封面

独创性 + 使用 声明

中文摘要

英文摘要

目录

1 绪论

1.1 研究背景及意义

1.2 研究现状

1.2.1 基于先验知识的去雾算法

1.2.2 基于神经网络的去雾算法

1.3 研究目的和研究内容

2 基于先验知识的去雾算法研究

2.1 暗通道

2.2 透射图估计

2.3 引导滤波

2.4 大气光估计

2.5 无雾图像计算

2.6 算法评价（如局限性、创新）

3 基于神经网络的去雾算法研究

3.1 DehazeNet

3.2 MSCNN

3.3 AOD-Net

3.4 Pixel2Pixel？

4 算法实现及评估

4.1 RESIDE数据集简介，生成简介，为什么不好获取图片对

4.2算法实现参数设置

数据准备、损失函数、优化器等等

4.3 算法评估

4.3.1 去雾效果

4.3.2 去雾速率

4.3.3 去雾指标

5 结论

5.1 总结（做了啥，结论是啥）

5.2 展望（研究改进 + 应用前景）  
1、atmospheric scattering model may be inadequate to represent the reality.

2、图片如何获取

致谢

参考文献

附录A 程序代码（mean filter, guided filter, DCP, DehazeNet, MSCNN, AOD-Net）

附录B 去雾示例

绪论

* 1. 研究背景和意义

信息化时代的一大特征即是海量数据的存储，例如图片、视频、音频、文本数据等。以图片数据为例，绝大部分的图片都采集于室外，但室外采集的图片相比于室内图片，更容易受到空中浑浊物，如水滴、悬浮颗粒的降质影响。降质图片无法反映原景物的对比度和真实颜色，降低观赏性，某些图片还会丢失重要信息，如交叉口抓拍图片。

图像去雾化处理是一个热门的研究领域。一方面，图像去雾能够有助于高层次的计算机视觉研究，如目标检测、无人驾驶等；另一方面，经过去雾处理的图片能够较好地纠正色差，更具观赏性。此外，在去雾处理中生成的深度图也会有助于图像编辑和其他视觉算法[1]。

含雾图片形成可以由大气散射模型来描述，这一模型最初是由McCartney[2]提出来的，之后由Narasimhan和Nayar[3, 4]进一步发展。模型表达式如下：

 （1）

其中，I是获取的含雾图片，J是对应的无雾图片，t是介质散射率，A是大气光值。去雾的目的即为，根据已有的I，获取对应的J，A，t。若介质时均质的，介质散射率可以表示为

（2）

其中，β是大气散射系数，为未知值，d为景深。

* 1. 研究现状

由大气散射模型可知，图像去雾的关键步骤在于估计出透射图和大气光。含雾图片的形成及其逆过程图像去雾都与景深有关。对于一张随意采集的图像，我们无法获取其景深，因此单幅图片的去雾是一个约束不足的问题。

针对这一困境，许多学者提出使用多幅图像获取附加信息进行去雾[3, 5, 6, 7, 8]，但这些方法的应用场景会受到限制。以自动驾驶为例，行驶中的车辆无法对同一处景物进行多次拍照，而多个摄像头同时拍照则无法保证拍照取景完全一致，难以获取同一景物的多幅图像；此外，此类算法需要对多幅图片进行处理，速率较慢，而自动驾驶车辆获取的数据量巨大，如特斯拉自动驾驶汽车8个摄像头每秒有2100帧输入图像，多幅图像去雾无法满足其处理速度要求。

理想的去雾算法可以对摄像头拍摄的图片进行实时去雾，因此单幅图片的高效、快速去雾成为研究的热点。目前，单幅图像去雾主要有两类较高效的方法：基于先验知识的去雾方法和基于神经网络的去雾方法，这两类方法都基于上述的大气散射模型。除此之外，还有不基于物理模型的图像增强方法，如直方图均衡化算法[9]、Retinex算法[10]等，但此类方法不考虑有雾图像的生成原因，直接对关注的细节进行增强，虽然简便易行，但容易丢失图像信息，使图像失真，本论文不对此类算法进行深入研究。

1.2.1 基于先验知识的去雾算法

Tan [11] 等人观察到无雾图像比其对应有雾图像的对比度高，于是采用了最大化对比度的方法来去除图像中的雾，该方法在视觉上能取得一定的效果，但是容易使图像过饱和及颜色失真。Fattal [12] 假设透射率局部不相关、反射率局部为常数，估计出景物的反射率并推导出透射率，进而计算出原图像，这一方法假设太强并且未考虑到景物深度的结构，导致无法处理浓雾图像并且不能准确估计景物深度。He [1] 等人在统计分析大量无雾图像后，于2009年提出了基于暗通道先验的去雾算法。这一方法非常简单，其去雾效果却非常好。He最初提出该方法时采用了软抠图的方法来优化透射率，但是该方法计算效率太低，其在2013年提出的引导滤波 [13] 方法可以大幅度提升计算效率。在He之后，有许多学者研究如何改进He的算法以使其适用于天空区域，并尝试优化透射率的估计以及计算效率。这一方法主要问题是对于含天空区域的图像，暗通道会失效，天空区域进行去雾处理后会失真。Zhu [14] 等人提出了基于颜色衰减先验的去雾方法，该方法提出了一个线性模型来表示颜色衰减先验下含雾图像中的景物深度，之后采用监督学习的方法学习到模型参数，从而获取图像中的深度信息，并以此来估计透射图从而给图像去雾。Berman [15]提出了基于非局部先验的去雾方法，该方法假设无雾图像中每个颜色团簇都是RGB空间中的一条雾线。

1.2.2 基于神经网络的去雾算法

随着卷积神经网络的兴起及其在计算机视觉领域的大规模应用，一些去雾算法也开始应用卷积神经网络。Cai [16] 等人提出了一个可训练的系统DehazeNet，用以估计含雾块的透射率，而后将此系统应用到整张图片以获取含雾图片的透射图，从而根据大气散射模型计算出无雾图片。Ren [17] 等人提出了一个多尺度卷积神经网络 (MSCNN) 用以直接估计与含雾图片对应的透射图，该方法先使用粗尺度模型对含雾图像进行处理，之后将此模型输出和原含雾图像作为细尺度模型的输入，而后使用细尺度模型输出的透射图进行无雾图像的计算。Li [18] 等人提出了一个端到端的去雾网络AOD-Net，该方法先将透射率t和大气光A重建为一个新的变量，而后使用两个模块构成的AOD-Net对含雾图像进行端到端的处理。

Pixel2pixel？

* 1. 研究目的和研究内容

本论文旨在探索图像去雾化的研究前沿，实现去雾算法，并对其效果进行对比。此外，本论文提出将去雾算法应用到交叉口视频监控以为交通管控提供更好的数据支撑。

本论文主要内容是实现去雾领域较为代表性的算法，例如暗通道法（DCP），DehazeNet, MSCNN, AOD-Net。第一章介绍去雾研究的背景意义，去雾研究的现状综述。第二章以暗通道法为例介绍基于先验知识的图像去雾算法，第三章以DehazeNet, MSCNN, AOD-Net为主要内容介绍基于神经网络的去雾算法。第四章介绍领域内广泛使用的数据集，以及上述各算法具体实现和其性能评估。第五章总结全文并提出后续研究方向和本研究应用前景展望。

基于先验知识的去雾算法研究

本章以暗通道法为例介绍基于先验知识的去雾算法，并对该算法进行描述性的评价。先验知识为不对问题进行具体处理即可知道的知识，即先于经验之前的知识。与之对应的是后验知识，即在有经验之后的知识。去雾处理中的先验知识，即不对图像进行去雾，只凭观察对比含雾图像或去雾图像就能获得的知识。在应用神经网络的去雾方法变得较为流行之前，单幅图像的去雾方法主要是基于不同的先验知识，如本章介绍的暗通道先验。

2.1 暗通道

暗通道是He [1] 等人观察分析大量无雾图片所得出的。他们发现，大多数户外无天空区域的图片块中存在某些像素点，这些像素点有至少一个很小的RGB分量值，也就是说这些图片块中最小的像素值很小，甚至接近于0。这一发现可由下式表示

 （3）

其中，J是一张图片，Jc是其中一个颜色通道，Ωx是以x为中心的一个图片块。他们观察到Jdark非常小，且接近于0，于是将Jdark称为图片J的暗通道，而将上式(3)称为暗通道先验。

暗通道的发现也有其物理背景。He等人提出，户外图片常常具有很多阴影，并且是具有多种颜色的，而阴影的存在会使得各通道像素值较小，具有鲜明颜色的物体则往往会缺少某一通道的颜色，使得该通道像素值较小，因此户外无天空区域图像的暗通道确实应该是暗的。理论分析之外，他们从网上搜集了5000张无雾风景图，裁剪掉天空区域并统计了他们的暗通道像素强度，统计数据也支持了他们提出的暗通道理论。

2.2 透射图估计

估计透射图之前，He等人先假设大气光值A已给定且为常数，并且假设以x为中心的局部区域内透射率为常数，记之为~tx，对大气散射模型(1)两端取最小值则有

(4)

两端同时除以大气光值有

(5)

在三个通道之内取最小值有

(6)

根据暗通道先验有

(7)

而大气光值Ac恒为正值，则有

(8)

联式(6)和(8)有

(9)

在含雾图片中，天空区域的颜色和大气光值很相近，因此由(9)计算出天空区域透射率为



而天空是在无限远处的，透射率应当为0，这一结果与(9)式计算结果一致，因此计算透射率时无需区分开天空区域和非天空区域。

在实际中，即便是无雾天气空中也会存在小颗粒，在远眺时图像仍会被降质。而倘使将小颗粒全部去除，即令~tx为0，则人眼无法感知景物深度，使看到的景物不真实，因此可在(9)中引入一个常数w以保留空中的部分颗粒，结果如下

(10)

其中，w取值为0.95。

2.3 引导滤波

He等人初始提出暗通道法时，采用的是软抠图方法对(10)计算出的透射图进行优化，但该方法计算效率太低，难以应对快速的去雾处理。其在2013年提出的引导滤波[13]方法能够较好地提升计算效率，同时保持了暗通道法较好的去雾效果。

2.3.1 滤波

一张图片可以视作一个函数，自变量为图片中像素点的位置，因变量为像素值的大小，由此我们可以得到与每张图片唯一对应的一个函数。其中，如是灰度图片，自变量为二元坐标值(x, y)，因变量f(x, y)则为处于0-255之间的一个整数值；如是彩色图片，自变量仍为二元坐标值(x, y)，因变量则为3\*1的向量，该向量分量值为每个颜色通道的像素值（0-255之间的整数）



滤波的作用即改变像素点位置对应的像素值，但不改变像素点的位置，是常用的图像增强、提取图像信息的手段。常用滤波器有线性时不变滤波器，如均值滤波、高斯滤波，这类滤波器有显式声明的核函数，不因图片内容产生变化。以下以均值滤波为例简要介绍滤波器原理，

这里拿一张含噪声图片作为例子，将代码作为附录

均值滤波算法简单，对图像处理速度较快，但在降低噪声的同时容易使图像变模糊，尤其是图像边缘部分。除线性时不变滤波器之外，有时候我们想利用图片的信息进行对图片进行滤波处理，因而需要使用一张引导图像，即引导滤波。

2.3.2 引导滤波

引导滤波的关键假设是引导图像I与滤波输出q之间存在线性关系，在以像素点k为中心的窗口wk中，滤波输出是引导图像的线性转换

(11)

其中，(ak, bk)是在窗口wk中为常数的线性系数。此滤波器使用以r为半径大小的窗口。

我们通过最小化滤波器输入p与输出q之间的差异来确定线性系数的大小，由此定义的损失函数为

(12)

其中ε是正则化系数，用以惩罚过大的系数ak。最小化上述损失函数，得

(13)

(14)

其中，uk和σk²是引导图像在窗口wk中的均值和方差，|w|是窗口内像素点的个数，而是窗口内被引导图像得均值。得到线性系数(ak, bk)值后，我们即可通过(11)式计算滤波器输出。

值得注意的是，图片中某个像素点i不为某一个窗口专属，因此在不同窗口下根据(11)式计算出的qi并不一样。对此，一个较为简单可行的方法是对qi所有可能值取平均。因此，在计算出每个窗口中的线性系数之后，我们可以通过下式计算滤波器输出

(15)

也即

(16)

其中，和 是包含像素点i的所有窗口中线性系数的均值。式(13), (14), (16)即是引导滤波的定义。

2.3.3 算法伪代码

由以上定义的引导滤波，我们给出算法的伪代码



其中，fmean是均值滤波，其窗口大小为r。corr是相关，var是方差，cov是协方差。

由以上介绍的算法，我们可以对2.2节中估计出的透射图进行优化，以此得到更精确的透射图。

2.4 大气光估计

He等人提出的暗通道可用以估计图片中雾的厚度：暗通道中亮度值越高，表明原图中该区域雾越厚。由此，He等人提出，先选取暗通道中前0.1%亮度值最高的像素点，在这些像素点中选取原图中的最高亮度值作为大气光A。

2.5 无雾图像计算

如前所述，根据大气散射模型(1)进行图像去雾处理，最关键步骤是计算出透射图和大气光。值得注意的是，在介质透射率较小甚至趋近于0时，根据模型获取的无雾图像易产生较大噪声，因此需要对透射率最小值进行限制，可由下式进行无雾图像的计算

(17)

其中，t0取值0.1。

2.6 算法评价（如局限性、创新）

He等人提出的暗通道法毫无疑问是单幅图像去雾的标志性工作。该方法原理非常简单，但是去雾效果非常好。在暗通道法提出后，许多学者对此方法表现出极大兴趣，并致力于改进该算法，以使其更好的适用于天空区域。

与此同时，我们也应当意识到这一方法的局限性，其一，是多次提到过的，该方法不适用于存在和天空颜色较为相似的物体的图片，因为在此类图片中，暗通道会失效。其二，是所有基于大气散射模型的去雾算法共有的局限性——该大气散射模型也许无法准确描述含雾图片的生成。

基于神经网络的去雾算法研究

本章着重介绍几种广为人知的基于神经网络的去雾算法，DehazeNet, MSCNN, AOD-Net。神经网络以其强大的非线性拟合能力在学界掀起了一股热潮，而卷积神经网络以其对图像的卓越处理能力成为了计算机视觉领域的标志性算法。本章将先介绍卷积神经网络的基本结构，如卷积层、池化层等，再深入讲解以卷积神经网络为基础的去雾算法。

3.1 卷积神经网络

卷积神经网络最早是由Yang LeCun提出的，但当时并未引起较大注意。之后的十几年，计算能力显著提高，CPU越来越快，GPU成为常用的计算工具，智能手机、摄像头的广泛应用使人们可以获取海量数据，大量的数据和不断提高的计算能力让神经网络的训练变得越来越易行。2012年由Alex Krizhevsky提出的AlexNet在ImageNet图像识别比赛中大幅获胜，激起了卷积神经网络的热潮，并由此拉开了深度学习的时代序幕。自那时起，卷积神经网络便成为了深度学习的一把利器。

典型的卷积神经网络一般含有卷积层、池化层和全连接层。

3.1.1 卷积层

卷积操作和上述线性时不变滤波器对图片进行的操作是类似的，如下图所示，其中I为需进行卷积操作的图层，K为卷积核，卷积结果为I \* K。卷积层一般用于提取前一层的局部特征，卷积核即特征提取器。



以彩色图片为输入层时，其大小为M \* N \* D，其中M为图片高度，N为图片宽度，D为图片深度，即颜色通道个数，彩色图片通道个数为3（D = 3）。不失一般性，不妨假设输入卷积层的张量大小为M \* N \* D，记为X，进行卷积操作后输出层张量大小为 M’ \* N’ \* P，记为Y，则卷积核张量大小为m \* n \* D \* P，记为W，其中，m，n为二维卷积核大小，P为输出特征图的个数。

卷积层计算输出特征图Yp时，用卷积核Wp（大小为m \* n \* D）对输入张量X进行卷积操作，加上偏置标量bp即得到卷积层的净输入Zp，Zp经过激活函数得到特征图Yp。其计算过程如下

(18)

(19)

其中，f为激活函数，常用的激活函数有tanh, ReLU, sigmoid, Leaky ReLU。一个卷积层需要P \* D \* （m \* n） + P个参数。

3.1.2 池化层

池化层，又叫下采样层，一般置于卷积层之后，目的是进行特征选择，减少特征数量。常用的池化函数有最大池化和平均池化，两者作用原理如下图所示。



池化层的一般操作是将每张特征图划分为2\*2的不重叠区域，对每个区域使用最大池化或平均池化函数。其中，划分区域及池化函数可以自定义，但过大的划分区域会导致信息的过分损失。

3.1.3 上采样层

上采样层与池化层对应，类似于池化层的逆操作。其过程如下图所示，将一个像素点的值扩充至以该像素点为中心的窗口，窗口大小可以自己定义，一般选择2\*2的窗口。



3.2 DehazeNet

由大气散射模型可知，图片去雾最关键步骤在于透射图的估计，由是Cai [16] 等人提出了一个端到端预测透射率的系统，该网络架构如下图所示。值得注意的是，该网络输入层是大小为16 \* 16 \* 3的图片块，输出结果是该图片块的透射率，这一性质将在第四章详细讨论。



3.2.1 特征提取层

特征提取层通过卷积操作和非线性激活函数实现。卷积操作使用16个大小为5 \* 5 \* 3的卷积核，不同于一般使用ReLU或tanh激活函数，这一层使用Maxout [19] 激活函数，并设置其参数为k=4。以下使用一个简单的例子讲解Maxout激活函数。



上图中，(a)表示传统的激活函数，(b)表示Maxout激活函数。(a)中，我们有

Y = f（W \* X + b）

其中，f为常用的激活函数，如tanh，ReLU，sigmoid，参数个数为 2 + 1 = 3个。(b)中，我们有

Y = max{W1 \* X + b1, W2 \* X + b2, W3 \* X + b3, W4 \* X + b4} (20)

其中参数个数有 (2 + 1) \* 4 = 12个。由此知，Maxout单元会成倍的增加参数个数，成倍因子为Maxout单元的参数k。

由(20)式可知，Maxout激活单元是分段的线性函数，因而可以拟合任意的凸函数，如下图所示。



3.2.2 多尺度映射

多尺度特征在图像去雾上被证明是较为有效的，于是本网络也采用了多尺度卷积。具体而言，三个尺度的卷积核大小分别为3\*3， 5\*5， 7\*7，每个尺度均选用16个卷积核。由于上一层输出特征组中有四张特征图，因此每个卷积核的深度均为4。

3.2.3 局部极值

局部极值采用最大池化，窗口大小为7 \* 7。上一层输出特征组有16 \* 3 = 48张特征图，大小均为12 \* 12。因此本层局部极值的输出大小为48 \* 6 \* 6。

3.2.4 非线性回归

本层在卷积操作之后采用双边线性整流函数(BReLU)作为激活函数。一方面，sigmoid激活函数容易导致梯度消失，进而导致较慢的收敛速度和不理想的局部极值；另一方面，ReLU激活函数比较适用于分类问题，不是非常适合回归问题，如这里的图像复原。受sigmoid函数和ReLU函数启发，DehazeNet提出使用BReLU激活函数，该函数可以保持双边约束以及局部线性。下图展示了BReLU和ReLU激活函数的对比，在本方法中，tmin和tmax分别取值为0和1。



本层仅使用一个卷积核，其大小为6 \* 6，因而输出张量大小为1 \* 1，经过BReLU激活函数之后输出一个0-1之间的标量值，即为该图片块的透射率值。

3.2.5 网络参数

下表显示了DehazeNet中各层的参数设置。在本网络中，卷积层的偏置恒为0，只设置权重参数，该网络中可训练参数共有16 \* 4 \* 5 \* 5 + 16 \* 4 \* （3\*3+5\*5+7\*7） + 48\*6\*6 = 8240个。



3.2.6 无雾图像计算

如前所述，DehazeNet的输入是16 \* 16 \* 3的图片块，输出该图片块的透射率值。在网络训练好之后，对含雾图像去雾时，需要先划分出图片块，然后预测每个图片块的透射率值从而得到整张图片的透射图，之后使用引导滤波的方法对该透射图进行优化。估计大气光时，需先选出透射图中前百分之0.1像素强度值的像素点位置，并在这些位置中选取最高像素值作为大气光值。最后按照大气散射模型采用下式计算无雾图像

(21)

3.3 MSCNN

与DehazeNet类似，Ren [17]等人提出了多尺度卷积神经网络(MSCNN)以预测透射图。该网络结构如下图所示。由图可知，该网络分为粗尺度网络和细尺度网络，粗尺度网络的输出作为细尺度网络的一个输入，该网络包含有卷积层、池化层、上采样层、线性组合层以及融合层。



3.3.1 网络各层介绍

输入层为彩色含雾图片，具有三个通道。

卷积层采用线性整流函数(ReLU)作为激活函数，并且含有偏置。

池化层至于卷积层之后，采用的下采样因子为2，也即窗口大小为2 \* 2。

上采样层采用2\*2的窗口。

线性组合层旨在融合最后一个上采样层的特征，并采用sigmoid激活函数。具体而言，该层将上采样层输出的特征图进行线性组合，并添加偏置，之后应用sigmoid激活函数。

细尺度网络将粗尺度网络的输出和第一个上采样层的输出融合，共同作为第二个卷积层的输入。

3.3.2 网络参数

本网络在训练时，需先训练粗尺度网络，后训练细尺度网络。粗尺度网络训练时，可训练参数个数为(5 \* 3 \* 11 \* 11 + 5) + (5 \* 5 \* 9 \* 9 + 5) + (10 \* 5 \* 7 \* 7 + 10) + (10 + 1) = 6321个，其中包含权重和偏置；细尺度网络训练时，可训练参数个数为(4 \* 3 \* 7 \* 7 + 4) + (5 \* 5 \* 5 \* 5 + 5) + (10 \* 5 \* 3 \* 3 + 10) + (10 + 1) = 1693个，包含权重和偏置。

3.3.3 无雾图像计算

由该网络结构可知，输出透射图与输入RGB图像具有同样的高度和宽度，因此无需对输入图像进行划分。在使用本网络进行无雾图像计算时，大气光值的估计与DehazeNet中估计方法一致，并按照下式进行无雾图像的计算

 (22)

3.4 AOD-Net

上述两种基于神经网络的方法都采用了同一种思路，即先估计透射图，后根据大气散射模型对图片去雾。这种思路存在两方面的问题，其一，这种思路并不是直接最小化去雾生成图片和无雾图片的差异，因而生成的去雾图片会与真实无雾图片存在较大偏差；其二，分别估计透射率和大气光值产生的误差会累积，进而加剧估计偏差。由此，Li [18]等人提出了AOD-Net，完全将去雾转化为一个端对端的问题，不需要任何的中间步骤估计参数，如大气光值。

3.4.1 问题重构

根据大气散射模型，无雾图片可以由以下方程计算得到

(23)

为了不用分开估计透射图t和大气光值A，Li等人将式(23)重写为下式



(24)

其中b为一个常数偏置，默认值为1。经过上述转换，我们便无需分开估计t和A，而是最小化上式输出J(x)和真实无雾图之间的误差。

3.4.2 网络结构



如上图所示，该网络结构分为两个模块，一个模块用来估计K（x），另一个模块用来逐元素计算生成无雾图片。前一个模块为全卷积网络，采用五个卷积层和三个融合层。其中，每一个卷积层都只使用三个卷积核，网络中使用不同大小的卷积核以获得多尺度特征，而使用融合层则用以弥补在卷积操作中丢失的信息。不同于DehazeNet，本网络中每个卷积层之后均使用ReLU激活函数。该网络的输入为含雾图片，输出为生成的无雾图片，训练以及预测未知含雾图片军比较简便。

3.5 本章小结

由上可知，基于神经网络的去雾算法有更大的灵活度，网络结构的选择层出不穷，但我们也应注意到，以上介绍的三种算法均依据大气散射模型进行去雾，是否存在不基于大气散射模型、端对端、高效快速的去雾算法，这不得而知。本论文中，我们着重研究各种方法的去雾效果，以及他们的泛化能力，并将这些方法应用到视频监控中以观察其实际效果。

算法实现及评估

本章我们具体讲解每种算法的实现，以及他们在统一测试数据上的表现。注意到，前面介绍的几种方法中，只有暗通道法是直接对含雾图像进行处理，不需要其对应的无雾图像，而另外几种方法除含雾图像外，均需要其他信息，如该图像对应无雾图像，或该图像对应透射图，或该图像中图片块对应的透射率值。在实际中，我们几乎不可能找到大量的含雾—无雾图片对，而任意给定一张含雾图像，我们也不知道该图像的透射率或投射图，因此，现存文献中均使用图片合成的方法对上述神经网络进行训练。

含雾图片的合成基于两个假设[20]：1) 图片内容与景物深度或介质透射率无关，也即同一图片内容可以出现在不同图片的不同深度；2) 局部范围内，景物深度相近，也即在一个较小的窗口里，该窗口中所含像素点对应的景物深度值相近。依据上述两条假设，我们可以采用下述方法生成含雾图片块：给定无雾图片块pj，以及大气光值A，我们随机生成透射率值t∈[0,1]，根据计算含雾图片块。同时，为减少参数学习的不确定性，设置大气光值为A=[1,1,1]。

以上生成含雾图片的方法影响了许多后续研究，这让大规模含雾—无雾图片对的获取成为可能，是一项开创性的工作。在此基础上，Li[21]等人提出了RESIDE数据集—一个去雾算法学习评估的基准数据集，下面会介绍。

此外，由第二条假设可知，一个较小的图片块具有相同的透射率值，我们可以通过图片块预估该图片块中所有像素点的透射率值，此即DehazeNet所采用的方法。

4.1 RESIDE[[1]](#footnote-1)数据集

RESIDE数据集旨在解决去雾领域含雾—无雾图片对难以获取的困境，通过大规模合成室内、室外含雾图片，为研究人员提供一个综合全面的去雾数据集。此外，该数据集提供了一个评价去雾算法的平台，由学者提出的去雾算法可以使用该数据集的数据进行训练、评估，从而界定各种算法的效果。评估算法时，该平台提出了一系列的评价指标，包括但不限于峰值信噪比PSNR、结构相似性SSIM，以及人眼评价去雾效果。

根据公式(1)，(2)可得出下式

(25)

由上式可知，含雾图片的生成需要无雾图片J，该图片对应的深度数据d，大气光值A，大气散射系数β。RESIDE数据集的生成使用了NYU2[22]和Middlebury stereo[23]深度数据集，并从该数据集获取了对应的无雾图片，因而获取了J和d的信息。在此基础上，Li等人在[0.7, 1.0]之间对大气光值A随机取值，在[0.6, 1.8]之间对β随机取值，从而生成含雾图片。

在该数据集网站上可以看到，RESIDE数据集含有五个共训练测试的子集，分别为：室内训练集ITS，共110000张含雾图片；室外训练集OTS，共313950张含雾图片；客观合成测试集SOTS，共1000张含雾图片；应用驱动测试集RTTS，共4322张；综合主观测试集20张。该数据集数据量之大，图片内容之全，以及评价指标之多，均为去雾领域之最。

4.2算法实现与参数设置

本节我们讲解各算法的具体实现及其参数设置。对于部分算法，作者在提出算法时已经给出参数值，我们不做更改。而另外一部分算法，作者并未给出全部参数的参考值，因此我们自行调优给出。

4.2.1 暗通道法 (DCP)

暗通道法是直接对含雾图片进行处理的算法，不需要对应的无雾图片，也不需要对图片大小进行修改，算法输入是任意的含雾图片。算法具体实现步骤为：

1. 根据公式(3)计算图像暗通道
2. 根据公式(10)计算透射图
3. 根据引导滤波算法优化透射图
4. 根据2.4提出方法估计大气光值
5. 根据公式(17)计算无雾图像

其中，计算暗通道时需给定窗口大小，我们设置该大小为16\*16。引导滤波中使用盒子滤波的半径值设为40，正则化参数设为0.001。大气光最大值设置为220

4.2.2 DehazeNet

如前所述，DehazeNet输入数据为16\*16\*3的图片块，输出为该图片块的透射率值。我们采用RESIDE数据集的ITS子集作为训练数据。值得注意的是，ITS中给出的数据是含雾图片和该图片对应的投射图，因此我们需要对该图片进行划分。以一张320 \* 240的彩色图片及其对应投射图为例，我们用16\*16的窗口在该图片及其对应投射图上不重复随机采样，将该采样图片块作为一个样本。由于样本标签是该采样图片块对应透射率值，而我们采样出投射图的图片块是一个16\*16的数值矩阵，因此我们对投射图图片块取均值作为样本标签。经此处理，一张320\*240的彩色图片可以生成300个无关联样本（因为我们采用不重复随机采样）。由于训练样本数据大小为16\*16\*3，与ITS数据集中图片大小无关，因此我们不需要统一图片大小尺寸，将图片高度和宽度均调整为距自身最近的16的整数倍即可。原作者训练网络时使用了100 000个样本，本论文一方面扩大训练数据量，另一方面考虑到机器性能以及运行时间，最终决定在ITS数据集中随机抽取10 000张图片生成样本进行训练。

DehazeNet使用均方误差作为损失函数

(26)

其中，F(I)为网络输出，t为实际透射率值，N为使用图片块个数。本网络使用随机梯度下降优化器，momentum设置为0.9，decay设置为5e-4，学习率初始化为0.005，每训练10个epoch让学习率减半，共训练40个epoch。权重初始值从均值为0，标准差为0.001的正态分布中随机取值，偏置恒设置为0，共8240个参数。训练过程中，我们按4:1的比例划分训练集和交叉验证集，并将batch size设为40。

训练好DehazeNet之后，我们采用划分图片块的方法对测试图片进行预测，并使用引导滤波优化投射图，之后如前所述估计大气光值并计算无雾图像。

4.2.3 MSCNN

MSCNN的输入为含雾图像，输出为该图像对应投射图。我们使用RESIDE数据集中的ITS子集对该网络进行训练。原作者训练时使用了6 000张无雾图片并生成18 000个样本作为训练集，本论文采用更大的数据量并考虑到机器性能和处理时间，最终从ITS数据集中随机抽取30 000张图片用作训练。此外，该网络在训练时将所有图片统一尺寸为320\*240。本论文采取类似做法。

MSCNN按下式计算损失函数

(27)

其中，q为含雾图片张数，tix为网络生成投射图，tix\*为实际使用的投射图。值得注意的是，本网络中粗尺度网络和细尺度网络是分开训练的，两次训练均采用此损失函数。网络在训练时，先训练好粗尺度网络，而后将粗尺度网络的输出作为细尺度网络的一张特征图，继续训练细尺度网络。本网络使用随机梯度下降优化器，初始学习率设置为0.001，momentum设置为0.9，decay设置为5e-4，每十个epoch让学习率减至原值的10%，共设置四十个epoch。训练过程中，我们按4:1的比例划分训练集和交叉验证集，并设置batch size为48。

训练好MSCNN之后，我们用之对测试图片进行预测，并按照传统方法估计大气光值，进而计算无雾图像。MSCNN对传入图片的限制是该图片的高度值和宽度值均为偶数，基于此，我们将测试图片大小调整为距自身最近的偶数值，以尽量保存测试图片信息。

4.2.4 AOD-Net

AOD-Net的输入为含雾图片，输出为去雾图片。我们使用RESIDE数据集中的OTS子集对该网络进行训练。原作者训练时使用27 256个样本，我们采用更大的数据量并考虑机器性能和处理时间，最终从OTS数据集中随机选取40 000张图片用于训练。AOD-Net是一个全卷积网络，因此对输入图片大小没有限制，但是考虑到我们训练时采取分批次的方法，而每一批次的数据应当具有相同大小，于是我们将所有训练数据统一尺寸为320\*240。训练过程中，我们按4:1的比例划分训练集和交叉验证集，并设置batch size为32。

AOD-Net使用MSE作为损失函数，使用随机梯度下降优化器，其中momentum设置为0.9，decay设置为0.0001，并且将梯度值限制在[-0.1, 0.1]之间。本网络采用高斯分布随机变量初始化权重。

网络训练好之后，即可直接对测试图片进行去雾。由于该网络是全卷积网络，因此不用调整图片大小。

4.3 算法评估

本节我们对上述介绍的算法进行评估。我们使用SOTS集对各算法进行测试，分析对比由算法计算出的各指标，之后观察各算法在实拍图片上的效果。最后，我们将各算法应用于雾天监控视频/行车记录仪拍摄视频，以观察应用效果。

所有算法均用python/keras实现，在同一台Linux机器上运行，机器CPU为Intel(R) Core(TM) i7-6700k 4GHz, 24GB RAM，Nvidia GeForce GTX 1080 Ti GPU, 8G显存。

4.3.1 去雾指标介绍

图片在获取、处理、压缩、储存或传输时都有可能引入噪声，本论文研究的雾属于噪声的一种。本论文对图片的处理最终是以人为受众的，因此处理效果由由受众主观评价最为准确。然而，主观评价需要花费大量的时间精力，且评价结果极易收到评价者的影响。同一评价者在不同的环境下有可能对同一张图片给出不同的评价结果，因而主观评价是不稳定的。

由此，学者们展开了客观评价指标的研究。客观指标可以分为三类[24]：全参考指标，即可以获得真实图片时对处理图片进行评价的指标；无参考指标，即无法获取真实图片时进行评价的指标；部分参考指标，旨在能获取部分真实图片信息时对处理后图片进行评价。为尽可能全面评价去雾算法，我们采用全参考指标PSNR，SSIM和无参考指标SSEQ，BLIINDS-II。下面对各指标逐一介绍。

1) 峰值信噪比PSNR

在通信领域，信噪比是指信号功率与噪声功率之比，峰值信噪比表示信号最大可能功率与噪声功率之比值。在图像处理领域，峰值信噪比则通过均方误差MSE进行定义，即像素点最大像素值与MSE的比值。

假设有一张去雾图像I和其对应的无雾真实图像J，其大小为m \* n \* 3，其中3是颜色通道个数。则均方误差定义为

 前面还有个三分之一，K改成J(28)

峰值信噪比为

(29)

其中，MAXI表示图像点像素最大值。由上可知，峰值信噪比越大，图像失真越小，去雾效果越好。

MSE和PSNR的计算非常简便，也具有清晰的物理含义，因此被广泛应用于重建图像的评估，但这两个指标对图像的的评价结果与人眼的评价结果不完全一致，因此存在应用的局限性。

2) 结构相似性SSIM

与PSNR不一样，SSIM更强调与评价指标与人眼判断具有较好的一致性。人眼在观看图片时，会潜意识将图片内容结构化，也即认为相邻像素之间是存在关联的，因此Wang[24]等人在设计SSIM时着重考虑了图片间的结构性信息。

Wang[24]等人设计的SSIM系统将图像相似性的测量分为亮度、对比度以及结构的测量，按照下式进行SSIM的计算

(30)

(31)

(32)

(33)

（维基）

用SSIM评估图像相似度时，计算的SSIM越大，则两张图片相似度越高。在实际使用时，为简化上述表达式，一般令α=β=γ=1并使C3 = 3/2 C2，得

(34)

此外，由于视觉的局部性，我们使用固定大小的窗口在图片上滑动，计算每个窗口的SSIM值，并使用高斯加权函数对每个窗口值进行加权平均以防止出现块状现象。

3) SSEQ

4) BLIINDS-II

4.3.1 算法评估

1）四个指标的表格

2）实拍图片的示例

3）视频的示例

4.3.2 去雾速率

比较测试时候的单张图片速率即可

需要与作者给出的代码进行对比，不能比他们的差，直接看他们论文或RESIDE里面给出的结果，不用实现他们的算法。而且使用的数据不一样，不具对比性。

DehazeNet代码得改，scheduler

训练集，OTS选一部分，ITS选一部分？

视频去雾处理 行车记录仪 + 监控视频

<https://www.youtube.com/watch?v=VSrQ1qMqEac>

5 结论

5.1 总结（做了啥，结论是啥）

5.2 展望（研究改进 + 应用前景）  
1、atmospheric scattering model may be inadequate to represent the reality.

2、图片如何获取

致谢

参考文献

附录A 程序代码（mean filter, guided filter, DCP, DehazeNet, MSCNN, AOD-Net）

附录B 去雾示例

[1] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” in IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[2] E. J. McCartney, “Optics of the atmosphere: scattering by molecules and particles,” New York, John Wiley and Sons, Inc., 1976

[3] S. G. Narasimhan and S. K. Nayar. Chromatic framework for vision in bad weather. In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, volume 1, pages 598-605. IEEE, 2000.

[4] S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. International Journal of Computer Vision, 48(3):233-254, 2002.

[5] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. PAMI, 25:713–724, 2003.

[6] S. K. Nayar and S. G. Narasimhan. Vision in bad weather. ICCV, page 820, 1999.

[7] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. CVPR, 1:325, 2001.

[8] S. Shwartz, E. Namer, and Y. Y. Schechner. Blind haze separation. CVPR, 2:1984–1991, 2006.

[9] Kim J Y, Kim L S, Hwang S H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(4): 475−484

[10] Parthasarathy S, Sankaran P. A RETINEX based haze removal method[C]//2012 IEEE 7th International Conference on Industrial and Information Systems (ICIIS). IEEE, 2012: 1-6.

[11] R.Tan. Visibility in bad weather from a single image. CVPR, 2008.

[12] Fattal R. Single image dehazing [J]. ACM transactions on graphics (TOG), 2008, 27(3): 72.

[13] He K, Sun J, Tang X. Guided image filtering [J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(6): 1397-1409.

[14] Q. Zhu, J. Mai, and L. Shao, “A fast single image haze removal algorithm using color attenuation prior,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3522–3533, 2015.

[15] D. Berman, S. Avidan et al., “Non-local image dehazing,” in IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[16] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, “Dehazenet: An end-to-end system for single image haze removal,” IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187–5198, 2016.

[17] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, “Single image dehazing via multi-scale convolutional neural networks,” in European Conference on Computer Vision, 2016.

[18] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “Aod-net: All-in-one dehazing network,” in IEEE International Conference on Computer Vision, 2017.

[19] I. Goodfellow, D. Warde-farley, M. Mirza, A. Courville, and Y. Bengio, “Maxout networks,” in Proceedings of the 30th International Conference on Machine Learning (ICML-13), 2013, pp. 1319–1327.

[20] K. Tang, J. Yang, and J. Wang, “Investigating haze-relevant features in a learning framework for image dehazing,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2995– 3002.

[21] B. Li *et al*., "Benchmarking Single-Image Dehazing and Beyond," in *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492-505, Jan. 2019.

[22] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, “Indoor segmentation and support inference from rgbd images,” in European Conference on Computer Vision. Springer, 2012, pp. 746–760.

[23] D. Scharstein and R. Szeliski, “High-accuracy stereo depth maps using structured light,” in Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on,vol.1. IEEE, 2003, pp. I–I.

[24] Z. Wang, A. C. Bovik, H. R. sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transaction on Image Processing, 2004, 13(4):600–612

参考文献格式、正确性全部需要检查

1. <https://sites.google.com/view/reside-dehaze-datasets> [↑](#footnote-ref-1)