

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2016)04-464-11

论文引用格式: Mai J M, Wang M H, Liang Y, Cai R C. Single image dehazing algorithm by feature learning[J]. Journal of Image and Graphics, 2016, 21(4): 464-474. [麦嘉铭, 王美华, 梁云, 蔡瑞初. 特征学习的单幅图像去雾算法[J]. 中国图象图形学报, 2016, 21(4): 464-474. ] [ DOI: 10.11834/jig.20160408 ]

## 特征学习的单幅图像去雾算法

麦嘉铭<sup>1</sup>, 王美华<sup>1</sup>, 梁云<sup>1</sup>, 蔡瑞初<sup>2</sup>

1. 华南农业大学数学与信息学院, 广州 510642; 2. 广东工业大学计算机学院, 广州 510006

**摘要:**目的 雾霾天气使图像降质,严重影响军事、交通和安全监控等领域信息系统的正常运作,因此图像去雾具有重要研究意义。目前主流的单幅图像去雾算法主要利用各种与雾相关的颜色特征实现,但不同的颜色先验知识往往存在各自的场景局限性。为提高图像去雾的普适性,提出一种特征学习的单幅图像去雾方法。**方法** 首先通过稀疏自动编码器对有雾图像进行多尺度的纹理结构特征提取,同时抽取各种与雾相关的颜色特征。然后采用多层神经网络进行样本训练,得到雾天条件下纹理结构特征及颜色特征与场景深度间的映射关系,并估算出有雾图像的场景深度图。最后结合大气散射模型,根据场景深度图复原无雾图像。**结果** 与主流去雾算法的结果定性对比,复原后的图像细节更清晰,颜色更自然。采用均方误差和结构相似度定量评价各算法去雾结果与真实无雾图像的相似度,本文算法结果与真实无雾场景间的相似度最高,达到 99.9%。**结论** 对实验结果的定性及定量分析表明,本文算法能有效获取有雾图像的场景深度,复原出视觉效果理想的无雾图像,且具有很好的场景普适性。**关键词:**图像去雾;特征学习;深度学习;神经网络;稀疏编码

## Single image dehazing algorithm by feature learning

Mai Jiaming<sup>1</sup>, Wang Meihua<sup>1</sup>, Liang Yun<sup>1</sup>, Cai Ruichu<sup>2</sup>

1. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China;  
2. School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China

**Abstract:** **Objective** Fog-degraded images seriously affect the normal operation of the information systems of the military, traffic, safety monitoring, and other fields. Hence, research on image dehazing has become important and practical. Most existing single image dehazing approaches focus on investigating haze-relevant color features. However, different color priors have their respective limitations. To improve the adaptability of the image dehazing algorithm, we propose a single image dehazing algorithm by feature learning. **Method** First, we extract multi-scale textural and structural features from a hazy image by using a sparse auto encoder; the haze-relevant color features are obtained as well. Second, we model the scene depth with the extracted textural, structural, and color features by using multi-layer neural networks. By training samples with the neural network model, the relationship between scene depth and the features is determined, and the depth map of the hazy image is restored. Finally, the haze-free image with the depth map is restored according to the atmospheric

收稿日期: 2015-09-28; 修回日期: 2015-11-17

**基金项目:**国家自然科学基金项目(61202269, 61472089, 61202293); 国家自然科学基金—广东联合基金项目(U1501254); 广东省自然科学基金(2014A030308008); 广东省科技计划项目(2013B051000076, 2014A050503057, 2015A020209124, 2015B010131015, 2015B010108006)

**第一作者简介:**麦嘉铭(1991—),男,华南农业大学计算机应用技术专业硕士研究生,主要研究方向为计算机视觉、机器学习。

E-mail: jiamingmai@163.com

**通信作者:**王美华,副教授,硕士生导师, E-mail: wangmeihua@scau.edu.cn

**Supported by:** National Natural Science Foundation of China(61202269, 61472089, 61202293)

scattering model. **Result** Compared with the results of mainstream algorithms, those of the proposed algorithm are more natural and have more details. Qualitative evaluation of the similarity between our results and the ground truth haze-free images indicates that our results have high similarity to the ground truth situation, reaching 99.9%. **Conclusion** Experiments show that the proposed algorithm can effectively restore the depth map of a hazy image and produce a high-quality haze-free image with good adaptability.

**Key words:** image dehazing; feature learning; deep learning; neural networks; sparse coding

## 0 引言

在雾、霾等恶劣天气下,大气中存在较多的水滴或尘粒,致使光在传播过程中发生散射或被吸收,此时户外拍摄的图像往往模糊不清,对比度较低。监控设备等依赖于户外图像的系统性能急剧下降,甚至无法正常运转。因此,对图像进行去雾处理具有重要意义。

目前图像去雾的方法主要分为两类:一类是使用图像增强的方法,即通过增强图像的对比度来达到去雾目的<sup>[1-4]</sup>,这类方法可以应用并有针对性地改良已有的成熟图像处理算法,对薄雾图像能够取得较好的去雾效果,但由于没有针对成雾的客观原因,所以不能从根本上实现去雾。另一类是基于物理模型的去雾方法<sup>[5-14]</sup>,该类方法通过研究大气成雾以及图像退化的客观机理,构建出大气散射模型,再根据模型对雾天图像复原,从而实现去雾。这一类方法效果理想,不容易使图像信息丢失。

近年来,在基于物理模型的单幅图像去雾方面有了重要进展。Tan 等人<sup>[5]</sup>以大气散射模型为基础,假设局部区域大气光平滑,并根据有雾图像的对比度较无雾图像低得多的先验信息,提出了局部对比度最大化的去雾算法。但使用该算法去雾后的图像容易出现光晕效应和过饱和。Fattal 等人<sup>[6]</sup>利用了独立成分分析的方法对图像进行去雾处理,该方法对薄雾图像的处理效果理想,但存在时间复杂度较高和不能处理灰度图像等缺点。He 等人<sup>[7]</sup>通过对大量无雾图像进行实验,发现了暗原色先验统计规律,并以此提出一种简单有效的去雾算法。该算法在大部分情况下能取得理想的去雾效果,但效率较低,且对于天空或白色物体等不具有暗原色的区域失效。针对 He 方法<sup>[7]</sup>的不足,研究人员在此基础上提出了一些改进的方法<sup>[8-10,13-15]</sup>。例如,Tarel 等人<sup>[8]</sup>、Wang 等人<sup>[9]</sup>和 He 等人<sup>[15]</sup>分别采用中值

滤波、双边滤波和引导滤波替代文献[7]中的软抠图算法<sup>[16]</sup>以提高去雾算法的效率;Meng 等人<sup>[13]</sup>通过研究透射率函数的内在边界约束,提出一种有效的正则化方法来进行图像去雾;Tang 等人<sup>[14]</sup>采用随机森林使暗原色和多种颜色特征相结合,提高透射率的估计精度。该方法的局限性在于两方面:一方面只利用了颜色特征,而没有考虑到与雾霾相关的纹理结构特征,致使透射率的估计精度受到限制;另一方面,训练随机森林所使用的数据是基于场景深度与纹理不相关的假设生成的,由于该假设只在少数情况下成立,从而导致训练数据不完备,训练后的模型具有一定的场景局限性。

为更好地利用图像纹理、结构和颜色等特征以提高去雾效果,本文提出一种特征学习的单幅图像去雾算法。首先通过稀疏自动编码器对有雾图像进行多尺度的纹理结构特征提取,同时抽取各种与雾相关的颜色特征。然后采用多层神经网络进行样本训练,学习得到雾天条件下纹理结构特征及颜色特征与场景深度间的映射关系,并估算出有雾图像的场景深度图。最后结合大气散射模型,根据场景深度图复原无雾图像。实验结果表明,与主流的去雾算法相比,该算法能够复原出更高质量的无雾图像,且具有更好的场景普适能力。

## 1 大气散射模型

大气散射模型由 McCartney<sup>[17]</sup>于1978年提出,经过 Narasimhan 和 Nayar<sup>[18-20]</sup>的进一步推导后,被广泛用于计算机视觉和图像处理等领域。该模型描述了雾天图像退化的客观机理,即

$$I(x) = J(x)t(x) + A[1 - t(x)] \quad (1)$$

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

式中, $I(x)$ 表示有雾图像在位置 $x$ 上的强度值, $J(x)$ 表示对应的无雾条件下的场景光照度,即要复原的无雾图像的像素值, $A$ 为大气光照度, $t(x)$ 表示场景光

的透射率,  $\beta$  为大气散射系数,  $d(x)$  表示场景深度。在图像去雾中, 由于  $I(x)$  已知, 所以去雾处理的本质就是要求出  $A$  和  $t(x)$ , 然后通过式(1)求出  $J(x)$ 。

根据式(1)(2)可知, 场景深度  $d$  是十分重要的信息。文献[20]指出, 在同质大气条件下,  $\beta$  可以看做为较小的常量, 所以在  $d$  已知的情况下,  $t$  可以直接通过式(2)求得。另外, 根据式(1), 当  $t(x) = 0$  时, 可以得到

$$I(x) = A \tag{3}$$

式(3)表明, 大气光照度  $A$  的值可在  $t(x) = 0$  处取得。因为  $t(x) = e^{-\beta d(x)}$ , 所以当  $d(x)$  的值足够大时, 可近似认为  $t(x) \approx 0$ 。因此, 如果已知场景深度图  $d$ , 可求得大气光照度为

$$A = I(x), x \in \{x | \forall y: d(y) \leq d(x)\} \tag{4}$$

综上, 图像去雾的问题可进一步转化为场景深度  $d$  的求解问题。然而, 在一般情况下, 一幅有雾图像的 场景深度图  $d$  往往难以估算。

2 雾天图像景深估计方法

本文结合特征学习, 提出一种有效估计有雾图

像场景深度的方法。图 1 给出了该方法的流程, 其主要由两部分组成: 一是对有雾图像进行特征提取; 二是基于提取后的特征, 利用多层神经网络进行场景深度的估计。

2.1 训练样本的生成

图 1 所示的场景深度估计方法包含了两个需要进行训练的模型(稀疏自动编码器、多层神经网络)。因此, 训练样本的来源是该方法的关键。稀疏自动编码器和多层神经网络的训练, 均需要一组有雾图像及其对应的真实深度图作为训练样本。由于有雾图像的真实深度图在实际中难以获取, 所以本文根据大气散射模型式(1)(2), 用一组无雾图像及其真实深度图生成对应的有雾图像作为训练样本。

图 2 给出了有雾图像的生成过程。首先, 从 Saxena 等人<sup>[21-23]</sup>建立的数据集中获取大量的户外无雾图像及其真实场景深度图。该数据集是一个包含不同场景真实景深的标准训练数据集, 目前已广泛应用于 3 维重建<sup>[21-23]</sup>的研究, 其真实场景深度由激光设备测量得到, 具有较高的精确度, 因此该数据集可用于本文的模型训练。为了使生成的有雾图像更加细致, 在生成有雾图像之前, 采用引导滤波<sup>[15]</sup>

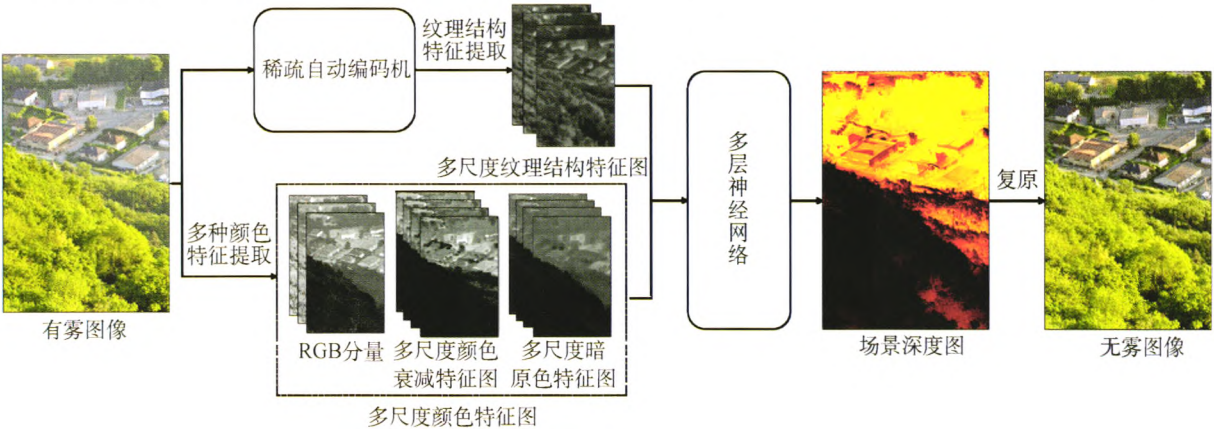


图 1 有雾图像的场景深度估计方法

Fig. 1 The depth estimation approach for hazy images

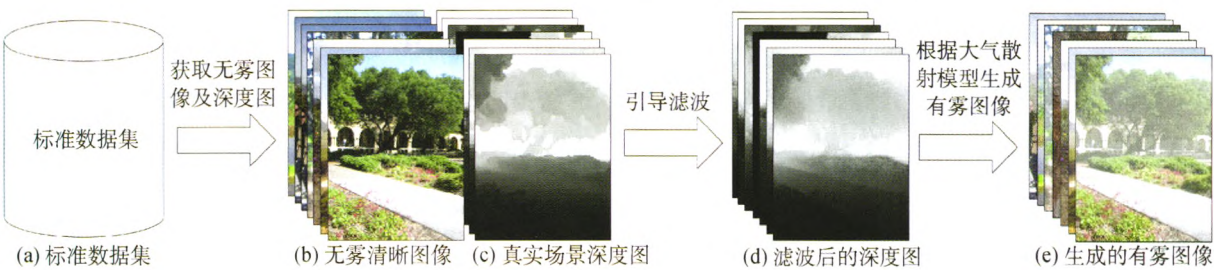


图 2 有雾图像的生成

Fig. 2 Hazy images generation ((a) ground truth data set; (b) haze-free images; (c) ground truth depth maps; (d) the filtered depth maps; (e) the generated hazy images)



对真实场景深度图进行平滑处理,得到如图 2(d) 所示的深度图。然后,根据文献[20]的假设,即同质大气条件下大气散射系数  $\beta$  可看做为一个较小常量,通常令  $\beta=2.0$ 。接着,随机生成大气光照度  $A$ 。由于一般情况下,  $A$  在有雾图像的光亮区域取得<sup>[20]</sup>,所以令大气光照度  $A=[k+\lambda_1, k+\lambda_2, k+\lambda_3]$ ,其中  $k, \lambda_1, \lambda_2$  和  $\lambda_3$  是随机数,  $k$  服从区间  $(0.85, 1)$  上的均匀分布,  $\lambda_1, \lambda_2$  和  $\lambda_3$  均服从均值为 0, 方差为 1 的高斯分布。最后给定无雾图像  $J$ 、场景深度图  $d$ 、大气光照度  $A$  和大气散射系数  $\beta$ , 根据式(1)(2)生成有雾图像  $I$ 。

采用上述方法,可以根据 1 组  $n$  幅无雾图像  $J^{set}=\{J_1, J_2, \cdots, J_n\}$  及其场景深度图  $d^{set}=\{d_1, d_2, \cdots, d_n\}$ ,生成一组有雾图像  $I^{set}=\{I_1, I_2, \cdots, I_n\}$ 。最终使用  $I^{set}$  和  $d^{set}$  作为模型的训练样本集。

2.2 多尺度颜色特征提取

有雾图像与无雾图像在色彩上存在较大差异。因此,在图像去雾中,颜色特征常用于有雾图像的场景深度估计。诸如 Tan<sup>[5]</sup> 提出的局部最大对比度、Ancuti 等人<sup>[11]</sup> 提出的色调差异、He 等人<sup>[7]</sup> 提出的暗原色和 Zhu 等人<sup>[12]</sup> 提出的颜色衰减性等各种基于颜色的先验知识已充分用于单幅图像去雾。这些不同类型的先验知识可看做为一系列与雾霾浓度相关的颜色特征。事实表明,这些特征在针对特定场景独立使用时已取得不错的效果<sup>[5,7,11-12]</sup>,但不同的

特征具有各自的场景局限性。例如,单独使用暗原色处理具有大面积白色物体的区域时效果不佳。结合多种颜色特征可提高场景深度的估计精度,本文对有雾图像进行多种颜色特征的提取,以便进一步估计其场景深度。

2.2.1 多尺度暗原色特征

在图像去雾中,暗原色<sup>[7]</sup>是一种具有重要价值的颜色特征,本文对暗原色特征的定义为

$$D_r(x)=\min_{y\in\Omega_r(x)}\min_{c\in\{R,G,B\}}\frac{I^c(y)}{A^c}\tag{5}$$

式中,  $\Omega_r(x)$  是以  $x$  为中心,大小为  $r\times r$  的局部块,  $I^c$  和  $A^c$  分别表示有雾图像  $I$  和大气光照度  $A$  在颜色通道  $c$  下的分量,  $D_r$  表示局部块尺寸为  $r\times r$  的暗原色特征。  $r$  的取值与场景深度  $d(x)$  的估计直接相关:一方面,较小的  $r$  将导致  $d(x)$  大于真实值;另一方面,较大的  $r$  容易使  $d(x)$  的值过小。因此,采用 Tang 等人<sup>[14]</sup> 提出的方法,使用多尺度的暗原色特征以提高估计的精度,令  $D^M=\{D_{r_1}, D_{r_2}, \cdots, D_{r_M}\}$ ,其中  $D^M$  表示一个由  $M$  种不同尺度  $r_1, r_2, \cdots, r_M$  的暗原色特征所形成的集合。在本文的实验中,取  $M=4, r_1=1, r_2=4, r_3=7, r_4=10$ , 即  $D^4=\{D_1, D_4, D_7, D_{10}\}$ 。图 3(b) 给出了有雾图像的暗原色特征图  $D_{10}$ , 该特征图近似地反映了雾气在图像中的分布情况,图中越亮的区域代表其雾的浓度越高,场景深度值越高。

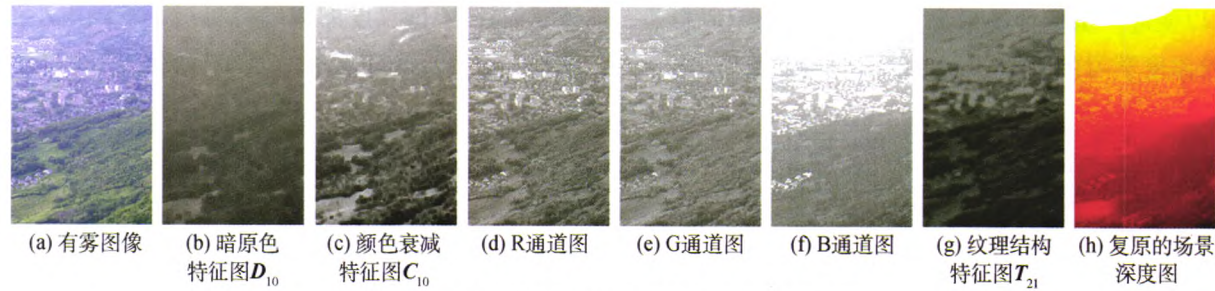


图 3 有雾图像的特征图及复原后的场景深度图

Fig. 3 Different feature maps and restored depth map of the hazy image ((a) the hazy image; (b) the dark channel feature map  $D_{10}$ ; (c) the color attenuation feature map  $C_{10}$ ; (d-f) the R, G and B components of the hazy image; (g) texture structure feature map  $T_{21}$ ; (h) restored depth map)

2.2.2 多尺度颜色衰减特征

颜色衰减先验<sup>[12]</sup>是根据雾天条件下图像的成像机理所提出的一种特征,即在户外数字图像中,雾浓度越高的区域,其亮度均值越高,饱和度均值越低。由于在大部分情况下,有雾图像中的场景深度与雾的浓度正相关,所以可进一步扩展文献[12]对

颜色衰减先验的定义,根据区域中亮度均值和饱和度均值的线性关系来决定其场景深度,即

$$C_r(x)=\min_{y\in\Omega_r(x)}[\theta_0+\theta_1I^{val}(y)+\theta_2I^{sat}(y)]\tag{6}$$

式中,  $I^{val}$  和  $I^{sat}$  分别为有雾图像在 HSV 颜色空间下的亮度分量和饱和度分量,  $\theta_0, \theta_1$  和  $\theta_2$  为线性系数, 本文采用了文献[12]的参数设置,令  $\theta_0=0.189\ 3$ ,



$\theta_1 = 1.026\ 7, \theta_2 = -1.296\ 6$ 。与暗原色特征类似, 本文采用多尺度颜色衰减特征:  $C^d = \{C_1, C_4, C_7, C_{10}\}$ 。图 3(c) 给出了有雾图像的颜色衰减特征图  $C_{10}$ , 该特征图大致反映了有雾图像的场景深度, 越明亮的区域具有越高的场景深度值。

2.2.3 RGB 分量

由于大部分的颜色特征都以 RGB 颜色模型为基础, 所以有雾图像在 RGB 空间上的 3 个分量 ( $R$  通道分量、 $G$  通道分量和  $B$  通道分量) 对场景深度的估计有重要价值其定义为

$$R = I^{red} \tag{7}$$

$$G = I^{green} \tag{8}$$

$$B = I^{blue} \tag{9}$$

式中,  $I^{red}$ 、 $I^{green}$  和  $I^{blue}$  分别表示有雾图像  $I$  在 RGB 颜色空间上的 3 个分量。本文将式 (7) — (9) 中的  $R$ 、 $G$ 、 $B$  看做特殊的特征图, 成为场景深度估计的重要参考信息。图 3(d—f) 分别展示了图 3(a) 的  $R$ 、 $G$ 、 $B$  3 个分量。

2.3 多尺度纹理结构特征提取

传统的单幅图像去雾算法单纯依赖颜色特征估计场景深度并实现去雾, 因此具有一定的局限性。物体的纹理和结构特征是估计场景深度的重要信息, 特别是对于形状规则或颜色较浅的物体, 其纹理和结构在估计中将起关键作用。本文采用稀疏自动编码器对有雾图像中景物的纹理和结构特征进行提取, 以便进一步估计场景深度。稀疏自动编码器是一种深度学习的算法, 最初用于去噪处理<sup>[24]</sup>, 由于

它能够有效提取物体的主要特征, 被广泛应用于物体检测与识别等领域。

图 4 给出了用于提取景物纹理和结构特征的稀疏自动编码器及其训练过程。该自动编码器由两部分组成: 第 1 部分为稀疏编码, 用于提取局部块的主要纹理与结构特征, 通过一个无监督的自学习神经网络实现, 如图 4(c) 所示, 网络的输入是向量化的局部块,  $S$  是对其降维后所得到的特征,  $\hat{I}$  为  $S$  再次映射到高维时所得到的输出向量; 第 2 部分为一个 2 层神经网络, 用于获得纹理结构特征与场景深度之间的映射关系。稀疏自动编码器中的网络使用 sigmoid 函数  $f(x) = 1/(1 + e^{-x})$  作为传输函数, 其训练时, 首先从训练样本集  $I^{set}$  中, 随机组成选取 20 000 个大小为  $r \times r$  的局部块作为第 1 部分的网络输入, 进行无监督学习。然后, 以第 1 部分隐含层输出作为第 2 部分的网络输入。最后, 对每一个  $r \times r$  的局部块, 在训练样本集  $d^{set}$  中找到相对应的局部块, 以其中中心位置的场景深度作为标签, 进行有监督学习。神经网络中参数的学习通过反向传播算法实现, 其详细过程可参考文献[25]。如果采用不同大小的局部块进行训练, 可形成针对多种尺度的自动编码器。本文分别训练出针对  $11 \times 11$ 、 $15 \times 15$  和  $21 \times 21$  这 3 种尺度的稀疏自动编码器, 以便进行多尺度的纹理结构特征提取。

经训练后的稀疏自动编码器可看做为一个非线性滤波核, 不同尺度的编码器相当于不同大小的滤波核。对有雾图像进行纹理结构特征提取的过程,

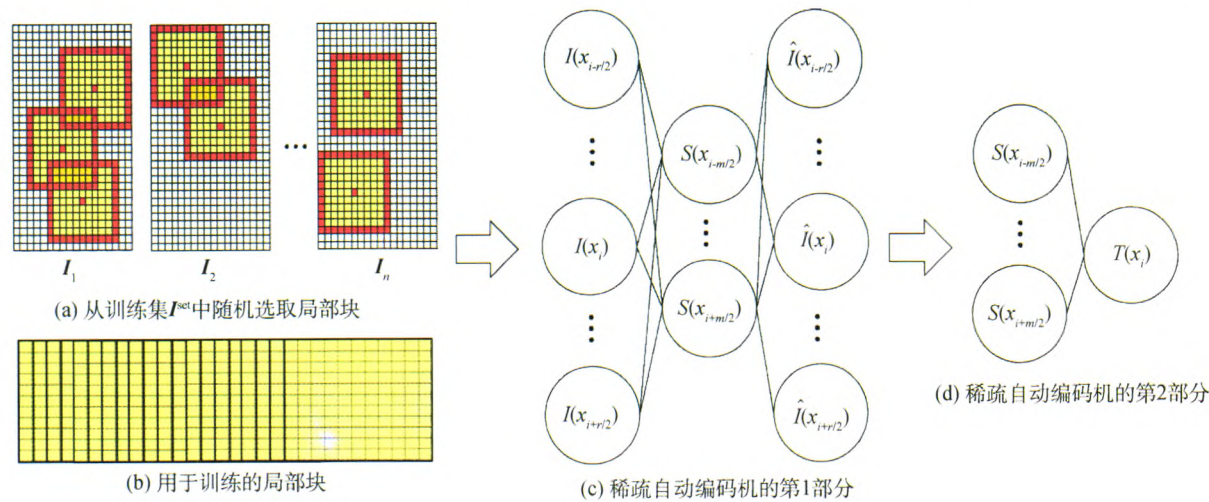


图 4 稀疏自动编码器及其训练过程

Fig. 4 Sparse autoencoder and its training process ((a) randomly selecting the local patches from the training data set  $I^{set}$ ; (b) local patches for training; (c-d) the first and the second parts of the sparse autoencoder)



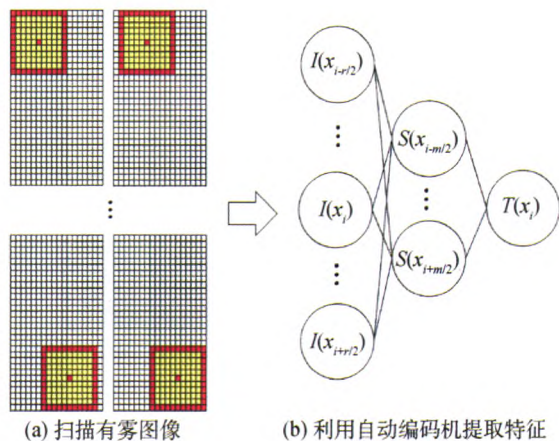


图5 有雾图像的纹理结构特征提取

Fig. 5 Extraction of textural and structural features for the hazy image ((a) scanning the hazy image; (b) feature extraction with autoencode)

实质是对有雾图像的一次非线性滤波。下面以尺寸为  $r \times r$  的稀疏自动编码器为例,给出该过程的详细步骤:如图5所示,给定一幅  $w \times h$  的有雾图像,采用一个  $r \times r$  的矩形框以一定的步长对其进行扫描。当步长为  $s$  时,对图像扫描后得到的局部块个数为

$$\left(\left\lfloor \frac{w-r}{s} \right\rfloor + 1\right) \times \left(\left\lfloor \frac{h-r}{s} \right\rfloor + 1\right) \quad (10)$$

式中,  $s$  不大于  $\lfloor r/2 \rfloor$ 。对于一个特定的局部块,设  $x_i$  表示其中心在有雾图像  $I$  中的位置,则该局部块与场景深度相关的纹理结构特征值为

$$S(x_k) = f\left(\sum_{j=i-\lfloor \frac{r}{2} \rfloor}^{i+\lfloor \frac{r}{2} \rfloor} I(x_j) W_{1j-(i-\lfloor \frac{r}{2} \rfloor)+1,k}^{(1)}\right) \quad (11)$$

$$T(x_i) = f\left(\sum_{j=i-\lfloor \frac{m}{2} \rfloor}^{i+\lfloor \frac{m}{2} \rfloor} S(x_j) W_{1j-(i-\lfloor \frac{m}{2} \rfloor)+1,1}^{(2)}\right) \quad (12)$$

式中,  $f(\cdot)$  为 sigmoid 函数  $W_{1p,q}^{(l)}$  表示自动编码器第  $l$  层网络中第  $p$  个神经元与第  $l+1$  层网络中第  $q$  个神经元之间的权值,  $S$  为稀疏编码器中隐含层的输出,  $T$  是与雾相关的纹理结构特征图,  $T(x_i)$  表示以  $x_i$  为中心的局部块的纹理结构特征值。与多尺度颜色特征类似,定义多尺度纹理结构特征  $T^3 = \{T_{11}, T_{15}, T_{21}\}$ , 其中  $T_{11}$ 、 $T_{15}$  和  $T_{21}$  分别表示在  $11 \times 11$ 、 $15 \times 15$  和  $21 \times 21$  这3种尺度的稀疏自动编码器下提取的纹理结构特征图,  $T^3$  是由特征图  $T_{11}$ 、 $T_{15}$  和  $T_{21}$  组成的集合。

图3(g)给出了有雾图像的纹理结构特征图  $T_{21}$ ,与图3(b)和图3(c)相比,该特征图能更准确

地反映雾的分布情况,从图中可以看出,山体到房屋之间存在着明显的景深突变。

## 2.4 场景深度估计

结合不同类型的特征,可有效估计有雾图像的场景深度,进而根据大气散射模型复原出无雾图像。首先假设各种特征与场景深度之间存在着一种非线性的映射关系,即

$$d(x) = f_1(f_2(\cdots f_n(F_1(x), \cdots, F_k(x)))) \quad (13)$$

式中,  $x$  表示图像中的一个特定位置,  $F_k$  表示第  $k$  个特征图,  $f_1(\cdot), \cdots, f_n(\cdot)$  是一系列未知的非线性函数,  $d(x)$  表示  $x$  位置上的场景深度。然后,建立多层神经网络模型,并利用样本集  $I^{\text{set}}$  和  $d^{\text{set}}$  对其进行训练以进一步确定  $f_1(\cdot), \cdots, f_n(\cdot)$ 。最后,给定一幅有雾图像及其不同类型的多尺度特征图,通过训练好的多层神经网络估计其对应的场景深度图。

图6给出了用于场景深度估计的神经网络模型及其训练过程。该网络包含3层(输入层、隐含层和输出层),输入层中神经元的个数与特征图的数目相等,隐含层中的神经元个数是输入层的2倍,输出层只有一个神经元,表示最终输出的场景深度。如图6所示,网络的训练过程较为直观:首先,分别对训练样本集  $I^{\text{set}}$  中的有雾图像  $I_1, I_2, \cdots, I_n$  进行特征提取,每一幅有雾图像  $I_i$  对应  $k$  幅特征图  $F_1, \cdots, F_k$  (实验中,  $k=14$ ,  $F_1, \cdots, F_{14}$  分别为2.1—2.3节中定义的  $R, G, B, C_1, C_4, C_7, C_{10}, D_1, D_4, D_7, D_{10}, T_{11}, T_{15}, T_{21}$ )。然后,以这些特征作为输入,以训练样本集  $d^{\text{set}}$  中的真实场景深度  $d_i$  作为标签,采用反向传播算法<sup>[25]</sup>对网络进行有监督学习,最后得到训练后的神经网络模型。

图7给出了不同特征组合在训练集上的均方误差直方图。观察图7可以看出,单独使用  $RGB, T^3 = \{T_{11}, T_{15}, T_{21}\}, C^4 = \{C_1, C_4, C_7, C_{10}\}$  和  $D^4 = \{D_1, D_4, D_7, D_{10}\}$  作为特征对模型进行训练时,具有较高的均方误差,分别为0.0616、0.0547、0.0428和0.0422。通过结合不同类型的特征对模型进行训练,可不同程度地降低均方误差,提高模型的性能。组合  $RGB, T^3, C^4$  和  $D^4$  这4种类型的特征训练后,能够有效地将模型的均方误差降至0.0379。

利用训练好的网络模型可复原出有雾图像的场景深度图。给定一幅有雾图像,采用2.1—2.3节的方法提取出14幅特征图  $F_1, \cdots, F_{14}$  (分别为  $R, G, B, C_1, C_4, C_7, C_{10}, D_1, D_4, D_7, D_{10}, T_{11}, T_{15}, T_{21}$ ), 计



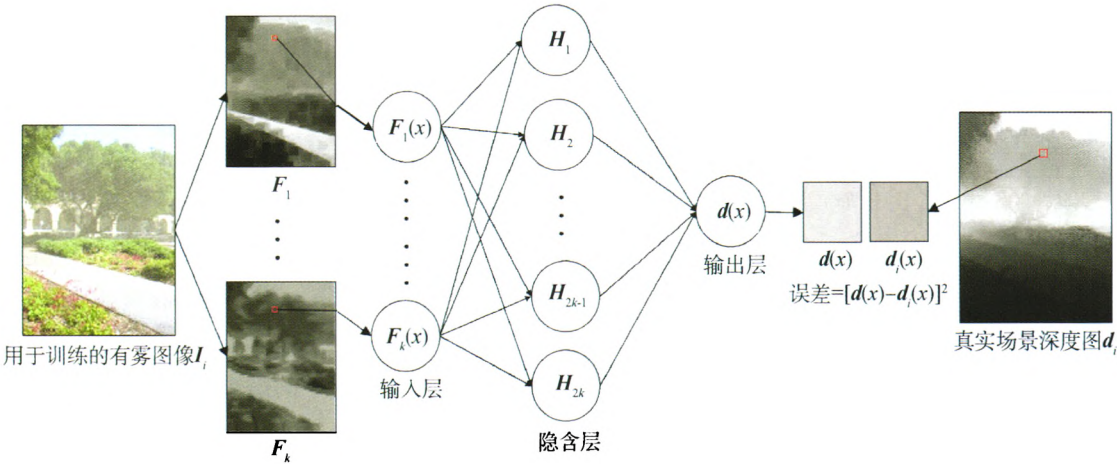


图6 用于场景深度估计的多层神经网络及其训练过程  
Fig. 6 Multi-layer neural network for depth estimation and its training process

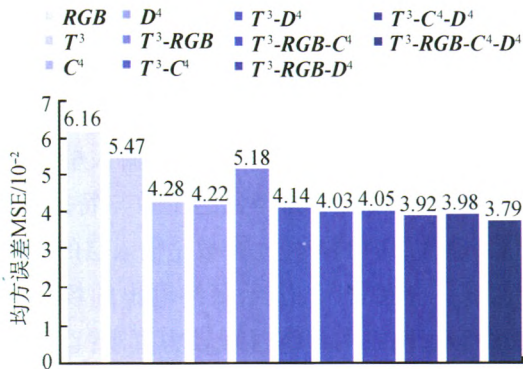


图7 不同特征组合在训练集上得到的均方误差  
Fig. 7 Mean square error of different combinations of features on the training set

算对应的场景深度图

$$H_j = f(\sum_{i=1}^k F_i(x) W_{2i,j}^{(1)}) \tag{14}$$

$$d(x) = f(\sum_{j=1}^{2k} H_j W_{2j,1}^{(2)}) \tag{15}$$

式中,  $W_{2p,q}^{(l)}$  表示训练后得到的第  $l$  层网络中第  $p$  个神经元与第  $l+1$  层网络中第  $q$  个神经元之间的权值,  $f(\cdot)$  是 sigmoid 函数,  $H_j$  为隐含层中的第  $j$  个输出,  $d$  是待求解的场景深度图。

图 3(h) 给出了该方法对图 3(a) 进行复原后的场景深度图。图中越亮的区域表示该处离摄像机越远, 对应的场景深度值越大。可以看出, 复原后的场景深度与真实情况较为接近。

3 图像去雾

在获得有雾图像的场景深度图之后, 可根据文献

[20] 的假设, 即同质条件下大气散射系数可看做为较小的常量, 令  $\beta=2.0$ , 并进一步通过式 (2) 和式 (4) 分别求得透射率  $t$  和大气光照度  $A$ 。在  $t$  和  $A$  已知的情况下, 可根据大气散射模型式 (1) 复原无雾图像

$$J(x) = \frac{I(x) - A}{\max(0.1, t(x))} + A \tag{16}$$

式中, 限制  $t(x)$  的最小值为 0.1 的目的在于避免透射率  $t(x)$  过小而产生较多噪声。

4 实验结果与分析

为了验证本文方法的有效性, 将实验结果与主流的去雾算法<sup>[7-8, 13-14]</sup> 作对比, 并进行定性、定量分析以及算法效率分析。本文的实验平台为 MATLAB R2014a, 程序的运行环境为 64 位 Windows 8.1, 计算机配置为 CPU Intel(R) Core(TM) i7-4710M 2.50 GHz, 内存为 16 GB。

4.1 定性分析

图 8 给出了 He 方法<sup>[7]</sup>、Tarel 方法<sup>[8]</sup>、Meng 方法<sup>[13]</sup>、Tang 方法<sup>[14]</sup> 及本文方法在自然图像上的去雾结果对比组图。实验 1 中, 各种主流算法都较好地去除了图像中的雾气, 但其去雾结果均存在不同程度的过饱和现象。例如, 经过去雾处理后, 女士的肤色明显加深了。出现过饱和的主要原因是单纯依赖颜色信息来估计透射率, 由于女士的肤色接近白色, 主流算法在估计透射率时, 误认为该区域具有较低的透射率, 最终造成局部区域去雾过度。而在本文的去雾结果中, 女士的肤色保持雪白。这是由于本文方法在估计透射率时, 不仅使用了传统的颜色



特征,同时还考虑了物体的纹理结构特征。另外,本文的结果中复原了更多纹理细节,例如图像右侧的黄色花朵清晰可见,相比之下,在 He 的结果及 Tang 的结果中几乎无法看清花朵的颜色。实验 2 中,主流算法的结果明显偏暗,图像中下方模糊一片,信息丢失严重,而在本文的结果中该区域清晰明亮。实

验 3 中,各种主流算法的结果明显偏蓝紫色,其原因在于各种主流算法在去雾时均以某种颜色特征作为唯一的线索,因此,在对存在轻微偏蓝色的图像进行去雾时,将误认为蓝色是物体原本的颜色,致使去雾后出现严重色偏。而在本文的结果中,景物的颜色自然逼真,具有较好的视觉效果。



图 8 不同算法在自然图像上的去雾效果图

Fig. 8 Results of different dehazing algorithms on real-world images ((a) the hazy images; (b) He's; (c) Tarel's; (d) Meng's; (e) Tang's; (f) our results)

为了更直观地比较不同算法的去雾效果,本文沿用文献[6,13-14]的评估方法,在带有标准结果的有雾图像上进行实验对比。实验的图像来源于用于评估立体匹配算法的标准数据库<sup>[26-28]</sup>,该数据库提供了大量清晰图像及对应的高精度场景深度图。利

用 2.1 节的方法生成一系列的有雾图像,并将清晰图像作为标准结果,应用不同的算法对有雾图像进行去雾,再通过对比去雾结果与标准结果来评估不同算法的去雾效果。图 9 给出了不同算法去雾结果与标准结果的部分对比组图。实验 1 中,Tarel 方法

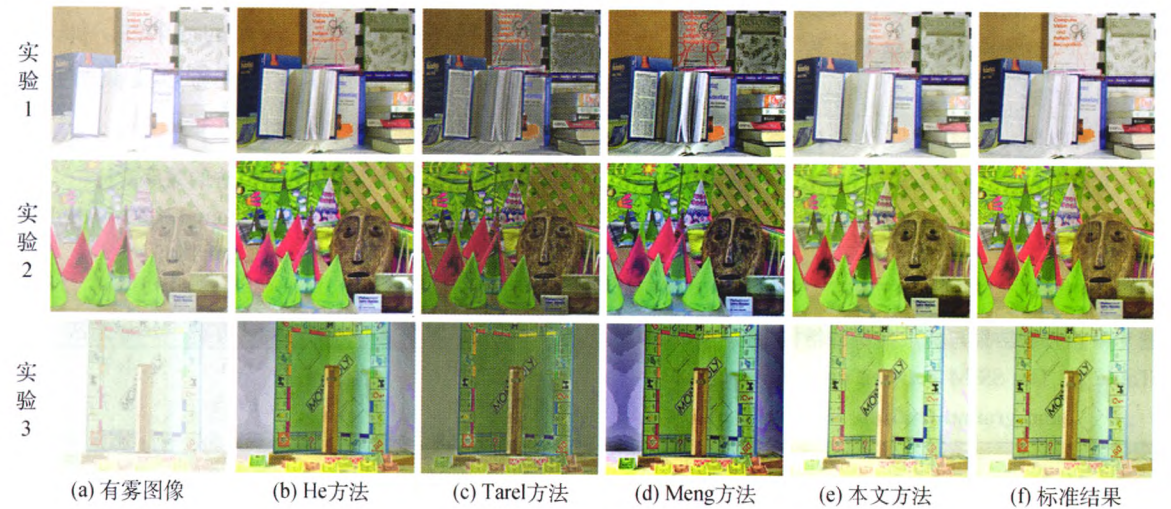


图 9 不同算法的去雾结果与标准结果的对比组图

Fig. 9 Visual comparison of different results on the stereos where the ground truth are known ((a) the hazy images; (b) He's; (c) Tarel's; (d) Meng's; (e) our results; (f) the ground truth haze-free images)



的结果和 Meng 方法的结果出现严重的过饱和,从图 9(c)和图 9(d)中可以看出,其背景的颜色明显过深,书籍封面及纸张均呈黑灰色,与标准结果差异较大。相比之下,He 方法的结果相对较好,但依然存在局部颜色过深的情况。而本文结果与标准结果最为接近。对于实验 2 而言,通过观察各算法结果中面具的颜色可以看出,Tarel 方法的结果和 Meng 方法的结果与标准结果差别较大,He 方法的结果则略佳于 Tarel 方法的结果和 Meng 方法的结果,而本文结果色彩相对更柔和,更接近真实场景。类似地,在实验 3 中,各主流算法的结果均存在背景过暗、失真严重的现象,而本文结果背景清晰明亮。

4.2 定量分析

为了对不同的去雾算法进行定量评估,采用均方误差(MSE)和结构相似度(SSIM)<sup>[29]</sup>两种指标来评价图 9 中各算法的去雾结果与标准结果的相似度。MSE 定义为

$$E = \sqrt{\frac{1}{3N} \sum_{c \in \{red, green, blue\}} \|J^c - G^c\|^2} \quad (17)$$

式中, $N$  为图像中像素点的个数, $J^c$  和  $G^c$  分别为去雾结果和标准结果在通道  $c$  上的分量。

表 1 和表 2 分别给出了图 9 中各去雾结果与标准结果的 MSE 和 SSIM。MSE 的值越小,表示去雾结果与标准结果越接近,SSIM 的值越大表示去雾结果与标准结果之间结构相似度越大,因此,去雾效果较好的结果应该具有较小的 MSE 和较大的 SSIM。

表 1 图 9 中各去雾结果与标准结果的均方误差 (MSE)

Table 1 The MSE between dehazed results and ground truth in Fig. 9

实验	He 方法	Tarel 方法	Meng 方法	本文方法
1	0.021 5	0.066 4	0.071 5	0.007 5
2	0.008 9	0.029 5	0.023 4	0.005 8
3	0.028 8	0.081 8	0.079 2	0.008 3
总体	0.019 7	0.059 2	0.058 0	0.007 2

表 2 图 9 中各去雾结果与标准结果的结构相似度 (SSIM)

Table 2 The SSIM between dehazed results and ground truth in Fig. 9

实验	He 方法	Tarel 方法	Meng 方法	本文方法
1	0.997 2	0.991 2	0.991 1	0.999 1
2	0.999 1	0.996 3	0.997 2	0.999 5
3	0.996 2	0.989 0	0.989 4	0.999 1
总体	0.997 5	0.992 2	0.9926	0.999 2

从表 1 和表 2 可以看出,Tarel 方法的结果在实验 1、实验 2 和实验 3 上都具有最高的 MSE 和最小的 SSIM,表明其与标准结果相差最大。而 Meng 方法的结果比 Tarel 方法的结果具有更小的 MSE 和更大的 SSIM,但差别不大,原因在于 Tarel 方法和 Meng 方法在该组实验上去雾过度,致使结果严重失真。相比之下,He 方法的结果的 MSE 显著小于 Tarel 方法的结果和 Meng 方法的结果,同时其 SSIM 也明显更大。这与定性分析中所得出的结论一致。与前 3 种主流算法的结果相比,本文方法结果的 MSE 明显更小,SSIM 更大,表明本文方法的去雾结果与标准结果更为接近。

4.3 算法效率分析

经过对本文算法的模型离线训练时间进行 20 次测量后,计算得出其平均训练时间为 104.33 s。在模型离线训练完毕的情况下,本文方法的时间复杂度为  $O(P^3N)$ ,其中  $P$  表示算法中使用到的最大扫描窗口大小, $N$  为图像的像素个数。为了进一步对比分析各去雾方法的效率,在不同尺寸的图像上对各方法进行测试,并记录其平均运行时间,结果如表 3 所示。从表 3 中可以看出,Tarel 方法受图像尺寸的影响较大,在图像尺寸较小的情况下,Tarel 方法所需的运行时间较小,但随着图像尺寸的增大,其运行时间急剧增大。总体而言,He 方法的效率相对更高,而 Meng 方法在不同图像尺寸的情况下均具有最高效率。相比之下,本文方法在效率上并不占优势,尤其是对于尺寸较大的图像,运行时间过长,这主要是由于在图像尺寸较大的情况下,所采用的多层神经网络本身具有较高的复杂性。加入并行计算和 GPU 加速等技术有望大幅度提高算法的运行效率,这将成为下一阶段的研究工作。

表 3 不同去雾算法的效率对比

Table 3 Time consumption comparison with different dehazing algorithm

图像尺寸/像素	He 方法	Tarel 方法	Meng 方法	本文方法
320 × 240	6.1	1.6	4.7	7.2
640 × 480	21.7	19.5	8.7	28.8
1 024 × 768	54.7	133.5	18.8	291.1
1 680 × 1 050	124.8	793.0	36.0	1 343.3
1 920 × 1 080	143.1	1 145.7	41.8	1 894.8

4.4 算法的局限性

尽管本文方法能够适用于大多数自然场景,但

对于雾气分布极度不均匀的图像,去雾效果并不理想,如图 10 所示。这是由于大气散射模型仍然相对较简单,不足以描述这种情况。然而,这将是所有基于物理模型的去雾方法需要解决共同难题。结合不同的气象模型<sup>[30]</sup>以建立一个更为鲁棒的物理模型,将会是进一步提高去雾效果的可行方法。



(a) 有雾图像



(b) 本文算法的去雾结果

图 10 本文算法对雾气分布不均图像的去雾效果  
Fig. 10 Our dehazing result of uneven hazed image  
( (a) the hazy image; (b) our dehazing result )

## 5 结 论

本文提出了一种特征学习的单幅图像去雾算法。该算法先采用稀疏自动编码器对有雾图像进行多尺度的纹理结构特征提取,同时抽取各种与雾相关的颜色特征。然后利用多层神经网络进行样本训练,学习得到雾天条件下纹理结构特征及颜色特征与场景深度间的映射关系,并估算出有雾图像的场景深度图。最后根据场景深度图复原出清晰的无雾图像。大量实验结果表明,与传统的单幅图像去雾算法相比,该算法能复原出质量更高的无雾图像,且有更好的场景宽适性。

## 参考文献 (References)

[ 1 ] Stark J A. Adaptive image contrast enhancement using generali-

zations of histogram equalization[ J ]. IEEE Transactions on Image Processing, 2000, 9 ( 5 ): 889-896. [ DOI: 10. 1109/83. 841534 ]

[ 2 ] Zhu P, Zhu H, Qian X M, et al. An image clearness method for haze[ J ]. Journal of Image and Graphics, 2004, 9 ( 1 ): 124-128. [ 祝培, 朱虹, 钱学明, 等. 一种有雾天气图像景物影像的清晰化方法[ J ]. 中国图象图形学报, 2004, 9 ( 1 ): 124-128. ] [ DOI: 10. 11834/jig. 20040120 ]

[ 3 ] Yang W T, Wang R G, Fang S, et al. Variable filter Retinex algorithm for foggy image enhancement[ J ]. Journal of Computer-Aided Design and Computer Graphics, 2010, 22 ( 6 ): 965-971. [ 杨万挺, 汪荣贵, 方帅, 等. 滤波器可变的 Retinex 雾天图像增强算法[ J ]. 计算机辅助设计与图形学学报, 2010, 22 ( 6 ): 965-971. ]

[ 4 ] Ma Z L, Wen J. Single-scale Retinex sea fog removal algorithm fused the edge information[ J ]. Journal of Computer-Aided Design and Computer Graphics, 2015, 27 ( 2 ): 217-225. [ 马忠丽, 文杰. 融合边缘信息的单尺度 Retinex 海雾去除算法[ J ]. 计算机辅助设计与图形学学报, 2015, 27 ( 2 ): 217-225. ]

[ 5 ] Tan R T. Visibility in bad weather from a single image[ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Alaska, America: IEEE, 2008: 1956-1963. [ DOI: 10. 1109/cvpr. 2008. 4587643 ]

[ 6 ] Fattal R. Single image dehazing[ J ]. ACM Transactions on Graphics, 2008, 27 ( 3 ): 1-9. [ DOI: 10. 1145/13995504. 1360671 ]

[ 7 ] He K, Sun J, Tang X. Single image haze removal using dark channel prior[ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33 ( 12 ): 2341-2353. [ DOI: 10. 1109/tpami. 2010. 168 ]

[ 8 ] Tarel J P, Hautiere N. Fast visibility restoration from a single color or gray level image[ C ] // Proceedings of IEEE 12th International Conference on Computer Vision. Washington DC, America. 2009: IEEE, 2009: 2201-2008. [ DOI: 10. 1109/iccv. 2009. 5459251 ]

[ 9 ] Wang Y F, Yin C L, Huang Y M, et al. Image haze removal using a bilateral filter[ J ]. Journal of Image and Graphics, 2014, 19 ( 3 ): 386-392. [ 王一帆, 尹传历, 黄义明, 等. 基于双边滤波的图像去雾[ J ]. 中国图象图形学报: 2014, 19 ( 3 ): 386-392. ] [ DOI: 10. 11834/jig. 20140307 ]

[ 10 ] Sun X M, Sun J X, Zhao L R, et al. An improved algorithm for single image haze removing using dark channel prior[ J ]. Journal of Image and Graphics, 2014, 19 ( 3 ): 381-385. [ 孙小明, 孙俊喜, 赵立荣, 等. 暗原色先验单幅图像去雾改进算法[ J ]. 中国图象图形学报: 2014, 19 ( 3 ): 381-385. ] [ DOI: 10. 11834/jig. 20140306 ]

[ 11 ] Ancuti C O, Ancuti C, Hermans C, et al. A fast semi-inverse approach to detect and remove the haze from a single image [ C ] // Proceedings of Asian Conference on Computer Vision.



- Queenstown, New Zealand; Springer, 2010: 501-514. [DOI: 10.1007/978-3-642-19309-5]
- [12] Zhu Q, Mai J, Shao L. Single image dehazing using color attenuation prior[C]//Proceedings of British Machine Vision Conference. Nottingham, Britain; BMVA, 2014. [DOI: 10.5244/C.28.114]
- [13] Meng G, Wang Y, Duan J, et al. Efficient image dehazing with boundary constraint and contextual regularization[C]//Proceedings of IEEE Conference on Computer Vision. Sydney, Australia; IEEE, 2013: 617-624. [DOI: 10.1109/ICCV.2013.82]
- [14] Tang K, Yang J, Wang J. Investigating haze-relevant features in a learning framework for image dehazing[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, America; IEEE, 2014: 2995-3002. [DOI: 10.1109/CVPR.2014.383]
- [15] He K, Sun J, Tang X. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1397-1409. [DOI: 10.1109/tpami.2012.213]
- [16] Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 228-242. [DOI: 10.1109/tpami.2007.1177]
- [17] McCartney E J. Optics of the atmosphere: scattering by molecules and particles[J]. IEEE Journal of Quantum Electronics, 1978, 14(9): 698-699. [DOI: 10.1109/JQE.1978.1069864]
- [18] Nayar S K, Narasimhan S G. vision in bad weather[C]//Proceedings of IEEE Conference on Computer Vision. Kerkira, Greece; IEEE 1999: 820-827. [DOI: 10.1109/iccv.1999.790306]
- [19] Narasimhan S G, Nayar S K. Vision and the atmosphere[J]. International Journal of Computer Vision, 2002, 48(3): 33-254. [DOI: 10.1109/1508044.1508113]
- [20] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(6): 713-724. [DOI: 10.1145/1508044.1508114]
- [21] Ashutosh S, Sung H C, Andrew Y N. Learning depth from single monocular images[C]//Proceedings of International Conference on Neural Information Processing System. Vancouver, British Columbia, Canada; MIT press 2005: 1161-1168.
- [22] Ashutosh S, Sung H C, Andrew Y N. 3-D depth reconstruction from a single still image[J]. International Journal of Computer Vision. 2008, 76(1): 53-69. [DOI: 10.1007/s11263-007-0071-y]
- [23] Ashutosh S, Sun M, Andrew Y N. Learning depth from single monocular images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 30(5): 824-840.
- [24] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. Journal of Machine Learning Research, 2010, 11: 3371-3408.
- [25] Barnard E. Optimization for training neural nets[J]. IEEE Transactions on Neural Networks. 1992, 3(2): 232-240. [DOI: 10.1109/72.125864]
- [26] Scharstein D, Szeliski R. High-accuracy stereo depth maps using structured light[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Madison, Wisconsin, America; IEEE 2003: 195-202. [DOI: 10.1109/CVPR.2003.1211354]
- [27] Scharstein D, Pal C. Learning conditional random fields for stereo[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, America; IEEE 2007: 1-8. [DOI: 10.1109/CVPR.2007.383191]
- [28] Hirschmüller H, Scharstein D. Evaluation of cost functions for stereo matching[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, America; IEEE, 2007: 1-8. [DOI: 10.1109/CVPR.2007.383248]
- [29] Wang Z, 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. [DOI: 10.1109/tip.2003.819861]
- [30] Preetham A J, Shirley P, Smits B. A practical analytic model for daylight[C]//Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. Los Angeles, America; ACM 1999: 91-100. [DOI: 10.1145/311535.311545]