



华南理工大学
South China University of Technology

博士学位论文

图像去雾与增强算法的研究

作者姓名	蔡博仑
学科专业	信息与通信工程
指导教师	徐向民 教授
所在学院	电子与信息学院
论文提交日期	2019 年 4 月 09 日

Researches on Image Dehazing and Enhancement

A Dissertation Submitted for the Degree of Doctor of Philosophy

Candidate: Cai, Bolun

Supervisor: Prof. Xu, Xiangmin

South China University of Technology

Guangzhou, China

分类号: TP391

学校代号: 10561

学 号: 201610101563

华南理工大学博士学位论文

图像去雾与增强算法的研究

作者姓名: 蔡博仑

指导教师姓名、职称: 徐向民 教授

申请学位级别: 工学博士

学科专业名称: 信息与通信工程

研究方向: 计算机视觉

论文提交日期: 2019 年 4 月 09 日

论文答辩日期: 2019 年 5 月 30 日

学位授予单位: 华南理工大学

学位授予日期: 年 月 日

答辩委员会成员:

主席: 李定海

委员: 赵建 殷瑞祥 杨春晓 梅文

摘 要

雾是一种常见的大气现象，空气中悬浮的水滴、灰尘、细沙或其他颗粒等都会引起成像清晰度的降低。在雾天的成像过程中，远处物体的反射光无法穿过稠密大气到达摄像头，大气散射导致成像对比度的下降。因此，图像去雾已经成为图像处理和机器视觉领域的研究重点和热点。本文主要研究图像/视频的去雾和增强，主要贡献如下：

1) 图像去雾的关键是大气透射率的准确估计。本文提出了可训练的端到端系统 DehazeNet 来估计透射率图，并通过大气散射模型重构清晰图像。DehazeNet 采用特殊的卷积神经网络，其中利用 Maxout 单元实现与雾相关的特征提取，并提出双边修正线性单元 (BReLU) 以提高重构质量。实验表明，相比于现有方法 DehazeNet 具有更好的性能，同时保持了高效性和易用性。

2) 视频去雾具有更广泛的应用场景，但其额外的挑战主要来自时空相关性和计算实时性。本文提出亮度值先验并建立基于时空马尔可夫随机场的实时视频去雾框架。此外，该框架采用积分图像技术和下采样技术进行加速，显著地降低了主要的计算负担。实验结果表明，本文方法在保持实时性的同时，可有效地去除区块效应和闪烁伪影：在单 CPU 下实现对 352×288 尺寸视频流约 240 帧每秒的实时处理。

3) 在大气散射模型中，非灰白色的大气光照会导致图像的颜色失真。当在 RGB 颜色空间中分解出光强度，反射率仅保留物体的本征颜色，因此 Retinex 分解可用于颜色校正。本文提出了一种联合先验的 Retinex 模型，实现光强度和反射率的同时估计。场景中三维物体在二维空间的成像受其内在属性（形状和纹理）和外在属性（光照）的影响。与现有方法相比，联合先验模型可以通过形状先验保存结构信息，通过纹理先验估计精细的反射率，通过光照先验捕获环境光源。

4) 浓雾会导致图像严重的对比度下降和细节丢失，信息退化可通过基于尺度感知平滑的多尺度细节增强来改善。无论是在细节还是结构区域，强的边缘梯度使边缘感知滤波器自然地保留它们。对此，本文提出了一种基于高斯相对性 (RoG) 的尺度感知方法——RoG 通过对尺度特征进行局部分析，并全局优化得到分段平滑的滤波结果。此外，本文引入可分的递归优化方法将全局优化的计算复杂度降低到 $O(N)$ 。

关键词： 图像/视频去雾；卷积神经网络；时空马尔可夫随机场；颜色校正；Retinex；细节增强；高斯相对性。

Abstract

Haze is a traditional atmospheric phenomenon where dust, smoke and other dry particles obscure the clarity of the atmosphere. Haze causes issues in the area of terrestrial photography, where the light penetration of dense atmosphere may be necessary to image distant subjects. This results in the visual effect of a loss of contrast in the subject, due to the effect of light scattering through the haze particles. For these reasons, haze removal is desired in both consumer photography and computer vision applications. This paper focuses on image/video dehazing and enhancement, and the main contributions are as follows:

1) The key to image dehazing is to estimate the medium transmission. In this paper, a trainable end-to-end system called DehazeNet is proposed to estimate the medium transmission map that is subsequently used to recover a haze-free image via atmospheric scattering model. DehazeNet adopts convolutional neural networks, whose layers are specially designed to embody the established dehazing priors. Specifically, layers of Maxout units are used for feature extraction, and Bilateral Rectified Linear Unit (BReLU) is proposed to improve recovering quality. Experiments show that DehazeNet achieves superior performance over existing methods, yet keeps efficient and easy to use.

2) Video dehazing has a more wide range of real-time applications, but additional challenges mainly come from spatio-temporal coherence and computational efficiency. In this paper, a spatio-temporal Markov random field is built with an intensity value prior for real-time video dehazing. Moreover, to facilitate real-time applications, integral image and down sampling technique are approximated to reduce the main computational burden. Experimental results demonstrate that the proposed method is effectively to remove haze and flickering artifacts, and sufficiently fast for real-time applications.

3) In atmospheric scattering model, non-gray airlight will result color distortion. When the illumination is decomposed in the RGB-color space, the reflectance retains the original color information of the object, meaning that the Retinex decomposition has the effect to correct color distortion. In this paper, a joint intrinsic-extrinsic prior model is proposed to estimate both illumination and reflectance. The 2D image formed from 3D object in the scene is affected by the intrinsic properties (shape and texture) and the extrinsic property (illumination). Better than

conventional Retinex models, the proposed model can preserve the structure information by shape prior, estimate the reflectance with fine details by texture prior, and capture the luminous source by illumination prior.

4) Heavy haze results contrast decreasing and detail losing, which can be improved by multi-scale detail enhancement based on scale-aware smoothing. Both in detail and structure regions, the gradient magnitudes are high enough to result edge-aware filters preserving them naturally. This paper presents a novel scale-aware image smoothing via Relativity-of-Gaussian (RoG). As a simple local measure, RoG performs the local analysis of scale features and globally optimizes its results into a piecewise smooth. In particular, a separable recursive optimization is introduced to improve the computational complexity to $O(N)$ and achieve a fast speed.

Keywords: Image/video dehazing, CNN, spatio-temporal MRF, color correction, Retinex, detail enhancement, Relativity-of-Gaussian.

目 录

摘要	I
Abstract	II
表格目录	VII
插图目录	VIII
第一章 绪论	1
1.1 引言	1
1.2 研究背景和意义	2
1.3 相关研究现状	3
1.3.1 基于图像增强的方法	3
1.3.2 基于 Retinex 的方法	4
1.3.3 基于物理模型的方法	6
1.4 主要研究内容	8
1.5 本文结构安排	9
第二章 基于卷积神经网络的图像去雾	11
2.1 概述	11
2.2 去雾卷积网络 DehazeNet	13
2.2.1 DehazeNet 的网络结构	13
2.2.2 DehazeNet 与传统方法的联系	16
2.2.3 DehazeNet 的训练过程	19
2.3 实验与分析	20
2.3.1 网络结构分析	21
2.3.2 滤波器个数与尺度	23
2.3.3 合成图像块的定量分析	24
2.3.4 合成图片的定量分析	25
2.3.5 RESIDE 数据库的定量分析	28
2.3.6 真实图片的定性分析	29
2.4 小结	31

第三章 基于时空马尔科夫随机场的视频去雾	33
3.1 概述	33
3.2 基于 ST-MRF 的实时视频去雾	34
3.2.1 单帧图像去雾	34
3.2.2 实时视频去雾	36
3.3 实验与分析	40
3.3.1 时间连续性分析	40
3.3.2 合成视频的定量分析	42
3.3.3 真实视频的定性分析	43
3.3.4 算法实时性分析	43
3.4 小结	45
第四章 基于联合先验 Retinex 模型的颜色校正	46
4.1 概述	46
4.2 局部变化方差	48
4.2.1 边缘/结构保持平滑	48
4.2.2 局部变化方差	50
4.2.3 理论分析	50
4.3 联合内外先验模型	51
4.3.1 形状先验	52
4.3.2 光照先验	53
4.3.3 纹理先验	54
4.3.4 联合优化	54
4.4 实验与分析	55
4.4.1 Retinex 分解	56
4.4.2 颜色校正	57
4.4.3 亮度自适应	59
4.4.4 先验分析	60
4.5 小结	62
第五章 基于高斯相关性 (RoG) 的细节增强	64
5.1 概述	64

5.2 基于高斯相关性的图像平滑	66
5.2.1 高斯相关性	66
5.2.2 平滑优化	67
5.2.3 理论分析	69
5.3 快速可分递归优化	71
5.3.1 一维递归优化	71
5.3.2 二维可分近似	72
5.4 实验与分析	73
5.4.1 细节增强	73
5.4.2 结构提取	75
5.4.3 JPEG 伪影消除	76
5.4.4 HDR 色调映射	78
5.4.5 其他应用	80
5.4.6 运算效率分析	81
5.5 小结	82
第六章 总结与展望	84
6.1 全文总结	84
6.2 学术贡献	85
6.3 工作展望	85
参考文献	87
攻读博士学位期间取得的研究成果	100
致谢	101

表格目录

2-1	DehazeNet 的参数设置	20
2-2	DehazeNet 在不同滤波器参数下的性能	24
2-3	在合成图像块中不同方法的大气透射率估计	25
2-4	在合成图片中不同方法的性能评估	26
2-5	在合成图片中不同方法的鲁棒性评估	27
2-6	不同雾浓度下在 RESIDE 数据库中 SOTS 的评估	29
2-7	在 RESIDE 数据库中 HSTS 的评估	29
3-1	不同视频去雾方法的平均亮度相关性分析	42
3-2	在合成视频中不同方法的定量分析	43
3-3	不同视频去雾方法的实时性分析	45
4-1	在 Color-Checker 中的颜色校正对比	59
4-2	在 35 幅测试图像中光照自适应的定量对比	62
4-3	在 35 幅测试图像中不同先验对光照自适应的影响	63
5-1	运算效率对比	82

插图目录

1-1	雾天和晴天的成像对比	2
1-2	颜色直方图与去雾增强	4
1-3	颜色恒常性	5
1-4	论文结构关系图	10
2-1	雾天成像系统与大气散射模型	12
2-2	DehazeNet 的网络结构	13
2-3	修正线性单元 (ReLU) 和双边修正线性单元 (BReLU)	15
2-4	去雾相关的特征	16
2-5	DehazeNet 中第一层的滤波器权值与 Maxout 单元	18
2-6	互联网上采集到的清晰图片示例	19
2-7	F_1 在不同降维映射下的训练过程	22
2-8	F_4 在不同激活函数下的训练过程	23
2-9	F_4 在不同激活函数下的回归曲线	23
2-10	基于立体图像合成的有雾图像与 DehazeNet 的去雾结果	25
2-11	DehazeNet 对光晕图像的增强	27
2-12	DehazeNet 重构的清晰图像与深度图	30
2-13	在真实图片中不同方法的定性对比	32
3-1	基于时空 MRF 的实时视频去雾	35
3-2	暗通道先验示例	36
3-3	亮度值先验	37
3-4	雾浓度的时间相关性	38
3-5	积分图计算示例	40
3-6	不同视频去雾方法的平均亮度曲线对比	41
3-7	基于立体视频合成的有雾视频与本文去雾结果	42
3-8	在合成视频中不同方法的定性分析	44
4-1	图像去雾与颜色校正	46
4-2	不同边缘/结构保持的平滑方法	48

4-3	局部变化方差在真实图像上的分析	51
4-4	局部变化方差在一维信号上的分析	52
4-5	局部变化方差在噪声图上的分析	52
4-6	真实环境中光照的形状先验	53
4-7	真实环境中反相图的光照先验	53
4-8	Retinex 分解的对比	57
4-9	颜色校正的对比	58
4-10	Color-Checker 数据库示例	59
4-11	光照自适应的定性与定量对比	61
4-12	不同先验的性能分析	62
4-13	不同先验的收敛分析	63
5-1	图像去雾与细节增强	64
5-2	具有强边缘的小尺度内容的自然图像示例	65
5-3	RoG 在噪声图上的平滑结果	66
5-4	RoG 在噪声图上的分析	67
5-5	边缘保持策略分析	69
5-6	结构保持策略分析	70
5-7	尺度感知策略分析	70
5-8	细节增强度对比	74
5-9	基于 RoG 的多尺度细节增强	75
5-10	结构提取的对比	76
5-11	在 BSD300 上的结构提取对比	77
5-12	JPEG 伪影消除的对比	77
5-13	压缩后 Emoji 生肖图与 RoG 伪影消除结果	78
5-14	在 100 张 Emoji 图片下 JPEG 伪影消除的对比	78
5-15	HDR 色调映射的对比	79
5-16	图像摘要与铅笔草图	80
5-17	纹理编辑	81
5-18	内容感知的图像缩放	82

第一章 绪论

1.1 引言

图像是一种视觉符号，是人类对视觉感知的物质再现。近 30 年来，数字采集技术和信息处理理论的日新月异，越来越多的图像以数字媒体的形式加以存储，从而带动数字图像处理技术取得了令人瞩目的进展。随着计算机视觉（Computer Vision）和模式识别（Pattern Recognition）技术的发展，智能的视觉感知系统得到广泛的应用，其中包括：无人驾驶、安防监控、工业机器人、医学影像等。图像的质量对视觉的感知、分析和理解都会产生重要的影响，但现实的成像条件常常不能满足复杂感知系统的要求。也就是说，传感器捕获的图像一般不能直接地用于视觉感知。因为，成像过程中会受到很多不确定因素的影响，如系统噪声、失焦模糊、运动模糊、恶劣天气等，这都会导致数字图像的不同程度退化。其中，雾霾天气影响户外成像的主要因素。本文重点研究恶劣天气中雾霾所导致的图像退化和质量改善。

雾霾，是雾和霾的组合词。雾是由悬浮在近地面大气中的微小水滴组成的湿气溶胶系统；霾是大气中的灰尘、硫酸盐、硝酸盐等固体颗粒物组成的干气溶胶系统。雾霾是由于特定气候与人类活动的相互作用——高密度的经济活动和社会活动会排放大量的细颗粒物（Fine Particulate Matter, PM_{2.5}），一旦污染物的排放量超过大气环境的承载量且受到静稳天气的影响，细颗粒物将持续累聚并出现雾霾天气。通常情况下，湿气溶胶（雾）和干气溶胶（霾）同时存在于空气中，且不易区分，因此被称为“雾霾”现象。本文对雾和霾统一称为雾（Haze）。

在雾天，传感器受大气中细颗粒物对光线传输的影响，从而导致其成像质量出现对比度下降、饱和度降低、清晰度变差、动态范围缩小、颜色失真等退化现象。此外，这些退化现象还会随着场景物体与传感器距离的变化而变化。图1-1给出了一组北京多个地标性建筑物在雾天和晴天的对比图片^{注1}。可以看出，成像质量受雾的严重影响——雾天图像（左半边）的场景颜色单调，能见度与对比度较低，对象细节模糊；晴天图像（右半边）场景颜色丰富，色彩饱和度高，纹理清晰可见。

图像去雾技术（Image Dehazing）是通过图像处理方法，消除或减少雾对成像系统的影响，从而提高图像的能见度、对比度、饱和度和清晰度，并获得携带丰富信息的高质量图像。图像去雾增强在计算机视觉系统中属于低层次图像处理过程，是进行高层

^{注1}<https://www.theguardian.com>

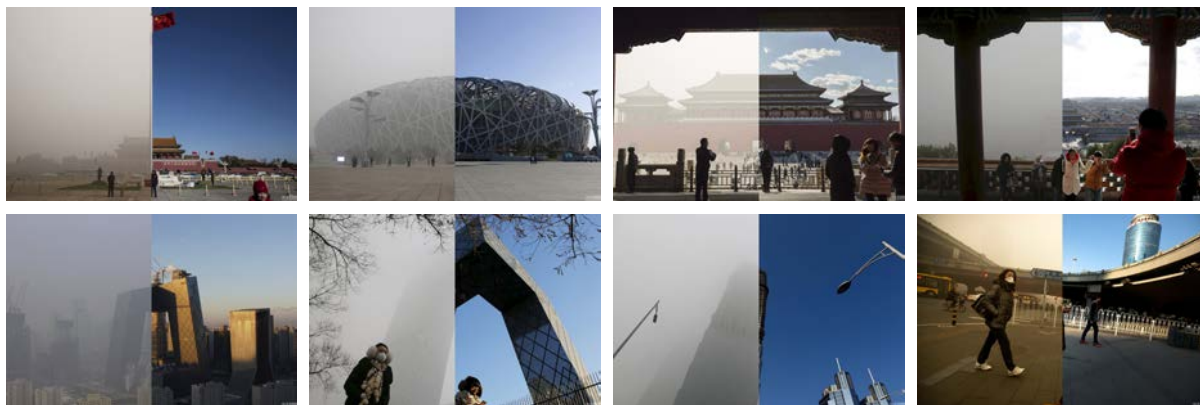


图 1-1 雾天和晴天的成像对比

次视觉感知任务的基础，如：物体跟踪、目标检测、图片分类、行为识别和语义分割等。图像去雾作为常用的图像处理技术，其有着非常广泛应用场景，已然成为图像处理领域的研究热点。然而，由于图像去雾受多样化环境的随机性与复杂性影像，现有去雾方法仍存在一定的局限性，需要进一步的改进和完善。

1.2 研究背景和意义

雾是一种常见的天气现象，尤其在秋冬季节出现频繁。近年来，我国环境污染较为严重，全国各地多处出现雾霾、沙尘暴等恶劣天气状况。有报告显示，全国仅有不到一成的城市达到世界卫生组织（WHO）的空气质量标准，全球十座空气污染最为严重的城市有七座在中国。雾天严重地影响了人们的社会生活、身体健康、交通出行等多个方面。2013 年，“雾霾”作为年度关键词受到民众广泛关注，当年 1 月雾霾笼罩 30 多个地区，北京市仅有 5 天是晴天。2014 年，国家减灾办首次将雾霾天气列入自然灾情报告；2015 年，习近平在京考察间指出，应对雾霾污染的首要任务是控制 PM_{2.5}；2017 年，李克强将“打好蓝天保卫战”写入年度的政府工作报告。因此，图像去雾技术也逐渐成为图像处理和机器视觉领域的研究热点。

雾霾会带来图像的能见度降低和视觉质量下降，其所导致的重大交通事故、公路交通阻断、航班延误等情况时有发生。在道路交通方面，大雾不仅给驾驶员造成视线模糊，也给道路监控带来成像干扰。在军事侦察方面，烟雾会影响指战员对战场的实时观察和战术的及时调整，云雾会限制无人机的全天候侦查和监视。在航空航天方面，雨雾影响飞行员视线并危害航班安全，雾霾影响遥感卫星的成像质量。同时，雾霾对智能视觉感知系统也造成巨大影响，其导致成像的对比度低、动态范围小、细节不清晰等，从而影响图像识别、目标检测、物体跟踪等高层次的视觉感知任务。

1.3 相关研究现状

近十几年, 计算机视觉技术高速发展, 期间出现多种可用于雾天图像复原的图像增强方法。根据其增强方式大致可划分为三类: 1) 基于图像增强的方法, 主要使用数字图像处理技术来消除或减弱雾霾对图像质量的影响; 2) 基于 Retinex 的方法, 应用人类视觉系统对颜色恒常特性来优化和提高图像的视觉效果; 3) 基于物理模型的方法, 结合物理原理和成像原理实现清晰图像的重构。

1.3.1 基于图像增强的方法

雾天图像的主要特点是低对比度, 早期研究工作主要有针对性地采用传统图像增强技术进行增强。基于图像增强的方法不考虑雾天的成像机理, 主要采用通用图像处理技术进行增强处理——通过一系列的图像增强操作, 提高对比度、饱和度、清晰度、锐度等有效信息, 从而达到减小雾对图像的影响。

直方图均衡作为经典的对比度增强算法, 最早地被应用于图像去雾。受空气中悬浮颗粒的散射影响, 有雾图像通常对比度较低且亮度较高, 表现为大部分像素集中在高亮度区域 (如图1-2a)。直方图均衡能有效地平衡各像素的强度分布, 实现图像动态范围的扩展 (如图1-2b)。全局的直方图均衡可一定程度上增强有雾图像, 但当雾浓度分布不均时, 全局增强方法的去雾效果并不理想。对此, Kim 等人^[1] 采用局部子块重叠的直方图均衡^[2] (Partially Over-lapped Sub-block Histogram Equalization, POSHE) 实现有雾图像的局部对比度增强。出于相同的目的, 广义局部直方图均衡^[3] 可通过均衡参数的自适应调节实现对比度的动态控制。为更好地适应局部增强, Khan 等人^[4] 使用直方图分割算法对图像进行分块, 并针对性地对子区域进行增强。在直方图均衡的基础上, Rudolf Richter^[5] 通过直方图匹配和局部加权融合实现边缘保持的图像增强。以上方法可在一定程度上提高图像对比度, 进而实现图像的去雾增强。

除此之外, 基于同态滤波^[6]、小波变换^[7]、曲波变换^[8] 等增强算法也在图像去雾中得到了应用。同态滤波^[9] (Homomorphic Filter) 是融合频率滤波与灰度变化的增强方法, 其能抵抗光照不均的影响, 比较适合于薄雾图像的增强。然而, 同态滤波仅考虑图像的频域特征, 忽略空域信息, 因此其对浓雾图像的增强效果并不理想。小波变换^[10] (Wavelet Transform) 融合频域与空域信息, 具有多尺度、去相关、局部的特征表达能力。基于小波变换的去雾方法可通过细节增强实现图像能见度的提高, 主要应用于热感图像的去雾和医疗影像的增强。曲波变换^[11] (Curvelet Transform) 是小波变换的一种

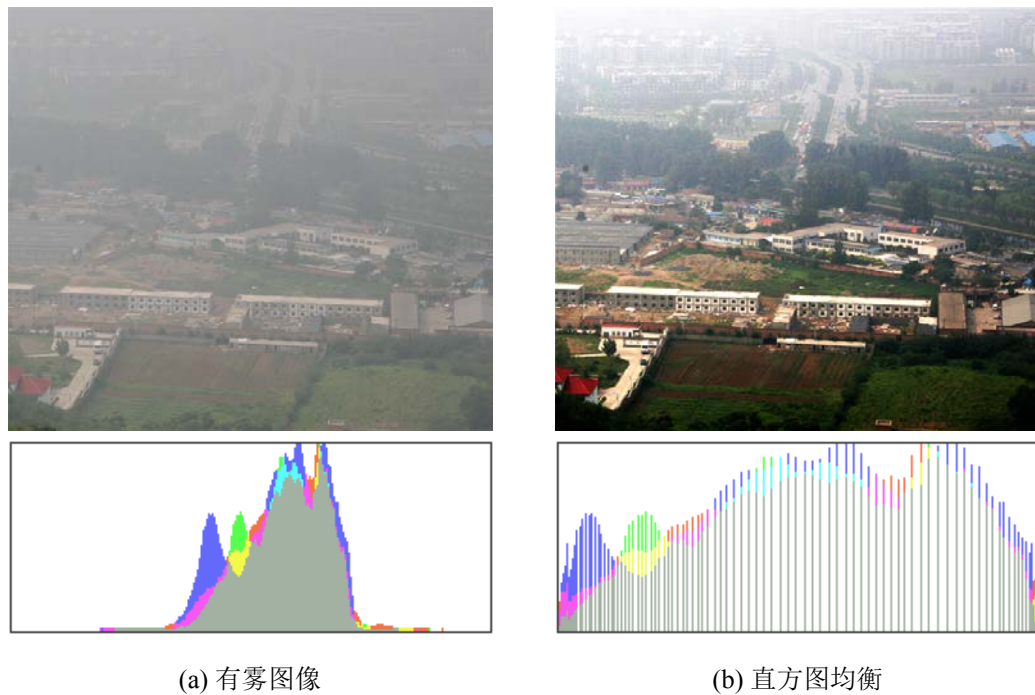


图 1-2 颜色直方图与去雾增强

高维泛化，可刻画不同尺度、不同角度的图像特征，对有雾图像也有明显的去雾效果。

基于图像增强的去雾方法具有效率高、易实现的优点。然而，由于其本质是对图像的对比度的增强，缺乏考虑有雾图像的成像机理和退化模型。因此，图像增强只能在一定程度上改善视觉效果，并不能从本质上去除雾对图像的影响。此外，由于缺乏对图像退化机理的研究，在对比度增强的同时可能导致颜色的失真。

1.3.2 基于 Retinex 的方法

颜色恒常性^[12]（Color Constancy）是人类视觉系统（Human Visual System, HVS）的知觉特性——当照射物体表面的环境光发生变化时，人们对该物体颜色的知觉仍保持不变，这使得人类视觉能不受环境光影响地感知物体的真实颜色。如图1-3所示，左魔方中的蓝色方块和右魔方中的黄色方块实际上是相同的灰色。1971 年，Edwin Land^[13]基于人类视觉对光照和颜色的感知模型提出了一种颜色恒常性的计算理论——Retinex 理论。Retinex 是由“Retina（视网膜）”和“Cortex（脑皮层）”构成的合成词，其模拟人类视觉系统的处理机制——人类视觉系统具有很好的颜色恒常性，其可主动地剔除外界光源的强度、颜色、方向等外征信息，有效地保留物体的纹理、形状、色彩等本征信息。不同于研究雾天的物理成像机理，Retinex 理论从人类视觉系统的颜色感知对图像失真进行建模，刻画的是图像失真的一个可能原因。图像去雾是一种特殊的颜色恒常性问题，其目的是去除大气光的散射对于物体成像的影响，因此可通过 Retine 模型对

其进行恢复和增强。

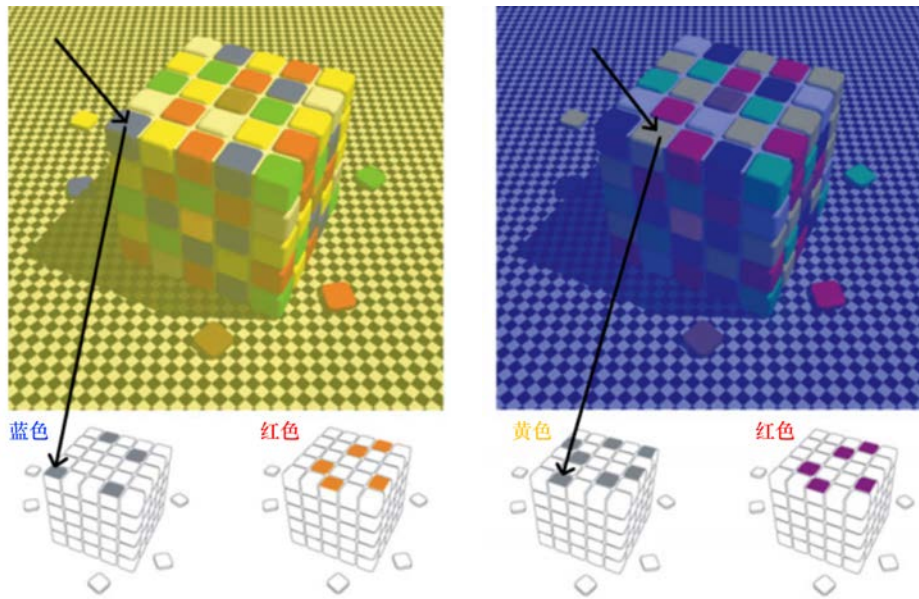


图 1-3 颜色恒常性

在 Retinex 理论的早期发展过程中，出现了基于迭代^[14,15]和基于路径^[16,17]的分解算法，但递归求解过程的高复杂度限制其应用，当前去雾研究主要集中在基于滤波器^[18,19]的 Retinex 分解。雾天对成像的影响可等效为不同光强度对成像的影响，因此芮等人^[20]结合多尺度 Retinex (Multiple Scale Retnax, MSR) 和伽马 (Gamma) 变化实现对有雾图像的对比度增强和颜色校正。Liu 等人^[21]改进 MSR 并提出利用加权多尺度 Retinex (Weight Multi-scale Retinex, WMSR) 来快速、有效地实现有雾图像的增强处理。然而，单通道的 MSR 仅考虑亮度通道的对比度增强，因此其存在较为严重的过饱和与视觉失真现象。针对此问题，Rahman 等人^[22]提出彩色多尺度 Retinex (Multi-scale Retinex with Color Restoration, MSRCR)，通过颜色通道处理补偿彩色丢失，在提高图像对比度的同时避免颜色失真。根据神经生物学研究，Land^[23,24]引入基于空间拮抗特性的中心环绕算法，Jobson 等人^[25]利用其将 Retinex 理论重写为现常见的中心/环绕对数形式，Provenzi 人等^[26]在该形式下建立图像驱动的颜色校正。

基于人类视觉系统建立的 Retinex 模型对雾天的退化图像有较好的去雾效果，但由于缺乏对雾浓度分布的建模，该方法对于复杂场景下的去雾效果不佳。因此，Retinex 方法常用于航拍、遥感等远距离且浓度分布均匀的图像增强，而对于复杂度地面成像系统效果并不理想，如监控、摄影、自动驾驶等。结合有雾图像特点并引入雾天的先验假设，将极大提升了 Retinex 模型对有雾图像的增强效果。2010 年，Xie 等^[27]提出一种

基于 MSR 和暗通道先验结合的图像去雾方法：MSR 先快速地实现快速的透射率估计，再利用暗通道先验和大气传输模型实现图像的退化修复。2012 年，嵇晓强^[28] 结合大气传输模型和暗原色先验实现透射率的估计和清晰图像的重构，并利用 Retinex 理论和非均匀的雾浓度分布估计对图像亮度拉伸，从而改善暗通道方法的图像偏暗现象。2017 年，Zhang 等人^[29] 结合自动偏振分解提出实时的图像去雾方法。

1.3.3 基于物理模型的方法

基于物理模型的方法属于图像复原的范畴——在分析大气散射对成像过程的基础上，针对雾天图像的特点构建了不同的图像退化模型，并利用其恢复清晰图像。这类方法着眼于去雾模型的光学基础和图像退化的机理，有着其自身的优越性。早期的研究工作主要针对调制传递函数^[30]（Modulation Transfer Function, MTF），由于该函数与雾浓度以及成像设备的光学特性紧密相关，因此对 MTF 的求解十分困难。1976 年，Mccartney 等人^[31] 从光学角度研究不同天气条件下大气粒子的散射路径和物理特性，并分析其对能见度的影响。直到 1998 年，Oakley 等人^[32] 利用米氏散射定律实现对图像的有效去雾，使得基于成像模型的去雾方法受到广泛地重视。1999 年至今，以 Narasimhan 和 Nayer 等人^[33] 为代表的计算机视觉研究者针对米氏（Mie）散射^[34] 和 Mccartney 衰减模型^[31] 进行深入研究，分析了大气散射对图像降质的影响，并提炼出了大气散射模型（Atmospheric Scattering Model）。此后，几乎所有基于物理模型的图像去雾方法都是建立在该模型之上。

• 单帧图像去雾

基于大气散射模型，从单幅图像中提取物理变量的关键是对雾浓度的准确估计。现有基于物理模型的图像去雾方法普遍利用观察经验的先验知识来构建去雾模型。基于清晰图像具有高对比度的先验知识，Tan 等人^[35] 利用马尔可夫随机场（Markov Random Field, MRF）提高场景图像局部对比度，从而恢复出清晰的场景图像。基于独立性假设，Fattal 等人^[36] 通过独立成分分析（Independent Component Analysis, ICA）实现对有雾图像的信息增强和颜色校正。基于正则概率先验，假设场景深度的统计独立，Kratz 等人^[37] 假设场景深度的统计独立，并通过最大化后验概率实现图像去雾。基于条件独立性假设，Nishino 等人^[38] 采用贝叶斯统计同时实现清晰图像和场景深度的重构。基于空间几何先验，Carr 等人^[39] 结合相机内外参数和场景空间关系实现大气透射率的优化。然而，这些算法在实际使用中都具有

的限制和约束，存在噪声放大和过饱和度的问题。

2009 年，He 等人^[40] 基于暗通道先验（Dark Channel Prior，DCP）提出场景深度估计和清晰图像重构的图像去雾方法。在此基础上，Gibson 等人^[41] 使用主成分分析（Principal Component Analysis，PCA）对暗通道先验进行分析，验证其在物理上的合理性。然而，仅基于 DCP 的去雾方法并不能完全满足复杂场景的去雾需求，因为白色物体和天空区域不满足暗通道先验，因此后续出现很多基于 DCP 的改进方法^[42-44]。随着机器学习理论的发展，基于先验特征学习^[45-47] 的方法也逐步受到重视，通过学习获取雾天图像与大气透射率之间的映射关系。

• 多帧图像去雾

单帧图像去雾是一个欠约束的病态问题，深度信息对大气透射率的估计有很好的辅助作用。Oakley 等人^[48] 通过二维的航拍图像重构出三维的场景地图来辅助遥感图像去雾；Narasimhan 等人^[49] 通过交互式的景深标记方法获取有雾场景的三维结构信息；Kopf 等人^[50] 通过地理信息库对有雾场景的深度进行估计。然而，在真实户外场景中的深度信息较难获得，因此基于辅助深度信息的去雾方法并未得到广泛的应用。总的来说，多帧图像去雾方法需要较多的场景图像，同时，为保证场景的一致还需要对有雾图像和深度图像进行像素级的对齐。

视频去雾作为一种特殊的多帧处理，是单帧图像去雾的扩展与延伸。与大量的单帧图像去雾工作相比，视频去雾方面的研究相对较少。视频序列中场景的不断变化和大量的图像帧，对去雾算法的性能提出了更高的要求。将单帧图像去雾算法应用到视频的每一帧中^[51-53] 可取得较好的去雾效果，虽能够适应视频场景的变化但效率较低。将单帧估计的雾浓度信息应用到后续视频中^[54-56] 有很高的去雾效率，但视角不变假设不能适应复杂变化的应用场景，一般仅适用于定点视频监控。Tarel 等人^[57] 在车载视频流中应用分割技术对运动物体和平面道路进行有效的切分，并应用单帧图像去雾方法^[58] 在线更新深度估计。为引进时空相关性，光流法^[59] 也被应用于大气透射率估计，但光流法计算的复杂度和光流信息的融合限制了光流在实时视频中的应用。Kim 等^[60] 为了消除视频去雾中的闪烁效应，通过最小化时序相关性来优化并增强对比度，但当对比度过度拉伸时会导致饱和度截止。Li 等人^[61] 联合深度图估计和浓度图估计实现对清晰视频的重构，但基于运动结构估计的深度图估计同样需要大量匹配计算。此外，由于远景物体的视差较小，其在远场景的去雾中并

不能取得很好效果。

相比基于图像增强和基于 Retinex 的去雾方法,物理去雾模型具有更好的理论基础。当大气透射率和大气光得到准确的估计,基于物理模型的方法能获得很好的去雾效果,特别对于非均匀分布的有雾图像也能取得较好的效果。因此,如何准确的估计大气透射率是本类方法的难点和重点。然而,基于物理模型的去雾方法也同样存在过度去雾的问题,从而使得恢复的图像的颜色整体偏离和颜色失真。

1.4 主要研究内容

本文从基于物理模型的去雾方法出发,提出基于卷积神经网络的大气透射率估计方法,并进一步针对视频去雾提出基于时空 MRF 的透射率优化方法。同时,在物理模型去雾的基础上,结合 Retinex 模型和高斯相对性分别实现颜色校正和细节增强。本文对图像去雾与增强问题做出了以下贡献:

(1) 单帧图像去雾是一个极具挑战性的不适定问题。现有的方法普遍使用各式各样的条件约束或观测先验来实现图像增强。图像去雾的关键是对有雾图像中大气介质的透射率进行准确估计。本文提出了一种可端到端训练的卷积神经网络——DehazeNet 用于透射率估计。DehazeNet 以有雾图像作为输入,输出透射率图,然后通过大气散射模型重构出清晰图像。DehazeNet 采用基于卷积神经网络的深度结构,通过特别设计网络结构学习有雾图像的假设和先验。其中,特征提取层的“Maxout+ 卷积”可生成几乎所有与雾相关的传统特征。此外,非线性回归层中提出了一种新的非线性激活函数——双边修正线性单元(BReLU)用于提高清晰图像的重构质量。在多个实验上表明,DehazeNet 比现有方法具有更好的性能,同时还保持了高效性和易用性。

(2) 相比于图片去雾,实时视频去雾具有更为广泛的应用前景,其挑战主要来自时空相关性和计算效率。本文提出了一种基于时空优化框架的实时视频去雾方法,可有效减少区块效应和闪烁效应,并获得高质量的增强效果。该框架建立了一个基于亮度值先验的马尔可夫随机场(Spatio-temporal Markov Random Field, ST-MRF)来处理空间一致性和时间连续性,通过最大化 MRF 的似然函数实现雾浓度图的估计和优化。此外,为了保证算法的实时性,本文通过积分图像技术实现 $O(N)$ 时间复杂度的算法加速,并应用降采样方法进一步减少冗余计算量。实验结果表明,该框架可有效地去除区块和闪烁伪影,并在单 CPU 下实现对通用影像传输格式(Common Intermediate Format, CIF) 352×288 尺寸视频流约 240 帧每秒的实时处理。

(3) 雾的形成是由于悬浮在空气中的微小颗粒, 大气光的估计误差会影响去雾结果的颜色平衡, 因此对复杂光照下的图像去雾需要对其进行颜色校正。对此, 本文提出了一个基于联合先验的 Retinex 模型, 提取并消除图像中的光照效果, 从而实现颜色校正。根据图形学研究, 三维物体在二维空间映射形成受场景属性(包括形状、照明和纹理)的影响。首先, 本文提出一种新的结构保持方法——局部变化方差作为结构保持先验。相比于传统的 Retinex 模型, 本文方法通过形状先验保持光照的结构, 通过光照先验捕获光源的信息, 通过纹理先验估计更精细的反射率。实验结果表明, 该方法在模拟数据和实际数据的处理效果良好。与其他的 Retinex 算法相比在主观和客观评测上都获得更好的结果。

(4) 在雾天环境下, 由于大气中悬浮颗粒使得光线在传播的过程中会发生散射、折射、吸收等现象, 从而导致图像细节信息的丢失。因此, 细节增强作为去雾后增强操作, 可在去雾结果的基础上进一步优化可视化效果。自然图像包含各个尺度的边缘和结构, 如何有效地捕捉小尺度细节是图像细节增强的关键。本文提出了一种基于高斯相关性(Relativity-of-Gaussian, RoG)的图像平滑方法。RoG 是一种具有尺度感知的图像平滑方法, 在保持主图像内容的同时, 可有效地捕获细节信息。大量实验表明, 该平滑方法在保持主图像内容的同时, 可有效地捕获细节信息。效率方面, 本文引入一种可分离的递归优化方法减少了求解稀疏线性系统的主要计算负担, 实现在 $O(N)$ 时间复杂度下的有效平滑, 即使在单核 CPU 上也可实现实时处理——处理百万像素图像约 0.15 秒。

1.5 本文结构安排

全文共分为六章节, 其主要内容与组织结构如图1-4所示, 具体安排如下:

第一章概述图像去雾的课题背景与研究意义, 并简述国内外相关的研究现状与工作进展, 最后阐述本文的主要研究内容。

第二章从单帧图像(图片)去雾方面展开研究, 针对大气透射率的估计问题, 提出一个新的基于卷积神经网络的图像去雾方法。特别地, 针对于图像去雾这特殊应用场景, 提出了“卷积+Maxout”的特征提取结构和 BReLU 的新型非线性回归函数。

第三章从多帧图像(视频)去雾方面展开研究, 针对大气透射率的时空优化问题, 提出一个基于时空马尔可夫随机场的实时视频去雾方法。在保证空间一致性和时间连续性的同时, 通过积分图和下采样技术快速优化和估计大气透射率。

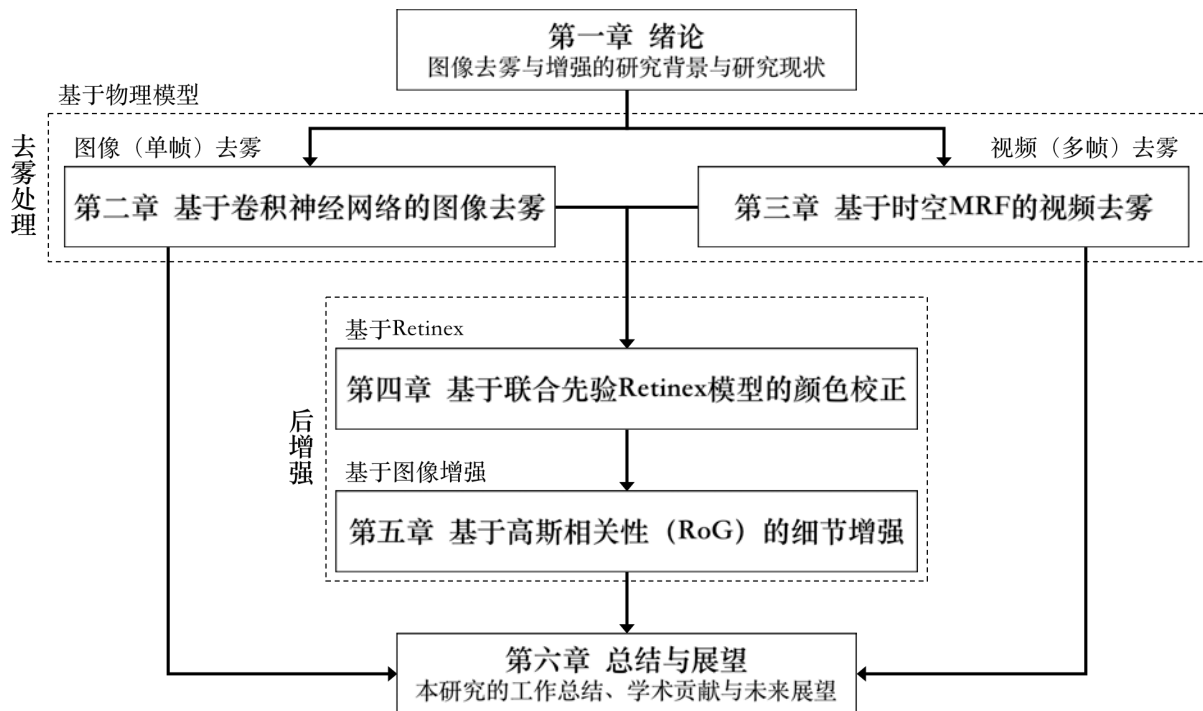


图 1-4 论文结构关系图

第四章从颜色校正方面展开研究，针对物理模型去雾可能面临的颜色失真问题，提出一种基于联合先验的 Retinex 模型。本文提出局部变化方差的度量方法刻画物体的形状信息，并结合亮度先验、纹理先验构建联合的 Retinex 求解函数。

第五章从细节增强方面展开研究，针对雾导致成像对比度下降的问题，提出一种基于高斯相对性的细节增强方法。本文讨论现有的非线性平滑方法，提出具有尺度感知的非线性平滑优化，并通过快速的可分递归优化保证其效率。

第六章对本课题的研究内容和学术贡献进行了总结，并思考和展望该领域的未来发展方向。

第二章 基于卷积神经网络的图像去雾

2.1 概述

雾是一种常见的大气现象，空气中悬浮的水滴、灰尘、细沙或其他颗粒等都会引起成像清晰度的降低。单幅图像去雾是一个病态问题，其关键是如何准确地估计成像过程中的大气透射率。现有的图像去雾方法普遍利用基于观察经验的先验知识来构建去雾模型。基于有雾图像具有低局部对比度的观察先验，Tan 等人^[35]通过马尔科夫随机场（Markov Random Field, MRF）实现局部对比度的最大化，但局部对比度方法同时会带来颜色过饱和现象。在暗物体减影技术的启发下，清晰图像的观测经验发现了暗通道先验^[40]（Dark Channel Prior, DCP）：在大多数清晰图像中，至少有一个颜色通道具有很低低的强度，因此可利用暗通道先验与大气散射模型实现清晰图像的重构。为了进一步提高去雾重构的质量，Nishino 等人^[38]使用阶乘 MRF 对图像进行建模，以更准确地估计大气透射率；Meng 等人^[62]提出了正则化去噪方法以约束图像边缘并提高重构精度。此外，中值过滤^[63]、双中值滤波器^[58]、双边滤波器^[64]和导向滤波器^[65]等取代耗时的软匹配^[66]操作以提高计算效率。近几年，机器学习方法被逐步应用于图像去雾的研究中。Tang 等人^[47]结合四种不同的图像去雾特征，并构建随机森林来回归大气透射率。Zhu 等人^[46]基于颜色衰减先验线性训练有雾图像到场景深度的线性映射模型。然而，以上图像去雾方法均以启发式的先验特征为基础，无法满足多样化的去雾场景需求。

众所周知，人类的大脑可以在无附加信息的条件下从单个自然场景中迅速地识别出有雾区域。近几年，生物启发式的卷积神经网络在高层次视觉任务已取得成功的应用，如：图像分类^[67]、人脸识别^[68]、物体检测^[69]等。基于卷积神经网络的深度学习方法也开始逐步被应用于图像修复等低层次视觉任务^[70-72]。然而，现有的网络结构并不能被直接应用于单幅图像去雾。因此，本文将根据图像去雾这一应用场景，针对性地设计一个基于先验启发的卷积神经网络。

在雾天的成像过程中，光线在物体表面的反射光在穿过大气的过程中发生散射，导致只有部分能到达摄像头，同时在悬浮颗粒表面散射形成的大气光也会被摄像头接收，因此导致图像亮度的增加和对比度的下降。如图2-1所示，McCartney^[31]和Narasimhan^[33]等提出了大气散射模型（Atmospheric Scattering Model）对雾天成像过程

进行建模——有雾图像可表示为：

$$I(x) = J(x)T(x) + \alpha(1 - T(x)). \quad (2-1)$$

其中, $I(x)$ 是观测到的有雾图像, $J(x)$ 是真实的清晰图像, $t(x)$ 是大气透射率, α 是全局大气光常数。可见, 式(2-1)中存在三个未知量, 当 $T(x)$ 和 α 被估计则可重构出清晰图像 $J(x)$ 。大气透射率 $T(x)$ 表示未被散射的反射光透过大气到达摄像头的比例: $T(x) = e^{-\beta d(x)}$, 其中 $d(x)$ 是场景中物体到摄像头的距离, β 是大气散射系数。当 $d(x)$ 趋于无穷远时, $T(x)$ 趋于零, 因此根据式(2-1)可得 $\alpha = I(x), d(x) \rightarrow \inf$ 。在真实成像中, 远景物体到摄像头的距离 $d(x)$ 并不存在真正的无穷大, 远视场带来的是非常低的透射率 t_0 。因此, 普遍采用以下规则取代式(2-1)实现更鲁棒的全局大气光估计:

$$\alpha = \max_{y \in \{x | T(x) \leq t_0\}} I(y). \quad (2-2)$$

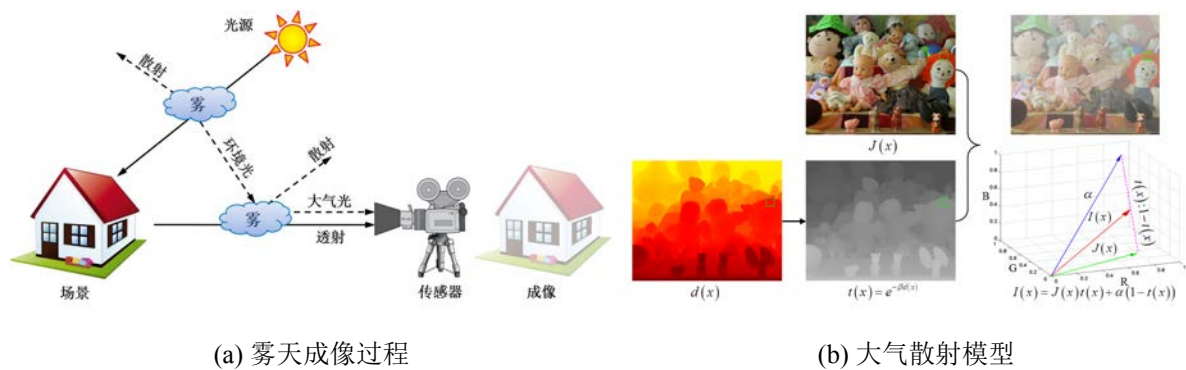


图 2-1 雾天成像系统与大气散射模型

综上所述, 图像去雾的关键是精确地估计大气透射率 $T(x)$ 。为此, 本章中提出了一个基于卷积神经网络的端到端系统 DehazeNet 来估计大气透射率。DehazeNet 将有雾图像 $I(x)$ 作为系统输入, 估计并输出对应的透射率图 $T(x)$, 并根据大气传输模型重构出清晰图像 $J(x)$ 。DehazeNet 的网络结构设计借鉴了传统图像去雾方法中的先验知识。有别于传统方法, DehazeNet 的参数可通过有雾与清晰图像的训练样本对自动学习得到。在多个基准测试数据库上的实验表明, DehazeNet 表现出比现有去雾方法更好的性能, 同时保持高效性和易用性。本工作的主要贡献总结如下:

- 提出了一个大气透射率的端到端估计系统 DehazeNet。该系统通过特别设计网络结构, 直接学习并估计有雾图像与透射率图之间的映射关系。
- 提出了一种新的非线性激活函数——双边线性修正单元 (Bilateral Rectified Linear

Unit, BReLU)。BReLU 是线性修正单元^[73] (Rectified Linear Unit, ReLU) 的扩展, 在双边约束的同时保持局部线性, 从而减少参数搜索空间并提高收敛速度。

- 分析传统方法中去雾先验与 DehazeNet 网络结构的关系, 并从理论上解释 DehazeNet 能实现自动参数学习和取得很好去雾效果的原因。

2.2 去雾卷积网络 DehazeNet

在节2.1中, 大气散射模型指出透射率估计是重构清晰图像的关键。本文提出一个可端到端训练的系统 DehazeNet, 学习有雾图像和透射率图的映射关系。本节将详细阐述 DehazeNet 的结构设计, 并讨论其与传统图像去雾方法的联系。最后, 根据式(2-1)和透射率图 $T(x)$, 可简单地通过像素级操作实现清晰图像 $J(x)$ 的重构。

2.2.1 DehazeNet 的网络结构

DehazeNet 的网络结构由卷积、池化和非线性函数共同组成 4 个连续的操作层实现透射率估计, 分别是: 特征提取、多尺度映射、局部极值和非线性回归。其中, 包含了 3 个卷积、1 个最大池化、最大输出单元和 BReLU 激活函数, 其详细设计细节如图2-2所示。

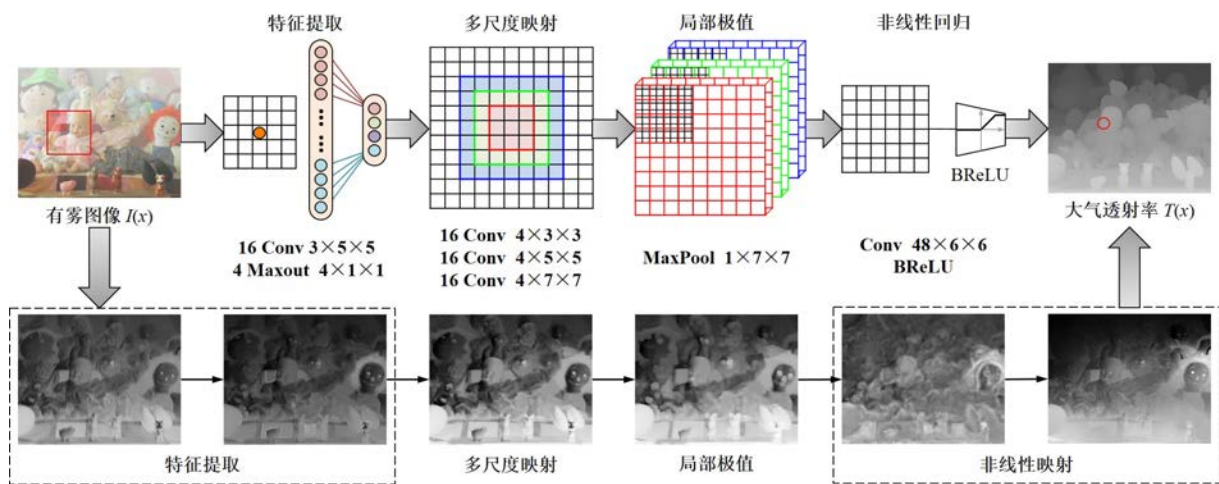


图 2-2 DehazeNet 的网络结构

2.2.1.1 特征提取

为解决图像去雾这一病态问题, 传统去雾方法基于先验假设在输入图像中密集地提取与雾相关的特征, 如: 暗通道^[74] (Dark Channel)、色相差异^[75] (Hue Disparity)、颜色衰减^[46] (Color Attuation) 等。注意到, 密集的特征提取可等价于线性的卷积操作

和独特的非线性映射。因此，受经典先验特征中颜色通道极值的启发，DehazeNet 采用独特的激活函数——最大输出（Maxout）单元^[76]用于非线性映射与特征降维。Maxout 单元是一种简单的前馈非线性激活函数，常用于多层感知机（Multilayer Perceptron, MLP）和卷积神经网络（Convolutional Neural Network, CNN）。当应用于 CNN 时，Maxout 单元将输入特征图的 k 个通道分为一组，分别对各组执行像素级的最大值操作。基于“卷积+Maxout”，DehazeNet 第一层的特征提取可表示为：

$$F_1^i(x) = \max_{j \in [1, k]} g^{i,j}(x), g^{i,j} = W_1^{i,j} * I + B_1^{i,j}. \quad (2-3)$$

其中， $\mathcal{W}_1 = \{W_1^{i,j}\}_{(i,j)=(1,1)}^{(n_1,k)}$ 和 $\mathcal{B}_1 = \{B_1^{i,j}\}_{(i,j)=(1,1)}^{(n_1,k)}$ 分别表示滤波器的权值和偏置， $*$ 表示卷积操作。 $W_1^{i,j} \in \mathbb{R}^{3 \times f_1 \times f_1}$ 是 $k \times n_1$ 个 $f_1 \times f_1$ 的卷积滤波器，其中 3 是输入图像的通道数（参数设置详见表 2-1）。Maxout 单元映射 kn_1 维的卷积结果并输出 n_1 维的特征图，取代传统先验特征实现自动化的雾相关特征提取。

2.2.1.2 多尺度映射

在前期去雾工作^[47,77]中，多尺度特征被证明在图像修复中的有效性。多尺度特征不仅可刻画不同尺度的物体，还可有效地应对物体的尺度变化。例如，GoogLeNet^[78]中的 inception 结构并联多个不同尺度的滤波器，解决了图片分类中物体的尺度对齐问题，并在 ILSVRC14^[79]中取得最好效果。受多尺度特征提取的启发，DehazeNet 在第二层结构中并联三组个数相同、尺度不同的卷积操作，滤波器尺度包括 3×3 、 5×5 和 7×7 。数学形式上，DehazeNet 第二层的多尺度映射可表示为：

$$F_2^i = W_2^{[i/3],(i \setminus 3)} * F_1 + B_2^{[i/3],(i \setminus 3)}. \quad (2-4)$$

$\mathcal{W}_2 = \{W_2^{p,q}\}_{(p,q)=(1,1)}^{(3,n_2/3)}$ 和 $\mathcal{B}_2 = \{B_2^{p,q}\}_{(p,q)=(1,1)}^{(3,n_2/3)}$ 分别包含 n_2 个可分解为 3 组的权值和偏置。其中， n_2 是第二层网络的输出， $i \in [1, n_2]$ 是输出特征图的索引， $\lceil \cdot \rceil$ 表示向上取整， $\setminus$ 表示取余操作。

2.2.1.3 局部极值

为获取空间不变性，视皮层复杂细胞对简单细胞的响应进行非线性的特征整合。Ilan 等人^[80]提出复杂细胞的空间整合机理可描述为一系列的池化操作。根据 CNN 的经典结构^[81]，邻近像素间的局部最大值操作可增强特征空间变化的鲁棒性。此外，局部极值操作符合透射率的局部一致性假设，此假设被普遍地应用于透射率的估计误差消

除。因此，局部极值操作为 DehazeNet 的第三层，表示为：

$$F_3^i(x) = \max_{y \in \Omega(x)} F_2^i(y). \quad (2-5)$$

其中， $\Omega(x)$ 是以 x 为中心的 $f_3 \times f_3$ 邻域，第三层特征的输出维度 $n_3 = n_2$ 。不同于标准卷积网络中最大池化降低特征图的分辨率，局部极值操作是最大值滤波器，其密集地应用于特征图中的每个像素点，因此更好地保持了图像的细节分辨率。

2.2.1.4 非线性回归

在深度学习中，常用的非线性激活函数主要包括：Sigmoid 函数^[82] 和修正线性单元^[73]（Rectified Linear Unit, ReLU）。Sigmoid 函数由于两端的响应饱和，容易出现梯度弥散现象，并可能导致网络收敛缓慢或陷入局部最优解。为解决梯度弥散的问题，ReLU^[73] 提供了稀疏的特征响应。然而，ReLU 的提出主要针对图像识别的分类问题，而非图像重构的回归问题。ReLU 的输出只抑制了小于 0 的值，这可能会导致响应溢出，特别是图像重构问题中 $[0, 1]$ 的有限响应输出。为了解决响应溢出的问题，本文提出了双边修正线性单元（Bilateral Rectified Linear Unit, BReLU），如图2-3所示。

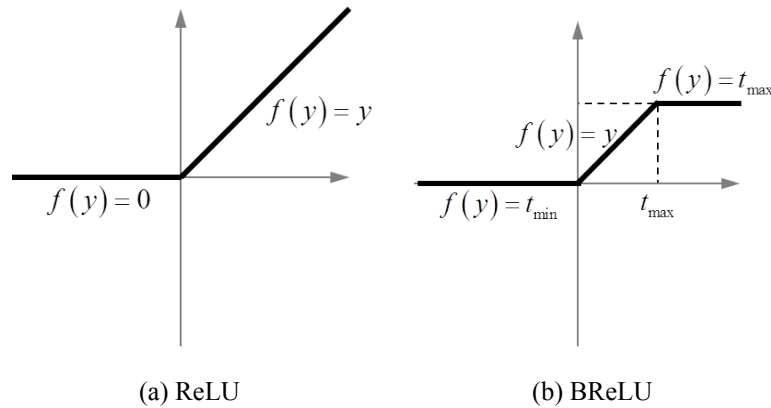


图 2-3 修正线性单元（ReLU）和双边修正线性单元（BReLU）

受 Sigmoid 和 ReLU 启发，BReLU 是一种新的激活函数，既保持双边修正又保证局部线性。基于 BReLU，输出层可定义为：

$$F_4 = \min(t_{\max}, \max(t_{\min}, W_4 * F_3 + B_4)). \quad (2-6)$$

其中， $\mathcal{W}_4 = \{W_4\}$ 包含一个尺度为 $n_3 \times f_4 \times f_4$ 的滤波器， $\mathcal{B}_4 = \{B_4\}$ 是其偏置， $t_{\min, \max}$ 是 BReLU 的修正边界（本文中 $t_{\min} = 0$ ， $t_{\max} = 1$ ）。根据式(2-6)，BReLU 激活函数的

梯度可表示为

$$\frac{\partial F_4(x)}{\partial F_3} = \begin{cases} \frac{\partial F_4(x)}{\partial F_3}, t_{\min} \leq F_4(x) < t_{\max} \\ 0, otherwise \end{cases} \quad (2-7)$$

2.2.2 DehazeNet 与传统方法的联系

上述四个级联模块共同组成了端到端的大气透射率估计系统 DehazeNet，其中卷积层中滤波器和偏差的参数可通过训练数据学习得到。本小节将讨论 DehazeNet 的各层结构与现有传统图像去雾方法的联系。

2.2.2.1 去雾相关特征

首先简要介绍现有传统图像去雾方法。基于观察经验，现有传统方法提出各种可用于刻画雾浓度的先验特征，从而实现图像的有效去雾。如图2-4所示，去雾相关的特征主要包括：暗通道、最大对比度、颜色衰减和色相差异。

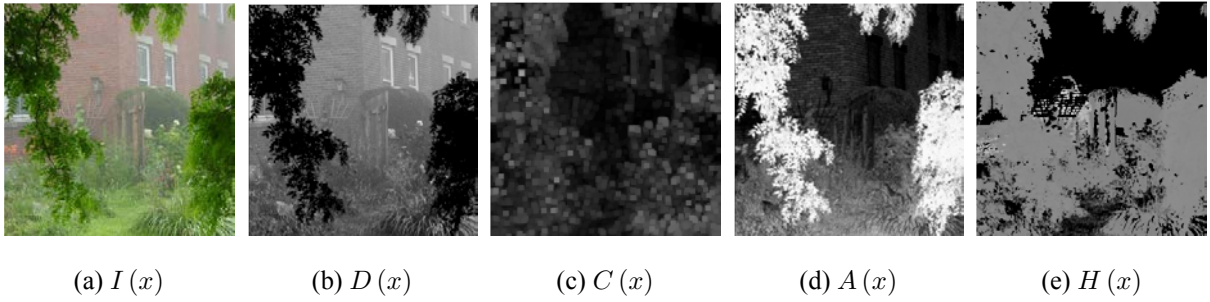


图 2-4 去雾相关的特征

- 暗通道^[74] (Dark Channel)

暗通道先验基于对大量户外清晰图像的观察得到的。在大部分清晰的图像块中，至少存在一个像素点在某一颜色通道具有很低的亮度值，甚至接近于 0。在图像块中，暗通道^[74] 被定义为所有像素的通道最小值：

$$\begin{cases} D(x) = \min_{c \in \{r, g, b\}} I^c(x) \\ D^r(x) = \min_{y \in \Omega_r(x)} D(y) \end{cases} \quad (2-8)$$

其中， I^c 是图像 I 的 RGB 颜色通道， $\Omega_r(x)$ 是以 x 为中心 $r \times r$ 邻域的局部块。暗通道特征与图像中的雾浓度有着强相关性，并可直接应用于大气透射率的估计 $T(x) \propto 1 - D^r(x)$ 。

- 最大对比度^[35] (Maximum Contrast)

根据大气散射模型,雾会导致图像对比度的降低: $\sum_x \|\nabla I(x)\| = T(x) \sum_x \|\nabla J(x)\| \leq \sum_x \|\nabla J(x)\|$ 。基于这一观测推论,尺度为 $r \times r$ 区域 Ω_r 内的局部对比度^[35]刻画像素变化的强度,并可用于雾的浓度估计。因此,在尺度 $s \times s$ 的局部块 Ω_s 中的最大局部对比度定义为:

$$C(x) = \max_{y \in \Omega_r(x)} \sqrt{\frac{1}{|\Omega_s(y)|} \sum_{z \in \Omega_s(y)} \|I(z) - I(y)\|^2}. \quad (2-9)$$

其中, $|\Omega_s(y)|$ 是局部邻域内的像素个数。根据式(2-9)可见,对比度特征 $C(x)$ 与大气透射率 $T(x)$ 有着强的相关性,因此可通过最大化局部对比度实现图像的去雾增强。

• 颜色衰减^[46] (Color Attenuation)

当彩色场景受雾霾的影响,大气透射率的下降会导致饱和度 $I^s(x)$ 的降低,同时大气光的散射会带来亮度 $I^v(x)$ 的提高,因此产生颜色衰减。根据以上颜色衰减先验^[46],图像亮度与饱和度的差值可用于雾的浓度估计:

$$A(x) = I^v(x) - I^s(x). \quad (2-10)$$

其中, $I^v(x)$ 和 $I^s(x)$ 分别表示转换到 HSV 颜色空间的亮度 (V) 和饱和度 (S) 通道:

$$\begin{cases} I^v(x) = \max_{c \in \{r, b, g\}} I^c(x) \\ I^s(x) = \frac{\max_{c \in \{r, b, g\}} I^c(x) - \min_{c \in \{r, b, g\}} I^c(x)}{\max_{c \in \{r, b, g\}} I^c(x)} \end{cases}. \quad (2-11)$$

综上所述,颜色衰减特征与景深 $d(x) \propto A(x)$ 成正比,因此可方便地被应用于大气透射率的估计。

• 色相差异^[75] (Hue Disparity)

色相差异是原始图像 $I(x)$ 与其半反图像 $I_{si}(x) = \max[I^c(x), 1 - I^c(x)]$ ($c \in \{r, g, b\}$) 之间的色相差异,其可被应用于雾的检测。对于清晰图像,半反图像中的三个通道的像素值并不会完全反转,所以 $I_{si}(x)$ 和 $I(x)$ 之间存在较大的色相差异。因此,色相差异^[75]特征定义为

$$H(x) = |I_{si}^h(x) - I^h(x)|. \quad (2-12)$$

其中,上标 h 表示转换到 HSV 颜色空间的色相 (H) 通道。根据式(2-12),大气透射率 $T(x)$ 与色相差异特征 $H(x)$ 成反比。

2.2.2.2 相关特征讨论

DehazeNet 第一层 F_1 中的“卷积 +Maxout”结构可提取与雾相关的特征。以暗通道^[74]特征为例，当滤波器权值 W_1 是反向滤波器（中心值为 -1 的稀疏矩阵，如图2-5a）、偏置 B_1 是单元 1 偏置时，Maxout 的通道最大值等价于原图的通道最小值，因此等价于暗通道^[74]特征（见式(4-1)）。同理，当滤波器权值 W_1 是环形滤波器（如图2-5c）， F_1 等价于最大对比度^[35]特征（见式(2-9)）；当滤波器权值 W_1 同时包含反向滤波器（如图2-5a）和全通滤波器（如图2-5a）， F_1 分别等价于通道最小值与通道最大值，这是 RGB 到 HSV 颜色空间转换的两个基础原操作，对应于基于 HSV 颜色空间的去雾特征——颜色衰减^[46]（见式(2-10)）和色相差异^[75]（见式(2-12)）。图2-5e所示，有效学习得到的 W_1 符合上述分析，与雾相关特征均可通过 DehazeNet 的第一层结构学习得到。此外从数学上分析，Maxout 是分段线性的激活函数函数，具有很强的拟合能力并可逼近任意的凸函数。如图2-5d所示，本文选择四个特征图（ $k = 4$ ）的最大值输出即可有效地近似任意的凸函数。

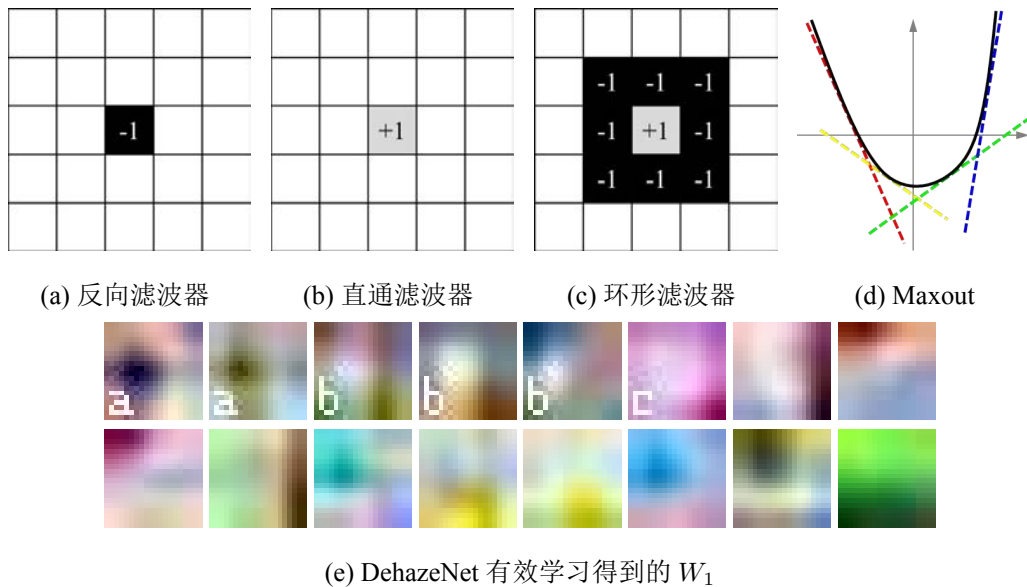


图 2-5 DehazeNet 中第一层的滤波器权值与 Maxout 单元

在自然图像中，白色物体与有雾场景具有相似的视觉特点——低饱和度和高亮度。因此，现有的大多数去雾方法都倾向于将白色物体当做远处物体，从而导致大气透射率的估计误差。基于景深的局部一致性假设，局部极值滤波器^[35,46,74]被应用于解决这一问题。在 DehazeNet 中，第三层操作中的局部极值操作对应于此假设，实现局部估计误差的消除。此外，当大气透射率 $T(x)$ 接近于 0 时，大气散射模型中的衰减

项 $J(x)T(x)$ 趋近于 0, 从而导致重构的清晰图像 $J(x)$ 非常容易受到透射率的估计噪声的影响。因此在 DehazeNet 中, BReLU 激活函数限制了透射率的范围于 $t_{\min}$ 和 $t_{\max}$ 之间, 可有效地消除噪声敏感的问题。值得注意的是, BReLU 等价于传统图像去雾方法^[46,74] 中的边缘限制。

2.2.3 DehazeNet 的训练过程

2.2.3.1 训练数据

深度模型的训练通常需要收集大量的有标签数据^[67], 这是一项耗时且耗力的工作。具体到 DehazeNet 的网络训练中, 数据收集工作变得更加困难, 因为真实场景中很难大量地采集到成对的有雾图像和清晰图像(或透射率图)。因此, 本文基于大气散射模型^[47] 合成大量成对的训练样本。

具体而言, 本文合成的有雾与清晰图像块基于两个假设^[47]: (1) 图像内容与大气透射率无关(同一物体可出现在任意场景深度); (2) 大气透射率具有局部一致性(小图像块内的像素点位于相同的深度)。基于这两个假设, 可以为任意的小图像块加上随机的雾浓度。给定清晰图像块 $J^P(x)$ 、大气光 α 和随机的大气透射率 $t \in (0, 1)$, 可合成对应的有雾图像块 $I^P(x) = J^P(x)t + \alpha(1 - t)$ 。为了降低变量的不确定性, 本文中将全局大气光 α 设成 1。

本文从互联网上收集大量的清晰图片, 并从中随机地采样成尺度为 16×16 的图像块。不同于^[47], 这些清晰图片不仅来源于人们的日常生活, 也包括大量的自然和城市景观。因为, 期望多样化的样本将有助于 DehazeNet 的参数学习, 从而提高 DehazeNet 对不同场景去雾的泛化性。图2-6展示本文收集的一些清晰图片。



图 2-6 互联网上采集到的清晰图片示例

2.2.3.2 训练方法

DehazeNet 需要有监督地学习有雾图像 RGB 值与大气透射率之间的映射关系 $\mathcal{F}$ 。网络参数 $\Theta = \{\mathcal{W}_1, \mathcal{W}_2, \mathcal{W}_4, \mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_4\}$ 的学习可通过最小化输入样本 $I^P(x)$ 与对应大气透射率 t 的距离。给定一系列有雾图像块与其对应的大气透射率，使用均方误差（Mean Squared Error, MSE）作为损失函数最小化误差：

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|\mathcal{F}(I_i^P; \Theta) - t_i\|^2. \quad (2-13)$$

借助于深度学习框架 *Caffe*^[83]，DehazeNet 采用随机梯度下降法（Stochastic Gradient Descent, SGD）完成模型的训练。DehazeNet 的详细参数设置总结见表2-1，其中包含 3 个卷积层和 1 个最大池化层，Maxout 和 BReLU 分别应用于第一层和最后一层的输出。

表 2-1 DehazeNet 的参数设置

结构	类型	输入尺度	滤波器个数 n	滤波器尺度 $f \times f$	扩边
特征提取	Conv	$3 \times 16 \times 16$	16	5×5	0
	Maxout	$16 \times 12 \times 12$	4	—	0
多尺度映射	Conv	$4 \times 12 \times 12$	16	3×3	1
			16	5×5	2
			16	7×7	3
局部极值	Maxpool	$48 \times 12 \times 12$	—	7×7	0
非线性回归	Conv	$48 \times 6 \times 6$	1	6×6	0
	BReLU	1×1	1	—	0

2.3 实验与分析

为验证 DehazeNet 结构的有效性，本文分析了其收敛性并与当前最优的去雾方法进行对比，包括：FVR^[58]、DCP^[74]、BCCR^[62]、ATM^[84]、RF^[47]、BPNN^[45] 和 CAP^[46]。关于训练数据，本文通过互联网图片随机收集 10,000 个清晰图像块，并对每一图像块随机地合成 10 个不同透射率 $t \in (0, 1)$ 的训练样本。因此，总共为 DehazeNet 的训练生成 100,00 个训练样本对。在 DehazeNet 中，每一层的滤波器的权值均采用标准正态分布初始化（均值 $\mu = 0$ ，方差 $\sigma = 0.001$ ），偏置初始化为 0。网络训练的批大小为 128，学习率从 0.005 到 3.125×10^{-4} 每 100,000 次迭代减半递减。基于以上训练参数设置，DehazeNet 在一台配备 Nvidia GeForce GTX 780 GPU 的计算机上完成 500,000 完成训练。

基于 DehazeNet 估计得的大气透射率图和大气散射模型, 清晰图片的重构过程与传统图像去雾方法一致。由于第三层中的局部极值, DehazeNet 估计的透射率图存在块效应 (Blocking Artifact)。为了进一步优化透射率图, 本文采用导向滤波器^[85] (Guided Image Filter) 对透射率进行边缘保持的平滑处理。根据式(2-2), 本文选择透射率图最低的 0.1% 值作为 t_0 , 并选择其对应的有雾图像 $I(x)$ 中最亮的像素点 $x \in \{y | t(y) \leq t_0\}$ 作为大气光 α 。因此, 在给定大气透射率 $T(x)$ 和大气光 α 的情况下, 可以简单地重构出清晰图像 $J(x)$ 。重构公式(2-1)被重写如:

$$J(x) = \frac{I(x) - \alpha(1 - T(x))}{T(x)}. \quad (2-14)$$

虽然 DehazeNet 是基于卷积神经网络, 但其轻巧的网络结构有效地保证了其实时性。在 MATLAB 2014A 的测试环境中, 完整的去雾框架在单 CPU (Intel i7 3770, 3.4GHz) 下对 640×480 图像的处理只需大约 1.5 秒。

2.3.1 网络结构分析

在 DehazeNet 中, 特征提取 F_1 和非线性回归 F_4 是两个针对大气透射率估计而专门设计的模块。为了验证其有效性, 本文以两个同样具有 3 层结构的传统卷积神经网络 (SRCNN^[71] 和 CNN-L^[86]) 作为基准, 其中 DehazeNet、SRCNN 和 CNN-L 的参数量分别是 8,240、18,400 和 67,552。

2.3.1.1 特征提取中的 Maxout

DehazeNet 特征提取模块中的激活函数是一种非线性的降维过程, 近似于传统的去雾特征提取。在图像处理领域, 低维映射是发掘主要成分和减少模式噪声的重要手段。例如, 主成分分析^[87] (Principal Component Analysis, PCA) 和线性判别分析^[88] (Linear Discriminant Analysis, LDA) 作为经典的低维映射, 被广泛地应用于计算机视觉和数据挖掘。SRCNN^[71] 采用 ReLU 的非线性稀疏映射实现高分辨率图像的重建。Maxout 作为特殊的低维映射, 通过最大化特征响应值以发掘有雾图像的先验知识。因此, 以下实验旨在验证 Maxout 单元的有效性。根据 SRCNN^[71], 线性映射 (Linear) 将 16 维向量映射到 4 维向量 (相当于应用 4 个 $16 \times 1 \times 1$ 的过滤器), 稀疏映射通过 ReLU 实现线性输出的非线性变换。

图2-7展示 DehazeNet 第一层采用不同激活函数 (Maxout、ReLU 和 Linear) 的训练过程。如图所示, 采用 Maxout 的网络的收敛速度相比 ReLU 和 Linear 更快。此外, 图2-7内括号中的值表示收敛结果。相比于 ReLU 和 Linear, Maxout 的性能提高了约

0.30×10^{-2} 。如节2.2.2.2分析，Maxout 可近似地等效于与雾相关的特征，并解决了 ReLU 简单二分段的缺点。

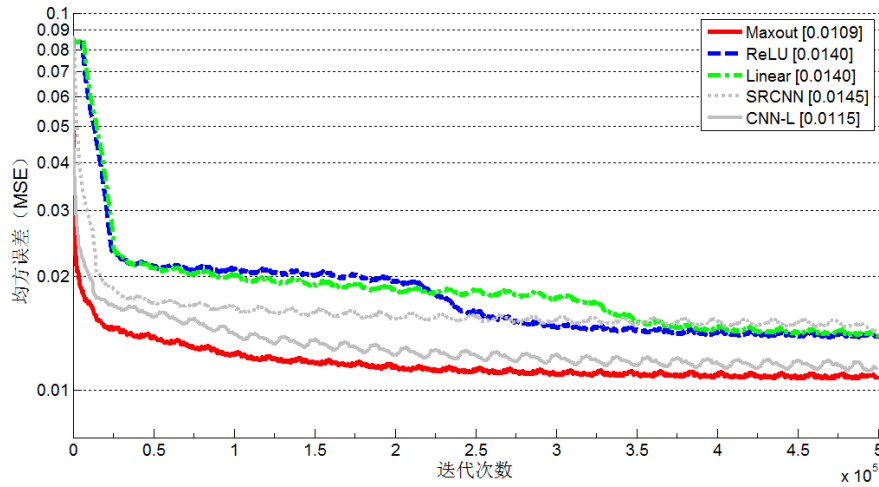
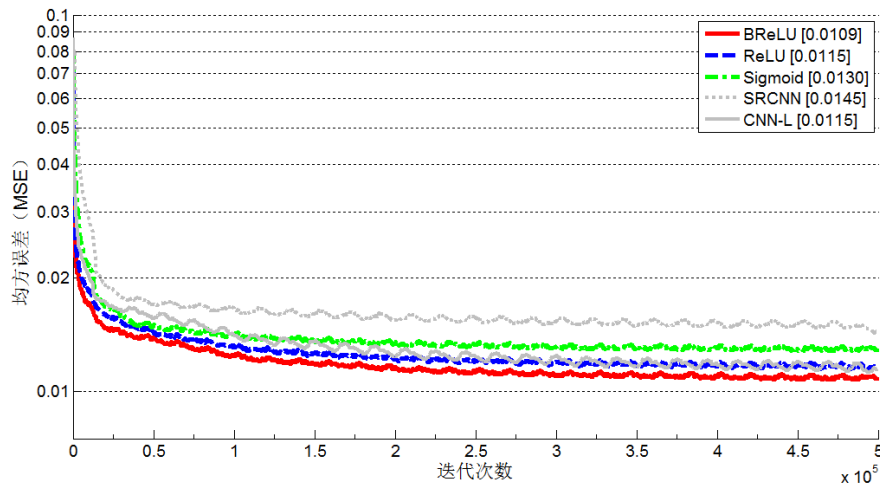
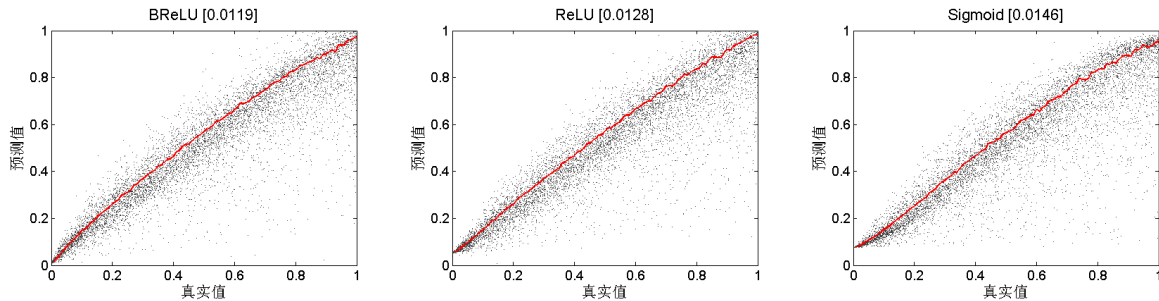


图 2-7 F_1 在不同降维映射下的训练过程

2.3.1.2 非线性回归中的 BReLU

本文提出一种新的激活函数 BReLU 用于图像恢复和重建。受到 ReLU 和 Sigmoid 函数的启发，BReLU 借鉴 Sigmoid 函数的双边约束和 ReLU 的局部线性设计。双边约束符合重构先验并缩小求解空间；局部线性克服梯度弥散以获得更好的结果。在对比实验中，ReLU 和 Sigmoid 函数分别代替 BReLU 在非线性回归层中的作用。对于 ReLU， F_4 可以重写为 $F_4 = \max(0, W_4 * F_3 + B_4)$ ；对于 Sigmoid 函数，其可重写为 $F_4 = 1/(1 + \exp(-W_4 * F_3 - B_4))$ 。

图2-8展示了 F_4 在不同激活函数下的训练过程。从收敛速度上，BReLU 具有比 ReLU 和 Sigmoid 更好的收敛效率，特别是在前 50,000 次迭代期间。从收敛精度上，BReLU 的性能相比于 ReLU 提高了约 0.05×10^{-2} ，相比于 Sigmoid 提高了 0.20×10^{-2} 。图2-9展示了在图像块上透射率估计的预测值与真实值对应的散点图和回归曲线。明显地可见，BReLU 的预测中心线在 45 度附近。ReLU 的预测值均高于真实值，且存在超过极限值 $t_{\max} = 1$ 的不合理结果。由于 Sigmoid 函数的两端饱和区，其在 0 和 1 附近区间的预测值远远偏离真实值。从定量分析上，在测试集上 BReLU 的均方误差为 1.19×10^{-2} ，ReLU 为 1.28×10^{-2} ，Sigmoid 为 1.46×10^{-2} 。

图 2-8 F_4 在不同激活函数下的训练过程图 2-9 F_4 在不同激活函数下的回归曲线

2.3.2 滤波器个数与尺度

为研究网络性能和网络容量之间的最佳平衡，本节逐步修改 DehazeNet 的网络参数并分析其对精度的影响。根据 DehazeNet 的默认设置，进行了两个实验：（1）增加滤波器的数量；（2）采用不同的滤波器尺度。本实验在同一数据集上训练和测试（如小节 2.2.3），表 2-2 展示不同的参数设置与训练和测试过程中的均方误差。

通常地，增加网络宽度会带来拟合能力的提高。从表 2-2 中明显可见，增加滤波器的数量可带来性能的显著提高。然而，在保证运算效率的情况下，轻巧的网络结构仍可优于现有的其他方法。因此，以下实验中采用“4 - (16 × 3)”的网络结构。

本节着重分析多尺度滤波器对网络性能的影响。DehazeNet 的默认参数设置（详见表 2-1）表示为“5-M-7-6”。首先，分析第二层 F_2 中滤波器尺度对性能的影响。根据表 2-2 分析如下：在 F_2 中，大的滤波器尺度可抓取更丰富的结构信息，从而获得更好的预测精度。在 DehazeNet 中， F_2 采用了 3/5/7 三个尺度的滤波器组合实现丰富的结构信息提取，其在更小网络容量下实现了与 7×7 单尺度相当的测试精度。此外，小

表 2-2 DehazeNet 在不同滤波器参数下的性能 ($\times 10^{-2}$)

滤波器	结构	训练 MSE	测试 MSE	参数量
个数 (n_1-n_2)	4-(16$\times$3)	1.090	1.190	8,240
	8-(32 $\times$ 3)	0.972	1.138	27,104
	16-(64 $\times$ 3)	0.902	1.112	96,704
F_2 尺度 ($f_1-f_2-f_3-f_4$)	5-3-7-6	1.184	1.219	4,656
	5-5-7-6	1.133	1.225	7,728
	5-7-7-6	1.021	1.184	12,336
	5-M-7-6	1.090	1.190	8,240
F_4 尺度 ($f_1-f_2-f_3-f_4$)	5-M-6-7	1.077	1.192	8,864
	5-M-7-6	1.090	1.190	8,240
	5-M-8-5	1.103	1.201	7,712

节2.2.1将展示多尺度映射对尺度鲁棒性的效果。

进一步分析第三层 F_3 和第四层 F_4 中不同尺度的滤波器对网络性能的影响。保持网络的整体感受野不变，同时调整 F_3 和 F_4 的滤波器尺度。结果表明，在非线性回归 F_4 中采用的大尺度滤波器可增强拟合能力，但同时也可能导致过拟合；局部极值 F_3 中的最大池化可提升网络在测试集上的泛化性。因此，在本文中选取“5-M-7-6”作为 F_3 和 F_4 的最佳参数设置。

2.3.3 合成图像块的定量分析

近年来，基于特征学习的图像去雾方法主要包括三种：（1）基于颜色衰减先验^[46]（Color Attenuation Prior, CAP）通过线性模型学习颜色衰减与场景深度的映射关系；（2）反向传播神经网络^[45]（Back Propagation Neural Network, BPNN）通过多层感知机从训练样本中挖掘颜色和深度之间的映射关系；（3）随机森林^[47]（Random Forests, RF）融合多种传统去雾特征实现大气透射率的估计。本实验中，上述所有方法和 DehazeNet 均采用与 RF^[47] 相同的测试设置。根据此测试基准，从清晰图片中随机采样 2000 个图像块，通过 10 个随机的大气传输率 $t \in (0, 1)$ 合成 20,000 个用于测试样本，并评估预测值和真实值之间的均方误差。其中，采用经典的去雾方法——暗通道先验^[74]（Dark Channel Prior, DCP）作为对比基线。

表2-3展示在合成图像块上的定量分析结果。可见，DehazeNet 取得了最好的预测结果 (1.19×10^{-2})，相比于当前最好方法 RF^[47] 的均方误差降低了 0.07×10^{-2} 。因

为，RF^[47] 通过图像块的特征值重排列来打破特征与图像内容之间的相关性。相反地，DehazeNet 认为图像的内容信息（天空区域和白色物体）对于大气透射率的估计是有效的。此外，在后续实验中取得不错效果的 CAP^[46] 由于线性回归带来的边界异常值（大于 1 或小于 0），从而导致在本测试中表现不佳。

表 2-3 在合成图像块中不同方法的大气透射率估计（ $\times 10^{-2}$ ）

方法	DCP ^[74]	BPNN ^[45]	CAP ^[46]	RF ^[47]	DehazeNet
均方误差	3.18	4.37	3.32	1.26	1.19

2.3.4 合成图片的定量分析

为验证 DehazeNet 对完整图像的去雾有效性，本文在已知深度图 $d(x)$ 的立体图像上合成有雾图像，并与 DCP^[74]、FVR^[58]、BCCR^[62]、ATM^[84]、CAP^[46]注 1 和 RF^[47] 进行定量分析。本文总共在 Middlebury 立体数据库（2001-2006）^[89-91] 中收集了 12 对立体图像。如图2-10所示，根据式(2-1)和立体图像合成有雾图像，并使用 DehazeNet 对其进行恢复。

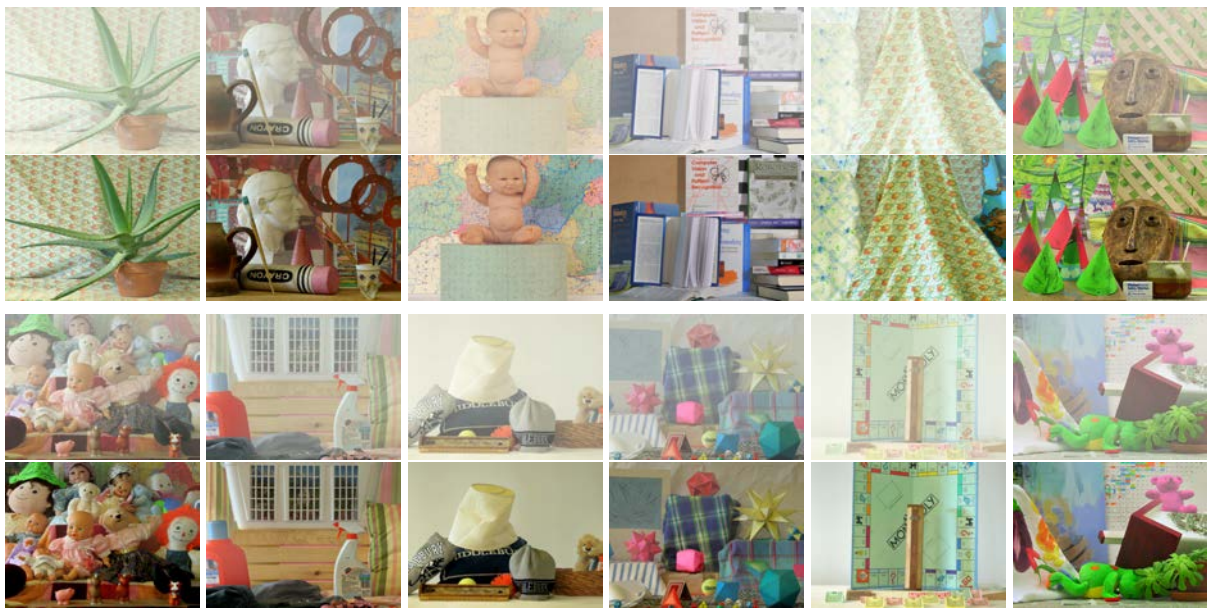


图 2-10 基于 Middlebury 立体数据库合成的有雾图像与 DehazeNet 的去雾结果

为定量评估以上方法，本文采用一系列的度量指标分析清晰图像和去雾结果之间的差异性，其中包括：均方误差（Mean Square Error，MSE）、结构相似性^[92]（Structural Similarity，SSIM）、峰值信噪比（Peak Signal-to-noise Ratio，PSNR）和加权峰值信噪

注 1 括号外为作者^[46] 提供代码实现结果，括号内为本文重新复现的结果。

比^[93] (Weighted Peak Signal-to-Noise Ratio, WPSNR)。在本文中, 将通用评估方法定义为一次性评估 (One-pass Evaluation, OPE)——使用标准参数设置 (散射系数 $\beta = 1$ 和大气光 $\alpha = 1$) 合成有雾图像并分析平均度量结果。如表2-4中所示, 在 OPE 评估中, DehazeNet 的去雾效果均优于当前 6 种最优方法。虽然 DehazeNet 仅采用最小均方误差进行优化, 但其在多项评价指标上均取得最佳性能。

表 2-4 在合成图片中不同方法的性能评估 ($\beta = 1$ 和 $\alpha = 1$)

指标	有雾图像	ATM ^[84]	BCCR ^[62]	FVR ^[58]	DCP ^[74]	CAP ^[46]	RF ^[47]	DehazeNet
MSE	0.0481	0.0689	0.0243	0.0155	0.0172	0.0075 (0.0068)	<u>0.0070</u>	0.0062
SSIM	0.9936	0.9890	0.9963	0.9973	0.9981	<u>0.9991</u> (0.9990)	0.9989	0.9993
PSNR	61.5835	60.8612	65.2794	66.5450	66.7392	70.0029 (70.6581)	<u>70.0099</u>	70.9767
WSNR	8.5958	7.8492	12.6230	13.7236	13.8508	16.9873 (17.7839)	<u>17.1180</u>	18.0996

图像去雾的效果对于雾的浓度非常敏感, 不同的散射系数 β 会影响方法的鲁棒性。因此, 本文提出系数鲁棒性评估 (Coefficient Robustness Evaluation, CRE) 来分析散射系数 $\beta \in \{0.75, 1.0, 1.25, 1.5\}$ 对去雾效果的影响。如表2-5所示, CAP^[46] 在轻雾 ($\beta = 0.75$) 上取得很好的表现, 但随着雾浓度的增加去雾性能逐渐降低, 其原因是 CAP 通过预测场景深度 $d(x)$ 和假设散射系数 $\beta = 1$ 来估算大气传输率 $T(x)$ 。RF^[47] 通过构建 200 棵树的随机森林进行非线性回归, 因此表现出很好的系数鲁棒性, 但随机森林的密集预测限制了其实时性。DehazeNet 通过 F_1 中的非线性激活函数 (Maxout) 对凸函数有非常优秀的拟合能力, 因此对散射系数有很强的鲁棒性。

由于悬浮颗粒和大气光源的颜色偏移, 大气光并不是完全的纯白色。因此, 本文提出大气光鲁棒性评估 (Airlight Robustness Evaluation, ARE) 来分析不同去雾方法对大气光 α 的泛化能力。尽管 DehazeNet 只通过大气光 $\alpha = 1$ 的样本进行训练, 但式(2-2)的大气光估计方法使得其对于多样化的大气光具有鲁棒性。特别地, 当大气光为 $[1.0, 1.0, 0.9]$ 时, DehazeNet 的性能优于现有的其他方法。因此, DehazeNet 也可应用于光晕的去除 (如图2-11所示)。

实际的成像过程中普遍存在视野变换和图像缩放, 尺度鲁棒性评估 (Scale Robustness Evaluation, SRE) 用于分析图像缩放对去雾性能的影响。不同于 OPE 仅采用单一尺度的测试图片, 在 SRE 中采用从 0.4 到 1.0 间四个不同的尺度比例的图像。如表2-5所示, 得益于 F_2 的多尺度映射, DehazeNet 表现出优秀的尺度鲁棒性。其中, 由于 CAP^[46], DCP^[74] 和 ATM^[84] 仅使用的单一尺度的特征提取, 导致其对于尺度变化敏

表 2-5 在合成图片中不同方法的鲁棒性评估 (MSE)

鲁棒性评估		输入	ATM ^[84]	BCCR ^[62]	FVR ^[58]	DCP ^[74]	CAP ^[46]	RF ^[47]	DehazeNet
CRE $\beta =$	0.75	0.0311	0.0581	0.0269	0.0122	0.0199	0.0043	<u>0.0046</u>	0.0063
	1.00	0.0481	0.0689	0.0243	0.0155	0.0172	0.0077	<u>0.0070</u>	0.0062
	1.25	0.0658	0.0703	0.0230	0.0219	0.0147	0.0141	<u>0.0109</u>	0.0084
	1.50	0.0833	0.0683	0.0219	0.0305	0.0134	0.0231	<u>0.0152</u>	0.0127
CRE 均值		0.0571	0.0653	0.0254	0.0187	0.0177	0.0105	<u>0.0094</u>	0.0084
ARE $\alpha =$	[1, 1, 1]	0.0481	0.0689	0.0243	0.0155	0.0172	0.0075	<u>0.0070</u>	0.0062
	[0.9, 1, 1]	0.0437	0.0660	0.0266	0.0170	0.0210	0.0073	0.0071	<u>0.0072</u>
	[1, 0.9, 1]	0.0435	0.0870	0.0270	0.0159	0.0200	0.0070	<u>0.0073</u>	0.0074
	[1, 1, 0.9]	0.0421	0.0689	0.0239	0.0152	0.0186	<u>0.0081</u>	0.0083	0.0062
ARE 均值		0.0443	0.0727	0.0255	0.0159	0.0192	0.0075	<u>0.0074</u>	0.0067
SRE $s =$	0.40	0.0478	0.0450	0.0238	0.0155	0.0102	0.0137	<u>0.0089</u>	0.0066
	0.60	0.0480	0.0564	0.0223	0.0154	0.0137	0.0092	<u>0.0076</u>	0.0060
	0.80	0.0481	0.0619	0.0236	0.0155	0.0166	0.0086	<u>0.0074</u>	0.0062
	1.00	0.0481	0.0689	0.0243	0.0155	0.0172	0.0077	<u>0.0070</u>	0.0062
SRE 均值		0.0480	0.0581	0.0235	0.0155	0.0144	0.0098	<u>0.0077</u>	0.0062
NRE $\sigma =$	10	0.0484	0.0541	0.0138	0.0150	0.0133	<u>0.0065</u>	0.0086	0.0059
	15	0.0488	0.0439	0.0144	0.0148	0.0104	<u>0.0072</u>	0.0112	0.0061
	20	0.0493	—	0.0181	0.0151	0.0093	<u>0.0083</u>	0.0143	0.0058
	25	0.0500	—	0.0224	0.0150	<u>0.0082</u>	0.0100	0.0155	0.0051
	30	0.0508	—	0.0192	0.0151	<u>0.0085</u>	0.0119	0.0191	0.0049
NRE 均值		0.0495	—	0.0255	0.0150	0.0100	<u>0.0088</u>	0.0137	0.0055

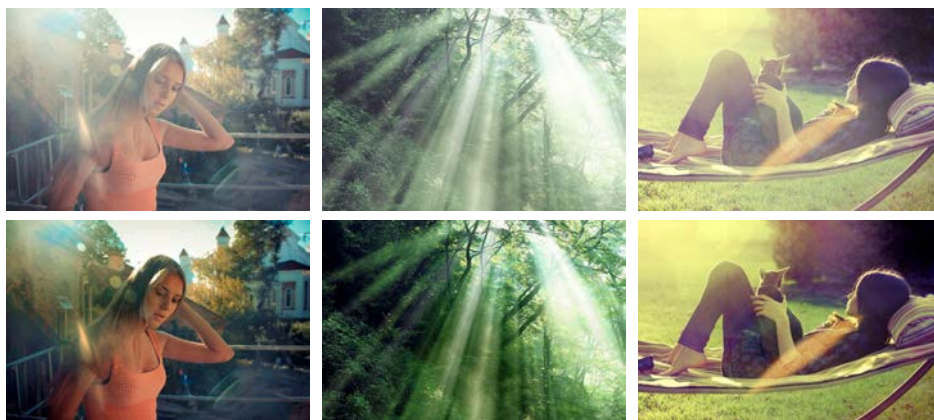


图 2-11 DehazeNet 对光晕图像的增强

感——当图像缩小时，过大的特征邻域将导致细节丢失；当图片放大时，过小的感受野将导致信息缺失。DehazeNet 的多尺度映射通过整合多样化特征空间，并在不同尺度下均获得最优效果。

在数字成像过程中，传感器电路所产生的随机热噪声会引起去雾方法的估计误差。因此，噪声鲁棒性评估（Noise Robustness Evaluation, NRE）将讨论不同程度的图像噪声对去雾方法的影响。本文选取标准差 $\sigma \in \{10, 15, 20, 25, 30\}$ 的加性高斯白噪声（Additive White Gaussian, AWG）作为基本噪声模型。受益于 F_1 中 Maxout 和 F_3 中 Max 池化的最值抑制，DehazeNet 展现出远优于其他方法的噪声鲁棒性。其中，在多个鲁棒性评估中性能优秀的 RF^[47] 采用特征值重排列打破特征与内容的相关性，从而放大异常值对于模型性能的影响，因此在 NRE 评估中严重受到噪声的影响。

2.3.5 RESIDE 数据库的定量分析

Li 等人^[94] 提出了一个大规模的真实图像去雾评价数据库 RESIDE（Realistic Single Image Dehazing）作为去雾质量评价标准。RESIDE 数据库包含 13990 张合成的有雾图像，其采用 1399 张包含深度图的室内立体数据库生成，包括：NYU2 数据库^[95] 和 Middlebury 数据库^[89-91]。其中，13000 张作为训练集，990 张作为验证集。RESIDE 数据库的测试集包括合成客观测试集（Synthetic Objective Testing Set, SOTS）和混合主观测试集（Hybrid Subjective Testing Set, HSTS）。SOTS 包括 500 张合成的室内有雾图像；HSTS 包括 10 张合成的室外有雾图像与 10 张真实的室外有雾图像。

在 RESIDE 的评测中，选择 9 种代表性的图像去雾算法，包括上节中的 DCP、FVR、BCCR、CAP，并增加了最新的图像去雾方法：非局部图像去雾^[96]（Non-local Image Dehazing, NLD）、多尺度卷积网络^[97]（Multi-scale CNN, MSCNN）、合一去雾网络^[98]（All-in-One Dehazing Network, AOD-Net）。其中，所有数据驱动的方法（CAP、MSCNN、AOD-Net、DehazeNet）均统一在 RESIDE 训练集上重新训练。基于经典重构度量指标（PSNR 和 SSIM），表2-6展示不同雾浓度下各个方法在 SOTS 测试集中的评估结果，表2-7展示不同方法在 HSTS 测试集中的评估结果。综上所述，DehazeNet 取得优异的去雾效果，甚至优于后续基于深度学习的图像去雾方法，如：MSCNN 和 AOD-Net。

表 2-6 不同雾浓度下在 RESIDE 数据库中 SOTS 的评估

指标	DCP ^[40]	FVR ^[58]	BCCR ^[62]	CAP ^[46]	NLD ^[96]	MSCNN ^[97]	AOD-Net ^[98]	DehazeNet
$\beta = 1.0$								
PSNR	16.62	15.72	16.88	19.05	17.29	17.57	<u>19.06</u>	21.14
SSIM	0.8179	0.7483	0.7913	0.8364	0.7489	0.8102	0.8504	<u>0.8472</u>
$\beta \in [0.6, 0.9]$								
PSNR	16.10	17.18	16.91	20.88	17.52	19.72	<u>22.40</u>	24.24
SSIM	0.8158	0.7682	0.7978	0.8597	0.7558	0.8489	<u>0.8980</u>	0.9044
$\beta \in [1.0, 1.4]$								
PSNR	16.58	16.00	17.07	<u>19.68</u>	17.37	17.25	19.61	22.02
SSIM	0.8210	0.7538	0.7942	0.8450	0.7487	0.8110	<u>0.8616</u>	0.8870
$\beta \in [1.5, 1.8]$								
PSNR	17.15	14.42	17.14	<u>17.21</u>	17.06	15.10	16.16	18.67
SSIM	<u>0.8259</u>	0.7289	0.7906	0.8120	0.7438	0.7723	0.8064	0.8454

表 2-7 在 RESIDE 数据库中 HSTS 的评估

指标	DCP ^[40]	FVR ^[58]	BCCR ^[62]	CAP ^[46]	NLD ^[96]	MSCNN ^[97]	AOD-Net ^[98]	DehazeNet
PSNR	14.84	14.48	15.08	<u>21.53</u>	18.92	18.64	20.55	24.48
SSIM	0.7609	0.7624	0.7382	0.8726	0.7411	0.8168	<u>0.8973</u>	0.9153

2.3.6 真实图片的定性分析

图2-12展示 DehazeNet 重构的清晰图像和深度图，更多结果详见<https://caibolun.github.io/DehazeNet/>。现有的去雾算法在大多数图像上均可获得较好的效果，因此难以在视觉上对其进行排序。为更好地对比，本文重点关注在相关研究中公认极具挑战的图像^[46,47,74]。因为现有的去雾算法普遍对于白色物体敏感，这些图像具有大面积难以处理的白色或灰色区域。图2-13展示了 DehazeNet 与 6 种现有优秀方法的定性比较，其中图2-13a 是有雾图像；图2-13b-g 分别是 ATM^[84]、BCCR^[62]、FVR^[58]、DCP^[74]、CAP^[46] 和 RF^[47] 的去雾结果；图2-13h 是 DehazeNet 的处理结果。

有雾图像的天空区域是去雾的挑战，因为云是和雾具有相同大气散射模型的自然现象。如前三组图像所示，图2-13b-g 方法均可很好地处理大部分区域，场景中的物体细节都得到很好的增强。然而，由于这些方法误判天空区域为有雾区域，导致天空区域的亮度降低和颜色过饱和。实际上，雾通常仅存在于大气层底部，因此天空区域一般并

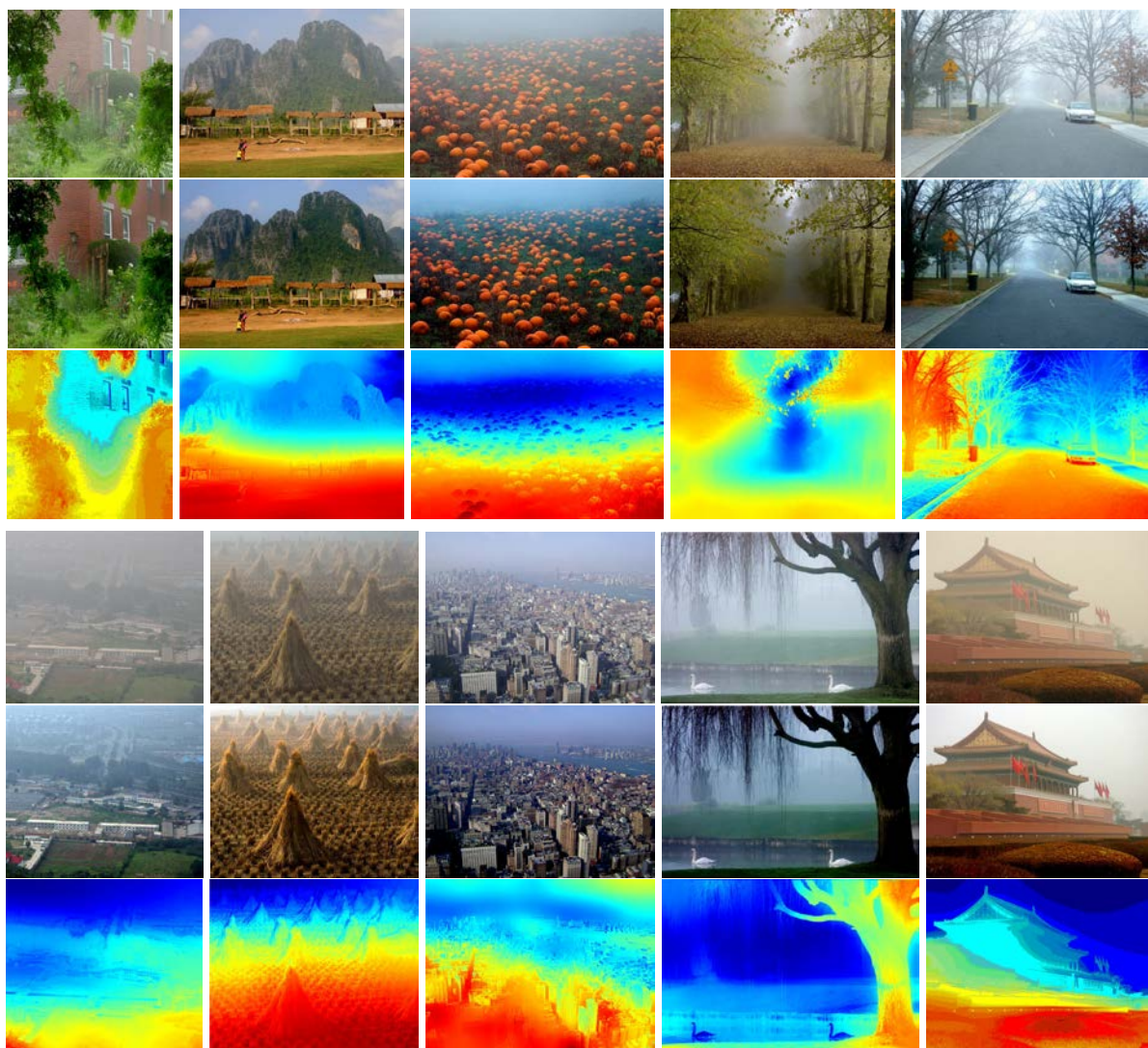


图 2-12 DehazeNet 重构的清晰图像与深度图

不需要去雾处理。基于学习的框架，CAP^[46] 和 RF^[47] 避免了天空中的颜色失真，但由于 CAP 中像素级的回归和 RF 中内容无关的特征学习，导致其非天空区域的增强效果较不理想，例如：第一组图像中的岩石和第三组图像中的草原。受益于 DehazeNet 隐含的局部属性学习，其网络似乎能够找到天空区域并保持原有颜色，同时确保在其他区域有良好的去雾效果。

基于先验的大气透射率估计方法的鲁棒性较弱。当场景中存在类似于大气光的物体，例如：第四组图中的白皙肤色和第五组图中的白色大理石，基于暗通道先验（DCP^[74]、BCCR^[62] 和 FVR^[58]）的大气透射率估计方法差强人意。因为，白色物体在暗通道下具有高的亮度值，其会导致雾浓度的高估。基于回归模型的学习方法（CAP^[46] 和 RF^[47]）解决了过饱和问题，但却低估了远景中雾的浓度（第四组图像中的棕色头发）。与其他六种算法相比，DehazeNet 的非线性回归避免了图像过饱和并取得

更好的去雾效果。

2.4 小结

在本章中，本文提出一种基于深度学习的去雾方法 DehazeNet。受传统方法中的大气散射模型和先验特征提取的启发，大气透射率的估计可通过一个特殊设计的端到端系统训练实现，其中特征提取（输入层）和非线性回归（输出层）针对性地在经典网络结构上做了改进。在输入层中，本文证明了 Maxout 单元与传统方法的去雾先验具有类似的效果，同时在提取去雾相关特征方面更具效率。在输出层中，本文提出一种新的激活函数 BReLU 替代了 ReLU 和 Sigmoid，以保持图像复原的双边限制和局部线性。得益于特殊设计的轻巧结构，DehazeNet 保持高效率的同时取得显著的去雾效果，综合性能上比现有算法都优秀。

需要指出的是，尽管本文第一次成功地将卷积神经网络应用于图像去雾中，但仍有许多需要改进的地方。首先，大气光并不是完全的全局不变，应该可以与透射率一起被模型学习；其次，大气散射模型应该也可在网络中被学习得到，从而实现有雾图像和清晰图像直接的端到端映射。在后续的研究工作中，已有相关工作基于 DehazeNet 的这几个问题进行改进，如：DTN^[99]、MSCNN^[97] 和 AOD-Net^[98]。

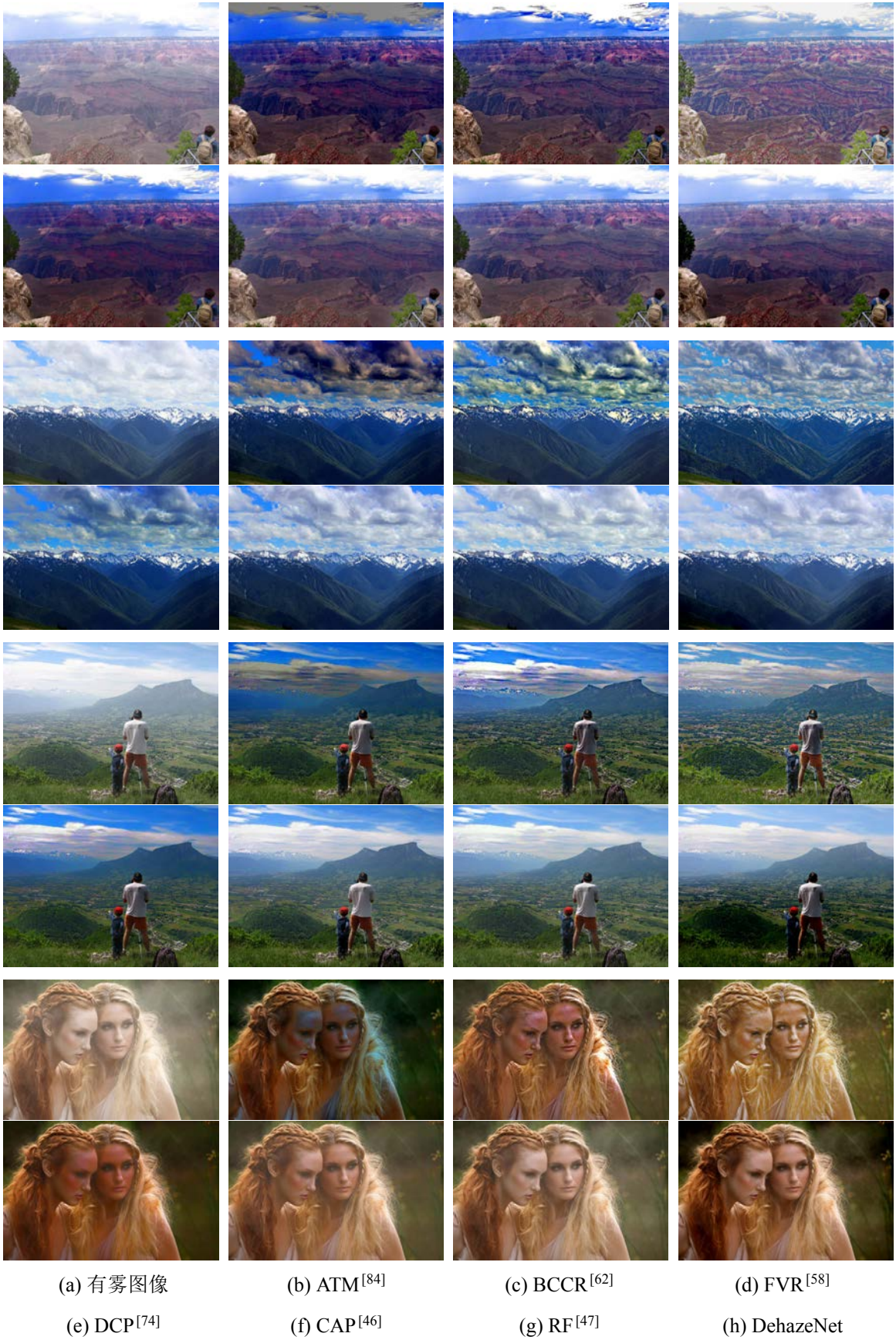


图 2-13 在真实图片中不同方法的定性对比

第三章 基于时空马尔科夫随机场的视频去雾

3.1 概述

雾作为一种常见的大气现象，其不仅影响静态图像成像过程，也影响动态视频的成像过程，从而导致视频的对比度和饱和度下降。相比于图像去雾，视频去雾有着更为广泛的应用场景，例如：自动驾驶、视频监控、行车记录仪等。然而，由于雾浓度具有时空相关性，使得视频去雾成为一个更具挑战的问题。

近几年，各种图像增强技术被广泛应用于单帧图像的去雾，例如直方图均衡化^[1]和伽马校正^[100]等。然而，这些基于颜色和纹理增强的方法，在不关注雾浓度的情况下改变图像的颜色或纹理分布，从而导致颜色失真或者对比度饱和。此外，部分去雾方法引入附加信息，如多帧图像^[101,102]或深度图像^[50,61]，但附加信息在真实场景中难以直接获得。最近，以暗通道^[74]、颜色衰减^[46]、色相差异^[75]为代表的先验特征在图像去雾上取得重大进展。然而，上述方法均仅关注于单帧图像去雾，并不考虑时序的相关性。因此，直接应用单帧图像去雾方法于视频去雾时会给结果带来闪烁和伪影。

与大量的单帧图像去雾工作相比，视频去雾方面的研究相对较少。Tarel 等人^[57]在车载视频流中应用分割技术对运动物体和平面道路进行有效的切分，并分别应用单帧图像去雾方法^[58]在线更新深度估计。受限于切割方法的场景特定性，因此其并无法广泛地应用于通用的视频去雾场景。为引进时空相关性，光流法^[59]也被应用于大气透射率估计，但光流法计算的复杂度和光流信息的融合限制了光流在实时视频中的应用。Kim 等^[60]为了消除视频去雾中的闪烁效应，通过最小化时序相关性来优化并增强对比度，但当对比度过度拉伸时会导致饱和度截止，从而进一步导致计算密集。Li 等人^[61]联合深度图估计和浓度图估计实现对清晰视频的重构，但基于运动结构估计（Structure-from-Motion, SfM）的深度图估计类似于光流法需要大量匹配计算。此外，由于远景物体的视差较小，其在远场景的去雾中并不能取得很好效果。

综上所述，将单帧图像去雾算法扩展至视频处理是一项有挑战性的工作，其主要挑战来自如下：

- 空间一致性。空间一致性包括两个方面的约束。（1）大气透射率的空间一致性：大气透射率具有局部一致从而克服估计噪声；（2）重构图像的空间一致性：去雾后的结果具有分段平滑以保证图像帧内的连续性和自然性。

- 时间连续性。人类的视觉系统对时序信息的不连续非常敏感。然而，逐帧地应用单帧图像处理方法很难保证增强后结果在时序上的连续性，并很可能给重构视频带来严重的闪烁和伪影。
- 计算有效性。一个实用的视频处理算法必须能够有效地、实时地处理视频序列中大量的图像帧和像素点。特别地，视频去雾方法至少应达到 24 帧每秒的实时处理速度，以满足人眼的视觉残留效应。

本文通过亮度值先验（Intensity Value Prior, IVP）构建时空马尔科夫随机场（Spatio-temporal Markov Random Field, ST-MRF）来优化基于暗通道先验（Dark Channel Prior, DCP）的雾浓度估计，从而实现实时的视频去雾。该方法有效地保证了视频去雾面临的时空一致性和时间连续性问题。此外，本方法通过积分图像技术^[103]实现 $O(N)$ 时间复杂度的算法加速，并应用降采样技术进一步减少冗余计算量。通过这一系列性能和速度优化，本文方法可在单 CPU 下实现对通用影像传输格式（Common Intermediate Format, CIF） 352×288 尺寸视频流约 240 帧每秒的实时处理。

3.2 基于 ST-MRF 的实时视频去雾

目前，大部分单帧图像去雾方法普遍基于大气透射率估计，并可在大多数去雾任务中取得较好的视觉增强结果。然而，对于视频去雾问题，若简单地逐帧应用单帧图像去雾方法，将破坏视频的时空相关性，并产生不和谐的区块效应和闪烁效应。此外，现有图像去雾方法所需的时间复杂度限制其在实时性场景中的应用。在本节中，本文提出了一个联合时空信息的优化框架实现实时的视频去雾。如图3-1所示，首先应用 DCP 实现对大气透射率的初步估计，并基于 IVP 联合建立帧内信息与帧间信息的 ST-MRF，从而实现对视频的实时性处理。

3.2.1 单帧图像去雾

单帧图像去雾是经典的图像增强问题。基于观测经验，近几年的图像去雾研究提出各种各样的先验特征以估计大气透射率，其主要包括：暗通道^[74]、最大对比度^[35]、色相差异^[75]、颜色衰减^[46]等。基于大气透射率估计，可以简单地通过大气散射模型实现对清晰图像的重构。本章重点讨论视频去雾中时空信息的融合，因此在此节中将采用并简要地介绍基于暗通道先验的大气透射率估计方法。

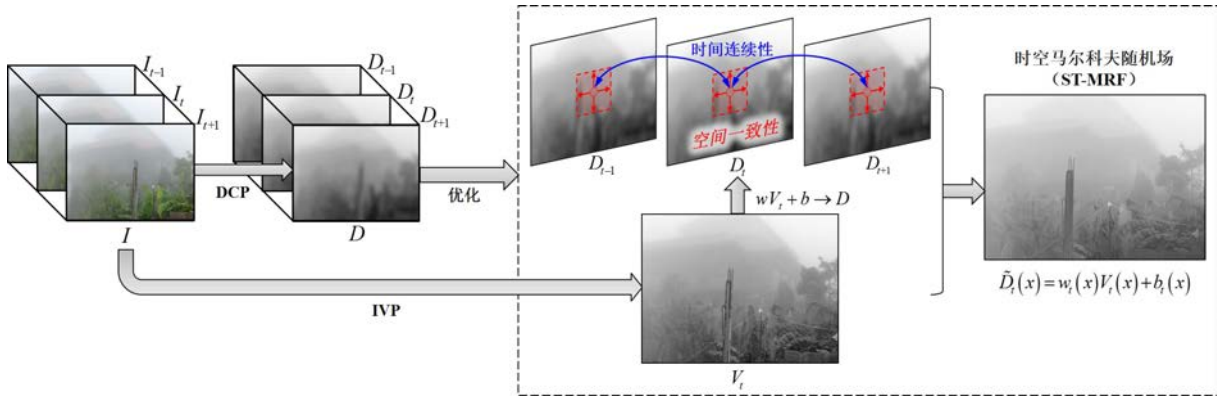


图 3-1 基于时空 MRF 的实时视频去雾

3.2.1.1 大气散射模型

为了对有雾图像建模，McCartney^[31] 提出大气散射模型以描述图像在雾霾环境中的成像过程。如节2.1所述，大气散射模型中对有雾图像描述为式(2-1)。当大气透射率 $T(x)$ 和大气光 α 得到很好地估计时，清晰图像 $J(x)$ 可通过有雾图像 $I(x)$ 有效地重构。其中，大气透射率 $T(x)$ 描述未被散射到达摄像头的光线，大气光 α 作为全局常数可通过最低大气透射率点进行估计（见式(3-12)）。因此，大气透射率估计是单帧图像去雾的重点和难点。

3.2.1.2 大气透射率估计

暗通道先验^[74] 是基于对大量户外清晰图像的统计发现。在绝大多数清晰图像块中，某一些像素的 RGB 颜色空间中至少有一个颜色通道具有很低的值，甚至接近于零。如式(4-1)，暗通道定义为 RGB 颜色空间中图片对应像素的最小通道： $D(x) = \min_{c \in \{r, g, b\}} I^c(x)$ 。根据暗通道的定义，清晰图像的暗通道在大部分区域都是黑的。如图3-2b所示，清晰图像 $J(x)$ 中的暗像素无处不在，包括：阴影、彩色物体、黑色物体等，可表示如下：

$$\min_{c \in \{r, g, b\}} J^c(x) \rightarrow 0 \quad (3-1)$$

如图3-2a所示，有雾图像的暗通道与雾浓度值高度相关，令 $\min_{c \in \{r, g, b\}} J^c(x) = 0$ ，其可推导如下：

$$\begin{aligned} D(x) &= \min_{c \in \{r, g, b\}} I^c(x) \\ &= T(x) \min_{c \in \{r, g, b\}} J^c(x) + \alpha(1 - T(x)) \\ &= \alpha(1 - T(x)) \end{aligned} \quad (3-2)$$



图 3-2 暗通道先验示例

综上所述，暗通道 $D(x)$ 等价于雾浓度，并与大气透射率成反比。因此，通过暗通道先验 $D(x)$ 可实现大气透射率 $T(x)$ 的估计：

$$T(x) = \frac{1 - \omega D(x)}{\alpha}. \quad (3-3)$$

其中，超参数 ω 是暗通道（雾浓度）到大气透射率的映射系数，在本章中的所有实验中其统一选取为 $\omega = 0.7$ 。

3.2.2 实时视频去雾

为了消除区块效应和闪烁效应，本文通过时空相关性来修正雾浓度估计。以暗通道 $D(x)$ 的雾浓度图为基础，本方法基于亮度值先验建立转换函数，通过时空 MRF 来微调浓度图估计，从而实现时空信息的融合并消除区块与闪烁效应。

3.2.2.1 亮度值先验

通过对有雾图像的观察发现，有雾图像中像素点的亮度值随着雾浓度的增加而增大。从主观上阐述亮度值先验，图3-3以一个通过已知雾度浓度合成的图像为例（此图引用自图2-10）。如图所示，雾浓度与亮度值的残差接近于零，此表明雾浓度与亮度值高度相关。从大气光散射模型 $\alpha(1 - T(x))$ 可同样推导出此结论，白色或灰色大气光

$\alpha \approx 1$ 将使有雾图像的亮度 $I(x)$ 与雾浓度相关 $(1 - T(x))$ 。因此, 受大气散射影响, 亮度值随雾浓度增加而增大。

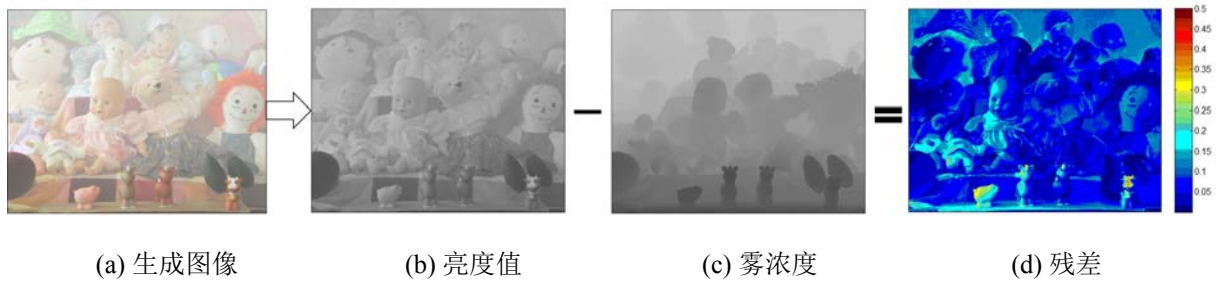


图 3-3 亮度值先验

3.2.2.2 时空马尔科夫随机场

基于亮度值先验, 本文建立时空马尔科夫随机场来修正雾浓度 $D(x)$, 从而联合时空信息保证时空一致性与时间连续性。

• 空间一致性

因为, 在真实环境中的图像成像经常存在噪声、灰尘等影响, 图像中的离群像素值将导致雾浓度的估计不准确。因此, 在大多数情况下, 像素级的雾浓度估计并不能取得很好的效果。基于雾浓度的局部一致性假设, 局部滤波器被广泛应用于解决此问题, 其中主要包括: 最小滤波器^[46], 最大滤波器^[35]和中值滤波器^[63]。然而, 局部滤波器的应用将给雾浓度图带来区块效应。为了解决局部一致性与帧内连续性问题, 本文基于亮度值先验构建空间的马尔科夫随机场。在空间邻域中, 亮度值 $V(x)$ 线性变化为雾浓度 $D(x)$, 变化场 $W = \{w(x)\}_{x \in \mathcal{V}}$ 和 $B = \{b(x)\}_{x \in \mathcal{V}}$ 仅与局部上下文信息相关。综上, 空间似然函数表示为:

$$P_s(w, b) \propto \prod_{y \in \Omega(x)} \exp \left(-\frac{\|w(x)V(y) + b(x) - D(y)\|_2^2}{\sigma_s^2} \right). \quad (3-4)$$

其中, $\Omega(x)$ 是以 x 为中心 $r \times r$ 大小的局部邻域, σ_s 是空间相关性系数。

• 时间连续性

联合帧间相关信息可有效避免闪烁效应。雾的浓度随着摄像机和物体运动而改变: 当物体接近摄像机时, 观测到的亮度值将逐渐接近清晰场景亮度; 相反地, 当物体远离摄像机, 观测到的亮度值将逐渐接近大气光强度。因此, 本文根据有雾图像的亮度值 $V(x)$ 自适应地修正场景的雾浓度 $D(x)$ 。如图3-4所示, 当前帧的亮度

值 $V_t(x)$ 可线性变换为相邻帧的雾浓度 $D_{t,t-1}(x)$ ，修正后相邻帧的雾浓度 $\tilde{D}_t(x)$ 和 $\tilde{D}_{t-1}(x)$ 的绝对误差接近于零。

此过程类似于光流估计的块匹配，相邻帧之间的雾浓度可通过亮度值先验转换为当前帧。如空间一致性，时间马尔科夫随机场实现雾浓度的时间连续性，其在时刻 t 的时间似然函数可表示为：

$$P_\tau(w_t, b_t) \propto \prod_{\tau \in [-f, +f]} \exp \left(-\frac{\|w_t(x) V_t(x) + b_t(x) - D_{t+\tau}(x)\|_2^2}{\sigma_\tau^2} \right). \quad (3-5)$$

其中， f 是邻近帧数， σ_τ 是时间相关性系数。

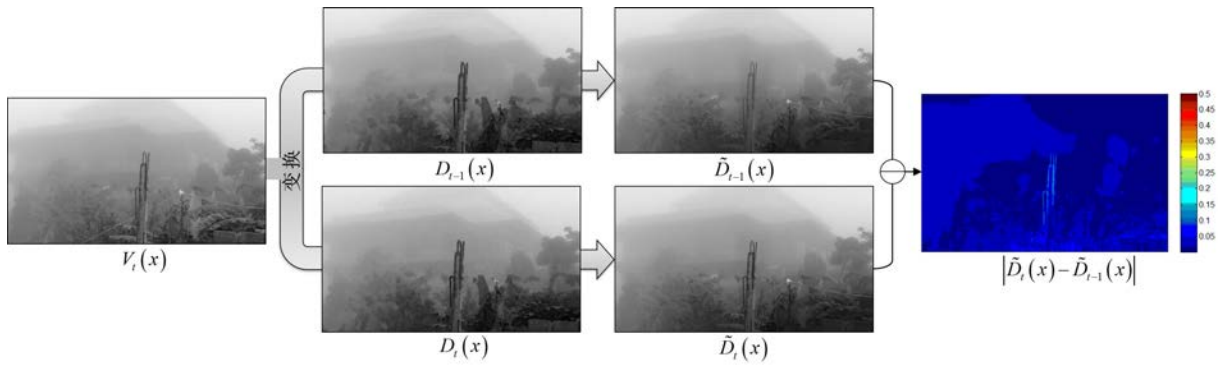


图 3-4 雾浓度的时间相关性

• 最大似然估计

分别沿着空间和时间维度，融合空间一致性（式(3-4)）和时间连续性（式(3-5)），联合似然估计表示如下：

$$P(w_t, b_t) = \prod_{\tau \in [-f, +f]} \prod_{y \in \Omega(x)} \exp \left(-\frac{\|w_t(x) V_t(y) + b_t(x) - D_{t+\tau}(y)\|_2^2}{\sigma_\tau^2} \right). \quad (3-6)$$

其中，式(3-4)中的 σ_s 是一常量，可被省略。

可见式(3-6)是单调函数，对数变换并不影响其最优点的位置，因此采用对数似然函数可更方便地实现求解——对数似然函数中仅存在变量加法，其求导会比原似然函数更容易。设时间权重 $\lambda_\tau = 1/\sigma_\tau^2$ (s.t. $\sum_\tau \lambda_\tau = 1$) 以方便公式表示，式(3-6)可重写为对数似然函数：

$$L(w_t, b_t) = \sum_{\tau \in [-f, +f]} \sum_{y \in \Omega(x)} -\lambda_\tau \|w_t(x) V_t(y) + b_t(x) - D_{t+\tau}(y)\|_2^2. \quad (3-7)$$

通过最大似然估计 $(w_t, b_t) = \arg \max L(w_t, b_t)$ 可找到最优的随机场 W 和 B 。因此，

令式(3-7)的偏导数为 0:

$$\begin{cases} \frac{\partial L(w_t, b_t)}{\partial w_t} = \sum_{\tau \in [-f, +f]} \sum_{y \in \Omega(x)} -\lambda_\tau (V_t^2(y) w_t(x) + b_t(x) V_t(y) - D_{t+\tau}(y) V_t(y)) = 0 \\ \frac{\partial L(w_t, b_t)}{\partial b_t} = \sum_{\tau \in [-f, +f]} \sum_{y \in \Omega(x)} -\lambda_\tau (b_t(x) + w_t(x) V_t(y) - D_{t+\tau}(y)) = 0 \end{cases} \quad (3-8)$$

求解此线性系统可得最大似然估计的最优点:

$$\begin{cases} w_t(x) = \frac{\sum_{\tau \in [-f, +f]} \lambda_\tau (\mathcal{U}_\Omega [V_t(x) D_{t+\tau}(x)] - \mathcal{U}_\Omega [V_t(x)] \mathcal{U}_\Omega [D_{t+\tau}(x)])}{\mathcal{U}_\Omega [V_t^2(x)] - \mathcal{U}_\Omega^2 [V_t(x)]} \\ b_t(x) = \sum_{\tau \in [-f, +f]} \lambda_\tau \mathcal{U}_\Omega [D_{t+\tau}(x)] - w_t(x) \mathcal{U}_\Omega [V_t(x)] \end{cases} \quad (3-9)$$

其中, $\mathcal{U}[\cdot]$ 是定义为 $\mathcal{U}[F(x)] = (1/|\Omega|) \sum_{y \in \Omega(x)} F(y)$ 的均值滤波器, $|\Omega|$ 是滤波器中像素的基数。最后, 时空 MRF 修正后的雾浓度图可表示为 $\tilde{D}_t(x) = w_t(x) V_t(x) + b_t(x)$ 。

3.2.2.3 实时性加速

本文构建的时空 MRF 的一个重要优点是其本身自然具有 $O(N)$ 非近似的速度。其主要的计算瓶颈来自局部邻域统计的均值滤波器 $\mathcal{U}[\cdot]$ 。幸运的是, 使用积分图技术^[103]可在 $O(N)$ 的时间复杂度下有效地实现局部均值计算。积分图又称总和面积表, 统计原点和当前位置形成的矩形区域, 也就是说积分图的每一点的值是原图中对应位置的左上角区域的所有值总和:

$$\mathcal{I}(x_w, x_h) = \sum_{y=(0,0)}^{(x_w, x_h)} F(y). \quad (3-10)$$

其中, (x_w, x_h) 是坐标点 x 对应的二维坐标。如图3-5所示, 一旦预先计算出积分图, 仅需三次加减运算便可求得任何矩形区域上的数值之和:

$$\mathcal{U}[F(x)] = \frac{1}{\Omega} (\mathcal{I}(x_w + w, x_h + h) + \mathcal{I}(x_w, x_h) - \mathcal{I}(x_w + w, x_h) - \mathcal{I}(x_w, x_h + h)). \quad (3-11)$$

因此, 均值滤波器 $\mathcal{U}[\cdot]$ 的计算时间与其滤波器尺度无关, 且节3.2.2.2中的最大似然估计自然是 $O(N)$ 时间复杂度。

此外, 本文采用下采样技术进一步地加速算法实时性。如上所述, 式(3-6)中的优化参数 $w_t(x)$ 和 $b_t(x)$ 是局部平滑的, 修正后的雾浓度图 $\tilde{D}_t(x)$ 的主要内容信息由亮度值先验 $V_t(x)$ 提供。因此, 变换场 $w_t(x)$ 和 $b_t(x)$ 并不需要在全分辨率下进行优化。本文将初始雾浓度图 $D_t(x)$ 和亮度值先验 $V_t(x)$ 采用 s 比率的下采样 (最近邻插值或双

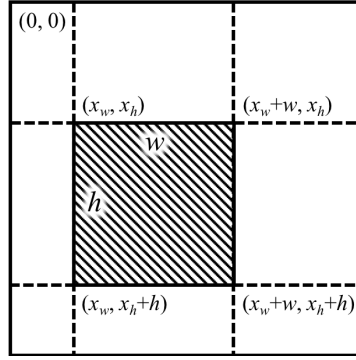


图 3-5 积分图计算示例

线性插值), 所有的均值滤波器均在低分辨率的图像上进行优化。最后, 再将低分辨率的 $w_t(x)$ 和 $b_t(x)$ 重新上采样到原始尺寸, 并与原分辨率的亮度值先验 $V_t(x)$ 修正高分辨率的雾浓度图 $\tilde{D}_t(x)$ 。综上, 所有均值滤波器的时间复杂度从 $O(N)$ 降为 $O(N/s^2)$, 其中增加的采样操作仅占少用运算量。实验表明, 当 $s = 4$ 时, 完整的算法流程大约可获得两倍的加速。

3.3 实验与分析

在本节中, 实验分析基于时空 MRF 的视频去雾方法的有效性, 并与当前最新的视频去雾方法进行对比分析, 其中包括 DCP^[74]、BCCR^[62]、MDCP^[63]、IVR^[57]、OCE^[60] 等。基于大气散射模型和大气透射率估计, 即可根据式(2-1)恢复出清晰的视频。在本章中, 采用图像中最亮区域的颜色作为大气光估计:

$$\alpha = \max_{x \in \mathbb{V}} \left(\min_{y \in \Omega(x)} V(y) \right). \quad (3-12)$$

其中, Ω 为 15×15 的空间邻域。在第 t 帧, 大气光通过 $\alpha_t = \rho\alpha + (1 - \rho)\alpha_{t-1}$ 进行更新, 设大气光更新的学习率 $\rho = 0.1$ 。节3.2中提到的其他参数设置如下: 空间邻域 $r = 41$, 时间邻域 $f = 1$, 时间相关系数 λ_r 为汉宁窗系数。

3.3.1 时间连续性分析

与单帧图像去雾相比, 时间连续性是视频去雾的主要挑战。为了验证本文方法可以有效地抑制闪烁效应, 本文通过对五组清晰视频^{注1}添加均匀雾浓度 $T(x) = 0.6$, 并比较连续帧之间的平均亮度值 (Mean Intensity Value, MIV)。

图3-6展示在 *Suzie* 和 *Foreman* 两个视频中连续帧的平均亮度值。当单帧去雾算法 (包括 DCP^[74]、BCCR^[62]、MDCP^[63]) 独立应用于连续帧时, 去雾后视频序列的 MIV

注¹<http://trace.eas.asu.edu/yuv/>

曲线与原始视频序列相比出现相对较大的波动，尤其是图3-6a中的 50-75 帧和图3-6b中的 175-225 帧。这些闪烁波动是由于缺少时间连续性约束，连续帧之间的色调突然变化引起。此外，表3-1通过去雾图像和原始图像之间 MIV 的相关性分析来定量地分析闪烁效应的强度。相比之下，本文提出的视频去雾方法可以缓解波动并有效地减少闪烁效应。

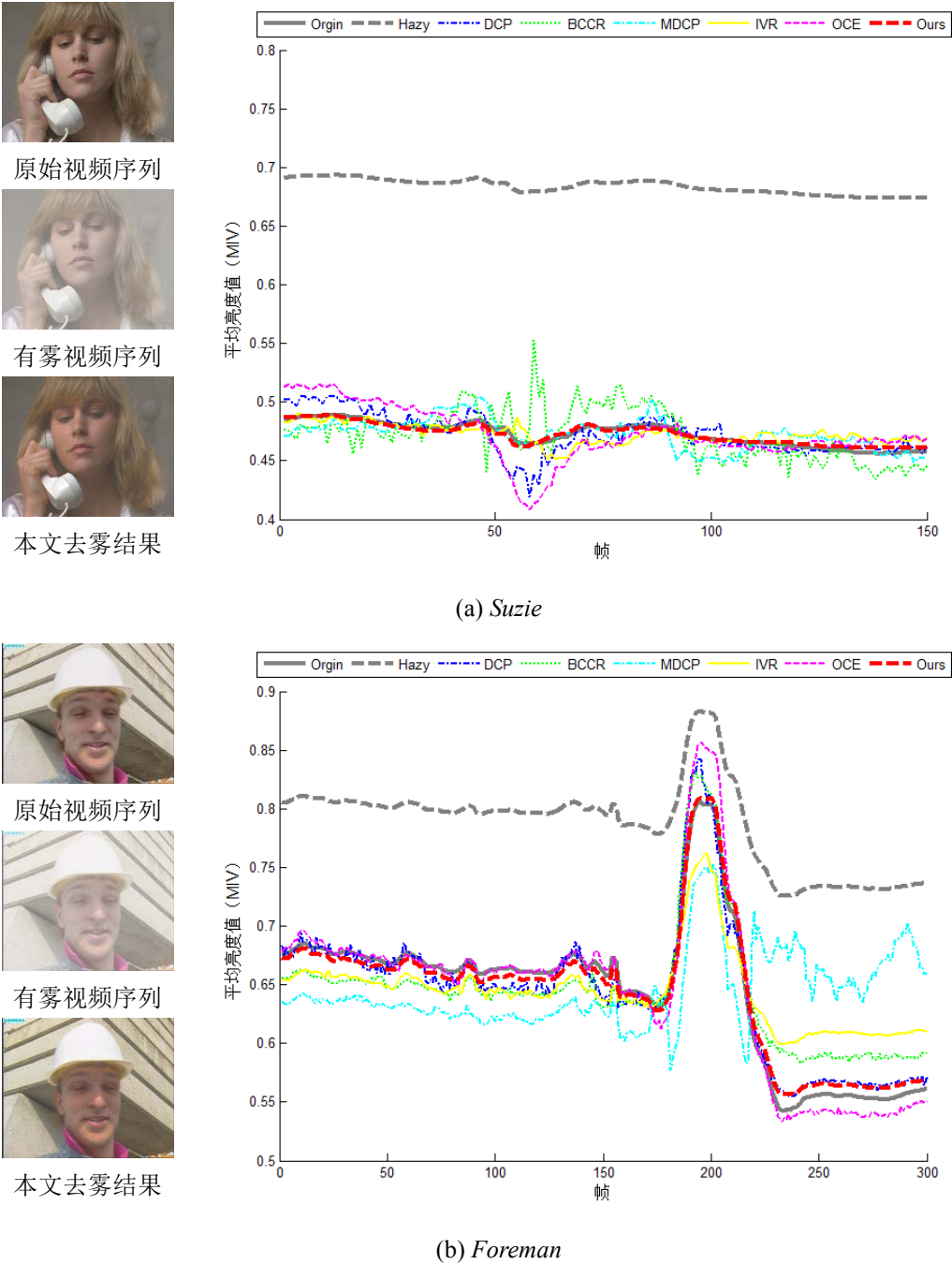


图 3-6 不同视频去雾方法的平均亮度曲线对比

表 3-1 不同视频去雾方法的平均亮度相关性分析

视频序列	DCP ^[74]	BCCR ^[62]	MDCP ^[63]	IVR ^[57]	OCE ^[60]	Ours
<i>Suzie</i>	<u>0.783</u>	0.612	0.641	0.584	0.649	0.976
<i>Foreman</i>	0.980	0.920	0.015	0.949	<u>0.994</u>	0.995
<i>Container</i>	0.929	0.703	0.927	<u>0.998</u>	1.000	0.955
<i>Hall</i>	0.784	0.429	0.444	0.824	<u>0.845</u>	0.991
<i>Silent</i>	0.853	0.898	<u>0.936</u>	0.892	0.770	0.990
平均相关性	<u>0.866</u>	0.712	0.592	0.849	0.851	0.982

3.3.2 合成视频的定量分析

为了验证去雾效果，本文在已知深度图^{注2}的立体视频^[104]上合成有雾视频。图3-7展示本文方法与5种代表性方法定量分析结果，其中：MDCP^[63]、IVR^[57]和OCE^[60]是最新的视频去雾方法；DCP^[74]和BCCR^[62]是经典的单帧图像去雾方法。有雾视频基于式(2-1)生成，本文假设大气光为纯白常量 $\alpha = 1$ 。

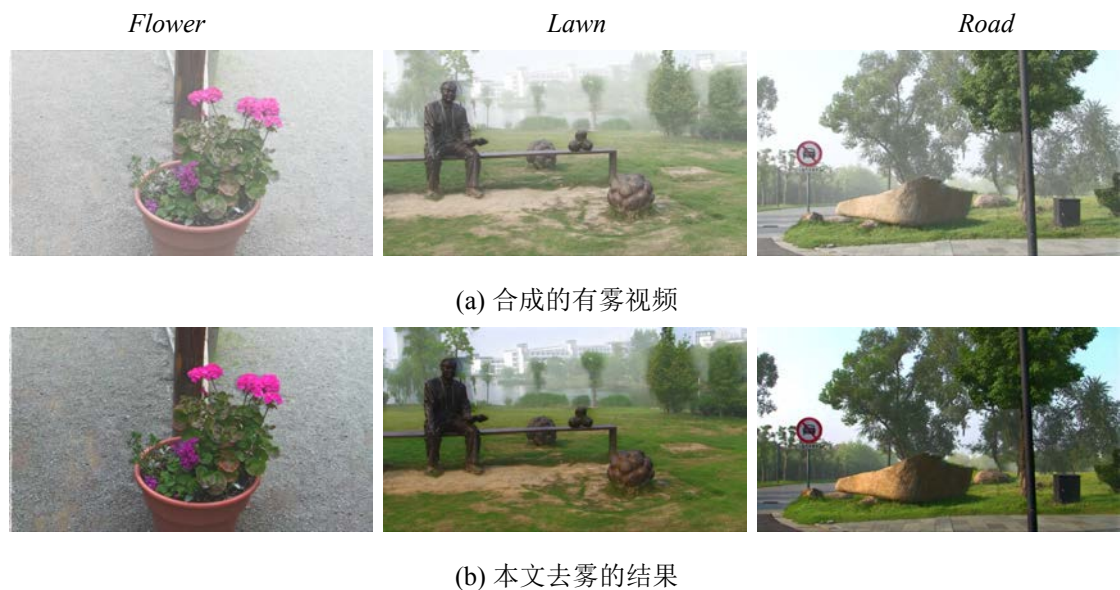


图 3-7 基于立体视频合成的有雾视频与本文去雾结果

在本文中，通过均方误差（Mean Square Error, MSE）定量评估清晰视频和去雾结果的偏差。低的 MSE 表示去雾结果令人满意，反之高的 MSE 表示去雾效果较差。在表3-2中，在三组合成视频的对比结果表明，本文方法与5种最新的方法相比取得最低的 MSE 和最好的去雾结果。

^{注2}<http://www.cad.zju.edu.cn/home/gfzhang/projects/videodepth/data/>

表 3-2 在合成视频中不同方法的定量分析

视频序列	DCP ^[74]	BCCR ^[62]	MDCP ^[63]	IVR ^[57]	OCE ^[60]	Ours
<i>Flower</i>	0.0228	0.0240	0.0257	0.0479	<u>0.0174</u>	0.0034
<i>Lawn</i>	0.0198	0.0176	0.4902	0.0141	0.0408	<u>0.0166</u>
<i>Road</i>	0.0141	0.0191	<u>0.0108</u>	0.0364	0.0274	0.0092
平均 MSE	<u>0.0189</u>	0.0202	0.1756	0.0328	0.0285	0.0097

3.3.3 真实视频的定性分析

此外，本文还在真实视频上进行定性的分析。定性分析了四种代表性方法的去雾效果，其中包括 DCP^[74]、BCCR^[62]、IVR^[57]、OCE^[60]。图3-8中展示了 6 组最近相关工作中采用的视频序列和各种方法的去雾结果，更多对比视频与实验代码详见<http://caibolun.github.io/st-mrf/>。如图可见，基于对比度最大化的去雾方法，如：BCCR^[62]，IVR^[57]，OCE^[60] 等，能够取得鲜明的颜色和锐利的结果，但它们往往会产生过度饱和和空间不一致，例如 *Bali* 中的山脉和 *Playground* 中的天空光晕。在图3-8e和图3-8d中，直接应用静态图像去雾方法（DCP^[74] 和 BCCR^[62]）会带来严重的闪烁和伪影，例如 *Cross* 中的道路区域和 *Hazeroad* 中的天空区域。虽然，视频去雾方法 OCE^[60] 使用重叠时间滤波器来引入时序信息，但由于缺少空间相关性导致结果中存在少量区块效应。与以上方法相比，本文提出的方法有效地避免了过饱和并保持时空一致性。

3.3.4 算法实时性分析

我们评估了各类视频去雾方法在不同大小的视频格式下的实时性。实验在一台 Intel i7 3770 CPU（3.4GHz）的个人计算机上运行，表3-3报告了本文方法与 DCP^[74]、BCCR^[62]、MDCP^[63]、IVR^[57] 和 OCE^[60] 的平均速度（帧每秒，fps）的对比。在表3-3中，本文方法分为两个版本：仅采用积分图技术加速的时空 MRF 表示为 ST-MRF，采用积分图与下采样技术共同加速的版本表示为 ST-MRF+。从实验结果可见，本文方法明显优于其他方法，即使给定的有雾视频很大，也能实现高效处理。特别地，ST-MRF 在通用影像传输格式（CIF，352）下可实现了大约 120fps 的处理速度，ST-MRF+ 可超过 230fps 的处理速度。因此，本文框架为其他视觉感知处理预留了大量的扩展空间，并且可方便地移植到嵌入式系统中。

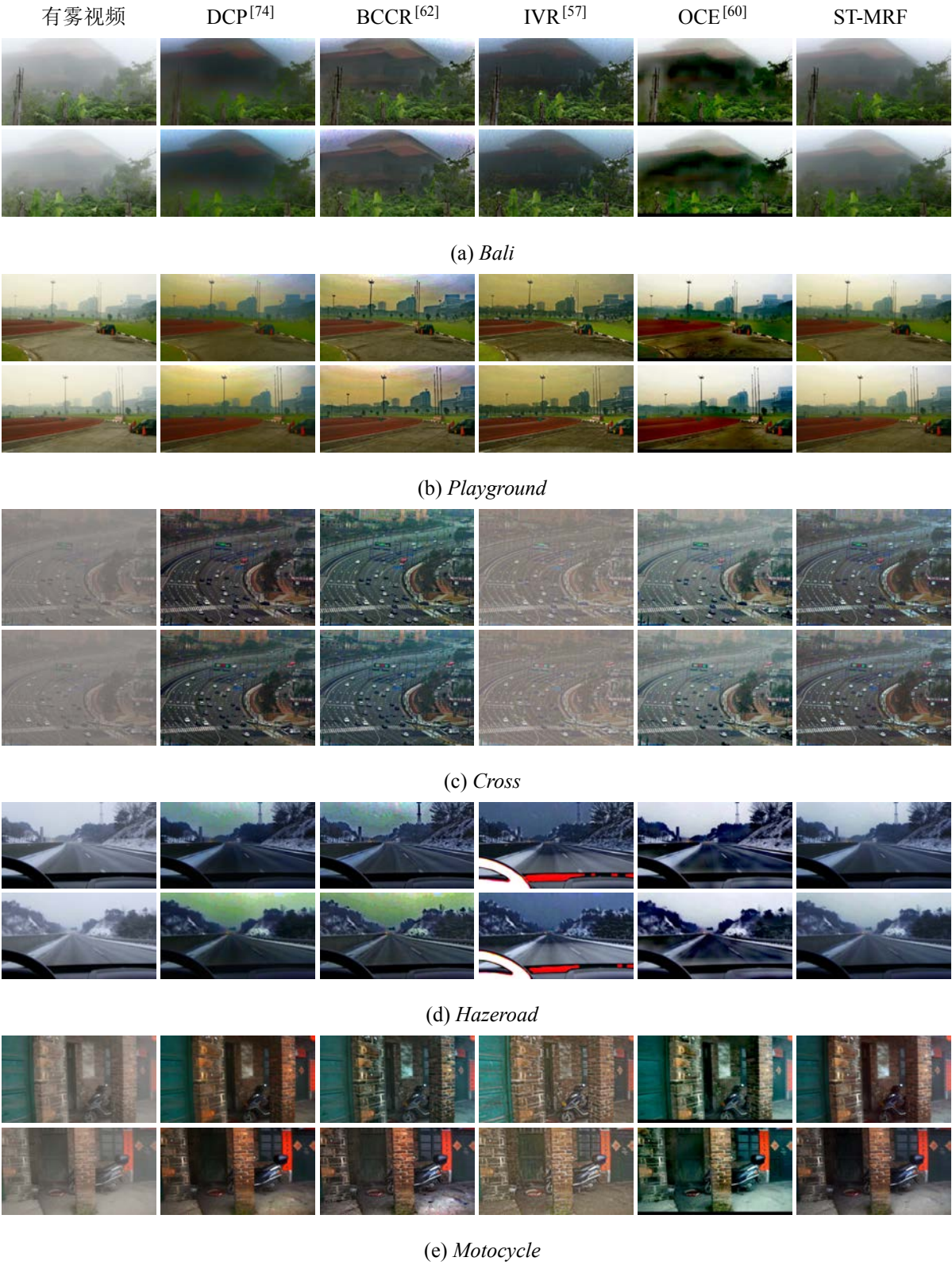


图 3-8 在合成视频中不同方法的定性分析

表 3-3 不同视频去雾方法的实时性分析 (fps)

视频格式	DCP ^[74]	BCCR ^[62]	MDCP ^[63]	IVR ^[57]	OCE ^[60]	ST-MRF	ST-MRF+
CIF (352 × 288)	1.485	1.322	7.343	1.205	97.076	116.371	234.732
VGA (640 × 480)	0.566	0.467	2.430	0.171	30.539	36.609	73.118
D1 (704 × 576)	0.414	0.358	1.830	0.102	22.930	27.493	55.003
XGA (1024 × 768)	0.216	0.197	0.842	0.028	12.106	14.515	29.015

3.4 小结

本章提出了一个基于时空 MRF 的实时视频去雾框架。本方法引入了空间一致性和时间连续性，以保证时空相关性并生成一个无区块效应和闪烁效应的去雾增强视频。此外，此框架可采用积分图像技术和下采样技术进行加速，显著地降低了计算复杂度并保证算法的实时性。实验结果表明，该算法能够在较低的计算复杂度下有效地恢复有雾视频。

需要指出的是，尽管本章提出了一个有效的联合时空信息的优化方法，但受限于暗通道先验并不能高精度地估计雾浓度。后续的研究工作将着重解决初始的雾浓度问题，通过多先验融合的线性学习模型在保证效率的同时对雾浓度进行更高精度的估计。此外，本文提出的时空优化框架还可进一步扩展到更为广阔的实时视频处理中，如低光增强、自动光照、高光动态等。这将是本工作后续的研究方向。

第四章 基于联合先验 Retinex 模型的颜色校正

4.1 概述

雾的形成是由于悬浮在空气中的微小颗粒，包括水滴、灰尘、细尘等对于大气光的散射引起。通常情况下，大部分悬浮颗粒（如：水滴和灰尘）普遍呈现灰白色，并不影响有雾图像的颜色平衡。然而，现实环境中还包括其他颜色悬浮颗粒形成的雾，如黄色的沙尘暴、蓝色的水雾、暗红色的烟雾等。因此，对于特殊场景下的图像去雾需要对其结果进行颜色校正，如图4-1所示。



图 4-1 图像去雾与颜色校正

Retinex 理论是模拟人类视觉系统的颜色感知模型，可用于提取并消除图像中的光照效果，从而实现颜色校正。Retinex 的主要目标是将观察到的图像分解为光照成分和反射成分。在 Retinex 分解中，光照代表光线在物体上的光强度，反射率表示物体的物理特性。除了颜色校正^[105]，Retinex 理论还可以应用于背光消除^[106]和低光增强^[107]中。

从单幅观测图像中估计和分解出光强度和反射率是一个不适定问题。基于 Retinex 理论^[13,23,24]，学界陆续提出各种算法以解决此问题，其主要可分为三类算法：基于路径（Path-based）的算法、基于偏微分方程式（PDE-based）的算法和基于中心/环绕（Center/Surround）的算法。基于路径的算法^[16,17]通过随机游走找到对应的反射率，但随机路径求解将带来高的计算复杂度。在路径递归算法^[14,15]中，递归矩阵求解取

代随机路径搜索过程提高搜索效率，但递归迭代次数缺少明确的定义且非常影响最终的收敛结果。在基于偏微分方程式的算法^[108,109]中，通过求解偏微分方程（Partial Differential Equation, PDE）来估计反射率，因此可通过快速傅立叶变换（Fast Fourier Transform, FFT）进行高效计算。然而，非稀疏散向量场的自由求解会导致反射率估计不满足分段连续的先验。基于中心/环绕的算法，主要包括：单尺度 Retinex^[18]（Single Scale Retinex, SSR）、多尺度 Retinex^[19]（Multi-scale Retinex, MSR）和彩色多尺度 Retinex^[22]（Multi-scale Retinex with Color Restoration, MSRCR）等，在约束光强度平滑的同时约束反射率锐利。然而，此类方法采用缺少结构保持的高斯滤波器作为光强度的平滑约束，导致其光强度在结构和边缘附近容易产生光晕和伪影。

最近几年，变分（Variation）方法被逐步引进以改善光强度和反射率估计。第一个 Retinex 的变分框架（Variational Framework, VF）由 Kimmel 等人^[110]提出，但仅基于光强度平滑假设的目标函数缺乏考虑反射率的物理先验。针对此问题，全变分模型^[111]（Total Variation Model, TVM）通过 Bregman 迭代实现对反射率的优化，但由于采用对数域来简化求解，从而导致反射率估计过度平滑和细节丢失。因此，Fu 等人^[106]提出了一种线性域的联合概率求解方法（Simultaneous Reflectance and Illumination Estimation, SRIE）同时实现反射率和光照度的估计，相比于对数域处理可更好地保留细节。同样为解决对数域问题，加权变分模型^[105]（Total Variation Model, TVM）通过加强对数域中高亮区域的正则化系数，实现对细节信息的保护。然而，SRIE 和 WVM 都仅基于光强度的平滑假设，但不同物体表面的不一致性可能导致光强度在边缘附近的估计误差，此外场景中的光源信息也会因此被破坏。

根据成像理论^[112,113]，现实中的成像过程受外界环境和内在本征的共同影响，包括形状、纹理和光照。在本文中，提出了一种联合内外先验（Joint intrinsic-extrinsic Prior, JieP）的 Retinex 模型在线性域中实现对光强度和反射率的分解。这是第一次将形状先验引入 Retinex 模型中。对此，本文提出一种新的结构保持统计量——局部变化方差（Local Variation Deviation, LVD）作为形状先验；纹理先验用于保持反射率的丰富细节和分段连续；基于光照先验的亮通道有效地捕获光源信息。最后，采用基于迭代重加权最小二乘（Iteratively Re-weighted Least Square, IRLS）的坐标下降法同时优化光强度和反射率。此外，预条件共轭梯度^[114]（Preconditioned Conjugate Gradient, PCG）的快速求解器使得每次迭代的时间复杂度从 $O(N^3)$ 降低到 $O(N)$ 。

4.2 局部变化方差

受物体形状影响，光强度呈现分段平滑。本节先简述现有的边缘/结构保持平滑方法，并以此提出一种新的结构选择性度量。

4.2.1 边缘/结构保持平滑

根据不同的应用场景，边缘/结构保持平滑（如图4-2）可通过不同的方法实现，其主要可分为：基于滤波器的方法、基于统计的方法和基于优化的方法。

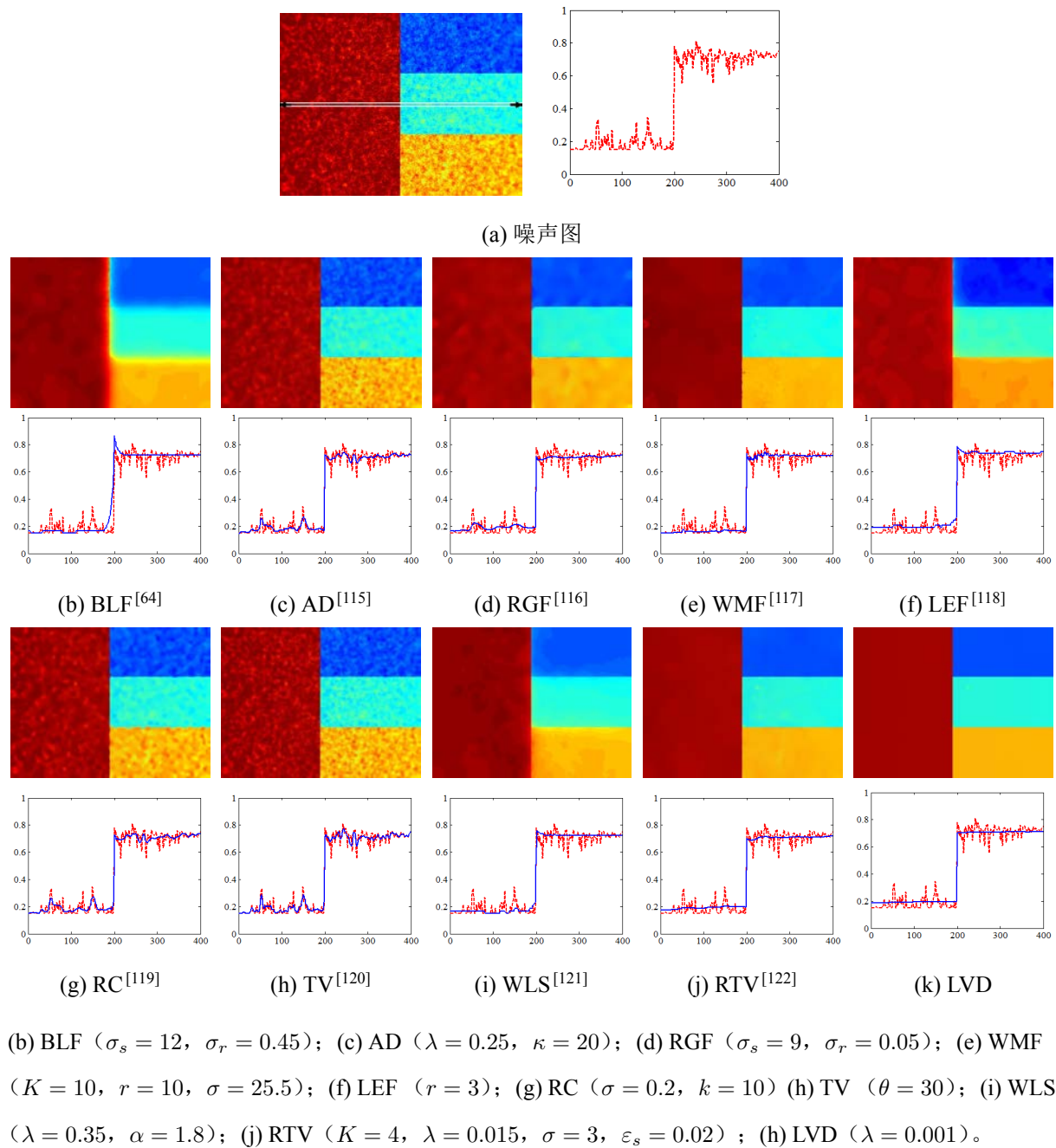


图 4-2 不同边缘/结构保持的平滑方法

• 基于滤波器的方法

基于滤波器的方法采用不同的滤波策略，主要包括双边滤波器^[64]（Bilateral Filter, BLF）、各项异性扩散^[115]（Anisotropic Diffusion, AD）和循环导向滤波器^[116]（Rolling Guidance Filter, RGF）。双边滤波器^[64]在不考虑交叉边缘的情况下，联合两个高斯核实现相邻像素的加权平均。各项异性扩散^[115]通过迭代求解基于局部变换的扩散方程，实现边缘保持的相邻像素平滑。循环导向滤波器^[116]级联多个导向滤波器实现递归引导，逐步地保留边缘信息。

这些基于滤波器的方法通过不同的滤波器组合，在相邻像素间平衡细节平滑和边缘保持的作用。然而，由于局域滤波器的频率泄露，局部滤波的方法会引起的吉布斯（Gibbs）现象，从而导致边缘附近产生显著的振铃效应（图4-2b中明显观察到）。

• 基于统计的方法

基于统计的方法，如：加权中值滤波^[117]（Weight Median Filter, WMF）、局部极值滤波器^[118]（Local Extrema Filter, LEF）和区域协方差^[119]（Region Covariance, RC），在局部块中计算其分布模式而不是关注其数值响应，因此可有效地剔除椒盐噪声。加权中值滤波^[117]通过加权统计和中值滤波器从高对比度的结构中平滑和剔除纹理。Subr 等人^[118]通过局部极值区分小尺度的震荡与显著的边缘。区域协方差^[119]在局部块中度量变化的统计信息以提取结构边缘。

但对于高频变化的输入信号，局部的统计方法仍然会带来震荡效应，且不能有效去除细节。不同于局部加权的滤波，基于局部统计的方法不可避免地会带来更高的计算成本。

• 基于优化的方法

全变分^[120]（Total Variation）、加权最小二乘法^[121]（Weighted Least Squares, WLS）和相对全变分^[122]（Relative Total Variation）是三种经典的基于优化的方法。这些方法通过优化全局目标函数（分别基于全变分范数、加权二阶范数和相对全变分范数）来实现对图像的平滑。图像梯度的 L2 范数最小化^[120]约束图像的边缘渐变以增强边缘感知的平滑质量。Farbman 等人^[121]提出加权最小二乘法，通过求解大规模线性系统实现多尺度的图像分解。Xu 等人^[122]最小化相对全变分的度量项以精

确地分离结构和纹理。

然而，全局优化方法只能处理相对较小的方差，并容易受局部纹理的影响。此外，相比于基于滤波器的方法，基于优化的方法更难实现和加速。

4.2.2 局部变化方差

为解决上述问题，本文提出一种新的结构选择性度量——局部变化方差（Local Variation Deviation, LVD）是一种**局部统计量**，并可应用于**全局优化**函数中实现结构保持的图像平滑方法。

在统计学中，标准差（Standard Deviation）反映一个数据集的离散程度，可用于度量数据的一致性。局部变化方差提供了一种简洁、自然、高效的方法来识别不同变化的统计特性。其中，局部变化代表图像的梯度特征，其方差代表局部邻域内变化的相关性。因此，局部变化方差在区分纹理（弱相关）和结构（强相关）方面提供了强大的识别能力。基于这些特性，可利用其来区分局部的图像形状。

在数学形式上， $\mathcal{D}_{x/y}$ 表示从图像 I 中提取的水平或垂直方向上的局部变化偏差：

$$\mathcal{D}_{x/y} = \left| \nabla_{x/y} I - \frac{1}{|\Omega|} \sum_{\Omega} \nabla_{x/y} I \right|. \quad (4-1)$$

其中， $\nabla_{x/y}$ 是梯度操作， Ω 是尺度 $r \times r$ 的局部邻域（本文中设置 $r = 3$ ）。为进一步增强纹理和结构的判别能力，本文采用相对方差 $\mathcal{R}$ 取代通用方差 $\mathcal{D}$ 以放大其差异程度，并重写如下：

$$\mathcal{R}_{x/y} = \left| \frac{\nabla_{x/y} I}{\frac{1}{|\Omega|} \sum_{\Omega} \nabla_{x/y} I} \right|. \quad (4-2)$$

如图4-3所示， $\mathcal{D}$ 和 $\mathcal{R}$ 是局部变化方差的均值，可见纹理区域（蓝色框）和结构区域（绿色框）的局部变化方差存在显著的差异。为选择性地平滑纹理，平滑结果 I 可由局部变化方差的目标函数表示：

$$\arg \min_I \|I - S\|_2^2 + \lambda \|\mathcal{R}_x\|_1 + \lambda \|\mathcal{R}_y\|_1. \quad (4-3)$$

其中，第一项 $\|I - S\|_2^2$ 是保真项，最小化平滑图像 I 和输入图像 S 的欧式距离。如图4-2k所示，本文提出的局部变化方差可有效地去除纹理，并很好地保护结构和边缘。

4.2.3 理论分析

局部变化方差在结构保持平滑的作用可直观地解释如下。为简明阐述，令局部变化方差的均值为 $\overline{\nabla I} = \frac{1}{|\Omega|} \sum_{\Omega} \nabla I$ 。图像的局部变化方差主要可分为三种情况：

- 情况一：平面。当局部块 I 几乎是不变的常数，那么可得 $\nabla I \approx 0$ 和 $\overline{\nabla I} \approx 0$ ，因此

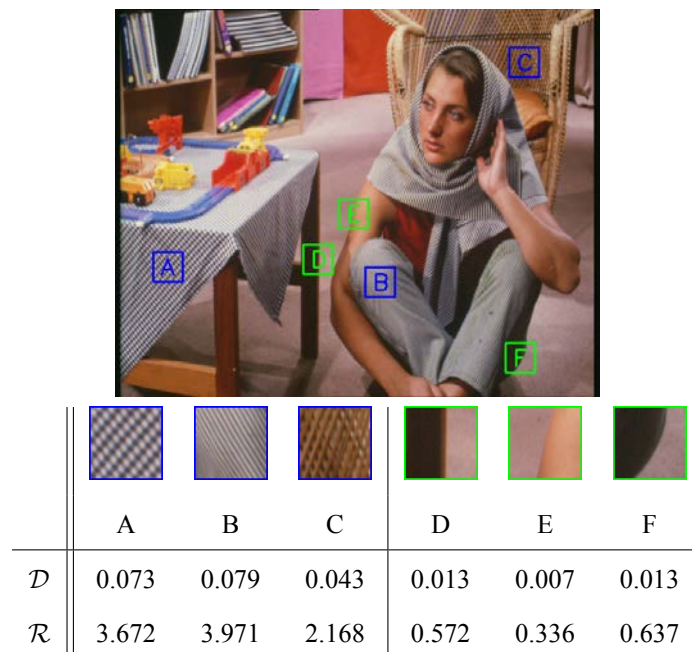


图 4-3 局部变化方差在真实图像上的分析

$\mathcal{D} \approx 0$ 和 $\mathcal{R} \approx 0$ 。

- 情况二：纹理。当局部块 I 的数值高频变化，那么 ∇I 的波动频率远大于 $\overline{\nabla I}$ ，因此 $\mathcal{D} > 0$ 和 $\mathcal{R} \gg 1$ 。
- 情况三：结构。当局部块 I 的变化是一致的，那么 ∇I 的方差很小，因此 $\mathcal{D} \approx 0$ 和 $\mathcal{R} \approx 1$ 。

为定量分析局部变化方差的有效性，图4-4展示一个包含纹理和结构的一维信号。如图所示，绝对变化 $|\nabla I|$ 包含纹理和结构的全梯度；平均变化 $|\overline{\nabla I}|$ 仅包含显著的结构；因此，相对方差 $\mathcal{R}$ 可从显著的结构中有效地捕获细节的纹理。

图4-5对噪声图（见4-2a）的局部变化方差 $\mathcal{R}$ 进行分析。如图4-5a和图4-5b所示，局部均值滤波器有效地抑制了纹理的响应，因此纹理区域的 $|\overline{\nabla I}|$ 远小于 $|\nabla I|$ 。另一个直观解释为，结构区域的梯度方向相比于复杂的纹理区域更为一致。

4.3 联合内外先验模型

Retinex 的物理模型可以描述为： $S = R \cdot L$ ，也就是说将观测图像 S 可分解为两个不同的成分：反射率 R 和光强度 I 。根据图形学理论^[112,113]，物体的表征受形状、光照和纹理的影响。在本节中将提出一个通用的联合内外先验的 Retinex 模型，通过最小化“形状、光照和纹理”的能量来约束观测图像的分解。

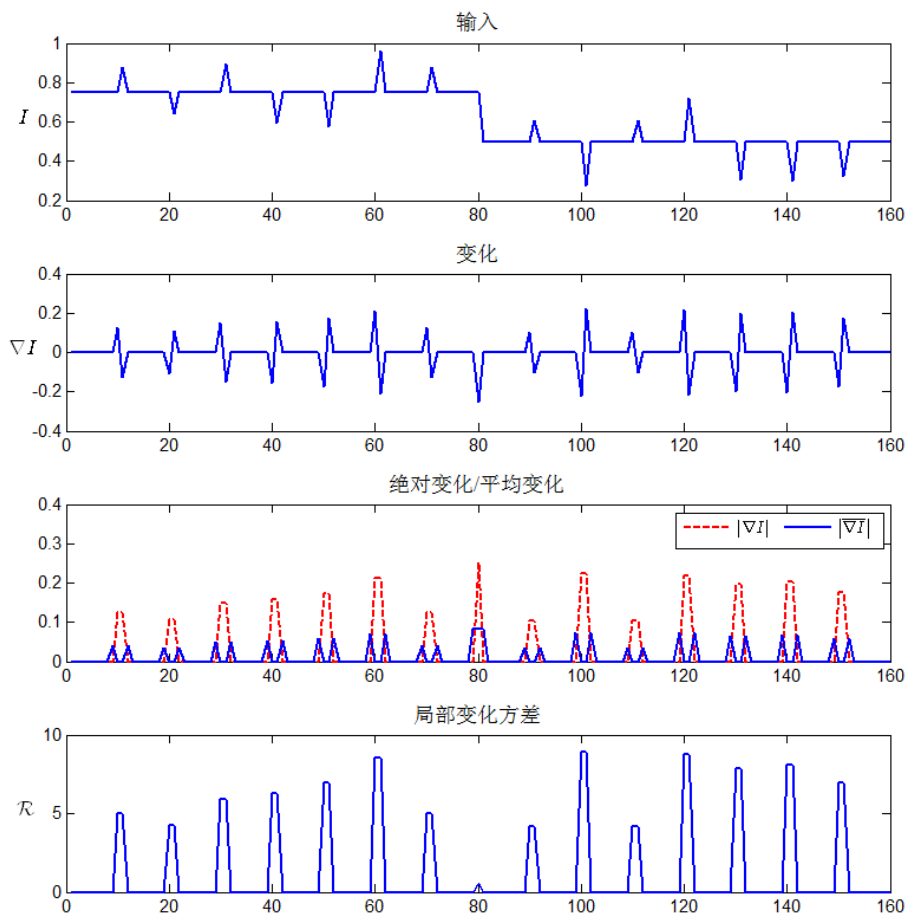


图 4-4 局部变化方差在一维信号上的分析

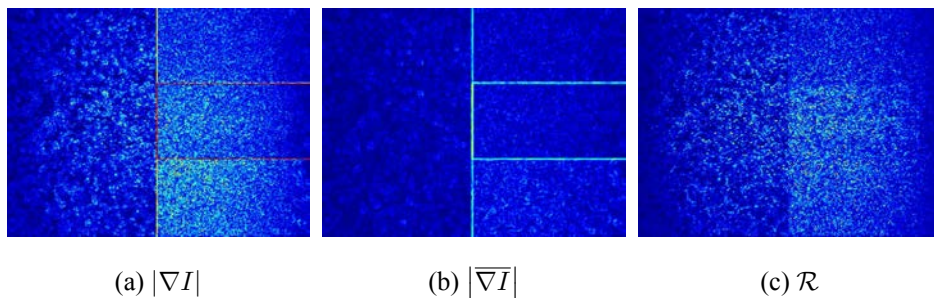


图 4-5 局部变化方差在噪声图上的分析

4.3.1 形状先验

因为光强度是未知的，本文提出形状先验来寻找影响光照条件的物体结构。形状先验是由物体不同方向的表面接受不同的场景光照而决定的——不同的表面以相同的概率面对不同的光照。如图4-6所示，不同的颜色蒙版表示不同照明条件的表面，可见形状先验在大多数环境中都是有效的，例如：地板、墙壁、地平面和天空等。

基于局部变化方差对结构的属性，形状先验可通过其简单并有效地提取显著结构。

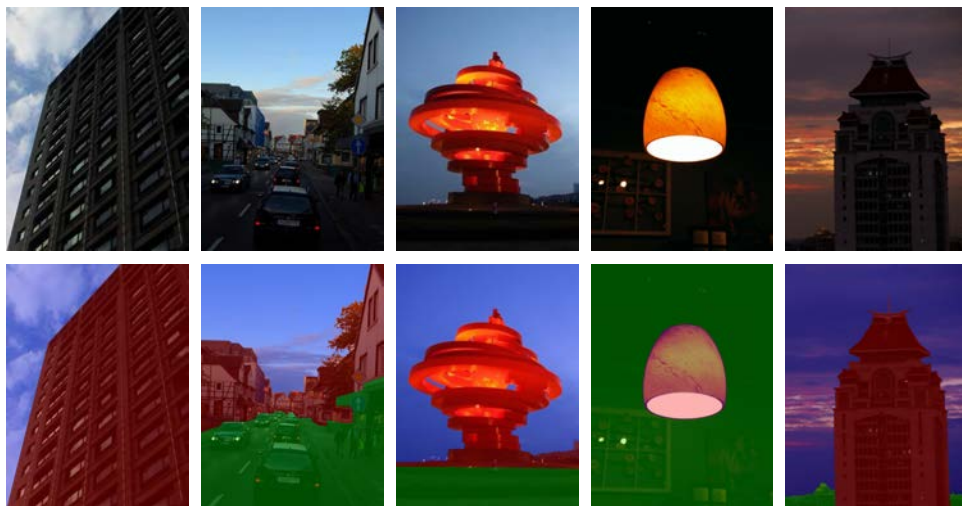


图 4-6 真实环境中光照的形状先验

本文结合 $\mathcal{R}_x$ 和 $\mathcal{R}_y$ 构成光强度 I 的形状先验，其能量函数可表示为：

$$E_s(I) = \left\| \frac{\nabla_x I}{\frac{1}{|\Omega|} \sum_{\Omega} \nabla_x I} \right\|_1 + \left\| \frac{\nabla_y I}{\frac{1}{|\Omega|} \sum_{\Omega} \nabla_y I} \right\|_1. \quad (4-4)$$

4.3.2 光照先验

若光强度 I 仅依赖分段平滑的形状先验，其并无法准确估计光源和白色物体的光照。如图4-7所示，光照先验基于观测图像的反相图 $1 - S$ 的视觉效果类似于有雾图像^[123]，其可表示为：

$$(1 - S) = 1 - I \cdot R = (1 - R) \cdot I + (1 - I) \quad (4-5)$$

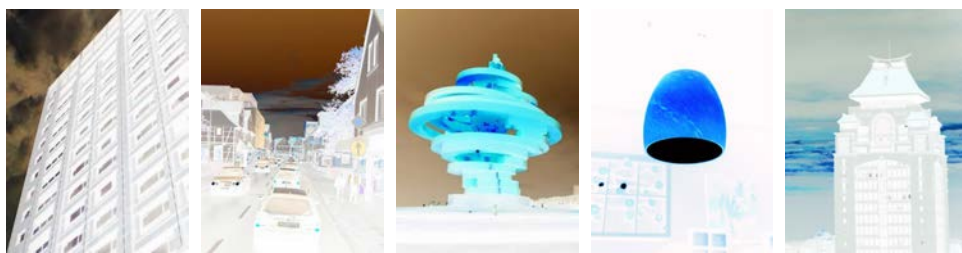


图 4-7 图4-6对应的反相图的光照先验

令 $H = 1 - S$, $J = 1 - R$, $A = 1$ 和 $T = I$, 可将式(4-5)转化为 $H = J \cdot T + A(1 - T)$, 其与大气散射模型具有相似的物理意义。其中, H 是观测到的有雾图像, J 是待恢复的清晰场景, T 是大气透射率, A 是全局大气光。在大多数清晰图像块中, RGB 颜色空间中至少存在一个颜色通道的强度值很低甚至接近于零。如节3.2.1.2所述, 暗通道先验

可用于大气透射率估计：

$$T = 1 - \min_{\Omega} \left(\min_{c \in \{r, g, b\}} \frac{H^c}{a} \right) \quad (4-6)$$

推理可得，可通过 RGB 三个颜色通道的最大值估计光强度 I ：

$$I = 1 - \min_{\Omega} \left(\min_{c \in \{r, g, b\}} (1 - S)^c \right) = \max_{\Omega} \left(\max_{c \in \{r, g, b\}} S^c \right). \quad (4-7)$$

综上所述，本文定义亮通道为 $B = \max_{\Omega} (\max_c S^c)$ ，并最小化光强度 I 和亮通道 B 之间的 L2 范数距离：

$$E_l(I) = \|I - B\|_2^2. \quad (4-8)$$

4.3.3 纹理先验

由于 Retinex 分解会导致原隐藏于阴影区域的噪声相应地放大，因此反射率估计面临的重要问题是如何有效地抑制阴影噪声。根据反射率的丰富细节和分段连续假设^[124]，反射率的梯度满足拉普拉斯分布——对应于全变分稀疏。因此，纹理先验强调反射率 R 的分段连续，可表示为：

$$E_t(R) = \|\nabla_x R\|_1 + \|\nabla_y R\|_1. \quad (4-9)$$

4.3.4 联合优化

为有效地估计光强度 I 和反射率 R ，融合上述内外先验（形状先验、光照先验、纹理先验），构建联合优化的目标函数：

$$E(I, R) = \|R \cdot I - S\|_2^2 + \alpha E_s(I) + \beta E_l(I) + \lambda E_t(R). \quad (4-10)$$

其中， α 、 β 和 γ 三个正则参数。第一项 $\|R \cdot I - S\|_2^2$ 是 L2 数据保真项，最小化 Retinex 分解 $R \cdot I$ 和观测图像 S 之间的 L2 范数距离。

由于 Retinex 模型中存在两个未知量，传统梯度下降方法并不能直接求解。因此，本文采用坐标下降法^[125]（Block Coordinate Descent, BCD）寻找非凸目标函数的最优解。另外，L1 范数的形状先验 E_s 和纹理先验 E_t 难以求解，对此引入迭代重加权的最小二乘法^[126]（Iteratively Re-weighted Least Square, IRLS）将式(4-4)和式(4-9)重写如下：

$$\begin{cases} E_s(I) = u_x \|\nabla_x I\|_2^2 + u_y \|\nabla_y I\|_2^2 \\ E_t(R) = v_x \|\nabla_x R\|_2^2 + v_y \|\nabla_y R\|_2^2 \end{cases}, \quad (4-11)$$

$$\text{其中} \begin{cases} u_{x,y} = \left(\frac{1}{\Omega} \sum_{\Omega} \nabla_{x,y} I \right) |\nabla_{x,y} I| + \epsilon)^{-1} \\ v_{x,y} = (|\nabla_{x,y} R| + \epsilon)^{-1} \end{cases}. \quad (4-12)$$

这里 ϵ 是一个极小正数以避免除数为零。因此，联合优化问题可分解为两个独立的

子问题迭代地循环优化。对于第 k 次迭代,

$$\begin{aligned}
 (\text{P1}) \quad I_k &= \arg \min_I \|R_{k-1} \cdot I - S\|_2^2 \\
 &\quad + \alpha (u_x \|\nabla_x I\|_2^2 + u_y \|\nabla_y I\|_2^2) + \beta \|I - B\|_2^2 \\
 (\text{P2}) \quad R_k &= \arg \min_R \|R \cdot I_k - S\|_2^2 \\
 &\quad + \lambda (v_x \|\nabla_x R\|_2^2 + v_y \|\nabla_y R\|_2^2)
 \end{aligned} \tag{4-13}$$

可看出, 此联合优化问题只包含二次项, 且两个子问题都具有闭合形式的全局最优解。具体求解流程如下:

1) P1 求解:

初始化 $I_0 = S$, 并采用矩阵形式重写损失函数如下:

$$(\mathbf{R}_{k-1} \cdot \mathbf{I} - \mathbf{S})^T (\mathbf{R}_{k-1} \cdot \mathbf{I} - \mathbf{S}) + \alpha (\mathbf{I}^T \mathbf{D}_x^T \mathbf{U}_x \mathbf{D}_x \mathbf{I} + \mathbf{I}^T \mathbf{D}_y^T \mathbf{U}_y \mathbf{D}_y \mathbf{I}) + \beta (\mathbf{I} - \mathbf{B})^T (\mathbf{I} - \mathbf{B}). \tag{4-14}$$

其中, $\mathbf{U}_x$ 和 $\mathbf{U}_y$ 分别是权值 u_x 和 u_y 的对角矩阵, $\mathbf{D}_x$ 和 $\mathbf{D}_y$ 分别是水平和垂直方向的离散微分矩阵。最小化子问题 (P1) 可得此线性系统中向量 $\mathbf{I}$ 的最优解:

$$\mathbf{I}_k = (\mathbf{R}_{k-1}^T \mathbf{R}_{k-1} + \alpha \mathbf{M}_k + \beta \mathbf{1})^{-1} (\mathbf{R}_{k-1}^T \mathbf{S} + \beta \mathbf{B}). \tag{4-15}$$

其中, $\mathbf{1}$ 是单位矩阵, $\mathbf{M} = \mathbf{D}_x^T \mathbf{U}_x \mathbf{D}_x + \mathbf{D}_y^T \mathbf{U}_y \mathbf{D}_y$ 是五点正定拉普拉斯矩阵^[127]。

2) P2 求解:

子问题 (P2) 同样是一个最小二乘问题, 如求解子问题 (P1) 先初始化 $\mathbf{R}_{k-1} = \mathbf{S}/\mathbf{I}_k$ 并如求解 $\mathbf{I}_k$ 一样迭代式更新 $\mathbf{R}_k$:

$$\mathbf{R}_k = (\mathbf{I}_k^T \mathbf{I}_k + \lambda \mathbf{N}_k)^{-1} (\mathbf{I}_k^T \mathbf{S}). \tag{4-16}$$

其中, $\mathbf{N} = \mathbf{D}_x^T \mathbf{V}_x \mathbf{D}_x + \mathbf{D}_y^T \mathbf{V}_y \mathbf{D}_y$ 。

迭代地循环更新 $\mathbf{I}$ 和 $\mathbf{R}$ 直到 $\|\mathbf{I}_k - \mathbf{I}_{k-1}\|/\|\mathbf{I}_{k-1}\| \leq \varepsilon$ 或 $\|\mathbf{R}_k - \mathbf{R}_{k-1}\|/\|\mathbf{R}_{k-1}\| \leq \varepsilon$ 。此外, 预条件共轭梯度^[114] (Preconditioned Conjugate Gradient, PCG) 的快速求解器使得每次迭代的时间复杂度降到 $O(N)$ 。完整的优化过程总结见算法4-1。

4.4 实验与分析

本节将从定性和定量两方面分析和验证基于联合先验的 Retinex 模型。在本节所有实验中, 正则参数 α 、 β 和 λ 分别设置为 0.001、0.25 和 0.0001, 收敛参数 ε 设置为 10^{-3} 。

算法 4-1: 基于联合先验的 Retinex 模型

Input: 观测图像 $\mathbf{S}$; 正则参数 α, β, λ ; 最大迭代次数 K ; 收敛参数 ε 。

Output: 光强度 $\mathbf{I}$; 反射率 $\mathbf{R}$ 。

```

1 初始化  $\mathbf{I}_0 \leftarrow \mathbf{S}$ ;
2 for  $k = 1$  to  $K$  do
3   根据式(4-12)计算权值  $u_{x,y}$ ;
4   根据式(4-12)计算权值  $u_{x,y}$  通过式(4-15)更新光强度  $\mathbf{I}_k$ ;
5    $\mathbf{R}_{k-1} = \mathbf{S}/\mathbf{I}_k$ ;
6   根据式(4-12)计算权值  $v_{x,y}$ ;
7   通过式(4-16)更新反射率  $\mathbf{R}_k$ ;
8   if  $\|\mathbf{I}_k - \mathbf{I}_{k-1}\|/\|\mathbf{I}_{k-1}\| \leq \varepsilon$  or  $\|\mathbf{R}_k - \mathbf{R}_{k-1}\|/\|\mathbf{R}_{k-1}\| \leq \varepsilon$  then
9     break;
```

4.4.1 Retinex 分解

为验证联合先验模型的有效性，图4-8对比不同的 Retinex 分解，其中包括：两种经典方法，变分框架^[110]（Variational Framework, VF）和全变分模型^[111]（Total Variation Model, TVM）；两种最新 Retinex 方法，联合概率方法^[106]（Simultaneous Reflection & Illumination Estimation, SRIE）和加权变分模型^[105]（Weighted Variation Model, WVM）。为展示彩色结果，对 HSV 颜色空间（色相、饱和度、亮度）中的亮度通道应用 Retinex 分解，然后分别将分解后的光照度 I 和反射率 R 取代原亮度通道，并转换回 RGB 彩色空间。

本文提出的 Retinex 模型对光强度的估计在保持图形结构的同时，相比于其他模型更加平滑。换句话说，本文结果更符合 Retinex 的形状先验和光照先验。图4-8中雕像面颊上的阴影区域是 Retinex 分解的一个挑战，因为此处光照的变化类似于反射率的纹理。TV 和 WVM 的光强度估计过于平滑并破坏光照信息；TVM 和 SRIE 通过反射率的稀疏假设间接地保护光照结构，但其反射率估计也因此丢失了细节。此外，由于光照的移除使得反射图中阴影区域的细节和噪音被放大，如图4-8b 和图 4-8e的背景。相反，本文的 Retinex 模型通过纹理先验的分段连续性实现对噪声的有效抑制。

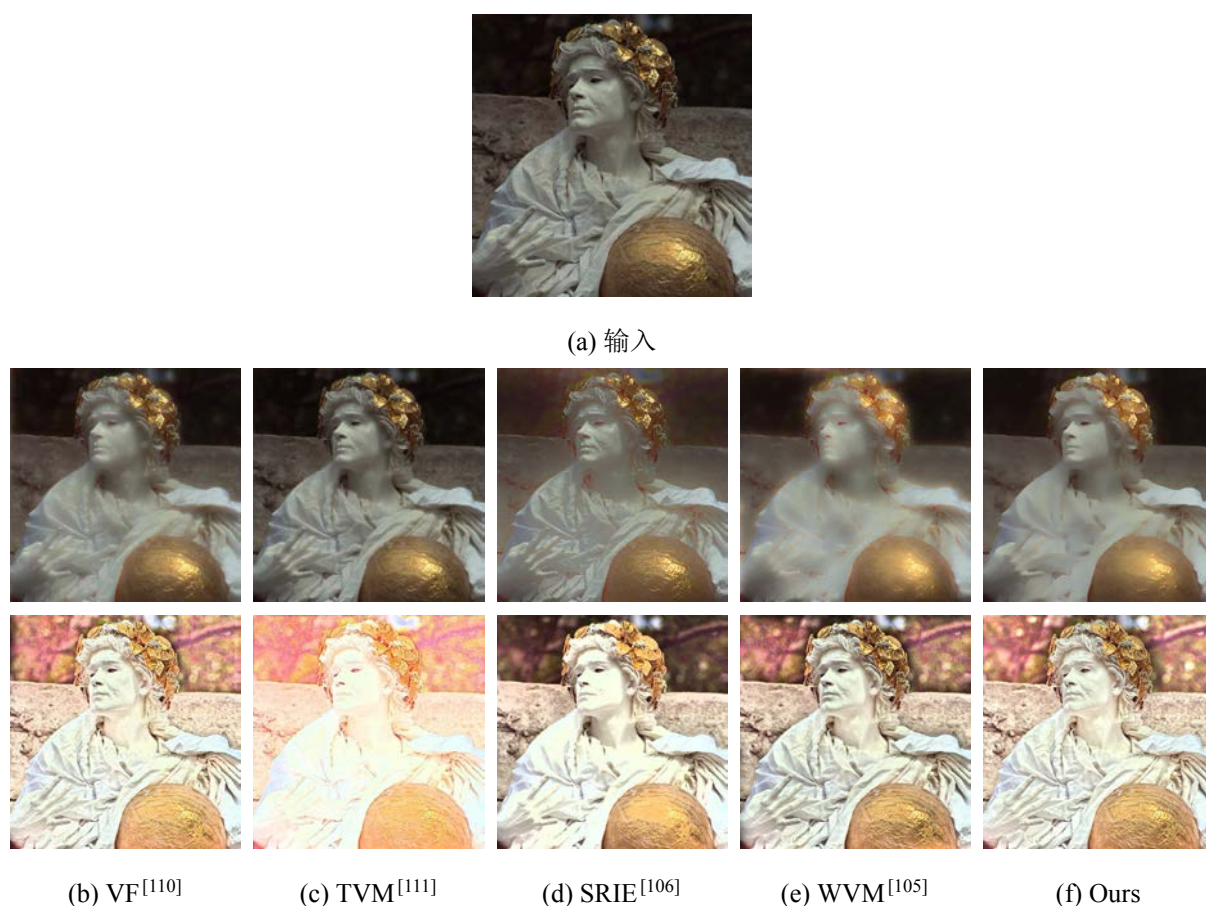


图 4-8 Retinex 分解的对比

4.4.2 颜色校正

分别对 RGB 颜色空间中各个通道进行 Retinex 分解，反射率将保留了物体的原始颜色信息，因此可将其应用于颜色校正。同时，通过颜色校正的性能评估也可验证光照度和反射率的估计精度。如图4-9所示，与其他四种 Retinex 算法相比，本文提出的 Retinex 模型展现了更为显著的颜色校正效果，例如第一幅图像中的橙色瓶子和第二幅图像中的蓝色书皮。

为定量地验证颜色校正的准确性，本文采用基于空间处理的 S-CIELAB 颜色度量^[128]评价颜色校正的误差。在图4-9中的第二行和第四行，绿色蒙版展示了颜色校正结果与真实值之间存在显著误差的空间位置。从 S-CIELAB 误差图可见，本文校正结果的误差区域（绿色区域）相较于其他方法更小，这表明本文的结果更接近于事实值。综上所述，联合先验的 Retinex 模型在处理颜色失真时能取得更准确的校正结果。

此外，本文还在 Color-Checker 数据集^[129]上与最新的颜色恒常性方法做定量的分析。Color-Checker 数据集包含由数码单反相机（佳能 5D 和佳能 1D）在自动模型下采

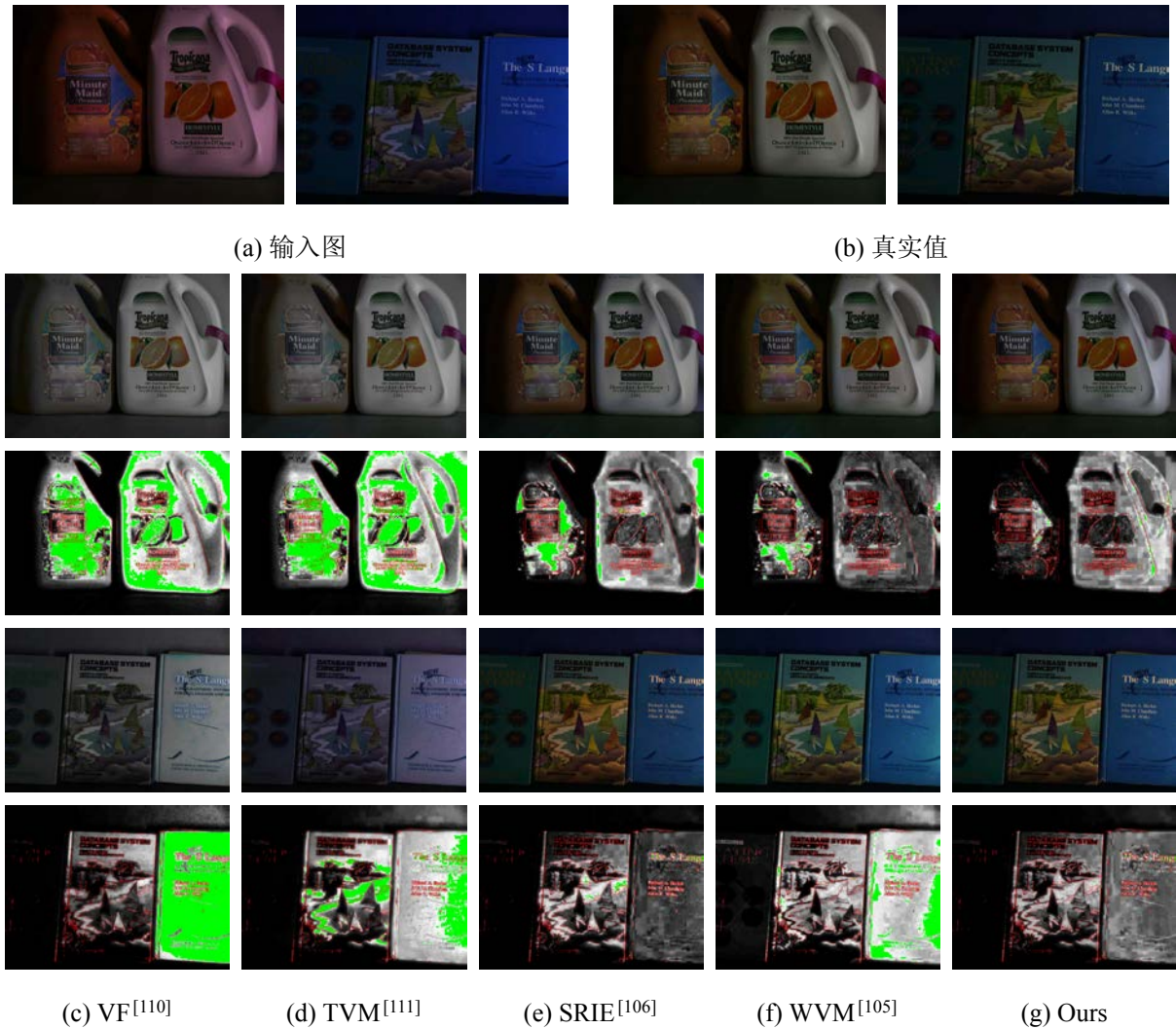


图 4-9 颜色校正的对比

集到 568 张室内和户外图像。每个图像中都包含一块标准 24 色卡作为参考，并提供手动测量的每个色卡在图中的位置坐标。为避免相机预处理（如：像素裁剪、Gamma 校正、色调校正、去马赛克、自动白平衡等）的影响，本文采用从佳能原始格式数据（.raw）中创建的无损压缩图像（.png）^{注 1}。根据每张图片中色卡的 6 个标准色块，通过 RGB 通道的中位数统计得到真实的环境光颜色。

实验中对比了近几年的颜色恒常性方法，包括：White-Patch^[130]、Grey-World^[131]、Gray-Edge^[132]、Shades-Gray^[133]、Bayesian^[129]、CNNs^[134]、Gray-World^[135] 和 Grey-Pixel^[136]。针对本文的 Retinex 方法，采用 RGB 各个颜色通道的光强度的全局平均值作

^{注 1}<http://www.cs.sfu.ca/~colour/data/>

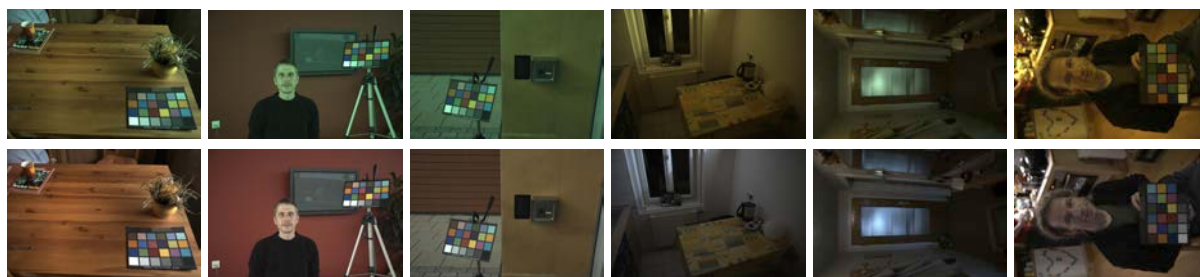


图 4-10 Color-Checker 数据库示例

为颜色向量 ρ 的估计，表示如下：

$$\rho = \left[\frac{\sum_{x \in \forall} I^r(x)}{N_{\forall}}, \frac{\sum_{x \in \forall} I^g(x)}{N_{\forall}}, \frac{\sum_{x \in \forall} I^b(x)}{N_{\forall}} \right]. \quad (4-17)$$

其中， x 是坐标索引， $N_{\forall}$ 是每个通道中像素点基数。根据 Hordley 和 Finlayson，比较环境光颜色的估计向量 ρ 和真实值 ρ_{gt} 的偏差角，表示如下：

$$\theta = \arccos \left(\frac{\rho^T \rho_{gt}}{\|\rho\| \|\rho_{gt}\|} \right). \quad (4-18)$$

如表4-1所示，本文提出的联合先验 Retinex 模型取得最优的颜色校正效果。

表 4-1 在 Color-Checker 中的颜色校正对比

方法	White-Patch ^[130]	Grey-World ^[131]	Gray-Edge ^[132]	Shades-Gray ^[133]	
平均误差角 (°)	7.55	6.36	5.13	4.93	
方法	Bayesian ^[129]	CNNs ^[134]	Gray-World ^[135]	Grey-Pixel ^[136]	Ours
平均误差角 (°)	4.82	4.73	4.66	4.60	4.32

4.4.3 亮度自适应

光强度 $\mathbf{I}$ 包含图像的亮度信息，因此可通过调整 $\mathbf{I}$ 实现低光增强和背光增强，如伽马（Gamma）变换。根据 SIRE^[106] 和 WVM^[105] 方法，可通过伽马变换调整光照度： $\mathbf{I}' = \mathbf{I}^{\frac{1}{\gamma}}$ ，其中经验参数 γ 设置为 2.2。自适应增强后的图像可表示为 $\mathbf{S}' = \mathbf{R} \cdot \mathbf{I}'$ 。为保留图像的颜色信息，伽马变换仅在 HSV 颜色空间中的亮度通道上进行处理。

在对比实验中，单尺度 Retinex^[18]（Single Scale Retinex, SSR）、彩色多尺度 Retinex^[22]（Multi-scale Retinex with Color Restore, MSRCR）、联合概率方法^[106]（Simultaneous Reflection & Illumination Estimation, SRIE）和加权变分模型^[105]（Weighted Variation Model, WVM）是基于 Retinex 的方法；自然保护增强^[137]（Naturalness Preserved Enhancement, NPE）、全局线性加窗优化^[138]（Globally Optimized Linear Window, GOLW）、多偏差融合^[107]（Multi-deviation Fusion, MF）和低光图像增

强^[139] (Low-light Image Enhancement, LIME) 是最新的图像增强方法; 直方图均衡^[140] (Histogram Equalization, HE) 和亮度保持的动态模糊直方图均衡^[141] (Brightness Preserving Dynamic Fuzzy Histogram Equalization, BPDFHE) 是两种作为基线的经典直方图均衡化。

图4-11展示五组光照自适应的例子, 其分别收集于最近的相关工作^[105,106,137,139], 并与五种最新的光照自适应方法进行对比。基于直方图均衡的方法 (HE 和 NPE) 增强后的亮度虽确有所改善, 但高光和阴影区域都出现明显的颜色失真, 如第二幅图中 HE 处理后的天空和 NPE 处理后的路基。其主要原因是, 全局的直方图均衡化并不考虑相邻像素间的相关性, 对每个像素独立地进行非线性的校正变换。虽然本文方法同样采用了非线性的伽玛校正, 但联合先验的光强度估计通过全局优化有效地结合形状和光照信息。如图4-11d, LIME 有效地通过光照自适应增强阴影细节, 但由于直接的光照剔除会导致过度增强和噪声放大, 如第三幅图像。本文方法的主观视觉效果与 SRIE 和 WVM 相似, 本文方法所产生的增强效果更为有效地强化了细节区域 (如第一幅图中的树枝)、增量了低光阴影 (第二幅图中的路基) 和保护了颜色自然性 (第四幅图中的椅子)。

大多数的光照自适应方法在一般户外图像都可获得较为有效的亮度增强, 此外增强效果缺少有参照的评价指标。在 SIRE^[106] 和 WVM^[105] 中, 自然图像质量评估^[142] (Natural Image Quality Evaluator, NIQE) 作为一种盲评估方法应用于效果评价, 其中 NIQE 值越低, 图像质量越高。如图4-11所示, 本文方法具有更低的 NIQE 值, 并与主观经验一致。此外, 本文从最近工作^[105-107,137-139] 中收集 35 幅公认有挑战性的图片以进一步验证算法性能。因 NIQE 仅对灰度图像进行评估, 本文增加自回归图像清晰度量^[143] (Autoregressive-based Image Sharpness Metric, ARISM) 作为彩色图像评估的补充。在表 4-2 中, 本文所提模型在 NIQE 和 ARISM 均取得更低的平均值, 这表明联合先验的 Retinex 模型在不同类型的图像上具有更为鲁棒的光照自适应性能。

4.4.4 先验分析

本小节详细分析本文提出的两种新的图形学先验: 形状先验和光照先验。为验证形状先验的影响, 这里将形状先验的局部变化方差替换为梯度的 L2 平滑正则。如图4-12c所示, 采用 L2 平滑正则的光强度估计会出现边缘模糊和结构缺失。为验证光照先验的影响, 在保持其他参数不变的情况下, 设正则化参数 $\beta = 0$ 。如图 4-12d 所示, 亮通道先验的缺少会导致光强度估计过于平滑, 且无法有效地捕获光源, 如图中月亮

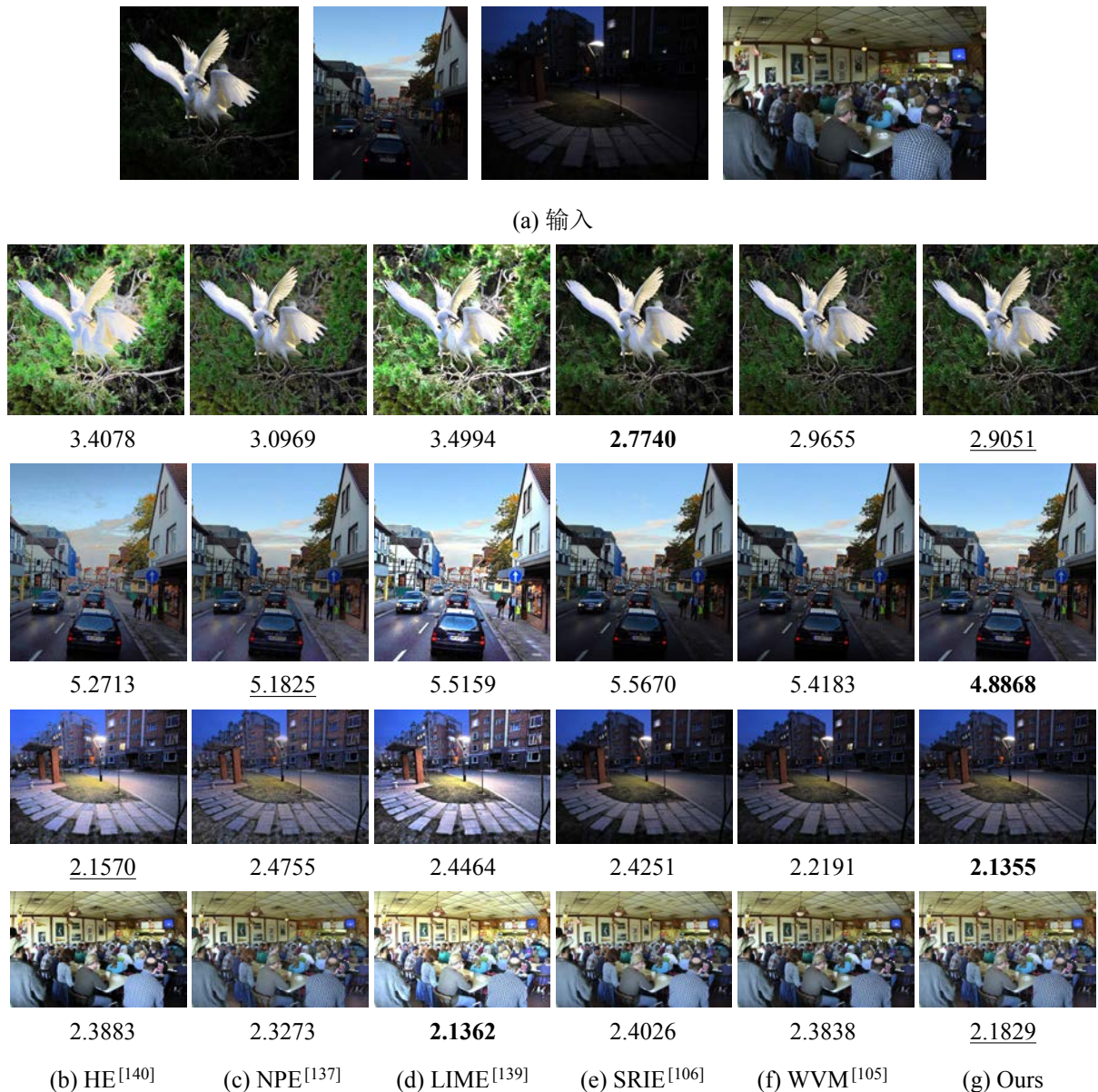


图 4-11 光照自适应的定性与定量对比 (NIQE)

和路灯。此外，根据表4-2的设置，验证不同先验组合对光照自适应的影响。如表4-3所示，联合先验模型在大多数情况下均取得更令人满意的结果。

有效的先验组合有助于 Retinex 模型的求解，因此图4-13分析不同的先验对于模型收敛速度的影响。图4-13 展示收敛误差 $\varepsilon_l = \|\mathbf{I}_k - \mathbf{I}_{k-1}\| / \|\mathbf{I}_{k-1}\|$ ， $\varepsilon_r = \|\mathbf{R}_k - \mathbf{R}_{k-1}\| / \|\mathbf{R}_{k-1}\|$ 与迭代次数 k 的关系。从收敛曲线可看出，联合先验模型的收敛速度远快于无形状先验和无亮度先验，因为有效的形状先验和亮度先验可缩小求解空间，更快地找到最优解。

表 4-2 在 35 幅测试图像中光照自适应的定量对比 (NIQE/ARISM)

方法	HE ^[140]	BPDFHE ^[141]	SSR ^[18]	MSRCR ^[22]	NPE ^[137]	
NIQE	3.4475	3.7267	3.3778	3.4295	3.4091	
ARISM	3.2902	3.3275	3.0469	3.1014	3.0891	
方法	GOLW ^[138]	MF ^[107]	LIME ^[139]	SRIE ^[106]	WVM ^[105]	Ours
NIQE	3.3647	3.5335	3.6155	3.4590	<u>3.3594</u>	3.3409
ARISM	3.3243	3.0200	3.1753	<u>2.9930</u>	2.9958	2.9917

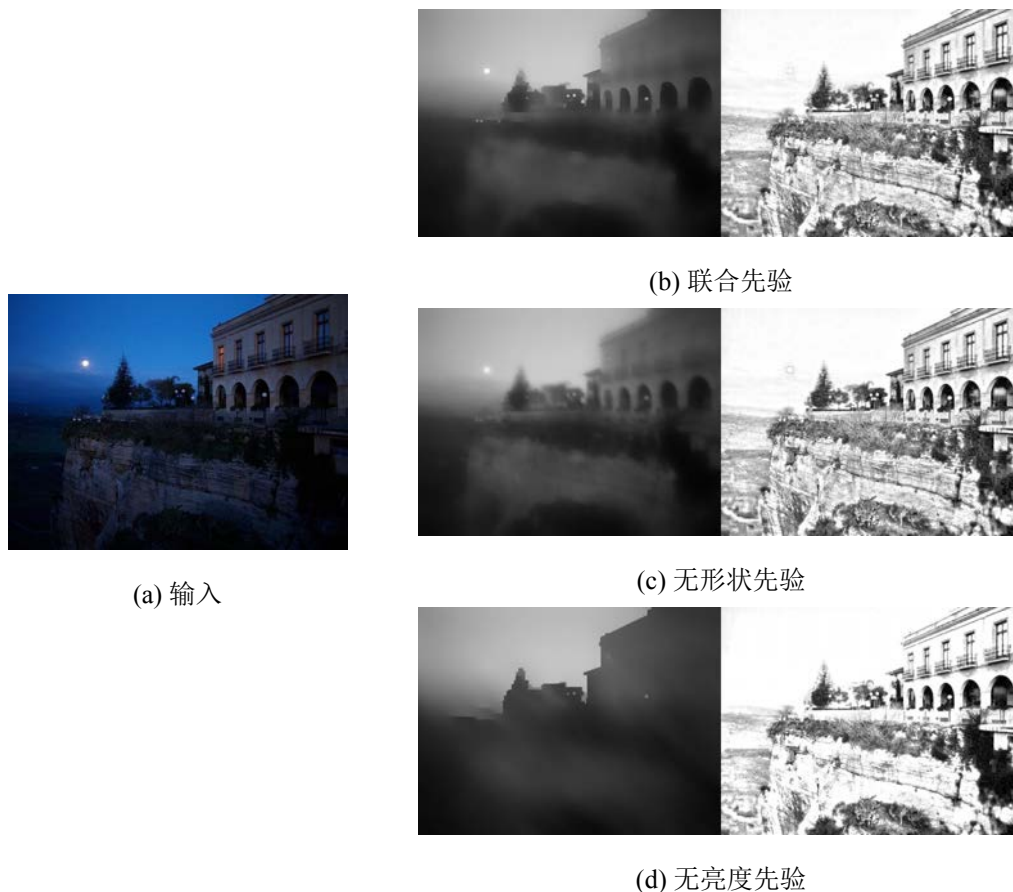


图 4-12 不同先验的性能分析

4.5 小结

本文提出了一种基于联合先验的 Retinex 模型，该模型可同时估计光强度和反射率。首先，为有效地去除纹理并保护结构，本文提出局部变化方差的结构度量方法。基于局部变化偏差，联合“形状、光照和纹理”先验的 Retinex 模型以改进 Retinex 分解的正则化约束。为了高效地分离光照和反射率，本文采用块坐标下降法将优化问题转化为迭代重加权最小二乘法，实现单次迭代 $O(N)$ 的时间复杂度。在实验分析中提出

表 4-3 在 35 幅测试图像中不同先验对光照自适应的影响 (NIQE/ARISM)

指标	NIQE	ARISM
联合先验	3.3409	2.9917
无形状先验	3.3964	2.9999
无光照先验	3.3791	3.0024

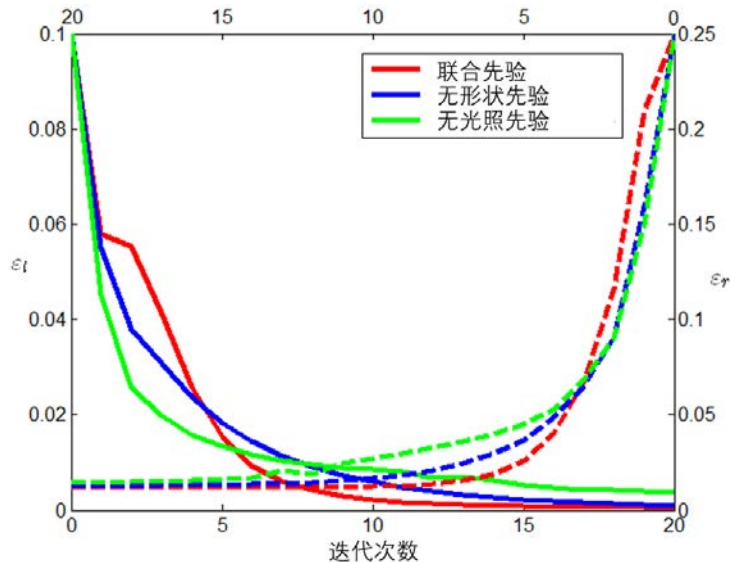


图 4-13 不同先验的收敛分析

了一个综合分析的方法，在 Retinex 分解、颜色校正、亮度自适应等应用方面进行主观和客观对比分析。与现有的多种算法相比，该基于联合先验的 Retinex 模型具有更好的收敛速度和收敛性。

需要指出的是，尽管本章提出了一个有效的联合先验的 Retinex 模型，但受限于多变量和非凸问题的最优化，迭代优化的算法效率并不能完全满足视频处理的实时性需求。此外，在纹理先验上仅采用拉普拉斯分布的弱先验，并不能解决复杂条件下的噪声消除问题。针对以上问题，可分的递归优化^[144]和低秩稀疏的先验^[145]将是本文后续的改进方向。

第五章 基于高斯相关性（RoG）的细节增强

5.1 概述

在雾天环境下, 由于大气中悬浮颗粒使得光线在传播的过程中发生散射、折射、吸收等现象, 从而降低成像的可见度和对比度。此外, 根据大气光散射模型也可推导出雾模型对于对比度的影响: $\sum_x \|\nabla I(x)\| = T(x) \sum_x \|\nabla J(x)\| \leq \sum_x \|\nabla J(x)\|$ 。因此, 细节增强作为去雾后处理的增强操作, 可在去雾结果的基础上进一步优化视觉效果, 如图5-1所示。



图 5-1 图像去雾与细节增强

自然图像提供了丰富视觉信息, 并包含各式各样的物体和对象。数字图像处理中的许多应用通常需要将图像分解为一个基础层和一个细节层来捕获不同的尺度信息: 分段平滑的基础层捕获大尺度的结构信息, 而局部变化的细节层包含小尺度的纹理信息。因此, 图像中的尺度分解不仅可应用于细节增强^[121,146], 还可应用于结构提取^[122,147]、JPEG 伪影消除^[64,148]、HDR 色调映射^[149,150] 等其他任务。

如图5-2所示, 自然图像由不同尺度的内容所组成。小尺度的内容表示细节信息, 如纹理、噪声、反射和小的物体的详细信息; 而大尺度的内容通常编码结构、边界、光

照和平面区域。近几年，基于尺度理论的内容分离技术被广泛应用于语义分割^[151]、目标检测^[152]、人脸识别^[153]等领域。然而，尺度选择性的图像分解仍然是一个难题。

图像金字塔是一种常用的多尺度图像分解的策略，如拉普拉斯金字塔^[154]、多尺度 Retinex^[19] 和尺度空间滤波器^[155]等。然而，这些尺度金字塔均是由线性滤波器构成，由于频谱泄露导致线性滤波器会在边缘附近产生光晕效应。为减少光晕和伪影，学界提出了非线性边缘感知的方法，如双边滤波^[64]、各项异性扩散^[115]和加权最小二乘法^[121]等。然而，边缘感知方法很难将结构和纹理有效地分开，因为边缘强度和内容尺度是两个完全不同的概念。如图5-2所示，因为其结构和纹理区域的梯度均较高，导致基于边缘感知的方法自然而然地保留它们。

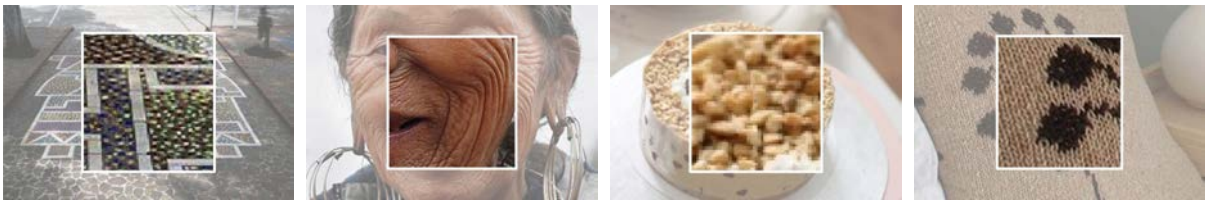


图 5-2 具有强边缘的小尺度内容的自然图像示例

对于纹理和结构的判别，近几年已经出现各种各样的方法在保留重要结构的同时实现纹理的平滑。为抑制高对比度的小尺度信息，Subr 等人^[118]提出局部极值包络来保持结构锐利的同时提取大尺度内容。Xu 等人^[122]提出基于相对全变分的优化框架实现从纹理中提取出主要的结构，但其中大尺度滤波器的重叠削弱了其在小尺度空间的特征识别能力。Karacan 等人^[119]使用区域方差和局部加权平均来实现纹理平滑，但局部统计量需要大量的计算。综上所述，现有的技术大部分能有效地实现结构和纹理的分离，但具有尺度感知的平滑方法仍存在两个主要挑战：

- 不同尺度下的空间变化不可避免地存在混叠，因此尺度信息难以简单地依赖局部像素来定义。
- 尺度感知平滑由于需要提取局部特征或求解大型线性系统，从而带来相对繁重的计算量。

为了解决上述问题，本文提出了一种有效的具有尺度感知的图像平滑方法，其可以选择性地去除不同尺度下的图像内容。该方法具有实现容易、效率高、计算成本低的优点，主要贡献如下：

- 1) 根据尺度空间理论，本文提出了一种新的尺度度量方法——高斯相对性，该度量函数实现在平滑过程中对不同内容尺度的识别。
- 2) 一种可分离的递归优化方法有效地减少了求解稀疏线性系统的主要计算负担，实现在 $O(N)$ 时间复杂度下的有效平滑，即使在单核 CPU 上也可实现实时处理。
- 3) 本文的平滑方法具有很强的通用性，可方便地扩展应用于多个领域，例如细节增强、结构提取、JPEG 伪影消除、HDR 色调映射等。

5.2 基于高斯相关性的图像平滑

为解决尺度感知图像平滑所面临的上述问题，本文提出一种基于**高斯滤波器的相对统计量**——高斯相关性（Relativity-of-Gaussian, RoG）可在**全局优化**中实现对于指定尺度的内容识别，可选择性地去除指定尺度的内容。如图5-3所示，相比于图4-2中的其他边缘/结构保持方法，本文方法在不破坏图像结构的情况下有效地定位并消除纹理（见图5-3b）。

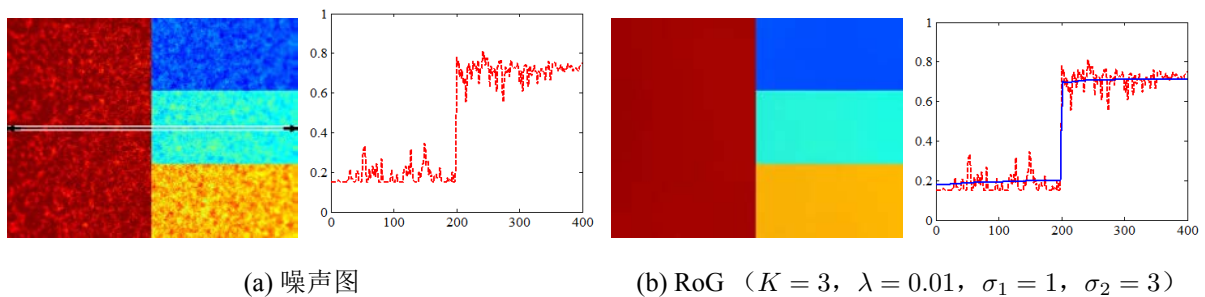


图 5-3 RoG 在噪声图上的平滑结果

5.2.1 高斯相关性

尺度空间表示是将原始图像嵌入到有尺度选择性的衍生图像中。对于视觉系统，尺度空间理论^[156,157]指出自然的尺度感知可采用不同尺度的高斯核卷积。为实现不同尺度的识别，高斯滤波器及其衍生^[158-160]是尺度分析的有效手段。

高斯差分^[161]（Difference-of-Gaussian, DoG）作为经典的尺度识别算法，可定位图像中的显著边缘和细节。在 SIFT^[158]中，DoG 被应用于关键特征点的定位与尺度空间判别。在 DoG 的启发下，本文采用相对性统计取代差分度量，以放大尺度度量之间的差异。因此，本文提出一种新的尺度度量方法——高斯相对性（Relativity-of-Gaussian,

RoG) 选择性地平滑梯度 ∇S 。数学形式上, RoG 度量可表示为:

$$\mathcal{R} = \left| \frac{G_{\sigma_1} * \nabla S}{G_{\sigma_2} * \nabla S} \right|, \text{s.t. } \sigma_1 < \sigma_2. \quad (5-1)$$

其中, 用于尺度选择的高斯核 $G_{\sigma}(x, y) = \exp\left(-\frac{1}{2\sigma^2}((x - x_0)^2 + (y - y_0)^2)\right)$, σ 是尺度参数, (x_0, y_0) 是高斯核的中心。

图5-8a展示合成噪声图, 其中包含小尺度细节(弱纹理)和大尺度结构(强边缘)。如图5-4所示, 可以直观地解释 RoG 的尺度感知属性:

- 1) 小尺度特征 $|G_{\sigma_1} * \nabla S|$ 包含几乎所有梯度变换, 包括弱纹理和强边缘, 如图5-4a所示。
- 2) 大尺度特征 $|G_{\sigma_2} * \nabla S|$ 仅包含强边缘如图5-4b, 因为弱纹理被大尺度的高斯核 G_{σ_2} 平滑。
- 3) 跨尺度的相对统计量 $\mathcal{R}$ 等价于 DoG, 选择性地定位特定尺度特征如图5-4c ——RoG 从显著的大尺度结构 ($\sigma_2 = 3$) 中捕获小尺度纹理 ($\sigma_1 = 1$)。

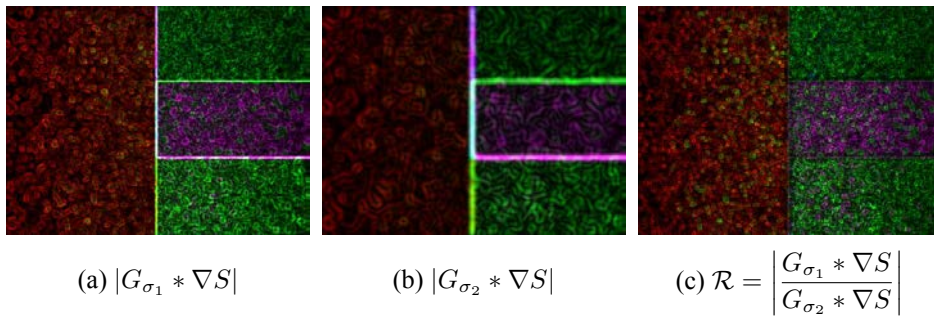


图 5-4 RoG 在噪声图上的分析

5.2.2 平滑优化

为了有效地过滤输入图像 I 中不同尺度的内容, 采用 RoG 度量的平滑优化函数可表示为:

$$S = \arg \min_S \|S - I\|_2^2 + \lambda \left(\left\| \frac{G_{\sigma_1} * \nabla_x S}{G_{\sigma_2} * \nabla_x S} \right\|_1 + \left\| \frac{G_{\sigma_1} * \nabla_y S}{G_{\sigma_2} * \nabla_y S} \right\|_1 \right). \quad (5-2)$$

其中, I 亮度强度或彩色通道, λ 是正则参数。这里的第一项 $\|S - I\|_2^2$ 是 L2 范数的数据保真项, 用于最小化平滑结果 S 和输入图像 I 。

因为基于 L1 范数的 RoG 度量是非光滑的正则项, 因此其不存在简单的闭合解。本文采用迭代重加权最小二乘法^[126] (Iteratively Re-weighted Least Square, IRLS) 求解此

非凸正则化问题。此处详细讨论水平方向（ x 方向）的 RoG 度量优化，对于 y 方向可做类似处理。 x 方向的 RoG 度量可重写为：

$$\begin{aligned} \left\| \frac{G_{\sigma_1} * \nabla_x S}{G_{\sigma_2} * \nabla_x S} \right\|_1 &= \left\| \frac{(G_{\sigma_1} * \nabla_x S)^2}{(G_{\sigma_2} * \nabla_x S)(G_{\sigma_1} * \nabla_x S)} \right\|_1 \\ &\approx \left\| \frac{G_{\sigma_1/2} * (\nabla_x S)^2}{(G_{\sigma_2} * \nabla_x S)(G_{\sigma_1} * \nabla_x S)} \right\|_1 \\ &= G_{\sigma_1/2} * \frac{1}{|(G_{\sigma_2} * \nabla_x S)(G_{\sigma_1} * \nabla_x S)|} \|\nabla_x S\|_2^2 \end{aligned} \quad (5-3)$$

式(5-3)中第二行是卷积分解的近似数值解。因此，RoG 度量可分解为二次项 $\|\nabla_x S\|_2^2$ 和非线性加权 $w_{x,y}$ ：

$$w_{x,y} = G_{\sigma_1/2} * \frac{1}{|(G_{\sigma_2} * \nabla_{x,y} S)(G_{\sigma_1} * \nabla_{x,y} S)| + \varepsilon}. \quad (5-4)$$

其中， ε 是极小值以避免被除数为零。

综上，平滑优化过程可迭代求解，对于第 k 次迭代表示为：

$$\begin{aligned} S^k &= \arg \min_S \|S - I\|_2^2 + \\ &\quad \lambda \left(w_x \|\nabla_x S^{k-1}\|_2^2 + w_y \|\nabla_y S^{k-1}\|_2^2 \right) \end{aligned} \quad (5-5)$$

初始化 $S^0 = I$ 并采用矩阵形式重写二次损失函数：

$$(\mathbf{S} - \mathbf{I})^T (\mathbf{S} - \mathbf{I}) + \lambda (\mathbf{S}^T \mathbf{D}_x^T \mathbf{W}_x \mathbf{D}_x \mathbf{S} + \mathbf{S}^T \mathbf{D}_y^T \mathbf{W}_y \mathbf{D}_y \mathbf{S}). \quad (5-6)$$

其中， $\mathbf{S}$ 和 $\mathbf{I}$ 分别是 S 和 I 的向量形式， $\mathbf{W}_{x,y}$ 是以加权值 $w_{x,y}$ 为对角线的对角矩阵，矩阵 $\mathbf{D}_x$ 和 $\mathbf{D}_y$ 是前向梯度算子的特普利茨（Toeplitz）矩阵。最小化式(5-6)可求得此线性系统的唯一解 $\mathbf{S}$ ：

$$\mathbf{S}^k = (\mathbf{1} + \lambda \mathbf{L}^{k-1})^{-1} \mathbf{I}. \quad (5-7)$$

其中， $\mathbf{1}$ 是单位矩阵， $\mathbf{L}^k = \mathbf{D}_x^T \mathbf{W}_x^k \mathbf{D}_x + \mathbf{D}_y^T \mathbf{W}_y^k \mathbf{D}_y$ 是五点正定拉普拉斯（Laplacian）矩阵^[127]。完整优化过程见算法5-1。

算法 5-1: 基于 RoG 的尺度感知平滑

Input: 输入图像 $\mathbf{I}$ ，尺度参数 $\sigma_{1,2}$ ，正则参数 λ ，最大迭代 K

Output: 平滑结果 $\mathbf{S}$

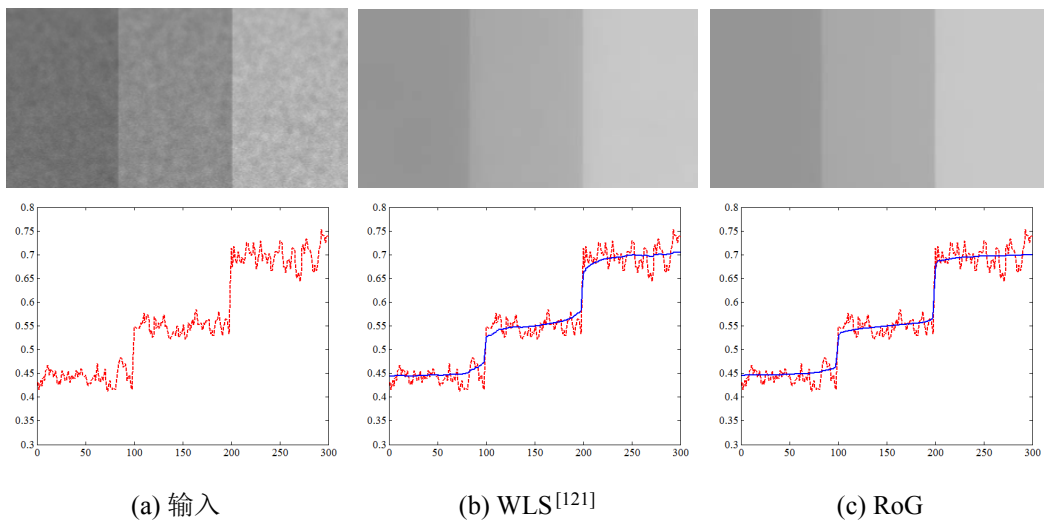
- 1 初始化 $\mathbf{S}^0 \leftarrow \mathbf{I}$;
 - 2 **for** $k = 1$ **to** K **do**
 - 3 根据式(5-4)计算权重 $w_{x,y}$;
 - 4 根据式(5-7)更新 $\mathbf{S}^k$;
-

5.2.3 理论分析

在本节中, 将对比和分析基于 RoG 的平滑方法与其他类似平滑方法, 包括边缘保持、结构保持和尺度感知。

5.2.3.1 边缘保持的平滑

在算法5-1中, 当最大迭代次数 $K = 1$ 时, 目标函数将近似转换成边缘保持的平滑方法 WLS^[121]。不同于 WLS 使用对数域加权, RoG 在原始的线性域对信号进行权值计算和平滑优化。虽然对数变换更符合模拟人类视觉感知机制的韦伯 (Weber) 定理^[162], 但会带来边缘感知的非线性。给定输入信号 S , 对数域中的 WLS 权值 $(\nabla \log S)^{-1} = S/\nabla S$ 。当 S 变大时, $(\nabla \log S)^{-1}$ 随之被放大, 这将导致在高亮度区域过度平滑。如图5-5所示, 在高光区域的第二个边缘相比于第一个更加平滑。相反地, RoG 采用原始线性域的相对函数可在高光和阴影区域平等地实现边缘保持, 如图5-5。



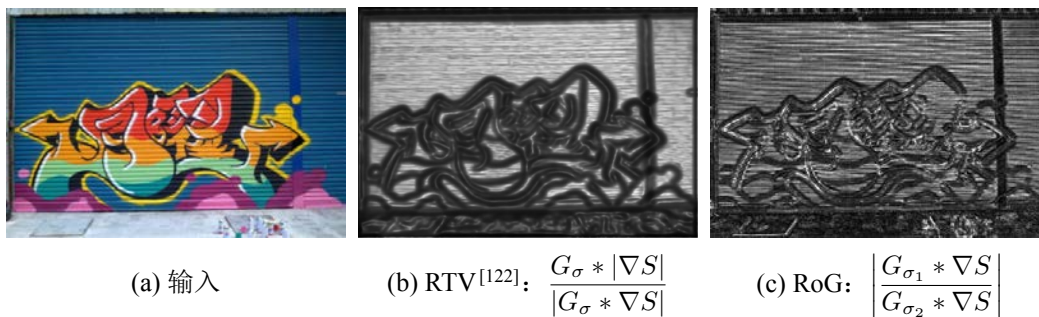
(b) WLS ($\lambda = 1.0$, $\alpha = 1.2$); (c) RoG ($K = 1$, 其他参数如图5-4)。

图 5-5 边缘保持策略分析

5.2.3.2 结构保持的平滑

当 $K > 1$ 时, 迭代求解过程类似于 RTV^[122] 从纹理中提取结构。RTV 的结构特征计算采用单尺度高斯核, 因此相邻两个局部块之间存在很大的像素重叠, 这将降低结构特征的分辨率。相反地, RoG 计算不同尺度高斯核 ($\sigma_1 < \sigma_2$) 的相对性函数, 实现高分辨率的特征提取。如图5-6所示, RoG 度量可准确地避开显著的结构, 并捕捉到更清晰的纹理。此外, 跨尺度的相对性是一种带通滤波器, 可以选择性地提取任意尺度的

结构。

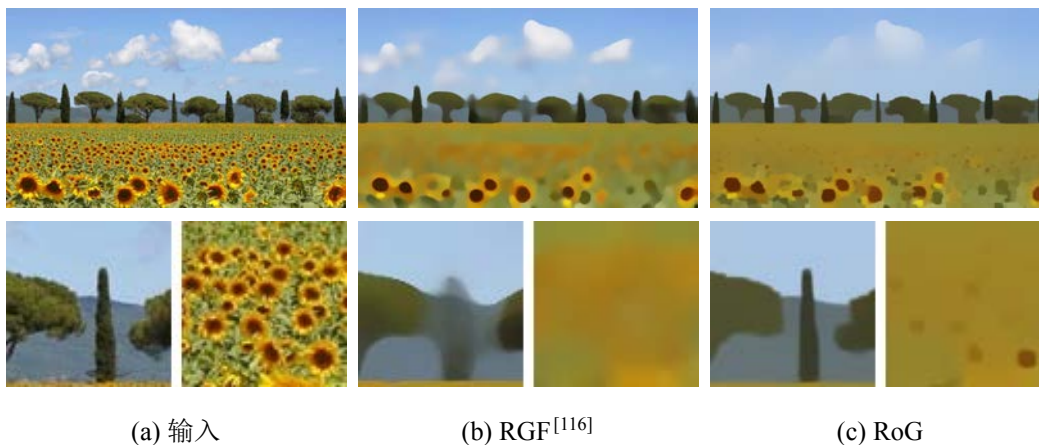


(b) RTV ($\sigma = 3$); (c) RoG ($\sigma_1 = 1, \sigma_2 = 3$)。

图 5-6 结构保持策略分析

5.2.3.3 尺度感知的平滑

本文采用的 IRLS 算法^[126] 的循环求解过程近似与尺度感知滤波器——循环导向滤波器^[116] (Rolling Guidance Filter, RGF)。在 RGF 中, 引导图首先通过指定尺度的高斯核滤波生成, 并通过联合滤波器迭代更新。然而, 由于第一步的高斯模糊导致 RGF 出现颜色扩散(树木融化到天空背景)和细节丢失(远处的向日葵过度平滑)。相比之下, RoG 更关注局部一致的梯度边缘, 并在无初始化模糊下进行迭代优化。因此, 如图5-7c所示, RoG 的平滑结果更好地保护了小尺度的显著特征。



(b) RGF ($\sigma_s = 5, \sigma_r = 0.1, K = 3$); (c) RoG (参数如图5-4)。

图 5-7 尺度感知策略分析

5.3 快速可分递归优化

基于 RoG 平滑的最大计算瓶颈在于求解大规模的稀疏线性系统，见式(5-7)。为提高其效率，预条件共轭梯度^[114] (Preconditioned Conjugate Gradient, PCG) 的快速求解方法可使得每次迭代的时间复杂度从 $O(N^3)$ 降低到 $O(N)$ 。然而，这仍然无法达到真正的实时处理。在本节中将采用可分递归优化^[144] 实现在单 CPU 下的实时处理。

5.3.1 一维递归优化

式(5-7)中定义的稀疏线性系统可沿水平和垂直分解。这里将详细讨论 x 方向输入信号 $\mathbf{I}_x$ 和输出信号 $\mathbf{S}_x$ 的求解， y 方向的处理与此类似。一维的输出结果 $\mathbf{S}_x$ 可通过式(5-7)求解，并重写如下：

$$(\mathbf{1} + \lambda \mathbf{L}_x) \mathbf{S}_x = \mathbf{I}_x. \quad (5-8)$$

其中， $\mathbf{1}$ 是尺寸 $W \times W$ 为单位矩阵 (W 是输入图像的宽)， $\mathbf{S}_x$ 和 $\mathbf{I}_x$ 分别是水平方向信号的向量形式。不同于式(5-7)中的五点拉普拉斯矩阵， $\mathbf{L}_x = \mathbf{D}_x^T \mathbf{W}_x \mathbf{D}_x$ 是稀疏的三点拉普拉斯矩阵。因此，求解式(5-8)远简单于求解式(5-7)，因为三点对角矩阵 $\mathbf{L}_x$ 仅在中心、左边和右边的对角线上存在非零值。因此，式(5-8)的线性系统可表示为：

$$\begin{bmatrix} b_0 & c_0 & 0 & 0 \\ a_1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & c_{W-2} \\ 0 & 0 & a_{W-1} & b_{W-1} \end{bmatrix} \begin{bmatrix} s_0 \\ \vdots \\ \vdots \\ s_{W-1} \end{bmatrix} = \begin{bmatrix} f_0 \\ \vdots \\ \vdots \\ f_{W-1} \end{bmatrix}. \quad (5-9)$$

这里 s_u 和 f_u 分别是 $\mathbf{S}_x$ 和 $\mathbf{I}_x$ 的元素，其中 $u \in [0, W-1]$ 是坐标索引。 a_u ， b_u 和 c_u 是 $(\mathbf{1} + \lambda \mathbf{L}_x)$ 中的三组非零对角线，可表示为：

$$\begin{cases} a_u = -\lambda w_u \\ b_u = 1 + \lambda(w_u + w_{u+1}) \\ c_u = -\lambda w_{u+1} \end{cases} \quad (5-10)$$

其中，边缘坐标位置 $a_0 = 0$ ， $c_{W-1} = 0$ 。

采用高斯消元法^[163]，这类矩阵自然可实现 $O(N)$ 时间复杂度的非近似加速。高斯消元法使用矩阵初等变换，对增广矩阵 $\mathbf{A} = [\mathbf{1} + \lambda \mathbf{L}_x | \mathbf{I}_x]$ 进行对角化。对角化的上三角

矩阵 $\tilde{\mathbf{A}}$ 通过自上而下的递归行操作, 可得

$$\begin{cases} \tilde{a}_u = 0 \\ \tilde{b}_u = 1 \\ \tilde{c}_u = \frac{c_u}{b_u - a_u c_{u-1}} \\ \tilde{f}_u = \frac{f_u - a_u f_{u-1}}{b_u - a_u c_{u-1}} \end{cases} \quad (5-11)$$

因此, 输出信号 s_u 可自下而上地求解:

$$s_u = \tilde{f}_u - \tilde{c}_u s_{u+1}. \quad (5-12)$$

其中, $s_{W-1} = \tilde{f}_{W-1}$ 。

5.3.2 二维可分近似

对于二维图像, 可将其分解为两个一维信号, 并分别沿水平和垂直方向递归地应用一维优化。然而, 两个维度独立地求解会导致在非正交的边缘产生条纹状的伪影, 这是可分平滑算法中的普遍缺点。因此, 为解决这一问题, 本文采用多次递归地应用一维平滑实现二维平滑。并逐步减小平滑参数 λ_t 以自适应(5-8)中的 λ 来控制空间平滑度, 从而避免了分解优化导致的条纹伪影。

根据快速全局优化方法^[144], 通过多次指数加权的一维平滑可近似地求解二维的平滑求解。平滑参数 λ_t 在每一次迭代 (包括水平方向和垂直方向的两个个子优化) 过程中逐步减半, 可得以下方程:

$$\lambda = \sum_{t=1}^T 2\lambda_t = \sum_{t=1}^T 2(0.5^{2t}\lambda_1). \quad (5-13)$$

其中, λ_t 是第 i 次迭代的平滑参数, T 是迭代求解的次数 (本文设迭代次数 $T = 3$)。

根据此公式, λ_t 可表示为:

$$\lambda_t = \frac{3}{2} \frac{4^{T-t}}{4^T - 1} \lambda. \quad (5-14)$$

综上所述, 快速可分递归优化过程可总结如算法5-2.

此外, 采用可分的权值计算可进一步提高了计算效率。在式(5-4)中, 涉及到三个复杂度为 $O(\sigma^2 N)$ 的卷积操作。快速傅立叶变换 (Fast Fourier Transform, FFT) 可以将卷积操作加速到 $O(N \log N)$, 但将引入大高斯核的边界伪影。在本文中, 应用高斯核的二维可分特性, 通过使用了两个一维的高斯滤波器 (水平方向和垂直方向) 将复杂度降低为 $O(\sigma N)$ 。

算法 5-2: 基于 RoG 的快速尺度感知平滑**Input:** 输入图像 $\mathbf{I} = \{f_{u,v}\}_{(u,v)=(0,0)}^{(W-1,H-1)}$, 尺度参数 $\sigma_{1,2}$, 正则参数 λ , 最大迭代 K **Output:** 平滑结果 $\mathbf{S} = \{s_{u,v}\}_{(u,v)=(0,0)}^{(W-1,H-1)}$

```

1 初始化  $\mathbf{S}^0 \leftarrow \mathbf{I}$ ;
2 for  $k = 1$  to  $K$  do
3   根据式(5-4)计算  $w_{x,y}$ ;
4   for  $t = 1$  to  $T$  do
5     根据式(5-14)计算  $\lambda_t$ ;
6      $f_{u,y} \leftarrow s_{u,v}$ ;
7     for  $v = 0$  to  $H - 1$  do
8       根据式(5-10)通过  $(\mathbf{1} + \lambda_t \mathbf{L}_x)$  计算  $a_u, b_u, c_u$ ;
9       根据式(5-12)求解水平方向  $s_u$ ;
10    for  $u = 0$  to  $W - 1$  do
11      根据式(5-10)通过  $(\mathbf{1} + \lambda_t \mathbf{L}_y)$  计算  $a_v, b_v, c_v$ ;
12      根据式(5-12)求解垂直方向  $s_v$ ;
13 更新  $\mathbf{S}^k = \{s_{u,v}\}_{(u,v)=(0,0)}^{(W-1,H-1)}$ ;

```

5.4 实验与分析

为验证尺度感知的图像平滑, 本文从细节增强、结构提取、JPEG 伪影消除、HDR 色调映射这几个方面对 RoG 的平滑方法进行评测。更多的应用代码和对比结果可见https://caibolun.github.io/rog_plus/。

5.4.1 细节增强

当 $K = 1$ 时, 基于 RoG 的平滑方法作为一种非线性的边缘, 可实现基础层和细节层的分解, 从而实现对于图像细节的增强。本文简单地应用 RoG 平滑方法替代经典细节增强框架中的边缘保持平滑。对于输入图像 I , 基于 RoG 的平滑方法提取并增强细节层信息 $I^{Detail} = I - S$ 并组合基础层信息 $I^{Base} = S$ 。增强后图像可表示为:

$$I^{Enhance} = I^{Detail} + \beta I^{Base}. \quad (5-15)$$

其中, 增强参数 β 用于平衡细节层与基础层的权重。

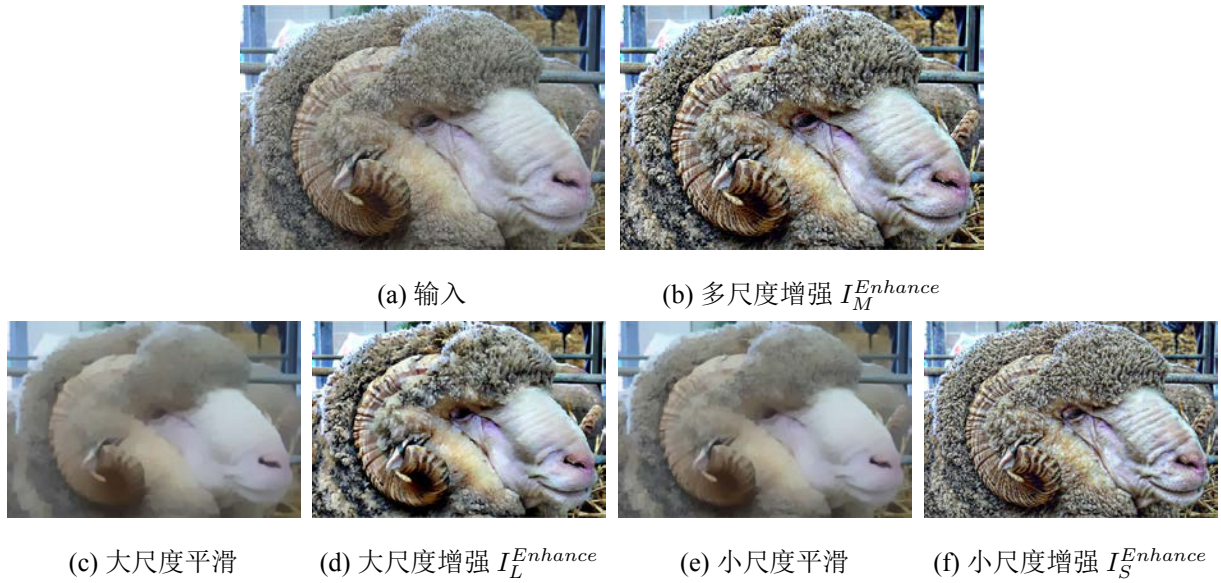
对于细节增强，光晕的伪影和噪声和放大是增强过程中需要解决的两个主要问题。本实验对比了 RoG 方法与最近几年的优秀细节增强方法，包括 WLS^[121]、LLF^[164]、RTV^[122] 和 mRTV^[165]。在图5-8中，第一行展示边缘保持平滑的基础层 I^{Base} ，第二行展示细节增强后的结果 $I^{Enhance}$ ，最后一行展示增强图的局部特写——蓝色箭头指向光环伪影，红色箭头指向噪声放大。如图所示，基于 RoG 的细节增强可有效地避免了过去方法中常见的光晕和噪音。



(b) WLS ($\lambda = 1.0$, $\alpha = 1.2$); (c) LLF ($\alpha = 2$, $\sigma_r = 0.4$); (d) RTV ($\lambda = 0.015$, $\sigma = 0.5$); (e) mRTV ($k = 3$, $n_{itr} = 5$); (f) RoG ($K = 1$, $\lambda = 0.001$, $\sigma_1 = 0.5$, $\sigma_2 = 1.0$)。

图 5-8 细节增强度对比

此外，因为 RoG 采用跨尺度的相对函数对尺度空间进行度量，因此其可实现不同尺度内容的选择性增强。为验证多尺度的细节增强效果，本文通过控制 $\sigma_{1,2}$ 来构造不同尺度空间的细节层分解，并调节基础层的强度和对比度。如图5-9所示，基于 RoG 的平滑方法可有效地实现不同尺度空间的平滑，并在指定的尺度上选择性地实现细节增强。融合大尺度增强 $I_L^{Enhance}$ 和小尺度增强 $I_S^{Enhance}$ 的结果，可得到联合多尺度的细节增强 $I_M^{Enhance} = 0.5I_L^{Enhance} + 0.5I_S^{Enhance}$ ，如图5-9b。



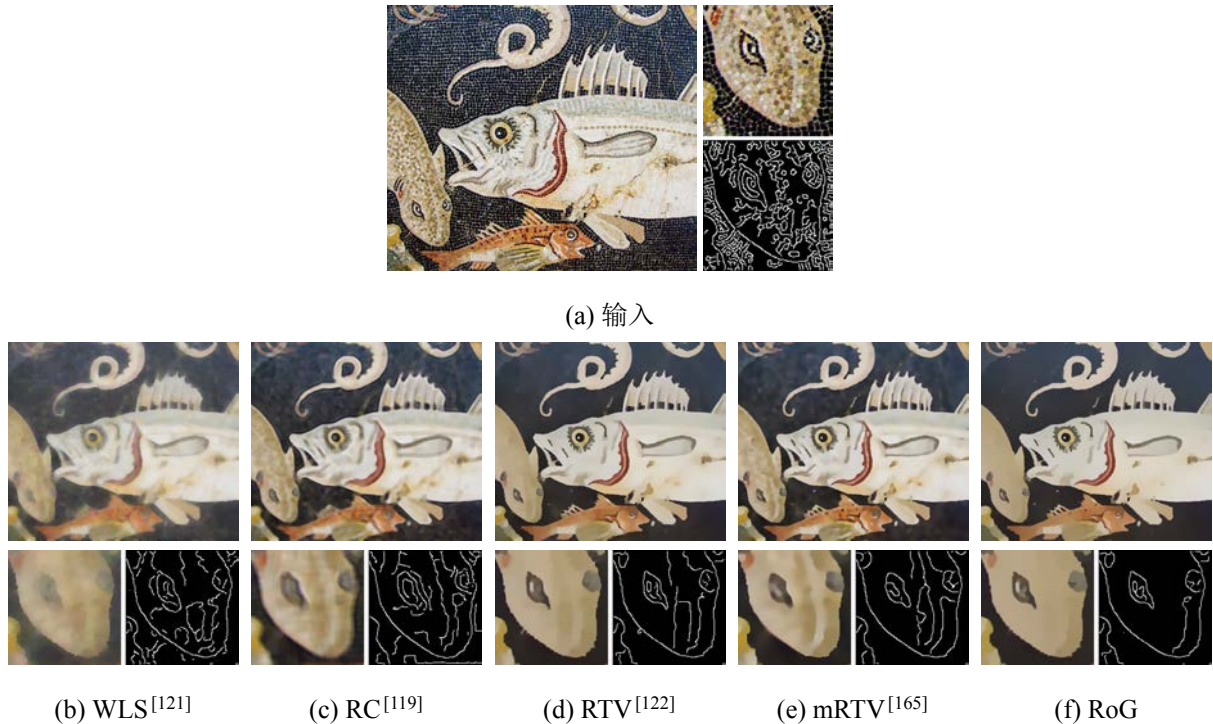
$K = 1$, $\lambda = 0.001$, 大尺度: $\sigma_1 = 1.0$ 和 $\sigma_2 = 1.5$; 小尺度: $\sigma_1 = 0.5$ 和 $\sigma_2 = 1.0$ 。

图 5-9 基于 RoG 的多尺度细节增强

5.4.2 结构提取

当 $K > 1$ 时, 基于 RoG 的平滑方法具有结构保持作用。本小节将在结构提取方面比较和分析 RoG 的平滑方法与当前最新的技术。图像中的结构由边缘所决定, 但图像中的小尺度纹理使得大尺度结构的提取极具挑战性。一个优秀的结构提取方法可以有效地避免纹理的干扰并捕捉真正的结构。如图5-10所示, 本文对比了 RoG 平滑方法和几个优秀的非线性平滑方法, 包括 WLS^[121]、LLF^[164]、RTV^[122] 和 mRTV^[165]。针对测试图像, 对不同方法均最优化地手动调整参数设置, 并通过 Canny 算子^[166] 生成结构图。由图5-10可见, 经典的非线性平滑方法 WLS 仅关注边缘保持而缺少结构保持, 从而导致图像的纹理和结构的模糊; 局部的统计滤波方法 RC 可一定程度保护结构信息, 但缺少全局上对纹理和结构的优化; 经典的结构保持平滑方法 RTV 和 mRTV 可有效地提取结构信息, 但缺少尺度感知作用使得对细节处理的分辨率不足。本文提出的基于 RoG 的平滑方法在保护鱼眼结构的情况下可更好地平滑鱼鳞区域, 如图5-10f。

为定量地评估结构提取的效果, 本文以结构保持的平滑算法作为预处理, 并使用最简单的边缘检测器 Sobel 算子^[167] 进行边缘提取。在伯克利分割数据集^[168] (Berkeley Segmentation Dataset, BSD300) 上, 使用基准评价代码对 100 个带有边缘标定的测试图像进行定量分析。图5-11展示了采用不同结构保持的平滑方法作为预处理后的边缘检测结果, 其中包括 L0^[148]、BLF^[64]、AD^[115]、WLS^[121]、RTV^[122] 和 RGF^[116]。定量评估



(b) WLS ($\lambda = 1.0$, $\alpha = 1.2$, $k = 13$); (c) RC ($\sigma = 0.2$, $k = 19$, $Mod\ 1$); (d) RTV ($\lambda = 0.015$, $\sigma = 6$); (e) mRTV ($k = 7$, $n_{itr} = 5$); (f) RoG ($K = 4$, $\lambda = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 4$)。

图 5-10 结构提取的对比

上，采用边缘检测邻域常用的评价指标——F-Score 度量精确率（Precision， P ）和召回率（Recall， R ）的加权调和平均：

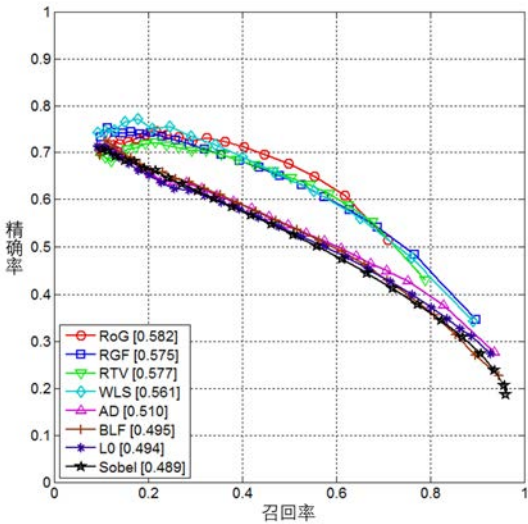
$$F = \frac{2PR}{P + R}. \quad (5-16)$$

从图5-11方括号中的 F-Score 可以清楚看出，基于 RoG 的平滑方法可有效地减少场景中的无效细节，从而生成简化的场景，以便更好地进行边缘检测。

5.4.3 JPEG 伪影消除

得益于 RoG 特殊的边缘保持能力，基于 RoG 的平滑方法可有效地对有损压缩后的剪贴画进行伪影消除。从图5-12可注意到，一般去噪（如 BM3D^[169]）或去振铃（如 Arti^[170]）方法并不适合于此类型退化的恢复，因为无损压缩的伪影与边缘具有很强的相关性。图5-12中 RoG 的参数设置为 $K = 4$ 、 $\lambda = 0.0075$ 、 $\sigma_1 = 1.0$ 和 $\sigma_2 = 1.5$ 。相比之下，基于 RoG 的平滑方法在无训练下也可有效地修复这些退化的剪贴画——在保持边缘的前提下有效地平滑压缩伪影。

此外，本文对比分析了最近相关的平滑方法，包括 AD^[115]、L0^[148]、BLF^[64]、WLS^[121]、RTV^[122] 和 RGF^[116]，用于消除有损图像压缩中基于块离散余弦变换带来的



L0: $\lambda = 0.1$, $\kappa = 2$; AD: $\lambda = 0.15$, $\kappa = 20$; BLF: $\sigma_s = 12$, $\sigma_r = 0.45$; WLS: $\lambda = 0.35$, $\alpha = 1.8$; RTV: $K = 4$, $\lambda = 0.015$, $\sigma = 3$, $\varepsilon_s = 0.02$; RGF: $\sigma_s = 9$, $\sigma_r = 0.05$; RoG: $K = 3$, $\lambda = 0.004$, $\sigma_1 = 1$, $\sigma_2 = 3$ 。

图 5-11 在 BSD300 上的结构提取对比

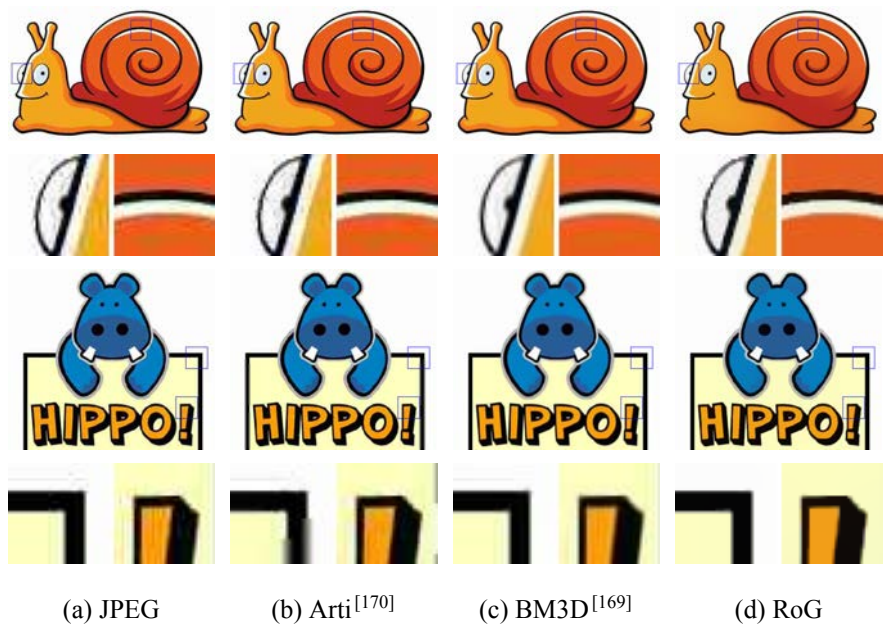


图 5-12 JPEG 伪影消除的对比

压缩伪影。为定量的分析和比较，本文收集 100 张 Emoji^{注 1}并使用质量参数 $Q = 10$ 的标准 JPEG 有损压缩格式（如图5-13）。应用不同平滑方法进行恢复后对比其峰值信噪比（Peak Signal Noise Ratio, PSNR）。图5-14展示各种方法的质量生存曲线和平均

^{注 1}这些 Emoji 图像收集自 Microsoft Windows 10 Creators Update（2017 年 4 月 11 日发布）。

PSNR 值, 实验结果表明 RoG 相比于其他平滑方法获得了最佳的总体性能。从图中可看出, RoG 平滑方法的平均 PSNR 相比 RGF^[116] 高了 0.213 db, RoG 的在 $\text{PSNR} \geq 77\text{dB}$ 下的生存率为 92%, 而次优方法仅为 85%。

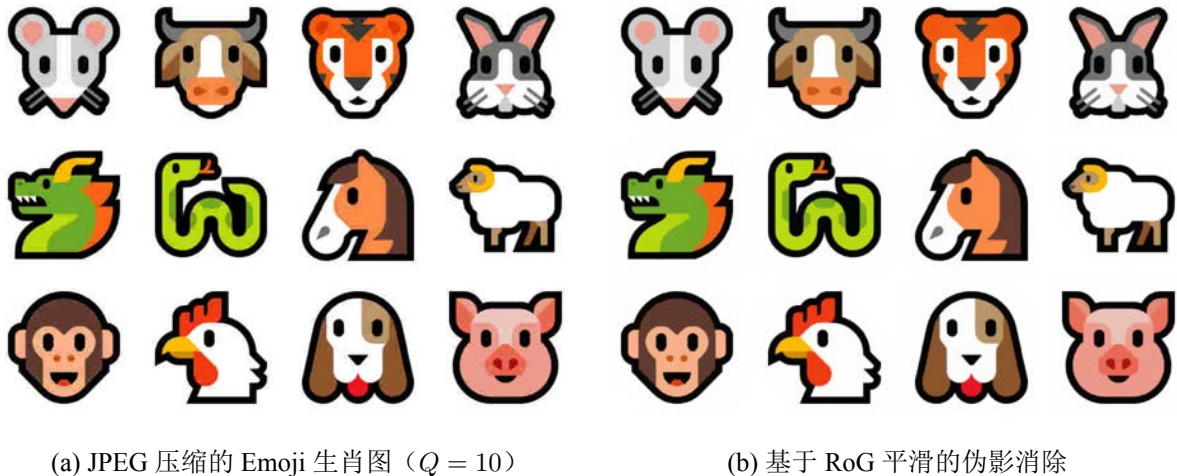
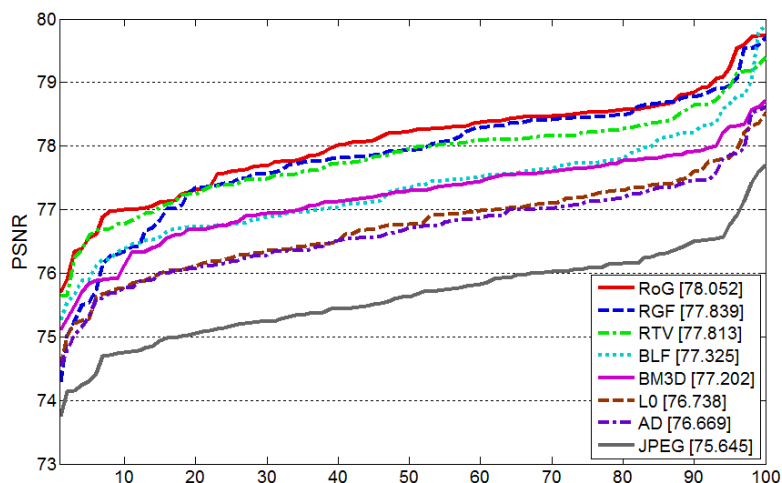


图 5-13 压缩后 Emoji 生肖图与 RoG 伪影消除结果



参数设置如下: AD ($\lambda = 0.15$, $\kappa = 20$); L0 ($\lambda = 0.1$, $\kappa = 2$); BLF ($\sigma_s = 12$, $\sigma_r = 0.45$); RTV ($K = 4$, $\lambda = 0.015$, $\sigma = 3$, $\varepsilon_s = 0.02$); RGF ($\sigma_s = 9$, $\sigma_r = 0.05$); RoG ($K = 4$, $\lambda = 0.0075$, $\sigma_1 = 1.0$, $\sigma_2 = 1.5$)。

图 5-14 在 100 张 Emoji 图片下 JPEG 伪影消除的对比

5.4.4 HDR 色调映射

HDR 色调映射是图像处理中的一个传统问题, 其解决在传统的低动态范围 (Low-dynamic Range, LDR) 显示设备上渲染高动态范围 (High-dynamic Range, HDR) 场

景。基于 RoG 的平滑方法可容易地被应用于 HDR 图像的色调映射。根据经典 HDR 色调映射方法^[149], RoG 平滑方法可简单地替换基于 LCIS 的多尺度分解。具体来说, 本文对 HDR 图像的亮度通道进行 4 个不同尺度空间的分解 (一个粗糙的基础层和三个精细的细节层), 并将每个分解层加权并重建成一个 LDR 的显示图像。

图5-15展示和对比不同方法生成 HDR 色调映射的结果。本文选取最新的色调映射方法作为对比, 其中包括: LCIS^[149]、TR^[171]、WLS^[121] 和 LLF^[164]。此外, 这些方法的饱和度和曝光度均手动地进行调整以达到最优的视觉效果。多曝光融合是 HDR 色调映射的主要问题, 因此精确的基础层分解是实现 LDR 图像渲染的关键。基于 RoG 的平滑方法较少地压缩基础层, 适度地提升细节层, 从而获得更好的视觉效果。与其他色调映射方法相比, 基于 RoG 的色调映射方法更好地渲染出木柱的油漆光泽。

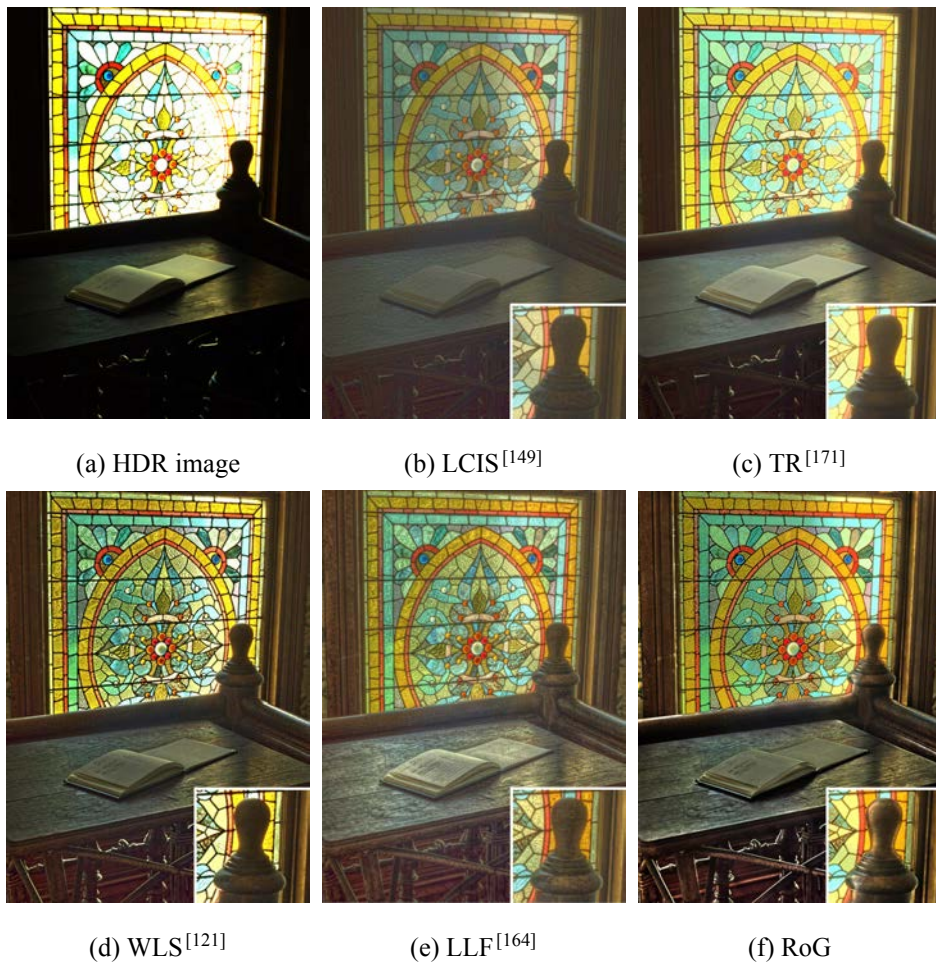


图 5-15 HDR 色调映射的对比

5.4.5 其他应用

如节5.2.3中分析，RoG 具有边缘保持、结构保持和尺度感知等特性，因此基于 RoG 的平滑方法可应用于各式各样的图像非线性平滑应用中，包括图像摘要、铅笔草图、纹理编辑和内容感知图像缩放。

5.4.5.1 图像摘要

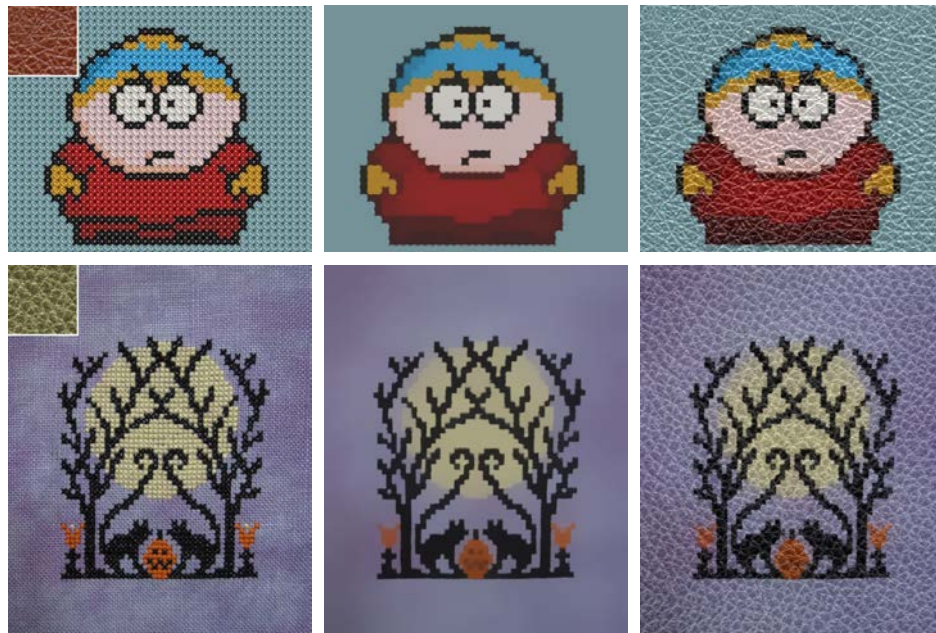
图像摘要旨在通过细节平滑和边缘增强来生成非真实感的图像。经典图像摘要方法包括两个主要步骤——基于边缘保持滤波的图像平滑和基于 DoG^[161] 的线条提取。如图5-16所示，基于 RoG 的平滑方法可在平滑优化的过程中同时实现线条的提取，因此在基于 RoG 的平滑图像 ($K = 3$, $\lambda = 0.005$, $\sigma_1 = 1$, $\sigma_2 = 2$) 上通过计算梯度图直接生成铅笔草图。



图 5-16 图像摘要与铅笔草图

5.4.5.2 纹理编辑

好的纹理和结构分离方法将有益于图像纹理的编辑。RoG 的尺度感知能力可有效地分离纹理层和结构层，并通过其他纹理的替换即可对纹理进行编辑。如图 5-17所示，本文使用基于 RoG 的平滑方法 ($k = 3$, $\lambda = 0.01$, $\sigma_1 = 1$, $\sigma_2 = 3$) 将输入图像分为纹理层和结构层，并将原始纹理替换为皮革纹理会产生合理的纹理编辑结果。



(a) 输入

(b) 结构分解

(c) 纹理替换

图 5-17 纹理编辑

5.4.5.3 内容感知的图像缩放

本文提出的 RoG 平滑方法还可应用于内容感知的接缝裁剪 (Seam Carving)。自然场景通常包含许多细节,但相比于感兴趣物体而言局部的细节并不重要。如图5-18所示,海浪和云彩具有较大的梯度强度并影响接缝裁剪对图像尺寸的调整。基于 RoG 的平滑方法 ($K = 1$, $\lambda = 0.001$, $\sigma_1 = 1$, $\sigma_2 = 1.5$) 可提取图像的主要内容,使得大部分的裁剪接缝不通过帆船,并取得更为有效的图像缩放效果。

5.4.6 运算效率分析

算法的效率分析实验在一台配置 Intel i7 (3.4GHz) 处理器的台式机中 Matlab 2014a 上进行。本文对比与之相关的非线性平滑方法在灰色图像的处理速度,其中包括 RTV^[122]、WLS^[121]、L0^[148]、AD^[115] 和 RGF^[116]。表5-1报告了采用 PCG 优化 (表示为 RoG) 和快速可分递归优化 (表示为 RoG+) 的运算速度,其中结构提取 (图5-11) 和伪影消除 (图5-14) 的结果作为性能的参考。在快速可分递归优化下,基于 RoG 的优化方法取得最快且最优的性能,即便是在单个 CPU 核上也能实时地处理——0.16 秒/百万像素。

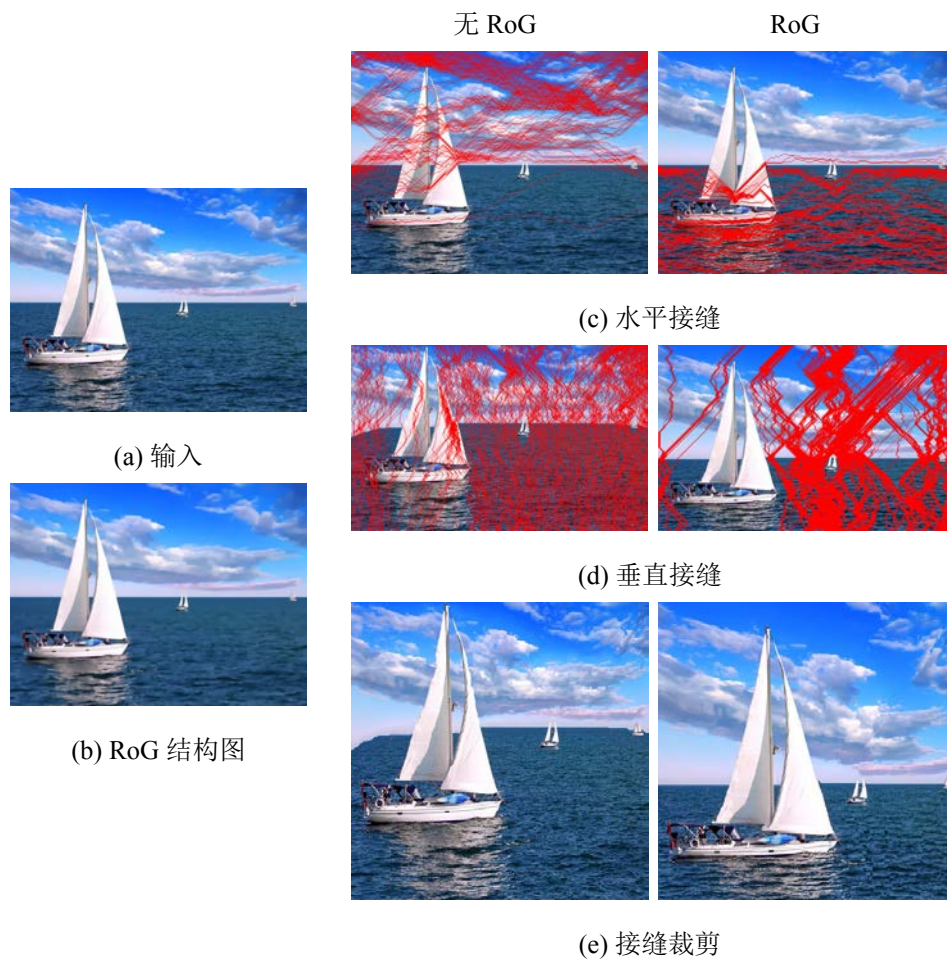


图 5-18 内容感知的图像缩放

表 5-1 运算效率对比

方法	RTV ^[122]	WLS ^[121]	L0 ^[148]	AD ^[115]	RGF ^[116]	RoG	RoG+
运算速度（秒/百万像素）	2.63	2.13	0.96	0.49	0.20	<u>1.09</u>	0.16
结构提取（F-score）	<u>0.577</u>	0.561	0.494	0.510	0.575	0.582	
伪影消除（PSNR）	77.813	70.839	76.738	76.669	<u>77.839</u>	78.052	

5.5 小结

本文提出了一种新的具有尺度感知的图像平滑方法——高斯相关性（RoG）。其作为尺度感知的度量，可以在保持图像主要内容的同时，有效地捕获细节信息。此外，引入可分的递归优化方法有效地减少了求解稀疏线性系统的主要计算负担，实现在 $O(N)$ 时间复杂度下的有效处理。基于 RoG 的图像平滑方法具有很强的普适性，在细节增强、结构提取、伪影消除、色调映射等方面均取得很好效果。

基于经典的空间尺度理论, RoG 是一种普适的尺度度量方法, 扩展其应用场景是未来的工作方向, 如: 物体检测、图像识别、语义分割等。此外需要指出的是, 受限于迭代重加权的求解方法, 基于 RoG 的平滑优化在中间尺度的选择平滑上表现稍有不足。应用交替方向乘子法^[172] (Alternating Direction Method of Multipliers, ADMM) 取代迭代重加权解决 L1 范数求解问题是未来工作的思路。

第六章 总结与展望

6.1 全文总结

雾是一种常见的天气现象，尤其在秋冬季节出现频繁。在恶劣天气下的成像过程中，观测图像受大气散射影响而产生对比度下降、饱和度降低、清晰度变差、动态范围缩小、颜色失真、细节丢失等退化现象。如何减弱或消除雾霾对光学成像系统的影响，提高图像和视频的清晰度与能见度，已经成为机器视觉和图像处理领域的研究热点。本文主要研究内容与取得的研究成果总结如下：

(1) 根据雾天成像的物理模型——大气散射模型，图像去雾的关键是大气透射率的估计。本文提出一种基于深度学习的去雾方法 DehazeNet，实现从有雾图像到大气透射率的端到端估计。DehazeNet 的网络受传统的先验特征启发，采用了“卷积+Maxout”的特征提取层和提出了 BReLU 的非线性回归层。得益于特殊设计的轻巧结构，DehazeNet 保持高效率的同时取得显著的去雾效果。

(2) 视频去雾的关键同样是大气透射率的精确估计，但有别于图像去雾，其更关注于如利用时空信息对透射率进行优化，时空相关性和实时性是其关注的主要问题。本文在大气散射模型的基础上，本文提出了一个基于时空 MRF 的实时视频去雾框架，保证透射率估计的空间一致性和时间连续性。此外，积分图像技术和下采样技术显著地降低了计算复杂度并保证算法的实时性。因此，该方法可在有限的计算资源下实现高效地视频去雾。

(3) 受大气光照的影响，对于部分应用场景下的去雾结果需要对其进行颜色校正的后增强。本文提出了一种基于联合先验的 Retinex 模型，可有效地分解光强度和反射率，并应用于颜色校正。基于本文提出的局部变化偏差的结构度量方法，联合“形状、光照和纹理”先验的 Retinex 模型以改进 Retinex 分解的正则化约束。实验表明，在主观和客观评价中本文方法在 Retinex 分解、颜色校正、亮度自适应等方面均取得优秀的结果。

(4) 在雾天环境下，大气中的悬浮颗粒降低成像的质量和导致图像细节的丢失。本文提出了一种基于高斯相关性 (RoG) 的图像平滑方法，其作为一种尺度感知的度量方法，可以在保持图像主要内容的同时，有效地捕获细节信息。此外，引入可分的递归优化方法有效地减少了求解稀疏线性系统的主要计算负担，实现在 $O(N)$ 时间复杂度下的有效处理。基于 RoG 的图像平滑方法在细节增强、结构提取、伪影消除、色调映射等

方面均取得很好效果。

6.2 学术贡献

本文的部分研究成果已在实际的工程项目中得到应用，已发表的学术成果已经得到同行的引用并作为支持性参考文献。

(1) DehazeNet (第2章) 作为第一个正式发表的基于卷积神经网络的图像去雾方法，受到学界的广泛认可和高度评价。《DehazeNet: An End-to-End System for Single Image Haze Removal》入选 2018 年度 ESI 高被引论文、热点论文 (SCI 前 0.1%)，Google Scholar 累计引用量已超 350 次。在 2018 年的大规模公开去雾评测库 RESIDE^[94] 中，DehazeNet 在 PSNR 和 SSIM 等多项客观评价指标中取得第一，仍优于后续相关的改进工作，如：MSCNN^[97] 和 AOD-Net^[98]。这得益于，DehazeNet 参考去雾先验特征而特殊设计的网络，在保证模型效率的同时获得高精度的重构。

(2) 在第3章中，本文提出的基于时空 MRF 的实时视频去雾方法具有较高的工程应用价值，相关学术论文《Real-time Video Dehazing based on Spatio-temporal MRF》获第 17 届环太平洋多媒体大会 (Pacific-Rim Conference on Multimedia 2016, PCM2016) 大会最佳学生论文。该成果已应用于广东军区某陆军作战部队、广东省某部队的科技装备动员演习中。

(3) 本文第4章中提出基于联合先验的 Retinex 模型。庞培法布拉大学 Marcelo Bertalmio 教授^[173] 指出，本文提出的 Retinex 模型“为 Retinex 图像增强模型提供理论支持”。合肥工业大学 Richang Hong 教授^[174] 以本文 Retinex 分解方法为基础引入反射率的 JPEG 伪影抑制，在压缩的低光图像中取得有效的增强效果。

(4) 本文第5章中提出的基于高斯相对性 (RoG) 的平滑方法获第 24 届 IEEE 国际图像处理大会 (IEEE International Conference on Image Processing 2017, ICIP2017) 大会最佳论文提名。印度 Kerala 大学 Vignesh 等人^[175] 应用局部极值和 RoG 平滑方法实现图像去雾与增强；韩国 Hanwha 集团 Kim^[176] 应用 RoG 作为联合引导滤波器实现红外图像的边缘保持平滑。

6.3 工作展望

本文从基于物理模型的图像去雾出发并扩展到视频去雾，此外从颜色校正和细节增强两方面实现去雾结果的后增强。然而，图像去雾的问题仍是当前的研究热点，后续

工作可从如下几方面进一步展开:

(1) 基于非成对训练的图像去雾。当前基于深度学习的图像重构方法普遍存在训练样本收集困难问题, 很难在雾霾天和晴天采集到同一场景下像素配准的训练样本对。因此, 如何利用跨域非成对样本进行特征学习是深度学习在图像去雾的重点和难点。生成对抗网络^[177] (Generative Adversarial Network, GAN) 为非成对样本训练提供了思路, 但基于分布的监督信号很难保证底层细节的重构精度。结合深度导向滤波器^[178] (Deep Guided Filter) 和循环 GAN^[179] (Cycle-GAN) 有望解决这个问题。

(2) 基于深度学习的视频去雾。深度学习方法普遍存在计算复杂度高的瓶颈, 因此其很难直接应用于实时视频去雾。在卷积神经网络的基础上保证去雾算法的实时性是当前视频去雾的研究重点。近几年, 自适应传播^[180] (Adaptive Propagation) 和记忆引导^[181] (Memory Guided) 通过引入前序帧的特征来辅助后续帧处理, 从而减轻当前帧处理复杂度。因此, 联合快慢网络^[181] (Fast-and-Slow Network) 是应用深度学习于实时视频去雾的突破口。

(3) 联合的后增强方法。颜色校正 (第4章) 和细节增强 (第5章) 是图像去雾普遍采用的后增强手段, 其中 Retinex 方法和非线性平滑方法具有相似的处理过程——边缘/结构保持的平滑优化。因此, 结合颜色校正和细节增强的共同特点, 联合优化光强度和反射率, 同时实现颜色校正和细节增强是本文后续的研究方向。此外, 若使用引导图计算迭代重加权的权值, 非线性平滑方法可扩展为联合滤波器 (Joint Filter) 并应用于大气透射率的优化。

参考文献

- [1] Kim T K, Paik J K, Kang B S. Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering[J]. IEEE Transactions on Consumer Electronics, 1998, 44(1):82–87.
- [2] Kim J Y, Kim L S, Hwang S H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization[J]. IEEE transactions on circuits and systems for video technology, 2001, 11(4):475–484.
- [3] Stark J A. Adaptive image contrast enhancement using generalizations of histogram equalization[J]. IEEE Transactions on image processing, 2000, 9(5):889–896.
- [4] Khan M F, Khan E, Abbasi Z. Segment dependent dynamic multi-histogram equalization for image contrast enhancement[J]. Digital Signal Processing, 2014, 25:198–223.
- [5] Richter R. Atmospheric correction with haze removal including a haze/clear transition region[A]. In: Algorithms for Multispectral and Hyperspectral Imagery II[C], 1996. 2758:254–263.
- [6] Seow M J, Asari V K. Ratio rule and homomorphic filter for enhancement of digital colour image[J]. Neurocomputing, 2006, 69(7-9):954–958.
- [7] Daubechies I. The wavelet transform, time-frequency localization and signal analysis[J]. IEEE transactions on information theory, 1990, 36(5):961–1005.
- [8] Ma J, Plonka G. The curvelet transform[J]. IEEE signal processing magazine, 2010, 27(2):118–133.
- [9] LI G, YANG W n, WENG T. A method of removing thin cloud in remote sensing image based on the homomorphic filter algorithm[J]. Science of Surveying and Mapping, 2007, 3.
- [10] Busch C, Debes E. Wavelet transform for analyzing fog visibility[J]. IEEE Intelligent Systems and Their Applications, 1998, 13(6):66–71.
- [11] Verma M, Kaushik V D, Pathak V K. An efficient deblurring algorithm on foggy images using curvelet transforms[A]. In: Proceedings of the Third International Symposium on Women in Computing and Informatics[C], 2015. 426–431.
- [12] Maloney L T, Wandell B A. Color constancy: a method for recovering surface spectral reflectance[J]. JOSA A, 1986, 3(1):29–33.
- [13] Land E H, McCann J J. Lightness and retinex theory[J]. Josa, 1971, 61(1):1–11.
- [14] Frankle J A, McCann J J. Method and apparatus for lightness imaging[J]. 1983. US Patent 4,384,336.

- [15] Funt B, Ciurea F, McCann J. Retinex in MATLAB[J]. Journal of electronic imaging, 2004, 13(1):48–57.
- [16] Land E H. Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image[J]. Proceedings of the National Academy of Sciences, 1983, 80(16):5163–5169.
- [17] Brainard D H, Wandell B A. Analysis of the retinex theory of color vision[J]. JOSA A, 1986, 3(10):1651–1661.
- [18] Jobson D J, Rahman Z u, Woodell G A. Properties and performance of a center/surround retinex[J]. IEEE transactions on image processing, 1997, 6(3):451–462.
- [19] Jobson D J, Rahman Z u, Woodell G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. IEEE Transactions on Image processing, 1997, 6(7):965–976.
- [20] 芮义斌, 李鹏, 孙锦涛. 一种图像去薄雾方法 [J]. 计算机应用, 2006, 26(1):154–0156.
- [21] Liu Q, Jian B, Yu F. An adaptive weight value-based multi-scale retinex algorithm for color image enhancement[A]. In: International Conference on Computer Sciences & Automation Engineering[C], 2016.
- [22] Rahman Z u, Jobson D J, Woodell G A. Retinex processing for automatic image enhancement[J]. Journal of Electronic Imaging, 2004, 13(1):100–110.
- [23] Land E H. Recent advances in retinex theory[M]//. In: Central and peripheral mechanisms of colour vision[C].[S.l.]: Springer, 1985. 5–17.
- [24] Land E H. An alternative technique for the computation of the designator in the retinex theory of color vision[J]. Proceedings of the national academy of sciences, 1986, 83(10):3078–3080.
- [25] Jobson D J, Rahman Z u, Woodell G A. Properties and performance of a center/surround retinex[J]. IEEE transactions on image processing, 1997, 6(3):451–462.
- [26] Provenzi E, Gatta C, Fierro M, et al. A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast[J]. IEEE transactions on pattern analysis and machine intelligence, 2008, 30(10):1757–1770.
- [27] Xie B, Guo F, Cai Z. Improved single image dehazing using dark channel prior and multi-scale retinex[A]. In: 2010 International Conference on Intelligent System Design and Engineering Application[C], 2010. 1:848–851.

- [28] 嵇晓强. 图像快速去雾与清晰度恢复技术研究 [D]: [博士学位论文]. 长春: 中国科学院研究生院 (长春光学精密机械与物理研究所), 2012.
- [29] Zhang W, Liang J, Ren L, et al. Real-time image haze removal using an aperture-division polarimetric camera[J]. *Applied optics*, 2017, 56(4):942–947.
- [30] Hufnagel R, Stanley N. Modulation transfer function associated with image transmission through turbulent media[J]. *JOSA*, 1964, 54(1):52–61.
- [31] McCartney E J. Optics of the atmosphere: scattering by molecules and particles[J]. New York, John Wiley and Sons, Inc., 1976, 1.
- [32] Oakley J P, Satherley B L. Improving image quality in poor visibility conditions using a physical model for contrast degradation[J]. *IEEE transactions on image processing*, 1998, 7(2):167–179.
- [33] 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.
- [34] Wiscombe W J. Improved Mie scattering algorithms[J]. *Applied optics*, 1980, 19(9):1505–1509.
- [35] Tan R T. Visibility in bad weather from a single image[A]. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]*, 2008. 1–8.
- [36] Fattal R. Single image dehazing[J]. *ACM transactions on graphics (TOG)*, 2008, 27(3):72.
- [37] Kratz L, Nishino K. Factorizing scene albedo and depth from a single foggy image[A]. In: *2009 IEEE 12th International Conference on Computer Vision[C]*, 2009. 1701–1708.
- [38] Nishino K, Kratz L, Lombardi S. Bayesian defogging[J]. *International journal of computer vision*, 2012, 98(3):263–278.
- [39] Carr P, Hartley R. Improved single image dehazing using geometry[A]. In: *2009 Digital Image Computing: Techniques and Applications[C]*, 2009. 103–110.
- [40] He K, Jian S, Tang X. Single image haze removal using dark channel prior[A]. In: *IEEE Conference on Computer Vision Pattern Recognition[C]*, 2009.
- [41] Gibson K, Võ D, Nguyen T. An investigation in dehazing compressed images and video[A]. In: *Oceans 2010 Mts/Ieee Seattle[C]*, 2010. 1–8.
- [42] Yu T, Riaz I, Piao J, et al. Real-time single image dehazing using block-to-pixel interpolation and adaptive dark channel prior[J]. *IET Image Processing*, 2015, 9(9):725–734.
- [43] Bui T, Tran H, Kim W, et al. Segmenting dark channel prior in single image dehazing[J]. *Electronics*

- Letters, 2014, 50(7):516–518.
- [44] Wang J B, He N, Zhang L L, et al. Single image dehazing with a physical model and dark channel prior[J]. Neurocomputing, 2015, 149:718–728.
- [45] Mai J, Zhu Q, Wu D, et al. Back propagation neural network dehazing[J]. IEEE International Conference on Robotics and Biomimetics (ROBIO), 2014:1433–1438.
- [46] Zhu Q, Mai J, Shao L. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior[J]. IEEE Transactions on Image Processing, 2015, 24(11):3522–3533.
- [47] Tang K, Yang J, Wang J. Investigating haze-relevant features in a learning framework for image dehazing[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2014. 2995–3002.
- [48] Oakley J P, Satherley B L. Improving image quality in poor visibility conditions using a physical model for contrast degradation[J]. IEEE transactions on image processing, 1998, 7(2):167–179.
- [49] Narasimhan S G, Nayar S K. Interactive (de) weathering of an image using physical models[A]. In: IEEE Workshop on color and photometric Methods in computer Vision[C], 2003. 6:1.
- [50] Kopf J, Neubert B, Chen B, et al. Deep photo: Model-based photograph enhancement and viewing[A]. In: ACM Transactions on Graphics (TOG)[C], 2008. 27:116.
- [51] Lin Z, Wang X. Dehazing for image and video using guided filter[J]. Appl. Sci, 2012, 2:123–127.
- [52] Yeh C H, Kang L W, Lin C Y, et al. Efficient image/video dehazing through haze density analysis based on pixel-based dark channel prior[A]. In: 2012 International Conference on Information Security and Intelligent Control[C], 2012. 238–241.
- [53] Kumari A, Sahoo S K. Real time image and video deweathering: The future prospects and possibilities[J]. Optik-International Journal for Light and Electron Optics, 2016, 127(2):829–839.
- [54] Xu Z, Liu X, Chen X. Fog removal from video sequences using contrast limited adaptive histogram equalization[A]. In: 2009 International Conference on Computational Intelligence and Software Engineering[C], 2009. 1–4.
- [55] John J, Wilscy M. Enhancement of weather degraded video sequences using wavelet fusion[A]. In: 2008 7th IEEE International Conference on Cybernetic Intelligent Systems[C], 2008. 1–6.
- [56] Archa S, Abdul A. A novel method for video dehazing by multi-scale fusion[J]. Int. J. Sci. Eng. Technol. Res., 2014, 3:4808–4813.

- [57] Tarel J P, Hautiere N, Cord A, et al. Improved visibility of road scene images under heterogeneous fog[A]. In: Intelligent Vehicles Symposium (IV), 2010 IEEE[C], 2010. 478–485.
- [58] Tarel J P, Hautiere N. Fast visibility restoration from a single color or gray level image[A]. In: IEEE International Conference on Computer Vision[C], 2009. 2201–2208.
- [59] Zhang J, Li L, Zhang Y, et al. Video dehazing with spatial and temporal coherence[J]. The Visual Computer, 2011, 27(6-8):749–757.
- [60] Kim J H, Jang W D, Sim J Y, et al. Optimized contrast enhancement for real-time image and video dehazing[J]. Journal of Visual Communication and Image Representation, 2013, 24(3):410–425.
- [61] Li Z, Tan P, Tan R T, et al. Simultaneous video defogging and stereo reconstruction[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2015. 4988–4997.
- [62] Meng G, Wang Y, Duan J, et al. Efficient image dehazing with boundary constraint and contextual regularization[A]. In: IEEE International Conference on Computer Vision (ICCV)[C], 2013. 617–624.
- [63] Gibson K, Vo D, Nguyen T. An investigation in dehazing compressed images and video[A]. In: OCEANS 2010[C], 2010. 1–8.
- [64] Tomasi C, Manduchi R. Bilateral filtering for gray and color images[A]. In: Computer Vision, 1998. Sixth International Conference on[C], 1998. 839–846.
- [65] He K, Sun J, Tang X. Guided image filtering[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(6):1397–1409.
- [66] Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting[J]. IEEE transactions on pattern analysis and machine intelligence, 2007, 30(2):228–242.
- [67] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[A]. In: Advances in neural information processing systems[C], 2012. 1097–1105.
- [68] Ma B, Su Y, Jurie F. Covariance descriptor based on bio-inspired features for person re-identification and face verification[J]. Image and Vision Computing, 2014, 32(6):379–390.
- [69] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C], 2014. 580–587.
- [70] Schuler C J, Burger H C, Harmeling S, et al. A machine learning approach for non-blind image decon-

- volution[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2013. 1067–1074.
- [71] Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[M]//. In: Computer Vision–ECCV 2014[C].[S.l.]: Springer, 2014. 184–199.
- [72] Eigen D, Krishnan D, Fergus R. Restoring an image taken through a window covered with dirt or rain[A]. In: IEEE International Conference on Computer Vision (ICCV)[C], 2013. 633–640.
- [73] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[A]. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10)[C], 2010. 807–814.
- [74] 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.
- [75] Ancuti C O, Ancuti C, Hermans C, et al. A fast semi-inverse approach to detect and remove the haze from a single image[M]//. In: Computer Vision–ACCV 2010[C].[S.l.]: [s.n.] , 2011. 501–514.
- [76] Goodfellow I, Warde-farley D, Mirza M, et al. Maxout Networks[A]. In: Proceedings of the 30th International Conference on Machine Learning (ICML-13)[C], 2013. 1319–1327.
- [77] Jia X, Xu X, Cai B, et al. Single Image Super-Resolution Using Multi-Scale Convolutional Neural Network[A]. In: Pacific Rim Conference on Multimedia[C], 2017. 149–157.
- [78] Szegedy C, Liu W, Jia Y, et al. Going Deeper With Convolutions[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2015. 1–9.
- [79] Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2014:1–42.
- [80] Lampl I, Ferster D, Poggio T, et al. Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex[J]. Journal of neurophysiology, 2004, 92(5):2704–2713.
- [81] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278–2324.
- [82] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504–507.
- [83] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[A]. In: Proceedings of the ACM International Conference on Multimedia[C], 2014. 675–678.

- [84] Sulami M, Glatzer I, Fattal R, et al. Automatic recovery of the atmospheric light in hazy images[A]. In: IEEE International Conference on Computational Photography (ICCP)[C], 2014. 1–11.
- [85] He K, Sun J, Tang X. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6):1397–1409.
- [86] Dong C, Loy C C, He K, et al. Image Super-Resolution Using Deep Convolutional Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015(99):1–1.
- [87] Chan T H, Jia K, Gao S, et al. PCANet: A Simple Deep Learning Baseline for Image Classification? [J]. IEEE Transactions on Image Processing, 2015, 24(12):5017–5032.
- [88] Scholkopf B, Mullert K R. Fisher discriminant analysis with kernels[J]. Neural networks for signal processing IX, 1999, 1:1.
- [89] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. International journal of computer vision, 2002, 47(1-3):7–42.
- [90] Scharstein D, Szeliski R. High-accuracy stereo depth maps using structured light[A]. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C], 2003. 1:I–195.
- [91] Scharstein D, Pal C. Learning conditional random fields for stereo[A]. In: IEEE Conference on Computer Vision and Pattern Recognition[C], 2007. 1–8.
- [92] 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.
- [93] Mannos J L, Sakrison D J. The effects of a visual fidelity criterion of the encoding of images[J]. IEEE Transactions on Information Theory, 1974, 20(4):525–536.
- [94] Li B, Ren W, Fu D, et al. Benchmarking single-image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2019, 28(1):492–505.
- [95] Silberman N, Hoiem D, Kohli P, et al. Indoor segmentation and support inference from rgbd images[A]. In: European Conference on Computer Vision[C], 2012. 746–760.
- [96] Berman D, Avidan S, et al. Non-local image dehazing[A]. In: Proceedings of the IEEE conference on computer vision and pattern recognition[C], 2016. 1674–1682.
- [97] Ren W, Liu S, Zhang H, et al. Single image dehazing via multi-scale convolutional neural networks[A]. In: European conference on computer vision[C], 2016. 154–169.
- [98] Li B, Peng X, Wang Z, et al. Aod-net: All-in-one dehazing network[A]. In: Proceedings of the IEEE

- International Conference on Computer Vision[C], 2017. 4770–4778.
- [99] Ling Z, Fan G, Wang Y, et al. Learning deep transmission network for single image dehazing[A]. In: 2016 IEEE International Conference on Image Processing (ICIP)[C], 2016. 2296–2300.
- [100] Schaul L, Fredembach C, Süsstrunk S. Color image dehazing using the near-infrared[A]. In: IEEE International Conference on Image Processing (ICIP)[C], 2009. 1629–1632.
- [101] Schechner Y Y, Narasimhan S G, Nayar S K. Instant dehazing of images using polarization[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2001. 1:I–325.
- [102] Narasimhan S G, Nayar S K. Chromatic framework for vision in bad weather[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2000. 1:598–605.
- [103] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[A]. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C], 2001. 1:I–511.
- [104] Zhang G, Jia J, Wong T T, et al. Consistent depth maps recovery from a video sequence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(6):974–988.
- [105] Fu X, Zeng D, Huang Y, et al. A weighted variational model for simultaneous reflectance and illumination estimation[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2016. 2782–2790.
- [106] Fu X, Liao Y, Zeng D, et al. A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation[J]. IEEE Transactions on Image Processing, 2015, 24(12):4965–4977.
- [107] Fu X, Zeng D, Huang Y, et al. A fusion-based enhancing method for weakly illuminated images[J]. Signal Processing, 2016, 129:82–96.
- [108] Horn B K. Determining lightness from an image[J]. Computer graphics and image processing, 1974, 3(4):277–299.
- [109] Morel J M, Petro A B, Sbert C. A PDE formalization of retinex theory[J]. IEEE Transactions on Image Processing, 2010, 19(11):2825–2837.
- [110] Kimmel R, Elad M, Shaked D, et al. A variational framework for retinex[J]. International Journal of computer vision, 2003, 52(1):7–23.
- [111] Ng M K, Wang W. A total variation model for Retinex[J]. SIAM Journal on Imaging Sciences, 2011, 4(1):345–365.

- [112] Fischler M A. Recovering intrinsic scene characteristics from images[R].[S.l.]: DTIC Document, 1982.
- [113] Olkkonen M, Hansen T, Gegenfurtner K R. Color appearance of familiar objects: Effects of object shape, texture, and illumination changes[J]. *Journal of Vision*, 2008, 8(5):13–13.
- [114] Barrett R, Berry M, Chan T F, et al. Templates for the solution of linear systems: building blocks for iterative methods[M].[S.l.]: SIAM, 1994.
- [115] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion[J]. *IEEE Transactions on pattern analysis and machine intelligence*, 1990, 12(7):629–639.
- [116] Zhang Q, Shen X, Xu L, et al. Rolling Guidance Filter[A]. In: *European Conference on Computer Vision[C]*, 2014. 815–830.
- [117] Zhang Q, Xu L, Jia J. 100+ Times Faster Weighted Median Filter (WMF)[A]. In: *IEEE Conference on Computer Vision and Pattern Recognition[C]*, 2014. 2830–2837.
- [118] Subr K, Soler C, Durand F. Edge-preserving multiscale image decomposition based on local extrema[J]. *ACM Transactions on Graphics (TOG)*, 2009, 28(5):147.
- [119] Karacan L, Erdem E, Erdem A. Structure-preserving image smoothing via region covariances[J]. *ACM Transactions on Graphics (TOG)*, 2013, 32(6):176.
- [120] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D: Nonlinear Phenomena*, 1992, 60(1-4):259–268.
- [121] Farbman Z, Fattal R, Lischinski D, et al. Edge-preserving decompositions for multi-scale tone and detail manipulation[A]. In: *ACM Transactions on Graphics (TOG)[C]*, 2008. 27:67.
- [122] Xu L, Yan Q, Xia Y, et al. Structure extraction from texture via relative total variation[J]. *ACM Transactions on Graphics (TOG)*, 2012, 31(6):139.
- [123] Dong X, Wang G, Pang Y, et al. Fast efficient algorithm for enhancement of low lighting video[A]. In: *Multimedia and Expo (ICME), 2011 IEEE International Conference on[C]*, 2011. 1–6.
- [124] Wang L, Xiao L, Liu H, et al. Variational Bayesian method for retinex[J]. *IEEE Transactions on Image Processing*, 2014, 23(8):3381–3396.
- [125] Tseng P. Convergence of a block coordinate descent method for nondifferentiable minimization[J]. *Journal of optimization theory and applications*, 2001, 109(3):475–494.
- [126] Candes E J, Wakin M B, Boyd S P. Enhancing sparsity by reweighted L1 minimization[J]. *Journal of*

- Fourier analysis and applications, 2008, 14(5-6):877–905.
- [127] Krishnan D, Fattal R, Szeliski R. Efficient preconditioning of laplacian matrices for computer graphics[J]. ACM Transactions on Graphics (TOG), 2013, 32(4):142.
- [128] Zhang X, Wandell B A. A spatial extension of CIELAB for digital color-image reproduction[J]. Journal of the Society for Information Display, 1997, 5(1):61–63.
- [129] Gehler P V, Rother C, Blake A, et al. Bayesian color constancy revisited[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2008. 1–8.
- [130] Brainard D H, Wandell B A. Analysis of the retinex theory of color vision[J]. JOSA A, 1986, 3(10):1651–1661.
- [131] Buchsbaum G. A spatial processor model for object colour perception[J]. Journal of the Franklin institute, 1980, 310(1):1–26.
- [132] Van De Weijer J, Gevers T, Gijsenij A. Edge-based color constancy[J]. IEEE Transactions on image processing, 2007, 16(9):2207–2214.
- [133] Finlayson G D, Trezzi E. Shades of gray and colour constancy[A]. In: Color and Imaging Conference[C], 2004. 2004:37–41.
- [134] Bianco S, Cusano C, Schettini R. Color constancy using CNNs[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops[C], 2015. 81–89.
- [135] Barnard K, Cardei V, Funt B. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data[J]. IEEE transactions on Image Processing, 2002, 11(9):972–984.
- [136] Yang K F, Gao S B, Li Y J. Efficient illuminant estimation for color constancy using grey pixels[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2015. 2254–2263.
- [137] Wang S, Zheng J, Hu H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images[J]. IEEE Transactions on Image Processing, 2013, 22(9):3538–3548.
- [138] Shan Q, Jia J, Brown M S. Globally optimized linear windowed tone mapping[J]. IEEE transactions on visualization and computer graphics, 2010, 16(4):663–675.
- [139] Guo X, Li Y, Ling H. LIME: Low-light Image Enhancement via Illumination Map Estimation[J]. IEEE Transactions on Image Processing, 2016, 26(2):982–993.

- [140] Cheng H, Shi X. A simple and effective histogram equalization approach to image enhancement[J]. Digital Signal Processing, 2004, 14(2):158–170.
- [141] Sheet D, Garud H, Suveer A, et al. Brightness preserving dynamic fuzzy histogram equalization[J]. IEEE Transactions on Consumer Electronics, 2010, 56(4).
- [142] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3):209–212.
- [143] Gu K, Zhai G, Lin W, et al. No-reference image sharpness assessment in autoregressive parameter space[J]. IEEE Transactions on Image Processing, 2015, 24(10):3218–3231.
- [144] Min D, Choi S, Lu J, et al. Fast global image smoothing based on weighted least squares[J]. IEEE Transactions on Image Processing, 2014, 23(12):5638–5653.
- [145] Zhou T, Tao D. Godec: Randomized low-rank & sparse matrix decomposition in noisy case[A]. In: International conference on machine learning[C], 2011.
- [146] Hao S, Pan D, Guo Y, et al. Image detail enhancement with spatially guided filters[J]. Signal Processing, 2016, 120:789–796.
- [147] Li Y, Huang J B, Ahuja N, et al. Deep joint image filtering[A]. In: European Conference on Computer Vision[C], 2016. 154–169.
- [148] Xu L, Lu C, Xu Y, et al. Image smoothing via L 0 gradient minimization[A]. In: ACM Transactions on Graphics (TOG)[C], 2011. 30:174.
- [149] Tumblin J, Turk G. LCIS: A boundary hierarchy for detail-preserving contrast reduction[A]. In: Proceedings of the 26th annual conference on Computer graphics and interactive techniques[C], 1999. 83–90.
- [150] Li Z, Zheng J, Zhu Z, et al. Weighted guided image filtering[J]. IEEE Transactions on Image Processing, 2015, 24(1):120–129.
- [151] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(5):898–916.
- [152] Felzenszwalb P F, Girshick R B, McAllester D, et al. Object detection with discriminatively trained part-based models[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(9): 1627–1645.
- [153] Liao S, Zhu X, Lei Z, et al. Learning multi-scale block local binary patterns for face recognition[J].

- Advances in biometrics, 2007:828–837.
- [154] Burt P, Adelson E. The Laplacian pyramid as a compact image code[J]. IEEE Transactions on communications, 1983, 31(4):532–540.
- [155] Witkin A. Scale-space filtering: A new approach to multi-scale description[A]. In: Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.[C], 1984. 9:150–153.
- [156] Lindeberg T. Scale-space theory: a basic tool for analyzing structures at different scales[J]. Journal of Applied Statistics, 1994, 21(1-2):225–270.
- [157] Sporring J, Nielsen M, Florack L, et al. Gaussian Scale-Space Theory[M].[S.l.]: Kluwer Academic Publishers, 1997.
- [158] Lowe D G. Object recognition from local scale-invariant features[A]. In: Computer vision, 1999. The proceedings of the seventh IEEE international conference on[C], 1999. 2:1150–1157.
- [159] Lindeberg T. Scale-space theory: A basic tool for analyzing structures at different scales[J]. Journal of applied statistics, 1994, 21(1-2):225–270.
- [160] Mikolajczyk K, Schmid C. An affine invariant interest point detector[A]. In: European conference on computer vision[C], 2002. 128–142.
- [161] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60(2):91–110.
- [162] Schreiber W F. Fundamentals of electronic imaging systems: some aspects of image processing[M]. Vol. 15.[S.l.]: Springer Science & Business Media, 2012.
- [163] Golub G H, Van Loan C F. Matrix computations[M]. Vol. 3.[S.l.]: JHU Press, 2012.
- [164] Paris S, Hasinoff S W, Kautz J. Local Laplacian filters: Edge-aware image processing with a Laplacian pyramid.[J]. ACM Trans. Graph., 2011, 30(4):68.
- [165] Cho H, Lee H, Kang H, et al. Bilateral texture filtering[J]. ACM Transactions on Graphics (TOG), 2014, 33(4):128.
- [166] Canny J. A computational approach to edge detection[J]. IEEE Transactions on pattern analysis and machine intelligence, 1986(6):679–698.
- [167] Ro D, Pe H. Pattern classification and scene analysis[J]. 1973.
- [168] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[A]. In: Computer Vision,

2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on[C], 2001. 2:416–423.
- [169] Dabov K, Foi A, Katkovnik V, et al. Image denoising with block-matching and 3D filtering[A]. In: Proceedings of SPIE[C], 2006. 6064:606414–606414.
- [170] Wang G, Wong T T, Heng P A. Deringing cartoons by image analogies[J]. ACM Transactions on Graphics (TOG), 2006, 25(4):1360–1379.
- [171] Reinhard E, Stark M, Shirley P, et al. Photographic tone reproduction for digital images[J]. ACM transactions on graphics (TOG), 2002, 21(3):267–276.
- [172] Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends® in Machine learning, 2011, 3(1):1–122.
- [173] Galdran A, Alvarez-Gila A, Bria A, et al. On the duality between retinex and image dehazing[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2018. 8212–8221.
- [174] Xu C, Hao S, Guo Y, et al. Enhancing Low-Light Images with JPEG Artifact Based on Image Decomposition[A]. In: Pacific Rim Conference on Multimedia[C], 2018. 3–12.
- [175] Vignesh R, Simon P. Single Image Defogging Based on Local Extrema and Relativity of Gaussian[M]//. In: Harmony Search and Nature Inspired Optimization Algorithms[C].[S.l.]: Springer, 2019. 431–439.
- [176] Kim I H. Edge Preserving Smoothing in Infrared Image using Relativity of Guided Filter[J]. Journal of the Korea Society of Computer and Information, 2018, 23(12):27–33.
- [177] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[A]. In: Advances in neural information processing systems[C], 2014. 2672–2680.
- [178] Wu H, Zheng S, Zhang J, et al. Fast end-to-end trainable guided filter[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2018. 1838–1847.
- [179] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[A]. In: Proceedings of the IEEE International Conference on Computer Vision[C], 2017. 2223–2232.
- [180] Li Y, Shi J, Lin D. Low-latency video semantic segmentation[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], 2018. 5997–6005.
- [181] Liu M, Zhu M, White M, et al. Looking Fast and Slow: Memory-Guided Mobile Video Object Detection[J]. arXiv preprint arXiv:1903.10172, 2019.

攻读博士学位期间取得的研究成果

已发表（包括已接受待发表）的论文，以及已投稿、或已成文打算投稿、或拟成文投稿的论文情况（只填写与学位论文内容相关的部分）：

序号	作者（按顺序排列）	题目	发表或投稿刊物、级别	发表的卷期、年月、页码	对应学位论文（章）	索引收录情况
1	Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, Dacheng Tao	DehazeNet: An End-to-End System for Single Image Haze Removal	IEEE Transactions on Image Processing (TIP)	25(11), 2016, 5187-5198	第二章	SCI
2	Bolun Cai, Xiangmin Xu, Dacheng Tao	Real-time Video Dehazing based on Spatio-temporal MRF	Pacific-Rim Conference on Multimedia (PCM)	17(01), 2016, 315-325	第三章	EI
3	Bolun Cai, Xiangmin Xu, Kailing Guo, Kui Jia, Bin Hu, Dacheng Tao	A Joint Intrinsic-Extrinsic Prior Model for Retinex	IEEE International Conference on Computer Vision (ICCV)	16(01), 2016, 4020-4029	第四章	EI
4	Bolun Cai, Xiaofen Xing, Xiangmin Xu	Edge/Structure Preserving Smoothing via Relativity-of-Gaussian	IEEE International Conference on Image Processing (ICIP)	24(01), 2017, 250-254	第五章	EI

致谢

华园的木棉花开花落，西湖的钟声日复一日，光阴荏苒，白驹过隙，十年求学，犹如昨日。值此拙文完成之际，回首在华南理工大学的点点滴滴，心中颇为感激，衷心地向各位老师、亲友和同学致以诚挚的敬意和谢意。

首先，衷心地感谢我的导师徐向民教授多年来对我的关心和帮助，对工作的指导和支持，言传身教，润物细无声。同时，感谢邢晓芬老师、贾奎老师对我工作上的指导和生活上的帮助。此外，特别地感谢陶大程教授在我科研上的传业授道解惑，其精湛的学术造诣和严谨的治学态度指引着我的科研道路。

其次，感谢实验室的同学伙伴、师兄师姐、师弟师妹，从3号楼到30号楼，从3205到626，从NGHCI到BDS，各自奔前程的身影渐行渐远，离别却不相忘于江湖。特别地感谢一起奋斗过的博士们，感谢苗总、郭博、索哥、刘芳、王琳在学术上的交流和工作上的帮助。回想起那些赶论文、写本子、叫外卖、挑灯夜战的日子，你们的陪伴让漫漫的科研路不再孤单。

最后，感谢家人朋友对我无微不至的关怀。感谢父母的养育之恩，关心我，照顾我，鼓励我，鞭策我，您们的支持和期许是前进的帆和灯塔。感谢兄长蔡博昆二十九年的陪伴和我学业的支持，看着那些可爱童年和青葱岁月变成一张张相片，懂得珍惜已过了多年，如今相聚的时光越来越少，海阔天高任鸟飞。此外，特别感谢徐桂从的理解和支持，磐石薄苇，金坚壁白。

云山苍苍，珠水泱泱，无意地翻看从前的相片，才不知不觉学会了怀念。北区的山坡建了博学楼，东区的空地修了喷水池，十年前的华园在记忆中渐渐模糊；西湖的鱼还很多，华园的路还在修，这里的故事还在继续。曾经的少年羡慕南飞的雁，岁月改变了当初的模样，放下可乐雪碧拿起枸杞保温杯，褪下一身稚气穿上了格子衫，留起偏分头遮住了发际线。而立之年，前路漫漫，背上行囊，全新出发。

蔡博仑

2019年4月9日

3. 答辩委员会对论文的评语

(主要内容包括: 1.对论文的综合评价; 2.对论文主要工作和创造性成果的简要介绍; 3.对作者掌握基础理论、专业知识程度、独立从事科研工作能力以及在答辩中表现的评价; 4.存在的不足之处和建议; 5.答辩委员会结论意见等)

图像去雾增强在计算机视觉系统中属于低层次图像处理过程, 是进行高层次视觉感知任务的基础, 面临大气透射率估计难、颜色易失真、图像对比度下降等挑战, 属于信息缺失的病态优化求解问题。本文从以下三个方面开展了创新研究:

1. 开展了基于大气透射率估计的图像去雾方法研究。提出了一种基于深度学习的去雾方法, 实现从有雾图像到大气透射率的端到端估计, 通过估计大气散射模型可以重构得到清晰图像, 其中特征提取层的“Maxout+卷积”可近似于现有雾相关的传统特征, 并在非线性回归层中提出了一种新的非线性激活函数 BReLU 用于提高清晰图像的重构质量。

2. 开展了实时的视频去雾方法研究。通过建立一个基于亮度值先验的马尔可夫随机场 (MRF) 来处理视频图像的空间一致性和时间连续性, 并最大化 MRF 的似然函数实现雾浓度图的估计和优化, 可有效减少区块效应和闪烁效应, 实现视频图像的有效增强和质量改善。此外, 采用积分图和下采样技术降低了计算复杂度。

3. 开展了复杂光照下图像去雾的颜色校正方法研究。提出了一种基于联合先验的 Retinex 模型, 可分解光强度和反射率, 实现图像的颜色校正。同时, 开展了图像细节信息增强的方法研究, 提出了一种基于高斯相关性的图像平滑方法, 可在保持图像主要结构信息的同时, 有效地捕获细节信息。

蔡博仑同学的论文工作量饱满, 格式规范, 层次清楚, 分析合理, 论证充分, 是一篇优秀的博士论文。在答辩过程中, 能清晰阐述博士论文工作, 清楚回答专家的提问。该生掌握了扎实的图像处理、计算机视觉和深度学习等相关研究领域基础理论, 并进行了系统性地深入研究, 具备独立从事相关研究领域科研工作的能力。相关研究成果已发表在国际知名的学术期刊和学术会议上。

答辩委员会经讨论表决, 认为论文达到了博士学位论文水平, 一致同意通过蔡博仑同学的博士学位论文答辩, 建议授予工学博士学位。

论文答辩日期: 2019 年 5 月 30 日

答辩委员会委员 5 人

表决票数: 同意毕业及授予学位 (5) 票;

同意毕业, 但不同意授予学位 () 票;

不同意毕业 () 票

表决结果 (打“√”): 通过 (√); 不通过 ()

决议: 同意授予博士学位 (√) 不同意授予博士学位 ()

答辩
委员
会成
员
签
名

蔡博仑

(主席)

赵建

殷瑞祥

杨智

徐心

答
辩
秘
书
签
名

邵明