

# 面向光束整形的自由曲面衍射光学设计方法 (特邀)

廖清明, 冯泽心\*

(北京理工大学 光电学院 北京市混合现实与新型显示工程技术研究中心, 北京 100081)

**摘要:** 光束整形技术能够实现激光光束能量分布的良好控制,在激光加工、医疗、激光核聚变等方面具有广泛的应用前景。光束整形方法主要分为几何光学方法和物理光学方法两大类,其中几何光学方法由于忽略了衍射效应,导致一些情形下的光束整形效果不够精细,但却能为物理光学方法提供良好的优化起点。结合几何光学方法和物理光学方法的复合设计方法能够设计出易于加工的自由曲面衍射光学元件,同时能抑制杂散光和散斑噪声。首先回顾了物理光学方法,然后综述了几何光学-物理光学复合设计方法的研究进展,最后对衍射光学元件设计方法的未来发展方向进行了一定的展望。

**关键词:** 光束整形; 衍射光学元件; 复合设计方法; 自由曲面; 相位恢复

**中图分类号:** O436 **文献标志码:** A **DOI:** 10.3788/IRLA20230430

## 0 引言

光束整形是一种重要的光学技术,它通过改变入射激光光束的相位分布或者振幅分布,使出射光束在目标面上具有特定的辐照度(或光强)分布。激光光束整形技术已经被广泛应用于材料加工、医疗、激光打印、光学数据存储、光谱学、激光核聚变等方面<sup>[1]</sup>。文中主要总结纯相位光束整形方法。

光束整形方法一般可分为几何光学方法和物理光学方法两大类。前者从几何光学的角度出发,将入射光束看作是一簇携带能量的光线,经过光学元件后光线方向发生偏折,最终在目标面上产生特定的辐照度分布。一般情况下,入射面和目标面的能量分布之间存在着一种特定的点对点映射关系,该映射关系决定了入射光线的偏折方向<sup>[2-4]</sup>。根据光线的偏折方向可以求解光学元件对应的相位分布。在薄元近似<sup>[5]</sup>下,可根据该相位分布直接确定光学曲面。非薄元近似和非傍轴情形下,基于几何光学构建光学自由曲面是一个更为复杂的反问题,文中不涉及,具体可参考文献<sup>[6]</sup>。一般情况下,几何光学方法能够产生连续的光学自由曲面,易于加工。然而在很多情况下,衍

射效应会很大程度上影响目标面辐照度分布的精度。特别在光学元件特征尺寸较小时,几何光学方法不再适用。

基于波动光学理论描述光场传播的物理光学方法是一种更精确的光束整形方法。该方法通过纯相位型衍射光学元件(Diffractive Optical Element, DOE)改变入射光场的相位分布,从而在目标面上产生特定的辐照度分布。相比于传统的折反射光学器件,DOE的质量更轻,体积更小,能实现传统折反射光学器件难以实现的集成化、阵列化等功能,更符合光电系统轻量化、微型化的发展趋势。例如一些智能手机采用 DOE 在有限空间内生成特定的点阵用于人脸识别。与此同时,DOE 对加工精度提出了更高的要求,制造缺陷容易产生高阶衍射级次、杂散光和散斑噪声<sup>[7]</sup>。而传统物理光学方法<sup>[8-25]</sup>设计的 DOE 通常具有复杂不规则的表面形状,加工误差较大,往往导致生成的辐照度图像质量退化严重。

为此,研究者们提出了结合几何光学方法和物理光学方法的复合设计方法<sup>[26-30]</sup>,在考虑衍射效应的基础上,尽可能保持 DOE 表面(或其对应的相位面)光

收稿日期:2023-04-20; 修订日期:2023-05-20

基金项目:国家自然科学基金项目(11704030);中国人工智能学会-华为 MindSpore 学术奖励基金项目(CAAIXSJLJJ-2021-014A, CAAIXSJLJJ-2022-010A)

作者简介:廖清明,男,硕士生,主要从事衍射光学元件设计方面的研究。

导师(通讯作者)简介:冯泽心,男,副教授,博士生导师,博士,主要从事计算光学和自由曲面光学方面的研究。

滑。首先利用几何光学方法求解一个连续光滑的相位分布,随后将其作为物理光学方法的初值进行优化,得到最终的相位分布。基于几何光学方法计算得到的相位分布一般是连续的,以其作为优化起点可使物理光学算法快速收敛,且通常具有相对较低的粗糙度,大大简化了对应 DOE 的制造难度<sup>[31]</sup>。为与传统 DOE 相区别,这里将复合型设计方法获得的 DOE 称为自由曲面衍射光学元件(也有研究者称其为自由曲面全息图<sup>[26]</sup>)。

文中对面向光束整形的 DOE 设计方法进行了概述。首先简单回顾了传统的物理光学设计方法,然后着重介绍了结合几何光学方法和物理光学方法的复合设计方法,最后进行了总结与展望。

### 1 传统物理光学设计方法

薄元近似下,DOE 的面型高度  $h(x,y)$  与其对应的相位分布  $\varphi(x,y)$  成线性关系,一般可写为<sup>[5]</sup>:

$$h(x,y) = \frac{\lambda}{2\pi(n-1)}\varphi(x,y) \quad (1)$$

式中:  $\lambda$  为波长;  $n$  为 DOE 的折射率;  $(x,y)$  为 DOE 采样点坐标。此时 DOE 的设计问题等价于根据入射面和目标面辐照度分布恢复中间调制元件的相位分布问题。目前,相位恢复问题的物理光学求解方法主要包括迭代傅里叶变换算法 (Iterative Fourier Transform Algorithm, IFTA)<sup>[8-21]</sup> 和最优化方法<sup>[22-25]</sup>。

1972 年, Gerchberg 和 Saxton 提出了经典的 GS 算法<sup>[8]</sup>, 开创了 IFTA 的先河。GS 算法预先设定一个相位初值,通过快速傅里叶变换 (Fast Fourier Transform, FFT) 和逆傅里叶变换 (Inverse Fast Fourier Transform, IFFT) 在空域约束集和频域约束集来回迭代从而恢复出未知的相位信息,原理如图 1 所示。

将 GS 算法中的傅里叶变换过程改为不同的衍射

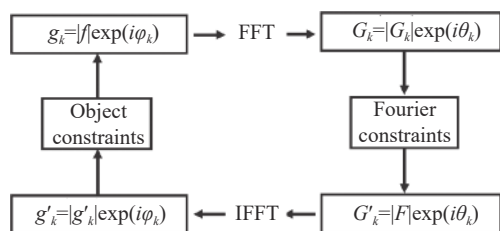


图 1 GS 算法原理图

Fig.1 Schematic diagram of GS algorithm

传播过程能够实现不同场景下的相位恢复,如迭代角谱算法<sup>[9]</sup>就是将傅里叶变换过程替换为角谱传播过程从而能够适用于更广泛的衍射场景。

GS 算法在第  $k$  次迭代的过程概括如下:

(a) 对输入面光场  $g_k(x,y)$  进行傅里叶变换得到频域面光场  $G_k(u,v)$ ;

$$G_k(u,v) = \text{FFT}[g_k(x,y)] = |G_k(u,v)|\exp[i\theta_k(u,v)] \quad (2)$$

式中:  $(x,y)$  和  $(u,v)$  分别为输入面和频谱面点的坐标。

(b) 将计算得到的频域面光场振幅  $|G_k(u,v)|$  替换为已知的光场振幅  $|F(u,v)|$ ;

$$G'_k(u,v) = |F(u,v)|\exp[i\theta_k(u,v)] \quad (3)$$

(c) 对  $G'_k(u,v)$  作逆傅里叶变换得到传播回来的输入面光场  $g'_k(x,y)$ ;

$$g'_k(x,y) = \text{IFFT}[G'_k(u,v)] = |g'_k(x,y)|\exp[i\varphi_k(x,y)] \quad (4)$$

(d) 对  $g'_k(x,y)$  作振幅限制,并将其作为第  $k+1$  次迭代的输入光场,继续执行第 (a) 步。

$$g_{k+1}(x,y) = |f(x,y)|\exp[i\varphi_k(x,y)] \quad (5)$$

GS 算法一般用迭代次数大于特定次数或者损失函数小于特定值作为停止迭代条件,其中第  $k$  次迭代的损失函数常用频域的均方根误差  $E_{Fk}^2$  和空域的均方根误差  $E_{Ok}^2$  表示:

$$E_{Fk}^2 = N^{-2} \sum_{u,v} [|G_k(u,v)| - |F(u,v)|]^2 \quad (6)$$

$$E_{Ok}^2 = \sum_{x,y} [|g'_k(x,y)| - |f(x,y)|]^2 \quad (7)$$

GS 算法虽然能够简单有效地恢复出入射光场的相位分布,但是存在着收敛速度慢、容易陷入局部最优解、依赖初始值等问题。为此研究者们基于 GS 算法提出了一系列改善收敛性的改进算法。

我国的两位学者杨国桢、顾本源提出了非幺正变换系统中振幅和相位恢复问题的一般描述方法—杨顾算法<sup>[10-11]</sup>,相较于 GS 算法适用范围更广。

Fienup 在 GS 算法的基础上对空域的幅值约束作了修改,提出了误差下降 (Error-Reduction, ER) 算法<sup>[12]</sup>,其空域估计值的更新方式为:

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y), & (x,y) \in \gamma \\ 0, & (x,y) \notin \gamma \end{cases} \quad (8)$$

式中:  $\gamma$  为  $g'_k(x,y)$  满足设定的空间域约束的区域。可以证明 ER 算法迭代过程中频域的损失函数一直在减小或者不变<sup>[13]</sup>,故整个迭代过程是收敛的。事实上,ER

算法只在开始迭代的几个周期内能够较快的收敛,但在后面的迭代过程中收敛缓慢。为进一步加快收敛速度, Fienup 在 ER 算法的基础上引入负反馈机制,提出了混合输入输出 (Hybrid Input Output, HIO) 算法<sup>[13]</sup>, 相比于 GS 算法同样更改了空域估计值的更新方式:

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y) & (x,y) \in \gamma \\ g_k(x,y) - \beta g'_k(x,y) & (x,y) \notin \gamma \end{cases} \quad (9)$$

式中:  $\beta$  为一个取值在 0 和 1 之间的常数, 通常取值 0.9<sup>[14]</sup>。

HIO 算法中,  $f_k(x,y)$  的估计值并不是  $g_k(x,y)$ , 而是  $g'_k(x,y)$ , 因此只有空域损失函数  $E_{0k}^2$  才有意义。但是  $E_{0k}^2$  无法衡量输出面上的图像质量, 鉴于该原因 Fienup 等人建议 ER 算法和 HIO 算法交替使用<sup>[15]</sup>, 即在一个周期内前 5~10 次迭代采用 ER 算法, 后 20~50 次迭代采用 HIO 算法, 重复多个周期。随后 Guo 等人<sup>[16]</sup> 又发现 ER 算法和 HIO 算法交替使用能够提升算法的抗噪声能力和收敛速度。由于 HIO 算法的输入估计值不是输出估计值的连续函数, 迭代过程中图像会出现振荡现象。为此, Fienup 提出了连续的 HIO 算法<sup>[17]</sup>, 进一步提高了收敛精度。

除了引入负反馈机制到空域振幅限制, 许多研究者提出另一种放松振幅限制的方法——将目标平面划分为信号区域和背景区域, 并赋予不同的限制策略, 从而有效改善信号区域的图像质量<sup>[18-20]</sup>。其他形式的面向光束整形的 IFTA 可参考 Ripoll 等人的综述文章<sup>[21]</sup>。

除 IFTA 外, 相位恢复问题可以视为一个最优化求解问题, 即优化相位分布以实现目标面上辐照度分布尽可能接近于所需的辐照度分布。不同的优化算法相继在衍射光学元件设计领域得到了应用, 如模拟退火 (Simulated Annealing, SA) 算法<sup>[22-23]</sup>、遗传 (Genetic Algorithm, GA) 算法<sup>[24]</sup>、粒子群 (Particle Swarm Optimization, PSO) 算法<sup>[25]</sup> 等。这些算法拥有跳出局部最优解的能力, 但收敛速度往往较慢。

## 2 几何光学-物理光学复合设计方法

传统的物理光学设计方法往往都是以随机相位作为优化起点, 优化过程中也没有考虑相位的平滑性, 导致最终得到的 DOE 表面常常包含许多高频特

征, 一方面难以精确加工, 另一方面也非常容易受杂散光的影响。同时, 随机相位分布包含许多相位突变和相位涡旋, 这会使得目标面重建图像出现许多黑点状的散斑<sup>[32]</sup>。而优化过程中相位涡旋难以自湮灭, 导致收敛停滞<sup>[33]</sup>。

利用二次相位作为物理光学方法的优化起点能有效抑制散斑噪声从而避免收敛停滞问题<sup>[34-39]</sup>, 二次相位的系数可根据目标面所需光斑尺寸确定<sup>[38]</sup>。事实上, 二次相位可以看作是一个粗略的几何光学相位。进一步地, Schmidt 等人提出使用低阶  $xy$  多项式的形式近似表示几何光学相位<sup>[26]</sup>:

$$\varphi_{\text{Ref.}}(x,y) = ax + by + cx^2 + dxy + ey^2 \quad (10)$$

式中: 系数  $a \sim e$  可以通过优化得到, 以在目标面上产生一个最接近所需辐照度分布的椭圆轮廓光斑。尽管如此, 这些初始二次相位只能在目标面上生成一个模糊光斑, 并不是精确的几何光学相位解。

2005 年, Kaempfe 等人<sup>[27]</sup> 首先采用迭代网格自适应算法求解一个入射面和目标面能量之间的映射关系, 然后根据这个映射关系求解了一个平滑的几何光学相位解。之后, 用该相位作为 IFTA 优化的起点进行迭代, 同时为了保证迭代过程中的相位分布尽可能保持平滑, 将目标辐照度分布设置为所需辐照度分布和重建辐照度分布的加权叠加, 并随着迭代次数的增加缓慢增大所需辐照度分布的权重系数。最终优化得到的相位分布比物理光学方法更平滑, 同时重建图像质量比几何光学方法更好, 体现了复合设计方法的优势。

2015 年, Feng 等人<sup>[28]</sup> 首先数值求解几何光学近似下的蒙日-安培 (Monge-Ampère, MA) 方程获得几何光学相位分布, 随后采用了一种适合光束整形的改进过补偿 (Modified Over Compensation, MOC) IFTA 进一步迭代优化, 得到最终的相位分布。该方法由于采用了先进的最优输运求解算法<sup>[40]</sup>, 针对复杂形式的目标面辐照度分布, 也可以较为快速地求解出相应的几何光学相位分布。另外, 该方法采用的对传统 IFTA 的过补偿策略可以进一步加快算法收敛速度, 提升光束整形质量。

根据文献 [3,28], 下面将简单介绍几何光学相位分布的求解过程。图 2 给出了光束整形问题的几何



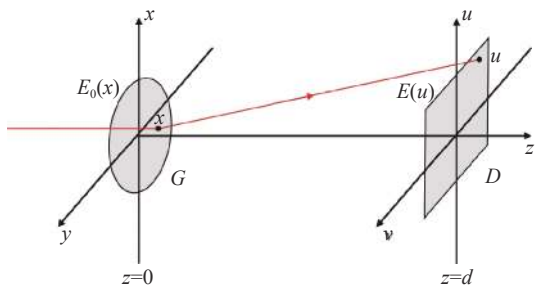


图 2 光束整形问题的几何示意图

Fig.2 Geometry diagram of the problem of beam shaping

示意图,入射光经过光学元件调控后在焦平面或者指定目标面上形成特定的照度分布。在非涅尔近似条件下,目标面上的光场 $E(\mathbf{u})$ 与入射面光场 $E_0(\mathbf{x})$ 之间的关系为:

$$E(\mathbf{u}) = \frac{e^{ikd}}{i\lambda d} \iint_{\Omega} E_0(\mathbf{x}) e^{i\varphi(\mathbf{x})} e^{i\frac{k}{2d}|\mathbf{u}-\mathbf{x}|^2} d\mathbf{x} \quad (11)$$

式中: $\mathbf{x}=(x,y)$ 为入射平面区域 $G$ 上的点; $\mathbf{u}=(u,v)$ 为目标平面区域 $D$ 上的点; $d$ 为光场传播距离; $\varphi(\mathbf{x})$ 为DOE的相位分布; $k$ 为波矢。公式(11)所述积分可以改写为:

$$\iint g(\mathbf{x}) \exp[ikq(\mathbf{x})] d\mathbf{x} \quad (12)$$

根据稳定相法<sup>[41]</sup>可知对公式(12)积分产生显著贡献的点为鞍点,有 $\nabla q(\mathbf{x})=0$ ,从而可以推导出相位函数的梯度:

$$\nabla \varphi(\mathbf{x}) = \frac{k}{d}(\mathbf{u}-\mathbf{x}) \quad (13)$$

$$\begin{aligned} \text{令 } \psi(\mathbf{x}) &= \frac{d}{k}\varphi(\mathbf{x}) + \frac{1}{2}|\mathbf{x}|^2, \text{ 则有:} \\ \nabla \psi(\mathbf{x}) &= \mathbf{u} \end{aligned} \quad (14)$$

此时,根据能量守恒可得:

$$I_0(\mathbf{x})d\mathbf{x} = I(\mathbf{u})d\mathbf{u} \quad (15)$$

式中: $I_0(\mathbf{x})$ 和 $I(\mathbf{u})$ 分别为入射面和目标面上的辐照度分布。

将公式(14)代入公式(15)能够得到一个带有边界条件的标准MA方程<sup>[28]</sup>:

$$\begin{cases} \det(D^2\psi(\mathbf{x})) = I_0(\mathbf{x})/I(\nabla\psi(\mathbf{x})) \\ \nabla\psi: G \rightarrow D \end{cases} \quad (16)$$

该方程可采用最优输运光线映射方法<sup>[40]</sup>进行求解,最终计算出所需的几何光学相位函数。随后采用MOC算法作进一步迭代优化,该算法在传统IFTA的

基础上对目标区域的光场振幅进行过补偿,可有效提升目标面照度图像的质量。

图3给出了不同方法设计的自由曲面和对应的目标面辐照度分布。几何光学相位分布对应的自由曲面面形和辐照度分布分别如图3(a)和图3(b)所示,可见衍射效应会导致整形效果退化。经过MOC算法进一步优化得到的自由曲面面形和对应辐照度分布如图3(c)和图3(d)所示。虽然此时曲面光滑度有所下降,但是目标面上的整形效果非常精细,包含了许多细节信息。复合设计方法相比于纯MOC算法(图3(e)和图3(f))和GS算法(图3(g)、图3(h))整形质量更好,曲面更平滑。

类似地,2021年Yang等人<sup>[29]</sup>介绍了一种映射型傅里叶对合成方法,可通过最优输运算法求解出平滑的相位分布,并通过仿真表明其对于随后的IFTA是合理的初始估计。

尽管上述复合方法设计得到的几何光学相位是连续的,但并没有控制后续IFTA迭代过程中相位的变化,导致最终优化得到的DOE表面仍然包含一些高频特征。为此,Schmidt等人<sup>[26]</sup>提出能够严格控制IFTA迭代过程相位表面起伏的复合算法,并将该复合算法应用在自由曲面全息图的设计上。首先采用一种光线映射方法<sup>[42]</sup>得到全息图几何光学相位分布 $\varphi_{\text{Ref.}}(x,y)$ ,随后为保持后续IFTA迭代过程中全息图相位的平滑,对迭代过程中相位的变化量 $\varphi_{\text{Diff.}}(x,y)$ 做高斯滤波处理,相位分布的更新方式为<sup>[26]</sup>:

$$\begin{aligned} \varphi^{k+1}(x,y) &= \varphi_{\text{Ref.}}(x,y) + \tilde{\varphi}_{\text{Diff.}}^k(x,y) * e^{-\frac{x^2+y^2}{2\sigma^2}} \\ \tilde{\varphi}_{\text{Diff.}}^k(x,y) &= \tilde{\varphi}^k(x,y) - \varphi_{\text{Ref.}}(x,y) \end{aligned} \quad (17)$$

式中: $\varphi^{k+1}(x,y)$ 为第 $k+1$ 次IFTA迭代的初始相位; $\varphi_{\text{Ref.}}(x,y)$ 为几何光学相位; $\tilde{\varphi}_{\text{Diff.}}^k(x,y)$ 为第 $k$ 次IFTA迭代后的相位变化量,可以视为在几何光学相位上叠加的物理光学相位; $\tilde{\varphi}^k(x,y)$ 为第 $k$ 次IFTA迭代后的相位;“\*”代表卷积运算。该更新方式保证了IFTA迭代过程中全息图相位的平滑,同时能够根据加工水平选择合适的平滑权重 $\sigma$ 以控制表面的起伏程度。

图4(a)为几何光学相位对应的折射表面,此时远场强度图像比较模糊(图4(b))。经过IFTA平滑迭代后得到最终的自由曲面全息图(图4(e)),该自由曲面全息图可看作是折射表面和衍射表面(图4(c))的叠

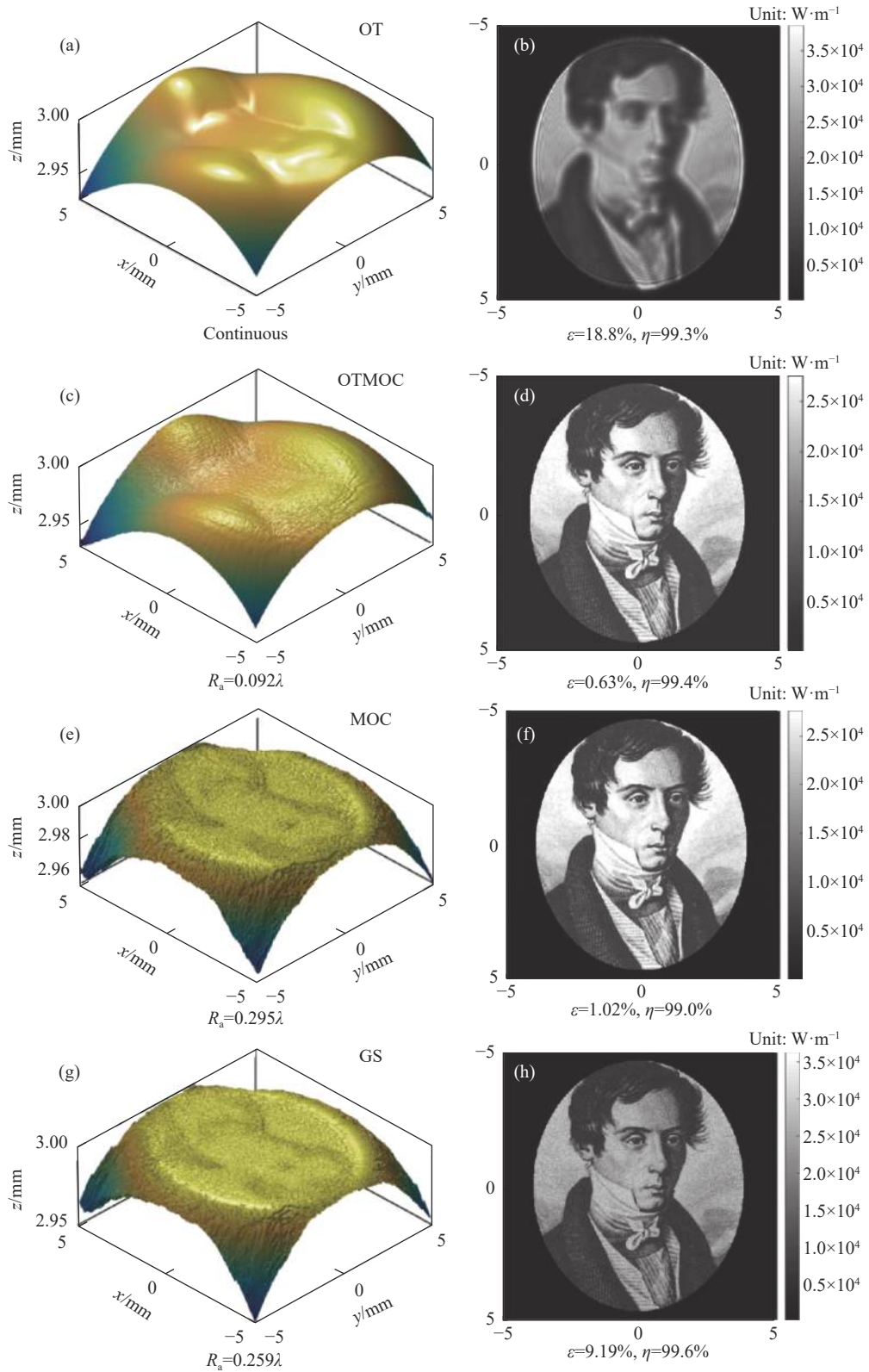


图 3 Feng 等人<sup>[28]</sup>提出的复合设计方法结果:不同方法设计的自由曲面和对应的辐照度分布,其中复合方法(OTMOC)同时实现了表面光滑和图像清晰

Fig.3 The results of the composite design method proposed by Feng et al<sup>[28]</sup>. Freeform surfaces designed by different methods and their corresponding irradiance distributions, where the composite method (OTMOC) achieves both surface smoothness and image clarity simultaneously. (Adapted with permission from<sup>[28]</sup>. © 2015 Optical Society of America)

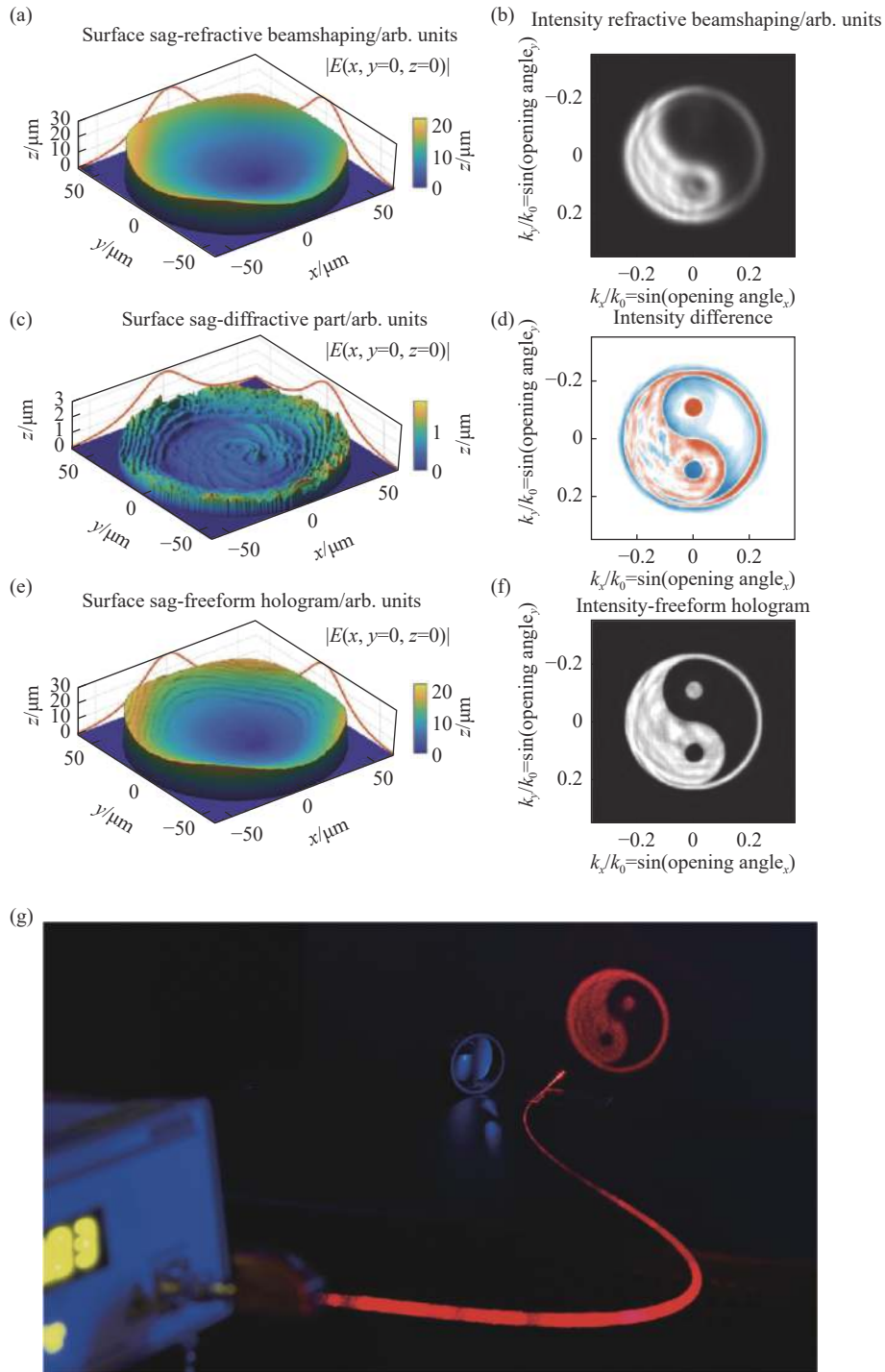


图 4 Schmidt 等人<sup>[26]</sup>提出的复合型设计方法结果: 将单模光纤出射的高斯光束在远场转化为太极阴阳图光斑的折射部分面形 (几何光学结果)(a)、纯衍射部分面形 (c) 和折射-衍射叠加的自由曲面全息图 (e); 几何光学方法和复合方法对应的强度分布 (b)、(f) 及两者之间的差异 (d); 实现的远场强度分布 (g)

Fig.4 The results of the composite design method proposed by Schmidt et al<sup>[26]</sup>. Surface sags of the refractive part (geometric optics solution) (a), the pure diffractive part (c), and the freeform hologram (e) as a refractive-diffractive superposition to redistribute the initial Gaussian eigenmode emitted from a single mode fiber into a Yin Yang pattern in the far field; Simulated intensity distributions of the geometric (b) and composite (f) methods, and the difference (d); The realized farfield intensity distribution (g). (Adapted with permission from<sup>[26]</sup>. © 2020 Optical Society of America)

加, 对应远场强度图像是一个轮廓清晰的阴阳图 (图 4(f)), 实验结果如图 4(g) 所示, 验证了复合方法的有效性, 图 4(d) 代表了复合方法和几何光学方法得到的强度图像之间的差异。

类似地, Doskolovich 等人提出了结合程函函数求解和 IFTA 的复合算法<sup>[30]</sup>, 其中程函函数等于相位函数除以波矢  $k$ 。该方法基于费马原理将程函函数求解问题表述为一个最优化问题:

$$\Phi(\mathbf{x}) = \max_{\mathbf{u} \in G} (\Psi(\mathbf{u}) - \rho(\mathbf{x}, \mathbf{u})) \quad (18)$$

式中:  $\mathbf{x} = (x, y)$  为入射平面区域  $G$  上的点;  $\mathbf{u} = (u, v)$  为目标平面区域  $D$  上的点;  $\Phi(\mathbf{x})$  为入射面即 DOE 的程函函数;  $\Psi(\mathbf{u})$  为目标面的程函函数,  $\rho(\mathbf{x}, \mathbf{u})$  是入射面点  $\mathbf{x}$  和目标面点  $\mathbf{u}$  之间的距离。采用梯度下降算法求解

了程函函数  $\Phi(\mathbf{x})$ , 进而得到几何光学相位分布。随后, 与 Schmidt 等人类似, 采用了结合高斯滤波的 GS 算法进行了迭代优化。

设计过程中 DOE 的相位分布和对应的目标面辐照度分布结果如图 5 所示。图 5(a) 为几何光学相位, 此时产生的六芒星形式的辐照度分布 (图 5(b)) 边缘不够锐利。以此初始相位进行 GS 算法迭代得到的相位分布 (图 5(c)) 引入了一些高频噪声, 而采取结合高斯滤波的 GS 算法迭代则能够保证迭代后的相位分布 (图 5(e)) 保持平滑, 对应的辐照度分布 (图 5(f)) 相比几何光学方法 (图 5(b)) 有较大的提升。相比之下, 以简单二次相位作为初值进行 GS 算法迭代后的相位分布 (图 5(g)) 比较杂乱, 加工难度较大。

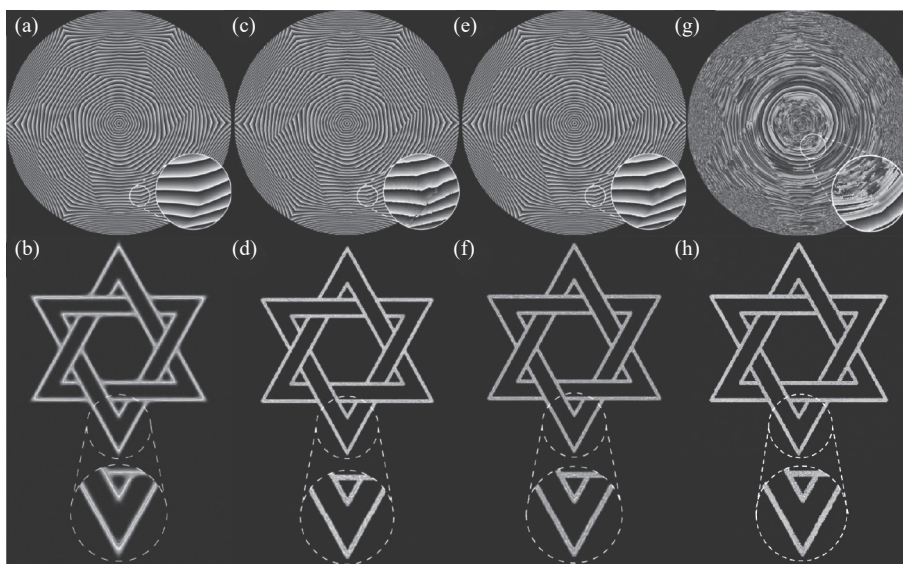


图 5 Doskolovich 等人<sup>[30]</sup>提出的复合设计方法结果: 几何光学相位分布 (a), 使用 GS 算法迭代得到的相位分布 (c), 使用相位平滑 GS 算法得到的相位分布 (e), 以及使用二次相位作为初始相位进行 GS 算法迭代得到的相位分布 (g); (b), (d), (f), (h) 是 (a), (c), (e), (g) 在目标面上产生的归一化辐照度分布

Fig.5 The results of the composite design method proposed by Doskolovich et al<sup>[30]</sup>. (a), (c), (e), (g) Geometric optics phase distribution (a) and the phase distributions obtained using the GS algorithms without (c) and with (e) phase smoothing. Phase distribution calculated using the GS algorithm with the quadratic initial phase (g); (b), (d), (f), (h) Normalized irradiance distributions generated in the target plane for phase distributions of Figs. (a), (c), (e) and (g), respectively. (Reprinted with permission from<sup>[30]</sup>. © 2021 Optical Society of America)

### 3 结束语

文中综述了光束整形 DOE 的物理光学设计方法和结合几何光学方法和物理光学方法的复合设计方法。其中复合设计方法既能克服几何

光学方法忽略衍射效应的缺点, 又可解决传统物理光学方法收敛易停滞、相位不规则的问题, 能够设计出易于加工、光束整形质量高的自由曲面衍射光学元件。表 1 总结了文中所述各类设计方法的优缺点。



表 1 文中所述 DOE 设计方法对比

Tab.1 Comparisons of the design methods mentioned in this paper

Methods	Advantages	Limitations
IFTAs <sup>[8-21]</sup>	Simple implementation; High calculation speed	Slow convergence; Rely on initial estimate; Complex and irregular phase
Global optimization methods <sup>[22-25]</sup>	The ability to jump out of local optima; Strong robustness	Time-consuming; Sensitive to parameters; Complex and irregular phase
Composite methods <sup>[26-30]</sup>	Fabrication-friendly; High performance	Time-consuming

从科学和工程的角度而言, 光束整形 DOE 设计方法仍具有研究价值和发展空间, 存在许多值得探索的方向, 简单概括如下:

(a) 几何光学相位求解过程复杂耗时, 需要研究更简洁、更有效的数值求解方法。

(b) 研究更快速准确的光场传播模型和数值计算方法, 特别是大衍射角场景的衍射计算。

(c) 目前 DOE 的常规加工方式有刻蚀技术、灰度掩膜法、单点金刚石车削技术<sup>[43]</sup>、激光直写技术<sup>[44]</sup>等。不同加工方式引起的加工误差有所不同, 需要对特定的加工方式研究对应的加工误差灵敏度低的设计方法, 以使加工后的 DOE 指标最大限度满足设计需求。

除上述复合型设计方法之外, 近年来基于物理光传输模型和自动微分技术<sup>[45-46]</sup>的梯度下降算法<sup>[47-51]</sup>开始应用于相位恢复领域并取得了成功。该方法借助自动微分技术对梯度进行自动求解, 体现了灵活高效的特点。如何基于自动微分技术设计易于加工的 DOE 也是值得深入研究的方向。

总之, 灵活准确的 DOE 设计方法能够大幅提升 DOE 的光束整形性能, 进一步提高 DOE 在工业、医疗、军事、安防等领域的应用水平。

参考文献:

[1] Dickey F M, Lizotte T E. Laser Beam Shaping Applications[M]. 2nd ed. Boca Raton: CRC Press, 2017.

[2] Romero L A, Dickey F M. Lossless laser beam shaping [J]. *Journal of the Optical Society of America A*, 1996, 13(4): 751-760.

[3] Romero L A, Dickey F M. Mathematical and Physical Theory of Lossless Beam Shaping[M]//Laser Beam Shaping: Theory and Techniques. 2nd ed. Boca Raton: CRC Press, 2018.

[4] Feng Z X, Huang L, Gong M L. Focal-plane irradiance tailoring using the concept of woofer-tweeter deformable mirrors [J]. *Optics Express*, 2014, 22(8): 8871-8879.

[5] Goodman J W. Introduction to Fourier Optics[M]. 3rd ed. Colorado: Roberts and Company Publishers, 2005.

[6] Wu R M, Liu P, Zhang Y Q, et al. A mathematical model of the single freeform surface design for collimated beam shaping [J]. *Optics Express*, 2013, 21(18): 20974-20989.

[7] Swanson G J. Binary Optics Technology: The Theory and Design of Multi-level Diffractive Optical Elements[M]. Lexington, Massachusetts: Massachusetts Institute of Technology, Lincoln Laboratory, 1989.

[8] Gerchberg R W. A practical algorithm for the determination of plane from image and diffraction pictures [J]. *Optik*, 1972, 35(2): 237-246.

[9] Misell D L. A method for the solution of the phase problem in electron microscopy [J]. *Journal of Physics D: Applied Physics*, 1973, 6(1): L6.

[10] Yang Guozhen, Gu Benyuan. On the amplitude-phase retrieval problem in optical systems [J]. *Acta Physica Sinica*, 1981, 30(3): 410-413. (in Chinese)

[11] Yang G Z, Dong B Z, Gu B Y, et al. Gerchberg-Saxton and Yang-Gu algorithms for phase retrieval in a nonunitary transform system: a comparison [J]. *Applied Optics*, 1994, 33(2): 209-218.

[12] Fienup J R. Reconstruction of an object from the modulus of its Fourier transform [J]. *Optics Letters*, 1978, 3(1): 27-29.

[13] Fienup J R. Phase retrieval algorithms: a comparison [J]. *Applied Optics*, 1982, 21(15): 2758-2769.

[14] Latychevskaia T. Iterative phase retrieval in coherent diffractive imaging: Practical issues [J]. *Applied Optics*, 2018, 57(25): 7187-7197.

[15] Fienup J R, Wackerman C C. Phase-retrieval stagnation problems and solutions [J]. *JOSA A*, 1986, 3(11): 1897-1907.

[16] Guo C L, Liu S, Sheridan J T. Iterative phase retrieval



- algorithms. I: optimization [J]. *Applied Optics*, 2015, 54(15): 4698-4708.
- [17] Fienup J R. Phase retrieval with continuous version of hybrid inut-output[C]//Frontiers in Optics. Optical Society of America, 2003.
- [18] Pasienski M, Demarco B. A high-accuracy algorithm for designing arbitrary holographic atom traps [J]. *Optics Express*, 2008, 16(3): 2176-2190.
- [19] Tao S H, Yu W X. Beam shaping of complex amplitude with separate constraints on the output beam [J]. *Optics Express*, 2015, 23(2): 1052-1062.
- [20] Chang C L, Xia J, Yang L, et al. Speckle-suppressed phase-only holographic three-dimensional display based on double-constraint Gerchberg-Saxton algorithm [J]. *Applied Optics*, 2015, 54(23): 6994-7001.
- [21] Ripoll O, Kettunen V, Herzig H P. Review of iterative Fourier-transform algorithms for beam shaping applications [J]. *Optical Engineering*, 2004, 43(11): 2549-2556.
- [22] Meister M, Winfield R J. Novel approaches to direct search algorithms for the design of diffractive optical elements [J]. *Optics Communications*, 2002, 203(1-2): 39-49.
- [23] Pang H, Yin S Y, Deng Q L, et al. Effective design of diffractive optical elements for beam splitting with SA algorithm[C]//SPIE, 2012, 8556: 334-341.
- [24] Zhou G Y, Chen Y X, Wang Z G, et al. Genetic local search algorithm for optimization design of diffractive optical elements [J]. *Applied Optics*, 1999, 38(20): 4281-4290.
- [25] Shokoooh-saremi M, Magnusson R. Particle swarm optimization and its application to the design of diffraction grating filters [J]. *Optics Letters*, 2007, 32(8): 894-896.
- [26] Schmidt S, Thiele S, Toulouse A, et al. Tailored micro-optical freeform holograms for integrated complex beam shaping [J]. *Optica*, 2020, 7(10): 1279-1286.
- [27] Kaempfe T, Kley E B, Tuennermann A. Hybrid approach to the design of refractive beam shaping elements[C]//SPIE, 2005, 5876: 163-175.
- [28] Feng Z X, Froese B D, Liang R G. Composite method for precise freeform optical beam shaping [J]. *Applied Optics*, 2015, 54(31): 9364-9369.
- [29] Yang L X, Badar I, Hellmann C, et al. Light-shaping design by a Fourier pair synthesis: the homeomorphic case [J]. *Optics Express*, 2021, 29(3): 3621-3630.
- [30] Dskolovich L L, Mingazov A A, Byzov E V, et al. Hybrid design of diffractive optical elements for optical beam shaping [J]. *Optics Express*, 2021, 29(20): 31875-31890.
- [31] Hessler T, Rossi M, Kunz R E, et al. Analysis and optimization of fabrication of continuous-relief diffractive optical elements [J]. *Applied Optics*, 1998, 37(19): 4069-4079.
- [32] Aagedal H, Schmid M, Beth T, et al. Theory of speckles in diffractive optics and its application to beam shaping [J]. *Journal of Modern Optics*, 1996, 43(7): 1409-1421.
- [33] Senthilkumar P, Wyrowski F, Schimmel H. Vortex stagnation problem in iterative Fourier transform algorithms [J]. *Optics and Lasers in Engineering*, 2005, 43(1): 43-56.
- [34] Shimobaba T, Ito T. Random phase-free computer-generated hologram [J]. *Optics Express*, 2015, 23(7): 9549-9554.
- [35] Pang H, Wang J Z, Cao A X, et al. High-accuracy method for holographic image projection with suppressed speckle noise [J]. *Optics Express*, 2016, 24(20): 22766-22776.
- [36] Chang C L, Qi Y J, Wu J, et al. Speckle reduced lensless holographic projection from phase-only computer-generated hologram [J]. *Optics Express*, 2017, 25(6): 6568-6580.
- [37] Wu Y, Wang J, Chen C, et al. Generation of phase-only Fourier hologram based on double phase method and quantization error analysis [J]. *IEEE Photonics Journal*, 2020, 12(5): 1-13.
- [38] Chen L Z, Zhang H, Cao L C, et al. Non-iterative phase hologram generation with optimized phase modulation [J]. *Optics Express*, 2020, 28(8): 11380-11392.
- [39] Tian S Z, Chen L Z, Zhang H. Optimized Fresnel phase hologram for ringing artifacts removal in lensless holographic projection [J]. *Applied Optics*, 2022, 61(5): B17-B24.
- [40] Benamou J D, Froese B D, Oberman A M. Numerical solution of the optimal transportation problem using the Monge-Ampère equation [J]. *Journal of Computational Physics*, 2014, 260: 107-126.
- [41] Bleistein N, Handelsman R A. Asymptotic Expansions of Integrals[M]. New York: Ardent Media, 1975.
- [42] Bösel C, Gross H. Ray mapping approach for the efficient design of continuous freeform surfaces [J]. *Optics Express*, 2016, 24(13): 14271-14282.
- [43] Huang Yuetian, Fan Bin, Li Shijie, et al. Error control of diffractive optical element fabricated by single point diamond turning [J]. *Infrared and Laser Engineering*, 2023, 52(3): 20220504. (in Chinese)
- [44] Liang Guanglei, Sun Shufeng, Wang Jin, et al. Research and application of diffractive optical element fabricated by laser direct writing [J]. *Infrared and Laser Engineering*, 2023, 52(4): 20220567. (in Chinese)

- [45] Baydin A G, Pearlmutter B A, Radul A A, et al. Automatic differentiation in machine learning: a survey [J]. *Journal of Machine Learning Research*, 2018, 18: 1-43.
- [46] Margossian C C. A review of automatic differentiation and its efficient implementation [J]. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2019, 9(4): e1305.
- [47] Peng Y F, Choi S, Padmanaban N, et al. Neural holography with camera-in-the-loop training [J]. *ACM Transactions on Graphics (TOG)*, 2020, 39(6): 1-14.
- [48] Chen C, Lee B, Li N N, et al. Multi-depth hologram generation using stochastic gradient descent algorithm with complex loss function [J]. *Optics Express*, 2021, 29(10): 15089-15103.
- [49] Wong A, Pope B, Desdoigts L, et al. Phase retrieval and design with automatic differentiation: Tutorial [J]. *JOSA B*, 2021, 38(9): 2465-2478.
- [50] Choi S, Kim J, Peng Y F, et al. Optimizing image quality for holographic near-eye displays with michelson holography [J]. *Optica*, 2021, 8(2): 143-146.
- [51] Chen L Z, Zhu R Z, Zhang H. Speckle-free compact holographic near-eye display using camera-in-the-loop optimization with phase constraint [J]. *Optics Express*, 2022, 30(26): 46649-46665.

## Design methods of freeform surface diffractive optics for beam shaping (*invited*)

Liao Qingming, Feng Zexin\*

(Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China)

### Abstract:

**Significance** Beam shaping plays an important role in many fields including laser material processing, medical treatment and laser fusion. The goal of beam shaping is to transform an incoming laser beam into a desired output irradiance (or intensity) distribution. Diffractive optical elements (DOEs) are one of the most promising ways for beam shaping. The design of DOEs plays a crucial role in high-quality beam shaping applications. To further promote the development of more advanced methods for designing DOEs which can better meet the requirements of different beam shaping applications, it is necessary to summarize the research progress of existing DOE design methods, discuss their advantages and disadvantages, and provide a necessary outlook.

**Progress** This review summarizes the design methods of phase-only DOEs for beam shaping. Since the DOE microrelief height function is lineally proportional to the phase function of the optical field generated by the DOE, the DOE design problem can be directly transferred into the calculation of the DOE phase distribution. There are two design methods of geometrical optics methods and physical optics methods to realize this goal. Geometrical optics methods usually generate continuous freeform optical surfaces. However, in many cases, the beam shaping quality can be degraded due to the diffraction effects. Physical optics methods, which describe the light propagation in a more accurate way, are commonly used to design phase-only DOEs for beam shaping. However, the iterative Fourier transform algorithms (IFTAs), which are the most commonly-used approach for designing DOEs, often generate complex and irregular shapes in DOE profiles. Such DOE profiles are difficult to fabricate and could generate speckles, which significantly impair the quality of the generated irradiance distribution. In addition, traditional IFTAs often suffers from slow convergence and iteration stagnation. Composite methods that combine the geometrical and physical optics methods have been proposed to address these issues. The freeform surfaces generated from the geometrical optics methods could provide good initial values for the following

IFTAs, significantly improving the convergence. The resulting optical surface profiles are more regular than those of the traditional IFTAs, which are easier to fabricate and could achieve high-quality beam shaping.

**Conclusions and Prospects** We have summarized (some of) the design methods of DOEs for beam shaping. After a brief recall of the traditional physical optics methods and their limitations, we have paid more attention to the review of the composite methods which can generate freeform DOEs that are easier to fabricate and could achieve high-quality beam shaping. Future directions of the DOE design methods include developments of fast geometrical optics solvers and wide-angle light propagation algorithms, more considerations of different fabrication techniques, and other promising methods based on auto-differentiation.

**Key words:** beam shaping; diffractive optical element; composite design methods; freeform surface; phase retrieval

**Funding projects:** National Natural Science Foundation of China (11704030); CAAI-Huawei MindSpore Open Fund (CAAIXSJLJJ-2021-014A, CAAIXSJLJJ-2022-010A)