

硕士学位论文

基于场景图像的空气品质等级估计

Scene Image Based Air Quality Level Estimation

学 科 专 业 控制科学与工程

学 科 方 向 模式识别与智能系统

作 者 姓 名 李沁

指 导 教 师 谢斌 副教授

中 南 大 学

2018 年 5 月

中图分类号 TP391.4
UDC 621.3

学校代码 10533
学位类别 学术学位

硕士学位论文

基于场景图像的空气品质等级估计

Scene Image Based Air Quality Level Estimation

作者姓名: 李沁
学科专业: 控制科学与工程
学科方向: 模式识别与智能系统
研究方向: 图像处理与模式识别
二级培养单位: 信息科学与工程学院
指导教师: 谢斌 副教授

论文答辩日期_____

答辩委员会主席_____

中 南 大 学

2018 年 5 月

学位论文原创性声明

本人郑重声明，所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了论文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中南大学或其他教育机构的学位或证书而使用过的材料。与我共同工作的同志对本研究所作的贡献均已在论文中作了明确的说明。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

作者签名：_____ 日期：_____年____月____日

学位论文版权使用授权书

本学位论文作者和指导教师完全了解中南大学有关保留、使用学位论文的规定：即学校有权保留并向国家有关部门或机构送交学位论文的复印件和电子版；本人允许本学位论文被查阅和借阅；学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用复印、缩印或其它手段保存和汇编本学位论文。

保密论文待解密后适应本声明。

作者签名：_____

导师签名_____

日期：_____年____月____日

日期：_____年____月____日

基于场景图像的空气质量等级估计

摘要：针对现有空气质量观测技术存在的问题，本文提出利用相机拍摄场景获取图像信息，并利用这一类信息学习有关场景空气质量的特征表示实现对空气质量等级的估计。具体过程如下：

（1）基于视觉特征的空气质量等级估计。视觉特征描述了场景图像的表面性质。本文引入 6 类视觉特征，包括 SIFT 特征、HOG 特征、LBP 特征、色彩饱和度、颜色直方图以及对比度。为了判断不同特征对空气质量估计的准确率，本文搭建了一个标注空气质量等级的场景图像数据库 VAQI-1。VAQI-1 为目前最大且唯一公开用于空气质量等级估计的场景图像数据库。基于 VAQI-1，本文分别提取上述 6 类特征并采用支持向量机（SVM）对空气质量等级进行估计。实验结果表明基于视觉特征的估计结果普遍不理想，准确率最高仅达 42%。这表明视觉特征无法对空气质量进行描述。

（2）基于能见度特征的空气质量等级估计。能见度描述了场景最大可视范围，可直观表征空气质量。本文提出一种仅用一张场景图像计算能见度的方法。该方法准确率较高且有较强的通用性。基于此方法，本文提取 VAQI-1 图像能见度特征并估计空气质量等级。结果表明能见度特征表现要优于视觉特征，准确率达 58%。然而极端天气下仅用能见度特征仍无法描述空气质量等级。因此，本文将能见度特征与视觉特征进行组合估计空气质量等级。实验结果表明能见度特征与色彩饱和度的组合分类准确率达 65%。这一结果证明能见度特征与色彩饱和度能基本表征场景空气质量。

（3）基于深度特征的空气质量估计。本文采用卷积神经网络提取场景图像深度特征，用到的模型有 LeNet 模型、AlexNet 模型、VGG 模型、GoogLeNet 模型和 ResNet 模型。基于 VAQI-1，本文分别利用上述模型获取深度特征，并利用 SVM 分类器估计空气质量等级。实验结果表明基于 ImageNet 预训练后的 GoogLeNet 模型得到的深度特征估计结果达 76%，高于视觉特征和能见度特征的估计结果。这一结果表明这类深度特征最能够表征空气质量成像机理，同时也证明了基于深度特征的空气质量等级估计方法有望成为现有场景空气质量观测技术的重要补充。

图 33 幅，表 9 个，参考文献 67 篇

关键词：空气质量；视觉特征；能见度特征；深度特征
分类号：TP391.4

Scene Image Based Air Quality Level Estimation

Abstract: With the aim to solve the problems of existing air quality measuring methods, this paper intends to estimate air quality using image data. The specific process is as follows:

(1) Estimate air quality level using the visual feature. The visual feature describes basic characteristics of an observed scene. This paper introduces 6 visual features - SIFT, HOG, LBP, saturation, color histogram and contrast. In order to judge the estimation accuracy of different features, a high quality image database named VAQI-1 is built in this paper, which contains 4444 images of various types of scenes. As far as we know, VAQI-1 is the largest database used for air quality estimation, and is the only one available in public. By using VAQI-1, this paper estimates the air quality level of the corresponding image based on above 6 visual features separately. The classifier we used is the multi-class support vector machine (multi-SVM). Experiments show that visual features can not obtain satisfying results.

(2) Estimate air quality level using the visibility feature. Visibility describes the furthest visual range in an observed scene, which directly reflects the atmosphere turbidity. Based on this, this paper develops a novel method to estimate the scene visibility using a single image as input. Experiments show that this method performs well without additional techniques. Therefore, this paper utilizes the visibility feature to estimate the air quality level based on VAQI-1. Results show that visibility feature performs better than the visual feature, which is up to 58%. Considering the limitation of visibility to describe the air quality in extreme weather conditions, this paper combines visibility and above 6 visual features separately to estimate the air quality level. Results show that the combination of visibility and color histogram performs best and the accuracy is 65%, which demonstrates that visibility and color feature can basically describe the scene air quality.

(3) Estimate air quality level by the deep feature extracted by the convolution neural network (CNN). The CNN models used in this paper are LeNet, AlexNet, VGG, GoogLeNet and ResNet. The classifier we used is softmax classifier. Experiments show that the accuracy obtained by fine-tuned GoogLeNet is 76%, which indicates that the deep feature is better than the visual feature and the visibility feature. This significant result not only shows

that the deep feature characterizes the air quality best, but also proves that air quality level estimation based on the deep feature is likely to be an important supplement in the field of air quality measuring technology.

Keywords: air quality; visual feature; visibility feature; deep feature

Classification: TP391.4

目 录

摘要.....	I
Abstract.....	II
1 绪论.....	3
1.1 研究背景.....	3
1.2 国内外研究现状.....	6
1.3 研究目的和内容.....	9
1.3.1 研究目的.....	9
1.3.2 研究内容.....	9
1.4 本章小结.....	9
2 基于视觉特征的空气质量等级估计.....	11
2.1 视觉特征.....	11
2.1.1 SIFT 特征.....	11
2.1.2 HOG 特征.....	12
2.1.3 LBP 特征.....	13
2.1.4 色彩饱和度.....	13
2.1.5 颜色直方图.....	15
2.1.6 对比度.....	15
2.2 实验设计与结果分析.....	17
2.2.1 实验方案设计.....	17
2.2.2 VAQI-1 数据库.....	19
2.2.3 结果分析.....	22
2.3 本章小结.....	22
3 基于能见度特征的空气质量等级估计.....	24
3.1 能见度特征.....	24
3.1.1 计算 β_0	25
3.1.2 计算 β/β_0	28
3.1.3 计算能见度 V_{met}	31
3.2 实验设计与结果分析.....	31
3.2.1 能见度估计结果分析.....	31
3.2.2 空气质量等级估计结果分析.....	33
3.3 本章小结.....	35

4 基于深度特征的空气质量等级估计.....	36
4.1 卷积神经网络.....	36
4.1.1 卷积神经网络理论基础.....	36
4.1.2 LeNet 模型.....	38
4.1.3 AlexNet 模型.....	39
4.1.4 VGG 模型.....	41
4.1.5 GoogLeNet 模型.....	42
4.1.6 ResNet 模型.....	44
4.2 实验设计与结果分析.....	45
4.3 本章小结.....	48
5 总结.....	49
参考文献.....	51
攻读学位期间主要的研究成果目录.....	56
致谢.....	57

1 绪论

1.1 研究背景

近三十年来,中国经济经历了高速增长,民众对生活质量的的要求不断提高,环境污染问题正日益成为全社会共同关注的一个热点话题。其中,空气污染给人们的影响尤为严重。

空气污染是指空气中大气污染物(如PM_{2.5}、PM₁₀、二氧化硫、二氧化氮、一氧化碳等)浓度超过环境可接受的极限程度,导致生态系统以及人类的生存环境遭到破坏,阻碍了人类社会的正常生存和发展。这些大气污染物主要来自工厂、汽车、家用燃煤炉等排出的废烟废气。空气污染对人类和环境的影响主要分为以下3个方面:

(1) 空气污染对人体和健康的影响。大气污染物会使人体呼吸系统受损、生理机能障碍。当污染物的浓度很高时,人体易造成急性污染中毒,严重者病状恶化,甚至失去生命。即使在污染物浓度不高的时候,长时间呼吸也会引起慢性支气管炎、支气管哮喘等疾病,严重者甚至致癌;

(2) 空气污染不仅会降低生物对病虫害的抗御能力,减缓生物的正常发育,还使生物中毒甚至死亡。生长期中的植物长期接触空气污染会损伤叶面,减弱光合作用,使植物枯萎直至死亡。在各类污染气体中,氯气、氟化氢和二氧化硫等对植物的危害最大。空气污染对动物的损害主要包括呼吸道感染和因食用被污染食物导致的胃道腐蚀,严重者甚至死亡。

(3) 空气污染对大气环境的影响。空气污染发展至今已超越国界,其危害遍及全球。首先,雨、雪等在形成和降落过程中吸收并溶解了空气中的二氧化硫、氮氧化物等污染颗粒形成了酸雨。酸雨对环境的危害是巨大的,不仅能使土壤酸化,危害农作物,还对仪器设备、建筑物有腐蚀作用。人类若长期生活在酸雨中,则会生成呼吸器官衰竭、肾病、癌症等一系列疾病。其次,空气中的二氧化碳等温室气体吸收地面辐射中的红外线导致全球温度上升,即温室效应。温室效应会使全球降水量重新分配,造成冰川和冻土消融、海平面上升等一系列严重的自然灾害,对人类生存构成巨大的威胁。

如何对空气污染进行治理是目前亟待解决的问题。而对空气质量进行有效监测是治理空气污染的重要前提。目前,空气质量监测主要依靠空气站,即空气质量监测站。空气站的功能是对空气中的污染粒子进行连续或定时采样,并通过技术手段进行测量和分析。然而,国内几乎所有城市仅有若干个空气站站点,无法覆盖整个城市,这使得空气站测量数据仅能对城市整体空气质量进行宏观表征,

无法反映城市局部区域的空气质量微观特征。

2015 年 2 月 28 日,前央视女记者柴静推出雾霾深度调查纪录片《穹顶之下》,该片通过视频网站、社交网络等共同的力量,让数千万乃至上亿人再次认识到雾霾的危害。为了对抗雾霾,该片提出呼吁:“世界上再强大的政府也无法独立治理好污染。”——雾霾治理仅靠政府是不够的,仅靠环保部门更不行,它一定需要公民个体和民间组织的参与。

民众已经开始在行动。自 2013 年 1 月 27 日起,北京市民邹毅开启了他的《一目了然天天晨报》之旅:每天坚持在同一时间、同一地点、同一角度,用镜头对准北京电视台拍摄一张照片制成北京视觉空气日志,分享到微博和微信朋友圈。邹毅希望用这样一种简单直接的方式,让大家看到我们生活的环境遭受到的环境污染和健康威胁,引发大家的关注和重视。截止到 2018 年 1 月 1 日,邹毅已经发布了近五年共 1800 期《一目了然天天晨报》(2013-2017 一目了然年报如图 1-1 至图 1-5 所示),《新京日报》、《财经网》、《新华网》等主流媒体争相报道,引起了广泛的社会媒体关注和影响。受此带动,全国各地乃至全球的环保志愿者纷纷行动起来,于是我们可以看到越来越多的城市视觉日记(部分城市记录如图 1-6 所示)。

通过相机拍摄观测场景,则可以提供相对长距离广区域(数公里级别)的空气质量信息。如果能便捷的记录,进而有效的分析这些视觉数据,获取图像特征与空气质量之间的关系,我们就可能通过便捷的视觉手段对空气质量进行估计,无疑将对提升全社会的环保意识有着巨大的推动作用。



图 1-1 2013 年一目了然年报

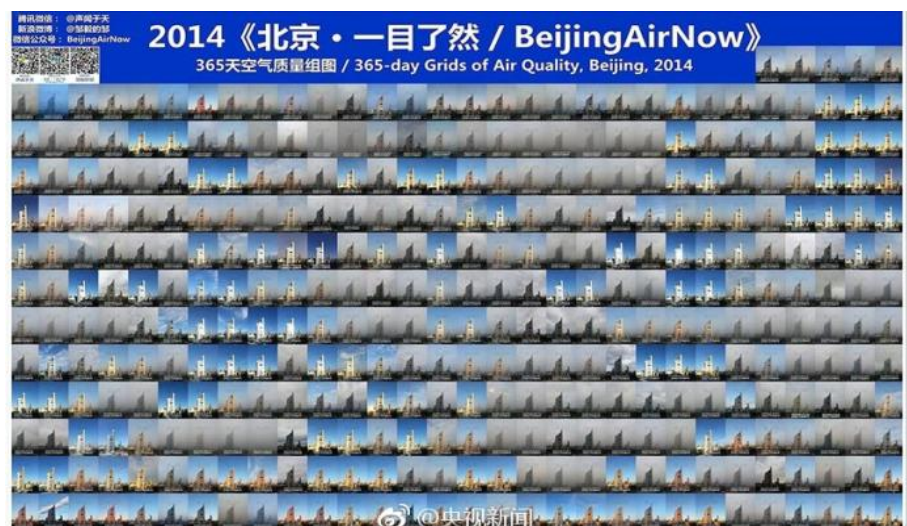


图 1-2 2014 年一目了然年报



图 1-3 2015 年一目了然年报

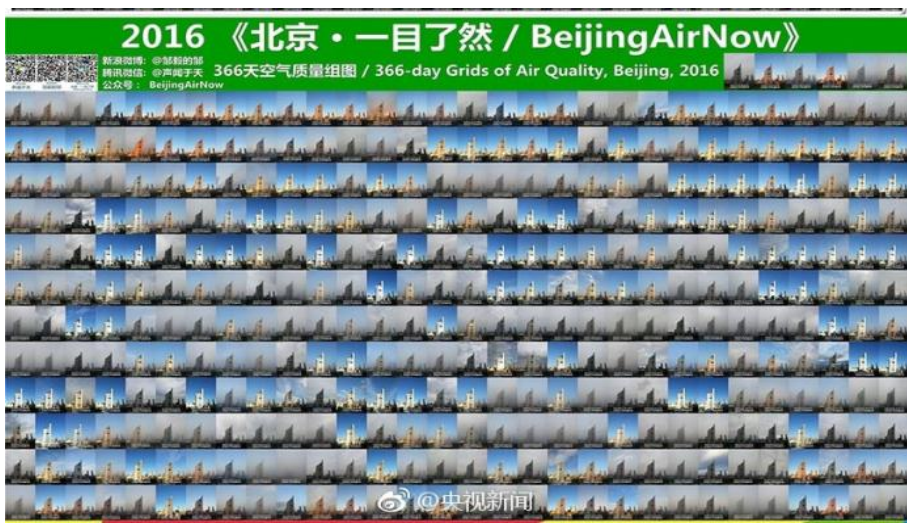


图 1-4 2016 年一目了然年报

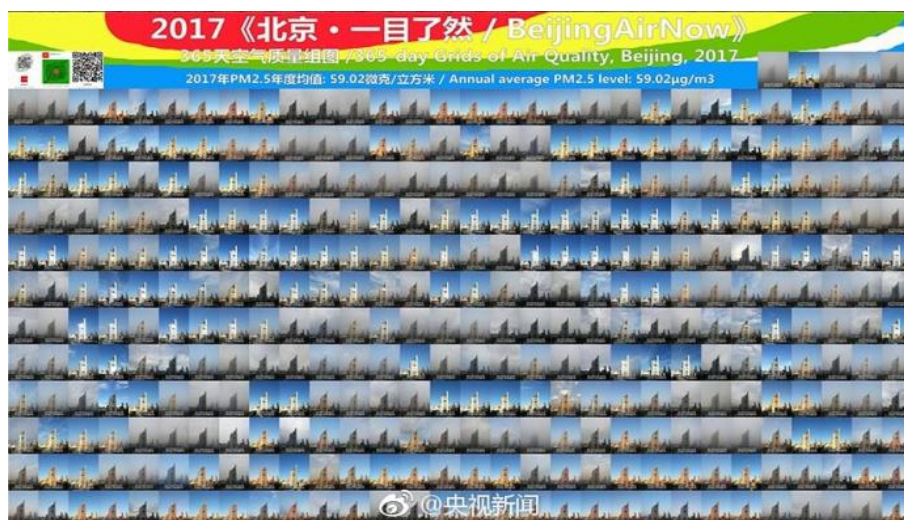


图 1-5 2017 年一目了然年报



图 1-6 部分城市视觉日志

1.2 国内外研究现状

早期人们采用直接测量法计算空气质量, 包括 β 衰减检测器^[1]、光散射法^[2]以及锥形元件震荡微天平法^[3]等。测量具体过程如下: 在待测量场景中放置采样滤膜, 在一段时间(即采样周期)内对空气中的颗粒物进行采样并计算重量变化。将重量变化量与采集到的空气颗粒物总重量做比值即得到待测场景在采样周期内空气颗粒物的浓度。然而, 该方法不对颗粒物种类进行区分, 无法得出具体污染颗粒物的浓度。因此, 研究者们又依据不同粒子的尺度、物理性质等特征设计了空气颗粒物分离系统。该系统能够成功得到 PM2.5、PM10 以及可吸入颗粒物的浓度。直接量测法主要依赖于滤网采集, 仅能在一段时间内工作, 并且滤网需及时更换, 系统也要进行保养, 成本较高。

目前对于空气质量检测主要基于空气质量监测站。然而, 每个城市内仅有若干个空气质量监测站且仅在很小一部分区域内进行采样, 无法对任意地区空气质量进行准确估计。以北京为例, 全市仅有 35 个空气质量监测站点, 且分布不均

（如图 1-7 所示）。同时，空气质量监测站需要花费上百万进行维修保养，成本颇高。



图 1-7 北京空气质量监测站配置及布局

为了利用有限的空气质量监测站数据对城市范围内的空气质量进行精准估计，不同领域的研究人员为此展开研究。

基于空间插值的空气质量等级估计方法是由环境和公共健康研究单位最早提出的空气质量检测方法，用以检测环境空气质量与人类身体健康效的关系^{[4][5]}。这一类方法包括：空间平均法，最近邻法等。这些方法主要基于对空气污染颗粒分布的空间连续性假设。然而，这类方法有两点不足：（1）不同差值方法得到的估计数值差异较大；（2）对于空间分布较稀疏的原始数据，其差值结果并不理想。

处理全市空气质量监测的更可靠的方法是基于辅助传感器（如遥感、车载传感器（VSN）、众包等）。Gupta 等^[6]通过从卫星成像数据中提取的气溶胶光学厚度估算全球 PM_{2.5} 质量。Yu 等人^[7]提出了基于 VSN 的城市环境监测制度，通过压缩感知技术实现数据采样。上述方法一般需要相对高成本的基础设施部署。越来越多的研究者将精力集中在众包计算。Li 等人^[8]通过便携式传感器和智能手机监测气体污染物和颗粒物。然而，目前的便携式传感器无法达到监测站的精度。此外，PM_{2.5} 检测设备需要超过 1 小时的数据采集过程。模型还需要尽可能避免因移动造成的误差。

一些研究人员提出模拟污染物排放模型进行空气质量等级估计。不同的空气污染物对应不同的排放模型。通过对燃料燃烧、农业、运输等部门进行详细划分来估算总体污染物水平。相关环保部门会对不同的模型设定相应的标准。例如，

根据不同的驾驶模式,美国移动部门推出机动车排放模拟器(MOVES)模型^[9]来模拟不同类型机动车排放污染物的过程。另一种方法是利用扩散模型^[10]描述污染物排放及消散过程进而对污染物浓度进行预测。一些研究人员采用理论气象散射模型^{[11][12]}模拟污染物排放过程^[13-15]。CMAQ 模型模拟 PM_{2.5} 排放以及对前体减排的响应过程^[16]。该方法可用来设计有效的污染物排放控制策略。WRF-CHEM 模型用来模拟气象模型和空气质量^[17]。这是一个完全耦合的线上模型,可以在气象模型运行的同时对空气质量进行模拟,提高预测潜力。然而,由于理论基础不完备,模拟过程易受未知污染物排放的影响,测量精度较差。其中一些模型通常需要很多可能难以满足的假设和复杂的参数设置。

一些研究人员采用基于数据驱动的统计方法进行空气质量等级估计。人工神经网络(ANN)^{[18][19]}、多元线性回归(MLR)^[20]和支持向量回归(SVM)均普遍应用于空气质量等级估计。考虑到不同空气质量站之间的高空间相关性,研究人员考虑应用时空预测模型进行空气质量等级估计,如时空人工神经网络(STANN)^[21]、时空支持向量机(STSVR)模型^[22]以及时空自动编码器。

最近,许多研究集中在通过时空(ST)异构城市大数据估计空气质量。通过分析城市大数据的相关性和模式,可以挖掘未知但有用信息,这涉及到包含空间,时间和类别信息的数据集^[23-26]。基本假设是空气质量受到城市的动态影响,如风、人群、移动车辆、污染物排放点等。通过分析诸如气象和交通等能源动力学数据之间的时间依赖性和空间相关性,未被监测站覆盖的场景空气质量能够被准确估计。Zheng^[27]通过使用城市地区的多个数据来推断细粒噪声污染。Shang 等人^[28]根据 GPS 轨迹数据估计气体消耗和污染物排放。Ryan 和 LeMasters^[29]回顾了基于周围土地利用和交通量预测空气质量的土地利用回归模型。Zheng^[30]提出了一种城市数据驱动的方法来预测未来 48 小时内的空气污染。然而,这些工作以耗时的复杂算法为代价取得了很好的结果。此外,在这些工作中大量地获取的信息很难获得。

随着计算机视觉的发展,室外摄像机的使用引起了人们的极大兴趣。尽管有显著的研究价值,但只有很少的研究人员进行基于图像数据的空气质量研究。北京邮电大学的 Z Zhang 等人建立两个观测点的图像数据库,提取多个图像特征作为空气质量预测的鲁棒表征。通过使用机器学习方法,他们学习了一个用于空气质量等级估计的自适应分类器^{[31][32]}。清华大学的 H Q Wang 等人选择一个观测站点通过相机收集场景图像,以此分析了 PM_{2.5} 浓度与图像清晰程度的关系^[33]。

1.3 研究目的和内容

1.3.1 研究目的

使用户外图像进行空气质量评估对于计算机来说是一项艰巨的任务,有时甚至对人来说也是如此。本文旨在通过收集任意场景图像,提取表征空气质量的特征,依靠统计学习方法研究空气质量变化对场景成像的影响,以此对任意场景空气质量进行估计。该方法有较广的观测范围、丰富的时间空间信息,较低的配置成本、简单易使用等优点,有望成为现有空气质量监测方法的一种重要补充。

1.3.2 研究内容

本文主要研究了如何利用图像特征对场景空气质量进行估计。本研究的核心在于相关性特征的提取。图像特征种类繁多,理想的特征应具有高效性、可区分性及可重复性等特点。基于此,本文提出利用图像视觉特征、能见度特征和深度特征分别对空气质量进行估计,具体如下:

- (1) 研究不同图像视觉特征与场景空气质量的关系,并利用机器学习方法实现基于视觉图像特征的空气质量等级估计;
- (2) 提出利用单一场景图像计算能见度的方法,并基于此方法提取图像能见度特征并结合机器学习方法实现空气质量等级估计;
- (3) 利用卷积神经网络提取场景图像的深度特征,研究其与场景空气质量的关系,并尝试利用深度特征估计场景空气质量等级。

1.4 本章小结

针对现有空气质量计算方法存在的成本高、效率低、测量范围窄等不足,本文提出通过拍摄场景图像获取相关性特征并进行空气质量等级估计。为了研究图像不同性质的特征与场景空气质量的关系,本文提出三类不同类型的图像特征:视觉特征、能见度特征和深度特征。视觉特征表征了图像的纹理、边缘、颜色、对比度以及与大气成像相关的表象特征;能见度特征描述了场景大气粒子分布特性,对场景空气质量有直观表征;深度特征表征了图像深层次的视觉机理,挖掘了场景空气质量的本质信息。本文分别通过这三类图像特征进行基于场景图像的空气质量等级估计,并探索能够表征场景空气质量的相关性图像特征。

本文共分五个部分,论文的组织结构如下:

第一章为绪论部分,介绍了目前空气质量检测的研究背景及国内外研究现

状。同时，对本文的研究目的和内容进行分析。

第二章对视觉特征进行分析，分别介绍了 6 类视觉特征及其计算方法，并对本文提出的场景数据库 VAQI-1 进行说明。基于此数据集，本章利用机器学习算法实现基于这 6 类视觉特征的空气质量等级估计并进行试验结果分析。

第三章对能见度特征进行分析，提出了一种基于单幅图像的场景能见度估计算法，并通过实验成功证明该方法的有效性。基于 VAQI-1，本章同样利用机器学习算法实现了基于能见度特征的空气质量等级估计，并与视觉特征结果进行比较分析得出实验结论并进行分析。

第四章对深度特征进行分析，概述了用于提取图像深度特征的卷积神经网络。基于 VAQI-1，本章利用不同类型的卷积神经网络提取深度特征进行空气质量等级估计，并与前两类特征即视觉特征和能见度特征的估计结果相比较得出实验结论。

第五章对本文进行总结，并指出本文的优势及创新点。

2 基于视觉特征的空气质量等级估计

视觉特征描述了图像拍摄区域内场景景物的表面性质，包括颜色、边缘、纹理、清晰度、局部特征等。这些性质几乎表征了场景成像过程中的各种特性。基于此，本章提取了 6 类能够描述上述特性的视觉特征，包括：SIFT 特征、HOG 特征、LBP 特征、色彩饱和度、颜色直方图和对比度。接下来，本文将对不同特征的计算方法进行详述。

2.1 视觉特征

2.1.1 SIFT 特征

SIFT 特征是 Lowe^[34]提出 SIFT 特征用于从图像缩放和旋转不变的图像中提取独特的不变特征，用于描述图像的局部特性。该特征被广泛应用于图像拼接，识别，检索等。

SIFT 算法具体过程如下：

- (1) 尺度空间极值检测：利用高斯模糊算法获取尺度空间，对所有尺度上的图像位置进行搜索，并利用高斯微分函数识别拥有尺度和旋转不变性的兴趣点；
- (2) 关键点定位：对于每个兴趣点位置都通过一个拟合模型来确定尺度和位置。拟合模型通常是三维二次模型函数，不仅可以剔除不稳定的边缘响应点和低对比度的关键点，还能提高匹配的稳定性及抗噪能力；
- (3) 方向确定：通过计算图像局部的梯度方向给每个关键点分配一个或多个方向。由于后续操作均是基于关键点的尺度、位置和方向进行变换，从而提供了对于这些变换的不变性。梯度的模值和方向计算如式 (2-1) 和式 (2-2)

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2-1)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (2-2)$$

L 为关键点所在的尺度空间值。完成关键点的梯度计算之后，使用直方图统计关键点邻域内像素的方向和梯度。方向直方图的峰值代表该特征点邻域梯度的方向，最大值代表该关键点的主方向。算法为了提高匹配的鲁棒性，仅保留大于主方向峰值 80% 的方向并设定该方向为关键点的辅方向。因此，对于同一梯度值多个峰值的关键点位置，在相同的尺度和位置将会创建多个不同方向的关键点。虽然仅有 15% 的关键点被赋予多个方向，但关键点匹配的稳定性可以明显提高。

- (4) 关键点描述：每个关键点都包含三个信息：位置、方向和尺度。接下来需

要用一组向量对每个关键点的三个信息进行描述,使其不随光照、视角等变化而发生改变。描述向量是对关键点邻域高斯图像梯度统计结果的表示。通过对关键点周围图像区域进行分块并计算块内的梯度直方图生成特征描述向量。对于描述向量的要求包括:包含关键点、包含关键点周围对该点有贡献的像素点以及应具有较高的独特性。这样的描述向量是局部区域图像信息的一种抽象性的表征,能够提高特征点正确匹配的可能性,具有唯一性。

2.1.2 HOG 特征

梯度方向直方图 (HOG) [35]假定图像中的局部对象外观和形状由强度梯度或边缘方向的分布来描述,通过计算中心的水平和垂直梯度方向和大小并创建单元格上的直方图。对于彩色图像,为每个像素选择具有最高梯度幅度的颜色通道。梯度方向和大小由公式 2-3 和公式 2-4 计算得到。

$$grad(x, y) = \sqrt{sub - gradient_x(x, y)^2 + sub - gradient_y(x, y)^2} \quad (2-3)$$

$$angle(x, y) = \tan^{-1} \frac{sub - gradient_y(x, y)}{sub - gradient_x(x, y)} \quad (2-4)$$

HOG 特征计算过程如图 2-1 所示。

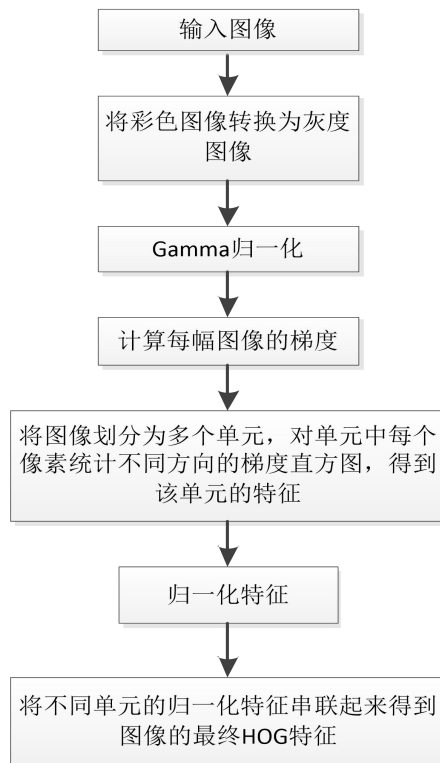


图 2-1 HOG 特征计算过程

2.1.3 LBP 特征

局部二值模式 (LBP) [36] 是一种特殊的纹理特征描述算子, 其在各种应用中取得了很好的效果。LBP 具有高度可区分性, 其关键优势在于对单调灰度级变化和计算效率的不变性, 这使得其该特征用于要求苛刻的图像分析任务。

对于给定图像, 其 LBP 特征通过比较图像中每个像素与其邻域像素的差异值计算得到:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (2-5)$$

其中 g_c 是中心像素的灰度值, g_p 是其邻域像素的灰度值, P 是邻域的像素个数。假设 g_c 的坐标是 $(0,0)$, 那么 g_p 的坐标为 $(R \cos(2\pi p / P), R \sin(2\pi p / P))$ 。不在图像网格中的邻域灰度值可以通过插值来估计。假设图像大小为 $I * J$ 。在识别每个像素的 LBP 值之后, 构建直方图表征图像纹理特征:

$$H(k) = \sum_{i=1}^I \sum_{j=1}^J f(LBP_{P,R}(i, j), k), k \in [0, K] \quad (2-6)$$

$$f(x, y) = \begin{cases} 1, & x = y \\ 0, & \text{else} \end{cases}$$

其中 K 是 LBP 特征最大值。这里定义 U 为空间转换的数量 (如公式 (2-7) 所示)。

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (2-7)$$

为了保证 LBP 特征的旋转不变性, 局部旋转不变模式定义如下:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{else} \end{cases} \quad (2-8)$$

2.1.4 色彩饱和度

色彩饱和度表征色彩的纯度, 越鲜明的色彩意味着饱和度越高。对于给定的 RGB 彩色图像。我们首先将 RGB 颜色空间转换为 HSV (色调-饱和度-亮度) 颜色空间, 再计算 HSV 空间中的色彩饱和度。

HSV 是圆锥形几何形状 (如图 2-2), 色调的角度尺寸从 0° 的红色原色开始, 通过 120° 的绿色原色和 240° 的蓝色原色, 然后回卷到 360° 的红色原色。

在每个几何图形中，中心垂直轴从下到上表示亮度的提升，0 为黑色，1 为白色。从边缘到中心表征饱和度的下降。

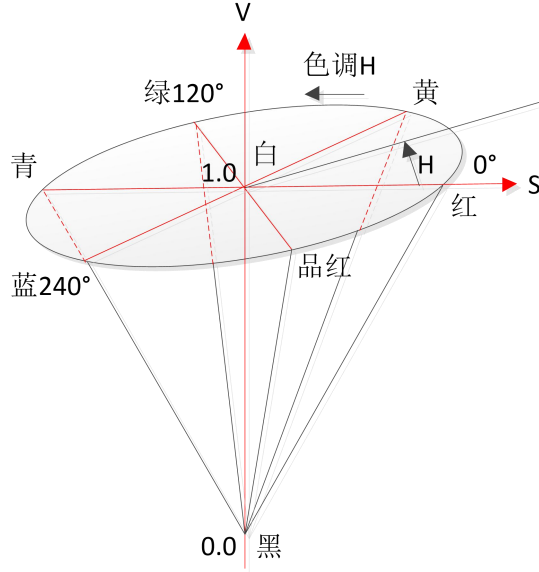


图 2-2 HSV 颜色空间模型

对于给定 RGB 彩色图像，通过公式将其转换为 HSV 通道图像。

$$h = \begin{cases} 0^\circ & \text{if } \max = \min \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, & \text{elseif } \max = r \text{ and } g \geq b \\ 60^\circ \times \frac{g-b}{\max-\min} + 360^\circ, & \text{elseif } \max = r \text{ and } g < b \\ 60^\circ \times \frac{g-b}{\max-\min} + 120^\circ, & \text{elseif } \max = g \\ 60^\circ \times \frac{g-b}{\max-\min} + 240^\circ, & \text{else } \max = b \end{cases} \quad (2-9)$$

$$s = \begin{cases} 0 & \text{if } l = 0 \text{ or } \max = \min \\ \frac{\max-\min}{\max+\min} = \frac{\max-\min}{2l}, & \text{elseif } 0 < l < \frac{1}{2} \\ \frac{\max-\min}{2-(\max+\min)} = \frac{\max-\min}{2-2l}, & \text{else } l > \frac{1}{2} \end{cases} \quad (2-10)$$

$$v = \max \quad (2-11)$$

式中 $\max(S_l)$ 表征饱和度最大值， $\min(S_l)$ 表征饱和度最小值。式中 $\max$ 表示 RGB 中的最大值， $\min$ 表示 RGB 中的最小值。我们利用公式（2-12）将饱和度归一化

$$S(x, y) = \frac{S_{x,y} - \min(S_I)}{\max(S_I) - \min(S_I)} \quad (2-12)$$

2.1.5 颜色直方图

颜色直方图描述了图像各种色彩在整幅图像中所占的比例,而对于每类颜色所处的位置并不关注,即该特征无法描述图像中具体的对象和物体。颜色直方图可以在不同的颜色空间和坐标系上进行计算。一般情况下图像的表征都是在 RGB 颜色通道上进行。然而,RGB 颜色空间结构不符合人们对于颜色相似性的主观判断。因而有研究人员提出基于 HSV 空间、LUV 空间和 LAB 空间的颜色直方图。这些颜色空间更接近人眼对于颜色的主观认识。

计算颜色直方图需要先进行颜色量化,即将颜色空间划分为若干个小空间。每个小区间可看作是直方图的一个像素统计量。然后统计每个小区间内所有颜色的像素数量作为直方图的统计数值。用于颜色量化的算法很多,如向量量化、聚类或神经网络。目前,最普遍的方法是将颜色空间的各个分量进行均匀划分。然而,这种方法会存在某分量的像素数量非常少的不均衡情况。相比之下,聚类算法考虑到了图像颜色在空间中的分布情况,可以有效解决不同颜色的像素值分布不均的情况,更有益于颜色量化。然而,基于聚类的方法也会存在问题。假设两幅图像的颜色直方图几乎一样,只是错开了一个统计量,这是如果采用距离侧脸算法计算两幅图的相似度,会得到非常小的相似度值。为了解决这个问题,我们需要考虑相似但不相同的颜色之间的相似程度。一种方法是采用二次式距离;另一种方法是对颜色直方图先进行平滑过滤,即每个统计量对相邻统计量的贡献。这样的话相似但颜色不相同的相似度也会对小区间之间的相似度有一定贡献。

选择合适的颜色小区间数目和颜色量化方法是计算颜色直方图最重要的两个因素,这与具体应用的效率要求和性能有关。一般情况下,颜色空间分的越细,小区间的数目越多,直方图对颜色的分辨能力越强。然而,小区间的数目过大会增加计算负担。一种有效减少小区间个数的方法是只选取像素数目较多的直方图来构造图像特征,因为这些表征图像主要颜色的小区间能够近似表征图像的颜色特征。

2.1.6 对比度

对比度是指图像明暗区域中最亮的区域与最暗的区域之间不同亮度等级的差距。图像亮度亮区域和暗区域的差距越大代表对比度越大,同理差距越小即对比度越小。目前有以下几种对比度定义。

(1) 韦伯对比度。

韦伯定律：又名感觉阈值定律，即在相同刺激下人能感受到的刺激的动态范围与标准刺激的强度成正比

$$K = \frac{\Delta I}{I} \quad (2-13)$$

K 为给定常数， I 为刺激， ΔI 为能感受到的刺激的动态范围。

将韦伯定律应用到人的视觉刺激，得韦伯对比度定义如下：

$$C_W = \frac{I - I_b}{I_b} \quad (2-14)$$

I 为物体的亮度值， I_b 为背景的亮度值。

(2) Michelson 对比度

Michelson 对比度^[37]又称为视觉能见度，其定义如下：

$$C_M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (2-15)$$

式中 $I_{\max}$ 和 $I_{\min}$ 分别表示图像最亮的亮度和最暗的亮度。从理论上讲，Michelson 对比度与人眼视觉感受中视锥细胞对视场光通频率的感受程度是近似一致的。

(3) 均方根对比度

均方根对比度用于描述图像内像素值的均方根（即标准差）：

$$C_\delta = \delta_{I_{w \times h}} = \sqrt{\frac{1}{w \times h} \cdot \sum_{I_{w \times h}} (I(x, y) - \mu_{I_{w \times h}})^2} \quad (2-16)$$

$$\mu_{I_{w \times h}} = \frac{1}{w \times h} \cdot \sum_{I_{w \times h}} I(x, y) \quad (2-17)$$

式中 $I_{w \times h}$ 表示宽为 w 高为 h 的像素图， $I(x, y)$ 是图像中 (x, y) 位置的像素值， $\mu_{I_{w \times h}}$ 代表图的像素平均值， $\delta_{I_{w \times h}}$ 是像素标准差。

本文采用韦伯对比度，分别对图像计算前景亮度平均值和背景亮度平均值并代入公式（2-15）得到对比度值。

2.2 实验设计与结果分析

2.2.1 实验方案设计

为找出最能描述场景空气质量的传统特征，本章采用支持向量机（SVM）^[38]作为分类器，利用上述 6 类传统特征分别进行空气质量等级估计。实验框图如图 2-3 所示

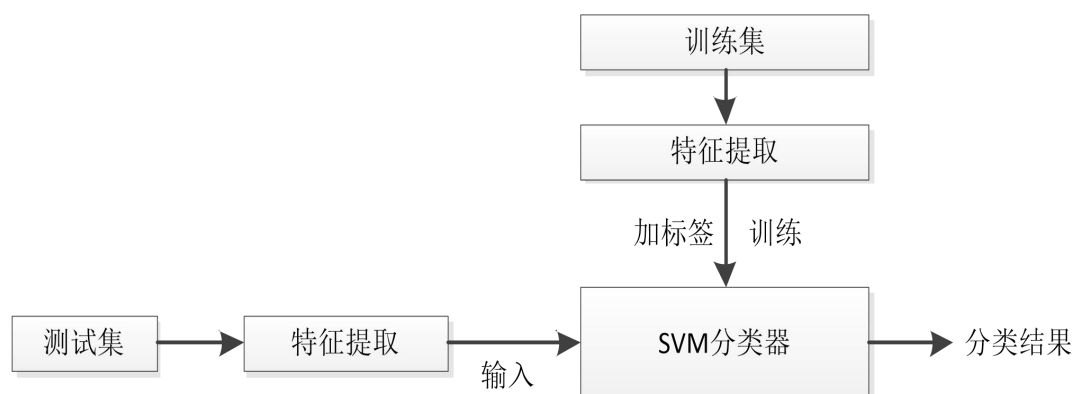


图 2-3 基于传统特征的空气质量等级分类

支持向量机是目前比较流行的二类分类器，其模型为特征空间上最大间隔的线性分类器。同时，支持向量机还包括核操作，使得它也成为了一种非线性分类器。支持向量机的基本目标是求解能够分类数据集并使几何间隔最大的分离超平面。对于线性可分的数据集来说存在无穷多个可分离超平面，但仅存在唯一的几何间隔最大的分离超平面。间隔最大化可理解为：寻找最大几何间隔的超平面等同于以较大的置信度对数据进行分类，即不仅分离了间隔较远的数据，同时也对间隔很近的数据也能做到较好的分类。如果支持向量机能学习到这种超平面，则意味着它能对未知的测试数据也能有较好的分类作用，即有较强的泛化性。

由于本实验属于多项分类，无法直接使用 SVM。因此，本文结合多个 SVM 构造多分类分类器。

结合多个 SVM 构造多类分类器的方法主要有两类：

（1）一对多法（one-versus-rest）。一对多法，简称 OVR SVMs，主要通过将多个 SVM 分类器进行组合构造多类分类器。在训练时依次提取某一类的样本作为一类，把其余类的样本当作第二类，代入一个 SVM 分类器进行分类。 k 个类别的训练数据会需要 k 个 SVM。在进行测试时，将测试数据分别代入这 k 个 SVM 分类器，每个分类器都会得到一个输出结果。取输出结果里的最大值作为最终的分类结果。

（2）一对一法（one-versus-one）。一对一法，简称 OVO SVMs，主要是对任意

两类训练样本设计一个 SVM。 k 个样本需要 $k(k-1)/2$ 个 SVM。在对测试样本进行分类时, 将该测试样本分别代入这些 SVM 中进行分类, 并记录样本被分类到的所有类别的个数。最后统计得票最高的类别即为该测试样本的类别结果。

以上两种方法中, 一对多法存在缺陷, 因为每个 SVM 对应的两类样本数据量差别较大, 学习的分类超平面存在偏差, 容易过拟合。而第二种方法实现效果较好, 但唯一不足的是当分类类别较多时, SVM 的个数会大大增多, 使得计算压力巨大。

本文实验需要对输入数据进行空气质量等级分类。空气质量等级共 6 类, 分别为: 优、良、轻度污染、中度污染、中度污染以及严重污染。由于分类类别较少, 因此本文采用一对一法构造 6 类分类器。

在训练时候将 6 类空气质量等级两两匹配, 并选择对应等级的训练数据训练对应的 SVM 分类器。在测试的时候, 将测试数据代入所有的 SVM 进行测试, 并统计投票结果。投票过程如下:

- (1) 令 A、B、C、D、E、F 分别为空气质量 6 个等级的统计个数;
- (2) 代入(优, 良)一分类器, 如果分为优, 则 $A = A + 1$; 如果分为良, 则 $B = B + 1$;
- (3) 代入(优, 轻度污染)一分类器, 如果分为优, 则 $A = A + 1$; 如果分为轻度污染, 则 $C = C + 1$;
- (4) 代入(优, 中度污染)一分类器, 如果分为优, 则 $A = A + 1$; 如果分为中度污染, 则 $D = D + 1$;
- (5) 代入(优, 重度污染)一分类器, 如果分为优, 则 $A = A + 1$; 如果分为重度污染, 则 $E = E + 1$;
- (6) 代入(优, 严重污染)一分类器, 如果分为优, 则 $A = A + 1$; 如果分为严重污染, 则 $F = F + 1$;
- (7) 代入(良, 轻度污染)一分类器, 如果分为良, 则 $B = B + 1$; 如果分为轻度污染, 则 $C = C + 1$;
- (8) 代入(良, 中度污染)一分类器, 如果分为良, 则 $B = B + 1$; 如果分为中度污染, 则 $D = D + 1$;
- (9) 代入(良, 重度污染)一分类器, 如果分为良, 则 $B = B + 1$; 如果分为重度污染, 则 $E = E + 1$;
- (10) 代入(良, 严重污染)一分类器, 如果分为良, 则 $B = B + 1$; 如果分为严重污染, 则 $F = F + 1$;
- (11) 代入(轻度污染, 中度污染)一分类器, 如果分为轻度污染, 则 $C = C + 1$; 如果分为中度污染, 则 $D = D + 1$;
- (12) 代入(轻度污染, 重度污染)一分类器, 如果分为轻度污染, 则 $C = C + 1$;

如果分为重度污染, 则 $E = E + 1$;

(13) 代入(轻度污染, 严重污染)一分类器, 如果分为轻度污染, 则 $C = C + 1$;

如果分为严重污染, 则 $F = F + 1$;

(14) 代入(中度污染, 重度污染)一分类器, 如果分为中度污染, 则 $D = D + 1$;

如果分为重度污染, 则 $E = E + 1$;

(15) 代入(中度污染, 严重污染)一分类器, 如果分为中度污染, 则 $D = D + 1$;

如果分为重度污染, 则 $F = F + 1$;

(16) 代入(重度污染, 严重污染)一分类器, 如果分为重度污染, 则 $E = E + 1$;

如果分为严重污染, 则 $F = F + 1$;

(17) 选择 A 到 F 中最大的值, 其对应的类别即为测试样本的类别。

2.2.2 VAQI-1 数据库

本文搭建了一个为期三年的任意场景高质量图像数据库 VAQI-1。VAQI-1 共包含 4444 幅图像, 来自中国 26 个城市的 75 个观测点。观测城市分布如图 2-4 所示。VAQI-1 是不同空气质量下的场景图像集合, 采用空气质量指数(AQI)作为图像数据标签。空气质量指数表征空气质量好坏, 分为六个等级。等级越高表明空气污染物浓度增加(如表 2-1 所示)。VAQI-1 在六种空气污染等级下的图像数量统计如图 2-5 所示。图 2-6 显示了若干场景在不同空气污染等级下的场景图像。



图 2-4 观测城市分布

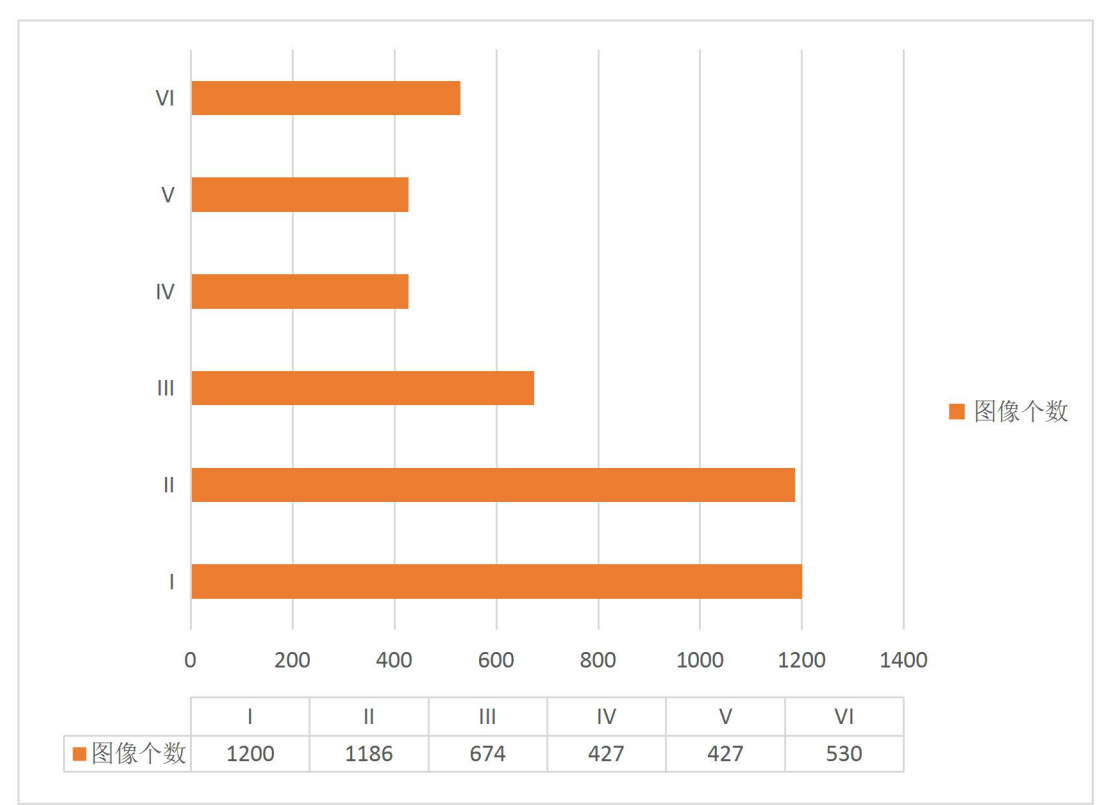


图 2-5 不同空气污染等级下的图像数量统计

表 2-1 空气质量等级分类

空气质量指数 (AQI)	空气质量等级	表示颜色	描述
0-50	I	绿	没有空气污染
51-100	II	黄	空气质量很好仅存在轻微污染
101-150	III	橙	空气污染略加重不适宜敏感人群
151-200	IV	红	空气污染对健康人群有一定影响
201-300	V	紫	健康人群普遍出现不适症状
301-500	VI	栗	健康人群出现强烈不适症状

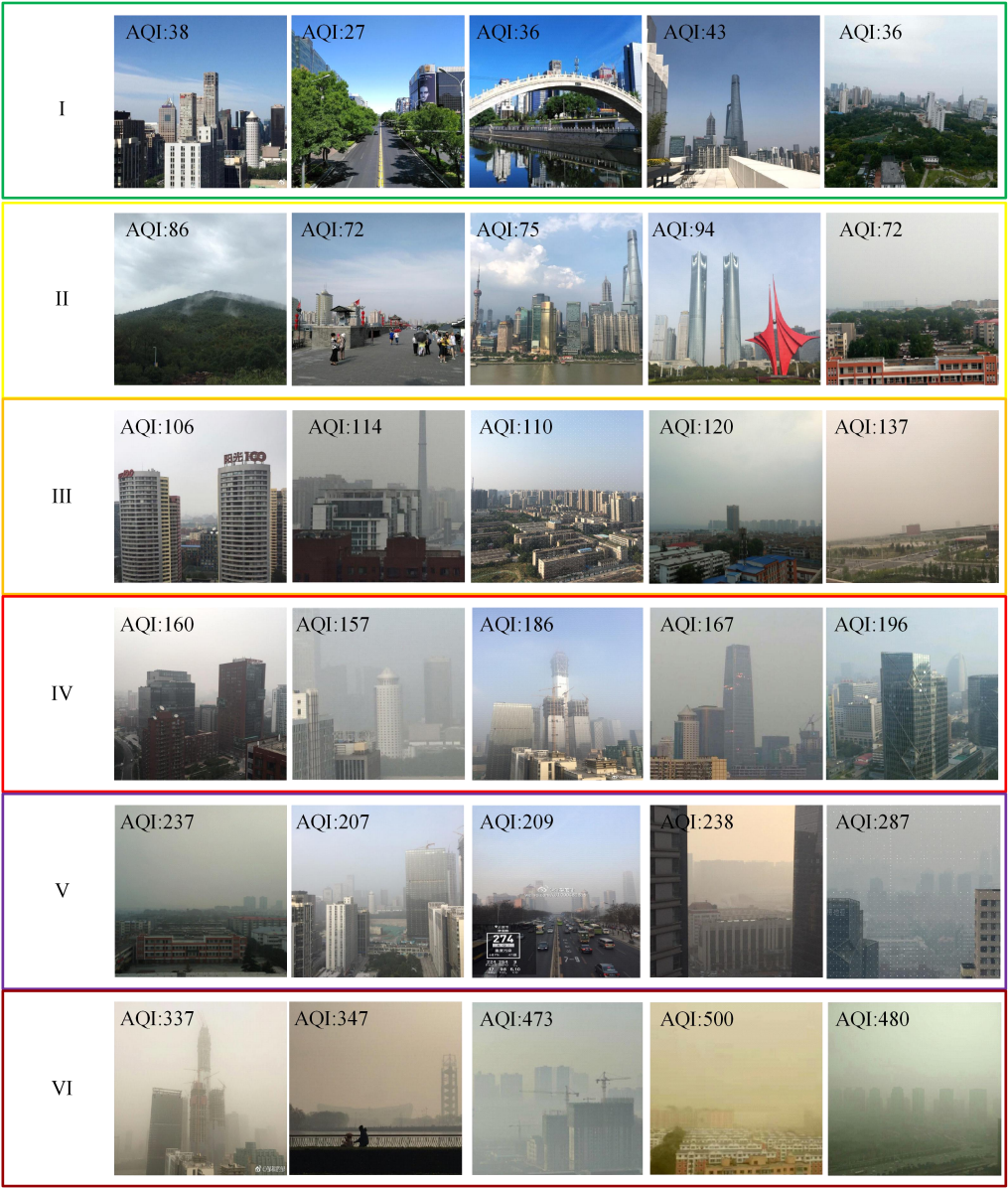


图 2-6 VAQI-1 部分示例图像

我们将 VAQI-1 与现有户外场景空气质量数据库进行比较（见表 2-2）。可以发现，VAQI-1 明显扩充了观测点的数量和数据库的大小。并且目前只有 VAQI-1 是公开数据集。为吸引更多研究者参与到基于场景图像的空气质量评估这一研究来，我们会持续更新并完善 VAQI-1。

表 2-2 现有户外场景空气质量数据库比较

数据库	[31]	[33]	VAQI-1
图片数量	2000	<500	4444
观测场景数量	2	1	75
观测时间	2014 年	2013 年-2014 年	2014 年到现在
是否公开	否	否	是

本文按 3:1 的比例从 AQI-1 中抽取 3000 张为训练图像, 1444 张为测试图像用于训练分类模型。设定类别个数为 6 类, 即 6 类空气质量等级: 优、良、轻度污染、中度污染、重度污染和严重污染。

2.2.3 结果分析

基于 VAQI-1 数据集, 我们对训练库和测试库中的图像提取 6 类特征传统特征并利用多分类 SVM 进行分类。准确率计算公式如式 2-19 所示

$$\text{准确率} = \frac{\text{测试集分类正确的图像数量}}{\text{测试集图像总量}} \times 100\% \quad (2-19)$$

实验结果如表 2-3 所示

表 2-3 基于视觉特征的空气质量等级估计结果对比

	视觉特征	准确率[%]
1	SIFT	8
2	HOG	12
3	LBP	14
4	色彩饱和度	35
5	颜色直方图	20
6	对比度	42

从表 2-3 可以看出, 基于视觉特征的空气质量分类效果不理想。这是因为图像视觉特征(如边缘、纹理、颜色等)仅仅描述了场景图像的表象特征。由于 VAQI-1 中的场景图像来自任意场景和任意拍摄角度, 所以此时的图像特征只是尽可能的描述了图像的场景类别信息、颜色信息、景物位置信息等。而这些信息的变化要比场景空气质量的变化大很多。因此利用这些特征几乎完全估计不出场景的空气质量。相比较而言, 对比度特征表现要明显优于其他特征。这是因为对比度特征表征了场景的清晰度, 随大气浑浊程度的变化而发生改变, 对场景成像介质有较直观描述。然而, 基于对比度的空气质量等级估计仍无法得到满意的结果, 这也许是因为对比度的计算受图像局部信息影响较大, 无法准确描述场景空气质量。

2.3 本章小结

本章主要介绍了 6 类图像视觉特征, 分别是: SIFT 特征、HOG 特征、LBP 特征、色彩饱和度、颜色直方图以及对比度。这些特征描述了图像拍摄区域内场景景物的表面性质, 如颜色、边缘、纹理、清晰度、局部特征等。为了探究基于图像的空气质量等级估计, 本文搭建了一个任意场景高质量图像数据库 VAQI-1, 并记录了每张图像对应的 AQI 值。本章利用多分类 SVM 方法, 分别基于上述 6

类特征对场景图像的空气质量等级进行估计。实验结果表明，基于视觉特征的估计效果普遍不理想，最高准确率仅达 42%。这说明了依靠视觉特征无法描述场景空气质量。

3 基于能见度特征的空气质量等级估计

由第2章我们可以看出,仅依靠视觉特征无法对空气质量进行估计,这是由于视觉特征仅能够描述场景图像的表面性质,对于成像介质中的粒子分布特性没有涉及。基于此,本章提出利用一类专门描述大气粒子分布特性的特征——能见度对空气质量进行估计。接下来,本文将对能见度特征的计算方法进行详述。

3.1 能见度特征

能见度描述了场景最大的可视范围,对空气质量有直观反应。当能见度值较大时,场景可视范围较广,说明此时大气清晰度较好,空气质量好;当能见度较小时,场景可视范围较小,说明大气浑浊度较高,空气质量差。依靠能见度检测,人们可以对当前场景空气质量有一个粗略判断。

目前,能见度的测量主要通过仪器完成。世界上普遍使用的能见度仪是透射仪和前向散射仪。透射仪在雾滴引起的中低能见度条件下测量结果较为准确,常被用于民航系统,但在有降水(如雨、雪等)或漂尘(如扬沙等)现象发生时,透射仪自身光源引起的前向散射光的影响不容忽视^{[39][40]}。前向散射仪因其体积小、安装简单、容易维护、测量范围广等特点,在航空、码头、高速公路等部门得到了广泛地应用,但在不同类型气溶胶条件下前向散射仪需要不同的修正,且非常小的采样体积的代表性也是需要考虑的问题^[41]。

随着计算机视觉技术的发展,基于数字摄像技术测量能见度的方法得到了国内外学者的广泛关注。美国明尼苏达大学运输系的 Tack Mu Kwon 使用视频照相机和多个标靶,在高速公路上进行日间能见度检测,通过将标靶的对比度值进行能见度非线性曲线拟合得出能见度值^[42]。吕伟涛等^[43]利用2组目标亮度差的比值计算气象能见度,即双亮度差方法。该方法能有效地消除系统暗电流和背景杂散光的影响,但是仍需增加人工标志物,且测量点过少带来的检测误差也不可忽视。Hautière N 等^[44]将雾天图像成像模型和大气对比度的衰减模型相结合,实现了无需人工标记的能见度测量^[45-47]。但该方法时间计算复杂度较高,不能满足实时性的需求。且在实际环境下的能见度检测需要场景中只存在天空和路面。南京大学陈启美等提出了基于视频图像对比度的检测方法^[48],解析各像素与其四邻域的对比度,所选取的最大值若大于给定阈值,即为人眼分辨像素,再结合摄像机标定换算能见度,由于阈值划分,易受噪声干扰,算法不够稳定。该团队还提出了基于路面亮度特征的能见度检测算法^[49],通过分析路面像素因大气消光系数导致的亮度变化趋势,找出其中的变化特征点,再结合摄像机标定求解大气消光系数和能见度值。与能见度的对比度算法相比较,该方法操作简单,准确率高,但仍需

相机标定技术，实时性不强。

本文提出一种基于单幅图像的能见度计算方法。能见度由场景大气消光系数唯一确定，而大气消光系数主要受空气分子及气溶胶粒子影响，其中大尺度气溶胶粒子是影响大气消光系数的最主要因素。基于对大气粒子分布规律的观测，本文提出了一种假想：理想大气（由空气分子及少量天然气溶胶构成）消光系数描述地球表面空气粒子分布情况，近似为恒定值，用 β_0 表示。本文将对场景图像大气消光系数 β 的计算分为两部分：计算 β_0 和 β / β_0 。 β_0 可通过大气辐射学理论计算得到； β / β_0 可通过图像去雾算法分别计算理想大气伪深度信息 βd 和当前大气伪深度信息 $\beta_0 d$ 并做比值得到。计算 β_0 与 β / β_0 的乘积可得场景图像大气消光系数 β ，进而代入能见度计算公式得到场景图像能见度值。本算法能对不同天气下的场景图像能见度 V_{met} 进行估计， V_{met}^{ob} 为人眼能见度值。实验结果表明，该方法估计结果与人眼观测结果基本一致，不仅简单有效，且无需人工干预及仪器辅助，适用于任意场景，有较强的创新性和实用性。

3.1.1 计算 β_0

大气主要由空气分子及气溶胶粒子构成。其中大尺度气溶胶粒子是影响能见度的最主要因素，如工业生产活动形成的沙尘气溶胶，化石燃料燃烧释放的碳气溶胶及森林火灾产生的烟雾等。一般情况下，地球大气层分布复杂，粒子种类多且粒子分布多是无规律变化，大气消光系数很难通过公式直接计算得到。而理想天气即无雾无霾晴朗天气下大气粒子主要由空气分子及少量天然气溶胶构成^[50]，粒子散射以瑞利散射为主。由于地球表面任意地点在任意时刻下空气分子的组成成分基本不变。天然气溶胶即使受地形、风向等方面的影响而发生变化，也因占比小基本不影响大气粒子的整体组成^[51]。基于此，本文提出理想大气消光系数恒常性假设：理想大气（由空气分子及少量天然气溶胶构成）消光系数近似为恒定值，并且该值可依据大气辐射学理论计算得到。

为了计算理想大气消光系数，本文引入能够表征大气层对竖直方向射入的太阳辐射削弱程度的物理量——大气透过率^[52]，其计算公式如式（3-1）：

$$T(\lambda, L) = e^{-\int_0^L \beta(\lambda, l) dl} \quad (3-1)$$

式中 $\lambda = 0.55 \mu\text{m}$ 。 L 为太阳辐射路径长度。大气消光系数 β 由两部分构成：空气分子消光系数 β_R 及气溶胶消光系数 β_a ，如公式（3-2）

$$\beta = \beta_R + \beta_a \quad (3-2)$$

式中瑞利散射消光系数 β_R 仅由空气分子决定，一般情况下为常值^[53]： 1.159×10^{-5}

($0.55\mu\text{m}$ 处, 单位为 m^{-1})。因此, 为了得到理想大气消光系数 β_0 , 我们仅需计算理想大气气溶胶消光系数 β_{α_0} 。将公式 (3-2) 代入大气透过率计算公式可得:

$$T(L) = e^{-\int_0^L \beta_R(l)dl} \times e^{-\int_0^L \beta_\alpha(l)dl} = T_R(L) \times T_\alpha(L) \quad (3-3)$$

由公式 (3-3) 可知, 大气透过率由两部分组成: 空气分子透过率 $T_R(L)$ 和气溶胶透过率 $T_\alpha(L)$ 。为了计算理想大气气溶胶消光系数 β_{α_0} , 我们引入理想大气假设条件对 $T_\alpha(L)$ 进行简化:

(1) 假定 $H=10^4\text{m}$ 为大气层的有效高度。这是因为 $H=10^4\text{m}$ 为对流层高度平均值, 包含 80% 的大气粒子以及几乎所有的水汽, 且在大于 10^4m 的大气区间内, 大气粒子数密度系数 $N(z)$ 与气象视距无关^[54];

(2) 假定理想大气消光系数 β_0 随高度不变。这是因为无云或逆温层 (即理想大气) 时, 云层、大气对流、湍流等活动对大气粒子分布的影响可以忽略不计^[55]。基于理想大气假设条件, 我们得到简化后的理想大气气溶胶透过率计算公式:

$$T_\alpha(L) = e^{-\beta_{\alpha_0} L} \quad (3-4)$$

即

$$\beta_{\alpha_0} = -\frac{\log(T_{\alpha_0})}{L} \quad (3-5)$$

式中 $L = mH$, $H=10^4\text{m}$ 。 m 是相对光学质量^[56], 即来自天体的光线穿过大气层到达海平面的路径长度除以整层大气的垂直距离, 如图 3-1 所示:

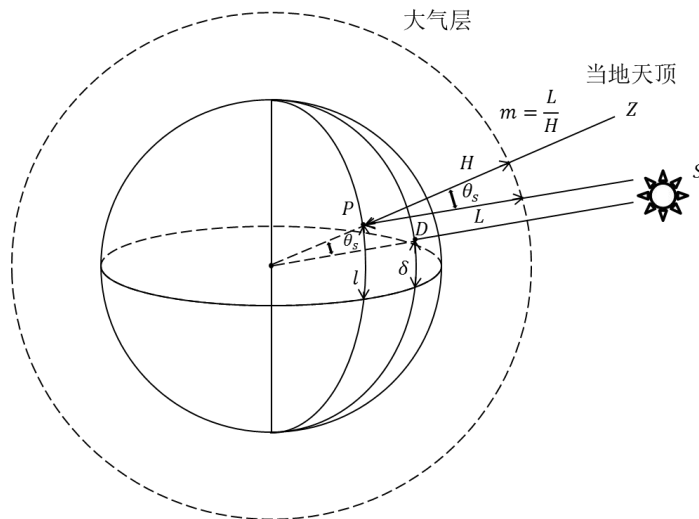


图 3-1 相对光学质量示意图

图 3-1 中, P 和 D 分别是观测点及日下点 (太阳正下方的点), l 为观测点的纬度, δ 为太阳赤纬度, θ_s 为太阳天顶角。相对光学质量 m 由 θ_s 唯一确定, 如公式

(3-6)

$$m = \frac{1}{\cos \theta_s + 0.15 * (93.885 - \theta_s \frac{180}{\pi})^{-1.253}} \quad (3-6)$$

太阳天顶角 θ_s 表征了太阳与地球的位置关系, 可由场景拍摄信息计算得到^[57], 如公式(3-7)-(3-9):

$$t = t_s + 0.170 \sin(\frac{4\pi(J-80)}{373}) - 0.129 \sin(\frac{2\pi(J-8)}{355}) + \frac{12(SM-L)}{\pi} \quad (3-7)$$

$$\delta = 0.4093 \sin(\frac{2\pi(J-81)}{368}) \quad (3-8)$$

$$\theta_s = \frac{\pi}{2} - \arcsin(\sin l \sin \delta - \cos l \cos \delta \cos \frac{\pi t}{12}) \quad (3-9)$$

式中 t_s 为十进制观测时间, J 为儒略日期, L 和 l 分别为观测点的经纬度, SM 为观测点所在时区弧度。

为计算理想大气溶胶消光系数 β_{α_0} , 本文还需计算 T_{α_0} 。Angstrom^[58]提出理想大气气溶胶透过率计算公式如式(3-10)所示:

$$T_{\alpha_0} = e^{-\beta_A m \lambda^{-a}} \quad (3-10)$$

式中 β_A 为 Angstrom 浑浊度系数, 表征大气浑浊程度; a 为波长指数, 表示气溶胶粒子的尺度分布^[59]。由于人类活动的影响, 对流层气溶胶已经有了显著增加, 即使是理想大气能见度也无法达到理想程度。通过反复试验证实, 目前环境下理想大气 Angstrom 浑浊度系数 β_A 约为 0.09, 波长指数 a 约为 3.7。将理想大气气溶胶透过率计算公式代入式(3-5)得气溶胶消光系数计算公式(3-11):

$$\beta_{\alpha_0} = \frac{\beta_A m \lambda^{-a}}{L} \quad (3-11)$$

由于 $L = mH$, 将其代入公式(3-11)得

$$\beta_{\alpha_0} = \frac{\beta_A \lambda^{-a}}{H} \quad (3-12)$$

由公式(3-12)可得, 当令 $\lambda=0.55$ 、 $H=10^4$ 、 $\beta_A=0.09$ 、 $a=3.7$ 时可得理想大气气溶胶消光系数 β_{α_0} 恒等于 8.22×10^{-5} , 单位为 m^{-1} 。则依据公式(3-2)得理想大气消光系数为: $\beta_0 = 9.379 \times 10^{-5}$, 单位为 m^{-1} 。这也从侧面证实理想大气消光系数恒常性假设成立。

3.1.2 计算 β/β_0

场景图像可由大气散射模型表示（如图 3-2 所示），到达相机的光线由两部分组成：（1）直接传输项，物体表面光经过大气颗粒物散射后到达相机，其强度随着目标距离增大呈指数衰减；（2）环境光项，环境中的自然光（如太阳和天空等）经过大气颗粒物散射，有部分会到达相机，其强度随着距离增加而增大。因此，在相机处观察到的图像亮度可以表示为：

$$I = I_{\infty}\rho e^{-\beta(d)d} + I_{\infty}(1 - e^{-\beta(d)d}) \quad (3-13)$$

其中 ρ 为场景反照率， I_{∞} 为环境光强度，令 $J = I_{\infty}\rho$ ， J 表示场景固有亮度值； β 表示距离为 d 处的消光系数； $e^{-\beta(d)d}$ 为距离 d 处的透射率，通常用 $t(\beta, d)$ 表示， βd 为距离 d 处的伪深度信息。

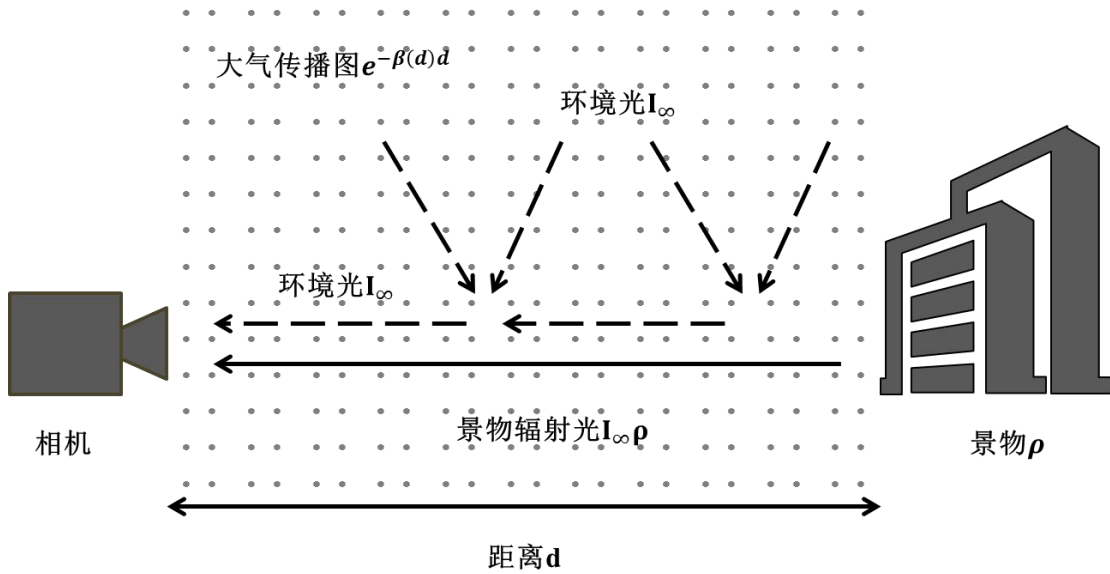


图 3-2 大气散射模型示意图

如果假设拍摄范围内大气粒子分布均匀，即 β 恒定不变，则公式（3-13）可简化为：

$$I = J e^{-\beta d} + I_{\infty}(1 - e^{-\beta d}) \quad (3-14)$$

将场景图像划分为以 x 为中心的若干区域， $x = 1, 2, 3 \dots n$ 。对每个区域的大气散射模型取最小运算符，并对 RGB 三通道使用最小运算符得：

$$\min_{c \in \{r, g, b\}} \left(\min_{y \in \Omega(x)} (I^c(y)) \right) = t(x) \min_{c \in \{r, g, b\}} \left(\min_{y \in \Omega(x)} (J^c(y)) \right) + I_{\infty}^c(1 - t(x)) \quad (3-15)$$

式中, $\Omega(x)$ 是以 x 为中心的区域, $x=1,2,3\cdots n$ 。 $I^c(y)$ 为目标区域内像素点 y 在 c 通道的像素值, $J^c(y)$ 为目标区域内像素点 y 在 c 通道的固有亮度值, I_∞^c 为 c 通道的大气光值, 令 $J^{dark}(x) = \min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} (I^c(y)))$ 记为 I 的暗通道图。由暗原色先验假设可知: 任意户外无雾图像总是存在某些像素, 其三通道中至少有一个通道的像素值非常小, 即:

$$\min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} (J^c(y))) \approx 0 \quad (3-16)$$

当景物颜色与天空颜色接近且没有阴影时, 暗原色先验假设不成立。因此, 为了得到该假设成立的区域, 本文在公式 (3-15) 的基础上再取一次最小运算符得到暗通道值最小区域 $\Omega_{\min}(x)$, 如公式 (3-17) 所示:

$$\min_{\Omega(x) \subset I} (\min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} (I^c(y)))) = t(x) \min_{\Omega(x) \subset I} (\min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} (J^c(y)))) + I_\infty^c (1 - t(x)) \quad (3-17)$$

由于满足 $\min_{\Omega(x) \in J} (\min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} (J^c(y))))$ 的区域 $\Omega_{\min}(x)$ 对应的无雾图像像素近似满足公式 (3-16), 表明该区域暗原色先验假设成立。因此可作为计算大气传播图的目标区域, 如图 3-3 所示。

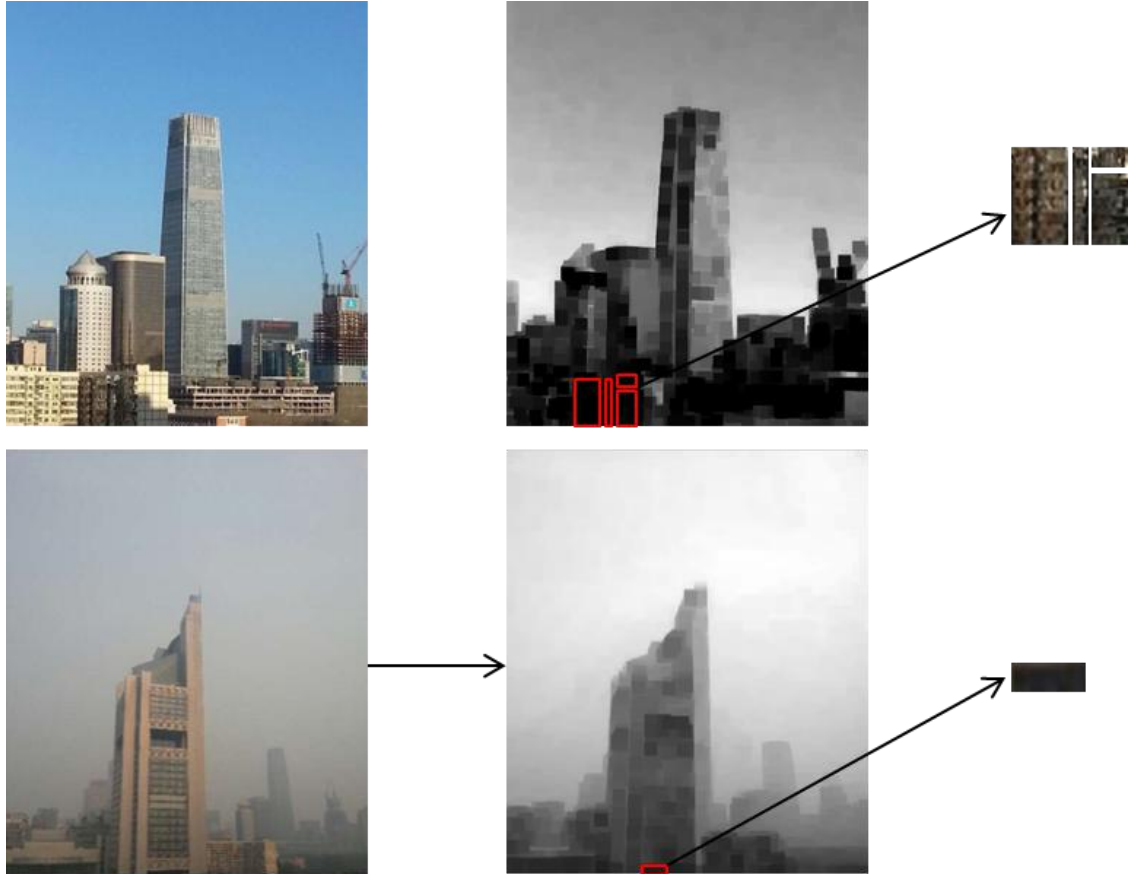


图 3-3 目标区域选取

得到目标区域后, 我们利用暗原色先验算法计算目标区域的透射率。分别令 $t(x)$ 为当前场景目标区域透射率, $t_0(x)$ 为理想大气透射率。透射率计算公式如

(3-18) 所示

$$\bar{t}(x) = 1 - \omega \min_{\Omega_{\min}(x) \subset I} \min_{c \in \{r, g, b\}} \left(\min_{y \in \Omega_{\min}(x)} \left(\frac{I^c(y)}{I_{\infty}^c} \right) \right) \quad (3-18)$$

$I^c(y)$ 为目标区域在 c 通道的像素值, I_{∞}^c 为 c 通道的环境光值。这里引入 ω 是为了保留一部分覆盖遥远景物的雾, 保留图像深度感, 使图像看起来真实。最后应用引导滤波器^[60]对 $\bar{t}(x)$ 进行完善。

由于场景图像易受噪声点干扰, 且场景不同深度交界处大气传播图易发生突变, 使得大气传播图准确性受到影响。为此, 本文提出一种区域收缩算法, 将目标区域内每一点的 $t(y)$ 与区域内大气传播图最大值 $t_{\max}(x)$ 进行差分, 并将差分结果与给定阈值进行比较, 保留差分结果小于给定阈值的子区域, 计算该子区域大气传播图的平均值作为改进后的大气传播图 $t(x)$ 。

为了计算理想大气透射率 $t_0(x)$, 我们需要对图像目标区域进行去雾操作, 如公式 (3-19) :

$$J(x) = \frac{I(x) - I_{\infty}}{t(x)} + I_{\infty} \quad (3-19)$$

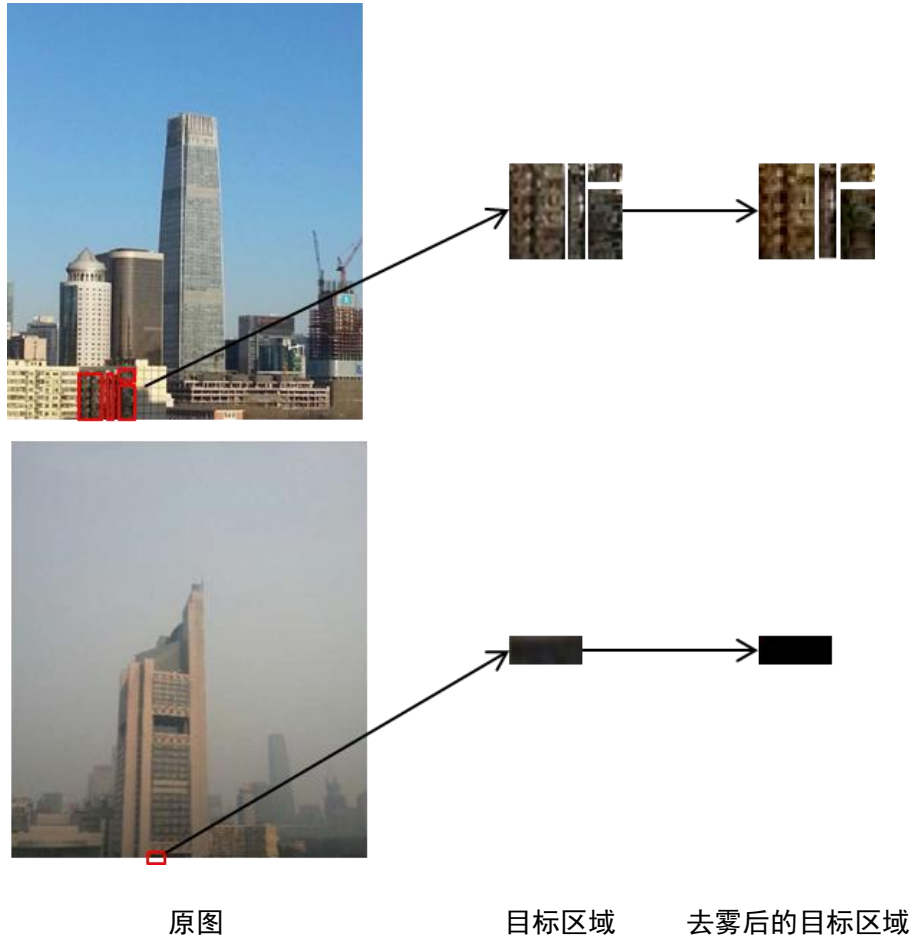


图 3-4 目标区域去雾

将目标区域的无雾亮度值分布 $J(x)$ 近似为理想天气场景亮度值分布 I_0 ，再次利用传播图计算方法得到理想大气透射率 $t_0(x)$ 。为了得到 β / β_0 ，通过 $t(x)$ 和 $t_0(x)$ 分别求得伪深度信息 βd 和 $\beta_0 d$ 并计算比值得：

$$\frac{\beta}{\beta_0} = \frac{\ln(t(x))}{\ln(t_0(x))} \quad (3-20)$$

3.1.3 计算能见度 V_{met}

将理想大气消光系数常值 β_0 和比值 β / β_0 相乘得到场景大气消光系数 β 。世界气象组织(WMO)基于气象光学视程 V [61] (Meteorological Optical Range, MOR) 提出能见度计算公式。光学视程计算公式如公式 (3-21) 所示

$$V = \frac{1}{\beta} \ln \frac{1}{\varepsilon} \quad (3-21)$$

式中 ε 是对比度域值, 等于 2%, 以人眼最敏感的波长 $0.55\mu\text{m}$ (绿光) 为依据。基于公式 (3-21) 可计算得到能见度 V_{met}

$$V_{met} \approx \frac{V}{3} \quad (3-22)$$

由公式 (3-21) 和 (3-22) 可知, 利用大气消光系数 β 可计算得能见度 V_{met} 。

3.2 实验设计与结果分析

3.2.1 能见度估计结果分析

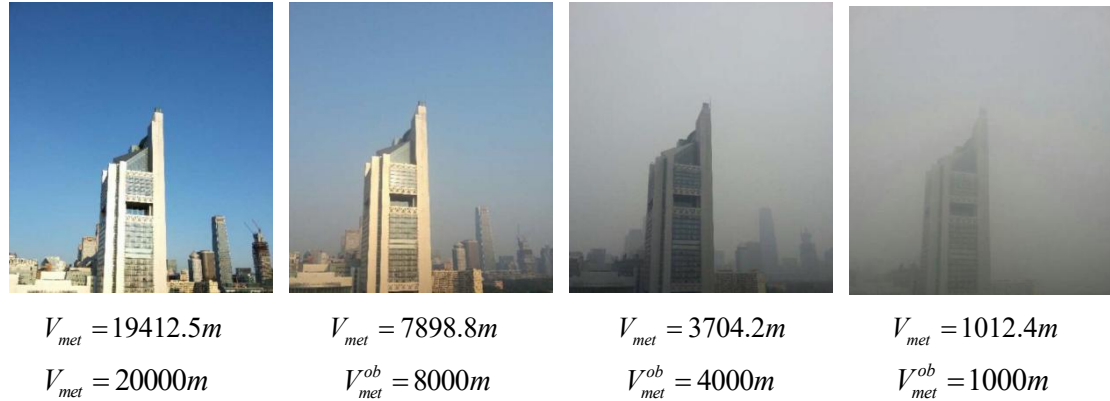
本节首先通过实验验证基于单一场景图像计算能见度方法的有效性。

本文选取三个主要场景, 分别是: 城市场景、公路场景及自然场景。通过对场景任意角度进行拍摄, 利用本文算法估算能见度。同时, 为了证明本文算法的准确性, 本文引入能见度仪测量值 V_{met}^{ob} 作为算法结果的评判标准。

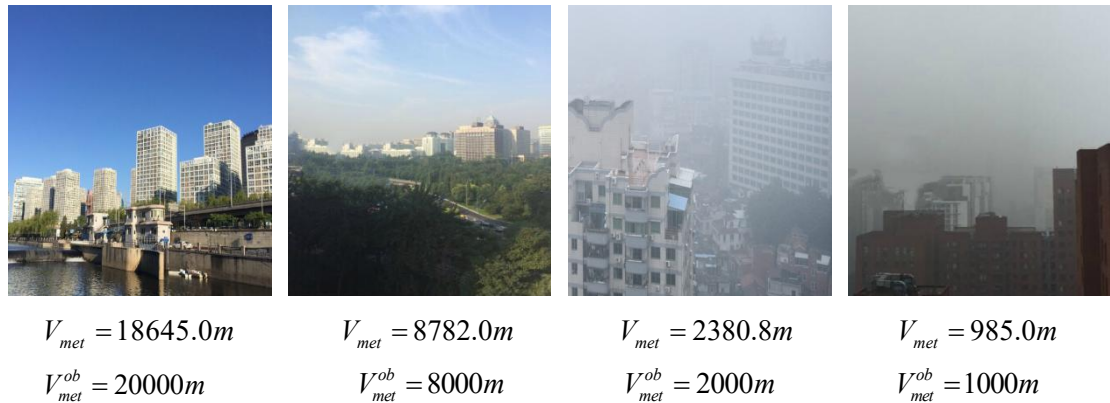
图 3-5 针对城市场景进行拍摄。图 3-5a 为固定场景分别在高能见度、较高能见度、较低能见度与低能见度 4 类天气下的能见度估计结果, 图中前一栋楼为北京市中央电视台, 后一栋楼为中国国际贸易中心第三期, 两栋楼相距约 1000m。图 3-5b 为任意城市场景分别在高能见度、较高能见度、较低能见度与低能见度 4 类天气下的能见度估计结果。图 3-6 为任意公路场景分别在高能见度、较高能见度、较低能见度与低能见度 4 类天气下的能见度估计结果。图 3-7 针对自然场景进行拍摄, 分别为高能见度、较高能见度、较低能见度与低能见度 4 类天气下的

能见度估计结果。实验结果表明，本文算法在任意场合均能适用，且估测结果与人眼观测能见度值基本吻合。

为验证本文算法的合理性，本文引入 Hautière N 的实验结果 V_{met}^H [44][46] 并与本文算法结果进行比对。由图 3-8 可知，本文基于单幅图像的能见度检测结果与 Hautière N 基于人工标记物^[46]及相机标定技术^[44]得到的能见度结果基本一致，这表明本文算法在保证算法准确性的基础上无需任何人工干预及仪器辅助，有较强的鲁棒性和通用性。



(a) 固定城市场景能见度估计



(b) 任意城市场景能见度估计

图 3-5 城市环境能见度估计结果

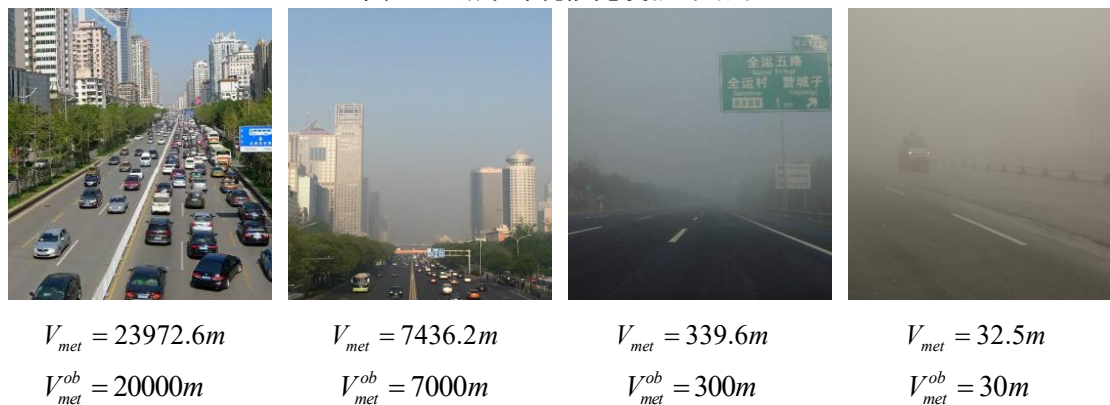


图 3-6 城市环境能见度估计结果

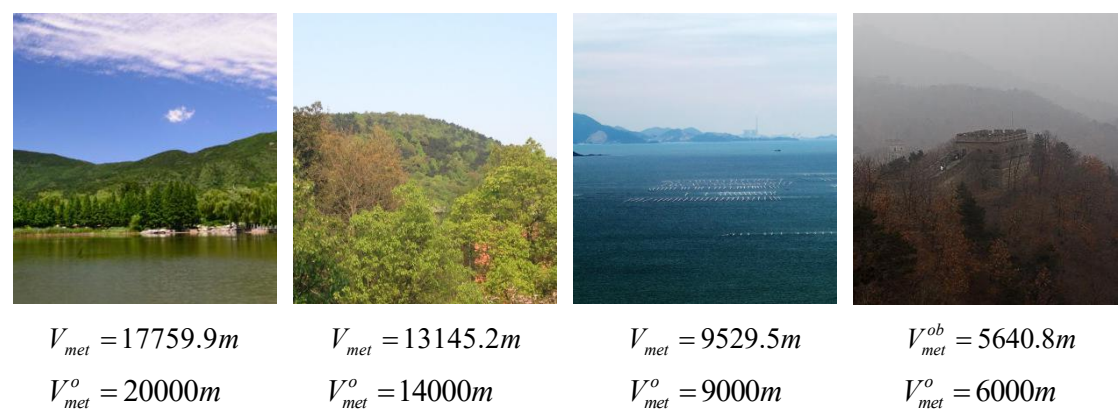


图 3-7 自然环境能见度估计结果

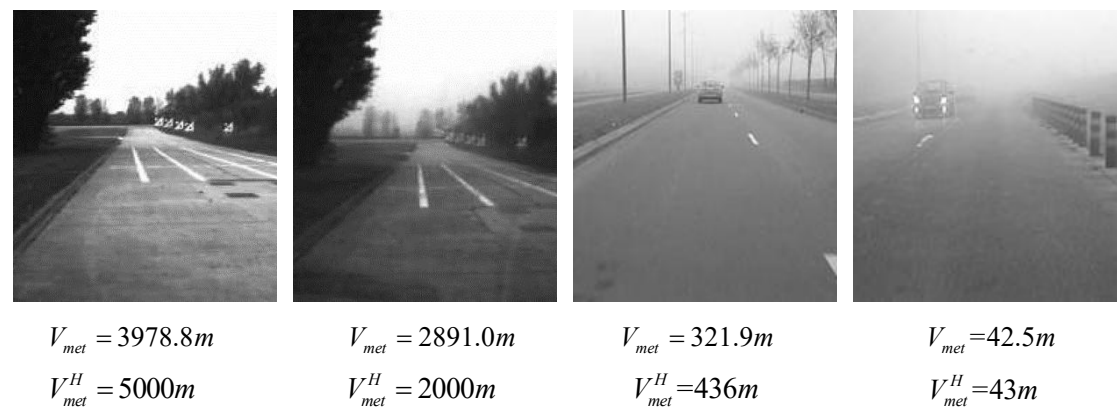


图 3-8 本文算法与 Hautiére N[44] [46]的结果对比

3.2.2 空气质量等级估计结果分析

接下来，我们基于图像能见度特征对空气质量等级进行估计。我们提取图像不同距离 ROI 并计算能见度值作为特征。

利用 VAQI-1 数据集，我们对训练库和测试库中的图像提取能见度特征并利用多分类 SVM 进行分类。表 3-1 展示了基于能见度特征的估计结果以及与基于视觉特征的分类结果比较。

表 3-1 基于能见度特征和视觉特征的空气质量等级估计结果对比

	传统特征	准确率[%]
1	能见度	58
2	SIFT	8
3	HOG	12
4	LBP	14
5	色彩饱和度	35
6	颜色直方图	20
7	对比度	42

实验结果表明,能见度特征的表现要明显优于视觉特征。这是因为能见度特征是依据大气粒子分布特性来提取的,是反映大气浑浊度的最主要指标,与待测场景的空气质量密切相关。当场景空气质量好时,远处景物能够清晰的看见,视力范围较广,能见度高;而当空气质量不好时,远处景物清晰度明显下降,视力范围较窄,能见度低。由此可见,通过能见度人们可以判断甚至直接估计当前场景的空气质量。正因如此,基于能见度特征的空气质量等级估计在这 7 类特征中表现最好。

尽管基于能见度特征可以得到接近 60%的准确度,但离理想估计精度还有一定距离。这是因为:能见度表征了大气浑浊度,在一般情况下能够描述场景空气质量。但是,能见度低的场景未必空气质量差,如图 3-9 所示。

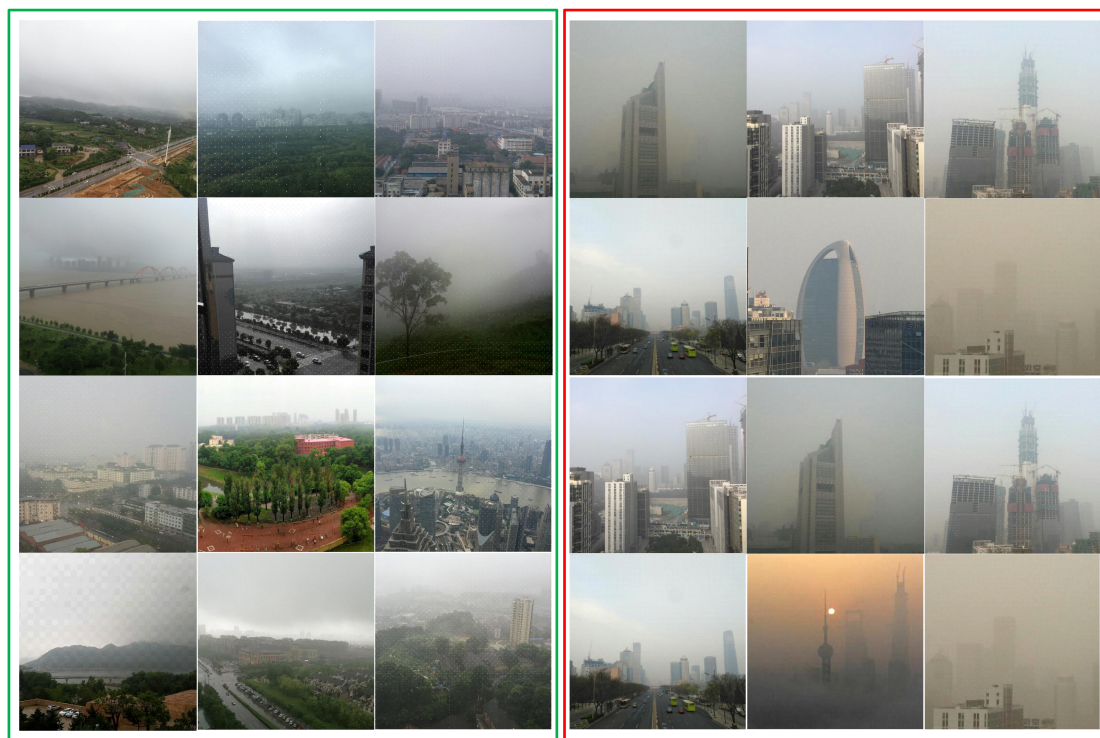


图 3-9 雾天图像(绿框)和霾天图像(红框)

图 3-9 展示了雾天与霾天的场景成像。红框中的图像为雾天图像, AQI 值均小于 50, 空气质量等级为优; 绿框中的图像为霾天图像, AQI 值均大于 200, 空气质量至少是重度污染。然而仅从图像看, 两者的能见度均比较低, 有些图像中雾天的能见度甚至比霾天的能见度还低, 但雾天的空气质量等级远远低于霾天的空气质量等级。这充分证明了仅用能见度这一特征仍无法准确描述场景的空气质量。基于此, 我们将第 2 章中的 6 类视觉特征分别于能见度特征结合, 代入分类器进行分类。实验结果如表 3-2 所示。

表 3-2 能见度与其他视觉特征结合的空气质量等级估计结果对比

	传统特征	准确率[%]
1	能见度+SIFT	8
2	能见度+HOG	13
3	能见度+LBP	14
4	能见度+色彩饱和度	65
5	能见度+颜色直方图	51
6	能见度+对比度	54

从表 3-2 可以看出，能见度与色彩饱和度的结合分类效果最好。这可以从图 3-9 进行解释。在图 3-9 中，雾天图像整体颜色较鲜艳，饱和度较高。同时由于雾气中的小水滴呈团状分布，在图像中分布不均匀，使得图像颜色变化差异大。而霾天图像颜色整体偏暗淡，饱和度较低。且霾颗粒在大气中分布比较均匀，使得场景图像颜色变化相对不大。因此，我们可以得出一个结论：场景图像的能见度特征与色彩饱和度能基本表征场景的空气质量。

3.3 本章小结

本章主要介绍了一类能够表征大气粒子分布特性的特征——能见度。能见度描述了场景最大的可视范围，对空气质量有直观反应。依靠能见度检测，人们可以对当前场景空气质量有一个粗略判断。基于此，本章提出了一类基于单幅场景图像的能见度估计方法。实验表明本算法在保证算法准确性的基础上无需任何人工干预及仪器辅助，有较强的鲁棒性和通用性。基于 VAQI-1 数据库，本章实现了基于能见度特征的空气质量等级估计。实验结果表明能见度特征估计结果要远好于视觉特征，准确率达 58%。然而，由于能见度特征无法分类极端天气如浓雾天的空气质量，所以基于浓雾天图像特性将能见度特征与颜色特征相结合，得到的分类效果要优于仅依靠能见度特征的分类效果。这说明了场景图像的能见度特征与颜色特征能基本表征场景的空气质量特性。

4 基于深度特征的空气质量等级估计

深度学习是目前机器学习领域最前沿和热门的研究课题之一，它能建立模拟人脑的神经网络，模仿人脑的思维机制来解释数据。深度学习中的卷积神经网络是目前模式识别、计算机视觉领域用到的最主流方法之一。本文采用深度学习中的卷积神经网络对场景图像进行深度解释，获取图像深度信息，并设计基于图像深度特征的空气质量等级估计方法探究其与场景空气质量的关系。

4.1 卷积神经网络

卷积神经网络（CNNs）是目前解决语音分析、图像识别等领域问题的最主要的方法。其思想来源于生物脑部神经工作原理，并在此基础上构成权值共享网络。与传统神经网络相比，卷积神经网络不仅显著减少了网络参数的个数，同时也大大降低了模型复杂度，对比例缩放、平移、转换等数据形式的变形具有高度不变性。在图像识别领域，卷积神经网络的应用避免了传统图像识别算法中特征提取、清洗、降维等过程，大大降低了人工成本。

4.1.1 卷积神经网络理论基础

卷积神经网络（CNNs）由输入层、输出层和多个隐含层构成，每个隐含层包含多个由若干个独立神经元组成的二维平面。CNNs 网络结构如图 4-1 所示。

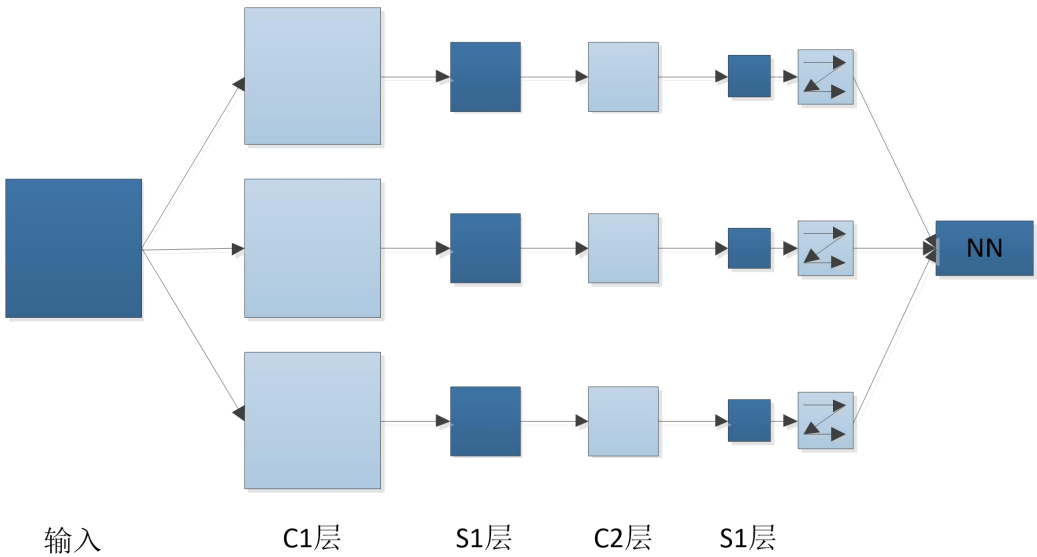


图 4-1 卷积神经网络结构

C 层为卷积层，S 层为池化层，F 层为全连接层。卷积层可看作是加入偏置的可训练的滤波器，对输入图像的像素进行加权求和和加偏置操作，即提取局部

特征。其主要功能如下：

(1) 局部感知。该功能启发于生物脑部的视觉系统结构。经研究发现，人类对外界环境的认知过程是从局部信息到全局信息，这是由于生物脑部视觉皮层中的神经元只接收外界局部信息，即生物脑部神经元只对一部分特定区域的刺激有响应。对于图像像素，邻接空间的像素相关性要比距离较远的像素相关性强。因此，对于图像识别问题，神经元仅需要对输入图像的局部区域进行感知学习，在网络后期将所有局部信息进行综合即得到输入图像的全局信息。

(2) 参数共享。局部感知操作可以适当的减少参数量，但对于深层网络来说参数量仍是巨大的。因此需要新的策略，即权值共享。权值共享的原理可表述为：图像中某部分的统计特性与其他部分一致，即这一部分的特征也能用于其他部分。因此，输入图像的各个位置均可以使用相同的学习特征进行学习。

(3) 多卷积核操作。对于图像识别任务，隐含层上只有一个卷积核显然是不够的。一般情况下都会加入多个卷积核进行卷积操作。每个卷积核的参数设置都不一样，对应的都会对图像提取不同性质的特征图。这些特征图可以看作是一张图像在不同通道的表征形式。

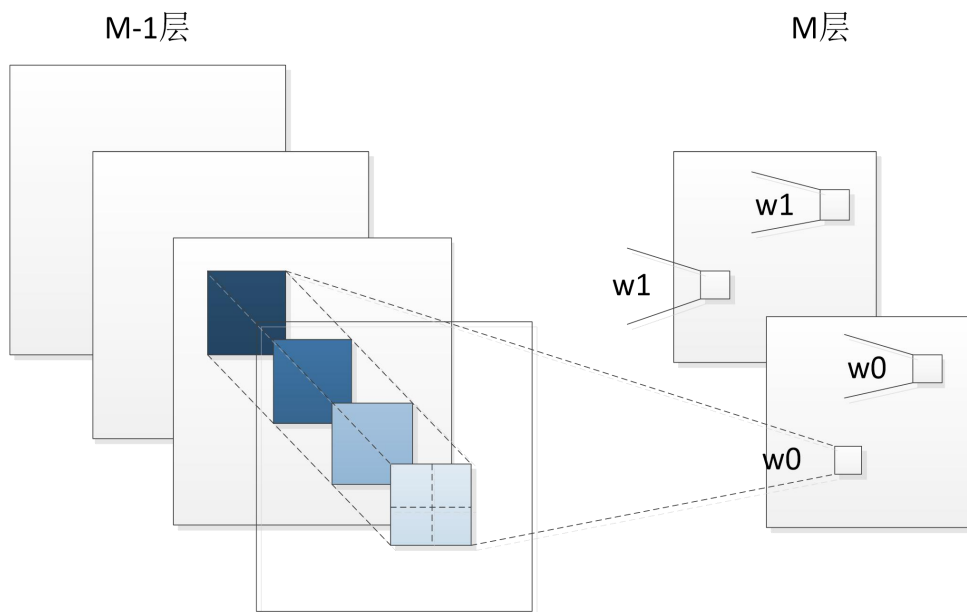


图 4-2 多卷积核

图 4-2 中描述了从 $m-1$ 层到 m 层的卷积操作。 $m-1$ 层有 4 个通道， m 层有两个卷积核，生成两个通道。需要强调的是，每个卷积核均会对 4 个通道进行卷积，卷积得到的结果是由四个通道的卷积结果相加再进行激活得到的。

池化层主要对卷积过程得到的特征进行聚合统计，其原理如下：当得到卷积特征时，我们需要利用这个特征完成具体的图像识别任务。但此时的卷积特征维度依然很大，用分类器进行分类时仍存在巨大的计算量，且极易出现过拟合。研

究人员发现,既然图像中某部分的统计特征可用于其他部分,那么某个区域的不同部分特征也应该可以通过某个具有聚合特性的特征进行替代,这样的操作既可以得到较低维度的特征,也可以有效防止过拟合。一般的池化操作是提取区域内所有特征的最大值、最小值和平均值作为具有聚合特性的特征。

全连接层用来计算输入特征与权重之间的点积和偏置,并通过激活函数计算输出值并进行图像识别任务。

接下来,本文列举几个常用的 CNN 模型。

4.1.2 LeNet 模型

LeNet 模型^[62]是早期卷积神经网络模型中最具代表性之一的模型,由 Yann LeCun 于 1998 年提出,被誉为卷积神经网络领域的开山之作。LeNet 具有较小的网络结构,但基本具备卷积神经网络的基本组件,即卷积层、池化层和全连接层。后续卷积神经网络的提出也是以 LeNet 为基础搭建。本文基于 LeNet 模型的深度特征提取过程如图 4-3 所示。



图 4-3 基于 LeNet 模型的深度特征提取过程

LeNet 共八层(包括输入层),每一层包含多个卷积核,其具体结构如下:

第一层:设定场景图像输入大小为 32×32 。固定输入层大小的目的是将图像潜在的显性特征能够集中于深度特征感受野中心。

第二层:C1 层是第一层卷积层,包含 6 个特征图(feature map), 5×5 大小的卷积核,故每个特征图有 $(32-5+1) \times (32-5+1)$ 个,即 28×28 个神经元。因为每个卷积核都与输入图像的 5×5 大小的区域相连,因此 C1 层有 $(5 \times 5+1) \times 6 = 156$ 个训练参数。输入层与 C1 之间的连接数为 $156 \times (28 \times 28) = 122304$ 个。第一层卷积

操作提取了输入图像的第一层特征，并且不同的卷积核提取到的特征不一样。

第三层：S2 层是以第一个下采样层。本文选取最大值池化进行下采样操作。下采样层包含 6 个特征图，每个特征图的大小为 14×14 。每个特征图与 C1 层每个特征图的 2×2 区域相连。LeNet 模型中的下采样层包含参数，即 S2 层中每个神经元是由 S2 层的 2×2 感受野中的 4 个输入相加再乘以一个参数并加上偏置最后通过 sigmoid 激活函数计算得到。S2 层有 $2 \times 6 = 12$ 个参数。C2 与 S2 之间的连接数为 $(4+1) \times (14 \times 14) \times 6 = 5880$ 个。通过下采样操作可以有效降低训练参数的个数，防止过拟合。

第四层：C3 是第二层卷积层，包含 16 个特征图， 5×5 大小的卷积核，每个特征图有 $(14-5+1) \times (14-5+1)$ 个，即 10×10 个神经元。通过计算得到 C3 层的参数个数为 $(5 \times 5 \times 3 + 1) \times 6 + (5 \times 5 \times 4 + 1) \times 9 + (5 \times 5 \times 6 + 1) \times 1 = 1516$ 个。

第五层：S4 层是第二层下采样层，包含 16 个特征图，每个特征图的大小为 5×5 。与 S2 类似，每个神经元与上一层感受野中 2×2 区域相连。S4 层有 $2 \times 16 = 32$ 个参数。

第六层：C5 是第三层卷积层，包含 120 个特征图， 5×5 大小的卷积核，每个特征图有 $(5-5+1) \times (5-5+1)$ 个，即 1×1 个神经元。通过计算得到参数个数为 48120 个。

第七层：F6 全连接层，包含 84 个特征图，每个特征图中仅有一个神经元与前一层相连，参数个数为 $(1 \times 1 \times 120 + 1) \times 84 = 10164$ 个参数。

第八层：输出层针对具体图像分识别任务设定神经元个数，采用径向基函数（RBF）的网络连接方式连接。

4.1.3 AlexNet 模型

AlexNet 模型^[63]由 Geoffrey 和 Alex 于 2012 年提出，并且在 2012 年 ILSVRC 竞赛（Large Scene Visual Recognition Challenge）中拿到冠军。ILSVRC 竞赛是全世界最权威的计算机视觉竞赛。AlexNet 也一举奠定了深度学习在计算机视觉领域的地位。

本文基于 AlexNet 模型的深度特征提取过程如图 4-4 所示，共有 8 层，约 60M 以上的参数量，前五层是卷积层和池化层的组合，后三层为全连接层，最后紧跟一个全连接层进行分类。



图 4-4 基于 AlexNet 模型的深度特征提取过程

AlexNet 实现过程与 LeNet 相似，但包含以下几点创新：

(1) AlexNet 输入图像尺寸更改为 224×224 ，与 32×32 相比包含更多的细节信息；

(2) 数据扩充。一般情况下，训练数据的增加能一定程度上提升算法的准确率。现有数据量有限，且没有有效渠道增加数据量。基于此，本文提出三种图像数据变换方法对数据量进行扩充。三种方法包括镜像变换、平移变换及光照和色彩变换。镜像变换是对图像进行镜像操作。平移变换指对平移提取图像左上、右上、左下、右下和中间 5 个部分的图像。光照和彩色变换是指对图像的颜色进行随机变换。

(3) 局部响应归一化。局部响应归一化可以帮助模型进行泛化，降低错误率。

(4) 重叠池化。令相邻池化窗口存在重叠区域，这使得模型效果有所提高。

(5) 避免过拟合。AlexNet 提出一个可以避免过拟合的模型组合——Dropout。

每输入一个训练样本，网络中进行“Dropout”的神经元不参加模型训练，相当于网络尝试了新的结构。不同的训练样本对应一个结构，并且这些结构之间共享权重信息，使得神经元复杂的互适应关系降低，避免了过拟合。

4.1.4 VGG 模型

VGG^[64]是由牛津大学 VGG（Visual Geometry Group）视觉几何组于 2014 年提出的卷积神经网络结构，并在 2014 年 ILSVRC 竞赛中夺得亚军。VGG 网络有 6 种结构，本文分别选取结构 D 和结构 E 提取场景图像深度特征(如图 4-5 所示)。



图 4-5 基于 VGG 结构 D 和结构 E 提取场景图像深度特征

其实现过程与 AlexNet 类似输入图像大小为 224×224，均有 5 个卷积层和

对应的池化层、2 个全连接层和一个分类层，且包含数据扩充、局部响应归一化等操作。不同点在于：VGG 在第一个卷积层使用了更小的 3×3 卷积核，这是因为尺寸为 3×3 的卷积核是能够捕获感受野位置特征的最小的卷积核尺寸，不仅能够有效减少参数量，而且比大尺寸卷积核有更强的非线性，是的判决函数可靠性更高。

4.1.5 GoogLeNet 模型

GoogLeNet^[65]是由谷歌于 2014 年提出，在当年的 ILSVRC 比赛中拿到冠军。该模型证明了卷积个数的增多以及网络层数的加深能有效提升模型的预测能力。

随着深度学习领域的快速发展，训练数据增多，网络模型也越来越深，同时对硬件设备性能要求也越来越高。但这一现象存在两个问题：（1）参数量急剧增长，尤其对于训练数据有限的情况极易出现过拟合；（2）计算量增加，同时消耗了更多计算资源。为解决以上两个缺陷，最根本的方法是将拥有大量参数的全连接层及包含大尺度卷积核的卷积层转化为稀疏连接。生物脑部的神经系统本身是稀疏连接，这样做可以更好的模仿生物脑部运作。同时，大规模的稀疏神经网络能够利用激活函数值的统计特性和对输出进行聚类操作得到最优的神经网络。这一观点表明复杂的稀疏神经网络可以通过统计和聚类方法进行简化，同时保证其最优性能。基于此，GoogLeNet 网络提出了 Inception 结构来解决此问题。

Inception 结构的提出是为了通过密集成分找出最优局部稀疏结构，如图 4-6。

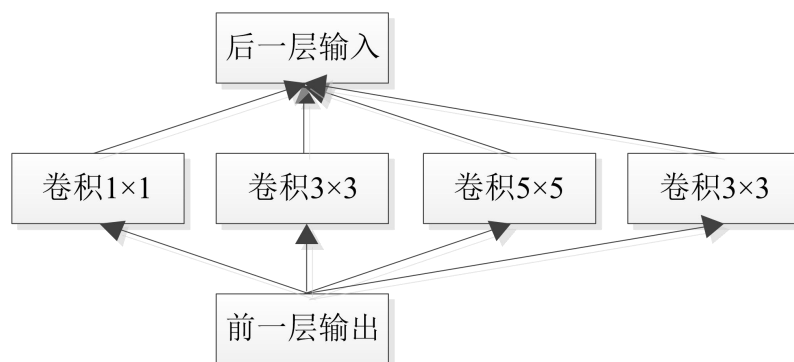


图 4-6 Inception 模型 A

如图所示，Inception 结构是一个层与层的结构，最后一层进行了聚合统计。这个结构与上一层的结构单元相连，并形成下一层的结构单元。在每一个结构单元中，卷积核大小不一，对图像的感受范围也不一样，使得得到的特征表征程度也不同，最后的聚合也意味着对不同程度的特征进行统计。

然而，越后期的网络模型提取的特征就越抽象，卷积核的使用比例也增加。在上述 Inception 模型 A 中，卷积核尺度为 5×5 和 3×3 ，这一级别的卷积核尺度

在模型后期会带来巨大的计算量。基于此，必须采用尺度为1×1的卷积核进行降维（如图 4-7）

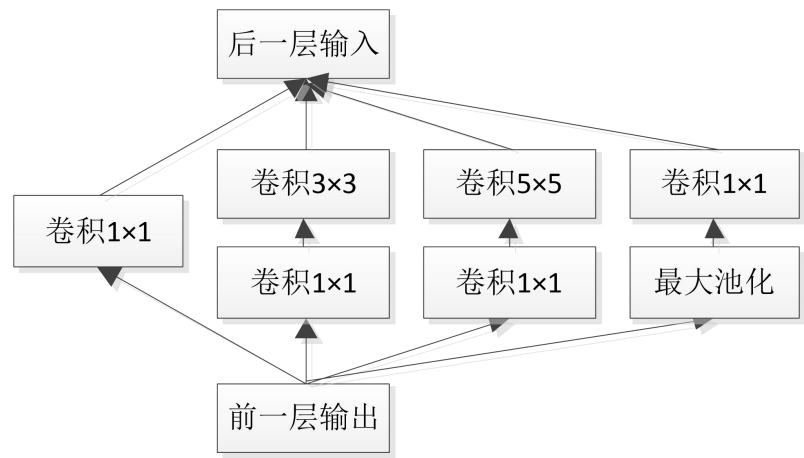


图 4-7 Inception 模型 B

本文采用 Inception 模型 B 设计 GoogLeNet 模型提取场景图像深度特征（如图 4-8 所示）。本方法同样用到数据扩充的三种方法，具体过程参照 4.1.3 节。



图 4-8 基于 GoogLetNet 模型的深度特征提取过程

GoogLeNet 结构具有以下特点：

- (1) 采用 Inception 模块化结构，显著减少参数量，提升模型预测能力；
- (2) 网络末端用平均池化层代替全连接层，提高了模型精度；
- (3) 网络将超大参数量的全连接层移除，保留了 Dropout 层，有效避免过拟合。

4.1.6 ResNet 模型

ResNet^[66]是由 MARA 何凯明团队于 2015 年提出，并在 2015 年的 ILSVRC 比赛中获得冠军。ResNet 网络的提出是为了解决深度学习的根本问题之一：梯度退化问题。梯度退化问题指当网络层数加深时，梯度特征退化使得分类准确率反而下降。造成这一问题的原因是当模型层数加深时，梯度下降法对参数的优化变的困难，无法学习到好的效果。

基于上述问题，作者提出 Residual 结构（如图 4-9 所示）

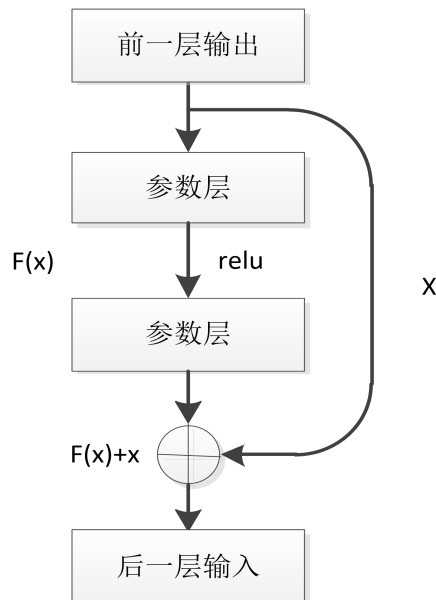


图 4-9 Residual 结构

Residual 结构提出增加一个恒等映射结构，即对于每一层卷积层的输出再加上该层的输入，即输出由 $F(x)$ 变为 $F(x)+x$ 。这一改进不改变特征提取的效果，不额外增加参数量，但大大减小了梯度的退化，提高了模型训练速度。

本文采用 50 层的 ResNet 模型提取场景图像空气质量深度特征（如图 4-10），输入图像大小为 224×224 ，同样用到数据扩充的三种方法，具体过程参照 4.1.3 节。



图 4-10 基于 ReNet 模型的深度特征提取过程

4.2 实验设计与结果分析

基于场景图像深度特征的空气质量等级估计算法设计如下：利用 4.1 节的各类卷积神经网络提取输入数据的深度特征，并采用 SVM 分类器进行空气质量等级分类。实验框架如图 4-11 所示

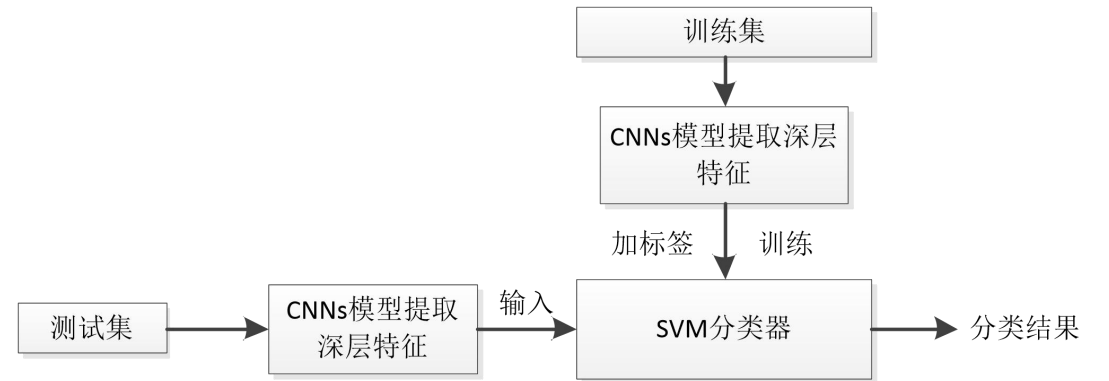


图 4-11 基于场景图像深度特征的空气质量等级估计

在特征提取之前，需设定各个网络的训练参数。训练参数设定如下：

- (1) 测试迭代次数 (test_iter) : 2
- (2) 测试间隔 (test_interval) : 50
- (3) 基础学习率 (base_lr) : 0.001
- (4) 屏幕日志显示间隔 (display) : 20
- (5) 最大迭代次数 (max_iter) : 500
- (6) 学习率变化规律 (lr_policy) : "step"
- (7) 学习率变化指数 (gamma) : 0.1
- (8) 动量 (momentum) : 0.9
- (9) 权值衰减 (weight_decay) : 0.005
- (10) 学习率变化频率 (stepsize) : 100
- (11) 是否使用 GPU (solver_mode) : GPU
- (12) 优化算法 (solver_type) : SGD

本次实验的硬件环境为: intel 酷睿 i7 处理器, 一个 NVIDIA GTX 1070 显卡 8G 显存、16GB 内存。

我们将 VAQI-1 数据库分别代入各类卷积神经网络模型提取深度特征, 并利用 SVM 分类器进行空气质量等级分类, 分类结果如表 4-1 所示

表 4-1 基于场景图像深度特征的空气质量等级估计结果对比

	卷积神经网络模型	准确率[%]
1	LeNet 模型	37
2	AlexNet 模型	32
3	VGG16 模型	41
4	VGG19 模型	42
5	GoogLeNet 模型	68
6	ResNet 模型	61

从结果可以看出, 基于前四类卷积神经网络模型得到的深度特征分类精度均不理想。这说明这 4 类模型无法利用场景图像学习空气质量特征。基于 GoogLeNet 模型提取的深度特征的估计准确率最高, 达 68%。这一结果表明 GoogLeNet 能基本学习到场景空气质量特征。

考虑到 VAQI-1 为小样本数据集, 可能会对上述 6 类卷积神经网络模型的参数学习过程带来较大的局限, 导致过拟合。基于此, 本文引入 ImageNet 数据集^[67]对模型进行预训练, 再利用 VAQI-1 对模型进行 fine-tuning, 实现空气质量等级分类, 如表 4-2 所示

表 4-2 基于 fine-tuned 模型的深度特征估计空气质量等级结果对比

	卷积神经网络模型	准确率[%]
1	fine-tuned LeNet 模型	52
2	fine-tuned AlexNet 模型	54
3	fine-tuned VGG16 模型	65
4	fine-tuned VGG19 模型	67
5	fine-tuned GoogLeNet 模型	76
6	fine-tuned ResNet 模型	70

从实验结果可以看出，基于微调后的深度特征的空气质量等级估计结果整体比没有微调的深度特征估计结果好，这证明了 fine-tuning 对于小样本数据模式识别问题有较大的帮助。不仅如此，基于 Fine-turned GoogLeNet 模型提取的深度特征的分类准确率达 76%。这一结果充分证明了利用以空气质量为标签的场景图像数据能够让卷积神经网络学习到表征图像空气质量的深度特征。不同于边缘、纹理、颜色等停留于表面且人眼很容易识别的浅层特征，这种深度特征具有更深层的含义，同时贴合了空气质量成像机理。这类特征可类似于拥有一定场景空气质量观测经验的人的视觉感受，是一般没有观测经验的人无法学习到的。

为了更加深入的分析深度特征的分类性能，我们统计了基于 fine-tuned GoogLeNet 模型的空气质量等级分类的混淆矩阵，如表 4-3 所示

表 4-3 基于 fine-tuned GoogLeNet 模型的空气质量等级分类的混淆矩阵

		预测空气质量等级					
		I	II	III	IV	V	VI
实际空 气质量 等级	I	92%	8%	0%	0%	0%	0%
	II	6%	86%	4%	4%	0%	0%
	III	3%	12%	41%	32%	7%	5%
	IV	1%	2%	15%	65%	9%	8%
	V	2%	0%	2%	20%	73%	3%
	VI	0%	0%	0%	2%	7%	90%

从混淆矩阵可以看出，第一类（优）和第六类（重度污染）的分类效果最好，分别为 92%和 90%。这是因为这两场景图像属于空气质量分布的两种极限状况，特征可区分性最高。第三类（轻度污染）、第四类（中度污染）和第五类（中度污染）的分类效果相对来说不理想，分别为 41%、65%和 45%。这是因为这类场景

空气质量变化特征不明显。不同等级的空气质量的变化程度要小于同种等级不同场景图像的场景变化程度。这说明了基于 fine-tuned GoogLeNet 模型的深度特征对于这种情况下的空气质量表征还有待加强。

为了充分比较基于不同特征的空气质量等级估计结果，表 4-4 列出基于各类特征的空气质量等级估计最好结果，并与现有基于场景图像的空气质量等级估计结果进行比较。

表 4-4 不同空气质量等级估计方法对比

	方法	分类精度[%]
1	对比度+SVM 分类器	42
2	能见度+SVM 分类器	58
3	能见度+颜色直方图+SVM 分类器	65
4	Fine-tuned GoogLeNet 模型+SVM 分类器	76
5	[31]	66
6	[32]	64

文献[31]提取了场景图像的透射率、天空色度、光谱率、对比度以及色彩饱和度作为可分类特征并利用多核学习方法学习得到一个自适应分类器，对空气质量等级进行分类。文献[32]与文献[31]类似，也是先提取了场景图像的透射率、对比度、光谱率等特征，然后结合支持向量机和决策树共同构造了一个多类分类器对空气质量等级进行分类。从对比结果可以看出，文献[31]与文献[32]的实验结果与本文基于能见度+色度特征的空气质量等级估计结果基本一致，这说明了场景图像的能见度特征以及色度特征是目前能够表征场景空气质量的最有效的图像特征。而利用 Fine-tuned GoogLeNet 模型的深度特征进行空气估计结果要好于其他方法，这充分证明了场景图像深度特征相比其他类型特征更能描述场景的空气质量。

4.3 本章小结

本节利用卷积神经网络模拟人脑的数据解释能力，对场景图像进行特征学习。为了能够提取与场景空气质量有较强相关性的深度特征，本节介绍了几个目前最具代表性的卷积神经网络结构，包括 LeNet 模型、AlexNet 模型、VGG 模型、GoogLeNet 模型以及 ResNet 模型，并利用这些卷积神经网络分别提取输入数据的深度特征，结合 SVM 分类器进行空气质量等级估计。实验结果表明，基于深度特征的空气质量等级估计效果要整体优于视觉特征和能见度特征，最高能达到 76%。这证明了深度特征最能表征空气质量成像机理。

5 总结

近几年来,全国各地均被雾霾侵袭,许多城市均不同程度地存在超过空气质量标准的情况。研究表明,雾霾中的大气颗粒物能吸附大量基因毒性诱变物质和致癌物质,会给人体健康带来巨大的威胁,如:恶化人体免疫结构、肺结构、呼吸系统和心脏系统、影响生殖能力、加剧慢性病的发生等。因此,对雾霾的防治已成为一个全社会共同关注的热点话题。

本项目提出一项空气质量等级估计的研究计划,旨在通过拍摄场景图像,并利用图像估计场景的空气质量等级。具体工作如下:

(1) 基于视觉特征的空气质量估计。本文引入 6 类图像视觉特征(SIFT 特征、HOG 特征、LBP 特征、色彩饱和度及颜色直方图)。基于 VAQI-1 数据集,本文利用多分类支持向量机方法研究这些视觉特征与场景空气质量的关系。实验结果表明,基于视觉特征的估计效果普遍不理想,最高准确率仅达 42%。这说明了依靠视觉特征无法描述场景空气质量。为此,本文需要探究更能描述空气质量的特征。

(2) 基于能见度特征的空气质量等级估计。由于视觉特征在空气质量等级估计的局限性,本文提出利用能见度特征对空气质量等级进行估计。能见度特征描述了场景最大的可视范围,对空气质量有直观反应。基于大气物理学和单幅图像去雾算法,本文提出了一种基于单幅图像的能见度计算方法。实验结果表明该算法在保证算法准确性的基础上无需任何人工干预及仪器辅助,有较强的鲁棒性和通用性。基于能见度特征,本文利用多分类支持向量机方法对场景空气质量等级进行估计。实验结果表明能见度特征相比视觉特征分类效果明显提升,准确率达 58%。考虑到能见度特征对极端天气如纯雾天、纯霾天等无法进行有效区分,本文将能见度特征与不同视觉特征进行结合,发现能见度与颜色直方图的组合准确率达 65%,该结果要优于仅依靠能见度特征的分类效果。这说明场景图像能见度特征与颜色特征的结合可以较准确的描述场景空气质量。

(3) 基于图像深度特征的空气质量等级估计。卷积神经网络是目前深度学习领域最前沿和热门的方法之一。尤其在计算机视觉方面,卷积神经网络的应用开启了一个全新的时代。本项目利用目前最主流的 5 类卷积网络模型(即 LeNet 模型、Alex 模型、VGG 模型、GoogLeNet 模型、ResNet 模型)分别提取场景图像的深度特征,并利用 SVM 分类器进行空气质量等级估计。考虑到 VAQI-1 为小样本数据集,本文引入 ImageNet 数据集对模型进行预训练,再利用 VAQI-1 对模型进行 fine-tuning,提高估计精度。实验结果表明利用 fine-tuned GoogLeNet 模型提取的深度特征进行空气质量等级估计准确率达到 76%。这个结果不仅优于其他

基于深度特征的空气质量等级估计方法,同时也优于所有基于视觉特征和能见度特征的空气质量等级估计方法。这充分证明了这类深度特征最能贴合空气质量成像机理。

本项目有以下几点贡献:

(1) 本文建立了建立目前已公开的最大的用于空气质量检测的场景图像数据库。该数据库共包含来自中国 26 个城市 75 个观测点共 4444 张场景图像。为吸引更多研究者参与到基于场景图像的空气质量评估这一研究来,我们会持续更新并完善 VAQI-1。

(2) 本文提出一种基于单幅图像计算场景能见度的方法。该方法估计结果与人眼观测结果基本一致,不仅简单有效,且无需人工干预及仪器辅助,适用于任意场景,有较强的创新性和实用性。

(3) 由于智能手机的普及,越来越多的人用智能手机便捷的拍摄身边的影像。这些图像数据不仅记录了不同区域的人们每天的生活,而且也记录了城市几乎每个区域的生态环境的改变。有数据表明,大数据时代最大数据量的数据即是由全球数十亿用户拍摄的各类视频、图像数据。基于这类图像数据,本项目提出的基于场景图像的空气质量检测技术能够实现用户用手机实时拍摄场景估计空气质量,这给人们提供了用以记录、改善我们生存的环境的前所未有的强大工具,大大提高了民众参与度,提高了环保意识。

在接下来的研究中,本文将针对 GoogLeNet 网络结构特性深入研究其提取空气质量相关特征的原理,并以此为基础对现有网络结构进行改进,以提高空气质量等级估计准确率。

参考文献

- [1] Nader J. Current Technology for Continuous Monitoring of Particulate Emissions[J]. Air Repair, 1975, 25(8):814-821.
- [2] Knollenberg R G. The Optical Array: An Alternative to Scattering or Extinction for Airborne Particle Size Determination.[J]. Journal of Applied Meteorology, 1970, 9(9):86-103.
- [3] HarveyPatashnick, Rupprecht E. Continuous PM-10 Measurements Using the Tapered Element Oscillating Microbalance[J]. Air Repair, 1991, 41(8):1079-1083.
- [4] Schwartz J. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II[J]. Environmental Research, 1989, 50(2):309-321.
- [5] Chestnut L G, Schwartz J, Savitz D A, et al. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I[J]. Archives of Environmental Health: An International Journal, 1991, 46(3): 135-144.
- [6] Gupta P, Christopher S A, Wang J, et al. Satellite remote sensing of particulate matter and air quality assessment over global cities[J]. Atmospheric Environment, 2006, 40(30): 5880-5892.
- [7] Yu X, Liu Y, Zhu Y, et al. Efficient sampling and compressive sensing for urban monitoring vehicular sensor networks[J]. IET wireless sensor systems, 2012, 2(3): 214-221.
- [8] Li L, Zheng Y, Zhang L. Demonstration abstract: Pimi air box: a cost-effective sensor for participatory indoor quality monitoring[C]//Proceedings of the 13th international symposium on Information processing in sensor networks. IEEE Press, 2014: 327-328.
- [9] Tong H Y, Hung W T, Cheung C S. On-road motor vehicle emissions and fuel consumption in urban driving conditions[J]. Journal of the Air & Waste Management Association, 2000, 50(4): 543-554.
- [10] Emetere M E, Akinyemi M L. Modeling of generic air pollution dispersion analysis from cement factory[J]. Analele Universităţii din Oradea–Seria Geografie, 2013, 23(1): 181-189.
- [11] Liang X Z, Xu M, Yuan X, et al. Regional climate–weather research and forecasting model[J]. Bulletin of the American Meteorological Society, 2012, 93(9): 1363-1387.
- [12] Jeong J I, Park R J, Woo J H, et al. Source contributions to carbonaceous aerosol concentrations in Korea[J]. Atmospheric environment, 2011, 45(5): 1116-1125.
- [13] Brauer M, Amann M, Burnett R T, et al. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution[J].

- Environmental science & technology, 2012, 46(2): 652-660.
- [14] Kim Y, Fu J S, Miller T L. Improving ozone modeling in complex terrain at a fine grid resolution: Part I—examination of analysis nudging and all PBL schemes associated with LSMs in meteorological model[J]. Atmospheric Environment, 2010, 44(4): 523-532.
- [15] Baklanov A, Molina L T, Gauss M. Megacities, air quality and climate[J]. Atmospheric Environment, 2016, 126: 235-249.
- [16] Chen J, Lu J, Avise J C, et al. Seasonal modeling of PM_{2.5} in California's San Joaquin Valley[J]. Atmospheric environment, 2014, 92: 182-190.
- [17] Saide P E, Carmichael G R, Spak S N, et al. Forecasting urban PM₁₀ and PM_{2.5} pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model[J]. Atmospheric Environment, 2011, 45(16): 2769-2780.
- [18] Chan K Y, Jian L. Identification of significant factors for air pollution levels using a neural network based knowledge discovery system[J]. Neurocomputing, 2013, 99: 564-569.
- [19] Cheng S, Li L, Chen D, et al. A neural network based ensemble approach for improving the accuracy of meteorological fields used for regional air quality modeling[J]. Journal of environmental management, 2012, 112: 404-414.
- [20] Li C, Hsu N C, Tsay S C. A study on the potential applications of satellite data in air quality monitoring and forecasting[J]. Atmospheric Environment, 2011, 45(22): 3663-3675.
- [21] Nguyen V A, Starzyk J A, Goh W B, et al. Neural network structure for spatio-temporal long-term memory[J]. IEEE Transactions on neural networks and learning systems, 2012, 23(6): 971-983.
- [22] Cheng T, Wang J, Li X. The Support Vector Machine for Nonlinear Spatio-Temporal Regression[J]. Ieice Trans.info. & Syst, 2013, 55(03):169-186.
- [23] Zheng Y, Liu F, Hsieh H P. U-air: When urban air quality inference meets big data[C]//Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013: 1436-1444.
- [24] Chen J, Chen H, Pan J Z, et al. When big data meets big smog: A big spatio-temporal data framework for China severe smog analysis[C]//Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data. ACM, 2013: 13-22.
- [25] Hasenfratz D, Saukh O, Walser C, et al. Pushing the spatio-temporal resolution limit of urban air pollution maps[C]//Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on. IEEE, 2014: 69-77.
- [26] Li Y, Zhu Y, Yin W, et al. Prediction of High Resolution Spatial-Temporal Air Pollutant Map from Big Data Sources[M]// Big Data Computing and Communications. Springer International Publishing, 2015:273-282.

- [27] Zheng Y, Liu T, Wang Y, et al. Diagnosing New York city's noises with ubiquitous data[C]//Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2014: 715-725.
- [28] Shang J, Zheng Y, Tong W, et al. Inferring gas consumption and pollution emission of vehicles throughout a city[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014: 1027-1036.
- [29] Eeftens M, Beelen R, de Hoogh K, et al. Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; results of the ESCAPE project[J]. Environmental science & technology, 2012, 46(20): 11195-11205.
- [30] Zheng Y, Yi X, Li M, et al. Forecasting fine-grained air quality based on big data[C]//Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015: 2267-2276.
- [31] Zhang Z, Ma H, Fu H, et al. Outdoor air quality level inference via surveillance cameras[J]. Mobile Information Systems, 2016, 2016.
- [32] Zhang Z, Ma H, Fu H, et al. Outdoor air quality inference from single image[C]//International Conference on Multimedia Modeling. Springer, Cham, 2015: 13-25.
- [33] Wang H, Yuan X, Wang X, et al. Real-time air quality estimation based on color image processing[C]// Visual Communications and Image Processing Conference. IEEE, 2015:326-329.
- [34] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [35] Surhone L M, Tennoe M T, Henssonow S F, et al. Histogram of Oriented Gradients[J]. Betascript Publishing, 2016, 12(4):1368 - 1371.
- [36] Guo Z, Zhang L, Zhang D. A completed modeling of local binary pattern operator for texture classification[J]. IEEE Transactions on Image Processing, 2010, 19(6): 1657-1663.
- [37] Bruni V, Rossi E, Vitulano D. On the Equivalence Between Jensen–Shannon Divergence and Michelson Contrast[J]. IEEE Transactions on Information Theory, 2012, 58(7):4278-4288.
- [38] Ukil A. Support Vector Machine[J]. Computer Science, 2002, 1(4):1-28.
- [39] Kerker M. The scattering of light and other electromagnetic radiation[M]. Elsevier, 2016.
- [40] Tomida T, Tsuyuguchi Y, Arai T, et al. The atmospheric transparency measured with a LIDAR system at the Telescope Array experiment[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2011, 654(1): 653-660.
- [41] Stewart I D, Oke T R. Local climate zones for urban temperature studies[J].

- Bulletin of the American Meteorological Society, 2012, 93(12): 1879-1900.
- [42] Kwon T M. Video camera-based visibility measurement system. US: 7016045 B2[P]. 2006.
- [43] 吕伟涛, 陶善昌, 刘亦风, 等. 基于数字摄像技术测量气象能见度——双亮度差方法和试验研究[J]. 大气科学, 2004, 28(4): 559-570.
- [44] Hautière N, Tarel_Affn J P, Lavenant_Affn J, et al. Automatic fog detection and estimation of visibility distance through use of an onboard camera[J]. Machine Vision & Applications, 2006, 17(1): 8-20.
- [45] Hautiere N, Labayrade R, Aubert D. Real-time disparity contrast combination for onboard estimation of the visibility distance[J]. IEEE Transactions on Intelligent Transportation Systems, 2006, 7(2): 201-212.
- [46] Hautière N, Aubert D, Dumont é, Tarel J P. Experimental validation of dedicated methods to in-vehicle estimation of atmospheric visibility distance. IEEE Transactions on Instrumentation and Measurement, 2008, 57(10): 2218-2225.
- [47] Boussard C, Hautiere N, D'Andrea-Novet B. Vehicle dynamics estimation for camera-based visibility distance estimation[C]// Ieee/rsj International Conference on Intelligent Robots and Systems. IEEE, 2008: 600-605.
- [48] 李勃, 董蓉, 陈启美. 无需人工标记的视频对比度道路能见度检测[J]. 计算机辅助设计与图形学学报, 2009, 21(11): 1575-1582.
- [49] 张潇, 李勃, 陈启美. 基于亮度特征的 PTZ 视频能见度检测算法及实现[J]. 仪器仪表学报, 2011, 32(2): 381-387.
- [50] Saunders R W. Atmospheric radiation: Theoretical basis. By R. M. Goody and Y. L. Yung. Oxford university press. second edition 1995. XIII + 519 pp. price 37.50 (paperback). isbn 0 19 510291 6[J]. Quarterly Journal of the Royal Meteorological Society, 2010, 123(539): 801-801.
- [51] Tomchuk P M, Starkov V M, Butenko D V. Integral equations in the general theory of light absorption and scattering[J]. Ukrainian Journal of Physics, 2017, 62(8): 705-716.
- [52] Nemoto Y, Mizuno M, Nagata A. An introduction to atmospheric radiation[M]. Academic Press, 2002.
- [53] Patterson M S, Chance B, Wilson B C. Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties[J]. Applied Optics, 1989, 28(12): 2331-2336.
- [54] Huang F, Zhou J, Tao J, et al. PMODTRAN: a parallel implementation based on MODTRAN for massive remote sensing data processing[J]. International Journal of Digital Earth, 2016, 9(9): 819-834.
- [55] 吴北婴. 大气辐射传输实用算法[M]. 气象出版社, 1998.
- [56] Gueymard C A. Solar Radiation, Spectrum[M]. Springer New York, 2012.
- [57] Chabane F, Moummi N, Brima A. Predictions of solar radiation distribution:

- Global, direct and diffuse light on horizontal surface[J]. European Physical Journal Plus, 2016, 131(4):1-8.
- [58] Aring, Ngström A. The parameters of atmospheric turbidity[J]. Tellus, 1964, 16(1):64-75.
- [59] Karplus R, Neuman M. The Scattering of Light by Light[J]. Physical Review, 1951, 83(4): 776-784.
- [60] He Kai-Ming, Sun Jian, Tang Xiaoou. Guided image filtering. In: Proceedings of the 11st European Conference on Computer Vision. Crete, Greece: Springer, 2010. 1-14.
- [61] 刘敏, 赵普洋. 气象光学视程(MOR)在民用航空地面气象观测中的应用[J]. 气象水文海洋仪器, 2012, 29(1):78-80.
- [62] 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.
- [63] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
- [64] Kim J, Lee J K, Lee K M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016:1646-1654.
- [65] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:1-9.
- [66] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [67] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]// Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009:248-255.

攻读学位期间主要的研究成果目录

参与科研项目情况

- [1] 国家自然科学基金青年项目（61602520）：基于能见度检测的空气质量等级估计，2017.01-至今，主要研究人员
- [2] 中南大学研究生科研创新项目（2017zzts500）：基于图像视觉特征的场景空气质量等级估计，2017.03-2018.03，负责人

已发表和录用的文章

- [1] Li Q, Xie B. Visibility Estimation Using a Single Image[C]// CCF Chinese Conference on Computer Vision. Springer, Singapore, 2017:343-355.

已申请的专利

- [1] 《一种基于单幅图像估计空气质量的方法》，国家发明专利申请，第二作者，申请号：201810047540.3
- [2] 《一种通过单幅图像估计能见度的方法》，国家发明专利申请，第二作者，申请号：201710540165.1

硕士期间参加学术交流活动

- [1] 2017 年 10 月 第二届中国计算机视觉会议（CCCV2017）（天津）

致谢

时光匆匆如流水，转眼便是研究生毕业时节。毕业日期已日趋渐进，毕业论文的完成也随之进入了尾声。从开始进入课题到论文的顺利完成离不开老师、同学、朋友给我热情的帮忙。在这里请接受我诚挚的谢意！感谢这一路走来支持过和帮助过我的导师、同学和好朋友们！

首先最想感谢我的研究生导师——谢斌老师。在学术方面，谢老师传授我专业知识，教我如何解决问题以及思考问题的方法。正是在谢老师的指导下，我开始接触了图像处理、模式识别等研究并产生了浓厚的兴趣。谢老师广博的学术知识和开阔的学术视野让我深受启发，产生了很多灵感。在生活方面，谢老师经常和我谈论人生理想以及未来的发展规划，让我对未来的道路有了更明确的方向。谢老师就是我的良师益友！

感谢智能所的各位老师给我的帮助！感谢蔡自兴老师、王勇老师、唐璘老师、刘丽珏老师、高琰老师、肖晓明老师、陈白帆老师、于伶俐老师、李仪老师、邹逸群老师、郭幡老师、肖晓明老师及谭平老师。

感谢谢门的小伙伴们！感谢喻仲斌师兄对我的帮助，感谢我的同门梁照栋在学术和生活中对我的照顾，感谢宋迪、何小宇、张文婷、秦觅、林学斌、赵新宇以及徐勇在我的研究生生活中给予的帮助和欢乐。

感谢我同一级的小伙伴们！感谢李佳朋、万明阳、罗舒宁、王琦、梁春芳、刘春发、薛喜辉、张财旺、刘昊、吴潇、顾磊、李凡、龙子威、莫斯尧和李路。因为有你们，我的研究生生活变得多姿多彩。

感谢我的舍友罗舒宁、党伟然和贺卓！研究生三年的日子里有你们的陪伴我觉得很幸福。

感谢 209 的小伙伴们平时对我的帮助！感谢刘角、王欣、于俭、王润宇、刘懿、晋周南、何志远、茹志阳、吴喆、刘志忠、黄佩秋、薛韬略、林剑清、陈春蓉和马忠伟。

感谢我的另一半张哲六年来对我的帮助与爱。

最后还要特别感谢我的父母和家人，谢谢你们这么多年的培养。没有你们的无私奉献，就没有今天的我。谢谢！