文章编号:1674-2974(2025)08-0023-10 DOI:10.16339/j.cnki.hdxbzkb.2025280

基于深度自监督学习的可微分半色调框架

刘登峰[1,2],朱佳伟[1,2],徐昊[1,2],杜晓凯[1,2],柴志雷[1,2†]
(1. 江南大学 人工智能与计算机学院,江苏 无锡 214122;
2. 康养智能化技术教育部工程研究中心,江苏 无锡 214122)

摘 要:针对当前数字半色调算法处理速度慢以及半色调效果不佳的局限性,提出一种基于数据驱动的半色调框架.通过引入Gumbel-Softmax重参数化策略解决半色调离散选择带来的不可微分问题,实现了网络反向传播过程中的梯度无偏估计.为进一步强化半色调图像的效果,设计出一种新型蓝噪声损失函数,对半色调网点的分布予以优化.同时,提出区域置信度聚合模块,通过结合像素的空间相关性,使网络在训练过程中更加注重像素之间的交互信息.基于以上策略,通过优化半色调质量评估的期望值,构建了一个不需要标签引导的自监督可微分半色调处理框架.实验结果表明,所提出的方法不需要图像标签,能够在保持较高处理速度和较低参数复杂度的前提下,生成高质量的半色调图像,有效保留图像的局部结构信息和纹理细节.并且,该框架可灵活扩展至多级半色调处理,以适应多级打印喷头的需求.

关键词:半色调;深度学习;梯度估计;蓝噪声特性
中图分类号:TP314 **文献标志码**:A

A Differentiable Halftoning Framework Based on Deep Self-supervised Learning

LIU Dengfeng[1,2],ZHU Jiawei[1,2],XU Hao[1,2],DU Xiaokai[1,2],CHAI Zhilei[1,2†]
(1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China;
2. Engineering Research Center of Intelligent Technology for Healthcare, Ministry of Education, Wuxi 214122, China)

Abstract: To address the limitations of current digital halftoning algorithms, such as slow processing speed and suboptimal halftoning effects, a data-driven halftoning framework is proposed. By introducing the Gumbel-Softmax reparameterization strategy, the non-differentiability issue caused by discrete halftone selection is resolved, enabling unbiased gradient estimation during network backpropagation. To further enhance the halftoning effects, a novel blue noise loss function is designed to optimize the distribution of halftone dots. Additionally, a Patch-wise Confidence Aggregation module is introduced to incorporate spatial correlations between pixels, allowing the network to focus more on pixel interactions during training. Based on these strategies, a label-free, self-supervised,

* **收稿日期**:2024-08-24
基金项目:国家重点研发专项计划(2022YFE0112400), National Key Research and Development Special Plan(2022YFE0112400); 国家自然科学基金资助项目(21706096), National Natural Science Foundation of China(21706096); “科技兴蒙”行动重点专项(NMKJXM202210), The “Technology Empowering Inner Mongolia” Key Special Initiative (NMKJXM202210)
作者简介:刘登峰(1980—),女,河南南阳人,江南大学副教授,博士
† 通信联系人,E-mail: zlchai@jiangnan.edu.cn

differentiable halftoning framework is constructed by optimizing the expected value of the halftone quality metric. Experimental results demonstrate that the proposed method, without requiring image labels, can generate high-quality halftone images and maintain high processing speed and low parameter complexity, effectively preserving local structural information and texture details. Moreover, this framework can be flexibly extended to multi-level halftoning to accommodate the requirements of multi-level printheads.

Key words: halftoning; deep learning; gradient estimation; blue noise characteristics

半色调作为将连续色调图像转换为其近似的离散版本的技术,在打印领域应用广泛.在半色调处理中,通过像素点的不同排列与像素密度来营造不同的灰度层次.由于人类视觉系统的低通滤波特性,肉眼在一定的观察距离下会将半色调图像感知为平滑过渡的连续色调图像.半色调算法依据像素点排布的周期性和其为聚集点还是分散点来分类,非周期性分散点半色调算法生成的半色调图片通常由均匀分布且离散的黑白点组成,往往具有蓝噪声特征,符合人眼感知特性,因此,与蓝噪声属性相匹配的半色调图案是生成高质量半色调图像的关键.彩色半色调的实现通常将彩色图像分解为多个颜色通道,对每个通道的颜色强度信息分别进行半色调处理.在印刷或显示过程中,通过特定方式对这些经过半色调处理的通道进行叠加,网点之间相互作用,会再现原始图像的色彩和亮度.因此,灰度图像的半色调处理方法是扩展到彩色半色调的基础,理解和优化灰度半色调处理对实现高质量的彩色半色调至关重要.在半色调技术研究中,图像质量与处理效率是主要的研究重点.常见的半色调算法可以分为三类:有序抖动算法[1]、误差扩散算法[2-4]以及基于搜索的算法[5-7].有序抖动算法将连续色调图像系统地划分为均匀的小区域,并将其与预先设计或生成的抖动矩阵进行比较处理.该方法凭借其高效的并行处理能力和较低的计算复杂度,为快速图像半色调处理提供了有效途径,然而,这种速度优势往往以牺牲半色调图像的质量为代价.误差扩散算法通过精细的像素级处理和误差传递机制,在视觉质量上取得了显著进步,但算法的串行本质[8]及潜在的视觉伪影问题成为进一步优化的阻碍.基于搜索的算法将半色调处理视为一个复杂的优化问题.这类方法首先定义一个综合考虑人类视觉系统(HVS)特性的半色调质量评估指标[9],然后利用启发式算法,如模拟退火算法或直接二分搜索算法对这些评估指标进行精确优化.这种方法直接对专门设计的指标(例如像素级的均方误差)进行调整,因此能够生成质量较优的半色调图像,然而其高昂的计算成本成为限制此项技术广泛应用的主要障碍.

深度学习模型特别是卷积神经网络,因其在图像识别、生成和转换任务中的出色表现而备受关注.监督学习、无监督学习和强化学习是深度学习中常见的三种范式,在半色调处理中,卷积神经网络通过学习成对图像数据之间映射关系,期望通过单次前向传播实现图像的高效抖动处理[10-13].然而,此学习过程面临两大主要挑战:首先,准备“真实”半色调图像作为训练标签成本高昂[14],寻求最佳半色调图像需要庞大的计算资源.现有的基于搜索的方法要么仅能优化特定的指标度量,要么依赖于需要针对每个实例进行调参的元启发式搜索算法.其次,半色调处理本质上是一个一对多的映射问题,即同一连续图像可能对应多种有效的半色调表现形式.若直接采用像素级损失函数(例如交叉熵)来优化,可能会导致模型仅学习到输入图像的平均表现形式,无法满足半色调处理的离散性要求.因此,如何在没有大规模标注数据集的情况下,依靠自监督的方式来实现半色调[15],是半色调研究的重要内容.在深度学习应用于半色调处理的过程中,一个核心问题源自半色调图像的强二值特性,这直接导致离散选择(如确定性地选择最佳像素状态)在训练过程中不可微,从而阻碍了梯度的反向传播.现有研究一方面利用生成对抗网络从预先准备好的半色调标签数据集进行训练,旨在最小化输出和标签之间的皮尔逊χ^2散度,但因其需要构建大量标签数据集,实用性受到一定限制.另一方面研究主要解决梯度回传问题,一种策略是梯度直通估计器[16].该方法在前向传播中执行实际的离散操作,而在反向传播过程中,则直接将梯度传递至离散变量的输入,忽略离散化带来的非连续性,这种方式削弱了二值化的特性,因此需要引入

离散化惩罚损失,将输出值推向最近的离散端[17].然而,这种贪婪的二值化规则可能会损害全局角度的优化.梯度直通估计器虽然简单易行,但其梯度估计的不准确性可能影响模型的优化质量;也有研究采用REINFORCE等基于策略梯度的方法,由于半色调图像处理需要对所有像素进行采样[18],其梯度估计通过累积往往伴随较高的方差,导致训练效率低下,模型难以收敛.

本文针对上述问题,提出一种基于自监督学习的可微分半色调框架.无须依赖预先准备好的半色调标签数据集进行训练,而是通过优化半色调评价度量的期望训练一个轻量级的残差网络,所做的主要贡献如下:

1) 采用重参数化策略提供了一种平滑可微的近似方法来模拟离散分布的采样过程,使模型能够有效地估计梯度,并进行端到端的优化.

2) 对半色调图像网点分布加以分析并提出蓝噪声损失,令半色调图像展现出理想的蓝噪声特性.

3) 提出区域置信度聚合评估机制.通过综合考量像素间的相互依赖性,有效地模拟人类视觉系统的感知特性,从而维持图像的局部结构信息和纹理细节.

4) 展示了该框架从二级到多色调处理的灵活拓展能力,显示了其在不同半色调场景下的应用潜力.

1　方法

本节将首先探讨基于数据驱动的深度学习半色调处理框架,该框架通过精确的梯度估计解决了由离散选择导致的不可微分问题.随后介绍了蓝噪声损失函数,此函数设计用于优化生成图像的视觉感知质量,并诱导图像展现出理想的蓝噪声特性.最后,提出了区域置信度聚合机制,这种机制综合考虑了局部区域信息,能够更全面地优化图像的纹理和结构细节.

1.1　网络架构

本文根据蓝噪声的抖动模式,提出了一种基于自监督学习的卷积残差网络的模型.由于卷积神经网络具有空间共享内核的卷积范式,其所带来的归纳偏置容易产生平坦性退化的现象.具体而言,常数信号$s(x)\equiv b$与任意核函数$k(x)$的卷积仍然是一个常数信号$\check{s}(y)=b\cdot\mu(k(x))$,其中$\mu(k(x))$表示核函数的均值,因此,给定平坦输入$X$,无论CNN参数如何,CNN的操作都会退化为缩放操作$Y=\alpha X$.于是本文在不破坏原始输入信息完整性的前提下,利用高斯噪声图作为空间变化,并将其引入特征空间中为模型提供足够的抖动依赖[17-20].图1展示了深度半色调的具体流程,网络输入由下述公式定义:

$$I_c = \mathrm{Concatenate}(c, z) \tag{1}$$

$$I_{c_g} = \mathrm{Concatenate}(c_g, z) \tag{2}$$

式中:c是连续调灰度图像;c_g是额外采样的恒定灰度图像;z则为动态采样的高斯噪声,引入的随机性噪声有助于引导模型聚焦于整体图案的统计分布特性而非单一像素值.

网络主体包含多个残差块,每个残差块处理输入特征并将处理后的特征图与原始输入相加,这些残差块有助于在深层网络中有效传递信息,从而避免梯度消失问题,使得网络在学习深层特征时保持

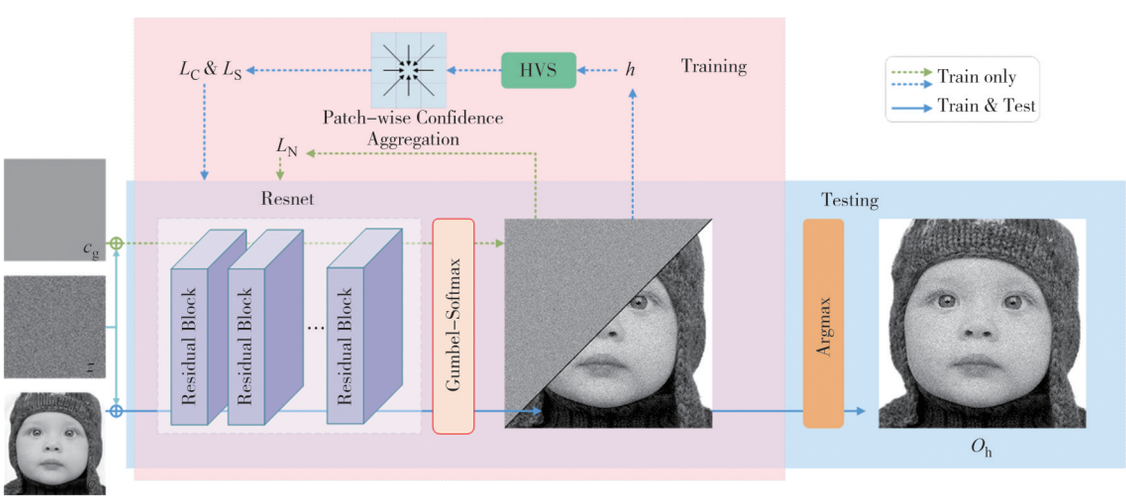


图1　深度可微分半色调框架

Fig.1　Deep differentiable halftoning framework

性能的稳定.

在深度框架网络输出的最后阶段,采用Gumbel-Softmax操作对模型处理结果进行离散化[21],并实现梯度的无偏估计,生成半色调图像.Gumbel-Softmax策略旨在解决无法直接将重参数化技巧应用于离散数据的问题,这种策略基于两个洞察:利用Gumbel分布能够实现离散分布的参数化;argmax函数本身不具备连续性,通过使用温度参数控制的Softmax函数,为离散选择过程提供了一个连续且可微的近似.具体来说,对于给定的离散随机变量X,设其具有K个类别,且第k个类别非标准化对数概率表示为$\log \pi_k$.对每个类别k都生成一个独立的Gumbel噪声:

$$G_k = -\log(-\log(U_k)) \tag{3}$$

其中U_k服从均匀分布$U\sim\text{Uniform}(0,1)$.对于这k个事件的类别分布,概率分别为$\pi_1,\cdots,\pi_k$,计算出$\log\pi_1,\cdots,\log\pi_k$,并且在这些项中都加入从Gumbel分布中采样的独立同分布噪声,然后通过argmax函数来表征随机样本X:

$$X = \underset{k}{\text{argmax}}\ (\log \pi_k + G_k) \tag{4}$$

在离散变量的处理中,直接基于argmax操作选择的过程是非连续的[22],则通过依赖于温度参数τ的Softmax函数将离散变量松弛为连续变量:

$$x_i = \frac{\exp\left((\log \pi_i + G_i)/\tau\right)}{\sum_{j=1}^{k}\exp\left((\log \pi_j + G_j)/\tau\right)} \tag{5}$$

其中i=1,$\cdots$,k,因此,该策略将参数$\pi_1,\cdots,\pi_k$的依赖性从不可微的随机采样函数转移到了由Softmax和log运算组成的可微函数上,而且其内在机制巧妙地延续了初始Concatenate操作中引入的噪声效益,即通过引入基于Gumbel分布的噪声,不仅有效地模拟和扩展了初始噪声的随机性影响,且进一步增强了模型处理离散输出的能力.

由于半色调的离散性,通过人类视觉系统(HVS)过滤后图像能够有效地模拟人眼的感知特性[23].本文在此基础上进行端到端的优化.

1.2 蓝噪声损失

在半色调处理中,蓝噪声纹理图案具有非周期性特征的同时避免了低频颗粒感,符合人眼感知[24].符合蓝噪声的半色调应保持以下频谱特性:1)含有较少的低频分量;2)高频区的能量分布相对平坦;3)各向异性在所有频率上都非常低.以往的研究表明,在处理恒定灰度图像时,仅仅依赖降低通过人类视觉系统过滤后图像的均方误差损失和结构相似性损失,并不足以确保较好的蓝噪声质量[25].此外,卷积神经网络的并行处理特性易引发全局一致性的棋盘伪影,这进一步导致了半色调图像质量下降.Xia等人[17]通过离散余弦变换惩罚半色调图像的低频分量,虽然在一定程度上最小化了这些分量,但并未显著解决过度的各向异性问题,这种过度的各向异性在视觉上可能导致不受欢迎的纹理偏向.为了优化生成图像的蓝噪声特性并评估其质量[18],避免模型倾向于极端输出,本文引入了基于功率谱方差的度量方法.首先,通过离散傅里叶变换(DFT)计算半色调h图像的功率谱$\hat{P}(f)$:

$$\hat{P}(f) \approx \frac{1}{N}\left|\text{DFT}(h)\right|^2 \tag{6}$$

其中N是样本的像素总数,进一步地,为了得到径向平均功率谱密度$P(f_\rho)$(RAPSD),将其定义为每个频率环f_ρ的平均功率谱:

$$P(f_\rho) = \frac{1}{n(r(f_\rho))}\sum_{f\in r(f_\rho)}\hat{P}(f) \tag{7}$$

其中$n(r(f_\rho))$是径向频率f_ρ周围宽度为$\Delta\rho=1$的圆环中离散频率样本的数量,此外,定义相同频率的样本方差为:

$$V(f_\rho) = \frac{1}{n(r(f_\rho))-1}\sum_{f\in r(f_\rho)}\left(\hat{P}(f)-P(f_\rho)\right)^2 \tag{8}$$

由于方差用于衡量数据离散程度,可以显式地表征半色调点的分布特性和波动程度,基于这些分析,设计了一种蓝噪声损失函数:

$$L_{\text{N}} = \text{MSE}(V_c(f_\rho), V_h(f_\rho)) \tag{9}$$

其中$V_c(f_\rho)$是理想的蓝噪声频谱方差,通过量化对比模型输出与理想蓝噪声频谱方差的均方误差(MSE),该损失函数能够精确评估模型在保持图像蓝噪声属性方面的效果.由于光谱分析只对恒定灰度的半色调图像有意义,因此本文在额外的小批量恒定灰度图像C_{g}上优化该损失函数.

为了综合评估并优化模型生成的半色调图像的质量,定义一个总体损失函数L_{total},结合不同的损失成分:

$$L_{\text{total}} = L_{\text{C}} + w_1\cdot L_{\text{N}} + w_2\cdot L_{\text{S}} \tag{10}$$

其中,L_{C}是基于HVS滤波(其核大小为11×11)后半色调图像h和原图c的均方误差[23]的损失,其定义为:

$$L_{\text{C}} = \text{MSE}(\text{HVS}(h), \text{HVS}(c)) \tag{11}$$

L_{S}是结构相似性指数(SSIM)损失,用于评估图像之间的相似性:

$$L_{\text{S}} = 1 - \text{SSIM}(h, c) \tag{12}$$

其中超参数w_1=0.02、w_2=0.01是经验设置值.在训练过程中,将总体损失函数L_{total}作为最终优化目标,指导模型在学习过程中同时优化图像的视觉相似度、结构相似性和蓝噪声特性.通过这种方式,不仅能够生成在视觉上与原图像相近的半色调图像,还能确保其具备理想的蓝噪声特性,从而在视觉效果上达到最优平衡.

1.3 区域置信度聚合

在半色调图像质量评估中,现有度量标准如均方误差和结构相似性指数虽然能够量化图像间差异,但它们往往受限于对像素级精确度的片面追求,忽略了像素间的相互作用对整体感知质量的影响.这些指标通过人类视觉系统模型对目标图像和参考图像进行预处理,以模拟人眼对图像细节的敏感度,继而生成反映图像差异的误差图并对其进行平均计算得到标量度量结果.像素的表现不仅仅依赖于其自身的值,还会受到其所在局部窗口内其他像素的影响,这一作用范围由HVS滤波器所限定的窗口大小决定.然而,仅通过广义平均池(generalized mean pooling)直接优化平均精度,将每个网络参数的梯度简化为所有像素处梯度的平均值,该处理方式将所有像素对的置信度成本等同对待,未能充分考虑到半色调图像中像素点排列模式对生成图像质量的关键影响[26-27].

针对上述局限,提出区域置信度聚合模块,即摒弃孤立审视像素的做法,转而采取一种更加全局化的视角,将局部像素的置信度聚合到统一的聚类中心来进行处理.旨在通过综合考虑像素所在局部区域的质量评估,更全面地优化图像的纹理和结构细节.具体而言,利用模型输出的均方误差构建逐像素匹配置信度特征图F,采用区域置信度聚合机制,从特征图中提取密集采样子区域(大小为11×11)的局部描述符$D_{i,j}$.每个描述符不仅包含当前像素的信息,还综合了其邻域内所有像素的置信度值.同时引入可学习的卷积滤波器权重w,用于调节不同像素在聚合过程中的贡献度,确保模型能够自适应地捕捉图像的局部特征,局部描述符由下式来定义:

$$D_{i,j}=\sum_{(u,v)\in N(i,j)} w_{u,v}\cdot F_{i+u,j+v} \tag{13}$$

式中:$N(i,j)$表示以(i,j)为中心、大小为11×11的邻域.通过将区域置信度聚合的邻域大小与HVS滤波器窗口保持一致,使得模型能够在和HVS相同的感知范围内,从而更有效地聚合邻域内的置信度信息.与此同时,w可以视为模型中卷积层的一部分,并且可与网络参数一同学习.为了更快达到收敛目标,将初始权重值设置为标准差为2的高斯分布.在这种机制下,模型的决策输出不仅基于当前像素点,而且通过标记像素之间的相关性,更好地捕捉图像的上下文信息,从而增强了模型的空间感知能力.

2 实验与分析

本节将介绍方法的实现细节,比较本文提出的工作与现有半色调方法的性能差异,进行相关消融实验,并探讨其在不同应用场景中的可扩展性.

2.1 实验设置

本文采用VOC2012数据集进行训练和评估,随机选取其中的13 758幅图像作为训练集,并保留其中的1 684幅和1 683幅图像作为验证集和测试集.基于自监督学习,通过设计特定的损失函数,使模型在训练时无须依赖标注数据,从而能够利用大量未标注的灰度图像进行学习.为了确保实验的一致性和可比性,所有图像在处理前都被转换为灰度图像,将所提出的方法与其他方法进行了全面的比较.为了进一步证明模型的有效性以及泛化能力,在多个公开数据集上进行测试,包括Set5,Set14,BSD100,BSD200,General100,Manga109,Urban100,DIV2K Val数据集.

本文采取残差网络作为半色调的主干网络,在训练中,将批量大小设置为64,并对训练图像进行64×64的随机裁剪.训练过程中通过最小化损失函数L_{total}来训练网络,取超参数w_1=0.02、w_2=0.01,重参数化的初始温度系数τ设为0.5.整个网络模型选取ADAM优化器进行训练,学习率α通过余弦退火计划从3×10^{-4}调整为1×10^{-5},整个训练过程在NVIDIA RTX 2080Ti GPU上进行400个轮次.

2.2 半色调质量评价

为了全面评估提出的方法在半色调图像生成任务中的性能,本文进行了系统的定量评估和视觉质量分析,对比了多种现有的半色调方法,包括基于Void-and-cluster方法(VAC)[1]、Ostromoukhov的方法(OVED)[4]、基于优化的搜索方法(DBS)[6],以及基于深度神经网络的方法(RVH[17]和TRH[12]).表1中详细列出了所有方法在测试数据集上的定量结果,其中base表示未采用区域置信度聚合机制.通过

HVS过滤的半色调和输入连续色调之间的峰值信噪比(PSNR)来测量色调一致性,并通过结构相似性指数(SSIM)来评估半色调图像的结构和纹理信息.

表1 半色调方法定量比较

Tab.1 Quantitative comparison of halftoning methods

方法	SSIM	PSNR	参数量	运行时间/ms
VAC	0.093 2	31.152	4 k	0.03
OVED	0.105 9	34.679	384	2.36
DBS	0.079 8	**36.577**	—	1 010
RVH	0.158 5	30.753	75.6 M	40.25
TRH	0.147 7	28.632	37.8 M	37.07
Ours /base	0.160 9	32.171	1.4 M	33.49
Ours	**0.166 9**	32.262	1.4 M	33.49

实验结果表明,所提方法在PSNR方面取得了具有竞争力的分数,并在SSIM方面获得了最佳表现.值得注意的是,尽管DBS方法通过优化搜索策略获得了最高的PSNR值,但极高的PSNR并不代表视觉上更佳的半色调效果,且极致追求PSNR指标会牺牲图像结构细节.图2(d)展示了DBS方法生成的结果在视觉上丢失局部细节.本文方法则更加重视图像结构纹理和色调一致性的平衡.

表1还比较了不同方法的参数量以及在512×512像素的"Lenna"测试图像上的运行时间.结果显示,VAC方法虽然运行时间最短,但其生成的半色调图像质量相对较低.另一方面,基于优化的搜索方法(DBS)采用计算密集型操作,较长的运行时间限制了其在实际应用中的可行性.相比之下,基于深度神经网络的方法能够生成细节更丰富、质量更高的半色调图像.同时,本文的方法在减少参数量和缩短运行时间方面均优于RVH和TRH方法,表现出更高的计算效率和更低的时间开销.总体上看,本文提出的框架在图像质量评估和计算成本方面处于领先地位.

图2展示了各方法对"Snail Shaped Organ"图像的半色调效果.其中VAC、OVED和DBS,未能有效保留图像的静脉结构细节,因为这些方法在生成半色调图像时忽视了图像的结构相似性,导致在处理边缘细节时表现不佳,使得半色调图片过度模糊化并丢失了细节.尽管这些方法能够在一定程度上抑制伪影,但这通常以牺牲细微纹理和边缘信息为代价.相比之下,本文的方法有效地平衡了伪影抑制与细节保留之间的关系,更准确地反映了原始图像的结构特征.不仅在抑制伪影方面表现出色,而且避免了过度模糊,成功地保留了图像的关键特征,这使得本文的方法在半色调图像的生成质量上显著优于其他方法.

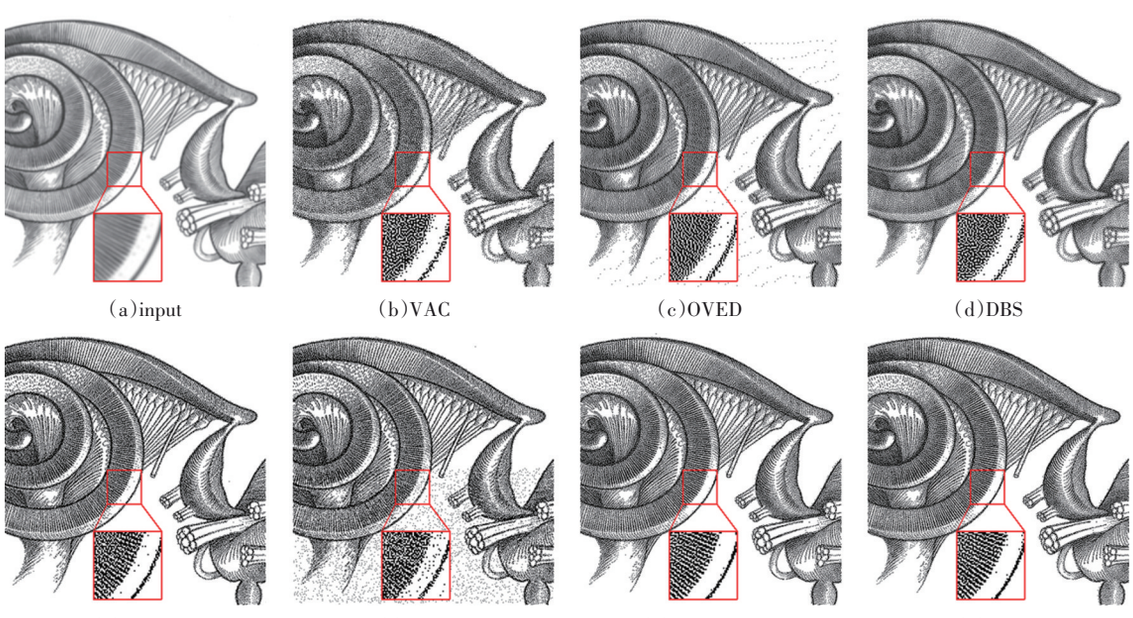
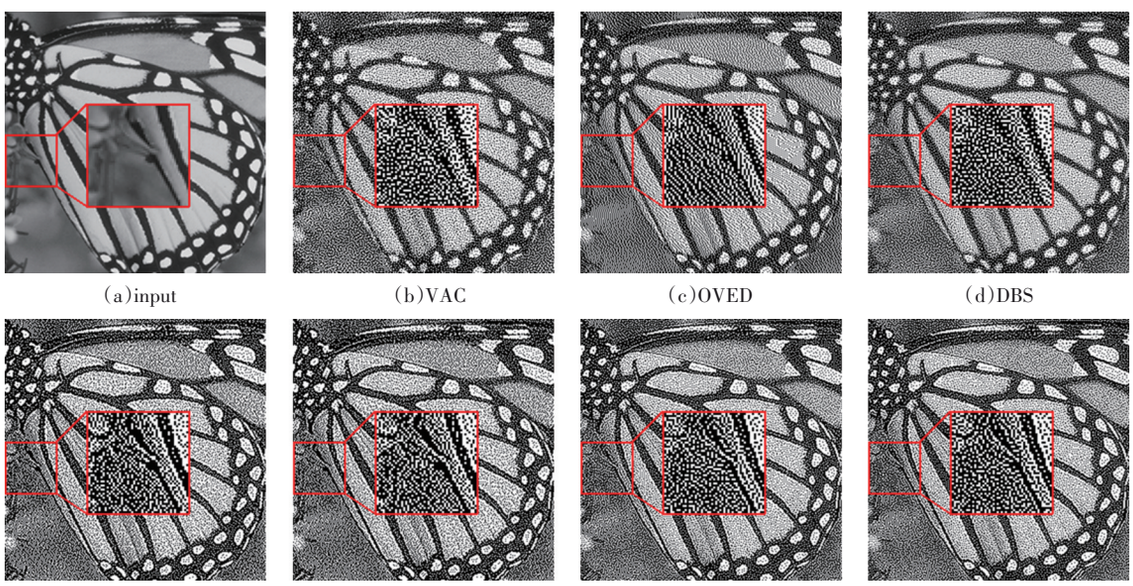


图2 "Snail Shaped Organ"的半色调效果

Fig.2 Halftoning result of the "Snail Shaped Organ"

图3展示了真实图像及对其采用不同方法生成的半色调结果.可以观察到,OVED、RVH和TRH方法生成的半色调图像中出现了明显的棋盘状网格伪影,且OVED具有相对严重的斜向纹理特征,尽管这

种模式在视觉上呈现出相对较低的低频特征,但它未能完全满足蓝噪声的理想特性;而RVH和TRH方法存在点聚集现象,即像素点倾向于在某些区域聚集,这样的点分布不属于理想中的随机均匀分布,这进一步影响了半色调图像的整体视觉效果.这些方法在优化低频特征的同时,未能有效地处理蓝噪声分布和点分散问题.相比之下,本文方法生成的半色调图像不仅有效避免了棋盘状网格图案的产生,还显著减少了点聚集现象,能够生成具有更自然纹理的半色调图像.



(a)input　(b)VAC　(c)OVED　(d)DBS

(e)RVH　(f)TRH　(g)Ours/base　(h)Ours

图 3　“Butterfly”的半色调效果

Fig.3　Halftoning result of the "Butterfly"

为了验证所提方法的有效性,在多个公开测试集上对不同方法的PSNR和SSIM指标进行了对比实验,定量指标比较如表2所示.本文方法在各个测试集上均表现出最好的SSIM指标和较好的PSNR指标,验证了该方法在保持图像细节和结构完整性方面具有明显优势,同时也表明了其具有较好的泛化性.

表 2　公开数据集的测试结果(SSIM/PSNR)

Tab.2　Test results on public datasets (SSIM/PSNR)

数据库	定量指标	半色调方法						
		VAC	OVED	DBS	RVH	TRH	Ours/base	Ours
Set5	SSIM	0.074 1	0.081 5	0.068 7	0.117 6	0.117 3	0.122 7	**0.126 2**
	PSNR	31.854	34.419	**36.431**	30.949	30.928	32.343	32.512
Set14	SSIM	0.104 4	0.109 0	0.097 2	0.165 7	0.150 5	0.165 4	**0.166 1**
	PSNR	31.343	34.890	**36.746**	30.607	29.798	32.375	32.406
BSD100	SSIM	0.078 3	0.087 2	0.069 1	0.145 9	0.120 1	0.146 3	**0.148 6**
	PSNR	31.051	34.858	**36.599**	30.452	27.013	31.988	32.226
BSD200	SSIM	0.066 4	0.076 0	0.058 6	0.126 5	0.109 0	0.126 8	**0.129 8**
	PSNR	31.218	34.909	**36.562**	30.497	27.048	32.104	32.215
General100	SSIM	0.084 8	0.090 5	0.079 8	0.130 5	0.122 9	0.131 3	**0.140 5**
	PSNR	31.780	34.603	**36.620**	30.815	28.861	32.122	32.219
Manga109	SSIM	0.202 5	0.187 1	0.198 1	0.260 7	0.198 0	0.284 7	**0.300 5**
	PSNR	32.143	34.283	**37.134**	31.415	27.457	31.879	32.057
Urban100	SSIM	0.130 2	0.148 3	0.115 4	0.207 4	0.207 4	0.209 1	**0.213 5**
	PSNR	30.810	34.461	**36.646**	30.350	30.347	31.923	32.043
DIV2K Val	SSIM	0.067 7	0.075 4	0.062 1	0.115 2	0.115 3	0.123 2	**0.130 6**
	PSNR	31.811	34.817	**36.583**	30.956	30.952	32.114	32.256

2.3 消融实验

为了在二值化过程中实现网络框架的可微性，RVH 和 TRH 方法使用了直通估计器来实现二值化过程中的梯度回传，并提出了二值化损失.该方法尽管在某种程度上促进了离散选择，但二值化损失的存在却限制了整体图像质量的优化.与此不同，本文采用了基于 Gumbel-Softmax 重参数化的策略，在保留离散选择特征的同时，允许模型在训练时探索连续的决策空间.这样不仅实现了离散选择过程的可微性，还提高了模型的训练稳定性和效率，从而显著改善了图像质量评价的优化效果.

在半色调过程中，蓝噪声特性是影响半色调图像质量的关键因素.图 4 展示了不同恒定灰度图像的半色调结果以及径向平均功率谱密度和各向异性度量.在首行图像半色调结果图中，每个子图的左、中、右分别为不使用蓝噪声损失、使用蓝噪声损失以及 DBS 方法生成的半色调效果，可以明显看出：不使用蓝噪声损失的半色调图像存在较为突出的条纹伪影，而使用该损失能够明显解决条纹伪影问题，同时经过优化后的半色调图像在径向频率功率谱上也呈现出理想的蓝噪声特性，即较少的低频分量和较高且平缓过渡的中高频分量.此外，具有蓝噪声特性的半色调频谱的各向异性曲线也相对平缓，避免了尖锐的频率峰值或突兀的变化，这种平滑的频率响应不仅提高了图像的视觉质量，还有效抑制了可能的视觉伪影.图 5 显示了恒定灰度等级为 127 的图像在不使用 L_N 和使用 L_N 情况下的傅里叶振幅谱.从图 5(a)可以明显观察到：不使用 L_N 时半色调点分布具有明显的方向偏好，而使用 L_N 后，这种偏好显著减少，显示出更理想的蓝噪声特性.

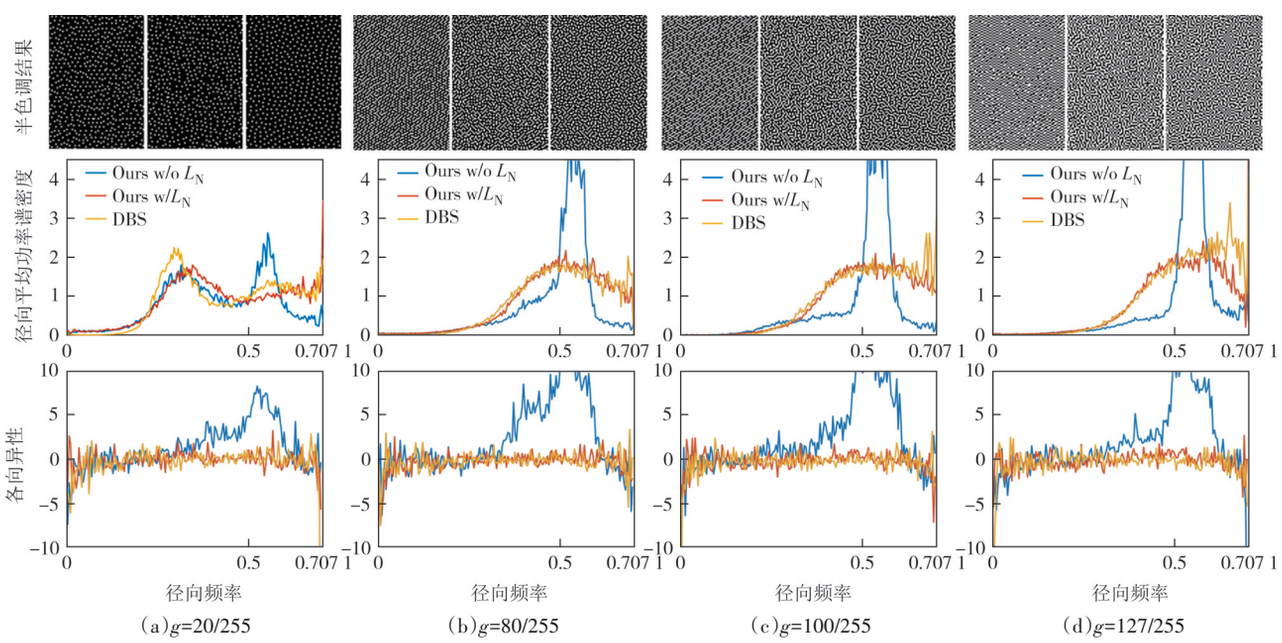


图 4 恒定灰度图像半色调表现及光谱分析

Fig.4 Halftoning performance and spectral analysis of constant grayscale images

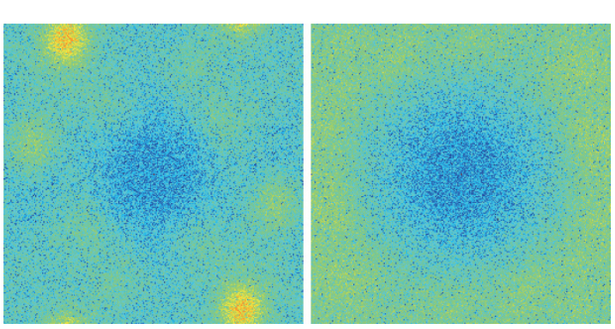


图 5 恒定灰度半色调图像的傅里叶振幅谱

Fig.5 Fourier amplitude spectrum of halftone images with constant grayscale

通过引入区域置信度聚合机制，能够更全面地捕捉和利用局部区域内的置信度信息.表 1 展示了该机制在测试数据集上的 SSIM 和 PSNR 得分均有所改进(base 表示未引入区域置信度聚合机制).图 6 展示了各种深度学习方法的 MSE 损失曲线，进一步表明通过综合考量像素间的置信度信息，模型能够更快速地收敛，从而提升训练效率.

此外，为进一步分析模型超参数对性能的影响，本文对损失函数中的超参数 w_2 进行了敏感性分析.图 7 绘制了不同 w_2 值下的优化结果，在本研究中选择 w_2=0.01，其生成的半色调图像在结构清晰度和一

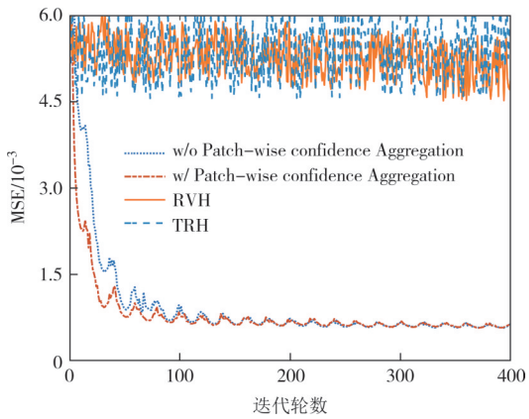


图 6　MSE 的训练损失曲线

Fig.6　Training loss curve of MSE

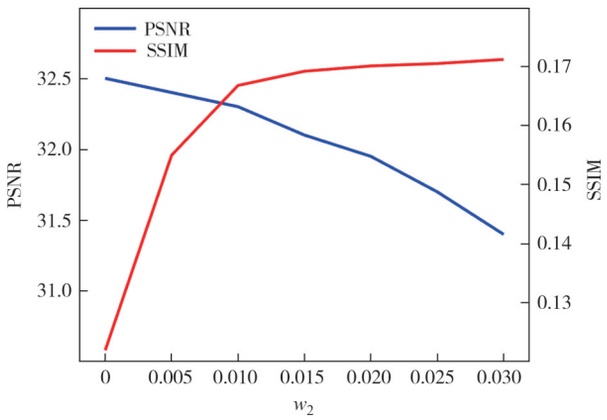


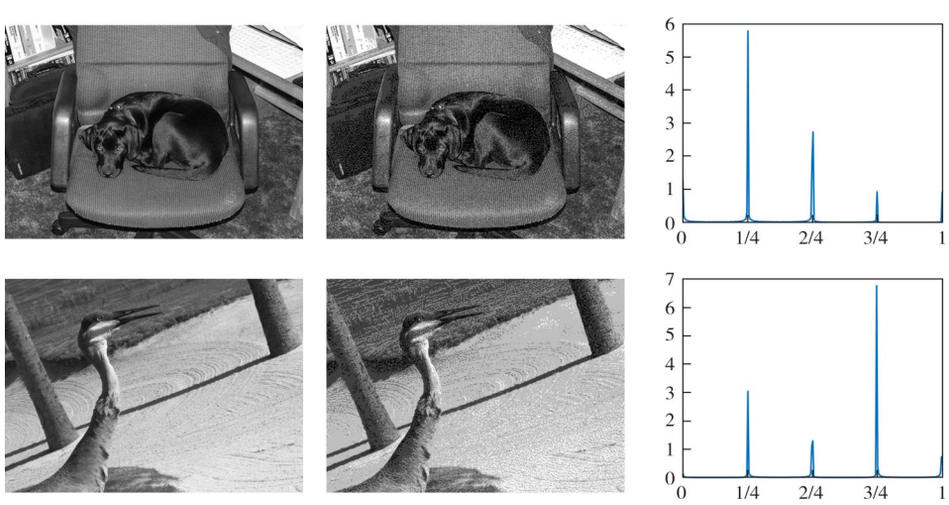
图 7　超参数 w_2 的敏感性分析

Fig.7　Sensitivity analysis for hyperparameter w_2

致性之间表现较好.根据图像质量不同偏好,可以通过降低 w_2 来获得更高的 PSNR 分数,但这可能以牺牲结构和纹理细节为代价.

2.4　多级半色调

本研究中介绍的框架虽然以基础的二级半色调问题为出发点,但其灵活性和扩展性使其能够适应更为复杂的图像处理场景,特别是对于多级半色调的拓展.多级半色调技术通过增加灰度层次,为图像每个区域提供了更多的网点选项,从而在视觉上实现了更加平滑和连续色调的模拟.为具体说明此点,本文在此提供了一个多级半色调的具体示例.通过调整模型的输出通道数,并利用 Gumbel-Softmax 技巧,将其输出的 one-hot 向量与预定义的多级半色调等级权重进行加权求和,由于所提出的蓝噪声损失仅针对二值离散化输出进行优化,因此不使用该蓝噪声解决方案,最终得到所需的离散半色调结果.经实验得出五级半色调在 VOC 测试集上的 SSIM 和 PSNR 分数分别为 0.3216 和 42.825.图 8 展示了两个五级半色调的实例,可以在折线图中观察到模型的输出依旧集中在离散的灰度级中心.



(a)input　　(b)multitoning image　　(c)output line chart

图 8　五级半色调实例

Fig.8　Five-level halftoning example

虽然本研究提出的扩展解决方案在多级半色调算法实现上展现出可行性,但为了进一步提升图像质量,必须注意到几个额外的复杂维度,包括选用合适的蓝噪声模型以优化视觉质量,以及有效缓解条纹伪影。为了实现这些目标,需要将这些考量因素融入现有框架中,以期达到更深层次的优化与完善[28].鉴于多维度问题的复杂性,要求更为精细的理论分析与实验验证,这些深入的研究工作可视为未来研究的方向.

3　结 论

本研究提出基于数据驱动方法的高效半色调方

法,引入Gumbel-Softmax重参数化策略,解决了半色调离散选择带来的不可微分问题,实现了有效的无偏梯度估计.设计新的损失函数,在训练阶段显著抑制了恒定灰度图像的各向异性现象,从而促进了蓝噪声特性的自然生成.提出区域置信度聚合机制,综合考虑邻域内像素间的相互关系,增强了网络对局部区域细节和全局结构的信息提取.总的来说,该可微分框架摆脱了对标签数据的依赖,直接针对半色调评价度量进行优化.与其他方法相比,本研究所训练的深度学习模型不仅能够生成富含细节的高质量半色调图像,而且在运算效率方面具有显著优势.此外,该框架具备良好的可扩展性,能够应用于多级半色调等场景中,为图像半色调提供了有效的解决方案.未来工作会进一步优化神经网络结构以及多级半色调和彩色半色调的蓝噪声问题.

参考文献

[1] ULICHNEY R A. Void-and-cluster method for dither array generation [C]//Human Vision, Visual Processing, and Digital Display IV. San Jose, CA. SPIE, 1993: 332-343.

[2] MAO Y F, ABELLO L, SARKAR U, et al. 4-row serpentine tone dependent fast error diffusion [C]//2018 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece. IEEE, 2018: 3973-3977.

[3] HU X Y. Simple gradient-based error-diffusion method [J]. Journal of Electronic Imaging, 2016, 25(4): 043029.

[4] OSTROMOUKHOV V. A simple and efficient error-diffusion algorithm [C]//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. ACM, 2001: 567-572.

[5] FRANK T, LIU J Y, GAT S, et al. A machine learning approach to design of aperiodic, clustered-dot halftone screens *via* direct binary search [J]. IEEE Transactions on Image Processing, 2022, 31: 5498-5512.

[6] LIAO J R. Theoretical bounds of direct binary search halftoning [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3478-3487.

[7] UFUK AGAR A, ALLEBACH J P. Model-based color halftoning using direct binary search [J]. IEEE Transactions on Image Processing, 2005, 14(12): 1945-1959.

[8] 王晓红, 刘丽丽, 陈豪, 等. 一种基于动态误差扩散系数的数字半色调算法[J]. 包装工程, 2017, 38(13): 199-203.
WANG X H, LIU L L, CHEN H, et al. A digital halftone algorithm based on the dynamic error diffusion coefficient [J]. Packaging Engineering, 2017, 38(13): 199-203. (in Chinese)

[9] FUNG Y H, CHAN Y H. Tone-dependent noise model for high-quality halftones [J]. Journal of Electronic Imaging, 2013, 22(2): 023004.

[10] SHAO L H, ZHANG E H, LI M. An efficient convolutional neural network model combined with attention mechanism for inverse halftoning [J]. Electronics, 2021, 10(13): 1574.

[11] JIANG H T, XIONG D L, JIANG X W, et al. Halftoning with multi-agent deep reinforcement learning [C]//2022 IEEE International Conference on Image Processing (ICIP). Bordeaux, France. IEEE, 2022: 641-645.

[12] LAU C K, XIA M H, WONG T T. Taming reversible halftoning *via* predictive luminance [J]. IEEE Transactions on Visualization and Computer Graphics, 2024, 30(8): 4841-4852.

[13] GUO J M, SANKARASRINIVASAN S. H-GAN: deep learning model for halftoning and its reconstruction [C]//2020 IEEE International Conference on Consumer Electronics (ICCE). Las Vegas, NV, USA. IEEE, 2020: 1-2.

[14] CHOI B, ALLEBACH J P. Mimicking DBS halftoning *via* a deep learning approach [J]. Electronic Imaging, 2022, 34(15): 158-1-158-7.

[15] LAU D L, ARCE G R. Modern digital halftoning [M]. Boca Raton: CRC Press, 2018.

[16] TITSIAS M K, LÁZARO-GREDILLA M. Local expectation gradients for black box variational inference [J]. Advances in Neural Information Processing Systems, 2015: 1-9.

[17] XIA M H, HU W B, LIU X T, et al. Deep halftoning with reversible binary pattern [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 13980-13989.

[18] JIANG H T, XIONG D L, JIANG X W, et al. Efficient halftoning *via* deep reinforcement learning [J]. IEEE Transactions on Image Processing, 2023, 32: 5494-5508.

[19] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778.

[20] 李硕士, 刘洪瑞, 甘永东, 等. 基于残差密集块与注意力机制的图像去雾网络[J]. 湖南大学学报(自然科学版), 2021, 48(6): 112-118.
LI S S, LIU H R, GAN Y D, et al. Image dehazing network based on residual dense block and attention mechanism [J]. Journal of Hunan University (Natural Sciences), 2021, 48(6): 112-118. (in Chinese)

[21] Unsupervised multi-object segmentation using attention and soft-argmax [C]//2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA. IEEE, 2023: 3266-3275.

[22] HUIJBEN I A M, KOOL W, PAULUS M B, et al. A review of the gumbel-max trick and its extensions for discrete stochasticity in machine learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(2): 1353-1371.

[23] KIM S H, ALLEBACH J P. Impact of HVS models on model-based halftoning [J]. IEEE Transactions on Image Processing, 2002, 11(3): 258-269.

[24] ULICHNEY R A. Dithering with blue noise [J]. Proceedings of the IEEE, 1988, 76(1): 56-79.

[25] ITOUA P, BEGHDADI A, VIARIS DE LESEGNO P. Objective perceptual evaluation of halftoning using image quality metrics [C]//10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010). Kuala Lumpur, Malaysia. IEEE, 2010: 456-459.

[26] FURUTA R, INOUE N, YAMASAKI T. PixelRL: fully convolutional network with reinforcement learning for image processing [J]. IEEE Transactions on Multimedia, 2020, 22(7): 1704-1719.

[27] VASSILO K, HEATWOLE C, TAHA T, et al. Multi-step reinforcement learning for single image super-resolution [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, WA, USA. IEEE, 2020: 2160-2168.

[28] ABEDINI F, GOORAN S. Structure-aware color halftoning with adaptive sharpness control [J]. Journal of Imaging Science and Technology, 2022, 66(6): 1-11.