

基于多分支融合复数网络的全息图像失真矫正

陈逸飞^{1,2}, 刘延伟^{1,2,3*}, 刘金霞⁴, 古晓艳^{1,2,3}

(1. 中国科学院信息工程研究所, 北京 100085; 2. 中国科学院大学网络空间安全学院, 北京 101408;
3. 网络空间安全防御全国重点实验室, 北京 100085; 4. 浙江万里学院, 浙江宁波 315100)

摘要: 全息显示技术可以再现出包含物体全部信息的三维成像, 能为用户带来高度逼真的视觉体验, 被认为是目前最理想的裸眼三维显示技术. 全息显示带来的独特三维沉浸式体验使得全息通信在医疗、教育、虚拟现实等领域具有广泛的应用前景. 但全息通信技术的大规模商业化应用还面临众多问题. 其中, 影响全息通信质量的一个主要问题就是在通信过程中压缩噪声与信道干扰等导致的多重混叠失真. 现有图像失真矫正技术大多聚焦单一失真类型, 难以应对复杂场景下的全息混合失真问题, 严重制约全息技术的实际应用效果. 针对这一难题, 本文提出一种基于多分支复数注意力网络的全息图像失真矫正方法, 通过构建分层并行的多分支网络结构, 实现对全息图像多尺度、多维度特征的深度提取与协同融合; 同时提出复数域自适应注意力机制, 强化网络对相位畸变、振幅衰减等关键失真特征的感知与抑制能力, 从而实现对压缩、传输等全链路失真的精准矫正. 在包含压缩和信道噪声等混合类型的全息失真矫正实验中, 相较于现有先进的深度学习失真矫正方法 SCUNet (Swin-Conv-UNet), 本文方法在峰值信噪比指标上提升 0.41 dB 以上, 结构相似性指标提升 0.006 以上, 有效抑制了振幅失真导致的亮度异常, 矫正了相位畸变, 显著提升了全息图像的重建质量.

关键词: 全息图像; 失真矫正; 复数网络; 复数注意力机制; 多分支融合网络

基金项目: 国家自然科学基金 (No.62371450)

中图分类号: TP391.41; TP18

文献标识码: A

文章编号: 0372-2112(2025)09-3319-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250549

Holographic Image Distortion Correction Based on Multi-Branch-Fused Complex-Valued Neural Network

CHEN Yi-fei^{1,2}, LIU Yan-wei^{1,2,3*}, LIU Jin-xia⁴, GU Xiao-yan^{1,2,3}

(1. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100085, China;

2. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 101408, China;

3. State Key Laboratory of Cyberspace Security Defense, Beijing 100085, China;

4. Zhejiang Wanli University, Ningbo, Zhejiang 315100, China)

Abstract: Holographic display technology can reproduce three-dimensional imaging that encompasses all information of an object, providing users with a highly realistic visual experience. It is regarded as the most perfect naked-eye 3D display technology currently available. The unique immersive 3D experience offered by holographic displays gives holographic communication has broad application prospects in fields such as healthcare, education, and virtual reality. However, the large-scale commercial application of holographic communication technology still currently confronts numerous obstacles. Among them, one major issue affecting the quality of holographic communication is the multiple aliasing distortions caused by compression noise and channel interference during hologram transmission. Existing image distortion correction techniques mostly focus on single distortion type and struggle to address the problem of mixed holographic distortions in complex scenarios, severely limiting the effectiveness of holographic technology in the practical applications. To tackle this issue, this paper proposes a holographic image distortion correction method based on a multi-branch complex-valued convolutional neural network. This method constructs a multi-level parallel multi-branch network architecture to achieve in-depth extraction and collaborative fusion of multi-scale and multi-dimensional distortion features of holographic images. Simulta-

neously, a complex-valued adaptive attention mechanism is proposed to enhance the network's perception and suppression capabilities for key distortion features such as phase distortion and amplitude attenuation, thereby achieving precise correction of end-to-end distortions caused during compression and transmission. In the experiments involving mixed-type holographic distortions including compression and channel noises, compared to the state-of-the-art deep learning distortion correction method SCUNet (Swin-Conv-UNet), the proposed method achieves an average improvement of over 0.41 dB in peak signal-to-noise ratio (PSNR) and an average increase of over 0.006 in structural similarity index (SSIM). These experimental results show that the proposed method can effectively suppress the brightness abnormalities caused by amplitude distortion, correct the phase distortions, and significantly enhance the reconstruction quality of holographic images.

Key words: holographic image; distortion correction; complex-valued neural network; complex-domain attention mechanism; multi-branch-fused complex-valued network

Foundation Item(s): National Natural Science Foundation of China (No.62371450)

1 引言

在沉浸式通信逐步向消费市场渗透的背景下,全息技术凭借革新用户与数字内容交互模式的潜力,成为推动其发展的关键技术^[1-3].相较于传统立体显示中普遍存在的调节距离与会聚距离失配问题,全息成像通过完整保留三维场景的真实特征,有效规避视觉不适,实现对现实3D场景的高度还原.

全息图作为电磁波与真实物体干涉的记录产物,不仅包含光场振幅信息,还保留了关键的相位信息,其再现通过参考波照射来实现.在全息图获取方式中,光学装置虽能精准获取真实物体的三维信息,但对设备与环境要求严苛^[4];而计算机生成全息图(Computer-Generated Holography, CGH)^[5]则基于菲涅耳衍射/角谱传播算法等理论,通过构建三维物体数学模型,计算物体光波空间分布并叠加参考光波生成全息图,最终借助空间光调制器(Spatial Light Modulator, SLM)等设备实现光学再现.这种技术摆脱了真实物体与光学条件的限制,能够灵活生成虚拟物体,在增强现实(Augmented Reality, AR)、3D显示、显微成像等领域展现出显著的应用价值.如图1所示,进一步发展的全息通信^[6,7]技术,则可实现真实三维场景或物体的远程传输,使接收方获得高度拟真的交互体验,在远程会议、教育、医疗及虚拟现实娱乐等场景中具备提升信息传递效率与质量、增强协作效果的巨大潜力.

尽管全息通信的应用前景非常广阔,但全息图像在通信过程中导致的内容失真使得全息体验变差.首先,全息压缩会导致一定的内容损失,这直接影响到全息图像的重构过程. JPEG (Joint Photographic Experts Group) 和 HEVC (High Efficiency Video Coding) 等压缩方法,由于对高频分量的量化误差比较敏感,在处理全息图密集的高频干涉条纹信息时,不可避免地导致重构图像出现边缘模糊与相位畸变^[8].在重构过程中,这些丢失的高频信息难以恢复,使得原本清晰的物体轮廓变得模糊,物体的三维空间关系也因相位畸变而发



图1 全息通信应用场景

生扭曲.基于卷积神经网络(Convolutional Neural Network, CNN)^[9,10]的压缩方案同样面临失真困境.全息图作为复振幅数据,包含振幅与相位双重关键信息,而传统CNN将复数拆分为实部和虚部进行分离处理,破坏了复信号内在关联性,致使编码三维结构与景深的相位信息无法被显式建模,在重构图像时,导致三维景深紊乱,细微结构无法正常呈现^[11].

其次,在传输过程中的信道失真^[12]对全息图像重构过程的影响同样不可忽视.如传输信道中的高斯白噪声干扰会导致全息信号质量下降,表现为全息图像上出现颗粒状的噪声.这些噪声会掩盖图像的细节信息,在重构时干扰算法对原始图像信息的提取,降低图像的清晰度和视觉效果.除了影响图像的清晰度,幅频损失会使颜色信号的幅值发生改变,进而导致颜色的饱和度、色调等出现偏差.在图像重构阶段,颜色信息的偏差使得最终呈现的全息图像颜色看起来不自然,严重影响全息重构图像的质量和可读性.

因此,在接收端进行全息失真矫正至关重要,它不仅能够修复图像质量,还能通过优化数据特征,变相提升数据压缩效率,减少网络传输压力与设备处理负担.遗憾的是,当前全息通信领域尚未形成成熟的失真矫正方案.传统图像失真矫正方法大多基于单模态数据,

仅关注振幅信息,难以精准捕捉全息图中复信号的相位-振幅耦合特性,导致矫正效果大打折扣.有鉴于此,本文首次提出基于深度学习的全息失真矫正方法,突破传统失真补偿方法的单一模态处理局限,提出复数域多分支网络协同的失真矫正思想,通过构建融合结构解析、多分支交互与全局特征感知的统一框架,实现对全息图像在失真场景下的有效矫正,具体贡献如下.

(1)本文通过分析全息失真在复数特征空间的本质表征——失真均表现为复振幅信号的相位-振幅耦合关系破坏,接着打破传统“失真类型→特定算法”的一对一设计模式,提出多分支协同与复数 Transformer 机制,同时修复振幅域局部细节与相位域全局结构,实现对多类失真的“同源异径”处理;通过自动适配不同失真类型的主导特征,避免重复建模,真正实现“一网络应对多失真”的泛化能力.

(2)针对传统失真矫正方法难以捕捉全息复振幅信号内在关联性、无法区分非焦平面天然模糊与真实失真的局限性,本文提出复数域自注意力机制并用于全息图像失真矫正.采用复数 Query-Key-Value 映射模块,通过实部与虚部联合编码直接处理全息图像的复振幅数据,并且动态地分配卷积权重,从而挖掘全息图像不同块之间的关系,实现失真特征与非焦平面重构模糊特征的有效区分.

2 相关工作

图像失真矫正是指在数码相机采集的受损图像或压缩失真图像等情况下恢复出清晰图像的过程,这一过程为自动驾驶和图像识别等高级视觉任务提供了重要的质量保障.目前,大多数失真矫正方法基于 $I_n = I_c + n$ 的退化模型^[13]来实现失真矫正.其中, I_n 被符号化为给定的失真图像, n 被定义为失真噪声, I_c 用于表示干净的图像.由此可见,图像失真矫正任务本质上是一个不适定^[14]问题.

2.1 传统图像失真矫正

基于端到端架构的 CNN^[15]凭借强大的学习能力被用于图像失真矫正^[16],主要可以归纳为单一网络架构和多网络架构两类方法.单一网络架构通过平移不变性和小尺寸滤波器提取图像局部特征.例如, DnCNN (Denoising Convolutional Neural Network)^[17]利用残差学习处理多种低级视觉任务中的噪声干扰,文献[18]引入递归与门控单元增强网络记忆能力来进行图像恢复,而文献[19]则采用复值变换恢复图像的细节信息.多网络架构通过整合多个 CNN,实现复杂场景下的特征挖掘与噪声滤除. Tian 等人^[20]构建异构网络捕捉互补信息,而文献[21]则融合多尺度操作和残差密集技术,在抑制噪声的同时保留纹理.稀疏思想指导的双

CNN^[22]平衡了矫正性能与效率.此外,在医学图像领域, CNLNet (Convolutional Non-Local Network)^[23]结合非局部自相似性与 CNN 提升彩色图像去噪效果, GCBD (Generative Color Blind Denoising)^[24]和 CBDNet (Convolutional Blind Denoising Network)^[25]分别基于生成网络和双网架构解决盲降噪与不成对数据处理问题.文献[26]提出了一种 Swin-Conv 块,将残差卷积层的局部建模能力和 Swin-Transformer 块^[27]的非局部建模能力结合起来,然后将其作为主要构建块插入 UNet 架构中,取得了较好的失真矫正效果.

上述的失真矫正神经网络架构原本就是面向传统图像而设计的,可以很好地对传统图像进行处理,但对全息图像在复数空间失真的本质认识不足,无法精确地捕捉复振幅信号的相位-振幅耦合关系,导致矫正效果不理想.

2.2 全息失真矫正

目前全息图像失真矫正方法较少.文献[28]通过对比光学重建点测试图与原始图估计失真系数,对输入图像进行预失真处理,实现了广角全息投影仪的失真最小化,可重建高质量图像.但光学系统的复杂性易导致失真系数估计偏差,且仅适用于特定类型设备,通用性不足.而基于角谱算法的方法^[29]利用相机和四步相移算法采集波前,计算失真分布,再通过空间光调制器消除失真,针对计算机生成全息图像显示的失真校正稳健性强.然而,该方法需要搭建复杂的采集和校正系统,实现成本和技术难度较大.针对透镜阵列全息光学的 3D 显示系统中透镜失真、横向失真和几何变形,文献[30]分别采用失真补偿和透视变换处理,能有效减少投影图像失真,提升 3D 重建效果.然而,面对多种失真相互影响的复杂情况,校正精度和效果易受制约.综上所述,现有方法都依赖于光学器件,针对全息图像失真矫正的局限源于对“全息图像失真物理本质-数据特征空间-网络架构设计”三元关系的耦合建模不足,未将光学传播影响、复分析理论等物理先验融入网络架构,导致效果不佳.

3 全息失真矫正网络

全息图像在相干光照射下可实现重建,当成像平面与计算机生成全息图像平面间距等于物体平面与全息图像平面间距,且采用理想平面波作为重建波时,可获得与原目标物体完全一致的重建图像.然而,复数实部和虚部格式的全息数据经过压缩传输后不可避免会造成信息丢失,导致理想平面波穿过失真全息图时相位分布改变,使失真从全息图像平面传播至重建图像,造成重建图像与原物体出现差异.因此,全息失真在复数域表现为相位和振幅信息耦合方面的误差,进

一步演化为全息重建图像振幅的损失. 为简化模型, 本文将压缩传输等过程所造成的失真对原始全息图像的影响等效为全息图像平面上透射率函数的作用, 即照明波可视为先穿过附加分布 $t(x, y, 0)$ 后再照射全息图像平面, 如图 2 所示. 压缩传输造成的全息图像失真具有非固定性, 附加分布会随图像变化而改变, 需要针对不同图像进行计算, 并在重建过程中将预估分布的逆失真分布 $t^{-1}(x, y, 0)$ 叠加至对应全息图像, 从而完成失真矫正. 针对这种动态的失真特性, 本文提出一种基于多分支融合复数网络的全息图像失真补偿机制, 利用神经网络自适应地学习逆失真分布以实现精准的失真校正.

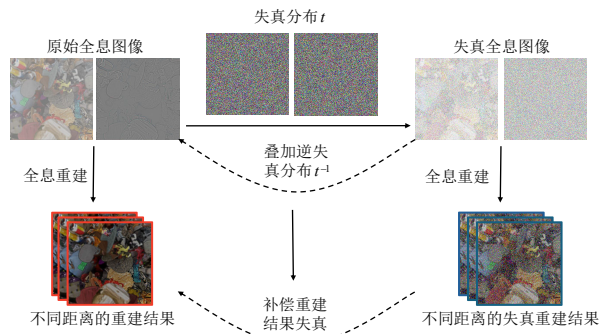


图 2 失真对全息重构影响分析

3.1 复数注意力机制

在全息图像重构过程中, 由于物理成像过程的限制, 不仅存在压缩失真等常规影响, 还面临非焦平面区域天然模糊的干扰. 若直接对重构图像进行传统失真矫正, 算法难以有效区分需要修复的失真部分与因重构产生的模糊区域, 极易将两者混淆处理, 导致在矫正失真的同时破坏全息图像特有的 3D 效果. 为此, 本文提出复数注意力机制, 进而构建复数 Transformer 模块 (Complex Transformer Block, CTB), 通过复数域的独特表征能力, 精准捕捉全息图像复振幅信号的相位-振幅耦合关系, 实现失真特征与重构模糊特征的有效分离.

给定复数输入 $X = A + iB$, 在不同的时间步长处理高维信息. 计算查询矩阵 $Q = XW_Q$ 、键矩阵 $K = XW_K$ 和价值矩阵 $V = XW_V$ (其中 Q, K, V 均是复值, W 矩阵也为复值) 后, 复数注意力机制定义如下: A' 和 B' 分别表示复注意力层输出的实部和虚部. 为了更好地解析实部和虚部之间的内部相似性, 对于公式中的每个项, 计算了多头注意力, 如式 (2) ~ 式 (5) 所示. 图 3 直观地给出了复数注意力机制的计算逻辑. 其中, 缩放因子 d_k 是 Q 和 K 的特征维数 (注意 Q, K, V 是复值矩阵, 而 $Q/K/V$ 则是可以插入 A 或 B 中的实值占位符矩阵). 基于 V 对 Q 和 K 的依赖关系, $\text{Attention}(Q, K, V)$ 可以直观地被视为 V

的扩展和权重调整的表示. $\text{MultiHead}(Q, K, V)$ 为不同子空间中的依赖关系提供了更好的解决方案, 其中 W^O 为点乘权重矩阵. 为了防止梯度爆炸, 式 (4) 中采用 Min - Max - Norm 而不是 Softmax.

$$\begin{aligned} QK^T V &= (XW_Q)(XW_K)^T(XW_V) \\ &= (AW_Q + iBW_Q)(W_K^T A^T + iW_K^T B^T)(AW_V + iBW_V) \\ &= (AW_Q W_K^T A^T AW_V - AW_Q W_K^T B^T BW_V \\ &\quad - BW_Q W_K^T A^T BW_V - BW_Q W_K^T B^T AW_V) \\ &\quad + i(AW_Q W_K^T A^T BW_V + AW_Q W_K^T B^T AW_V \\ &\quad + BW_Q W_K^T A^T AW_V - BW_Q W_K^T B^T BW_V) \\ &= A' + iB' \end{aligned} \quad (1)$$

$$\begin{aligned} \text{ComplexAttention}(X) &= (\text{MH}(A, A, A) - \text{MH}(A, B, B) \\ &\quad - \text{MH}(B, A, B) - \text{MH}(B, B, A)) \\ &\quad + i(\text{MH}(A, A, B) + \text{MH}(A, B, A) \\ &\quad + \text{MH}(B, A, A) - \text{MH}(B, B, B)) \end{aligned} \quad (2)$$

$$\begin{aligned} \text{MH}(Q, K, V) &= \text{MultiHead}(Q, K, V) \\ &= \text{Concat}(\text{Attention}(QW_i^Q, KW_i^K, VW_i^V)_{i=1}^n)W^O \end{aligned} \quad (3)$$

$$\text{Attention}(Q, K, V) = \text{Min - Max - Norm} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

$$\text{Min - Max - Norm}(X) = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (5)$$

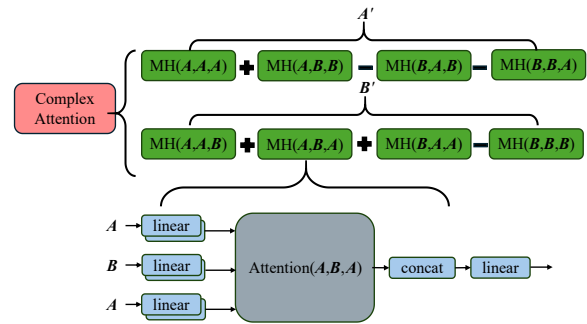


图 3 复数注意力机制的计算逻辑

基于复数注意力机制而设计的 Transformer 模块由编码和解码两部分构成. 复数 Transformer 编码器由 6 个相同的堆栈组成, 每个堆栈包含两个子层. 第一个子层是一个复数注意力层, 第二个子层是复数前馈网络, 这两个子层都有残差连接和层归一化. 在编码器中的残差连接之前采用层归一化, 即“Norm & Add”. 复数 Transformer 解码器也有 6 个相同的堆栈. 每个堆栈有 3 个子层: 复数注意力层、复数前馈网络和另一个复数注意力层. 第一个复数注意力层被额外的对角线掩码遮蔽, 以防止关注后续位置. 第二个复数注意力层的输入来

自编码器的输出和解码器的输入. 这种层级化设计旨在通过多层级的特征提取与交互, 逐层增强模型对全息图像复杂失真特征的解析能力, 同时利用复数注意力机制捕获真正需要进行失真矫正的部分, 从而避免对重构模糊区域的误判与误修复, 确保 3D 成像效果的完整性与准确性.

3.2 多分支融合复数网络架构

当压缩和传输等环节导致的全息失真附着在全息图像上一同进行角谱法衍射时, 很难对衍射后的多类型失真进行具体建模. 因此, 本文提出一种基于复数 Transformer 的多分支融合网络用于全息图像失真校正, 如图 4 所示. 基于复数 Transformer 的多分支融合网络利用从不同串行和并行复数网络中通过注意力方式获得的不同结构信息在广度和深度上进行交互来提取显著特征, 以更好地对失真图像进行矫正, 构建高保真的图像. 它的失真矫正能力主要通过复数串行块、混合并行块和复数残差块来实现. 复数串行块使用线性和非线性分量来深入搜索图像失真的关键信息. 为了获得更多来自不同视图的互补信息, 利用混合并行块从 3 个异构网络(即子网 1、子网 2 和子网 3)获得交互交叉特征并广泛搜索像素之间的相关性信息, 以提高失真矫正网络对复杂场景的适应性. 为了增强失真矫正网络的鲁棒性, 在复数串行块和混合并行块中嵌入了复数 Transformer 注意力机制, 以提取更显著的特征, 从而有效地滤除噪声. 最后, 基于复数 Transformer 的多分支融合网络利用复数残差块将获得的特征映射转换为矫正后的图像. 为了直观地理解上述过程, 使用以下公式描述:

$$\begin{aligned} I_c &= \text{CCNNCT}(I_N) \\ &= \text{CRB}(\text{HPB}(\text{CSB}(I_N))) \end{aligned} \quad (6)$$

其中, CSB 表示复数串行块(Complex Serial Block, CSB), HPB 表示混合并行块(Hyper Parallel Block, HPB), CRB 表示复数残差块(Complex Residual Block, CRB), I_N 表示失真全息图像, I_c 表示经过失真矫正网络后输出的结果.

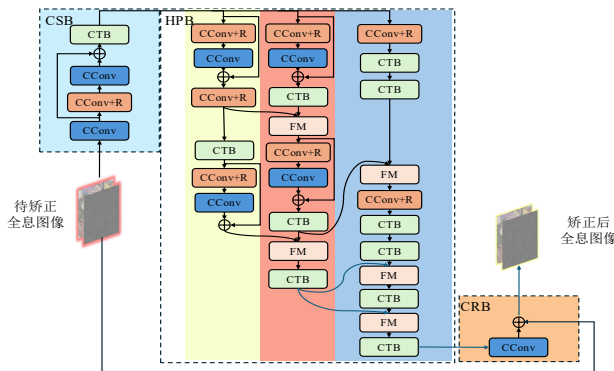


图 4 基于复值 Transformer 的多分支融合网络

具体地, 由 CSB、HPB 和 CRB 构成的多分支失真矫正网络的模块化设计如图 4 所示. CSB 通过深度结构信息搜索实现失真抑制, 包含复数卷积模块(CConv)、复数卷积-激活模块(CConv+R)和复数 Transformer 模块(CTB). 其中, CConv 负责将失真全息图像映射为复数域线性特征; CConv+R 通过线性与非线性特征组合提取多层特征; CTB 基于复数 Transformer 架构, 结合 JPEG 和 HEVC 的编码特性, 通过由两层复数多头自注意力(Complex-valued Multi-Head Self-Attention, CMHSA)构成的 4 层编码器和两层复数通道特征增强模块(Complex Channel Feature Enhancement, CCFE)组成挖掘图像块间复数域关系, 实现失真矫正. HPB 则融合了实数域与复数域信息, 包含 3 个异构子网络: 子网 1 通过卷积(Conv)与 Transformer 模块提取振幅信息, 子网 2 利用深度可分离卷积和特征融合模块(FM)实现相位信息提取与交互, 子网 3 通过复数残差学习与特征融合增强复数域特征. CRB 由单层复数卷积层和残差模块构成, 通过融合原始与预测复数图像实现清晰重构.

综上所述, 本文提出的网络架构设计真正体现出了以下理念: (1) 通过复数域运算直接处理复振幅数据, 能够有效捕捉相位信息; (2) 利用残差连接缓解深层网络梯度消失问题, 提升训练稳定性; (3) 结合多分支特征交互与模块化设计, 解决全息图像失真矫正中的多模态特征耦合与长距离依赖问题, 能够提升复杂场景适应性及失真矫正效率.

3.3 损失函数

多分支融合复数网络充分考虑了全息图像应用处理的优化诉求, 通过多目标联合优化的方式, 使得提出的失真矫正模型能够在训练过程中同时优化多个损失函数, 从而提高模型的整体性能.

全息失真矫正的误差损失. 使用均方误差(Mean Squared Error, MSE)作为损失函数来评估矫正的全息图的保真度, 有

$$L_{\text{holo}} = \left\| A_{\text{out}} - A_{\text{gt}} \exp\left(i\left(\delta(\phi_{\text{out}}, \phi_{\text{gt}}) - \bar{\delta}(\phi_{\text{out}}, \phi_{\text{gt}})\right)\right) \right\|_2^2 \quad (7)$$

其中, A_{out} 和 ϕ_{out} 分别表示复数域全息失真矫正网络的输出振幅和相位, A_{gt} 和 ϕ_{gt} 分别表示无失真全息图的振幅和相位, $\delta(\phi_{\text{out}}, \phi_{\text{gt}}) = \arctan\left(\frac{\sin(\phi_{\text{out}} - \phi_{\text{gt}})}{\cos(\phi_{\text{out}} - \phi_{\text{gt}})}\right)$ 用于相位偏移矫正, $\bar{\delta}(\phi_{\text{out}}, \phi_{\text{gt}})$ 表示 $\delta(\phi_{\text{out}}, \phi_{\text{gt}})$ 的均值项. 设置的损失函数不仅包含传统的振幅误差(如 MSE), 还引入相位一致性约束项, 要求矫正后的相位分布与参考相位的偏差控制在 $[-\pi, \pi]$ 内. 同时, 通过复数域的共轭对称性校验, 避免因特征优化导致的相位翻转或振幅异常.

全息重建损失. 为了实现重建目标平面上全息图的良好视觉感知效果, 本文使用角谱法模拟衍射过程,

重建全息图像,然后根据注意力掩膜计算 MSE 损失:

$$L_{\text{rec}} = \frac{1}{N} (O_{\text{gt}}(j) - O_{\text{out}}(j))^2 \quad (8)$$

其中, O_{gt} 和 O_{out} 分别代表真实无失真全息图像的重建图像振幅和复数域全息失真矫正网络输出的重建图像振幅, $N = H \times W$ 表示图像的大小.

全息失真矫正的主观质量损失. 为了保证矫正后的复数域全息图的主观视觉质量,使用多尺度结构相似性(MSSIM)损失和绝对值损失的组合来优化复数域数据的损失:

$$L_{\text{comp}} = \alpha \times L_{\text{MSSIM}} + (1 - \alpha) \times L_{l1} \quad (9)$$

其中,多尺度结构相似性(MSSIM)损失表示为

$$L_{\text{MSSIM}} = 1 - \text{MSSIM}(x_{\text{complex}}^{\text{out}} - x_{\text{complex}}^{\text{gt}}) \quad (10)$$

绝对值损失表示为

$$L_{l1} = \frac{1}{N} (x_{\text{complex}}^{\text{out}}(j) - x_{\text{complex}}^{\text{gt}}(j))^2 \quad (11)$$

其中, $x_{\text{complex}}^{\text{out}}$ 和 $x_{\text{complex}}^{\text{gt}}$ 分别表示复数平面上的真实无失真全息图像和经过复数域失真矫正网络的全息图像,权重因子 $\alpha = 0.84$ 用于平衡不同的损失.

最终通过最小化总损失来优化训练提出的全息失真矫正网络:

$$L = L_{\text{comp}} + L_{\text{holo}} + L_{\text{rec}} \quad (12)$$

4 实验结果与分析

4.1 数据集与实验配置

实验采用了两个公开的标准数据集,包括全息图像数据集(MIT Computer-Generated Holography dataset, MIT-CGH)^[31]和沉浸式视频数据集(MPEG immersive video dataset, MPEG)^[32]. MIT-CGH 共有 4 000 幅全息图像,每幅全息图像均由一对振幅图像和相位图像组成,其涵盖了丰富的场景内容和复杂的相位信息;MPEG 沉浸式视频数据集由 5 个超高清(分辨率 4 K 左右)的视频流组成,通过 Tensorholo^[31]将其转换为全息图像,其提供了多视角、高动态范围的视频序列,能够有效模拟真实场景下的高清全息成像环境. 上述全息图像数据集涵盖了室内和室外常见的对象类型,场景深度为 0.05~2 m,相位分布限定在 $[-\pi, \pi]$ 区间内;在采用 Python 工具实现全息图像角谱法模拟的过程中,重构角度无需人工额外设定,将由算法根据输入的全息图像核心参数(如像素分辨率、照明波长、物面与全息面间距等)自动推

导计算得出,确保模拟过程的客观性与结果的准确性. 实验使用 MIT-CGH 数据集中的 3 800 张作为训练集,100 张作为测试集,100 张作为验证集. 实验中,我们将全息图像分离为实部图像和虚部图像,再对其进行 JPEG 和 HEVC 压缩处理,并施加加性高斯白噪声(Additive White Gaussian Noise, AWGN)干扰. 具体的失真类型与失真强度如表 1 所示.

本文在 Intel (R) Xeon (R) Silver 4214 CPU 和 NVIDIA RTX2080 Ti GPU 上基于 PyTorch1.13.1 进行模型训练和性能测试. 实验中,批处理大小设置为 16,Adam 优化器的学习率的初始值为 1×10^{-4} ,并随着迭代进行衰减,在 100 轮数据迭代之后,将学习率降至 1×10^{-6} ,以便网络参数收敛.

4.2 全息失真矫正性能分析

本文采用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural Similarity index, SSIM)作为失真矫正方法的性能评价指标. 本文通过实验分别对比了含 JPEG 失真、含 HEVC 失真以及混合噪声失真的全息图像的矫正结果,具体包含了在不同观测距离(如近场、中场和远场)的全息重建图像的失真矫正结果,如表 1 所示. 所提出的复数多分支融合网络(Multi-Branch-Fused Complex-valued Convolutional Neural Network, MBF-CCNN)与先进的基于 Swin-Conv 的 SCUNet^[26]的方法以及使用全实数卷积的多分支融合网络(Multi-Branch-Fused Convolutional Neural Network, MBF-CNN)进行了对比. MBF-CNN 使用实数卷积分别单独对全息图像的实部和虚部数据进行失真矫正,而 SCUNet 在网络结构上也融入 Transformer 块,增强了局部和非局部建模能力.

表 2 和表 3 的实验结果表明,针对压缩与噪声混合失真场景,本文提出的 MBF-CCNN 的性能在所有方法中始终处于领先地位. 在 MIT-CGH 数据集上,MBF-CCNN 的 PSNR 指标和 SSIM 指标相比于 SCUNet 分别平均提升了 0.41 dB 和 0.007;在 MPEG 沉浸式视频数据集上,MBF-CCNN 的 PSNR 指标和 SSIM 指标相比于 SCUNet 分别平均提升了 0.46 dB 和 0.006. 这充分证明了 MBF-CCNN 的网络结构在处理多种失真联合干扰时的有效性. MBF-CCNN 的多分支特征融合结构通过并行处理不同尺度的特征图,分别捕捉图像的局部细节和全局结构,有效提升了网络对不同类型失真(如压缩伪

表 1 本文涉及的失真类型

失真类型	参数范围	失真程度描述	对全息图的影响
JPEG 压缩	QF=75,85	中等至较强压缩,明显块效应	高频条纹丢失、边缘模糊、相位畸变
HEVC 压缩	QP=26,28	中等压缩,更复杂的失真模式	预测误差导致的局部失真、结构破坏
AWGN 噪声	$\sigma=15,35$	中等到强噪声干扰	图像颗粒感增强、细节淹没、颜色偏移
混合失真	上述任意组合	复杂混合失真,难以单独建模	振幅与相位同时受损,重建图像质量显著下降

影和噪声干扰)的适应性. 这使得 MBF-CCNN 在面对复杂的混合失真场景时,能够协同各分支的优势,针对性地对不同失真成分进行矫正. 实验结果显示,在高噪声强度场景下,MBF-CCNN 的 PSNR 和 SSIM 相比于 SCU-

Net 提升振幅更大. 这表明复数注意力机制和多分支网络融合架构能够有效增强网络对压缩、传输干扰等多种失真的建模能力,更精准地恢复图像的高频细节和相位信息,从而达到更高质量的失真矫正效果.

表 2 MIT-CGH 数据集上失真矫正的平均峰值信噪比和平均结构相似性结果

单位: dB

(a) 各方法在 MIT-CGH 数据集混合 JPEG 失真与高斯白噪声的峰值信噪比结果									
σ/QF	15/无	35/无	无/75	无/85	15/75	15/85	35/75	35/85	平均提升
AWGN	24.83	22.24	—	—	21.31	22.85	19.31	20.97	—
JPEG	—	—	22.12	23.74					
MBF-CNN	27.31	25.16	25.09	27.17	24.41	25.57	23.15	24.73	3.15
SCUNet	30.04	28.83	28.24	29.54	27.01	28.24	26.21	27.29	6.00
MBF-CCNN	30.42	29.05	28.71	30.03	27.39	28.63	26.54	27.61	6.38

(b) 各方法在 MIT-CGH 数据集混合 HEVC 失真与高斯白噪声中峰值信噪比结果									
σ/QF	15/无	35/无	无/26	无/28	15/26	15/28	35/26	35/28	平均提升
AWGN	24.83	22.24	—	—	22.14	21.53	21.47	20.38	—
HEVC	—	—	26.24	24.39					
MBF-CNN	27.31	25.16	29.47	27.40	25.63	27.93	26.59	25.78	4.00
SCUNet	30.04	28.83	31.72	29.48	27.76	29.79	28.93	28.56	6.48
MBF-CCNN	30.42	29.05	32.21	30.29	28.11	30.09	29.56	28.80	6.92

(c) 各方法在 MIT-CGH 数据集混合 JPEG 失真与高斯白噪声的结构相似性结果									
σ/QF	15/无	35/无	无/75	无/85	15/75	15/85	35/75	35/85	平均提升
AWGN	0.821 3	0.751 9	—	—	0.765 1	0.783 7	0.691 0	0.713 9	—
JPEG	—	—	0.793 2	0.813 7					
MBF-CNN	0.852 7	0.781 5	0.811 3	0.832 5	0.782 5	0.794 1	0.711 4	0.721 8	0.019 2
SCUNet	0.871 3	0.799 4	0.828 1	0.845 3	0.787 9	0.803 3	0.716 7	0.726 5	0.030 6
MBF-CCNN	0.879 1	0.809 6	0.837 7	0.852 1	0.791 8	0.809 4	0.720 1	0.732 6	0.037 3

(d) 各方法在 MIT-CGH 数据集混合 HEVC 失真与高斯白噪声的结构相似性结果									
σ/QP	15/无	35/无	无/26	无/28	15/26	15/28	35/26	35/28	平均提升
AWGN	0.821 3	0.751 9	—	—	0.790 8	0.778 1	0.753 1	0.732 3	—
HEVC	—	—	0.867 3	0.847 1					
MBF-CNN	0.852 7	0.781 5	0.882 9	0.861 5	0.810 2	0.796 7	0.763 3	0.746 5	0.019 2
SCUNet	0.871 3	0.799 4	0.899 3	0.885 7	0.828 5	0.810 4	0.766 9	0.748 9	0.033 5
MBF-CCNN	0.879 1	0.809 6	0.903 1	0.891 9	0.835 7	0.819 0	0.770 4	0.751 8	0.039 8

注: σ 、QF 和 QP 分别表示高斯白噪声的强度(越大噪声越强)、JPEG 的压缩参数值(越小压缩失真越大)和 HEVC 的压缩参数值(越大压缩失真越大).

传统实数域网络需要将复数全息图拆解为相位、振幅两个实数通道分别处理,在特征优化时,易因通道间信息交互不足出现顾此失彼的情况,如过度压缩振幅通道而忽略相位通道的关键信息,进而导致信息割裂. 而本文提出的矫正网络在复数域内实现端到端处理,天然保留两者的物理关联性. 在与 MBF-CNN 的对比实验中,就证明了这一点. 当处理含复杂失真的全息图像时,MBF-CCNN 凭借复数注意力机制,能够更准确地识别并恢复被各种失真破坏的图像结构和纹理细节. 具体而言,在 MIT-CGH 数据集上,MBF-CCNN 的 PSNR 相比于 MBF-CNN 平均提升了 3.08 dB, SSIM 平均

提升了 0.02; 在 MPEG 沉浸式视频数据集上, MBF-CCNN 的 PSNR 相比于 MBF-CNN 平均提升了 3.46 dB, SSIM 平均提升了 0.014. 这一结果表明, MBF-CCNN 相较于 MBF-CNN 展现出了显著优势. 这得益于复数注意力机制. 复数注意力机制基于复数卷积的特性,能够自适应地分配不同区域的权重,精准聚焦于受失真影响严重的图像区域,而复数卷积通过引入实部和虚部交互运算,能够捕捉信号的相位信息,这对于全息图像这种依赖相位信息记录物体特征的数据类型尤为关键.

另一方面,当前矫正方法是数据驱动的深度学习算法,其泛化性在一定程度上依赖于数据集的丰富程

表3 MPEG数据集上失真矫正的平均峰值信噪比和结构相似性结果

单位: dB

(a) 各方法在MPEG数据集混合JPEG失真与高斯白噪音的峰值信噪比结果									
σ/QF	15/无	35/无	无/75	无/85	15/75	15/85	35/75	35/85	平均提升
AWGN	29.31	26.44	—	—	21.31	23.24	19.31	20.97	—
JPEG	—	—	24.83	26.79					
MBF-CNN	34.27	30.54	28.40	31.36	25.29	26.98	23.15	24.73	4.07
SCUNet	34.85	31.73	32.02	33.86	28.87	30.91	26.94	28.19	6.90
MBF-CCNN	35.14	32.35	32.69	34.51	29.40	31.52	27.39	28.68	7.44
(b) 各方法在MPEG数据集混合HEVC失真与高斯白噪音中峰值信噪比结果									
σ/QP	15/无	35/无	无/26	无/28	15/26	15/28	35/26	35/28	平均提升
AWGN	29.31	26.44	—	—	27.23	26.49	24.21	23.39	—
HEVC	—	—	29.14	28.26					
MBF-CNN	34.27	30.54	31.59	30.38	28.88	27.93	26.59	25.78	2.69
SCUNet	34.85	31.73	34.83	33.89	33.75	32.79	31.32	30.64	6.17
MBF-CCNN	35.14	32.35	35.06	34.19	34.01	33.26	31.76	30.97	6.54
(c) 各方法在MPEG数据集混合JPEG失真与高斯白噪音的结构相似性结果									
σ/QF	15/无	35/无	无/75	无/85	15/75	15/85	35/75	35/85	平均提升
AWGN	0.852 4	0.788 3	—	—	0.784 7	0.805 7	0.725 3	0.736 8	—
JPEG	—	—	0.814 2	0.832 9					
MBF-CNN	0.900 6	0.835 2	0.864 1	0.880 2	0.801 3	0.812 7	0.736 1	0.744 8	0.029 3
SCUNet	0.907 4	0.838 1	0.886 8	0.894 3	0.810 1	0.822 5	0.735 1	0.749 7	0.037 9
MBF-CCNN	0.911 9	0.843 7	0.897 1	0.901 6	0.815 1	0.829 1	0.743 3	0.755 2	0.044 6
(d) 各方法在MPEG数据集混合HEVC失真与高斯白噪音的结构相似性结果									
σ/QP	15/无	35/无	无/26	无/28	15/26	15/28	35/26	35/28	平均提升
AWGN	0.852 4	0.788 3	—	—	0.823 8	0.818 1	0.763 8	0.748 3	—
HEVC	—	—	0.867 3	0.847 1					
MBF-CNN	0.900 6	0.835 2	0.905 3	0.881 4	0.840 2	0.834 7	0.774 3	0.757 3	0.027 5
SCUNet	0.907 4	0.838 1	0.911 4	0.893 6	0.847 6	0.842 7	0.778 2	0.761 0	0.033 8
MBF-CCNN	0.911 9	0.843 7	0.914 3	0.905 9	0.855 7	0.849 0	0.781 2	0.764 6	0.039 6

注: σ 、QF和QP分别表示高斯白噪音的强度(越大噪音越强)、JPEG的压缩参数值(越小压缩质量越差)和HEVC的压缩参数值(越大压缩质量越差)。

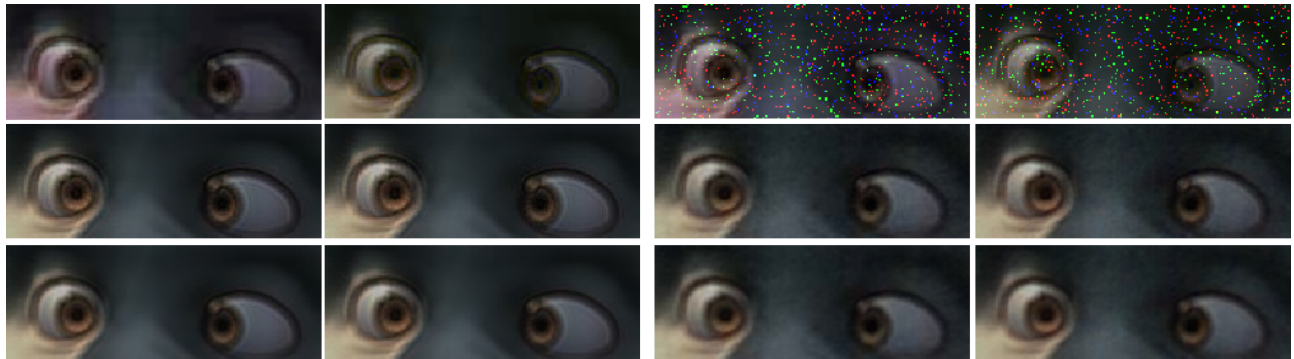
度. 由于训练数据集包含了室内外常见的全息对象类型, 表2和表3中两个不同数据集的结果表明, 当前算法可以很好地适配不同类型的全息场景, 在重构深度0.05~2 m, 相位分布 $[-\pi, \pi]$ 范围内可以取得很好的失真矫正性能.

图5给出了不同方法对Big Buck Bunny全息图像失真矫正后的主观质量对比结果. HEVC凭借先进的帧内预测与变换编码技术, 在保持相同视觉质量的前提下, 较传统JPEG算法实现了更高的压缩比. 然而, 尽管HEVC在编码效率上具有显著优势, 其重构的全息图像仍存在局部可见的块效应与边缘模糊现象. 通过图5(d)和图5(e)中的相位图像可以看到分别压缩全息图像实部和虚部会严重破坏其相关性, 从而导致相位图像严

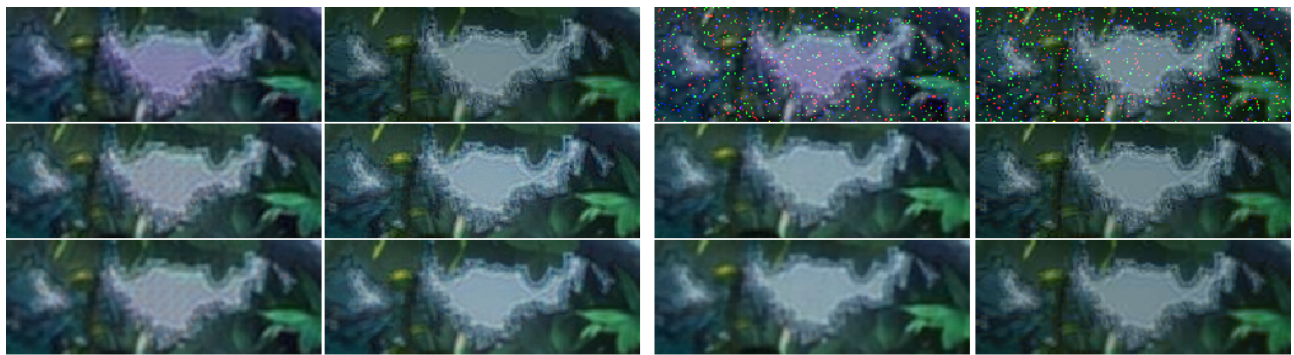
重失真, 进而影响全息图像的重构质量. SCUNet通过其独特的Swin-Conv块能够较好地捕捉全息图像失真, 为全息图像的实部和虚部分别进行恢复, 但是由于无法捕获全息图像实部和虚部之间的相关性, 其对最终观看的振幅图像的失真矫正效果不佳. 而本文提出的MBF-CCNN网络基于多分支特征融合架构, 通过设计专门的复数卷积模块, 能够精准捕获全息图像实部与虚部之间的耦合关系. 该网络不仅有效抑制了振幅失真导致的亮度异常, 还通过相位补偿机制, 显著改善了相位畸变问题, 在全息图像质量恢复方面展现出了很好的性能优势. 从图5可以看出, MBF-CCNN在主观质量评价方面依然领先于SCUNet, 这一点也恰好验证了与表2的客观评价结果的一致性.



(a) 整帧与局部原始全息振幅和相位图像



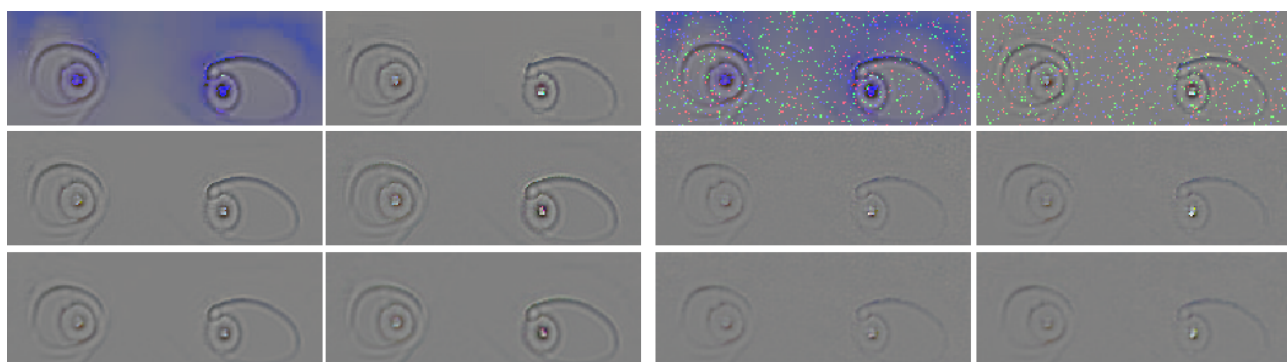
(b) 振幅失真矫正结果眼睛局部图



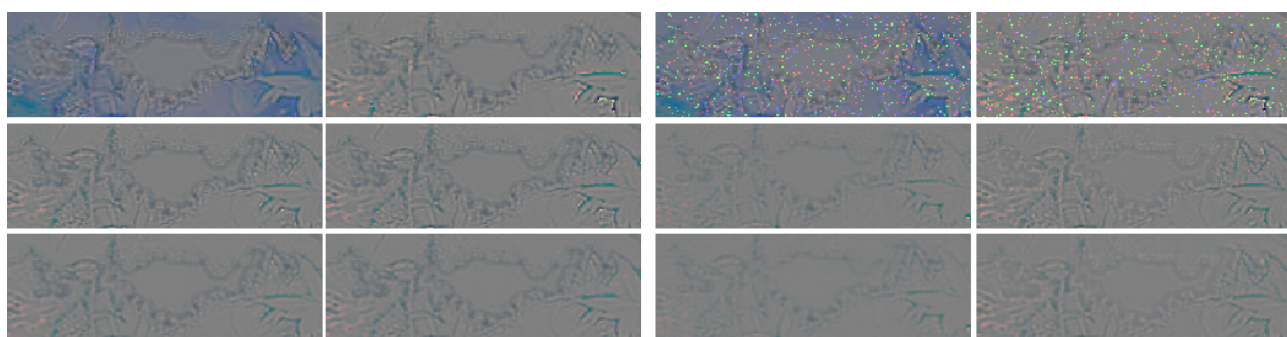
(c) 振幅失真矫正结果树叶局部图

注:在分图(b)~图(e)中,第一行中从左至右的全息图像分别对应为含 JPEG 失真、HEVC 失真、JPEG 与 AWGN 混合失真和 HEVC 与 AWGN 混合失真,从上至下的 3 行分别表示失真全息图像、经过 MBF-CCNN 网络矫正的全息图像以及经过 SCUNet 网络矫正的全息图像。

图 5 Big Buck Bunny 全息图像的失真矫正主观质量对比



(d) 相位失真矫正结果眼睛局部图



(e) 相位失真矫正结果树叶局部图

续图5

4.3 消融实验

本文还针对 MBF-CCNN 中的多分支结构和复数注意力机制进行了消融实验,实验结果如表 4 所示. 在针对多分支结构的消融实验中,去除子网 1 后,PSNR 的均值结果从 30.79 dB 降至 26.82 dB,SSIM 的均值结果从 0.827 2 降至 0.811 4;去除子网 2 后,PSNR 的均值结果降至 27.29 dB,SSIM 的均值结果降至 0.813 9;去除子网 3 后,PSNR 的均值结果降至 26.73 dB,SSIM 的均值结果降至 0.809 3. 这说明了各子网均对模型性能有正向贡献,子网的存在能提升图像质量还原与结构保持能力. 在去除子网 2 时,PSNR 和 SSIM 下降振幅相对较小(PSNR 下降 3.5 dB,SSIM 下降 0.013 3),这是由于子网 2 主要用于相位信息的提取,相位信息不是对最后观看的全息图像造成影响的主要因素. 但在去除子网 1 或子网 3 时,子网 1 和子网 3 分别对全息图像的振幅信息和复数域信息进行优化,这些部分对最后观看的全息图像能造成更关键的影响,因此指标降幅更显著. 而在针对复数注意力机制的消融实验中,无复数注意力机制的 MBF-CCNN 的 PSNR 和 SSIM 的均值结果分别为 28.45 dB 和 0.815 2,与带复数注意力机制的 MBF-CCNN 相比,PSNR 和 SSIM 分别下降了 2.34 dB 和 0.012. 这表明复数注意力机制相比于实数注意力机制更能捕捉到全息图像复数域数据的失真部分,从而更好地进行失真矫正.

表 4 消融实验结果

失真矫正网络	PSNR/dB	SSIM
MBF-CCNN(无子网 1)	26.82	0.811 4
MBF-CCNN(无子网 2)	27.29	0.813 9
MBF-CCNN(无子网 3)	26.73	0.809 3
MBF-CCNN(无复数注意力机制)	28.45	0.815 2
MBF-CCNN	30.79	0.827 2

5 结论

全息图像的压缩与通信引发的失真导致其重建质量变差. 针对这一问题,本文提出一种新颖的基于多分支混合注意力网络的全息图像失真矫正方法. 该方法系统探究了压缩与传输失真在全息图像重建过程中的传递机制与影响规律,紧接着利用多分支结构并行提取全息图像中不同尺度的特征信息,结合复数注意力机制,捕捉图像中块与块之间的关系、实部图像与虚部图像之间的关系,从而实现对压缩失真的有效抑制. 实验结果表明,被 JPEG 或 HEVC 压缩以及传输 AWGN 噪声干扰后的全息失真图像,经过本文所提出的 MBF-CCNN 网络进行失真矫正后,在峰值信噪比指标上比矫正前分别平均提升 6.38 dB 以上,结构相似性指标分别提升 0.037 以上;而相较于现有先进的深度学习失真矫正方法 SCUNet,本文方法在峰值信噪比指标上提升 0.41 dB 以上,结构相似性指标提升 0.006 以上,显著改

善了全息图像的重建质量,有效减少了失真带来的细节模糊与纹理丢失问题. 由于当前矫正网络需要与全息重建算法相结合,但不依赖于全息重建算法,在重建效率方面仍有不少空间可以提升. 在未来工作中,将考虑对接新的快速重建算法^[33,34],用动态阈值筛选关键特征,优化神经网络效率,在保证全息质量的同时更好地适配实时场景.

参考文献

- [1] LIU Z, MALEKIAN R, LI R, et al. Guest editorial: Holographic video communications[J]. *IEEE Wireless Communications*, 2025, 32(2): 94-95.
- [2] BLINDER D, AHAR A, BETTENS S, et al. Signal processing challenges for digital holographic video display systems[J]. *Signal Processing: Image Communication*, 2019, 70: 114-130.
- [3] CUERVO E. BEYOND REALITY: Head-mounted displays for mobile systems researchers[J]. *GetMobile: Mobile Computing and Communications*, 2017, 21(2): 9-15.
- [4] 刘娟, 皮大普, 王涌天. 实时全息三维显示技术研究进展[J]. *光学学报*, 2023, 43(15): 128-141.
LIU J, PI D P, WANG Y T. Research progress of real-time holographic 3D display technology[J]. *Acta Optica Sinica*, 2023, 43(15): 128-141. (in Chinese)
- [5] LIU N H, LIU K X, YANG Y X, et al. Propagation-adaptive 4K computer-generated holography using physics-constrained spatial and Fourier neural operator[J]. *Nature Communications*, 2025, 16: 7761.
- [6] 陈佳, 刘上, 郜帅, 等. 面向全息通信的智算融合网络[J]. *电子学报*, 2025, 53(3): 754-764.
CHEN J, LIU S, GAO S, et al. Smart computing integration network for holographic communication[J]. *Acta Electronica Sinica*, 2025, 53(3): 754-764. (in Chinese)
- [7] 朱原玮, 黄亚坤, 乔秀全. 面向全息视频通信的自适应分块传输方法[J]. *电子学报*, 2024, 52(4): 1144-1154.
ZHU Y W, HUANG Y K, QIAO X Q. Towards holographic video communications: An adaptive tiling solution[J]. *Acta Electronica Sinica*, 2024, 52(4): 1144-1154. (in Chinese)
- [8] JIAO S M, JIN Z, CHANG C L, et al. Compression of phase-only holograms with JPEG standard and deep learning[J]. *Applied Sciences*, 2018, 8(8): 1258.
- [9] SHIMOBABA T, BLINDER D, MAKOWSKI M, et al. Dynamic-range compression scheme for digital hologram using a deep neural network[J]. *Optics Letters*, 2019, 44(12): 3038-3041.
- [10] HU C C, YANG G L, XIE H Y. 3D information transmission of a computer-generated hologram using a quantum compensation hybrid neural network[J]. *Optics Express*, 2024, 32(13): 23736-23749.
- [11] JEE M, KIM H, YOON M, et al. Hologram super-resolution using dual-generator GAN[C]//2022 IEEE International Conference on Image Processing. Piscataway: IEEE, 2022: 2596-2600.
- [12] BOURTSOULATZE E, BURTH KURKA D, GÜNDÜZ D. Deep joint source-channel coding for wireless image transmission[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2019, 5(3): 567-579.
- [13] LEFKIMMIATIS S. Universal denoising networks: A novel CNN architecture for image denoising[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3204-3213.
- [14] 白勇强, 禹晶, 李一稔, 等. 基于深度先验的盲图像去模糊算法[J]. *电子学报*, 2023, 51(4): 1050-1067.
BAI Y Q, YU J, LI Y N, et al. Deep prior-based blind image deblurring[J]. *Acta Electronica Sinica*, 2023, 51(4): 1050-1067. (in Chinese)
- [15] YANG X H, XU Y, QUAN Y H, et al. Image denoising via sequential ensemble learning[J]. *IEEE Transactions on Image Processing*, 2020, 29: 5038-5049.
- [16] HONG Z W, FAN X C, JIANG T, et al. End-to-end unpaired image denoising with conditional adversarial networks[C]//2020 Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI, 2020: 4140-4149.
- [17] ZHANG K, ZUO W M, CHEN Y J, et al. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising[J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155.
- [18] TAI Y, YANG J, LIU X M, et al. MemNet: A persistent memory network for image restoration[C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 4549-4557.
- [19] QUAN Y H, CHEN Y X, SHAO Y Z, et al. Image denoising using complex-valued deep CNN[J]. *Pattern Recognition*, 2021, 111: 107639.
- [20] TIAN C W, XU Y, ZUO W M. Image denoising using deep CNN with batch renormalization[J]. *Neural Networks*, 2020, 121: 461-473.
- [21] LIU S Q, LEI Y, ZHANG L Y, et al. MRDDANet: A multiscale residual dense dual attention network for SAR image denoising[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5214213.
- [22] TIAN C W, XU Y, ZUO W M, et al. Designing and train-

- ing of a dual CNN for image denoising[J]. Knowl. Based Syst., 2020, 226: 106949.
- [23] LEFKIMMIATIS S. Non-local color image denoising with convolutional neural networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 5882-5891.
- [24] CHEN J W, CHEN J W, CHAO H Y, et al. Image blind denoising with generative adversarial network based noise modeling[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3155-3164.
- [25] GUO S, YAN Z F, ZHANG K, et al. Toward convolutional blind denoising of real photographs[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 1712-1722.
- [26] ZHANG K, LI Y W, LIANG J Y, et al. Practical blind image denoising via swin-conv-UNet and data synthesis[J]. Machine Intelligence Research, 2023, 20(6): 822-836.
- [27] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 9992-10002.
- [28] SZPYGIEL J, CHLIPALA M, KUKOLOWICZ R, et al. Distortion correction for wide angle holographic projector[J]. Photonics Letters of Poland, 2021, 13(4): 79-81.
- [29] HE Z H, SUI X M, JIN G F, et al. Distortion-correction method based on angular spectrum algorithm for holographic display[J]. IEEE Transactions on Industrial Informatics, 2019, 15(11): 6162-6169.
- [30] LI J H, ZHANG H L, JI Q L, et al. Distortion-corrected integral imaging 3D display system based on lens array holographic optical element[J]. Symmetry, 2022, 14(7): 1481.
- [31] SHI L, LI B C, KIM C, et al. Towards real-time photorealistic 3D holography with deep neural networks[J]. Nature, 2021, 591(7849): 234-239.
- [32] BOYCE J M, DORÉ R, DZIEMBOWSKI A, et al. MPEG immersive video coding standard[J]. Proceedings of the IEEE, 2021, 109(9): 1521-1536.
- [33] LIU K X, WU J C, HE Z H, et al. 4K-DMDNet: Diffraction model-driven network for 4K computer-generated holography[J]. Opto-Electronic Advances, 2023, 6(5): 220135.
- [34] LIU K X, WU J C, CAO L C. High-quality and high-speed computer-generated holography via deep-learning-assisted bidirectional error diffusion method[J]. Optics Express, 2024, 32(21): 37342-37354.

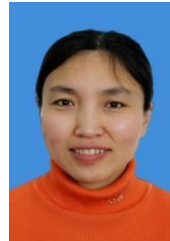
作者简介



陈逸飞 男, 2000年出生于福建省泉州市。现为华为技术有限公司工程师。主要研究方向为计算机生成全息图像的理论及应用。
E-mail: cyf_361@163.com



刘延伟 男, 1976年出生于黑龙江省克山县。现为中国科学院信息工程研究所副研究员。主要研究方向为全息通信、智能多媒体信息处理、人工智能安全与数据安全。
E-mail: liuyanwei@iie.ac.cn



刘金霞 女, 1969年出生于黑龙江省齐齐哈尔市。现为浙江万里学院教授。主要研究方向为沉浸式通信与多媒体通信。
E-mail: liujinxia1969@126.com



古晓艳 女, 1987年出生于山东省济宁市。现为中国科学院信息工程研究所正高级工程师。主要研究方向为数据安全、人工智能安全、深度学习。
E-mail: guxiaoyan@iie.ac.cn