

文章编号: 2095-2163(2023)12-0170-04

中图分类号: TP319

文献标志码: A

多尺度 CNN 与高频信息融合的图像去雾算法

罗晶¹, 肖德广²

(1 贵州高速公路集团有限公司, 贵阳 560009; 2 贵州中南交通科技有限公司, 贵阳 560018)

摘要: 针对现有去雾算法在恢复图像边缘区域效果不佳以及存在薄雾残留等问题, 本文提出了一种基于多尺度卷积神经网络(CNN)与高频信息融合的图像去雾算法(HF-MSNet), 采用拉普拉斯算子分离高频信息作为额外先验, 用于提高图像细节恢复。首先, 采用多尺度 CNN 作为主干网络, 提取图像的主要纹理特征; 其次, 采用高频特征融合分支对高频信息逐阶段地提炼, 并融合了通道注意力模块, 使用注意力机制对不重要的特征进行约束; 最后, 采用高频注意力模块对全局特征自适应地筛选, 整合高频特征和纹理特征的分布, 得到目标的无雾图像。实验结果表明, 本文的算法可以生成更自然真实的去雾图像, 并在峰值信噪比(PSNR)、结构相似度(SSIM)两种指标上都达到了理想的效果。

关键词: 图像去雾; 卷积神经网络; 高频信息; 注意力机制

Image dehazing algorithm based on multi-scale CNN and high frequency information fusion

LUO Jing¹, XIAO Deguang²

(1 Guizhou Expressway Group Co., Ltd., Guiyang 560009, China;

2 Guizhou Zhongnan Transportation Technology Co., Ltd., Guiyang 560018, China)

Abstract: This paper proposes an image dehazing algorithm based on multi-scale convolutional neural network and high-frequency information fusion (HF-MSNet) to address the issues of poor restoration of image edge regions and residual haze in existing dehazing algorithms. Using Laplace operator to separate high-frequency information as an additional prior for improving image detail restoration. Firstly, multi-scale CNN is used as the backbone network to extract the main texture features of the image. Secondly, the high-frequency feature fusion branch extracts high-frequency information step by step and integrates channel attention modules, using attention mechanisms to constrain unimportant features. Finally, a high-frequency attention module is used to adaptively filter global features, integrating the distribution of high-frequency and texture features, and obtaining a fog free image of the target. The experimental results show that the algorithm proposed in this paper can generate more natural and realistic dehazing images, and achieves ideal results in both peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) metrics.

Key words: image dehazing; convolutional neural network; high-frequency information; attention mechanism

0 引言

雾霾是一种常见的大气现象, 环境中存在大量的悬浮颗粒, 如沙尘和烟雾气体等, 在这样的场景下拍摄的图像, 往往质量不佳。在大数据的背景下, 许多计算机视觉任务需要大量的干净图像。由于雾霾的影响, 采集的数据中存在许多能见度低下的图像, 极大地影响后续的任务, 如目标检测、图像分类和分割等。近年来, 图像去雾已成为计算机视觉的重要研究领域。

早期的图像去雾方法大多都是基于大气散射模

型, 估计介质透射率和全球大气光照值, 通过逆转模型公式得到目标结果。然而, 雾霾的形成跟海拔、温度、湿度等因素有关, 在实际场景中, 雾霾的分布是不均匀的, 很难通过该方法得到理想的结果。

随着卷积神经网络在图像领域中展现出卓越的性能, 许多研究者也开始采用卷积神经网络(CNN)实现图像去雾。图像去雾算法大都使用深度卷积神经网络来估计介质透射率等参数, 恢复清晰图像, 如有的算法直接学习雾图与清晰原图之间的非线性映射, 得到最终预测图像^[1]; 有的算法基于整个图像预测整体透射图的粗尺度模型, 与一个局部细化结

基金项目: 贵州省交通运输厅重点研发计划。

作者简介: 罗晶(1992-), 男, 硕士, 高级工程师, 主要研究方向: 智慧交通领域的深度学习图像处理; 肖德广(1984-), 男, 学士, 工程师, 主要研究方向: 机电系统开发、图像处理。

收稿日期: 2022-12-16

果的细尺度模型结合,得到去雾预测图^[2]。但上述方式估计的透射率或大气光往往不够准确,严重影响图像恢复,泛化性较差。为了解决这一问题,许多研究者提出了端到端的 CNN 算法模型用于单幅图像去雾,去雾算法不再依赖于介质透射率和大气光照值的估计。

基于上述研究,本文提出一个端到端的基于多尺度卷积神经网络与高频信息融合的去雾算法,用于单幅图像去雾,不需要估计任何参数,能够直接恢复清晰无雾的图像,并有效地解决了算法在图像边缘区域的去雾能力不佳、指标低下等问题。首先,采用高频信息(High Frequency, HF),即在图像相邻区域中色彩或梯度处存在明显区别的分布特征,来作为额外先验。其次,采用空间与锐化函数获取高频信息,将其作为额外输入,并引用高频注意力模块(High Frequency Attention Module, HFAM)用于增强对边缘信息的恢复。

1 相关工作

1.1 基于大气散射模型

1976 年, John Wiley & Sons 等提出大气散射模型,解决图像恢复问题,式(1):

$$I(x) = J(x)T(x) + A(1 - T(x)) \quad (1)$$

其中, $I(x)$ 为观测到的雾天图像; A 为全局大气光照值; x 表示像素位置; $T(x)$ 为介质透射率; $J(x)$ 为无雾清晰的图像。

介质透射率描述了空气中的各项介质对光线传播的影响程度,式(2):

$$T(x) = e^{-\beta d(x)} \quad (2)$$

其中, β 和 $d(x)$ 分别为大气散射参数和场景深度。

基于大气散射模型公式(1)可以发现,在不考虑 A 和 $T(x)$ 的情况下,图像去雾是一个不确定的问题。

1.2 基于深度学习的去雾算法

随着大数据的发展以及高性能计算机的普及,研究者们提出了许多基于深度学习的算法来解决图像去雾问题,与传统基于先验知识算法相比展现出极大的优越性。首先,基于深度学习来估计介质传输率实现图像去雾,如 MSCNN^[2]、AodNet^[3]、DehazeNet^[4]等算法;其次,为了摆脱大气散射算法对光照条件、湿度等因素的依赖,研究者们进一步提出了新的算法,例如 GFN^[5]、GDN^[6]、EPDN^[7]等端到端的去雾算法,能够更准确地计算雾图和清晰图像之间的映射关系。然而这些算法在恢复清晰图像时,尤其在物体轮廓以及背景边缘区域的效果表现不佳。因此本文提出了基于多尺度卷积神经网络和 高频信息融合的去雾算法来解决这一问题。

2 多尺度 CNN 与高频信息融合的图像去雾算法

2.1 HF-MSNet 算法主要结构

HF-MSNet 算法的网络结构如图 1 所示。本文提出的算法输入由原始雾图、拉普拉斯锐化函数得到的高频信息,以及原始雾图通过上下采样生成全尺寸、1/2 尺寸以及 1/4 尺寸的特征图 3 部分组成。算法模型的主体结构采用多尺度的方式多阶段地完成特征提取,并结合了特征融合操作,以学习不同尺度的特征表达能力。因为高频特征的信息相对简单,本文只采用全尺寸卷积模块进行参数估计。同时,引入了通道注意力模块(CAM),适应性地学习通道上注意力权重,使算法模型更好地完成纹理特征和 高频特征的融合。在算法模型的末端,提出了高频信息注意力模块(HFAM),帮助算法自适应地学习不同频率信息的权重,抑制频率分支特征图输出不重要的特征信息,只学习有用的特征信息向下传播。

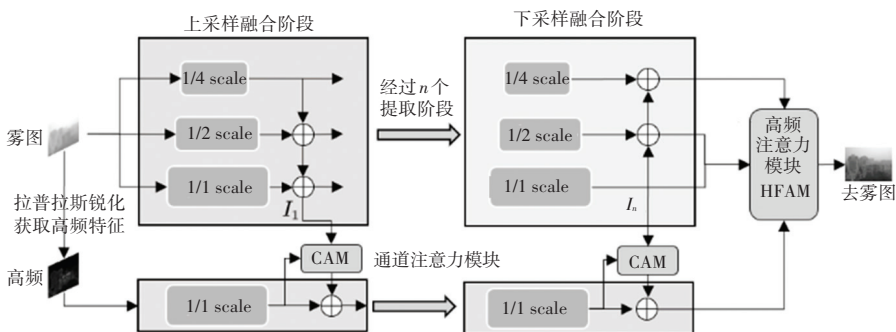


图 1 HF-MSNet 算法的网络结构

Fig. 1 Network structure diagram of the HF-MSNet

2.2 多尺度去雾网络

算法在上下采样融合阶段都采用了多尺度机制用于特征学习,以此在不同尺度间共享有效特征。本文嵌入了残差密度模块(Residual Dense Block, RDB),完成每个尺度的特征提取工作。用向上箭头 $\uparrow(\cdot)$ 和向下箭头 $\downarrow(\cdot)$ 分别表示尺度因子为2的上采样过程和下采样过程, \oplus 表示逐像素的特征融合。上下采样融合阶段,多尺度融合阶段中第*i*个阶段第*j*层尺度上生成的特征图 $L(i, j)$,式(3)和式(4):

$$L(i, j) = RDB(\downarrow(L(i, j-1)) \oplus L(i-1, j)) \quad (3)$$

$$L(i, j) = RDB(\uparrow(L(i, j+1)) \oplus L(i-1, j)) \quad (4)$$

通过上下采样融合的方式,有效显著的特征通过跨尺度的交换,算法表现出更强的特征提取能力。

2.3 高频特征融合

受一些采用额外先验(如场景深度信息、背景梯度)知识处理去雾的启发,本文提出了一种高频特征融合方案。对于一个给定的雾图 $I(x)$,采用拉普拉斯算子将其锐化,可以得到高频特征 $HF(x)$,将其作为主要前向传播特征。此外,将每一个阶段的全尺度特征图,传递到通道注意力模块(CAM)中,适应性学习通道上注意力权重,使算法模型更

好地完成纹理特征和低频特征的融合过程。

用 L_n 表示主干网络全尺度层第*n*个输出, HF_n 表示第*n*个高频特征权重图。将第*n-1*层的高频信息 GF_{n-1} 经过RDB模块提取特征,并采用通道注意力模块获取通道间的关联分布,最后和原始的高频特征在像素维度上融合,引导进一步的特征传播。过程如式(5)所示:

$$HF_n = CAM(RDB(HF_{n-1}), L_n) \oplus RDB(HF_{n-1}) \quad (5)$$

2.4 高频注意力模块

近年来,注意力机制在视觉领域展现了良好的性能,在许多图像增强任务中验证了其强大的特征提取能力。本文受到卷积注意力机制的启发,进一步提出用高频注意力模块(HFAM)筛选有用的特征,提升本文算法的性能,高频注意力模块结构如图2所示。一方面高频注意力模块引入了原始雾图和主体结构生成的纹理特征来提供有用的信息来帮助本文算法对高频区域边缘分布的感知;另一方面高频注意力模块能够生成注意力图,学习频率信息的有用程度与原有特征融合,可以抑制高频特征中不重要的特征,使得网络能够更加有效地学习到有用的信息来提升网络恢复的效果。

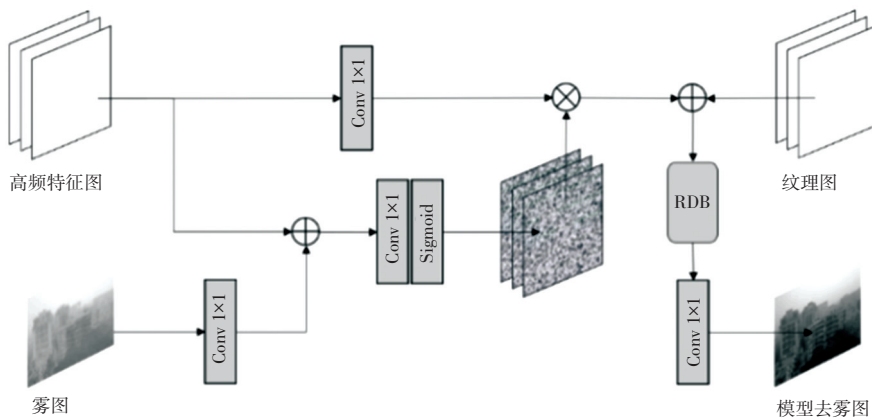


图2 高频注意力模块结构

Fig. 2 High frequency attention module structure

HFAM由3个输入组成:最后一阶段的高频特征图、纹理特征图以及原始雾图。 L_{in} 代表纹理特征图输入。通过卷积操作进行信息交互,得到带有高频边缘感知的初步融合结果,在Sigmoid激活函数的作用下得到权重图 W ,式(6):

$$W = \sigma(\text{Conv}(HF_{in} \oplus \text{Conv}(I_{hase}))) \quad (6)$$

其中, $\sigma(\cdot)$ 表示Sigmoid激活函数; HF_{in} 表示高频特征图; I_{hase} 表示原始雾图。

W 能够通过学习自适应地赋予有效信息相应的权值。因此将权值 W 与原始高频特征图进行点乘操作,得到高频特征带注意力权重的信息分布结果

L_{mid} ,式(7):

$$L_{mid} = \text{Conv}(HF_{in}) \otimes W \oplus L_{in} \quad (7)$$

其中, L_{in} 为纹理特征图。

最后再次采用残差密度模块(RDB)对高频特征和纹理特征融合的结果进行整合,在残差密度模块的学习下逐步得到最终的去雾结果。

3 实验与分析

3.1 数据集

本文采用RESIDE基准数据集用于算法的训练和测试。RESIDE的训练集包含室内场景训练集

(ITS)和室外场景雾图训练集(OTS)。ITS 包含1 399张清晰干净的室内图像,在此基础上生成不同程度的雾图,设置 $\beta \in [0.6, 1.8]$ 、大气光照强度 $A \in [0.7, 1.0]$,并采用数据增强的方式扩展为13 990张雾图;OTS 共包含 296 695 张雾图以及标签。此外,RESIDE 测试集 SOTS 包含了 500 张室内场景雾图和 500 张室外场景雾图分布作为对应测试集。

在训练过程中,算法的输入尺寸为 256×256 ,总共迭代 80 次。采用 Adam 优化器,学习率设置为 0.000 1, $\text{betas} = (0.9, 0.999)$ 。

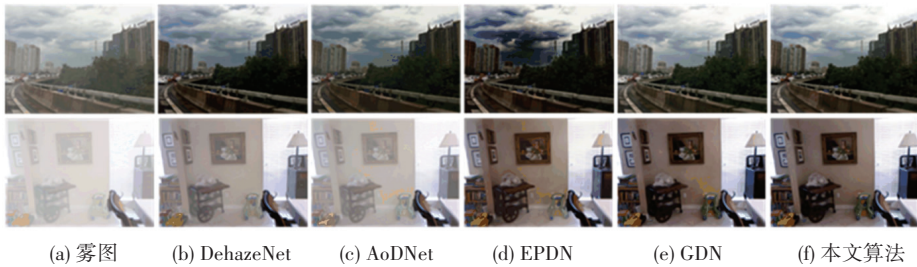


图 3 不同算法的去雾结果

Fig. 3 Comparison of defogging results of different algorithms

同时,采用峰值信噪比(PSNR)和结构相似度(SSIM)作为评价指标,实验结果见表 1。可见,在室内场景下,本文算法的 PSNR 值比第二名 GDN 的结果增长了 1.55 dB;而室外场景下,本文算法的 PSNR 值和 SSIM 值分别比第二名 GDN 的结果增长了 1.27 dB 和 0.004 5。本文算法在定量对比上有着良好的表现。

表 1 对比实验结果

Table 1 Comparison of experimental results

算法	SOTS 室内		SOTS 室外	
	PSNR	SSIM	PSNR	SSIM
DCP ^[8]	16.62	0.817 9	19.13	0.814 8
DehazeNet ^[4]	21.14	0.847 2	22.46	0.851 4
AODNet ^[3]	19.06	0.850 4	20.29	0.876 5
EPDN ^[7]	25.06	0.923 2	22.57	0.863 0
GDN ^[6]	32.16	0.983 6	30.86	0.981 9
本文算法	33.71	0.986 9	32.13	0.986 4

综上,本文的算法不论是 PSNR 和 SSIM 指标上的性能,还是在去雾结果的视觉效果上来说,都展现了巨大的优越性,都证明了本文算法的有效性。

4 结束语

本文提出了一个端到端的基于多尺度 CNN 与高频信息融合的去雾算法,应用于单幅图像去雾。多尺度特征能够学习图像基本的纹理知识,高频特征对图像的轮廓和细节具有更佳的精度把握,在多尺度结构以及注意力机制的引导下融合,恢复的图像具有更好

3.2 实验分析

为了验证算法的有效性,将本文算法与许多现有去雾算法如 DCP、DehazeNet、AODNet、EPDN、GDN 等进行了对比实验,不同算法的去雾结果如图 3 所示,可见本文的算法生成的去雾图像具有更好的视觉效果,DehazeNet 和 AoDNet 生成的去雾结果在部分区域仍然存在雾霾,EPDN 方法在某种程度上引起了严重的阴影和色差,而 GDN 的去雾结果虽然获得了较好的结果,但是仍然存在一些薄雾和噪声。

的视觉效果。将本文所提的 HF-MSNet 算法与许多现有的去雾算法做了对比试验,实验结果表明:本文算法不论是在定性还是定量上都展现了显著的优越性。

参考文献

- [1] ZHANG H, SINDAGI V, PATEL V M. Multi-scale single image dehazing using perceptual pyramid deep network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018: 902-911.
- [2] REN W, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks [C]//Proceedings of 14th European Conference on Computer Vision- ECCV 2016. Amsterdam, the Netherlands: Springer International Publishing, 2016: 154-169.
- [3] LI B, PENG X, WANG Z, et al. An all-in-one network for dehazing and beyond [J]. arXiv preprint arXiv:1707.06543, 2017.
- [4] CAI B, XU X, JIA K, et al. Dehazenet: An end-to-end system for single image haze removal [J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [5] REN W, MA L, ZHANG J, et al. Gated fusion network for single image dehazing [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3253-3261.
- [6] LIU X, MA Y, SHI Z, et al. Griddehazenet: Attention-based multi-scale network for image dehazing [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 7314-7323.
- [7] QU Y, CHEN Y, HUANG J, et al. Enhanced pix2pix dehazing network [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 8160-8168.
- [8] HE K, SUN J, TANG X. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(12): 2341-2353.