

研究论文

卷积误差消除的模型驱动相位型全息图生成网络

何泽浩¹, 刘珂瑄², 曹良才^{2*}, 张岩^{1*}

摘要 计算全息技术是动态三维显示的理想解决方案,具有广阔的应用前景。在当前的技术条件下,计算全息技术面临的挑战是全息算法难以同时兼顾计算的精度和速度。为此,提出了卷积误差消除的模型驱动相位型全息图生成网络,实现了高保真相位型全息图的快速生成。首先,研究了角谱法中卷积误差的产生机制,提出了无卷积误差的角谱法,开发了基于无卷积误差角谱法的迭代框架,证实了无卷积误差角谱法对于提升相位型全息图计算精度的有效性;其次,以无卷积误差角谱法作为编码器构建了卷积误差消除的模型驱动相位型全息图生成网络,将相位型全息图的计算时间减小了3个数量级。通过网络生成的相位型全息图,抑制了全息光学重建中的散斑噪声,提高了重建结果的细节质量,平均峰值信噪比高达20.38 dB。伴随着深度梯度显著性和通道效率一致性的持续提升,该网络有望广泛应用在虚拟现实、元宇宙和三维视频通讯等领域。

关键词 计算全息; 三维显示; 深度学习; 神经网络; 卷积误差

真实世界中的视觉信息多为三维信息,能充分调动人类的感知和运算潜能,帮助人类认识真实世界的多样性和复杂性^[1]。当前显示技术呈现的视觉内容多为二维图像,缺乏深度信息,难以充分调动人类的视觉感知和理解能力,限制了画面的真实性和沉浸感^[2]。因此,研究三维显示技术,提升显示的深度表达能力,一直是显示行业的重要追求。

深度线索是三维显示的关键评价指标,涵盖一切可能提供三维感知的因素,主要包括透视、遮挡、阴影、纹理、先验、视差、会聚和聚焦等种类^[3-4]。显示技术提供的深度线索种类越丰富,三维画面的立体感、沉浸感和真实感越突出。全息三维显示是基于波前复现的三维显示技术,能够提供所有种类深度线

索,理论上具有最为突出的三维视觉效果,被认为是三维显示的“圣杯”和终极呈现形式^[5]。然而,早期全息三维显示多基于全息光路和全息材料实现,光路的复杂性和材料的不可擦写性限制了全息三维显示的应用和推广^[6]。

计算机技术的发展为全息三维显示注入了新的发展活力。计算机技术和全息三维显示技术融合形成的交叉技术被称为计算全息显示技术。该技术极大降低了光路和材料对全息三维显示的限制,具有广阔的应用前景。当前,计算全息显示技术的挑战主要是如何同时保证全息算法的精度和速度^[7]。全息图的计算过程需要使用随机相位模拟三维场景的表面散射特性,但随机相位常导致全息重建结果的散斑噪声和细节丢失^[8-10]。迭代优化策略可以显著抑制随机相位对全息重建结果的影响,但迭代过程需反复执行衍射计算,严重限制了全息图的生成效率^[11-12]。近年来,

1. 首都师范大学物理系, 北京 100048

2. 清华大学精密仪器系, 北京 100084

收稿日期: 2024-12-22; 修回日期: 2025-01-09

基金项目: 国家自然科学基金青年科学基金项目(62205173); 国家自然科学基金专项项目(62441613)

作者简介: 何泽浩, 讲师, 研究方向为全息显示和通信, 电子邮箱: hezehao@cnu.edu.cn; 曹良才(通信作者), 教授, 研究方向为全息光学, 电子邮箱: clic@tsinghua.edu.cn; 张岩(共同通信作者), 教授, 研究方向为太赫兹光电子学、微纳光学、光学信息处理, 电子邮箱: yzhang@cnu.edu.cn

引用格式: 何泽浩, 刘珂瑄, 曹良才, 等. 卷积误差消除的模型驱动相位型全息图生成网络[J]. 科技导报, 2025, 43(5): 107-116; doi:10.3981/j.issn.1000-7857.2024.12.01771

深度学习技术的发展为全息图的快速生成提供了有力的工具。当前,基于深度学习的全息图生成网络可分为数据驱动和模型驱动2大类型^[13-16]。其中,数据驱动网络构造简单、推理高效,但推理精度依赖训练集的规模和质量;模型驱动网络使用了基于物理模型的解码器,降低了网络对训练集规模和质量依赖,是计算全息显示领域当前的热点研究方向。提升解码器精度、获得更高的光学重建精度是模型驱动网络的重要追求。

为了实现高保真相位型全息图的快速生成,本文提出了卷积误差消除的模型驱动相位型全息图生成网络。首先,为了获得更高的光学重建精度,研究了角谱法中卷积误差的产生机制,提出了无卷积误差的角谱法,开发了基于无卷积误差角谱法的迭代框架,证实了该方法在相位型全息图生成中的有效性;在此基础上,构建了卷积误差消除的模型驱动相位型全息图生成网络,实现了高保真相位型全息图的快速生成,显著降低了全息图的计算时间,并在光学重建中获得了理想的图像质量。

1 角谱法卷积误差的产生机制和消除方法

角谱法在任意物平面和全息平面上的采样间隔完全相同,适用于三维场景全息图的计算生成^[17]。然而,传统角谱法在数值计算时较少考虑卷积误差,即使经过迭代,获得的全息图也难以达到较高的重建质量^[18]。受限于同样的原因,基于角谱法设计的模型驱动网络也常常面临全息图光学重建质量受限的困扰。

本文将二维图像为例,首先阐明卷积误差的产生机制。对于二维图像,角谱法可由下式描述:

$$\tilde{E}_H(x, y) = \text{FT}^{-1} \left\{ \text{FT} \left\{ E_O(x, y) \exp[j\varphi(x, y)] \right\} \times \exp \left[j \frac{2\pi}{\lambda} z_0 \sqrt{1 - (\lambda u)^2 - (\lambda v)^2} \right] \right\} \quad (1)$$

式中, $E_O(x, y)$ 为物平面的振幅分布, z_0 为物平面到全息平面的距离, $\varphi(x, y)$ 为叠加的初始相位, λ 为入射波长, FT 为傅立叶变换, u 和 v 分别为水平和竖直方向的空间频率。

角谱法的基准面为全息平面,其深度设定为 $z = 0$ 。为便于推导,将式(1)进一步简化为一维形式:

$$\tilde{E}_H(x) = \text{FT}^{-1} \left\{ \text{FT} \left\{ E_O(x) \exp[j\varphi(x)] \right\} \times \exp \left[j \frac{2\pi}{\lambda} z_0 \sqrt{1 - (\lambda u)^2} \right] \right\} \quad (2)$$

使用 h_z 表示卷积核,式(2)可改写为卷积形式:

$$\tilde{E}_H(x) = \{ E_O(x) \exp[j\varphi(x)] \} * \text{FT}^{-1} \left\{ \exp \left[j \frac{2\pi}{\lambda} z_0 \sqrt{1 - (\lambda u)^2} \right] \right\} = \tilde{E}_O(x) * h_z(x) \quad (3)$$

式中, $*$ 为卷积符号。

式(3)可精确描述光学全息图的生成机制。但是,在计算全息中,全息平面的复振幅分布需要通过数值计算而获得。考虑到数值计算的离散性,物平面的振幅分布和卷积核均需要离散采样。此时,式(3)被改写为离散形式:

$$\tilde{E}_H(n) = \sum_{m=0}^{N-1} \tilde{E}_O(m) \cdot h_z(n-m) \quad (4)$$

式中, N 为数值计算的离散采样数, n 为 $[0, 2N-1]$ 之间的整数, m 为 $[0, N-1]$ 之间的整数。

考虑到卷积运算的延拓特性,式(4)得到的全息平面复振幅分布,其离散采样数最终被扩展至 $2N-1$ 。与此同时,为提升计算效率,数值计算中常使用快速傅里叶变换算法(fast Fourier transform, FFT)代替式(3)中的 FT 。在 FFT 算法中,非周期分布的离散函数会被周期延拓^[19]。在这种情况下,式(4)描述的线性卷积运算将转换为循环卷积运算^[20]:

$$\tilde{E}'_H(n) = \tilde{E}_O(n) \otimes h_z(n) = \left[\sum_{m=0}^{L-1} E_O(m) \cdot h_z((n-m))_L \right] R_L(n) \quad (5)$$

式中, $\otimes$ 为循环卷积符号, L 为循环卷积的运算周期, $((\cdot))_L$ 表示除以 L 后的余数, $R_L(n)$ 表示以 L 为周期的延拓。

由式(4)和式(5)可推导出线性卷积和循环卷积的转换关系:

$$\tilde{E}'_H(n) = \left[\sum_{r=-\infty}^{+\infty} \tilde{E}_H(n+rL) \right] R_L(n) \quad (6)$$

式中, r 表示任意整数。

可见,为防止周期延拓带来的混叠误差,循环卷积的延拓周期 L 应大于线性卷积的采样数:

$$L \geq 2N-1 \quad (7)$$

已有研究表明,以峰值信噪比(peak signal-to-noise ratio, PSNR)为评价指标,分辨率作为离散化参数,对全息光学重建质量的影响并不显著^[21-22]。不过,考虑到全息图分辨率过低时,全息重建结果的精细度会受到限制,本工作使用了 4K 分辨率的全息图。而

式(7)分析结果表明,在单个循环卷积周期内,采样窗口尺寸不小于目标图像尺寸的2倍时,才可以避免卷积误差的影响。此外,为了避免周期拓展引入额外噪声,拓展过程中,目标图像周围像素点的取值都应设置为0。

2 面向相位型全息图的无卷积误差角谱法迭代框架

2.1 基于无卷积误差角谱法的迭代框架

实验中,全息图需使用空间光调制器(spatial light

modulator, SLM)加载。当前,动态 SLM 难以同时加载振幅和相位,需要将式(6)得到复振幅进行编码。本文使用相位型 SLM 加载全息图,全息图编码时将复振幅的振幅分量完全舍弃:

$$P(x,y) = \exp\{j \cdot \arg[\tilde{E}_H(x,y)]\} \quad (8)$$

式中, $\arg$ 表示提取复振幅的辐角。

式(8)所述的相位型全息图由于振幅分量丢失,光学重建质量显著下降。为验证无卷积误差的角谱法能否用于相位型全息图的优化,本文开发了基于无卷积误差角谱法的迭代框架,该框架工作流程可分解为5个步骤,如图1所示。

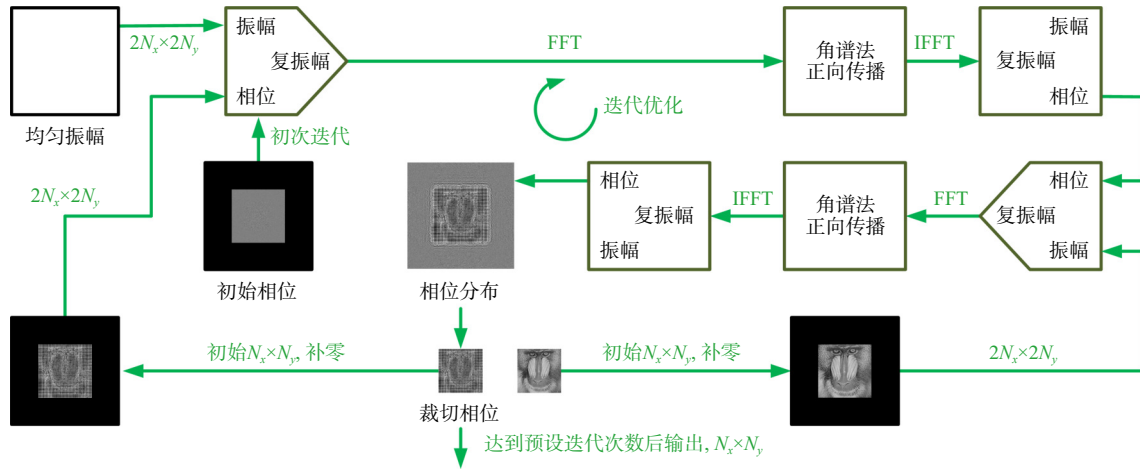


图1 基于无卷积误差角谱法的迭代框架

步骤1: 设置全息平面的复振幅。考虑到本文中全息图的纯相位特性,全息平面的强度分量被设置为均匀分布的单位强度矩阵。第一次迭代中,全息平面的相位分量可设置为任意的相位矩阵,其补零前分辨率为 $N_x \times N_y$, 补零后分辨率为 $2N_x \times 2N_y$ 。第二次及随后的迭代中,全息平面的相位分量由步骤5获得。

步骤2: 计算角谱法正向传播结果。角谱法正向传播指全息平面到物平面的传播,其数学形式为:

$$\tilde{E}_O(x,y) = \text{FT}^{-1} \left\{ \text{FT} \left\{ \tilde{E}_H(x,y) \right\} \times \exp \left[-j \frac{2\pi}{\lambda} z_0 \sqrt{1 - (\lambda u)^2 - (\lambda v)^2} \right] \right\} \quad (9)$$

正向传播的输入是步骤1得到的全息平面复振幅分布,输出是物平面的复振幅分布。数值计算时,式(9)中的FT和逆FT分别通过FFT和逆FFT算法实现。在逆FFT运算前,需对传递函数进行补零,将其由 $N_x \times N_y$ 扩展至 $2N_x \times 2N_y$ 。

步骤3: 设置物平面的复振幅。物平面的振幅分量是将目标图像补零扩展后得到的强度矩阵,其分辨率由 $N_x \times N_y$ 扩展至 $2N_x \times 2N_y$; 物平面的相位分量是步骤2计算获得的相位矩阵,该相位矩阵通过舍弃复振幅的振幅分量获取。

步骤4: 计算角谱法反向传播结果。角谱法反向传播指物平面到全息平面的传播,其数学形式由式(1)表示。与步骤2相同,本步骤中的FT和逆FT分别通过FFT和逆FFT算法实现,传递函数也需要由 $N_x \times N_y$ 扩展至 $2N_x \times 2N_y$ 。反向传播的输入是步骤3得到的物平面的复振幅分布,输出是全息平面的复振幅分布。

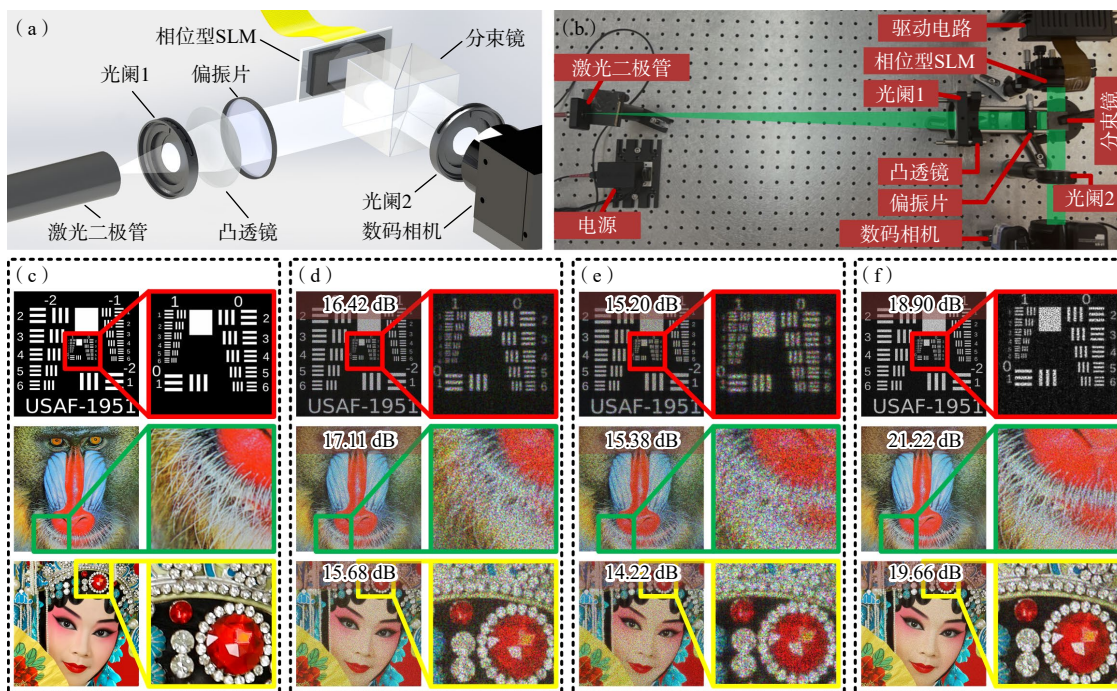
步骤5: 设置全息平面的相位。步骤4输出的复振幅分辨率为 $2N_x \times 2N_y$, 根据该复振幅提取得到的相位分量分辨率也为 $2N_x \times 2N_y$ 。该相位分量的边缘部分对应了目标图像的补零区域,对全息重建的影响较

为有限,是相位分量的无效区域。相位分量的中心部分主要用于重建目标图像,是相位分量的有效区域,该区域的分辨率仅为 $N_x \times N_y$ 。换句话说,步骤4获得的相位分布存在极大的空间带宽积冗余。为了提升系统的空间带宽积利用率,本步骤先提取复振幅的相位分量,随后舍弃该相位分量的边缘部分,只保留分辨率为 $N_x \times N_y$ 的中心区域。在迭代过程中,为避免步骤2产生卷积误差,本步骤通过补零将中心区域的分辨率拓展至 $2N_x \times 2N_y$;在迭代达到预设次数 N_{Layer} 时,最终输出的相位分量不做补零拓展,分辨率仍为 $N_x \times N_y$ 。这个未经拓展的相位分量就是迭代框架最终输出的相位型全息图。

2.2 光学重建结果

为了评估无卷积误差角谱法迭代框架对相位型

全息图的优化能力,本文搭建了彩色全息光学显示系统,该系统的原理和实物分别如图2(a)和图2(b)所示。系统使用的照明光源是Fisba ReadyBeam多波长激光二极管。光源发出的光束被光阑限制口径,随后被凸透镜准直。准直的平面波经过分束镜反射,照射型号为Holoeye Gaca-2 Vis的相位型SLM。SLM表面加载了目标图像的相位型全息图。考虑到相位型SLM的偏振敏感特性,彩色全息光学显示系统中还放置了偏振片。被相位型全息图所调制的光波经SLM反射后再次经过分束镜,最后在预设距离处实现光学重建,相应的重建结果被Canon 60D数码相机所采集。为减少杂散光对全息重建质量的影响,光学显示系统中分束镜和数码相机之间还放置了第二块光阑。



(a) 彩色全息光学显示系统原理图; (b) 彩色全息光学显示系统实物图; (c) 彩色目标图像; (d) 无卷积误差非迭代框架的彩色光学重建结果; (e) 有卷积误差迭代框架的彩色光学重建结果; (f) 无卷积误差迭代框架的彩色光学重建结果

图2 光学重建结果

待重建的目标图像如图2(c)所示。实验中,物平面到全息平面的距离为100 mm。为实现彩色重建,本文提取了目标图像红色、绿色和蓝色通道,以638 nm、520 nm和450 nm为中心波长计算获取了对应的相位型全息图,并时序加载到了相位型SLM的表面上。在时序显示方案中,由于红色、绿色和蓝色

通道对应的相位型全息图是分别单独计算的,不同颜色通道不存在因波长导致的重建尺寸差异,因此无需在计算时根据波长对目标图像做缩放预处理。全息光学显示系统的驱动电路保证了激光二极管发光通道和相位型全息图颜色通道的同步。数码相机时序采集红色、绿色和蓝色通道的全息重建结果,利用通

道合成技术实现了彩色目标的全息光学重建。为降低全息光学重建结果的色差,本文中激光二极管红色、绿色和蓝色通道的发光强度分别设置为最大发光强度的 100%、100% 和 10.1%。

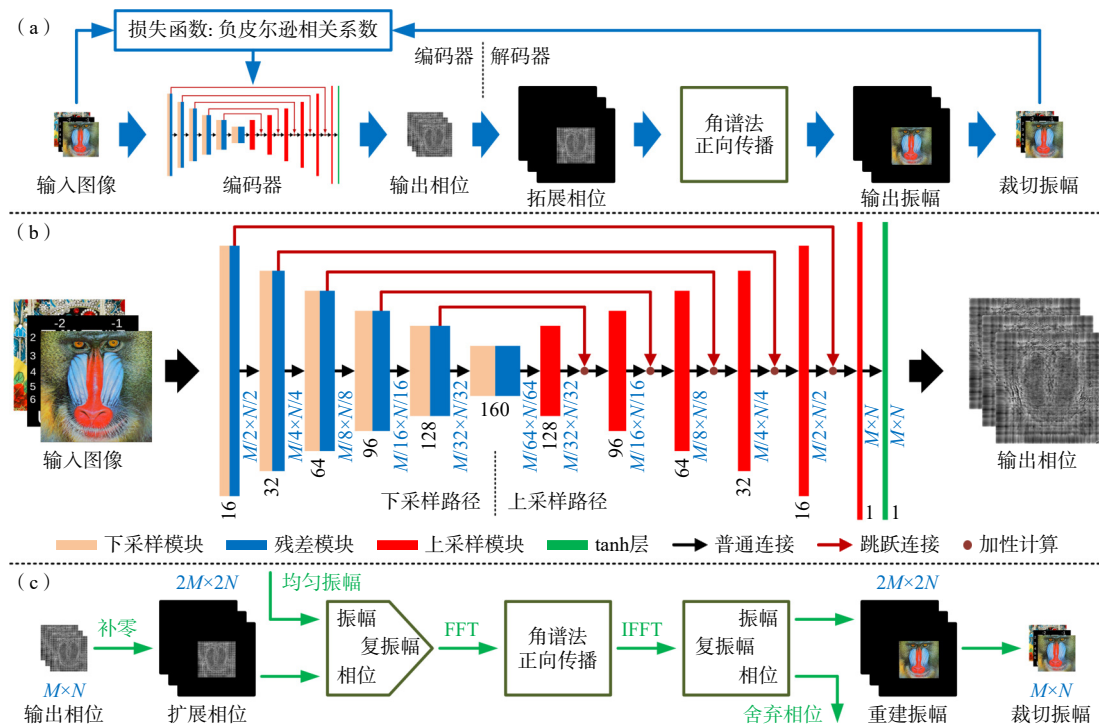
使用无卷积误差非迭代框架、有卷积误差迭代框架和无卷积误差迭代框架分别计算目标图像的相位型全息图并进行光学重建,对应的重建结果如图 2(d)~2(f)所示。其中,图 2(e)和 2(f)对应的相位型全息图均经过了 100 次迭代计算。非迭代框架下,全息光学重建结果的细节质量较差,散斑噪声和色差较为显著,图像平均 PSNR 值为 16.40 dB;有误差的迭代框架下,全息光学重建结果的细节质量更为受限,散斑噪声和色差也更为严重,图像平均 PSNR 值下降为 14.93 dB;无误差迭代框架下,全息光学重建结果的细节质量显著提升,散斑噪声和色差被显著抑制,图像平均 PSNR 值上升至 19.93 dB。可以看出,迭代过程不一定会提升相位型全息图的重建质量。在卷积误差没有消除时,迭代过程对全息重建质量的影响甚至是负面的。只有消除卷积误差后,迭代过程才会显著地提升相位型全息图的光学重建质量。比较光

学重建结果中 USAF-1951 分辨率板的最细解析线、狒狒的胡须和皇冠上的珠宝,可以验证无卷积误差角谱法迭代框架对相位型全息图重建质量的提升能力。也就是说,无卷积误差角谱法能够用于相位型全息图的优化。因此,以无卷积误差角谱法为支撑,本文构建了模型驱动的相位型全息图生成网络。

3 模型驱动的相位型全息图生成网络

3.1 卷积误差消除的模型驱动相位型全息图生成网络

无卷积误差迭代框架面临的最大的挑战是计算速度受限。计算一张 4K 分辨率的相位型全息图,该迭代框架消耗的时间为 611.24 s。为了提升全息图的计算效率,本文将无卷积误差角谱法应用在了模型驱动深度学习网络中,提出了卷积误差消除的模型驱动相位型全息图生成网络,如图 3(a)所示。该网络由 U-Net 编码器和无卷积误差角谱法解码器构成。网络训练完成后,U-Net 编码器可以单独取出,作为相位型全息图的生成工具。



(a) 模型驱动网络的整体架构; (b) 模型驱动相位型全息图生成网络的 U-Net 编码器;

(c) 模型驱动相位型全息图生成网络的无卷积误差角谱法解码器

图 3 卷积误差消除的模型驱动相位型全息图生成网络

U-Net 编码器由一条下采样路径和另一条上采样路径构成, 网络的输入是目标图像, 输出是对应的相位型全息图, 二者分辨率均为 $N_x \times N_y$, 如图 3(b) 所示。其中, 下采样路径包含 6 个下采样模块和 6 个残差模块, 每个下采样或残差模块均包含一个 3×3 卷积层、一个 ReLU 函数和两组批归一化(batch normalization, BN)运算, 且残差模块总是叠加在相应的下采样模块上; 上采样路径包含 6 个上采样模块, 每个上采样模块均基于亚像素卷积思想设计^[23], 可分解为卷积层和转换层, 设置卷积层是为了增加通道数, 而设置转换层则是为了重新设置数据结构。为了缓解网络训练的退化问题, 下采样路径中的残差模块通过跳跃连接到达相应的上采样模块。上采样路径的末端包含一个 tanh 函数层, 用以将输出结果的相位限制在 $[-\pi, \pi]$ 的范围内。

无卷积误差角谱法解码器本质上是式(1)所描述的角谱法正向传播模型。该模型的输入是补零后的相位型全息图, 输出是全息平面上的重建复振幅, 二者分辨率均为 $2N_x \times 2N_y$, 如图 3(c) 所示。提取全息重建复振幅中的振幅分量, 并将其中心分辨率 $N_x \times N_y$ 的区域裁剪取出, 即可获得相位型全息图的重建强度。该重建强度可以通过损失函数与输入图像的建立联系, 实现模型驱动相位型全息图生成网络的自监督训练^[24]。本文使用负皮尔逊相关系数作为网络的损失函数, 以提高网络的收敛概率。若 I 和 I_r 分别表示输入和重建图像的强度分布, 则负皮尔逊相关系数可由如下表达式描述:

$$L_{\text{NPCC}} = - \frac{\sum_{M,N} \{ [I_r(m,n) - \bar{I}_r] [I(m,n) - \bar{I}] \}}{\sqrt{\sum_{M,N} [I_r(m,n) - \bar{I}_r]^2 \sum_{M,N} [I(m,n) - \bar{I}]^2}} \quad (10)$$

式中, $\bar{I}_r$ 为重建图像像素值的平均值。

3.2 光学重建结果

模型驱动相位型全息图生成网络在 Matlab 中完成训练和测试。训练和测试使用的硬件平台 CPU 型号为 Intel Xeon Gold 6248R, GPU 型号为 NVIDIA QuadroGV100。DIV2K_train_HR 和 DIV2K_valid_HR 两个数据集分别作为网络的训练集和测试集。定义使用训练集全部样本对网络进行一次训练的过程为一个 epoch, 本文中训练过程包含了 40 个 epoch。对于彩色图像, 红、绿、蓝 3 个颜色通道的中心波长分别为 638 nm、520 nm 和 450 nm; 全息平面和物平面的间距为 300 mm。完成训练和测试, 随后的推理过程只使用 U-Net 编码器作为相位型全息图的生成网络。

使用图 2(a)所示的彩色全息光学显示系统对相位型全息图进行光学重建, 验证模型驱动相位型全息图生成网络的计算性能。为了实现全彩色光学重建, 网络分别计算了目标图像红、绿、蓝通道对应的相位型全息图, 并按照时间顺序分别加载在相位型 SLM 上。使用相机采集 3 个通道的光学重建结果, 通过通道合成获取目标图像的彩色光学重建结果。

结果表明, 训练完成的模型驱动网络可在 0.15 s 内推理获得 4K 分辨率单色图像对应的相位型全息图。相比非迭代框架和迭代次数为 100 的无卷积误差迭代框架, 模型驱动网络将相位型全息图的计算时间由 6.49 s 和 611.24 s 提升至 0.15 s, 计算速度分别提升了 43 倍和 4075 倍; 光学重建结果的平均 PSNR 值由 16.40 dB 和 19.93 dB 提升至 20.38 dB, 提升幅度分别为 3.98 dB 和 0.45 dB。如图 4 所示, 总体上看, 模型驱动网络的光学重建结果拥有更弱的散斑噪声, 细节重建质量和迭代框架相比也非常接近。



图 4 模型驱动相位型全息图生成网络的彩色光学重建结果

4 全息图计算生成方法在三维显示中的应用

4.1 面向三维显示的迭代框架和模型驱动网络

数值计算时, 式(2)所描述的目标图像振幅分布

和角谱传递函数均会被离散采样, FT 和逆 FT 运算分别由 FFT 和逆 FFT 算法实现。此时, 式(2)可写为离散运算形式:

$$\tilde{E}_H(n) = \text{IFFT} \left\{ \text{FFT} [\tilde{E}_O(n)] \exp \left[j \frac{2\pi}{\lambda} z_0 \sqrt{1 - (\lambda u_n)^2} \right] \right\} = \text{IFFT} \left\{ \text{FFT} [\tilde{E}_O(n)] H_{z_0}(u_n) \right\} \quad (11)$$

式中, u_n 为 u 的离散形式, $H_{z_0}(u_n)$ 为离散形式的角谱传递函数, IFFT 表示逆 FFT 算法。

通过补零, 采样窗口的尺寸由 N 扩大为 $2N$ 。设 Δ_x 为物平面的采样间隔, 经过第一次 FFT 运算, 傅里叶频谱面的采样间隔 Δ_u 和采样窗口尺寸 L_u 满足如下关系:

$$\Delta_u = 1/(2N \cdot \Delta_x), L_u = 2N \cdot \Delta_u = 1/\Delta_x \quad (12)$$

式中, Δ_x 为物面的采样间隔。

傅里叶频谱面上, FFT 的运算结果需要与角谱传递函数 H_{z_0} 相乘。为避免引入采样误差, 角谱传递函

数也采用式(12)的采样参数。

逆 FFT 运算可视为传播距离相等但方向相反的 FFT 运算。相应的, 逆 FFT 运算后全息平面的采样间隔 Δ'_x 和采样窗口长度 L'_x 可表示为:

$$\Delta'_x = 1/(2N \cdot \Delta_u) = \Delta_x, L'_x = 2N \cdot \Delta_x = L_x \quad (13)$$

式中, L_x 为物面的采样窗口。

可见, 全息平面和物平面的采样参数始终相等, 与传播距离无关。相比菲涅尔模型^[11,25], 角谱法用于三维显示时, 无需根据传播距离调整物平面的采样参数, 算法设计更加便捷。利用角谱法采样参数和传播距离无关的特性, 本文设计了面向三维显示的无卷积误差迭代框架和模型驱动相位型全息图生成网络。这些方法都可以分解为 3 个步骤, 如图 5 所示。

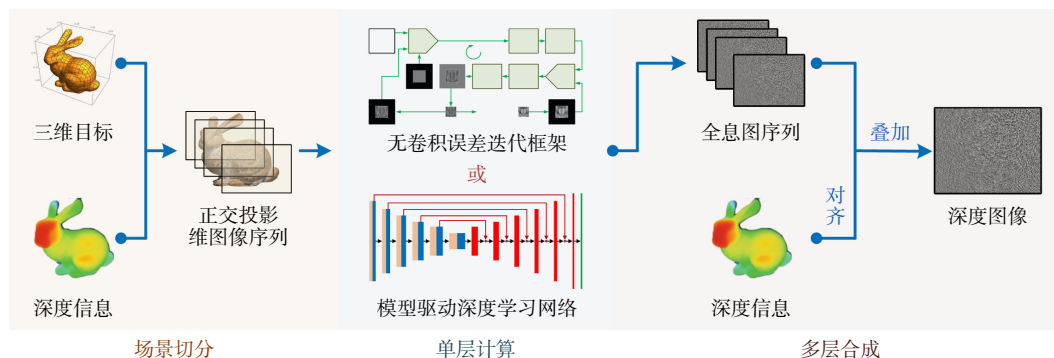


图5 面向三维显示的相位型全息图计算方法

步骤 1: 场景切分。在特定的波前传播方向上, 三维目标本质上就是投影在一系列正交平面上的二维强度分布。因此, 本方法基于这一特性, 首先将三维目标分割为到全息平面距离不同的 N_{Lay} 张二维强度图像。

步骤 2: 单层计算。使用无卷积误差迭代框架或模型驱动相位型全息图生成网络, 分别计算二维图像序列中每张二维图像的相位型全息图。

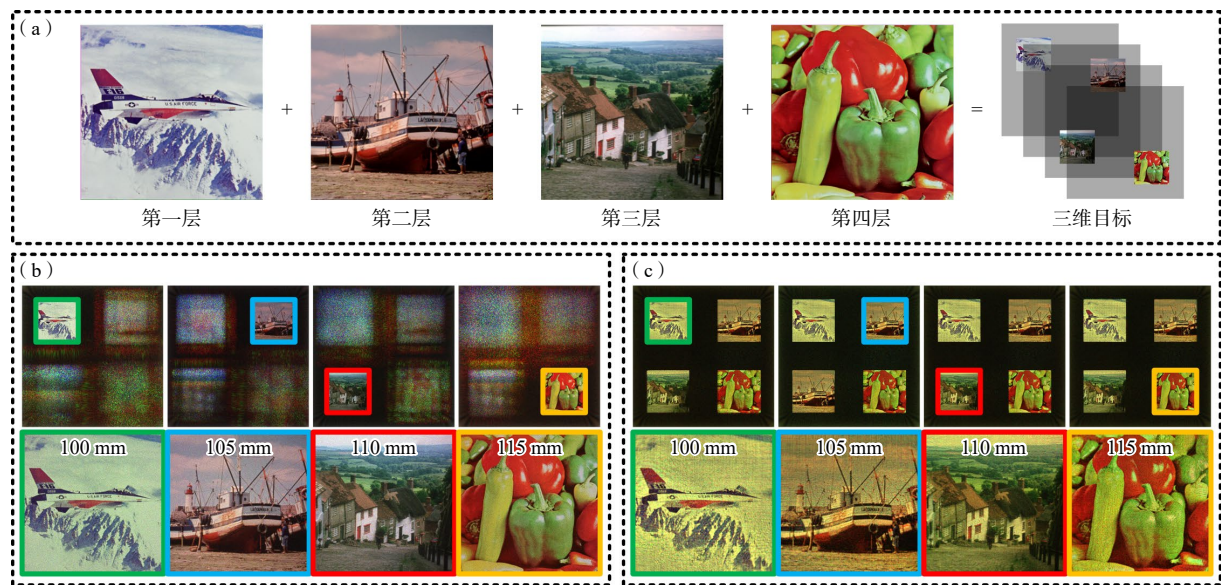
步骤 3: 多层合成。合成获取整个三维目标的相位型全息图时, 需要对每张二维图像的相位型全息图进行边界对齐和相位校准, 避免全息重建时产生像素串扰而降低显示质量。

4.2 光学重建结果

为了验证无卷积误差迭代框架和模型驱动相位型全息图生成网络在三维光学重建中的效果, 本文使用 4 张二维图像组成了一幅如图 6(a) 所示的三维目

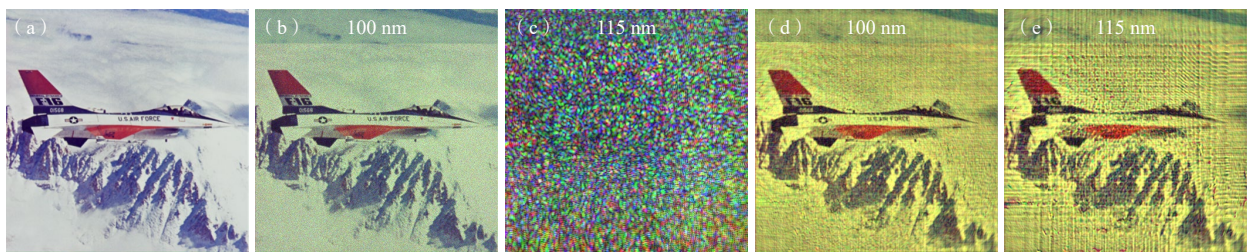
标图像。全息平面到 4 个离散物平面的距离分别为 100 mm、105 mm、110 mm 和 115 mm。利用迭代框架或生成网络分别计算 4 张二维图像在全息平面上的相位分布, 进行边界对齐和相位校准后进行叠加, 最终可得到三维目标的相位型全息图。迭代框架和模型驱动网络获得的相位型全息图在不同距离的光学重建结果分别如图 6(b) 和 6(c) 所示。可以看出, 当重建距离不同时, 不同深度上的目标图像得以清楚重建, 光学重建结果能够呈现出不同平面的离焦效果。全息光学重建结果具有较高的细节质量, 散斑噪声的抑制能力也相当优异。

值得注意的, 模型驱动网络生成的相位型全息图在 2 个方面明显区别于无卷积误差迭代框架。第一, 重建距离变化时, 通过模型驱动网络生成的相位型全息图, 其重建结果未随深度呈现显著的变化。如图 7 所示, 以左上角飞机图片为例, 当重建距离由 100 mm



(a) 三维目标图像的构建; (b) 无卷积误差迭代框架获取的相位型全息图的彩色三维光学重建结果;
(c) 模型驱动相位型全息图生成网络的彩色三维光学重建结果

图6 三维目标及光学重建结果



(a) 三维目标第一层使用的二维图像; (b) 无卷积误差迭代框架获取的相位型全息图在不同深度的光学重建结果;
(c) 模型驱动网络获取的相位型全息图在不同深度的光学重建结果

图7 无卷积误差迭代框架和模型驱动网络的光学重建结果对比

变化至 115 mm 时,无卷积误差迭代框架的重建结果变化剧烈,图像信息几乎无法辨认;但模型驱动网络的重建结果变化程度较小,可以在很大程度上辨认出图像信息。导致这一现象的主要原因是当前网络在下采样过程中舍弃了高频信息,抑制了重建结果的梯度变化。未来在设计深度学习网络时,可以考虑增加网络模块以保留尽可能多的高频成分,或者采用 Transformer 等网络构型避免过多的下采样操作。第二,当红、绿、蓝 3 个颜色通道的激光功率保持不变时,模型驱动网络的光学重建结果相比无卷积误差迭代框架存在色差。这一现象说明,当不同颜色通道的照明光强度不变时,通过模型驱动网络获得的相位型全息图目标重建级次的能量相对于无卷积误差迭代框架发生了改变,最终在通道合成时致使图像色彩发

生了偏移。这种色彩偏移现象很可能源于相位型全息图衍射效率的变化,我们计划在未来的研究中开展系统性探索。

5 结语

本文针对当前计算全息显示技术难以同时兼顾算法精度和速度的挑战,提出了卷积误差消除的模型驱动相位型全息图生成网络。该网络由编码器和解码器两部分组成。为了基于角谱法设计解码器,研究了角谱法中存在的卷积误差,提出了无卷积误差的角谱法;为了验证无卷积误差角谱法对于相位型全息图的有效性,开发了相应的迭代框架,在光学实验中将重建结果的平均 PSNR 值提升了 3.53 dB,证实了无

卷积误差角谱法对于相位型全息图计算任务的有效性。利用无卷积误差角谱法作为解码器,构建并训练了模型驱动的相位型全息图生成网络。该网络可在0.15 s内推理获得4K分辨率单色图像对应的相位型全息图,相比迭代次数为100的无卷积误差迭代框架,计算速度提升了3个数量级,并且在光学重建中获得了更高的平均PSNR值和更弱的散斑噪声。未来,我们将会进一步提升该网络的深度梯度显著性和通道效率一致性,推动其应用于虚拟现实、元宇宙和三维视频通讯等领域。

参考文献 (References)

- [1] 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 (特邀) [J]. 红外与激光工程, 2022, 51(1): 267–281.
- [2] 何泽浩, 曹良才. 面向沉浸式元宇宙的显示、交互和应用 [J]. 科技导报, 2023, 41(5): 6–14.
- [3] Blundell B, Schwarz A. Volumetric three dimensional display system[M]. New York: Wiley-IEEE Press, 1999.
- [4] Okoshi T. Three-dimensional imaging techniques[M]. New York: Academic Press, 1976.
- [5] Goodman J P. Introduction to Fourier optics[M]. 4th ed. New York: W. H. Freeman & Company, 2017.
- [6] Li C, Cao L, Wang Z, et al. Hybrid polarization-angle multiplexing for volume holography in gold nanoparticle-doped photopolymer[J]. *Optics Letters*, 2014, 39(24): 6891–6894.
- [7] Sui X, He Z, Chu D, et al. Non-convex optimization for inverse problem solving in computer-generated holography[J]. *Light: Science & Applications*, 2024, 13: 158.
- [8] Mengü D, Ulusoy E, Urey H. Non-iterative phase hologram computation for low speckle holographic image projection[J]. *Optics Express*, 2016, 24(5): 4462–4476.
- [9] Zea A V, Torroba R. Optimized random phase tiles for non-iterative hologram generation[J]. *Applied Optics*, 2019, 58(32): 9013–9019.
- [10] Zhao T, Liu J, Duan J, et al. Image quality enhancement via gradient-limited random phase addition in holographic display[J]. *Optics Communications*, 2019, 442: 84–89.
- [11] Chen C-Y, Chang H T, Chang T-J, et al. Full-color and less-speckled modified Gerchberg-Saxton algorithm computer-generated hologram floating in a dual-parabolic projection system[J]. *Chinese Optics Letters*, 2015, 13(11): 110901.
- [12] Liu K, He Z, Cao L. Double amplitude freedom Gerchberg-Saxton algorithm for generation of phase-only hologram with speckle suppression[J]. *Applied Physical Letters*, 2021, 120(6): 061103.
- [13] Yan X, Liu X, Li J, et al. Generating multi-depth 3D holograms using a fully convolutional neural network[J]. *Advanced Science*, 2024, 11(28): 2308886.
- [14] Liu K, Wu J, He Z, et al. 4K-DMDNet: Diffraction model-driven network for 4K computer-generated holography[J]. *Opto-Electronic Advances*, 2023, 6(5): 220135.
- [15] Chang C, Wang D, Zhu D, et al. Deep-learning-based computer-generated hologram from a stereo image pair[J]. *Optics Letters*, 2022, 47(6): 1482–1485.
- [16] Lee M H, Lew H M, Youn S, et al. Deep learning-based framework for fast and accurate acoustic hologram generation[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2022, 69(12): 3353–3366.
- [17] Kozacki T, Martinez-Carranza J, Kukolowicz R, et al. Accurate reconstruction of horizontal parallax-only holograms by angular spectrum and efficient zero-padding[J]. *Applied Optics*, 2020, 59(27): 8450–8458.
- [18] He Z, Sui X, Jin G, et al. Distortion correction method based on angular spectrum algorithm for holographic display[J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(11): 6162–6169.
- [19] Matsushima K, Shimobaba T. Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields[J]. *Optics Express*, 2009, 17(22): 19662–19673.
- [20] Williamson D. Discrete-time signal processing[M]. New York: Springer, 1999.
- [21] He Z, Sui X, Jin G, et al. Optimal quantization for amplitude and phase in computer-generated holography[J]. *Optics Express*, 2021, 29(1): 119–133.
- [22] Wang X, He Z, Cao L. Analysis of reconstruction quality for computer-generated holograms using a model free of circular-convolution error[J]. *Optics Express*, 2023, 31(12): 19021–19035.
- [23] Shi W, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, 2016: 1874–1883.
- [24] Wu J, Liu K, Sui X, et al. High-speed computer-generated holography using an autoencoder-based deep neural network[J]. *Optics Letters*, 2021, 46(12): 2908–2911.
- [25] Velez-Zea A, Torroba R. Mixed constraint in global and sequential hologram generation[J]. *Applied Optics*, 2021, 60(7): 1888–1895.

Convolution-error-free model-driven neural network for phase-only computer-generated hologram

HE Zehao¹, LIU Kexuan², CAO Liangcai^{2*}, ZHANG Yan^{1*}

1. Department of Physics, Capital Normal University, Beijing 100048, China

2. Department of Precision Instrument, Tsinghua University, Beijing 100084, China

Abstract Computer-generated holography is a distinguished and promising choice for dynamic three-dimensional display. Currently, the greatest challenge in computer-generated holography lies in the inability of holographic algorithms to simultaneously achieve both high precision and speed. To address this challenge, a convolution-error-free model-driven neural network is proposed in this work, enabling the fast generation of high-fidelity phase-only computer-generated holograms. Firstly, the mechanism of convolution errors in the angular spectrum method is analyzed. A modified angular spectrum method that eliminates convolution errors is proposed, along with an iterative framework based on this modified method. The effectiveness of the error-free angular spectrum method in enhancing the calculation precision of phase-only holograms is successfully demonstrated. Secondly, a model-driven neural network is developed by incorporating the convolution-error-free angular spectrum method as the decoder. This network achieves a reduction of calculation time for generating phase-only holograms by 3 orders of magnitude. The phase-only holograms generated by this network effectively suppress speckle noise and enhance detail quality in optical reconstructions, achieving an average peak signal-to-noise ratio (PSNR) of 20.38 dB. With ongoing advancements in depth gradient saliency and channel efficiency consistency, the proposed network holds significant potential for widespread application in areas including virtual reality, metaverse, and three-dimensional video communication.

Keywords computer-generated holography; three-dimensional display; deep learning; neural network; convolution error ●



(责任编辑 卫夏雯)