

基于双调制策略的真实感风格迁移方法

庾晨龙¹, 杨天², 唐贝贝³

1. 马鞍山学院 大数据与人工智能学院, 安徽 马鞍山 243100
2. 安徽理工大学 计算机科学与工程学院, 安徽 淮南 232001
3. 安徽理工大学 人工智能学院, 安徽 淮南 232001

摘要:目的 针对现有真实感风格迁移方法因缺乏有效的风格表示而导致内容失真和风格化强度不足的问题, 提出了一种名为 DMPST (Dual Modulation Photorealistic Style Transfer) 的真实感风格迁移方法。方法 为了有效地捕捉风格信息, DMPST 利用了浅层和深层风格信息互补的优势, 对多个尺度的风格特征进行渐进式融合。然后, 通过双调制策略将融合后的风格特征变换为两组风格信号, 并分别对内容特征进行特征调制和滤波调制, 从而实现全局性和局部性的风格化指导。此外, 为弥补风格化强度提高带来的内容信息丢失, 设计了自适应空间插值模块, 在解码阶段进行多尺度的内容细节修复。结果 通过定性和定量实验验证, DMPST 能够生成内容信息保存良好和高度风格化的结果, 在 DPST 数据集上, Content loss、Style loss、SSIM、PSNR 和 LPIPS 指标分别达到了 0.595、0.405、0.725、16.894 和 0.758。结论 所提出的方法在保持源图像内容细节的前提下, 实现了高度风格化的迁移效果, 对后续提高真实感风格化强度的研究具有参考价值。

关键词:真实感风格迁移; 双调制策略; 渐进式融合; 自适应空间插值

中图分类号:TP391.4 **文献标识码:**A **doi:**10.16055/j.issn.1672-058X.2026.0001.003

Photorealistic Style Transfer Method Based on Dual Modulation Strategy

YU Chenlong¹, YANG Tian², TANG Beibei³

1. School of Big Data and Artificial Intelligence, Ma' anshan University, Ma' anshan 243100, Anhui, China
2. School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001, Anhui, China
3. School of Artificial Intelligence, Anhui University of Science and Technology, Huainan 232001, Anhui, China

Abstract: Objective Aiming at the problems of content distortion and insufficient stylization intensity caused by the lack of effective style representation in existing photorealistic style transfer methods, a photorealistic style transfer method named dual modulation photorealistic style transfer (DMPST) is proposed. **Methods** To efficiently capture the style information, DMPST utilized the complementarity of the shallow and deep style information to progressively fuse stylistic features at multiple scales. Then, through a dual-modulation strategy, the fused style features were transformed into two sets of style signals, which were used for feature modulation and filter modulation of content features respectively, so as to achieve global and local stylization guidance. In addition, to make up for the loss of content information due to the increase in stylization intensity, an adaptive spatial interpolation module was designed to repair multi-scale content details

收稿日期:2024-03-05 **修回日期:**2024-06-20 **文章编号:**1672-058X(2026)01-0020-08

基金项目:安徽省自然科学基金(2008085MF220)资助。

作者简介:庾晨龙(2000—),男,安徽马鞍山人,硕士研究生,从事计算机视觉研究。Email:chenlong_yu@163.com.

引用格式:庾晨龙,杨天,唐贝贝.基于双调制策略的真实感风格迁移方法[J].重庆工商大学学报(自然科学版),2026,43(1):20-27.

YU Chenlong, YANG Tian, TANG Beibei. Photorealistic style transfer method based on dual modulation strategy[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(1): 20-27.

during the decoding stage. **Results** Qualitative and quantitative experiments verified that DMPST was able to generate results with well-preserved content information and a high degree of stylization. On the DPST dataset, the Content loss, Style loss, SSIM, PSNR, and LPIPS indicators achieved 0.595, 0.405, 0.725, 16.894, and 0.758, respectively.

Conclusion The proposed method achieves a highly stylized transfer effect while maintaining the content details of the source image, and it has reference value for subsequent research on improving the intensity of photorealistic stylization.

Keywords: photorealistic style transfer; dual-modulation strategy; progressive fusion; adaptive spatial interpolation

真实感风格迁移是计算机视觉领域中的一个重要研究课题,旨在真实图像之间实现风格的迁移并保持真实感。与艺术风格迁移不同,真实感风格迁移要求保持源图像的内容结构、纹理和笔触等,仅迁移色彩,因此真实感风格迁移可以看作采用深度学习技术的色彩迁移^[1]。

传统的色彩迁移方法^[2-3]根据风格图像的低级特征统计信息对源图像进行色彩迁移,由于算法缺少对深层信息的利用,风格化结果中容易出现色彩不一致和风格化强度不足等问题。近年来,艺术风格迁移利用深度学习技术得到了快速发展^[4-5], Gatys 等^[6]将艺术风格迁移看作一种纹理和色彩的迁移,提出了基于优化的风格迁移方法,利用深度卷积神经网络进行艺术风格迁移,开启了图像风格迁移的新纪元。然而,由于艺术风格迁移方法具有强大的纹理和笔触渲染能力,其风格化结果的内容会遭受一定程度的更改,如笔触、纹理、物体轮廓等,导致风格化结果缺少真实感。因此,艺术风格迁移无法直接用于真实图像风格迁移。

研究者在艺术风格迁移方法的基础上,通过增强内容细节保留和添加约束实现真实感风格迁移。Luan 等^[7]在基于优化的艺术风格迁移方法^[6]基础上,通过添加图像真实感正则化将迁移过程约束在色彩空间的局部仿射变换中,实现真实感风格迁移。然而,优化过程需要耗费大量的时间和计算资源。为加速真实感风格化的过程, Li 等^[8]提出了 PhotoWCT, 在基于前馈网络的艺术风格迁移方法 WCT^[9]的基础上使用反池化操作代替插值上采样,减少了解码过程中的细节损失,但由于由粗到细的特征变换实现了较强的风格化,需要内容图像和风格图像的语义掩模进行局部风格化指导,否则容易出现内容扭曲和色彩不一致现象。Yoo 等^[10]在 WCT 的基础上提出了 WCT², 使用小波矫正传输和多尺度风格化由细到粗地进行特征变换,保留了内容的精细细节,但风格化强度较低。Chiu 等^[11]结合

了 PhotoWCT 和 WCT² 的优点提出了 PhotoWCT², 引入高频残差连接来保留高频细节,并使用由粗到细的特征变换以保持风格化强度,但风格化结果中出现伪影和色彩不一致问题。

综上所述,目前的真实感风格迁移算法^[7-8,10-11]已经取得了较好的迁移效果,但在图像的风格化强度上仍存在局限性。由于网络缺乏对风格的有效表示,风格化结果往往无法在内容信息和风格信息之间取得平衡,容易出现内容失真和风格化强度低的问题。尽管在迁移过程中利用内容图像和风格图像的语义掩模指导区域风格化能够缓解这些问题^[8,12],但获取一个精确的语义掩模是困难的,并且使用语义掩模进行风格迁移增加了大量的计算开销,因此其实际应用受到限制。

针对以上问题,本文提出了一种基于双调制策略的真实感风格迁移方法 DMPST。该方法利用轻量级 Vision Transformer 编码器提取多尺度的特征,并输入到渐进式多尺度风格融合模块中,以融合浅层和深层的风格特征。通过调制器对融合后的风格特征进行处理,得到两组风格信号,并对最深层内容特征进行特征调制和滤波调制,从而获得风格化特征。最终,解码器对得到的风格化特征进行解码,并在解码阶段通过自适应空间插值模块对内容细节进行修复。实验表明:该方法在保留真实感的前提下,实现了风格化强度较高的迁移效果。

1 本文算法

1.1 网络概览

DMPST 网络主要由 6 个部分组成:内容编码器(Content Encoder, CE)、风格编码器(Style Encoder, SE)、渐进式多尺度风格融合模块(Progressive Multi-scale Style Fusion Module, PMSFM)、风格信号调制器(Modulator)、解码器(Decoder)和自适应空间插值模块(Adaptive Spatial Interpolation Module, AdaSIM)。网络主干为对称的编码器-解码器架构,如图 1 所示。

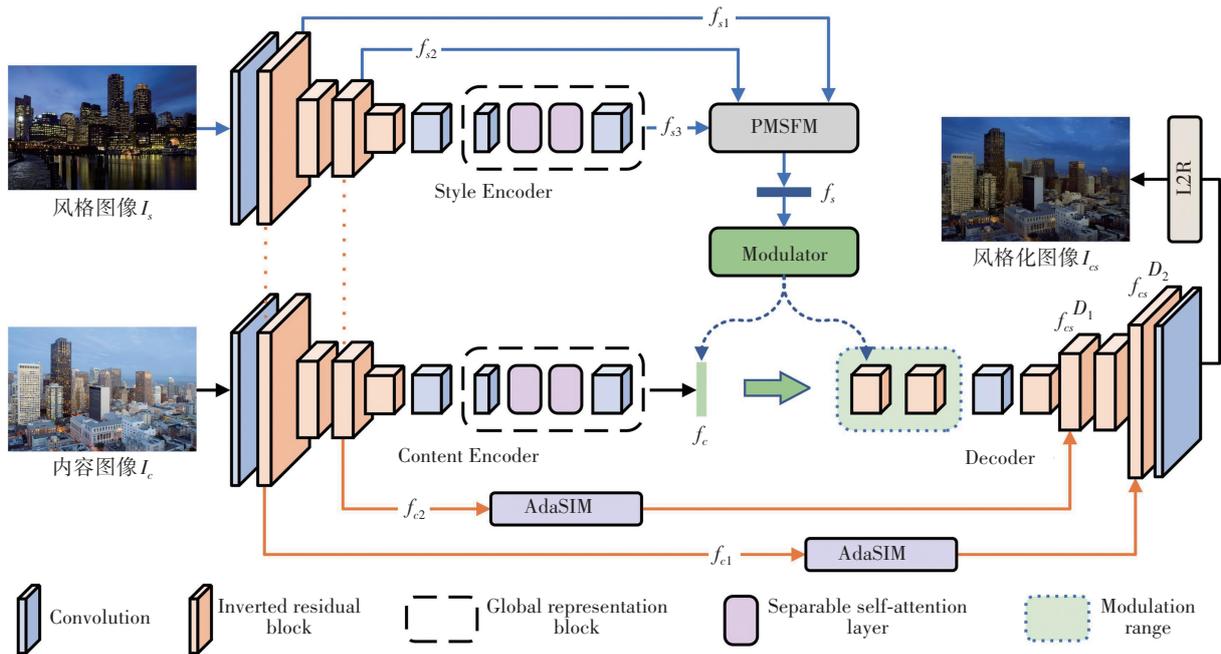


图 1 整体框架图

Fig. 1 Overall framework diagram

DMPST 的风格化过程遵从图像风格迁移中常用的“编码-迁移-解码”流程。输入内容图像 $I_c \in \mathbf{R}^{3 \times H \times W}$ 、风格图像 $I_s \in \mathbf{R}^{3 \times H \times W}$ (H 和 W 分别表示高和宽), 分别经过内容编码器和风格编码器得到多个尺度的内容特征集合 $\{f_{c1}, f_{c2}, f_c\}$ 和风格特征集合 $\{f_{s1}, f_{s2}, f_{s3}\}$, PMSFM 模块对 f_{s1} , f_{s2} 和 f_{s3} 进行融合得到 f_s , 利用调制器得到风格特征信号, 并注入内容特征 f_c 中得到风格化特征 f_{cs} 。然后将风格化特征 f_{cs} 输入到解码器中, AdaSIM 利用 f_{c1} 和 f_{c2} 在解码过程中进行自适应的多尺度内容细节修复, 最终由 L2R (Latent to RGB) 将隐空间的特征映射到 RGB 空间, 输出风格化图像 $I_{cs} \in \mathbf{R}^{3 \times H \times W}$ 。

1.2 渐进式多尺度风格融合模块

风格编码器在各尺度上提取到的风格特征包含不

同层次的风格信息, 聚合多个尺度的风格特征能够增强网络对风格的表达。在艺术风格迁移中, Gatys 等^[6]认为特征图的通道内积可以表达风格, 因此 Deng 等^[13]利用通道注意力机制增强了网络对风格的表达, 但没有利用到低层的风格特征, 风格化效果提升较小。因此, 针对真实感图像风格迁移, 提出了渐进式多尺度风格融合模块 PMSFM, 利用通道注意力机制进行多尺度的风格特征融合, 增强网络对风格的表达。

PMSFM 由平均池化层 (Average Pooling, AP)、 1×1 卷积层 (Conv1)、多层感知机层 (MLP) 和残差连接 (Residual Connection) 构成, 利用通道注意力机制将低层风格信息与高层风格信息融合, 如图 2 所示。

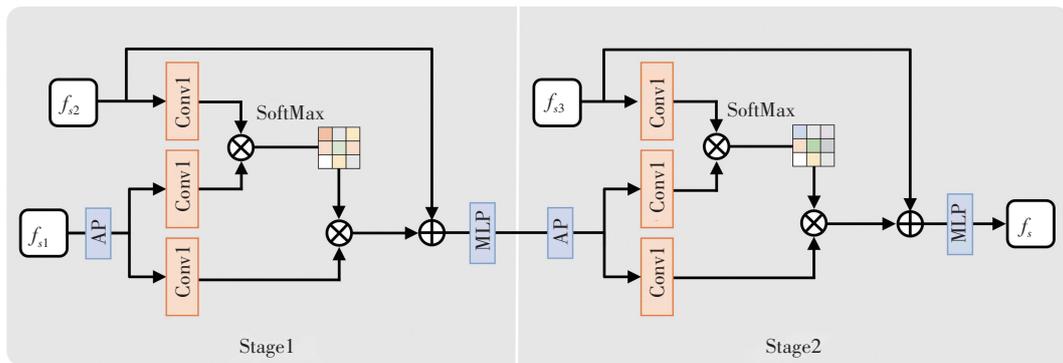


图 2 渐进式多尺度风格融合模块结构图

Fig. 2 Structure diagram of progressive multi-scale style fusion module

PMSFM 将融合分为两个阶段,首先,取出风格编码器中 3 个尺度的风格特征分别记为 $f_{s1} \in \mathbf{R}^{128 \times H/2 \times W/2}$, $f_{s2} \in \mathbf{R}^{256 \times H/4 \times W/4}$, $f_{s3} \in \mathbf{R}^{512 \times H/8 \times W/8}$, 第一阶段 (Stage1) 融合 f_{s1} 与 f_{s2} 得到 $f_{\text{imp}} \in \mathbf{R}^{256 \times H/4 \times W/4}$, 第二阶段 (Stage2) 融合 f_{imp} 和 f_{s3} 得到 $f_s \in \mathbf{R}^{512 \times H/8 \times W/8}$, 具体过程如式 (1)、式 (2) 所示。

(1) 融合 f_{s1} 和 f_{s2} :

$$\begin{aligned} Q_1 &= \text{Conv1}(f_{s2}) \\ K_1 &= \text{Conv1}(\text{AP}(f_{s1})) \\ V_1 &= \text{Conv1}(\text{AP}(f_{s1})) \\ A_{s1} &= \text{SoftMax}(Q_1 \otimes K_1^T) \\ f_{\text{imp}} &= \text{MLP}(A_{s1} \otimes V_1 + f_{s2}) \end{aligned} \quad (1)$$

其中, $Q_1 \in \mathbf{R}^{256 \times H/4 \times W/4}$, $K_1 \in \mathbf{R}^{128 \times H/4 \times W/4}$, $V_1 \in \mathbf{R}^{128 \times H/4 \times W/4}$ 。

(2) 融合 f_{s3} 和 f_{imp}

$$\begin{aligned} Q_2 &= \text{Conv1}(f_{s3}) \\ K_2 &= \text{Conv1}(\text{AP}(f_{\text{imp}})) \\ V_2 &= \text{Conv1}(\text{AP}(f_{\text{imp}})) \\ A_{s2} &= \text{SoftMax}(Q_2 \otimes K_2^T) \\ f_s &= \text{MLP}(A_{s2} \otimes V_2 + f_{s3}) \end{aligned} \quad (2)$$

其中, $Q_2 \in \mathbf{R}^{512 \times H/4 \times W/4}$, $K_2 \in \mathbf{R}^{256 \times H/8 \times W/8}$, $V_2 \in \mathbf{R}^{256 \times H/8 \times W/8}$ 。

PMSFM 考虑了多尺度风格特征包含的浅层和深层风格信息,通过两阶段的渐进式融合增强了网络对风格的表达。

1.3 双调制策略

真实感风格迁移方法通常使用 WCT 算法进行风格迁移^[8,10-11]。WCT 本质上是对齐内容特征与风格特征的二阶统计信息,属于全局性的特征调制 (FeatMod)。然而,全局风格化程度的增强容易破坏内容结构,现有的方法为了提高真实感而牺牲了风格化强度,导致风格化效果不佳。尽管使用语义掩模限制区域风格化能够缓解因风格化强度提高而导致的内容结构扭曲问题,但其计算开销较大,且获取精确的语义掩模是困难的。针对现有方法的缺陷,DMPST 将滤波调制作为一种局部风格化指导方法引入真实感风格迁移领域,提高网络对局部内容细节的风格化处理,进一步提高风格化强度。

特征调制主要捕捉色彩、对比度和饱和度等全局属性,而滤波调制 (FilterMod) 擅长捕捉局部属性的变化^[14]。DMPST 结合了特征调制和滤波调制的优势,在迁移过程中能够进行全局和局部风格化的指导,如图 3 所示。

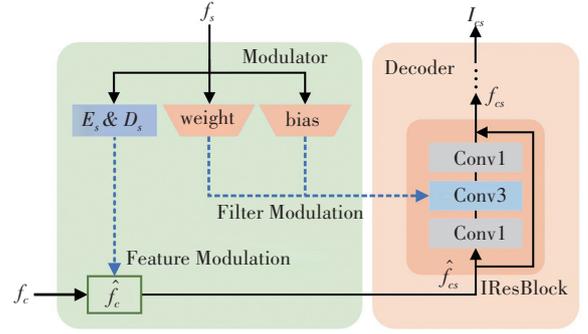


图 3 双调制示意图

Fig. 3 Schematic diagram of dual modulation

(1) 特征调制。WCT 中的白化和着色操作需要计算奇异值分解,若内容和风格特征图的协方差矩阵的奇异值几乎相等,会导致 WCT 模块失效^[15],因此 DMPST 采用了 cWCT^[16]用于特征调制。cWCT 同样包含两个步骤,白化和着色。白化过程可表示为式 (3):

$$\hat{f}_c = E_c D_c^{-\frac{1}{2}} E_c^T f_c \quad (3)$$

其中, \hat{f}_c 为白化结果, D_c 为内容特征值的对角矩阵, E_c 为内容特征向量的正交矩阵。

着色过程可表示为式 (4):

$$\hat{f}_{cs} = E_s D_s^{\frac{1}{2}} E_s^T \hat{f}_c \quad (4)$$

其中, \hat{f}_{cs} 为着色结果, D_s 为风格特征值的对角矩阵, E_s 为风格特征向量的正交矩阵。

(2) 滤波调制。特征调制能够进行全局风格化,但从全局角度来增强风格化强度会破坏内容完整性,导致风格化的真实感降低。为弥补全局风格迁移的缺陷,引入了滤波调制方法进行局部风格化指导,如图 3 所示。通过权重网络和偏置网络将风格特征转为权重 w_s 和偏置项 b_s ,然后将这些参数注入解码器 D 中调制特征。为了简化滤波调制,DMPST 采用等价的伪特征调制方式^[17],如式 (5) 所示:

$$\begin{aligned} \text{Conv}(\hat{f}_{cs}, (w_s, b_s)) &= (w_s * F + b_s) \otimes \hat{f}_{cs} = \\ &= (w_s * F) \otimes \hat{f}_{cs} + b_s \otimes \hat{f}_{cs} = \\ &= w_s * (F \otimes \hat{f}_{cs}) + b_s * \hat{f}_{cs} \end{aligned} \quad (5)$$

其中, F 表示卷积滤波器, $*$ 表示元素式乘法, \otimes 表示卷积。

双调制包含特征调制和滤波调制,利用全局风格化和局部风格化互补提高风格化强度,可表示为式 (6):

$$\begin{aligned} m_s &= (E_s, D_s, w_s, b_s) \\ f_{cs} &= \text{DualMod}(D, f_c, m_s) = \\ &= \text{FeatMod}(f_c, (E_s, D_s)) + \text{FilterMod}(D, (w_s, b_s)) \end{aligned} \quad (6)$$

其中, m_s 表示通过 Modulator 从风格特征 f_s 中获取的风格调制信号, f_{cs} 表示经过双调制得到的风格化特征。

1.4 自适应空间插值模块

由于浅层网络感受野较小,几何细节信息表征能力强,深层网络的感受野较大,语义信息表征能力强,在艺术风格迁移中通常使用深层内容特征进行风格化,达到了纹理、笔触、色彩等级别的迁移。真实图像风格迁移旨在保留所有内容细节以增强风格化的真实感,而风格化强度的提高会在一定程度上造成内容信息的丢失,导致解码过程中出现内容细节丢失的问题。因此,DMPST 使用 AdaSIM 将浅层内容特征作为内容细节修复源,从而修复多尺度的内容细节,以提高风格化结果的真实感,具体结构如图 4 所示。

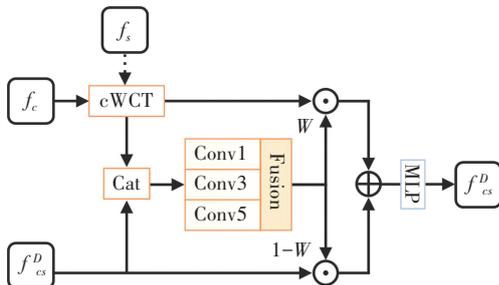


图 4 自适应空间插值模块结构图

Fig. 4 Structure diagram of adaptive spatial interpolation module

为了保证空间插值操作在同一特征空间中进行,首先利用 cWCT 将对应层的内容特征转换至风格特征空间,与对应尺度的解码层特征 f_{cs}^D (图 1 中的 f_{cs}^{D1} 或 f_{cs}^{D2}) 进行 Concat 操作,然后利用插值计算层识别相应特征之间的局部差异,得出合适的插值强度 W ,计算过程如式(7)所示:

$$W = \frac{1}{n} \sum_{i=1}^n \text{Sigmoid}(\psi_i(\text{Concat}(f_g, f_{cs}^D))) \quad (7)$$

其中, f_g 为 f_c 经过 cWCT 迁移得到的特征, ψ_i 表示插值计算层中的第 i 个卷积层, $\text{Concat}(\cdot, \cdot)$ 表示沿通道维度拼接两个特征。

最终, AdaSIM 利用得到的自适应权重 W 对解码层特征 f_{cs}^D 的内容细节进行自适应地修复,如式(8)所示。

$$f_{cs}^D = W * f_g + (1-W) * f_{cs}^D \quad (8)$$

其中, $*$ 表示元素式乘法。

风格化程度的增强会导致内容细节受损,自适应空间插值模块将内容细节保存良好的低层特征与解码层特征进行适应性融合。因此,在解码阶段利用自适应空间插值模块修复内容细节,能够保留更多内容信息,提高风格化的真实感。

2 损失函数

在训练阶段使用预训练的 VGG-19 作为损失网络。内容感知损失 L_c 度量风格化图像 I_{cs} 和内容图像 I_c 之间的内容差异。风格感知损失 L_s 度量风格化图像 I_{cs} 和风格图像 I_s 之间的风格差异。循环一致性损失

L_{cyc} 用于约束网络生成图像 I_{cs} 与内容图像 I_c 的内容一致性。

2.1 内容感知损失函数

内容感知损失定义如式(9),采用 VGG-19 的 ReLU4_1 作为内容特征层集合 C 。

$$L_c = \sum_{i=1}^{N_c} \|\phi_i(I_{cs}) - \phi_i(I_c)\|_2^2 \quad (9)$$

其中, N_c 表示内容特征层的数量, $\phi_i(\cdot)$ 表示 C 中第 i 个内容特征层的输出。

2.2 风格感知损失函数

风格感知损失定义如式(10),采用 VGG-19 的 ReLU1_1、ReLU2_1、ReLU3_1 和 ReLU4_1 作为风格特征层集合 S 。

$$L_s = \sum_{i=1}^{N_s} \|\mu(\phi_i(I_{cs})) - \mu(\phi_i(I_s))\|_2^2 + \|\sigma(\phi_i(I_{cs})) - \sigma(\phi_i(I_s))\|_2^2 \quad (10)$$

其中, N_s 表示风格特征层的数量, $\mu(\cdot)$ 和 $\sigma(\cdot)$ 分别表示特征层的均值和标准差。

2.3 循环一致性损失函数

输入风格化图像 I_{cs} 和内容图像 I_c 到内容编码器和风格编码器中,将内容图像 I_c 的风格迁移到风格化图像 I_{cs} 中,从而得到重建的内容图像 I_{csc} 。循环一致性损失定义如式(11)所示,通过惩罚 I_{csc} 与 I_c 之间的差异,使风格化图像 I_{cs} 在经过逆向风格化之后接近原始的内容图像 I_c ,从而提高 DMPST 生成图像的质量和内容一致性。

$$L_{cyc} = \|I_{csc} - I_c\|_1 \quad (11)$$

2.4 总损失函数

总损失表示为式(12),其中 λ_c 、 λ_s 和 λ_{cyc} 为权重超参数,分别设为 1、1 和 10。

$$L_{total} = \lambda_c L_c + \lambda_s L_s + \lambda_{cyc} L_{cyc} \quad (12)$$

3 仿真实验与结果分析

3.1 实验参数设置

采用 MS-COCO 数据集^[18]作为训练所需的内容数据集和风格数据集,并在 DPST^[7]提供的数据集上进行测试。其中训练数据集包含 82 783 张真实生活中的图片,在训练阶段将图像大小调整为 512×512,并随机剪裁至 256×256。测试数据集包含 60 组摄影照片,每组图像由内容图像和风格图像组成。采用 Adam 优化器,其中 $\beta_1 = 0.9$, $\beta_2 = 0.999$,初始学习率(lr)设置为 0.000 1,数据批尺寸(batch size)设置为 2。

3.2 定性结果与分析

为验证本文方法的有效性,将 DMPST 与 6 种方法进行对比,不同方法风格化结果如图 5 所示。

PhotoWCT^[8]的风格化结果中容易出现伪影和色彩不一致问题,使用后处理技术来平滑图像能够有效缓

解,但会造成图像细节丢失严重。尽管 WCT^[10] 利用小波变换缓解了内容信息丢失问题,但在部分结果中风格化效果不明显。SCCL^[19] 使用风格语料库约束不同样本间的风格一致性,提高了生成图像的真实感,但风格化强度较低。MAST^[20] 利用流行对齐进行风格迁移,提高了内容结构保存,但风格化结果中出现内容细节模糊和色彩不一致问题。PhotoWCT^[11] 利用高频残差

连接保留了更多的内容细节,但生成的图像中出现色彩不和谐现象。RAST^[21] 利用循环一致性约束提高了风格化结果中内容整体结构的保存,但其细节丢失严重,且存在色彩不一致问题。DMPST 增强了网络对风格的表达,并且在解码过程中对低层内容细节进行了自适应修复,生成结果在保留真实感的前提下,提高了风格化强度,风格化视觉效果较好。

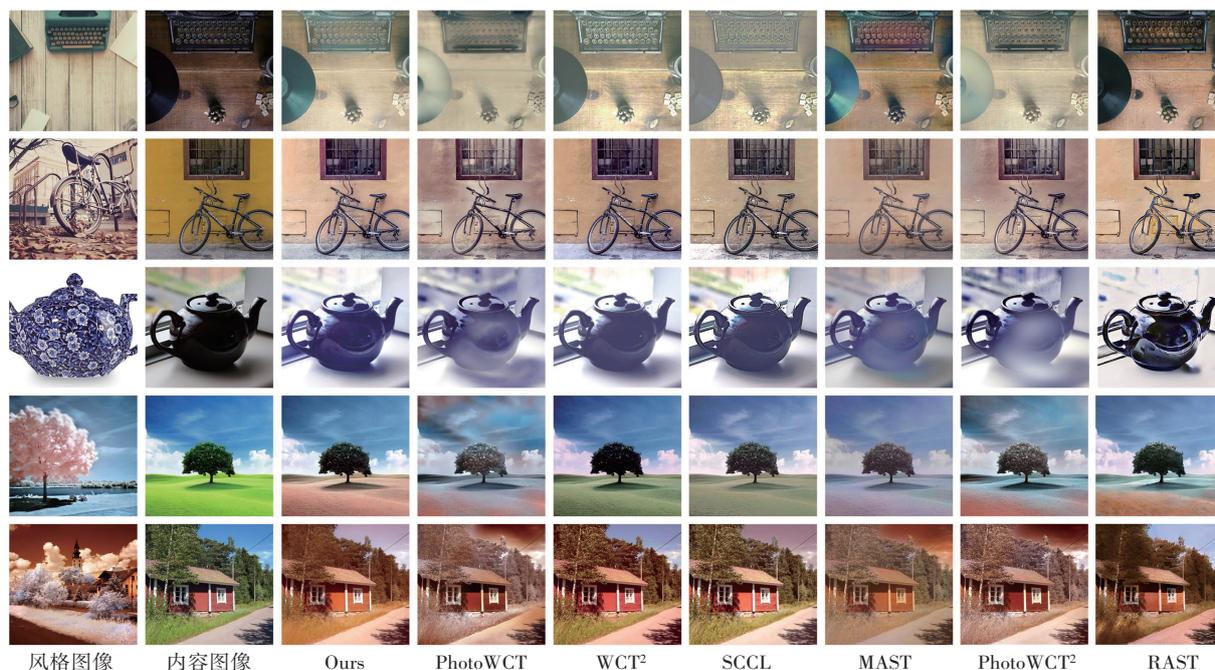


图 5 不同方法风格迁移结果图

Fig. 5 Stylization results of different methods

3.3 定量结果与分析

采用内容损失 (Content loss)、风格损失 (Style loss)、结构相似性 (Structure Similarity Index Measure, SSIM)、峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和可学习感知图像块相似度 (Learned Perceptual Image Patch Similarity, LPIPS) 作为指标评估网络风格化效果。其中,Content loss 用于衡量风格化图像和内容图像的内容特征差异,数值越小则说明内容特征越相似。Style loss 用于衡量风格化图像与风格图像之间的风格特征差异,数值越小则表明风格特征越相似。SSIM 用于衡量内容图像和风格化结果之间亮度、对比度和结

构上的相似性,数值越大则相似性越高。PSNR 用于衡量风格化图像与内容图像之间的图像质量差异,数值越大表明图像质量越接近于内容图像。LPIPS 用于衡量风格化图像与内容图像之间的内容保真度,数值越小则表明风格化图像的真实感越接近内容图像。

为兼容部分方法对图像输入尺寸的要求,将测试集中的图像统一调整至 512×512,然后计算各个方法风格化结果的指标。与其他方法相比,DMPST 在 DPST 数据集上的 Style loss 和 PSNR 达到最优,LPIPS 达到次优,表明风格化强度和图像质量最高,且内容保真度较好,如表 1 所示。

表 1 视觉质量的定量指标对比

Table 1 Comparison of quantitative indicators for visual quality

	Ours	PhotoWCT ^[8]	WCT ^[10]	SCCL ^[19]	MAST ^[20]	PhotoWCT ^[11]	RAST ^[21]
Content loss ↓	0.595	0.696	0.498	0.648	0.548	<u>0.550</u>	0.844
Style loss ↓	0.405	0.447	0.588	0.489	0.772	<u>0.428</u>	0.469
SSIM ↑	0.725	0.692	<u>0.745</u>	0.732	0.777	0.736	0.614
PSNR ↑	16.894	12.659	16.331	15.452	15.327	12.490	<u>16.543</u>
LPIPS ↓	<u>0.758</u>	0.784	0.759	0.756	0.761	0.768	0.771

注: ↑表示高指标更优, ↓表示低指标更优,加粗字体表示各列最优结果,加下划线表示次优结果

由于内容特征和风格特征无法完全解耦,在迁移风格或提高风格化强度时,内容信息丢失是不可避免的。因此,从数值方面来看,DMPST 的 Content loss 和 SSIM 指标处于可接受范围内。综上,DMPST 在风格化强度和风格化图像质量上优于其他方法,在保持图像真实感的前提下,提高了风格化强度。

3.4 消融实验

为了验证 PMSFM、FilterMod 和 AdaSIM 模块的有效性,分别将 3 个模块移除,然后重新训练网络,在 DPST 数据集上进行性能评估。网络去除各模块后的定量指标对比情况如表 2 所示。

表 2 消融实验定量指标对比

Table 2 Comparison of quantitative indicators for ablation experiments

	Full Model	w/o PMSFM	w/o FilterMod	w/o AdaSIM
Content loss ↓	0.595	0.525	0.668	0.886
Style loss ↓	0.405	0.414	0.411	<u>0.407</u>
SSIM ↑	<u>0.725</u>	0.742	0.711	0.680
PSNR ↑	16.894	<u>16.27</u>	14.699	15.631
LPIPS ↓	0.758	<u>0.760</u>	0.767	0.782

DMPST 去除 PMSFM 模块之后,风格表达能力变弱,无法充分利用参考图像的浅层风格信息,导致风格化结果仍保留一些内容图像原有的色调风格且风格化强度较低,因此保留的内容信息更多,Content loss 和 SSIM 较优,迁移效果如图 6(b)所示。若网络仅采用特征调制,迁移过程由于缺少局部风格化的指导,偏向于全局性的风格化增强,导致内容结构遭到破坏,进一步影响迁移效果,各项指标都变差,迁移效果如图 6(c)所示。若移除 AdaSIM 模块,由于风格化强度的提高在一定程度上会破坏内容信息的完整性,风格化结果中出现明显的细节丢失和伪影问题,且各项指标都变差,迁移效果如图 6(d)所示。

综上,PMSFM 可以帮助网络提取更丰富的风格信息,使风格化图像的色调风格更和谐。FilterMod 在迁移过程中利用 PMSFM 提供的风格信息进行局部风格化指导,从而在保留内容信息的前提下,增强了风格化强度。AdaSIM 作为解码阶段的内容细节修复工具,弥补了风格化强度提高带来的内容细节丢失问题。三者的协调运作提高了 DMPST 的风格化强度和生成图像的质量。

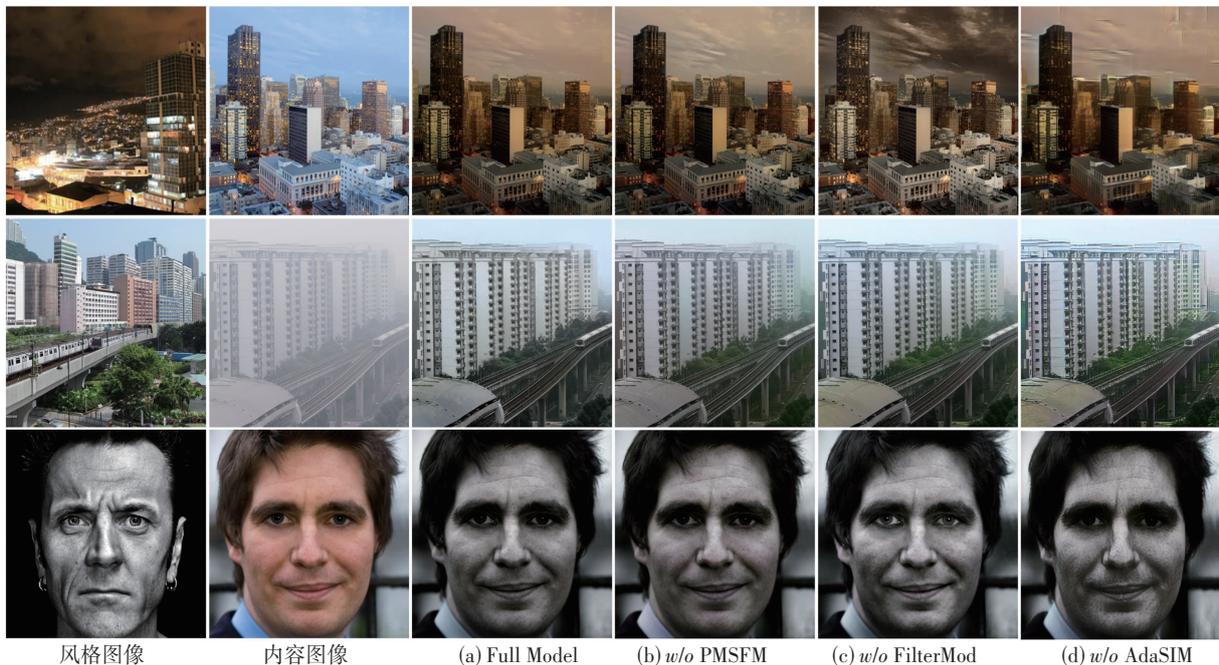


图 6 消融实验效果图

Fig. 6 Visual results of ablation experiments

4 结论

本文提出了一种基于双调制策略的真实感风格迁移算法 DMPST。该方法利用渐进式多尺度风格融合模块整合浅层和深层特征包含的风格信息,然后注入调

制器中获得更复杂、灵活的风格信号,并采用双调制策略对内容特征进行具有全局性和局部性指导的风格化。此外,还设计了一种自适应空间插值模块,用于解码过程中多尺度内容细节的修复,以降低风格化程度

的增强带来的内容细节丢失问题。实验表明:DMPST在保证图像真实感的前提下,提高了风格化强度,得到优于其他主流方法风格化结果。

参考文献(References):

- [1] KE Z, LIU Y, ZHU L, et al. Neural preset for color style transfer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 14173–14182.
- [2] REINHARD E, ADHIKHMIM M, GOOCH B, et al. Color transfer between images[J]. IEEE Computer Graphics and Applications, 2001, 21(5): 34–41.
- [3] WELSH T, ASHIKHMIM M, MUELLER K. Transferring color to greyscale images[C]//Proceedings of the 29th annual conference on Computer graphics and interactive techniques. 2002: 277–280.
- [4] 颜明强, 余鹏飞, 李海燕, 等. 语义风格一致的任意图像风格迁移[J]. 计算机科学, 2023, 50(7): 129–136.
YAN Ming-qiang, YU Peng-fei, LI Hai-yan, et al. Arbitrary image style transfer with consistent semantic style[J]. Computer Science, 2023, 50(7): 129–136.
- [5] 米志鹏. 基于风格特征分布匹配的图像风格迁移[J]. 重庆工商大学学报(自然科学版), 2023, 40(2): 51–56.
MI Zhi-peng. Image style transfer based on the distribution matching of the style features[J]. Journal of Chongqing Technology and Business University(Natural Science Edition), 2023, 40(2): 51–56.
- [6] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 2414–2423.
- [7] LUAN F, PARIS S, SHECHTMAN E, et al. Deep photo style transfer [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 6997–7005.
- [8] LI Y, LIU M Y, LI X, et al. A closed-form solution to photorealistic image stylization[C]//Proceedings of the Computer Vision-ECCV 2018: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part III. 2018: 468–483.
- [9] LI Y, FANG C, YANG J, et al. Universal style transfer via feature transforms[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 385–395.
- [10] YOO J, UH Y, CHUN S, et al. Photorealistic style transfer via wavelet transforms[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 9035–9044.
- [11] CHIU T Y, GURARI D. Photo WCT2: Compact autoencoder for photorealistic style transfer resulting from blockwise training and skip connections of high-frequency residuals[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2022: 2978–2987.
- [12] LI X, LIU S, KAUTZ J, et al. Learning linear transformations for fast image and video style transfer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3804–3812.
- [13] DENG Y, TANG F, DONG W, et al. Arbitrary style transfer via multi-adaptation network [C]//Proceedings of the 28th ACM International Conference on Multimedia. 2020: 2719–2727.
- [14] ALHARBI Y, SMITH N, WONKA P. Latent filter scaling for multimodal unsupervised image-to-image translation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 1458–1466.
- [15] WEN L, GAO C, ZOU C. CAP-VSTNet: Content affinity preserved versatile style transfer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 18300–18309.
- [16] KESSY A, LEWIN A, STRIMMER K. Optimal whitening and decorrelation[J]. The American Statistician, 2018, 72(4): 309–314.
- [17] WANG Z, ZHAO L, ZUO Z, et al. MicroAST: Towards super-fast ultra-resolution arbitrary style transfer[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(3): 2742–2750.
- [18] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//European Conference on Computer Vision. Cham: Springer, 2014: 740–755.
- [19] QIAO Y, CUI J, HUANG F, et al. Efficient style-corpus constrained learning for photorealistic style transfer[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2021, 30: 3154–3166.
- [20] HUO J, JIN S, LI W, et al. Manifold alignment for semantically aligned style transfer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 14841–14849.
- [21] MA Y, ZHAO C, BASU A, et al. RAST: Restorable arbitrary style transfer via multi-restoration[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023: 331–340.

责任编辑:陈 芳