质量,也为后续研究提供了生成高质量数据集的思路。基于上述构建全息图训练数据集的基本思想,Zheng等^[14]提出了一种基于多平面角谱迭代的CNN训练方法,并利用峰值信噪比和结构相似度等指

标验证了该方法不仅具有与传统的迭代法相当的精度,还能大幅缩短全息图计算时间。上述文献中的方法验证了DNN能用于全息图的生成任务,然而,受限于网络架构和训练策略的不成熟,上述方法的全息图重建质量不够高。

为了进一步提高全息图的重建质量,Khan等^[15]在全息图生成任务中使用了生成对抗网络(GAN),建立了GAN-Holo模型,并通过菲涅耳衍射(Fresnel diffraction)法制作训练数据集;然而,该方法存在训练数据合成复杂和训练不够稳定的问题。为此,Shi等^[16]提出了基于残差神经网络(ResNet)的张量全息术,该方法能从单张RGB-D图像实时合成彩色3D全息图,可在消费级GPU上以60 Hz的速度生成全息图,该方法的实验效果如图4(a)所示。使用峰值信噪比(PSNR)与结构相似性(SSIM)作为评价指标。



(a) 张量全息术的实验效果^[16]



(b) DEHNet的实验效果^[17]

图4 文献[16–17]数据监督方法的低噪声全息3D显示效果
Fig.4 Low-noise holographic 3D display effects of the data-supervised methods in Refs. [16–17]

虽然文献[16]方法能有效提高全息再现像的质量,但是仍存在深度渲染不准确的问题。为此,Yang等^[17]提出了衍射工程全息网络(DEHNet),通过网络优化波前使全息图在多平面传播后能模仿非相干光的自然离焦模糊,极大提升了全息再现像的真实感,该方法的实验效果如图4(b)所示。

文献[16–17]方法需要借助渲染器生成3D训练数据集,渲染过程往往耗时复杂,为此,Chang等^[18]提出了从立体图像到直接生成全息图的SHNet,免去了耗时的3D物体渲染过程;Liu等^[19]提出了全息调制器概念和通道式变分自编码器,通过学习输入全息图的解耦隐空间表征来生成特定3D物体的全息图;Yolalmaz等^[20]提出了CHoloNet,不仅能够多个观测平面和颜色通道上生成全息图,还具备从单幅全息图中直接恢复物体信息的能力。

为了进一步简化训练数据集,基于2D图像直接生成全息图的方案成为新的研究趋势。Chang等^[21]提出了一种从单张2D图像合成全息图的端到端深度学习方法,该方法先将高质量的2D图像转化为2D图像-目标全息图的数据集,再监督训练CNN;其计算和存储效率极高,为实时全息3D显示开辟了新的发展路径。

表2给出了典型的数据监督方法的特性。这些方法大多通过自行合成的训练数据集完成目标图像到全息图的映射,隐式地完成平滑相位/强度分布的学习,从而有效抑制了全息再现像中的散斑噪声。

需要注意的是,数据监督方法高度依赖大规模高质量训练数据集的构建,网络的泛化能力和模型性能与训练数据集质量密切相关。同时,高质量的训练数据集通常需要满足目标图像与全息图准确配对,且配对内容满足光学传播规律、数据量大等条件,这无疑增加了训练数据集的制作难度。未来,如何高效地融合衍射传播模型与数据监督法或是突破训练数据集制作难的关键。

表2 典型的数据监督方法的特性

方法	网络类型	损失函数	数据集
SHNet ^[17]	ResNet	均方误差损失	SceneFlow
CVAE ^[18]	CNN+全连接层	交叉熵损失	自建3D物体
DEHNet ^[19]	CNN+ResNet	均方误差损失	3 000个3D场景
CHoloNet ^[20]	CNN	相关性系数	51 200幅光学全息图
ResNet-based ^[21]	ResNet	均方误差损失+平均绝对误差损失	自建3D物体

2.2 模型监督方法

为了规避制作高质量训练数据集,研究者尝试使用物理模型来驱动网络参数的更新。为此,Peng等^[6]提出了相机在环训练策略,其实验效果如图5(a)所示。文献[6]方法在生成全息图的物理模型与真实硬件之

间直接建立反馈回路,利用相机拍摄光学全息重建结果,并通过与目标图像进行比较计算损失,因此在自动校准系统中不可避免会引入像差;此外,该方法依赖于复杂的硬件反馈系统,成本较高且难以部署至不同的平台。

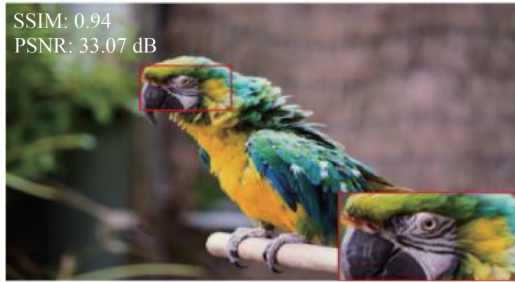
(a) 相机在环训练策略的实验结果^[6](b) IncepHoloRGB的实验结果^[24]

图5 文献[6,24]模型监督方法的低噪声全息3D显示效果
Fig. 5 Low-noise holographic 3D display effects of the model-supervised methods in Refs. [6,24]

为克服上述问题,研究者转向在软件层面实现模型监督方法。Qin等^[22]提出了一种复值生成对抗网络CV-GAN,用于实时生成高质量纯相位全息图;该方法中设计了一种复值U-Net,能够直接处理复振幅信息,有效提高了全息图的学习效率;此外,该方法还提出了复值亚像素卷积上采样策略与对抗学习机制来增强网络的拟合能力。Fang等^[23]提出了一种基于模型监督的训练网络SST-holo,该网络融合了Res-MSR模块以自适应学习多尺度图像特征,可在0.015 s内生成2K分辨率的高质量全息图。为实现彩色3D全息显示,Yu等^[24]提出了轻量化多波长网络模型IncepHoloRGB,该方法基于Inception采样模块和多波长传播模块,能够以191 fps的速度生成高质量彩色全息图,其实验效果如图5(b)所示。

然而,文献[22-24]中U-Net等传统网络架构在表征复杂特征方面仍存在局限。为此,Fu等^[25]提出了复值分层多融合神经网络CHMFNet,在U-Net框架中引入了复值多层感知机模块以增强网络对复杂特征的提取能力,并设计了复值分层多融合模块来实现多尺度分层特征融合;该方法实现了峰值信噪比达到34.11 dB的全息显示效果。Liu等^[26]从频域的角度切入,针对CNN在频谱学习中的低频偏置问题,提出了基于多级小波通道注意力的频率感知网络MW-CANet,通过多尺度小波变换独立捕获低频与高频特

征,并利用通道注意力机制聚焦高频分量,有效提升了全息重建质量。

此外,研究者也在优化网络结构和减小网络参数量方面做出了尝试。Yu等^[27]提出了一种非端到端的非对称神经网络,将全息图生成任务分解为相位分布预测和编码两个子任务,并根据各子任务的特性分别设计了两个子网络,而非采用传统的对称结构;该方法在保持网络轻量化的同时显著增强了网络的拟合能力,兼顾了全息图的生成速度和全息重建质量。Zhong等^[28]提出了一种仅含938个参数的轻量CNN用于全息图编码,该方法可在18 ms内生成单张4K全息图,每增加一个深度层推理时间仅增加12 ms,在160~210 mm的深度范围内,全息再现像的平均峰值信噪比均高于30 dB。

为进一步解决模型监督方法中多深度全息再现像的遮挡关系不准确的问题,Yan等^[29]提出了一种基于CNN的多深度全息图生成方法,设计了正反向衍射模型来计算多深度衍射场,并引入逐层替换方法来处理3D物体的遮挡关系,从而生成具有正确遮挡效果的全息图,全息再现像的峰值信噪比可达到31.8 dB。在此基础上,Zhang等^[30]进一步提出了基于残差块的复值卷积神经网络ResC-CNN,该方法在使用衍射模型监督训练的同时,引入分层图像-深度数据集以提升网络对3D物体的信息预测精度,在显著提高多深度全息图实时生成帧率的同时增强了全息再现像的保真度。

除了改进网络结构外,研究者还将模型监督方法与定制化的硬件相结合,构建面向真实3D场景的全息显示系统。Wang等^[31]提出了一种基于液体透镜的全息相机,其计算端采用端到端物理模型驱动网络EEPMD-Net,以真实3D场景信息作为输入,通过比较深度融合后的全息再现像与真实3D场景的强度差异来构建复合损失函数进行网络优化,实现了具有真实深度信息的高保真全息图生成。在此基础上,该团队进一步提出了实时全息相机^[32],在计算端设计了一种焦点堆栈网络FS-Net,通过焦点堆栈渲染器和可学习泽尼克(Zernike)相位进行训练,使全息图的计算时间缩短至13 ms,解决了全息再现像质量低和离焦模糊失真等难题。

模型监督方法通过将可微物理模型嵌入全息图的生成过程,并利用损失函数直接在全息再现像的强度上惩罚高方差分量。该机制能有针对性地抑制全息图中经衍射传播后产生的散斑噪声,同时,以物理一致性的方式提高全息再现像的保真度。

表3总结了近年来具有代表性的模型监督方法的特性。这些方法通过DNN来学习光的衍射传播过程,相比于数据监督方法具有全息重建质量高和可解释

性强的优势。然而,将物理模型融入网络架构会显著增加计算开销;此外,当前的主流方案主要采用平面光传播算法,全息图的计算被限制在固定视角的2D切片上,无法呈现具有多视角的全息3D显示效果。如

表3 典型的模型监督方法的特性

Tab. 3 Characteristics of typical model-supervised methods			
方法	网络类型	损失函数	数据集
MW-CANet+CCNN ^[26]	CNN	均方误差损失	DF2K
ResC-CNN ^[28]	CNN+ResNet	多尺度结构相似性损失	MIT-CGH-4K-V2
SST-holo ^[29]	CNN	负皮尔逊系数	MIT-CGH-4K
FS-Net ^[31]	CNN	多尺度结构相似性损失	DIV2K
EEPMD-Net ^[32]	CNN	多尺度结构相似性损失	CREStereo+DIV2K

2.3 混合监督方法

混合监督方法通过有机融合数据监督方法与模型监督方法,在避免数据依赖的同时可提升全息重建质量。为了突破网络模型的性能上限,Shi等^[33]提出了一种监督训练和无监督训练相结合的两阶段训练策略TensorHolo v2,用于生成高质量3D全息图;该方法构建了一个高质量的目标数据集MIT-CGH-4K-V2,先利用目标数据集进行预训练,再通过衍射模型进行无监督微调,突破了数据监督方法对全息重建质量的限制;该方法的实验效果如图6(a)所示。

何进一步提升全息图的生成速度与全息再现像质量、提升网络对不同场景的泛化能力,以及实现更自然的多深度遮挡与彩色显示效果,仍是模型监督方法需要继续关注的问题。

从而在紧凑的可穿戴设备中实现了高质量的彩色3D全息AR显示效果。

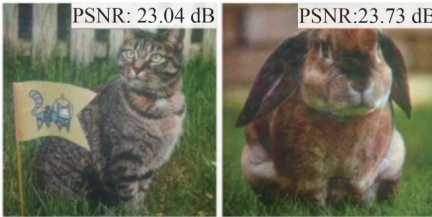
然而,文献[33–34]方法仍需要一定规模的配对数据集参与训练。为了降低网络模型对配对数据集的依赖,Sun等^[35]提出了一种基于CycleGAN的非配对学习方法,用于复振幅重建与全息图生成;该方法利用循环一致性损失,在无需配对数据的条件下学习物体与全息图之间的映射关系,训练后的网络不仅能够完成高质量的全息重建,还同时学习到了全息成像系统的物理特性。在此基础上,Zhang等^[36]提出了一种基于循环一致物理约束的非相干自干涉数字全息自适应重建网络,采用双输入全连接架构进行自适应特征优化,并通过嵌入非线性解卷积的循环一致性损失来强化物理一致性约束,实现了在数据稀缺条件下的高质量全息重建。

此外,研究者还尝试利用混合监督方法对衍射传播模型进行误差补偿。Yuan等^[37]提出了一种衍射传播误差补偿网络;该网络通过预测残差值来校正传播误差,使衍射过程逼近理想状态,从而抑制了吉布斯现象而导致的振铃伪影;将该网络应用于HoloNet^[6]和复值卷积神经网络(CCNN)^[28]时,全息再现像的峰值信噪比分别达到32.47 dB和29.53 dB。类似地,Tong等^[38]提出了U型残差密集网络U-RDN,通过深度学习方法获得低空间相干条件下的最优全息图生成方案;该方法使用光学实验生成大规模图像数据,并通过深度学习进行分析与复原,在其他数据集上展现出较强的鲁棒性。

混合监督方法通过加权组合数据损失与物理损失,能在有效抑制高频散斑噪声的同时,实现高保真的全息再现像,在数据稀缺或传播模型存在误差时,仍能表现出较强的鲁棒性。表4总结了近年来具有代表性的混合监督方法的特性。目前,基于混合监督的方法通过融合多种训练策略、非凸优化算法、引入弱监督学习范式等途径,在突破单一方法性能瓶颈方面



(a) TensorHolo v2的实验效果^[33]



(b) 彩色3D全息AR显示效果^[34]

图6 文献[33–34]混合监督方法的低噪声全息3D显示效果
Fig. 6 Low-noise holographic 3D display effects of the hybrid-supervised methods in Refs. [33–34]

进一步地,研究者将混合监督思想拓展至光学系统与算法的联合优化层面,Gopakumar等^[34]在Nature期刊上发表了一种彩色3D全息AR显示系统,该系统将彩色超表面光栅、紧凑型色散补偿波导结构与AI驱动的全息算法进行协同设计,消除了空间光调制器与波导之间对笨重准直光学器件的依赖;如图6(b)所示,该系统将波导模型与相机在环训练策略相结合,

展现了巨大潜力。然而,多阶段训练的复杂度、混合监督中的协调机制及非配对方法在极端条件下的稳定性,仍是该类方法需要进一步解决的问题。未来,随着

超构表面等新型光学元件与 AI 算法的深度协同设计,混合监督方法有望成为实现紧凑型全息 3D 显示系统的重要技术路径。

表 4 典型的混合监督方法的特性

Tab. 4 Characteristics of typical hybrid-supervised methods

方法	网络类型	损失函数	数据集
超表面全息波导 ^[34]	物理波导模型+CNN	相机反馈损失	自建实验数据
CycleGAN ^[35]	GAN	对抗损失	MNIST+自建实验数据
IHAR-Net ^[36]	CNN+全连接层	循环一致性损失	DIV2K+自建实验数据
误差补偿网络 ^[37]	复值 CNN	均方误差损失	DIV2K

3 结论与展望

理想的全息 3D 显示对全息图的计算速度和重建质量都提出了很高的要求,深度学习为全息 3D 显示的发展提供了强大助力,但仍面临以下难题:

1)在数据监督方法中,DNN 的训练效率低,对训练数据集依赖性强;

2)在模型监督方法中,全息图生成速度远未达到实时交互需求;

3)在混合监督方法中,单纯拓展 DNN 的深度或宽度带来的增益越来越有限。

虽然,第 2 节中介绍的部分研究尝试通过小批量数据集、并行加速硬件和多阶段训练策略来缓解上述问题,但这些方法仍然难以兼顾全息再现像的保真度。

为解决这些难题,未来的研究可以考虑从以下路线推进:

1)针对数据监督方法中训练效率低、对数据集依赖性强难题,可运用迁移学习方法,将预训练从大量自然图像或衍射传播模型的通用视觉 DNN 迁移到全息图生成任务中,使用少量全息图数据进行微调,便可在缩短训练时间的同时降低对数据集的依赖;

2)针对模型监督方法中全息图计算速度慢的难题,可以设计轻量级 DNN,借助知识蒸馏、模型剪枝、高效卷积算子与稀疏化表征等,在保持全息图计算精度的同时降低模型的复杂度,从而显著提升全息图生成速度;

3)针对混合监督方法中 DNN 性能增益有限,全息再现像质量难以提高的问题,可以探索新型 DNN 架构,如将 Transformer 模型、自注意力机制和扩散模型等引入全息图生成任务,突破传统 CNN 的性能边界。深度学习是推动全息 3D 显示技术从实验室走向应用的关键,随着全息图的训练效率、DNN 的推理速度与全息再现像质量的提升,全息 3D 显示技术有望带动医学成像、工业检测、教育与娱乐等产业的变革。

参考文献:

[1] Rivenson Y,Wu Yichen,Ozcan A.Deep learning in holography and coherent imaging[J].Light(Science & Applications),2019,8:85.

[2] Liu Kexuan,Wu Jiachen,He Zehao,et al.Progress of learning-based computer-generated holography[J].Chinese Journal of Liquid Crystals and Displays,2023,38(6):819–828.[刘珂瑄,吴佳琛,何泽浩,等.基于深度学习的计算全息显示进展[J].液晶与显示,2023,38(6):819–828.]

[3] Cao Liangcai,He Zehao,Liu Kexuan,et al.Progress and challenges in dynamic holographic 3D display for the metaverse (Invited) [J]. Infrared and Laser Engineering, 2022, 51(1): 20210935.[曹良才,何泽浩,刘珂瑄,等.元宇宙中的动态全息三维显示:发展与挑战[J].红外与激光工程,2022,51(1):20210935.]

[4] Gerchberg R W,Saxton W O.A practical algorithm for the determination of phase from image and diffraction plane pictures[J].Optik,1972,35:237–246.

[5] Chakravarthula P,Peng Yifan,Kollin J,et al.Wirtinger holography for near-eye displays[J].ACM Transactions on Graphics,2019,38(6):1–13.

[6] Peng Yifan,Choi S,Padmanaban N,et al.Neural holography with camera-in-the-loop training[J].ACM Transactions on Graphics,2020,39(6):1–14.

[7] Hsueh C K,Sawchuk A A.Computer-generated double-phase holograms[J].Applied Optics,1978,17(24):3874–3883.

[8] Wang Di,Li Zhaocong,Huang Qian,et al.Fast generation technology of computer-generated hologram[J].Journal of Signal Processing,2022,38(9):1863–1871.[王迪,李赵松,黄倩,等.计算全息图的快速生成技术[J].信号处理,2022,38(9):1863–1871.]

[9] Zhang Yixin,Zhang Mingkun,Liu Kexuan,et al.Progress of the computer-generated holography based on deep learning[J].Applied Sciences,2022,12(17):8568.

[10] Yu Xuan,Zhang Haomiao,Zhao Zhe,et al.On the use of deep learning for computer-generated holography[J].iScience,2025,

- 28(5):112507.
- [11] Shimobaba T, Blinder D, Birnbaum T, et al. Deep-learning computational holography: A review[J]. *Frontiers in Photonics*, 2022, 3: 854391.
- [12] Goi H, Komuro K, Nomura T. Deep-learning-based binary hologram[J]. *Applied Optics*, 2020, 59(23): 7103–7108.
- [13] Lee J, Jeong J, Cho J, et al. Deep neural network for multi-depth hologram generation and its training strategy[J]. *Optics Express*, 2020, 28(18): 27137.
- [14] Zheng Huadong, Hu Jianbin, Zhou Chaojun, et al. Computing 3D phase-type holograms based on deep learning method[J]. *Photonics*, 2021, 8(7): 280.
- [15] Khan A, Zhijiang Zhang, Yu Yingjie, et al. GAN-holo: Generative adversarial networks-based generated holography using deep learning[J]. *Complexity*, 2021, 2021: 6662161.
- [16] Shi Liang, Li Beichen, Kim C, et al. Towards real-time photorealistic 3D holography with deep neural networks[J]. *Nature*, 2021, 591(7849): 234–239.
- [17] Yang D, Seo W, Yu H, et al. Diffraction-engineered holography: Beyond the depth representation limit of holographic displays[J]. *Nature Communications*, 2022, 13: 6012.
- [18] Chang Chenliang, Wang Di, Zhu Dongchen, et al. Deep-learning-based computer-generated hologram from a stereo image pair[J]. *Optics Letters*, 2022, 47(6): 1482–1485.
- [19] Liu Shengchi, Chu Daping. Deep learning for hologram generation[J]. *Optics Express*, 2021, 29(17): 27373.
- [20] Yolalmaz A, Yüce E. Comprehensive deep learning model for 3D color holography[J]. *Scientific Reports*, 2022, 12: 2487.
- [21] Chang Chenliang, Dai Bo, Zhu Dongchen, et al. From picture to 3D hologram: End-to-end learning of real-time 3D photorealistic hologram generation from 2D image input[J]. *Optics Letters*, 2023, 48(4): 851–854.
- [22] Qin Haifeng, Han Chao, Shi Xuan, et al. Complex-valued generative adversarial network for real-time and high-quality computer-generated holography[J]. *Optics Express*, 2024, 32(25): 44437.
- [23] Fang Qiwei, Zheng Huadong, et al. Diffraction model-driven neural network with semi-supervised training strategy for real-world 3D holographic photography[J]. *Optics Express*, 2024, 32(26): 45406.
- [24] Yu Xuan, Teng Zhilin, Fan Xuhao, et al. IncepHoloRGB: Multi-wavelength network model for full-color 3D computer-generated holography[J]. *Opto-Electronic Advances*, 2025, 8(10): 250130.
- [25] Fu Jiahui, Wan Wenqiang, Wang Yunrui, et al. High-quality hologram generation based on a complex-valued hierarchical multi-fusion neural network[J]. *Chinese Optics Letters*, 2025, 23(9): 090501.
- [26] Liu Qingwei, Chen Jing, Yao Yongwei, et al. Frequency aware high-quality computer-generated holography via multilevel wavelet learning and channel attention[J]. *Optics Letters*, 2024, 49(19): 5559–5562.
- [27] Yu Guangwei, Wang Jun, Yang Huan, et al. Asymmetrical neural network for real-time and high-quality computer-generated holography[J]. *Optics Letters*, 2023, 48(20): 5351.
- [28] Zhong Chongli, Sang Xinzhu, Yan Binbin, et al. Real-time 4K computer-generated hologram based on encoding conventional neural network with learned layered phase[J]. *Scientific Reports*, 2023, 13: 19372.
- [29] Yan Xingpeng, Liu Xinlei, Li Jiaqi, et al. Generating multi-depth 3D holograms using a fully convolutional neural network[J]. *Advanced Science*, 2024, 11(28): 2308886.
- [30] Zhang Yuanzhe, Cheng Dewen, Beijing N N L, et al. Real-time multi-depth holographic display using complex-valued neural network[J]. *Optics Express*, 2025, 33(4): 7380.
- [31] Wang Di, Li Zhaosong, Zheng Yi, et al. Liquid lens based holographic camera for real 3D scene hologram acquisition using end-to-end physical model-driven network[J]. *Light: Science & Applications*, 2024, 13: 62.
- [32] Li Zhaosong, Liu Chao, Li Xiaowei, et al. Real-time holographic camera for obtaining real 3D scene hologram[J]. *Light: Science & Applications*, 2025, 14: 74.
- [33] Shi Liang, Li Beichen, Matusik W. End-to-end learning of 3D phase-only holograms for holographic display[J]. *Light: Science & Applications*, 2022, 11: 247.
- [34] Gopakumar M, Lee G Y, Choi S, et al. Full-colour 3D holographic augmented-reality displays with metasurface waveguides[J]. *Nature*, 2024, 629(7911): 791–797.
- [35] Sun Xiaotian, Xiong Chencen, Ren Zhenbo, et al. Unpaired learning for digital holographic reconstruction and generation[J]. *Optics Express*, 2025, 33(2): 3414.
- [36] Zhang Wanbin, Feng Yijian, Ren Yanchen, et al. Cycle-consistent physical constraint deep learning for incoherent self-interference digital holographic reconstruction[J]. *Optics and Lasers in Engineering*, 2026, 196: 109385.
- [37] Yuan Ganzhangqin, Zhou Mi, Peng Yifan, et al. Error-compensation network for ringing artifact reduction in holographic displays[J]. *Optics Letters*, 2024, 49(11): 3210–3213.
- [38] Tong Xin, Xu Renjun, Liu Ke, et al. A deep-learning approach for low-spatial-coherence imaging in computer-generated holography[J]. *Advanced Photonics Research*, 2023, 4: 2200264.

【校友简介】



王琼华,教育部高端人才、国家杰出青年科学基金获得者、国家“万人计划”科技创新领军人才,国际信息显示学会、美国光学学会、国际光学工程学会、中国光学学会、中国光学工程学会会士, *Photonix* 等期刊编委及国家重点研发

计划项目负责人,入选爱思唯尔与斯坦福大学“全球前 2% 顶尖科学家和中国高被引学者榜单”。研究方向为显示与成像技术,研制了裸眼集成光场 3D 显示器、全息 3D 显示系统、液体透镜和连续光学变焦显微镜等。发表 SCI 收录论文约 400 篇,出版著作 3 部,获准美国和中国发明专利近 200 件。以第一完成人荣获 2023 年度国家技术发明二等奖和 3 项省部级/学会科技一等奖,荣获“全国三八红旗手”称号。

Deep Learning-based Methods for Low-noise Hologram Generation

LI Zhaosong¹, FAN Yubo¹, SONG Dejiang¹, LI Yilong², WANG Di^{1,2}, WANG Qionghua^{1,2*}

(1.Hangzhou International Innovation Institute, Beihang University, Hangzhou 311115, China;

2.School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China)

Abstract:

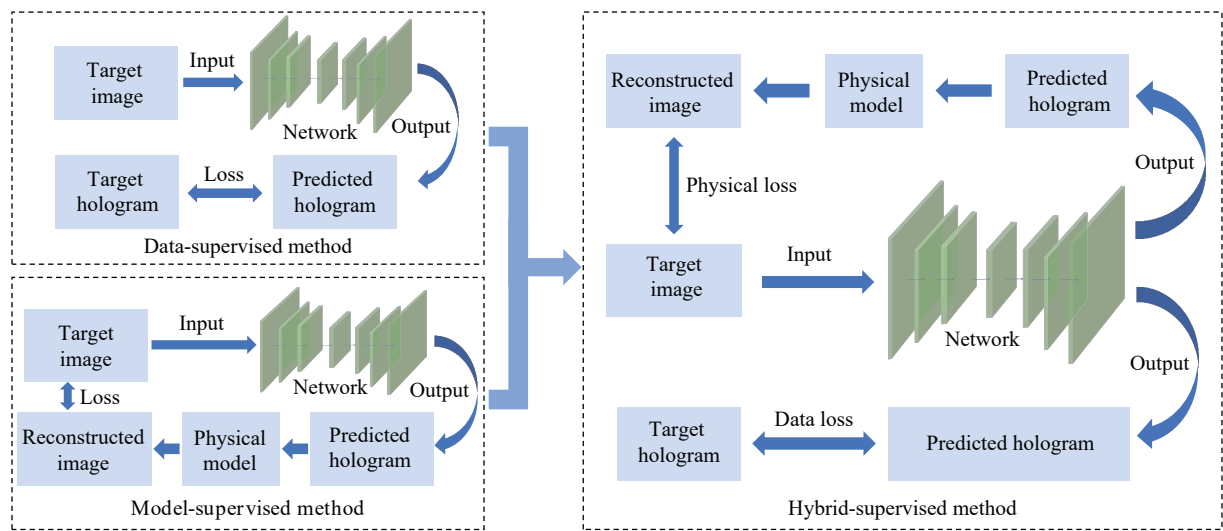
Significance Holographic 3D displays reconstruct the complete wavefront information of 3D objects and are therefore regarded as a key direction in the development of next-generation naked-eye 3D displays. However, traditional hologram generation methods struggle to achieve a balance between holographic reconstruction quality and computational speed. Among these challenges, the suppression of speckle noise remains a central issue in holographic 3D displays. Arising from coherent illumination and random phase distributions, speckle noise represents a primary factor that affects the quality of holographic reconstructions. In recent years, deep learning has provided a new computational paradigm for generating low-noise holograms. Therefore, this study systematically reviews deep learning-based methods for low-noise hologram generation and provides a detailed discussion of the fundamental principles, technical advancements, and future trends of data-, model-, and hybrid-supervised methods, with the aim of providing a reference for the further optimization of low-noise hologram generation methods.

Progress In recent years, low-noise hologram generation methods based on deep learning have experienced rapid development, and innovations in data-, model-, and hybrid-supervised methods have driven advances in holographic 3D displays. Data-supervised methods depended on paired datasets of target images and holograms, and low-noise holograms were generated by learning the mapping relationships between them. Early data-supervised studies focused on applying deep neural networks to hologram generation, in which tensor holography enabled the real-time synthesis of low-noise holograms, whereas diffraction-engineered networks further enhanced the visual realism of holographic 3D displays. One noteworthy strategy directly generated low-noise holograms from 2D target images, which simplified the construction of training datasets. However, data-supervised methods remained highly dependent on large-scale, high-quality datasets, and generalization ability and model performance were closely associated with dataset quality. Model-supervised methods embedded differentiable physical models into the training process, which avoided reliance on paired datasets and enabled unsupervised or self-supervised learning of low-noise holograms by minimizing the intensity difference between reconstructed and target images. Hardware-based model-supervised methods calibrated optical aberrations through system feedback and achieved high-quality color holograms. Software-based model-supervised methods introduced complex-valued neural networks or lightweight architectures and generated holograms with high accuracy and high frame rates. In addition, model-supervised methods were applied to address occlusion problems in complex 3D scenes. However, model-supervised methods required repeated diffraction propagation calculations, which resulted in high training costs. Hybrid-supervised methods integrated the advantages of data and model supervision and generated high-fidelity, low-noise holograms by combining data and physical losses. For example, the two-stage hybrid training method first pretrained networks on high-quality datasets and performed unsupervised fine-tuning using physical models, surpassing the performance ceiling of data-supervised methods. Under data-scarce conditions, hybrid-supervised methods based on nonpaired learning achieved object-to-hologram mapping through cycle-consistency structures. Hybrid-supervised methods were also applied to optical system optimization, propagation error compensation, and hologram generation under low spatial coherence. However, hybrid-supervised methods still encountered challenges associated with complex training processes and limited stability.

Conclusions and Prospects Although deep learning has driven the rapid development of holographic 3D displays, several challenges remain: 1) in data-supervised methods, the training efficiency of deep neural networks remains low, with a strong dependence on training datasets; 2) in model-supervised methods, the speed of hologram generation does not yet satisfy the requirements for real-time interaction; 3) in hybrid-supervised methods, the performance gains achieved by simply increasing the depth or width of neural networks are limited. Therefore, future research is expected to proceed in the following directions: 1) transfer learning methods can be applied to reduce the training time required for data-supervised methods and alleviate dependence on datasets; 2) lightweight neural networks can be designed to improve the generation speed of model-supervised methods; 3) novel architectures, such as Transformer models, self-attention mechanisms, and diffusion models, can be explored to extend the performance boundaries of neural networks. With continued improvements in training efficiency, inference speed, and reconstruction quality, holographic 3D displays are expected to facilitate transformative applications in medical imaging, industrial inspection, education, and entertainment.

Key words: holographic 3D display; hologram; deep learning; low-noise

Graphical abstract:



(编辑 赵 婧)

引用格式: Li Zhaosong, Fan Yubo, Song Dejiang, et al. Deep learning-based methods for low-noise hologram generation[J]. Advanced Engineering Sciences, 2026, 58(5): 252–261. [李赵松, 樊宇博, 宋德江, 等. 基于深度学习的低噪声全息图生成方法[J]. 工程科学与技术, 2026, 58(5): 252–261.]