

颜色感知的双通道反射消除算法^①

张凯威, 张 玲

(武汉科技大学 计算机科学与技术学院, 武汉 430081)

通信作者: 张 玲, E-mail: zhling@wust.edu.cn



摘 要: 为了解决现有反射消除算法在处理复杂场景时导致的图像质量下降问题, 本文提出了一种具有颜色感知的双通道反射消除算法. 首先, 设计了背景颜色生成器, 以准确预测图像的背景颜色信息, 为基础反射消除网络提供背景支持, 生成初步的反射消除结果. 接着, 提出了双通道反射消除网络, 对初步结果进行进一步的优化. 此外, 该算法在双通道反射消除网络中设计了稀疏 Transformer 模块、通道注意力模块以及特征融合模块, 从而提升了反射消除的精度和效果. 实验结果表明, 该方法在 RRID 和 Flash 数据集上表现出色, 能够有效去除反射光并显著提升图像的真实感.

关键词: 图像反射消除; 颜色感知; 双通道网络; 特征融合; 稀疏注意力

引用格式: 张凯威, 张玲. 颜色感知的双通道反射消除算法. 计算机系统应用, 2025, 34(3): 277-285. <http://www.c-s-a.org.cn/1003-3254/9793.html>

Color-aware Dual-channel Reflection Removal Algorithm

ZHANG Kai-Wei, ZHANG Ling

(School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430081, China)

Abstract: To address the issue of image quality decline caused by existing reflection removal algorithms when handling complex scenes, this study proposes a color-aware dual-channel reflection removal algorithm. First, a background color generator is designed to accurately predict the background color information of an image, provide background support for the basic reflection removal network, and generate preliminary reflection removal results. Subsequently, a dual-channel reflection removal network is proposed to further optimize these preliminary results. Additionally, the algorithm designs a sparse Transformer module, a channel attention module, and a feature fusion module within the dual-channel reflection removal network, thereby enhancing the precision and effect of reflection removal. Experimental results demonstrate that this method performs excellently on the RRID and Flash datasets, effectively removing reflected light and significantly enhancing image realism.

Key words: image reflection removal; color awareness; dual-channel network; feature fusion; sparse attention

在通过玻璃窗或镜面拍摄物体时, 反射光往往会掩盖图像中的关键细节, 并引入额外的噪声和伪影, 严重影响图像的可用性和分析精度, 从而降低图像分类^[1,2]、图像分割^[3,4]等视觉任务的准确性和有效性. 因此, 图像的反射消除已成为计算机视觉领域中的一个

重要问题.

通常, 反射图像可以通过以下模型表示^[5-7].

$$I = T + K \otimes R \quad (1)$$

其中, I 是包含反射的图像, T 是无反射的背景图像, R 是反射层图像, K 是用于建模反射效应的低通滤波

① 基金项目: 湖北省自然科学基金 (2023AFB615)

收稿时间: 2024-08-28; 修改时间: 2024-09-30; 采用时间: 2024-10-14; csa 在线出版时间: 2025-01-16

CNKI 网络首发时间: 2025-01-17

器, \otimes 表示离散卷积. 单图像反射消除的目标是从反射图像 I 中分离出反射层图像 R , 得到背景图像 T . 由于背景图像 T 和反射层图像 R 之间存在无限种可能的组合, 这使得反射消除成为一个高度不适定的问题.

传统的图像反射消除算法通常依赖额外的先验信息或辅助信息来处理特定类型的场景, 从而去除反射层^[8-13]. 然而, 这些先验信息仅适用于符合特定假设的反射图像, 在应对复杂的现实世界反射场景时效果不佳.

当前基于深度学习的图像反射消除算法, 通常依赖于 CNN 和 Transformer 等技术来自动识别和提取与反射相关的图像特征, 从而提高去除反射的效率^[14-22]. 然而, 由于其依赖单一的 CNN 模型或 Transformer 模型, 这些方法在处理复杂场景时仍然面临挑战. 尽管 CNN 模型在捕捉局部特征方面表现优异, 但在处理全局信息时存在不足, 容易引发细节丢失和伪影. 而 Transformer 模型, 虽然通过自注意力机制能有效捕捉全局信息, 但由于计算复杂度较高, 不仅降低了处理速度, 还限制了局部特征的提取效果. 此外, 现有方法往往忽略了颜色背景信息的关键作用, 导致反射消除后的图像缺乏细节和自然感. 这些问题在复杂场景下尤其明显, 直接影响了去反射后的图像质量和视觉信息的准确获取.

针对现有反射消除技术在处理复杂场景时的不足, 本文提出了一种具有颜色感知的双通道反射消除网络. 该网络由两个主要部分组成. 首先, 我们设计了一个背景颜色生成器, 用于预测图像的背景颜色信息, 并将其融入到基础反射消除网络中, 生成初步的反射消除结果. 接着, 提出了双通道反射消除网络, 对初步结果进一步优化, 以生成高质量的反射消除图像. 本文创新地将颜色背景生成器在不同阶段与解码器结合, 充分利用预测的真实颜色背景信息, 从而在反射消除过程中提供更为精确的背景支持. 此外, 在双通道反射消除网络中, 本文设计了稀疏 Transformer 模块、通道注意力模块和特征融合模块, 能够有效消除反射层, 提升在复杂场景中反射消除的能力. 本文的主要贡献如下.

(1) 设计了颜色感知的双通道反射消除网络, 结合颜色背景生成器和双通道反射消除网络, 利用背景颜色信息提高了反射消除的质量, 尤其是在复杂场景中能够更好地保留图像细节.

(2) 设计了稀疏 Transformer 模块, 该模块通过稀疏注意力机制精准聚焦关键区域, 同时有效捕捉长距离依赖关系, 提升了图像的恢复质量. 它能够减少无关

特征的噪声干扰, 适应不同场景下的特征变化, 进一步增强模型的鲁棒性.

(3) 提出了特征融合模块, 结合稀疏 Transformer 模块和通道注意力模块, 通过特征融合机制实现了局部特征与全局信息的高效结合, 从而提高了反射消除的整体效果.

1 相关工作

1.1 传统的图像反射消除算法

传统的反射消除算法通常依赖一些额外的先验信息或辅助信息来处理特定类型的场景, 从而去除反射层. 在某些特定场景中, 图像中的物体大致位于背景和反射层的同一焦平面上, 一些方法使用梯度稀疏先验来分解背景和反射层, 使梯度和局部特征(如边缘和角落)最小化^[8,9]. 例如, Li 等人^[10]假设反射层比背景层更模糊, 并将其建模为两个不同的梯度分布在两个图像层中以实现分离. Levin 等人^[11]使用稀疏先验来分解输入图像, 但这种方法需要用户手动标记背景和反射边缘, 这不仅需要大量的人工工作, 而且在纹理区域可能会失败. 受图像平滑算法的启发, Arvanitopoulos 等人^[12]提出了一个简单的优化公式, 其在背景图像上使用 l_0 范数对梯度进行惩罚. Wan 等人^[13]提出一种多尺度场深图来指导边缘分类, 并使用 Levin 等人^[11]中的方法进行后续的图像重建. 这些方法虽然在特定条件下能够一定程度上去除反射层, 但在不同成像条件和复杂场景下, 仍然难以有效消除反射层.

1.2 基于深度学习的图像反射消除算法

与传统的反射消除方法不同, 基于深度学习的方法通过数据驱动的方式实现, 能够更有效地分离反射和非反射部分, 生成更高质量的背景图像. 近年来, 深度学习在去反射任务中受到越来越多的关注. 其主要思路是构建特定的神经网络结构, 通过设计特征提取器来提取必要的特征, 并基于这些特征引导反射消除. Zhang 等人^[14]设计了一种结合感知损失、对抗损失和排除损失的神经网络模型, 并提出了一种线性方法合成用于训练的图像数据集, 同时使用边缘图作为辅助信息来指导反射消除. Yang 等人^[16]提出了一个由三元数据集驱动的网络结构, 先预测背景层, 再利用背景层辅助反射层的预测, 以双向方式观察和优化反射消除效果. Li 等人^[19]设计了一个级联网络, 使用反射图像、背景层和反射层作为输入进行训练, 通过迭代方式持

续优化背景层和反射层的分离. 关于 Yang 等人^[16]和 Li 等人^[19]提出的方法, 在每个阶段计算的误差都会累积, 容易导致伪影. Li 等人^[20]提出了一种新的两阶段反射消除网络 RAGNet, 通过 RAG 模块在解码器上计算背景图像和反射层特征之间的差异, 以抑制反射并增强解码过程中的特征表达能力. Hu 等人^[21]提出了一种双流交互机制, 分别使用不同的分支流来预测反射层图像和背景图像, 并在两个分子之间进行交互. Song 等人^[22]提出了一种集成跨尺度注意机制、多尺度融合和对抗性图像鉴别器的网络框架, 该模型在消除反射的同时表现出一定的稳定性. 然而, 这些方法在平衡全局与局部特征时存在不足, 且未能充分利用图像中的颜色信息, 导致关键细节的丢失和生成不自然的伪影.

2 本文方法

本文提出了一种具有颜色感知的双通道反射消除算法来消除图像中的反射. 首先, 使用背景颜色生成器预测图像的背景颜色信息, 为后续处理提供辅助信息. 在此基础上, 将背景颜色信息引入到基础反射消除网络中, 在 3 个不同的阶段融合背景颜色, 以生成初步的去反射结果. 随后, 利用设计出的双通道反射消除网络进一步优化初步的反射结果.

2.1 背景颜色生成器

颜色信息在图像处理中具有重要意义^[23-25], 尤其是在图像反射消除任务中. 反射层和背景层的颜色特征往往有明显的差异, 利用颜色信息可以有效地分离

和去除反射层. 传统的图像消除方法通常忽略了背景颜色信息的作用, 导致消除反射后的图像出现颜色失真和细节丢失. 在实际应用中, 背景颜色信息的充分利用可以进一步提高反射消除算法的性能.

为了充分利用颜色信息, 本文设计了一个背景颜色生成器, 旨在从输入图像中准确提取背景颜色信息, 进而辅助基础反射消除网络进行初步的反射消除. 该网络采用经典的 U-Net 结构, 结合多层次卷积神经网络 (CNN) 以提取和融合颜色特征. 具体设计包括 4 层下采样卷积, 6 个集成了通道注意力的残差块和 4 层上采样卷积. 在下采样阶段的瓶颈层集成了 6 个通道注意力的残差块, 这些残差块由通道注意力模块和标准残差块组成, 利用通道注意力模块增强对重要颜色特征的关注, 并通过残差连接缓解梯度消失的问题. 跳跃连接在编码器和解码器之间传递高分辨率颜色特征, 以保证颜色信息的充分保留和融合. 通过这一设计, 网络能够有效提取图像中的颜色特征, 最终生成高分辨率且细节丰富的背景颜色图, 用于辅助后续的反射消除任务.

2.2 基础反射消除网络

基础反射消除网络的结构如图 1 所示, 其结构与背景颜色生成器相似, 包含 4 层下采样卷积、6 个通道注意力的残差块以及 4 层上采样卷积. 在网络的输入阶段和解码器的第 3 层、第 4 层之前, 将相应大小的背景颜色信息融入到网络中. 该结构通过有效利用颜色特征, 增强了反射消除的效果.

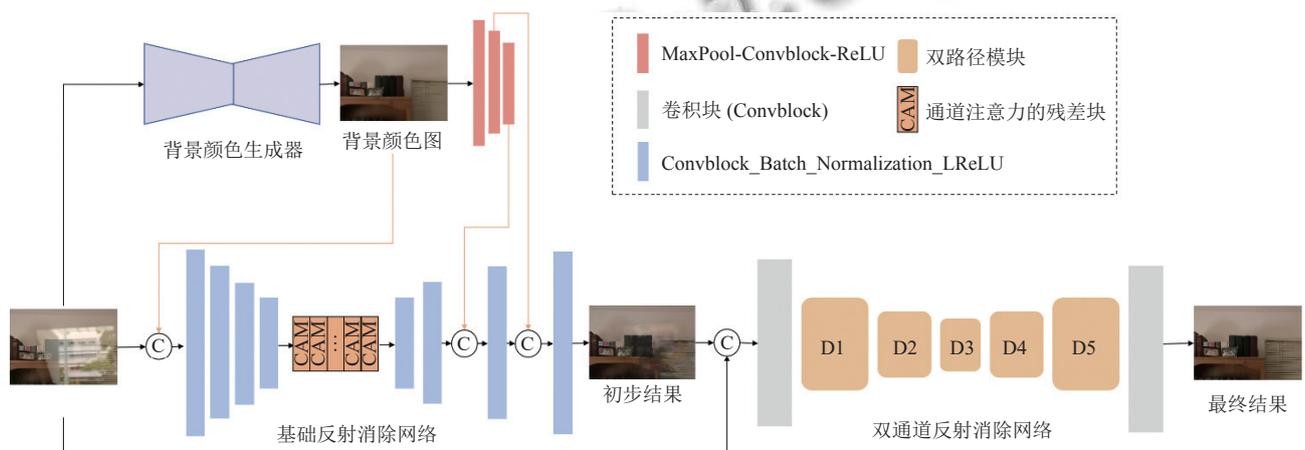


图 1 网络结构图

2.3 双通道反射消除网络

传统的反射消除网络主要依赖单一的 CNN 模型

或 Transformer 模型, 这存在一定的局限性. 基于 CNN 的方法擅长处理局部特征, 但难以捕捉全局信息, 可能

导致细节丢失和产生伪影. 基于 Transformer 的方法虽能捕捉长距离依赖, 但计算复杂度较高且可能忽视局部特征. 为克服这些不足, 本文设计了双通道反射消除网络, 将 CNN 和 Transformer 结合, 充分发挥 CNN 的局部特征提取能力和 Transformer 的全局信息处理优势, 从而实现了更准确的反射消除和更清晰的图像复原.

如图 1 所示, 双通道反射消除网络的整体结构包括卷积头、5 个双路径模块和卷积尾. 卷积头负责从输入图像中提取初步特征, 生成特征图. 随后, 这些特征图被送入 5 个双路径模块进行进一步的处理和优化. 最终, 通过卷积尾的整合处理, 生成去反射的图像. 在每个双路径模块中, 输入特征图被并行地传递给稀疏 Transformer 模块和通道注意力模块. 稀疏 Transformer 模块负责捕获特征中的长距离依赖信息, 以提取全局特征, 而通道注意力模块则擅长捕获局部特征的细节信息. 随后, 这两个模块的输出与原始特征图一起输入到特征融合模块. 特征融合模块通过结合稀疏 Transformer 模块和通道注意力模块的输出, 实现了局部特征与全局信息的有效融合, 从而生成更加精确的输出特征. 每个组件的设计将在后续部分详细描述.

2.3.1 稀疏 Transformer 模块

大多数 Transformer 模块中的注意力机制^[26,27]在计算自注意力时通常会处理全局信息, 但这种全局计算方式在图像恢复任务中存在一定的局限性, 因为它可能导致无关特征之间的噪声交互, 从而影响恢复效果. 为了解决这一问题, 本文提出了稀疏 Transformer 模块 (SATM). 该模块通过设定 top- k 算法中的 k 值来选择保留元素的比例, 动态选择注意力的关注区域. 这样, 网络能够避免无关特征的干扰, 聚焦于重要区域, 并捕捉长距离的依赖关系. 这种稀疏注意力机制不仅减少了全局计算的复杂度, 还提高了处理效率. 此外, SATM 能够适应不同场景和条件下的特征变化, 增强了模型的鲁棒性.

如图 2 所示, SATM 首先对输入特征进行 LayerNorm 操作, 以标准化特征分布, 增强训练的稳定性 and 收敛速度. 然后, 通过 1×1 卷积和 3×3 深度卷积对逐通道特征进行编码, 以捕捉上下文信息. 接着, 将自注意力机制应用到空间和通道维度, 以降低时间和空间复杂度. 随后, 计算所有重塑的查询和关键字之间的像素对的相似度, 并使用 top- k 算法筛选出重要的元素, 屏蔽关注度较低的不必要元素, 从而得到注意力掩码 M .

注意力掩码 M 是一个二值化的掩码, 其中 1 表示需要关注的位置, 0 表示屏蔽的位置. 在 top- k 算法中, k 是一个固定的可调参数, 用于控制稀疏性的程度, 并决定保留元素的比例. 在本文实验中, k 的值通过经验设定为 $3/4$. 然后, 将该掩码 M 与 $\text{Softmax}(QK^T)$ 相乘保留关系较强的元素, 得到 $(QK^T)'$. 将 $(QK^T)'$ 与 V 进行矩阵乘法运算, 得到重建的注意力图. 上述过程可以通过以下公式表示:

$$\begin{cases} M = \text{top-}k(QK^T) \\ \text{Attention}(Q, K, V) = \text{Softmax}(M \otimes (QK^T))V \end{cases} \quad (2)$$

其中, Q 、 K 和 V 分别表示查询、键和值. 这里的 \otimes 表示逐元素乘法. 最后, 本文使用 1×1 卷积来调整特征图的通道将结果与输入特征相加, 得到最终的特征.

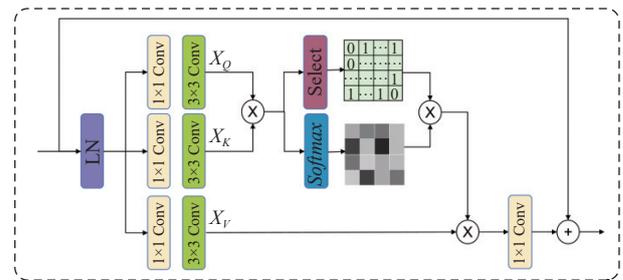


图 2 稀疏 Transformer 模块

2.3.2 通道注意力模块

通道注意力模块旨在强化图像的局部特征, 通过引入通道注意力机制来选择性地增强重要的通道特征. 如图 3 所示, 该模块首先对输入特征进行 LayerNorm 操作, 标准化输入特征以稳定训练过程并提高模型性能. 然后, 通过两个并行分支分别进行 1 次 1×1 卷积和 1 次 3×3 卷积, 总共进行 4 次卷积操作, 以提取多尺度的初步特征图: 1×1 卷积用于减少通道数量和计算量, 3×3 卷积用于捕捉局部细节. 接下来, 对上分支得到的特征图进行 ResidualBlock 操作. ResidualBlock 操作包括以下步骤: 首先, 进行全局平均池化, 将每个通道的特征图进行全局平均, 以获取通道级别的全局信息. 然后, 使用两个全连接层和 ReLU 激活函数计算注意力权重, 并通过 Sigmoid 激活函数生成通道注意力权重. 全局平均池化用于提取每个通道的全局特征, 全连接层和激活函数用于为每个通道生成对应的权重. 最终, 将生成的注意力权重与下分支得到的特征图相乘, 强调重要的通道特征, 再通过 1×1 卷积操作整合特征并

减少通道数量, 最后将结果与输入特征相加, 得到最终的特征. 通过这种设计, 该模块能够提升局部特征的提取能力, 提高网络对细节的捕捉和处理效果.

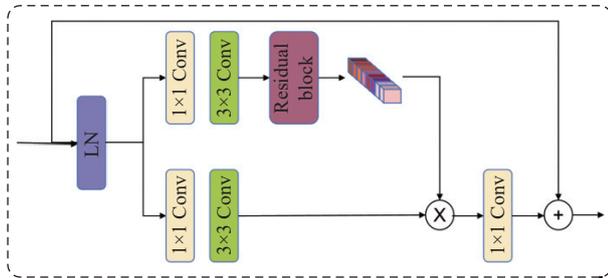


图3 通道注意力模块

2.3.3 特征融合模块

特征融合模块旨在有效整合通道注意力模块和稀疏 Transformer 模块的特征, 以充分利用两者的优势, 提高反射消除的效果.

如图4所示, 该模块首先将来自通道注意力模块、稀疏 Transformer 模块和上一个特征融合模块的特征图进行相加, 形成一个联合特征图. 接着, 对联合特征图进行两个分支的并行操作: 上分支进行全局平均池化, 提取全局特征信息, 然后通过两次卷积操作进一步处理这些全局特征. 并行的下分支则直接对联合特征图进行两次卷积操作, 提取局部细节特征. 随后, 将上分支处理后的全局特征与下分支处理后的局部特征相加, 进行 Sigmoid 激活操作, 生成融合权重. 最后, 将生成的融合权重与联合特征图逐元素相乘, 输出最终融合的特征图. 通过这种设计, 特征融合模块能够有效结合来自通道注意力模块的局部特征和稀疏 Transformer 模块的全局特征, 实现两者的优势互补.

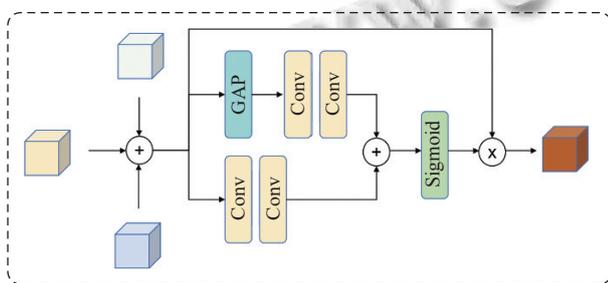


图4 特征融合模块

2.4 损失函数

本文所提出的反射消除算法中包括的背景颜色生成器、基础反射消除网络以及双通道反射消除网络, 均采用像素损失和感知损失两种损失函数进行约束.

2.4.1 像素损失

像素损失通过衡量预测的无反射图像与真实无反射图像之间的像素差异来评估重建质量. 本文采用 L1 损失函数计算像素损失, 以减少预测与真实结果之间的误差, 确保去反射图像在像素级别上的精确度. 像素损失表示为:

$$L_{\text{pixel}} = \|I_{\text{pred}} - I_{\text{gt}}\|_1 \quad (3)$$

其中, I_{pred} 表示各个网络预测的图像, I_{gt} 表示对应的真实图像.

2.4.2 感知损失

卷积神经网络的高层特征能够提取和编码图像的深层次语义信息, 包括图像的细节、质感和对象结构等关键要素. 为了增强模型对图像感知内容的理解, 本文引入了感知损失, 以确保生成的结果在语义层面上与目标图像更为一致. 感知损失表示为:

$$L_{\text{per}} = \|\varphi(I_{\text{pred}}) - \varphi(I_{\text{gt}})\|_2 \quad (4)$$

其中, $\varphi(\cdot)$ 为特征提取函数, 本文使用 VGG19 网络第 3 个卷积层的输出特征来评估图像的感知损失.

综合上述两种损失函数, 背景颜色生成器、基础反射消除网络以及双通道反射消除网络的总损失函数可以表示为:

$$L_{\text{total}} = \lambda_1 L_{\text{pixel}} + \lambda_2 L_{\text{per}} \quad (5)$$

其中, λ_1 、 λ_2 分别是各个损失函数的权重系数, 用于调节各部分损失对总损失的贡献. 在实验中分别设置为 0.5 和 1.

3 实验分析

3.1 实验配置

数据集: 本文使用 Flash 和 RRID 两个数据集. Flash 数据集^[28]是一个室内场景的反射数据集, 并且反射光偏强. 该数据集, 包括 950 个图像对, 每个图像对包含一张反射图像和一张对应的无反射图像, 其中 750 对用于训练, 200 对用于测试. RRID 数据集^[29]是在真实场景下拍摄的数据集, 场景集中在室外, 反射类型多样(弱反射和强反射). 该数据集, 包括 2 100 个图像对, 其中 1 900 对用于训练, 300 对用于测试.

评估指标: 为了评估论文中方法的性能, 本文采用了峰值信噪比 (PSNR)^[30]、结构相似性 (SSIM)^[31]两个常图像与真实无反射图像之间的差异. PSNR 值越高表

示图像失真越小,质量越高. SSIM 用于衡量图像的结构相似性,更好地反映了人类视觉系统的感知,SSIM 值越接近 1 表示图像质量越高.

实验环境:该模型使用 Adam^[32] 优化器优化,固定学习率为 1×10^{-4} . 本文首先训练背景颜色生成器、基础用的图像质量评价指标. PSNR 用于衡量去除反射光后反射消除网络各 100 个 epoch, 然后对双通道反射消除网络训练 100 个 epoch. 所有实验均在 PyTorch 环境中进行,并在一张 RTX 3090 GPU 的 Linux 系统上进行了训练和测试.

3.2 结果对比与分析

为验证算法在图像反射消除中的效果,本文将其与目前几种主流的反射消除方法进行比较,包括文献[33]、ERRN^[34]、文献[14]、IBCLN^[27]、RAGN^[20]、文献[22]和文献[21]. 为确保公平比较,在 RRID 和 Flash 这两个数据集上重新训练并测试了所有方法,实验结果见表 1.

表 1 评估得分对比

方法	RRID数据集		Flash数据集	
	PSNR (dB)	SSIM	PSNR (dB)	SSIM
文献[14]	24.201	0.847	24.940	0.922
IBCLN ^[27]	23.314	0.838	23.408	0.913
RAGN ^[20]	23.775	0.841	24.810	0.923
文献[22]	23.632	0.851	25.931	0.929
文献[33]	22.975	0.836	24.672	0.917
ERRN ^[34]	24.472	0.852	26.481	0.932
本文方法	25.172	0.865	27.649	0.944

可以得出,本文方法在两个数据集上都取得了更好的结果. 在 RRID 数据集上,对比文献[14]与 ERRN^[34],在 PSNR 指标上分别提升了 0.971 dB, 0.7 dB, 在 SSIM 指标上分别提升了 0.018, 0.013. 在 Flash 数据集上,对比文献[14]与 ERRN^[34],在 PSNR 指标上分别提升了 2.709 dB, 1.168 dB. 在 SSIM 指标上分别提升了 0.022, 0.012. 图 5 展示了一些代表性视觉反射消除的结果,以进一步验证本文方法的性能.

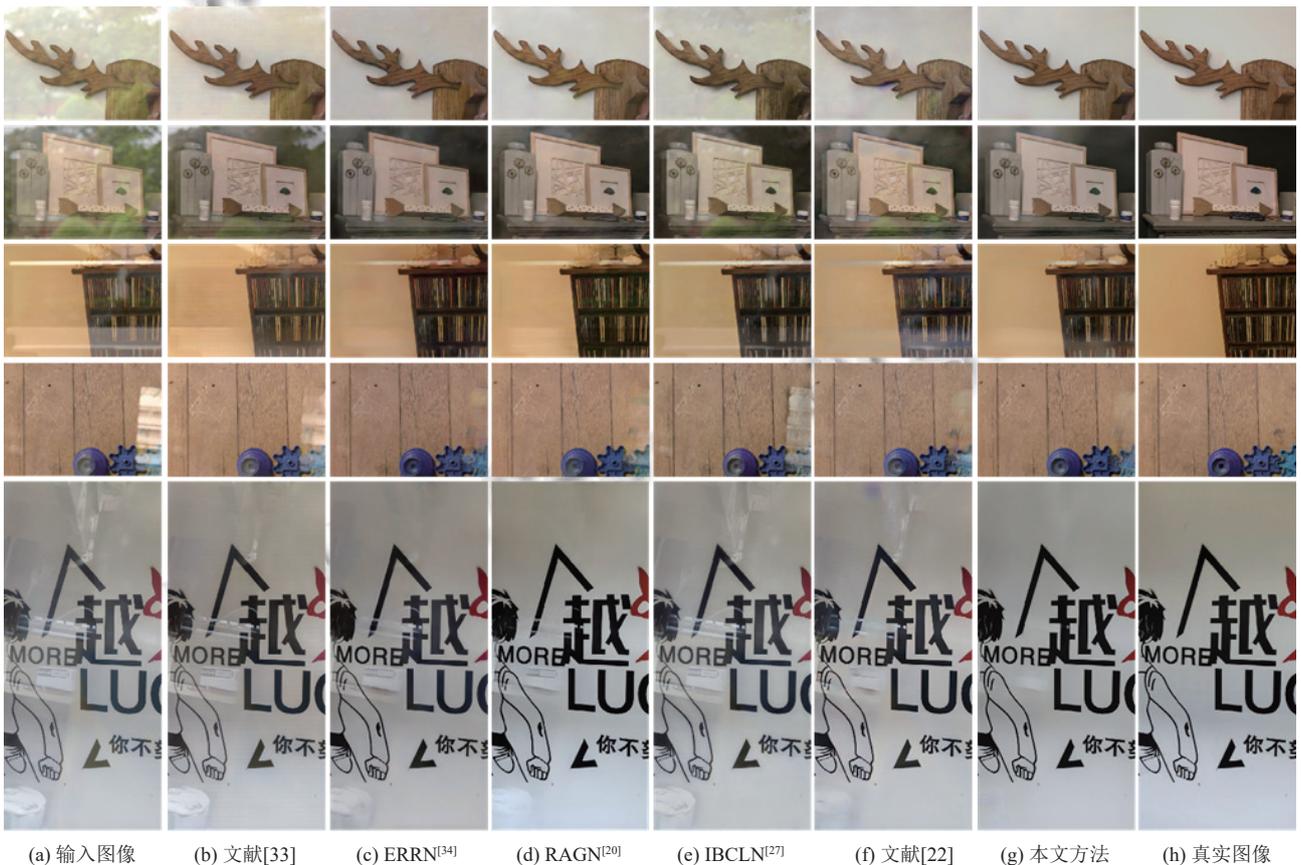


图 5 对比实验图

从图 5 中可以看到,文献[33]和 IBCLN^[27]虽然在处理大面积反射时能一定程度上减弱反射,但无法完

全去除,且在修复过程中产生了伪影,导致图像内容失真,如图 5(b) 和 (e) 所示. ERRN^[34]在较弱反射区域表

现较好,但在较强反射区域,修复信息与原始图像不一致,真实感较差,如图5(c)所示。RAGN^[20]通过估计反射层进行反射消除,但由于反射层的预测比背景层更具挑战性,不准确的反射层预测导致最终图像的反射消除效果不理想,如图5(d)所示。文献[22]的方法在简单场景下能够有效去除大部分反射光,但在复杂场景中,生成的图像仍然欠缺细节和真实感,如图5(f)所示。相比之下,本文提出的方法能够更加精确地去除反射层,并生成高质量的去反射图像。图像在细节保留、颜色准确性以及整体视觉真实感方面表现更好,如图5(g)所示。

3.3 消融实验

3.3.1 k 参数的选择与分析

在稀疏 Transformer 模块中,使用 top- k 算法来控制稀疏性的程度。 k 是一个固定且可调的参数,决定了保留元素的比例,取值范围在 0-1 之间。为确定合适的 k 值,本文进行了多组实验,分别在 k 取值为 1/4、3/8、1/2、5/8、3/4、7/8 和 1 的情况下,评估了模型在 RRID 和 Flash 数据集上的效果。实验结果如表 2 所示,随着 k 的增大,模型性能逐渐提升。当 k 为 3/4 时,模型在两个数据集上的表现最优,PSNR 分别为 25.172 dB 和 27.649 dB,SSIM 分别为 0.865 和 0.944。较小的 k 虽然降低了计算复杂度,但性能有所下降;而较大的 k 虽然包含更多信息,性能提升有限,且计算开销增加。因此,实验中, k 设置为 3/4 能够在性能和效率之间取得良好的平衡。

表 2 不同 k 值的得分对比

k 值	RRID数据集		Flash数据集	
	PSNR (dB)	SSIM	PSNR (dB)	SSIM
1/4	23.487	0.823	25.239	0.852
3/8	24.436	0.841	26.318	0.921
1/2	24.791	0.853	27.042	0.931
5/8	25.089	0.861	27.412	0.939
3/4	25.172	0.865	27.649	0.944
7/8	25.096	0.860	27.467	0.941
1	24.815	0.854	27.251	0.934

3.3.2 模块的性能验证

为了验证每个模块的性能,本文做了以下几组变体设置。

变体 1: 移除背景颜色生成器,直接使用原始图像进行处理,评估背景颜色生成器对去反射效果的影响。

变体 2: 只保留通道注意力模块的网络来处理反射

消除任务,评估双通道结合的优势。

变体 3: 移除通道注意力模块,仅使用标准的 Transformer 网络进行反射消除,评估双通道结合的优势。

变体 4: 移除通道注意力模块,仅使用本文设计的稀疏 Transformer 模块进行反射消除,以验证稀疏 Transformer 模块的有效性。

变体 5: 使用一个包含卷积和残差操作的常见融合模块替换本文设计的特征融合模块,评估特征融合模块对最终性能的影响。

本文在 RRID 和 Flash 数据集上对 5 种变体进行了评估,结果如表 3 所示,证明每个模块均提升了网络性能。对比变体 1、变体 2、变体 5 和本文方法,可以得出以下结论:在移除背景颜色生成器的情况下(变体 1),两个数据集的 PSNR 分别下降了 1.353 dB 和 1.887 dB,表明背景颜色对图像色彩保真度至关重要;在不使用注意力机制的情况下(变体 2),PSNR 在两个数据集上分别下降了 2.129 dB 和 2.821 dB,说明该机制在捕捉全局反射区域方面发挥了重要作用;在变体 5 中,使用常见的卷积和残差块组成的融合模块代替本文设计的特征融合模块,导致 PSNR 在两个数据集上也有所下降,表明特征融合模块对整体性能的提升效果显著。通过对比变体 3 和变体 4,可以看出稀疏 Transformer 模块相较于标准 Transformer 模块在反射消除任务中的优势更为明显。整体来看,完整的模型在两组数据集上均表现优异,验证了方法的有效性。图 6 展示了各变体在反射去除任务中的表现。变体 1 缺少背景颜色生成器,导致高反射区域的色彩还原不准确。变体 2 由于未使用注意力机制,反射光在强反射区域残留较多。变体 5 采用了普通的融合模块,虽然部分反射被去除,但边缘细节处理不足,影响了图像的质量。相比之下,完整的双通道网络在去除强反射光的同时,能够保留更多细节,生成的图像更加清晰且真实。

表 3 模块的消融实验

方法	RRID数据集		Flash数据集	
	PSNR (dB)	SSIM	PSNR (dB)	SSIM
变体1	23.819	0.822	25.761	0.928
变体2	23.043	0.819	24.828	0.842
变体3	23.313	0.828	25.121	0.903
变体4	24.186	0.842	25.379	0.927
变体5	24.752	0.854	25.817	0.933
本文方法	25.172	0.865	27.649	0.944

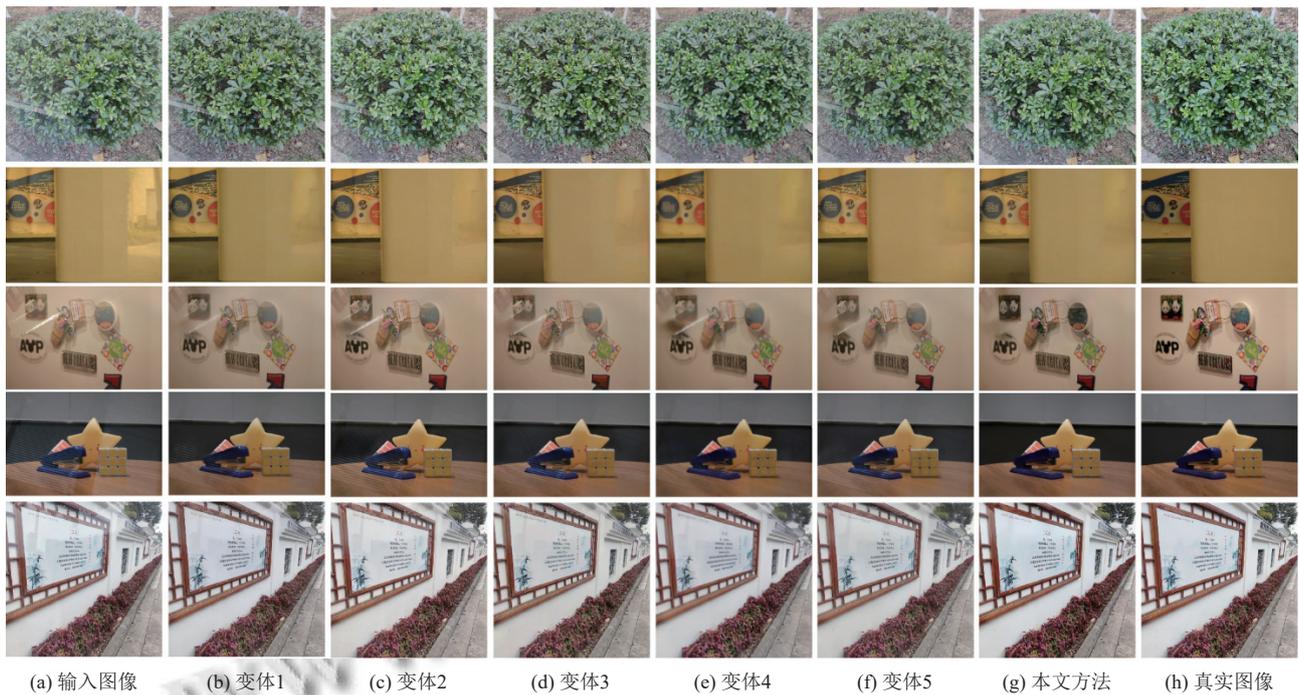


图6 消融实验的可视化结果

4 结论与展望

本文提出了一种基于颜色感知的双通道反射消除算法,解决了复杂反射光区域的细节保留和准确去除问题.该方法通过背景颜色生成器预测图像的背景颜色信息,帮助生成初步结果;随后,结合通道注意力模块的局部特征提取和Transformer模块的全局信息处理,进一步优化反射消除效果.特征融合模块有效整合了不同特征信息,提高了去除反射光的精度和自然度.实验结果表明,该算法在RRID和Flash数据集上取得了良好的表现.在未来工作中我们将侧重于提升模型的泛化能力和适应性,并探索其在更广泛的视觉任务中的应用潜力.

参考文献

- 王雷全,周家梁,林瑶.基于全局注意力信息交互的高光谱图像分类.计算机系统应用,2023,32(5):28-35.[doi:10.15888/j.cnki.csa.009120]
- 张琪东,迟静,陈玉妍,等.基于雾浓度分类与暗-亮通道先验的多分支去雾网络.计算机研究与发展,2024,61(3):762-779.[doi:10.7544/issn1000-1239.202220812]
- 金天虎,陶砚蕴,李佐勇.基于超像素图像分割的暗通道先验去雾改进算法.电子学报,2023,51(1):146-159.
- 孙超男,易芹,崔丽.小波变换结合模糊聚类在示温漆彩色图像分割中的应用.软件学报,2012,23(S2):64-68.
- Fan QN, Yang JL, Hua G, *et al.* A generic deep architecture for single image reflection removal and image smoothing. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 3258-3267.
- Wen Q, Tan YJ, Qin J, *et al.* Single image reflection removal beyond linearity. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3766-3774.
- Zheng Q, Shi B, Chen J, *et al.* Single image reflection removal with absorption effect. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2021. 13395-13404.
- Levin A, Zomet A, Weiss Y. Learning to perceive transparency from the statistics of natural scenes. Proceedings of the 15th International Conference on Neural Information Processing Systems. Vancouver: ACM, 2002. 1271-1278.
- Levin A, Zomet A, Weiss Y. Separating reflections from a single image using local features. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington: IEEE, 2004.
- Li Y, Brown MS. Single image layer separation using relative smoothness. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 2752-2759.
- Levin A, Weiss Y. User assisted separation of reflections

- from a single image using a sparsity prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(9): 1647–1654. [doi: [10.1109/TPAMI.2007.1106](https://doi.org/10.1109/TPAMI.2007.1106)]
- 12 Arvanitopoulos N, Achanta R, Susstrunk S. Single image reflection suppression. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1752–1760.
- 13 Wan RJ, Shi BX, Hwee TA, *et al.* Depth of field guided reflection removal. *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*. Phoenix: IEEE, 2016. 21–25.
- 14 Zhang XE, Ng R, Chen QF. Single image reflection separation with perceptual losses. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 4786–4794.
- 15 Dong Z, Xu K, Yang Y, *et al.* Location-aware single image reflection removal. *Proceedings of the 2021 IEEE International Conference on Computer Vision*. IEEE, 2021. 5017–5026.
- 16 Yang J, Gong D, Liu LQ, *et al.* Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal. *Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018. 675–691.
- 17 Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016. 694–711.
- 18 Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, 115(3): 211–252. [doi: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y)]
- 19 Li C, Yang YX, He K, *et al.* Single image reflection removal through cascaded refinement. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 3562–3571.
- 20 Li Y, Liu M, Yi YL, *et al.* Two-stage single image reflection removal with reflection-aware guidance. *Applied Intelligence*, 2023, 53(16): 19433–19448. [doi: [10.1007/s10489-022-04391-6](https://doi.org/10.1007/s10489-022-04391-6)]
- 21 Hu QM, Guo XJ. Single image reflection separation via component synergy. *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris: IEEE, 2023. 13092–13101.
- 22 Song ZB, Zhang ZY, Zhang KH, *et al.* Robust single image reflection removal against adversarial attacks. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 24688–24698.
- 23 Rother C, Kolmogorov V, Blake A. “GrabCut”: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (TOG)*, 2004, 23(3): 309–314. [doi: [10.1145/1015706.1015720](https://doi.org/10.1145/1015706.1015720)]
- 24 Maddalena L, Petrosino A. Background subtraction for moving object detection in RGBD data: A survey. *Journal of Imaging*, 2018, 4(5): 71. [doi: [10.3390/jimaging4050071](https://doi.org/10.3390/jimaging4050071)]
- 25 Li H, Wang S, Zhang W, *et al.* Image inpainting based on scene transform and color transfer. *Pattern Recognition Letters*, 2010, 31(7): 582–592.
- 26 Chen S, Atapour-Abarghouei A, Shum HPH. HINT: High-quality inpainting Transformer with mask-aware encoding and enhanced attention. *IEEE Transactions on Multimedia*, 2024, 26: 7649–7660. [doi: [10.1109/TMM.2024.3369897](https://doi.org/10.1109/TMM.2024.3369897)]
- 27 Song YD, He ZQ, Qian H, *et al.* Vision Transformers for single image dehazing. *IEEE Transactions on Image Processing*, 2023, 32: 1927–1941. [doi: [10.1109/TIP.2023.3256763](https://doi.org/10.1109/TIP.2023.3256763)]
- 28 Lei CY, Chen QF. Robust reflection removal with reflection-free flash-only cues. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 14806–14815.
- 29 张玲, 王国浩. 基于多角度感知和边缘引导的图像反射消除算法. *计算机技术与发展*, 2024, 34(9): 38–46.
- 30 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 2008, 44(13): 800–801. [doi: [10.1049/el:20080522](https://doi.org/10.1049/el:20080522)]
- 31 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612. [doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861)]
- 32 Kingma DP, Ba J. Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations*. San Diego: ICLR, 2015.
- 33 Chang YC, Lu CN, Cheng CC, *et al.* Single image reflection removal with edge guidance, reflection classifier, and recurrent decomposition. *Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa: IEEE, 2021. 2032–2041.
- 34 Wei KX, Yang JL, Fu Y, *et al.* Single image reflection removal exploiting misaligned training data and network enhancements. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 8170–8179.

(校对责编: 张重毅)