

从全局到局部：双注意力融合去雾网络

杨瑗玮，王华珂，侯兴松

(西安交通大学电子与信息学部，710049，西安)

摘要：为了处理现有的基于卷积神经网络去雾方法只使用单一的注意力、很难生成细节生动的清晰图像，且容易导致色彩失真的问题，提出了一个全局与局部注意力融合的图像去雾方法，以获得正常清晰度和无色彩失真的去雾图像。首先利用通道注意力将输入的有雾图像在通道维度切分为两部分，一部分送入通道像素注意力通道抽取局部特征，另一部分送入 Transformer 通道学习全局特征，然后利用像素注意力对两个通道学习的特征进行融合，将上述模块作为基本单元组合为一个多级 U 型去雾网络，增加残差连接缓解上下采样导致的细节信息丢失，最后在网络底层加入一个 Transformer 模块学习全局信息。在多个公开可用的去雾图像数据集 RESIDE SOTS Indoor、RESIDE SOTS Outdoor 上测试所提方法的有效性，结果表明：对比经典的去雾方法，所提网络生成的图像细节更丰富并且色彩失真最少；在 RESIDE SOTS Outdoor 数据集上，相比经典的 FFA-Net，峰值信噪比提高 1.16 dB，相比 GridDehazeNet，峰值信噪比提高 3.68 dB。提出的全局与局部注意力融合方法能有效地去除雾霾，提升图像的对比度与清晰度，设计的多级 U 型去雾网络和残差连接结构能够缓解细节丢失，提升去雾效果，获得清晰的图像。

关键词：图像去雾；全局与局部注意力融合；通道像素注意力；Transformer 模块

中图分类号：TP391 **文献标志码：**A

DOI：10.7652/xjtuxb202307018 **文章编号：**0253-987X(2023)07-0191-10

From Global to Local: A Dual-Attention Fusion Dehazing Network

YANG Aiwei, WANG Huake, HOU Xingsong

(Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: In terms of the problem that the existing dehazing methods based on the convolution neural network employs attention only from a single perspective, causing difficulties in generating a clear image with vivid details and the propensity to give rise to color distortion, this paper proposes a global and local attention fusion image dehazing network is proposed, in order to obtain a dehazing image with normal definition and no color distortion. The input haze image is first divided into two parts in the channel dimension by using the channel attention. One part is sent into the channel pixel attention channel to extract local features, and the other part is sent into the Transformer channel to learn global features. Then, the pixel attention is used to fuse the features learned by the two channels, and the above modules are combined as basic units into a multi-level U-shaped dehazing network, residual connection is added to alleviate the loss of detail information caused by upper and lower sampling, and finally, a Transformer module is added at the bottom of the network to learn global information. The effectiveness of the method proposed

is tested on several publicly available dehazing image data sets including RESIDE SOTS Indoor and RESIDE SOTS Outdoor. The results show that compared with the classical dehazing method, the image generated by the method proposed is more detailed and has the least color distortion. On the RESIDE SOTS Outdoor data set, the PSNR is 1.16dB higher than that of the classical FFA-Net, and 3.68dB higher than that of the GridDehazeNet. The global and local attention fusion method proposed in this paper can effectively remove haze and improve the contrast and clarity of the image; the designed multi-level U-shaped dehazing network and residual connection structure can alleviate the loss of details and improve the dehazing effect, so that clear images are obtained.

Keywords: image dehazing; local and global information fusion; channel pixel attention; transformer module

现代社会的发展总是伴随着环境的污染,雾霾和大雾天气出现得越来越频繁,很大程度上影响正常的生活。雾霾和大雾天气是由空气中的灰尘、烟雾和水气等小的漂浮颗粒产生的大气现象,这些漂浮的颗粒会吸收或散射大气光,导致拍摄图像质量下降。在雾天影响下,视频监控、自动驾驶等许多实际应用很容易受到影响,目标检测和分割等高级计算机视觉任务精度下降。因此,图像去雾成为一种越来越重要的技术,有很重要的研究价值,同时也是一项充满挑战性的课题。

目前已有的图像去雾方法可以分为传统方法和深度学习方法两大类。传统的去雾方法主要是基于先验知识的,例如 He 等^[1]提出的基于暗通道先验理论的去雾方法,通过大气散射模型复原出无雾图像,虽然该方法取得了显著的去雾效果,但是在恢复包含大片天空区域的图像中会出现失真现象;此后,高涛等^[2]为了解决传统暗通道不适用于大面积天空区域,容易造成去雾图像失真的问题,提出结合暗亮通道先验的远近景融合去雾方法,但该方法在计算速度上有所下降;近年来深度学习算法不断发展,卷积神经网络被大量应用于去雾任务,Cai 等^[3]提出的 DehazeNet 将深度学习首次应用到去雾任务中,利用神经网络对大气退化模型中的透射率进行估计,并提出了一种新的非线性激活函数,提高了恢复图像的质量。杨燕等^[4]提出结合自适应亮度变换不等式估计透射率的方法,将之前恒定的大气散射系数转换为动态的,使其能适应不同图像的场景需求,但是这些基于透射率估计的方法始终存在不可忽视的误差,因此 Li 等^[5]提出的 AOD-Net 首次端到端的直接生成清晰无雾图像,融合不同大小的滤波器形成多尺度特征,通过轻量级的卷积神经网络直接生成清晰图像,这样可以避免估计参数带来的误差相互放大。另外,随着 Hu 等^[6]提出的通道注意力、

Woo 等^[7]提出的卷积注意力模块以及 Transformer 网络^[8]在计算机视觉领域的大放异彩,更多研究人员基于特征注意力网络来设计去雾网络,例如 Qin 等^[9]提出的 FFA-Net 包含局部残差学习和局部特征注意力,局部残差学习允许通过多个局部残差连接绕过薄雾区或低频等不太重要的信息,让主网络结构关注更有效的信息,局部特征注意力自适应学习局部特征权重,赋予重要特征更多的权重。由于仅使用局部特征进行有雾图像的恢复,缺少全局信息的补充,容易引入色彩失真。

针对上述问题,为了恢复有雾图像,并得到无色失真和细节生动的清晰图像,本文对比了不同注意力机制^[10]对网络去雾性能的影响,发现相比于存在单一的局部或全局注意力机制的去雾网络,融合了局部和全局注意力的去雾网络既可以恢复图像边缘和细节信息,又可以减少色彩失真的出现。因此,本文提出一种全局与局部注意力融合的图像去雾方法。本文设计了一个双注意力模块,首先利用通道注意力将输入的有雾图像在通道维度切分为两部分,一部分送入通道像素注意力通道抽取局部特征,另一部分送入 Transformer 通道学习全局特征;然后利用像素注意力对两个通道学习的特征进行融合。利用上述模块作为基本单元,设计了一个多级 U 型去雾网络^[11]。为了提升去雾效果,在多级 U 型去雾网络底层加入一个 Transformer 模块学习全局信息,并增加残差连接^[12]缓解上下采样导致的细节信息丢失。该方法能保证在得到细节生动清晰图像的同时,避免因忽略全局特征而导致色彩失真,相比于其他经典去雾方法,在多个数据集上显著提高恢复图像的主观质量和客观指标。

1 基于双注意力模块的多级 U 型去雾网络

目前在深度学习中广泛使用的注意力机制源于

对人类视觉的研究,在认知科学中,由于信息处理的瓶颈,人们会选择性关注所有信息中的一部分,同时忽略其他可见信息。为了合理利用有限的视觉信息处理资源,人们需要选择视觉区域中的特定部分,然后重点关注它。

注意力机制没有严格的数学定义,例如传统的局部图像特征提取^[13-14]、滑动窗口方法等都可以看作是一种注意力机制。在神经网络中,注意力机制通常是一个额外的神经网络,能够硬性选择输入的某些部分,或者给输入的不同部分分配不同的权重。注意力机制能够从大量的信息中筛选出重要的信息。

在神经网络中引入注意力机制有很多种方法,以卷积神经网络为例,可以在像素维度引入像素注意力机制(PA),也可以在通道维度引入通道注意力机制(CA),当然也有混合维度(CPAM)即通道维度和像素维度同时增加注意力机制。另一方面,以Transformer 网络为例,可以引入全局自注意力^[15],以增加像素之间的关联性。

1.1 通道注意力

通道注意力机制在通道维度增加注意力机制,通过自动学习的方式,即使用另外一个新的神经网络,获取特征图每个通道的重要程度,然后用这个重要程度给每个特征赋予一个权重值,从而让神经网络重点关注某些特征通道。提升对当前任务有用的特征通道,并抑制对当前任务用处不大的特征通道。

本文通过使用全局平均池化^[16]将通道方面的全局空间信息转化为通道描述符 g_c ,然后可以得到通道注意力的模型表述如下

$$A_c = \sigma(\text{Conv}(\delta(\text{Conv}(g_c)))) \quad (1)$$

式中: σ 是 sigmoid 函数; δ 是 ReLu 函数。

通道注意力示意图如图 1 所示,在输入通道注意力机制之前(图 1 左侧第一个 $H \times W \times C_1$ 白色块),特征图的每个通道的重要程度都是一样的,通过通道注意力网络后(图 1 第一个 $H \times W \times C_1$ 彩色块),不同颜色代表不同的权重,每个特征通道的重要性变得不一样了,神经网络重点关注某些权重值大的通道。

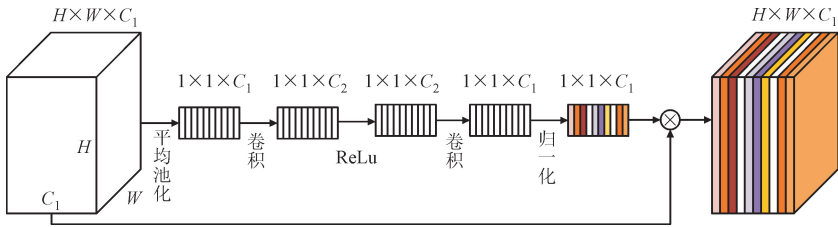


图 1 通道注意力示意图
Fig. 1 Schematic diagram of channel attention

1.2 像素注意力

像素注意力机制在像素维度增加注意力机制,通过自动学习的方式,即使用另外一个新的神经网络,获取到特征图的每个像素的重要程度,然后用这个重要程度去给每个特征赋予一个权重值,从而让神经网络重点关注某些像素。重点关注对当前任务有用的特征图像素,并抑制对当前任务用处不大的特征图像素。

得到像素注意力的模型表述如下

$$A_p = \sigma(\text{Conv}(\delta(\text{Conv}(F)))) \quad (2)$$

式中: F 是输入特征图。

像素注意力示意图如图 2 所示,在输入像素注意力机制之前(图 2 左侧第一个 $H \times W \times C_1$ 白色块),特征图的每个像素的重要程度都是一样的,通过像素注意力网络后(图 2 第一个 $H \times W \times C_2$ 彩色块),不同颜色代表不同的权重,每个像素的重要性变得不一样了,神经网络重点关注某些权重值大的像素。

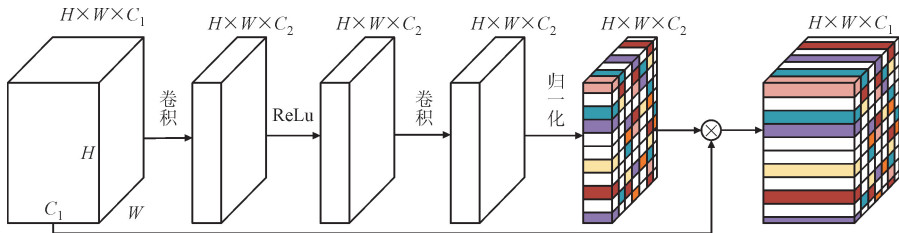


图 2 像素注意力示意图
Fig. 2 Schematic diagram of pixel attention

1.3 通道像素注意力模块

基于对通道注意力和像素注意力机制的分析,本文设计了通道像素注意力模块,同时使用通道注意力和像素注意力,模块结构如图 3 所示。像素注意力可使神经网络更加关注图像中对去雾起决定作用的像素区域而忽略无关紧要的区域,通道注意力则用于处理特征图通道的分配关系,同时对两个维度进行注意力分配,增强了注意力机制对模型性能的提升效果。

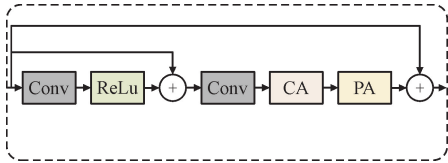


图 3 通道和像素注意力模块

Fig. 3 Channel and pixel attention module

1.4 双注意力模块 DAM

基于对通道像素注意力模块和 Transformer 模块^[17]各自优缺点的研究和分析,本文设计了一个双注意力模块,网络结构如图 4 所示,首先利用通道注意力将输入的有雾图像在通道维度切分为两部分,一部分送入通道像素注意力模块抽取局部特征,另一部分送入 Transformer 通道学习全局特征;然后利用像素注意力对两个通道学习的特征进行融合得到最终的特征图。相比于通常只使用局部注意力或者只使用全局注意力对所有通道进行相同处理的方

法,双注意力分通道处理的方法既保留局部特征信息以恢复图像的细节,又抽取全局信息以减少色彩失真的出现,充分融合两个注意力特征能有效增强去雾性能,减少网络学习难度。

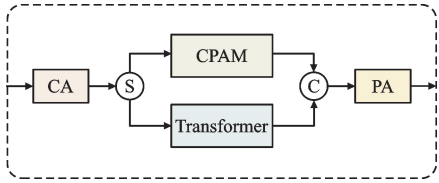


图 4 双注意力模块

Fig. 4 Dual attention module

1.5 双注意力融合去雾网络

利用上述模块作为基本单元,本文设计了一个多级 U 型去雾网络。为了提升去雾效果,在多级 U 型去雾网络底层加入一个 Transformer 模块学习全局信息,并增加残差连接缓解上下采样导致的细节信息丢失。本文去雾网络整体框架如图 5 所示。在每个基本模块中包含两个双注意力网络的串联使用,并将其输出在通道层面进行拼接后再使用通道注意力进一步融合其特征,这样做的好处是能够使网络充分学习到每个双注意力模块提取的全局和局部信息。

相比于仅使用单注意力模块的网络结构来说,使用双注意力模块的多级 U 型网络能更好地聚合全局信息和局部信息,从而在增加少量模型复杂度的基础上有效地提升了网络的去雾能力。

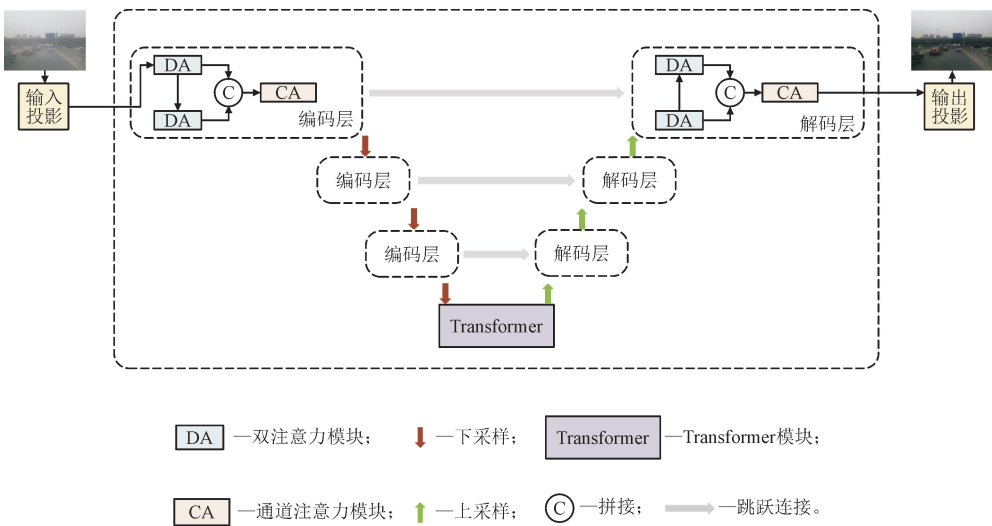


图 5 本文去雾网络整体框架

Fig. 5 The overall framework of the network

2 去雾实验

本文使用 PyTorch 框架构建网络结构,并在 NVIDIA GTX 2060 GPU、Intel(R) Core(TM) i5-8300H CPU 2.30 GHz 计算机上训练模型。下述实验验证了所提网络的优异性能以及泛化能力,通过消融实验证明了所提模块的有效性。

2.1 数据集及训练参数

为了验证所提方法的有效性,首先在 RESIDE ITS 和 RESIDE OTS 数据集^[18]上进行训练,在 RESIDE SOTS Indoor 和 RESIDE SOTS Outdoor 去雾基准测试数据集上进行测试,训练集分别包含 13 990对室内真实场景下的正常/有雾图像和 8 970对室外真实场景下的正常/有雾图像,测试集分别包含 500 对室内真实场景下的正常/有雾图像和 500 对室外真实场景下的正常/有雾图像。

本文在 RGB 通道中训练网络,训练图像被随机切分为 256×256 大小的图像作为网络的输入,并使用随机裁剪、随机反转和随机旋转等方式作为数据增强手段扩充样本量。训练使用 Adam 优化器,其中 β_1 和 β_2 分别设置为默认值 0.9 和 0.999。设置初始学习率为 1×10^{-4} ,采用余弦退火算法^[19]稳定地将学习率调整至 0。

2.2 图像去雾评价指标

为了评估不同图像去雾网络的性能,通常使用峰值信噪比(peak signal-to-noise ratio, R_{sn})^[20]和结构相似性指数(structural similarity index measurement,

S_{sim})^[21]来客观量化对比去雾图像恢复质量。

峰值信噪比的计算公式如下

$$R_{sn} = 20\lg\left(\frac{M_1}{\sqrt{M_{se}}}\right) \tag{3}$$

式中: M_1 是表示图像点颜色的最大数值; M_{se} 为原始图像和处理后图像的均方误差。 M_{se} 的计算公式如下

$$M_{se} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \tag{4}$$

结构相似性的计算公式如下

$$S_{sim}(X,Y) = \frac{(2u_Xu_Y + C_1)(2\sigma_{XY} + C_2)}{(u_X^2 + u_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{5}$$

式中: u_X 和 u_Y 分别表示图像 X 和图像 Y 的均值; σ_X^2 和 σ_Y^2 分别表示图像 X 和 Y 的方差; σ_{XY} 表示图像 X 和 Y 的协方差; C_1 和 C_2 为常数。

将本方法的 R_{sn} 和 S_{sim} 与之前最先进的方法进行比较,比较方法包括 FFA-Net(2019)、U2-Former(2021)^[22]、Uformer(2021)^[23]、GridDehazeNet(2019)^[24]、GMAN(2018)^[25]、AOD-Net(2017)和 Deep DCP(2018)^[26],比较结果见表 1。实验结果表明,本文所提方法在客观指标上以非常大的优势超越了以前的方法。

为了更好地评估提出方法的性能和泛化能力,本文在多个广泛采用的数据集,例如 RESIDE SOTS(synthetic objective testing set) Outdoor、RESIDE SOTS Indoor 数据集上对比了所提方法与主流方法之间的性能差异。

表 1 使用 RESIDE SOTS 室外数据集和室内数据集基于 8 种去雾方法去雾后的客观评价结果
Table 1 Objective evaluation results after dehazing using RESIDE SOTS outdoor dataset and indoor dataset based on 8 dehazing algorithms

方法	R_{sn}/dB		结构相似性	
	室外数据集	室内数据集	室外数据集	室内数据集
FFA-Net	<u>33.38</u>	36.39	<u>0.980 4</u>	<u>0.989 0</u>
U2-Former	31.10	<u>36.42</u>	0.976 0	0.988 0
Uformer	26.52	31.91	0.944 0	0.971 0
GridDehazeNet	30.86	32.16	0.982 0	0.984 0
GMAN	28.19	20.53	0.963 8	0.808 1
AOD-Net	24.14	20.51	0.920 0	0.816 0
Deep DCP	24.08	19.25	0.933 0	0.832 0
本文	34.54	36.85	0.985 4	0.990 8

注:加粗数据表示最优结果;下划线数据表示次优结果。

2.3 视觉效果分析

选取 GridDehazeNet、FFA-Net 以及 GMAN 与本文方法进行对比实验,首先在整理分析 RESIDE SOTS 室外数据集时,从中选取 4 张有代表性的图像进行去雾效果对比。图 6 和图 7 对比两种不同场景的图像去雾效果,图 8 和图 9 对比两种不同浓度的雾图去雾效果。

由图 6 可见:GridDehazeNet 方法的结果天空

区域和马路区域存在黑色块状伪影;FFA-Net 方法存在不均匀的路面黑色区域,但去雾效果良好;GMAN 方法在路面和天空区域出现大量失真现象,整张图像存在块状效应并且在车辆和道路指示牌的边缘出现模糊现象;本文方法的去雾结果最接近无雾原图,在带有光源的近景浓雾区域去雾效果良好,整体颜色明亮,在大面积黑色和白色的区域没有出现上述方法的颜色失真现象。



图 6 强光源有雾图像去雾效果对比
Fig. 6 Strong light source area scene

由图 7 可见:GridDehazeNet 方法的边缘区域信息损失严重,塔尖边缘模糊程度高,整体图像较为模糊;FFA-Net 方法的塔尖出现少量信息丢失,导致塔尖轮廓不够明显,但是去雾效果良好,图像较为清晰;GMAN 方法在水面区域出现大量伪影,塔尖

边缘模糊程度高,整体图像较为模糊,边缘区域信息损失严重;本文方法的去雾结果最接近无雾原图,在湖面和天空区域去雾效果良好,整体颜色最接近无雾图,在水面区域没有出现上述方法的失真现象,塔尖也没有信息丢失导致的模糊现象出现。

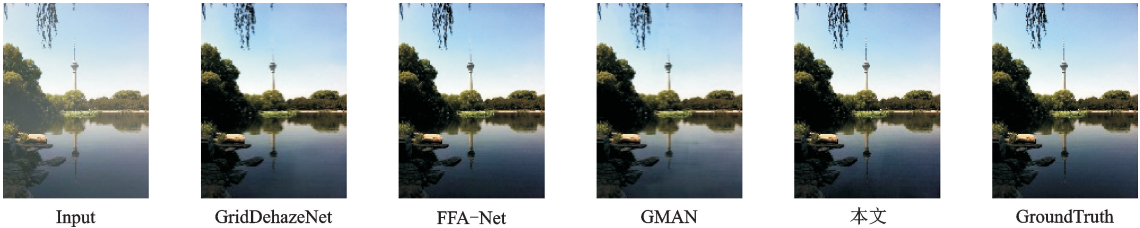


图 7 湖面有雾图像去雾效果对比
Fig. 7 Comparison of haze removal effect of haze image on lake surface

由图 8 可见:GridDehazeNet 方法的地砖区域存在大量白色块状失真,地面亮暗交界处存在黑色伪影;FFA-Net 方法在路面亮暗交界边缘区域存在黑色块状伪影,车牌附近的颜色较深,同时残留着一些雾未去除干净;GMAN 方法在天空和地

砖区域出现大量失真现象,整张图像存在块状伪影;本文方法的去雾结果最接近无雾原图,在近景浓雾区域去雾效果良好,整体颜色最接近无雾图,在地面亮暗交界区域没有出现上述方法的伪影现象。



图 8 浓雾图像去雾效果对比
Fig. 8 Comparison chart of dehazing effect of dense haze image

由图 9 可见:GridDehazeNet 方法的路面区域存在少量失真;FFA-Net 方法的路面黑色区域色块与原图稍有不同,去雾效果良好;GMAN 方法在天

空区域出现少量伪影,在路面黑色区域存在大量失真;本文方法的去雾结果最接近无雾原图,在近景薄雾区域去雾效果良好,整体颜色最接近无雾图,在路

面区域没有出现明显的失真现象。



图 9 薄雾图像去雾效果对比

Fig. 9 Comparison of haze image dehazing effect

在整理分析 RESIDE SOTS 室内数据集时,从中选取 4 张有代表性的图像进行去雾效果对比,图 10 和图 11 存在明暗交替场景的图像去雾效果,图 12 和图 13 对比室内暗光环境和亮光环境下的图像去雾效果。

由图 10 可见: GridDehazeNet 和 FFA-Net 方

法的桌面上方亮暗交替区域存在未恢复的薄雾; FFA-Net 和 GMAN 方法的墙面区域存在颜色失真; 本文方法的去雾结果最接近无雾原图,存在亮暗交替区域的图像去雾效果良好,整体颜色最接近无雾图,没有出现明显的失真现象。



图 10 存在亮暗交替区域的去雾效果对比图 1

Fig. 10 Comparison chart of dehazing effect with alternating bright and dark areas

由图 11 可见: GridDehazeNet 和 GMAN 方法的纯色墙面区域存在大量颜色失真; FFA-Net 方法的图像清晰度不高; 本文方法的去雾结果最接近无

雾原图,存在亮暗交替区域的去雾效果良好,整体颜色和清晰度最接近无雾图,在墙面区域没有出现明显的失真现象。

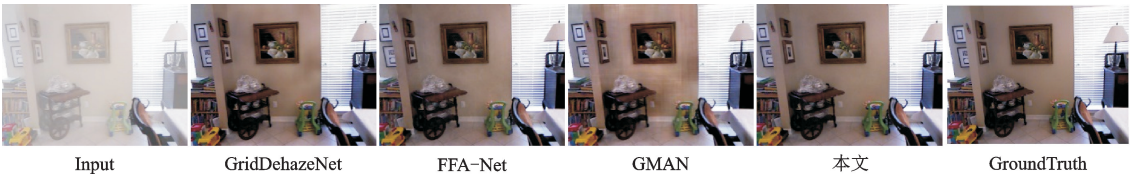


图 11 存在亮暗交替区域的去雾效果对比图 2

Fig. 11 Comparison chart of dehazing effect with alternating bright and dark areas

由图 12 可见: GridDehazeNet 方法的房门区域存在颜色失真; FFA-Net 方法的暗色区域还存在未去除的雾; GMAN 方法在墙壁区域出现大量失真;

本文方法的去雾结果最接近无雾原图,在室内暗光环境去雾效果良好,整体颜色最接近无雾图,在整体区域都没有出现明显的失真现象。



图 12 室内暗光环境去雾效果对比

Fig. 12 Comparison chart of dehazing effect in indoor dark light environment

由图 13 可见:GridDehazeNet 方法的墙壁区域和椅背都存在颜色失真;FFA-Net 方法的椅背出现失真色块;GMAN 方法在墙壁和地面区域出现大量

颜色失真;本文方法的去雾结果最接近无雾原图,在室内亮光环境去雾效果良好,整体颜色最接近无雾图,在整体区域都没有出现明显的失真现象。



图 13 室内亮光去雾效果对比

Fig. 13 Indoor bright light dehazing effect comparison chart

最后在 RESIDE SOTS Outdoor 测试数据集中选择 5 张浓雾图像进行去雾效果对比,结果如图 14 所示。由图 14 可以看出,本文所提出的方

法在浓雾图像去雾任务中表现是最好的,去雾后的图出现的色彩失真和细节丢失最少,最接近无雾图。



图 14 室外浓雾图像去雾效果对比

Fig. 14 Comparison of dehazing effects of outdoor dense haze images

2.4 消融实验

本文提出的基本编解码层引入了一个精心设计的双注意力模块。为了评估所提模块的有效性,在 RESIDE SOTS Outdoor 测试数据集上进行了 3 项消融实验。设定使用普通卷积作为提取局部特征的模块以及每层编解码层不使用通道注意力融合输出特征的模型作为基线。所有数据表中粗体标注出的数据为最优。

首先,将提取局部特征的普通卷积替换成双注意力模块以证明双注意力模块的作用。实验结果见表 2,结果表明,所提出的双注意力模块可以有效地

提高去雾性能。在本文提出的双注意力模块中,引入了两个注意力模块,即通道注意力和像素注意力,以结合局部相关性和全局依赖。分别拆解通道注意力和像素注意力,以研究它们对去雾性能的贡献,实验结果见表 3。可以看出,无论缺少通道注意力还是像素注意力,去雾效果都有一定程度的下降,符合双注意力模块的设计原则。

其次,研究编解码器中的基本层模型结构对整体网络性能的影响,分别对比了使用通道注意力融合基本层输出特征和不使用通道注意力融合基本层输出特征两个编解码模块的设计方式对去雾性能

的影响,实验结果见表 4。结果表明,本文所提出的使用通道注意力融合基本层输出特征的方式可以明显提高图像去雾效果。

当将双注意力模块和通道注意力融合编解码层输出特征的设计同时添加到基线中时即得到本文所提出的方法,最终取得了最佳结果。

2.5 模型复杂度分析

表 5 列出了各个模型的参数量、浮点数运算量和测试时间,测试采用 RESIDE SOTS Outdoor 测试数据集 500 张图像,图像大小均为 620×460 ,测试时间为 500 张图像的平均测试时间。其中,FFA-Net 的浮点数运算量最大,远远超过本文方法;本文方法可以在适度的浮点数运算量下获得最佳的去雾结果。此外,相比于 FFA-Net,本文方法还具有较快的推理速度处理相同的图像,本文提出的模型仅需 74.74% (0.71/0.95)的时间,能很好地兼顾效率与性能。

表 2 通道和像素注意力模块对网络去雾性能的影响
Table 2 Effects of channel and pixel attention modules on network dehazing performance

参数	数值	
	普通卷积	CPAM
R_{sn}/dB	33.26	34.54
结构相似度	0.981	0.985 4

注:加粗数据表示最优结果。

表 3 像素和通道注意力在双注意力模块中不同组合形式对网络去雾效果的影响
Table 3 Effects of different combinations of pixel and channel attention in the dual-attention module on network dehazing performance

参数	数值			
	普通卷积	通道注意	像素注意	像素注意
		力+普通卷积	力+普通卷积	力+通道注意力
R_{sn}/dB	33.50	33.92	33.89	34.54
结构相似度	0.982 5	0.984 0	0.983 9	0.985 4

注:加粗数据表示最优结果。

表 4 通道注意力融合输出特征对去雾性能的影响
Table 4 Effect of channel attention fusion output features on dehazing performance

参数	数值	
	不使用通道注意力	使用通道注意力
R_{sn}/dB	34.12	34.54
结构相似度	0.983 2	0.985 4

注:加粗数据表示最优结果。

表 5 模型复杂度比较				
Table 5 The model complexity comparison				
方法	参数量/ 10^6	浮点数 运算量/ 10^9	测试时间/ s	R_{sn}/dB
GMAN	0.96	188.15	0.39	20.49
Grid DehazeNet	0.82	154.97	0.52	32.13
FFA-Net	4.46	2 500.39	0.95	36.35
本文	5.01	193.22	0.71	36.82

3 结 论

本文提出了一个全局与局部注意力融合的图像去雾方法,以获得正常清晰度和无色彩失真的去雾图像。本文所提方法首先利用通道注意力将输入的有雾图像在通道维度切分为两部分,一部分送入通道像素注意力通道抽取局部特征,另一部分送入 Transformer 通道学习全局特征,然后利用像素注意力对两个通道学习的特征进行融合,将上述模块作为基本单元组合为一个多级 U 型去雾网络,增加残差连接缓解上下采样导致的细节信息丢失,最后在网络底层加入一个 Transformer 模块学习全局信息。实验结果表明:本文提出的全局与局部注意力融合方法能有效地去除雾霾,提升图像的对比度与清晰度,设计的多级 U 型去雾网络和残差连接结构能够缓解细节丢失,提升去雾效果,获得清晰的图像。

参考文献:

[1] HE Kaiming, SUN Jian, TANG Xiaoou. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.

[2] 高涛,刘梦尼,陈婷,等. 结合暗亮通道先验的远近景融合去雾算法 [J]. 西安交通大学学报, 2021, 55(10): 78-86.

GAO Tao, LIU Mengni, CHEN Ting, et al. A far and nearscene fusion defogging algorithm based on the prior of dark-light channel [J]. Journal of Xi'an Jiaotong University, 2021, 55(10): 78-86.

[3] CAI Bolun, XU Xiangmin, JIA Kui, et al. DehazeNet: an end-to-end system for single image haze removal [J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.

[4] 杨燕,陆鑫璇. 结合自适应亮度变换不等式估计透射率的图像去雾方法 [J]. 西安交通大学学报, 2021, 55(6): 69-76.

- YANG Yan, LU Xinxuan. An image dehazing method combining adaptive brightness transformation inequality to estimate transmittance [J]. Journal of Xi'an Jiaotong University, 2021, 55(6): 69-76.
- [5] LI Boyi, PENG Xiulian, WANG Zhangyang, et al. AOD-net: all-in-one dehazing network [C]//2017 IEEE International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2017: 4780-4788.
- [6] HU Jie, SHEN Li, SUN Gang. Squeeze-and-Excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [7] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham, Switzerland: Springer International Publishing, 2018: 3-19.
- [8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [9] QIN Xu, WANG Zhilin, BAI Yuanchao, et al. FFA-net: feature fusion attention network for single image dehazing [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2020: 11908-11915.
- [10] HU Jie, SHEN Li, SUN Gang. Squeeze-and-Excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [11] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham, Switzerland: Springer International Publishing, 2015: 234-241.
- [12] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [13] LOWE D G. Object recognition from local scale-invariant features [C]//Proceedings of the seventh IEEE international conference on computer vision. Piscataway, NJ, USA: IEEE, 1999: 1150-1157.
- [14] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [15] LIN Zhouhan, FENG Minwei, DOS SANTOS C N, et al. A structured self-attentive sentence embedding [EB/OL]. (2017-10-25). <https://arXiv.org/abs/1703.03130>.
- [16] LIN Min, CHEN Qiang, YAN Shuicheng. Network in network [EB/OL]. (2017-10-25). <https://arXiv.org/abs/1312.4400>.
- [17] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2021: 9992-10002.
- [18] LI Boyi, REN Wenqi, FU Dengpan, et al. Benchmarking single-image dehazing and beyond [J]. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.
- [19] LOSHCHILOV I, HUTTER F. SGDR: stochastic gradient descent with warm restarts [EB/OL]. [2022-10-10]. <https://arxiv.org/abs/1608.03983>.
- [20] WELSTEAD S T. Fractal and wavelet image compression techniques [M]. Bellingham, WA, USA: SPIE Optical Engineering Press, 1999.
- [21] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [22] JI Haobo, FENG Xin, PEI Wenjie, et al. U2-Former: a nested U-shaped transformer for image restoration [EB/OL]. [2022-10-10]. <https://arXiv.org/abs/2112.02279>.
- [23] WANG Zhendong, CUN Xiaodong, BAO Jianmin, et al. Uformer: a general U-shaped transformer for image restoration [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2022: 17662-17672.
- [24] LIU Xiaohong, MA Yongrui, SHI Zhihao, et al. GridDehazeNet: attention-based multi-scale network for image dehazing [C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2019: 7313-7322.
- [25] LIU Zheng, XIAO Botao, ALRABEIAH M, et al. Generic model-agnostic convolutional neural network for single image dehazing [EB/OL]. (2019-11-29). <https://arXiv.org/abs/1810.02862>.
- [26] GOLTS A, FREEDMAN D, ELAD M. Unsupervised single image dehazing using dark channel prior loss [J]. IEEE Transactions on Image Processing, 2019, 29: 2692-2701.

(编辑 武红江)