

中国研究生创新实践系列大赛  
“华为杯”第十七届中国研究生  
数学建模竞赛

学 校 郑州大学

参赛队号 20104590100

1.戴浩然

队员姓名 2.彭斐琳

3.智志洋

**中国研究生创新实践系列大赛**  
**“华为杯”第十七届中国研究生**  
**数学建模竞赛**

题 目      大雾演化规律的量化分析与预测模型研究

---

**摘                      要：**

随着人工智能技术的进步，近年来基于视频的路况能见度检测方法有了较大的创新，该方法可以克服传统激光能见度仪耗资巨大和探测范围小等问题，然而现有方法大多是只选取少量视频、截取图像中的某些固有特征进行能见度数值的间接估计。事实上，大雾的形成和消散有其自身的规律，视频资料中蕴含了丰富的雾变化过程信息，充分利用这些信息，不仅可以提高能见度估计精度，也可以对大雾的消散进行预测。这对我国的高速公路管理和航空飞行具有重要的意义。

基于此，本文旨在通过收集和分析与能见度相关的气象因子信息和图像特征信息，评估各变量特征与能见度之间的相关性程度，并以此为基础在构建地面气象观测与能见度之间的多变量非线性回归模型；同时，充分利用视频的连续信息建立基于视频数据的能见度估计深度学习模型。然后利用图像中不同景深的物体在不同能见度下的亮度差异构建不依赖实测数据的能见度估计模型；最后运用时间序列分析模型量化能见度随时间的变化规律，建立数学模型预测大雾变化趋势以及其消散时间。

针对问题一，要求建立地面气象观测数据与能见度的具体关系式。为此，需要寻找出影响能见度的气象因子，从而建立能见度和气象因子的多元非线性回归模型。研究中结合文献调研，初步选取地面气象观测数据中可能影响能见度的因子，进而运用相关性分析评估各因子之间的内在关联，最终确定回归模型的变量因子。然后运用 SPSS 和 Matlab 软件对历史实测数据进行时空变化趋势做统计分析研究能见度与各气象因子的数学关系，以此构建出不同气候模式下估计能见度的多元非线性回归模型；在求解参数时，利用遗传算法得到多变量非线性回归模型的最优解并检验参数合理性。结果表明，利用该方法可以从定量和定性两个角度建立能见度和气象因子的回归方程，具有一定的可靠性和正确性。

针对问题二，要求建立基于视频数据与实测能见度数据的深度学习反演模型。为此，需要寻找出影响能见度的图像特征因子，从而构建能见度与视频数据的深度学习模型。首先，通过已有研究初步选取影响能见度的图像特征，然后运用相关性分析筛选最能影响雾天能见度的图像特征向量以及图像特征的窗格大小，同时评估各窗格区域内的特征值与能见度的关系，选取最优兴趣窗口，进而构建回归模型的参数向量；然后以多元线性回归方程建立参数向量与能见度之间的统计关系，并利用 BP 神经网络对线性拟合的残差进行估计修正多元线性回归的预测结果。结果表明，该模型能够同时兼顾线性和非线性的特点，反演精度较高，可以实现对能见度的精准估计。

针对问题三，要求建立非直接测量能见度的能见度估计算法。为此，需要获取图片深度信息估计能见度。从图像数据来看，近点有较清晰的道路信息，所以可借助道路线进行梯度检测求得边缘信息；已知车道线的长度 6m，利用亮度对比法结合光学模型的传播可以求出摄像机参数；远点因为大雾而看不到道路的尽头，因此我们利用求定的相机参数来

获取离相机最远可分辨的像素，即人眼可分辨的亮度差异；然后使用**暗原色先验原理**获得目标点的透射率，求出大气折光系数，进而对能见度进行估计。结果表明其中 2016 年 4 月 14 日 6 点 30 分至 7 点 39 分时段能见度变化趋势较缓，集中在 47m 附近，波动范围±10m。

针对问题四，要求根据能见度随时间的变化规律，建立数学模型预测大雾变化趋势以及何时消散。为此，需要对能见度变化序列进行时序分析，主要包括倾向变化率、平稳性、周期性、偏相关和自相关，进而通过**线性回归模型**、**灰色预测模型**和 **ARIMA 模型**模拟大雾的变化趋势以及消散过程。结果表明，问题三等到的能见度随时间变化具有一定的周期性，且整体上保持恒定，上下波动不大；线性回归模型的结果显示下一时刻的大雾变化趋势逐渐减弱，且在 8 点 50 分 56 秒大雾开始消散即达到能见度为 150m 以上；ARIMA 时序预测模型和灰色预测模型的结果显示大雾的变化趋势在短时间内逐渐加重，但速率较缓。

**关键词：**BP 神经网络、多元非线性回归模型、相关性分析、亮度差异、暗原色先验、时间序列预测、灰色预测、ARIMA 时间序列模型

公众号关注：建模忠哥  
获取更多资源

## 目录

1. 问题重述 .....	5
1.1 问题背景 .....	5
1.2 问题提出 .....	5
2. 模型假设 .....	7
3. 技术路线 .....	8
4. 问题一建模与求解 .....	11
4.1 问题分析与思路 .....	11
4.2 数据预处理 .....	12
4.2.1 信息剔除与数据化 .....	12
4.2.2 数据标准化 .....	12
4.3 能见度时序变化特征分析 .....	13
4.4 气象因子和能见度的相关性分析 .....	14
4.5 能见度与各气象因子的非线性特征分析 .....	15
4.5.1 能见度与风速的关系 .....	15
4.5.2 能见度与本站气压的关系 .....	16
4.5.3 能见度与温度的关系 .....	17
4.5.4 能见度与相对湿度的关系 .....	18
4.6 能见度与气象因子的多变量非线性回归模型 .....	19
4.6.1 能见度与气象因子的多元回归方程的确定 .....	19
4.6.2 遗传算法求解多元非线性方程的最优回归系数 .....	19
4.6.3 结果分析 .....	20
5. 问题二建模与求解 .....	22
5.1 问题分析 .....	22
5.2 视频数据预处理 .....	23
5.2.1 视频转关键帧图像 .....	23
5.2.2 异常图像处理 .....	24
5.2.3 ROI 选取 .....	24
5.2.4 降噪处理和增强处理 .....	25
5.3 能见度时序变化特征分析 .....	26
5.4 图像特征因子分析 .....	26
5.4.1 图像特征因子 .....	26
5.4.2 图像子窗格划分 .....	27
5.4.3 特征因子与能见度的相关性分析 .....	28
5.4.4 不同窗格下图像特征因子与能见度的相关性分析 .....	28
5.5 兴趣窗格的选择 .....	29
5.6 能见度反演模型的构建 .....	31
5.6.1 三元线性回归建模 .....	31
5.6.2 BP 神经网络 .....	32
5.5.3 BP 神经网络修正线性残差模型的构建 .....	33
5.5.4 精度分析 .....	33
6. 问题三建模与求解 .....	35

6.1 问题分析 .....	35
6.2 数据预处理 .....	35
6.2.1 剔除无关信息 .....	35
6.2.2 图像增强 .....	36
6.3 边缘检测 .....	36
6.4 道路提取 .....	37
6.4.1 霍夫变换 .....	37
6.5 道路兴趣区的提取 .....	38
6.6 相机标定 .....	39
6.7 亮度对比度提取道路最远可视点 .....	39
6.8 暗原色法计算消光系数 .....	40
6.9 能见度估计 .....	41
7. 问题四建模与求解 .....	43
7.1 问题分析 .....	43
7.2 能见度变化时序特征分析 .....	43
7.2.1 能见度倾向率 .....	43
7.2.2 时间序列的平稳性及周期性检验 .....	44
7.2.3 自相关性与偏相关性 .....	45
7.3 ARIMA 时间序列模型 .....	45
7.4 GM(1,1)灰色预测模型 .....	47
7.5 线性回归模型 .....	48
参考文献 .....	49
附录 .....	50

公众号关注：建模忠哥 获取更多资源

# 1. 问题重述

## 1.1 问题背景

能见度是气象、公路行车、飞机飞行中常见指标，单位通常是米。一般意义下的能见度，是指视力正常的人能将目标物从背景中识别出来的最大距离。在科学研究中，能见度是大气环境研究领域的重要课题，是反应大气透明度的一个指标<sup>[1]</sup>。气象上所讲的大气能见度指白天情境下，正常视力所能看到并且辨认出的以地平线附近天空为背景的大小合适的黑色目标物所对应的最远的距离<sup>[2][5]</sup>。

能见度对高速公路行车安全非常重要，雾天能见度低对交通运输会产生很大影响，尤其对高速公路运输和航空服务等，在这样的情况下，人们的生命财产安全面临巨大的威胁。据相关部门调查研究发现，发生的交通事故中，由于雾天能见度降低造成严重后果的交通事故数量远远超过其他恶劣天气，且死伤人数占了交通事故总人数的  $1/3$ <sup>[6]</sup>。虽然在公路运输和航空服务中通行标准有所不同，但不可否认，气象条件的影响会给相应的管理和协调带来极大的困扰。因此，能见度预测一直是高速公路部门和航空公司十分重视的问题。

自从 20 世纪初期 Koschmieder 提出 Koschmieder 定律，能见度研究进展逐渐深入。目前，我国高速路网已逐步形成，若大量使用激光能见度仪对全国高速路网进行全覆盖将耗资巨大，同时激光能见度仪还存在对团雾检测精度不高，探测的范围很小，维护成本高等不足。近年来，基于视频的路况（跑道）能见度检测方法受到人们的关注，它在某种程度上克服了激光能见度仪的不足。视频能见度检测方法是大气光学分析与图像处理及人工智能技术结合，通过对视频图像分析处理，建立视频图像与真实场景之间的关系，再根据图像特征的变化，间接计算出能见度数值。这样做有以下几个优点：第一，通过使用道路监控数据提高了监控相机或者是影像采集设备的使用价值；第二，通过监控视频可以方便地、自动地、长期地、连续地、低成本地获取外界环境数据；第三，这种连续无间断的时间序列数据可以对气象变化进行很好地拟合和预测。然而，现有的基于视频图像的能见度检测方法，由于是间接计算，很难准确地估算能见度。特别地，这些方法中大多数只是选取少量视频、截取图像中的某些固有特征基于 Koschmieder 定律进行估计，并没有充分利用视频的连续信息，估计精度不高，存在较大的改进空间。因此，基于视频数据的能见度检测的研究是一个热点也是难点。

事实上，大雾的形成和消散有其自身的规律，通常与近地层的气象因素有关。而视频资料包含了丰富的信息，特别是涵盖了大雾的变化过程信息。充分利用这些信息，不仅可以提高能见度估计精度，也可以对大雾的消散进行预测，这将对国家、相关部门和人民的生命安全具有重要意义。

## 1.2 问题提出

本文以上述分析为背景，主要考虑一下几个问题：

（1）根据能见度与地面气象观测站（包括温度、湿度和风速等）的数据，揭示气象因子与能见度之间的数学关系。对各气象因子和能见度的变化规律进行分析，包括变化速率和时序性；然后评估各气象因子与能见度之间的相关关系，进而对各因子与能见度进行统计描述，分析之间的函数关系，以此构建出不通气候特征下的多元回归方程；

（2）根据已知的机场视频数据和能见度数据，建立起基于视频数据的能见度估计深度学习模型并进行相应的精度评价。通过对图像特征的提取分析各特征与能见度之间的相关关系，确定深度学习的特征参数向量，并以此构建深度学习模型。



(3) 根据提供的视频监控数据，建立通过视频数据估计能见度的算法；由于激光能见度仪不能大规模使用，而高速公路的监控视频数据作为一项简单易得的数据，因此建立不依赖于能见度仪观测数据的能见度观测方法具有十分重要的现实意义；

(4) 基于前述模型和算法分析能见度随时间变化规律，构建大雾变化的预测模型并根据现实情况下的一些因素对未来一定时间段内的变化曲线做出一定的说明。

公众号关注：建模忠哥  
获取更多资源

## 2. 模型假设

### 2.1 模型的基本

在本文的研究中心，采用基于暗通道先验理论所需要进行的问题假设：

- (1) 假设研究的道路是光滑的水平道路，即不存在减速装置和道路起伏；
- (2) 假设相机拍摄不同时间视频或图像时电子元件均正常；
- (3) 假设实际生活中大气本身并不是十分干净，即便是晴天图像中观察远距离物体仍然会存在少量雾霾；

### 2.2 模型的符号说明

表 2-1 模型符号说明

序号	符号	说明
1	$\sigma, meas$	图像方差
2	$g, grad$	图像平均梯度
3	$c, cons$	图像对比度
4	$\mu$	图像均值
5	$reg$	图像子窗格
6	$VIS$	能见度
7	$\hat{V}IS$	能见度线性拟合值
8	$e$	拟合残差
9	$\hat{e}$	修正拟合残差
10	$vs$	风速
11	$pa$	压强
12	$t$	温度
13	$u$	相对湿度



### 3. 技术路线

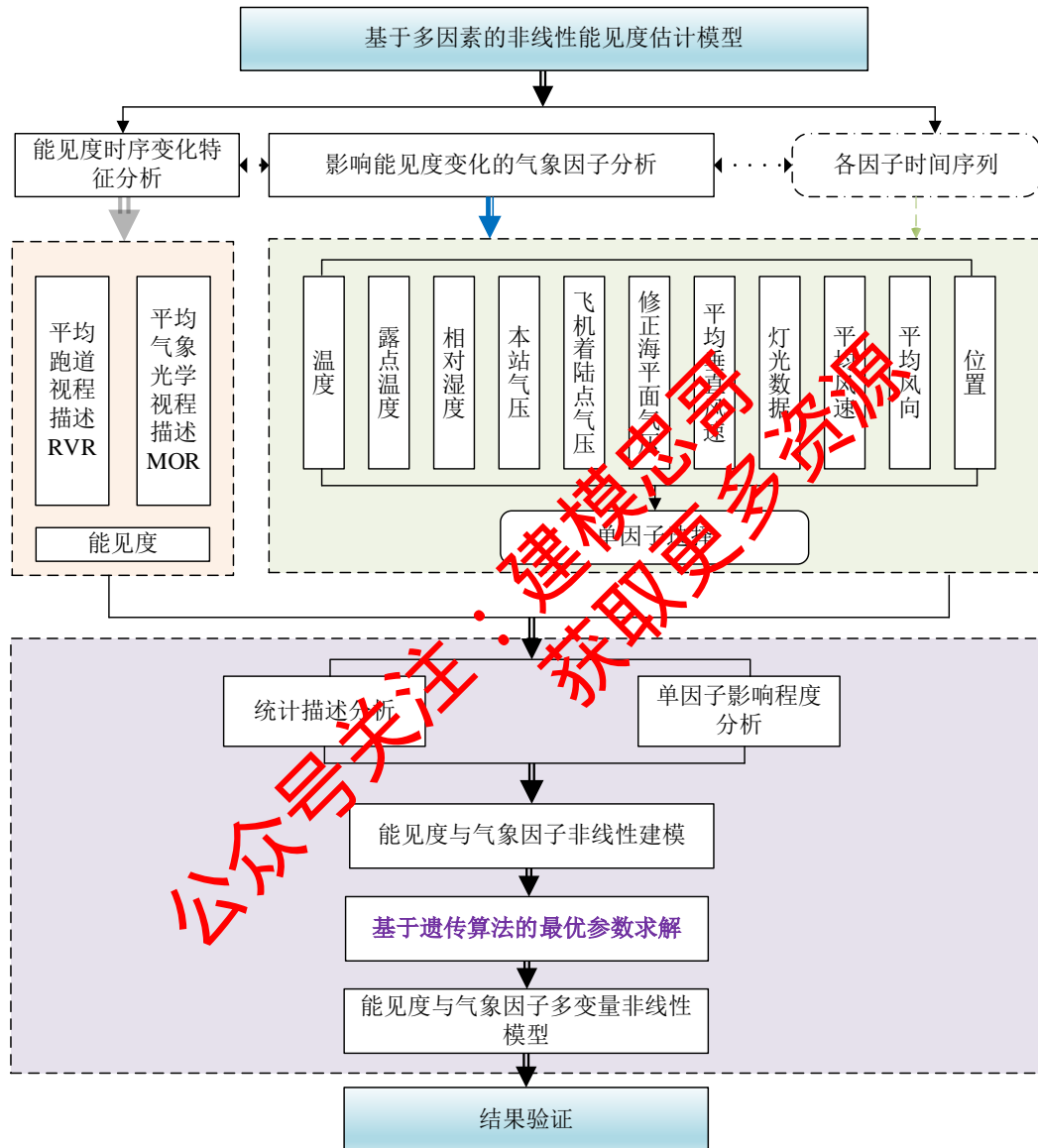


图 3-1 问题一技术流程图

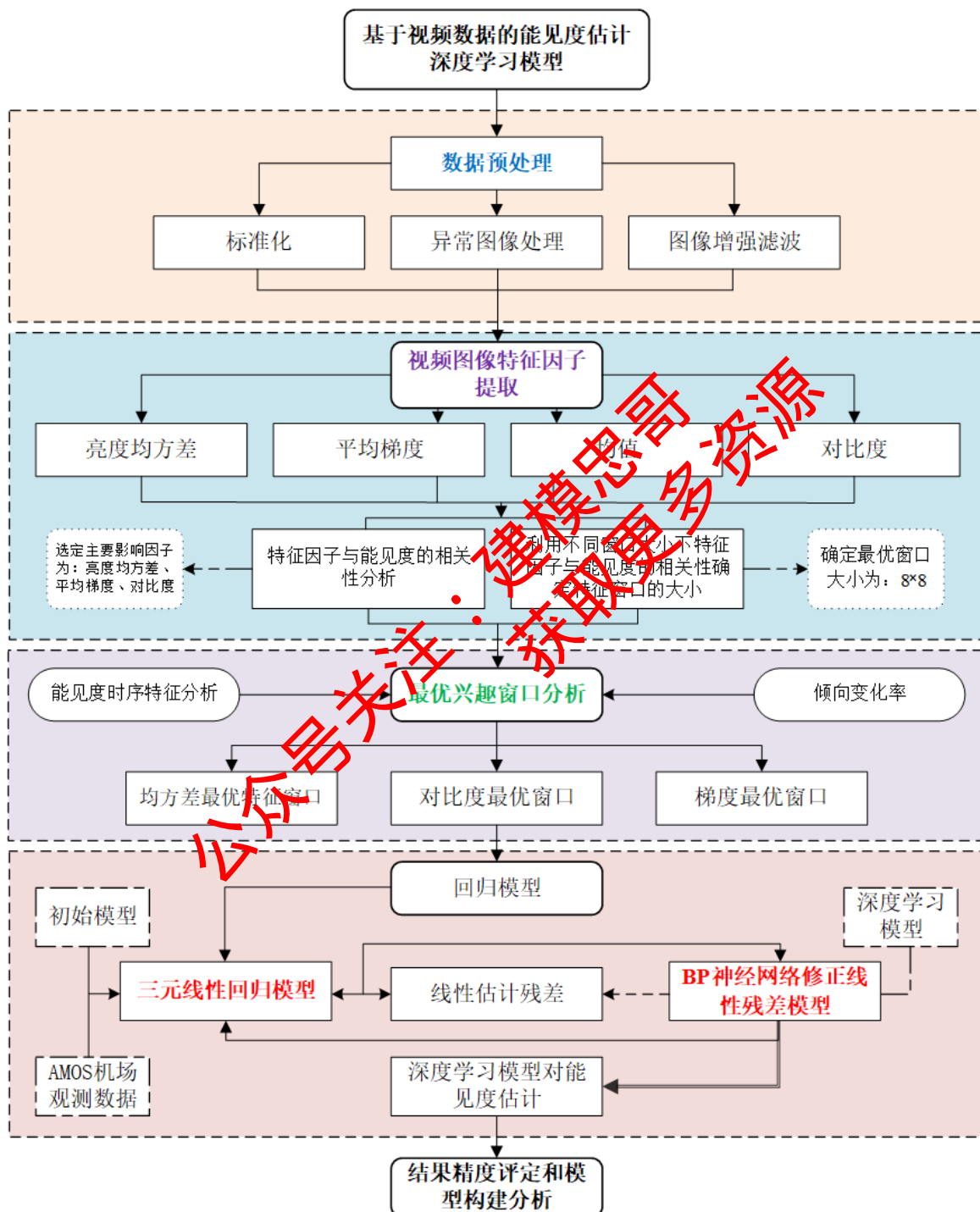


图 3-2 问题二技术流程图

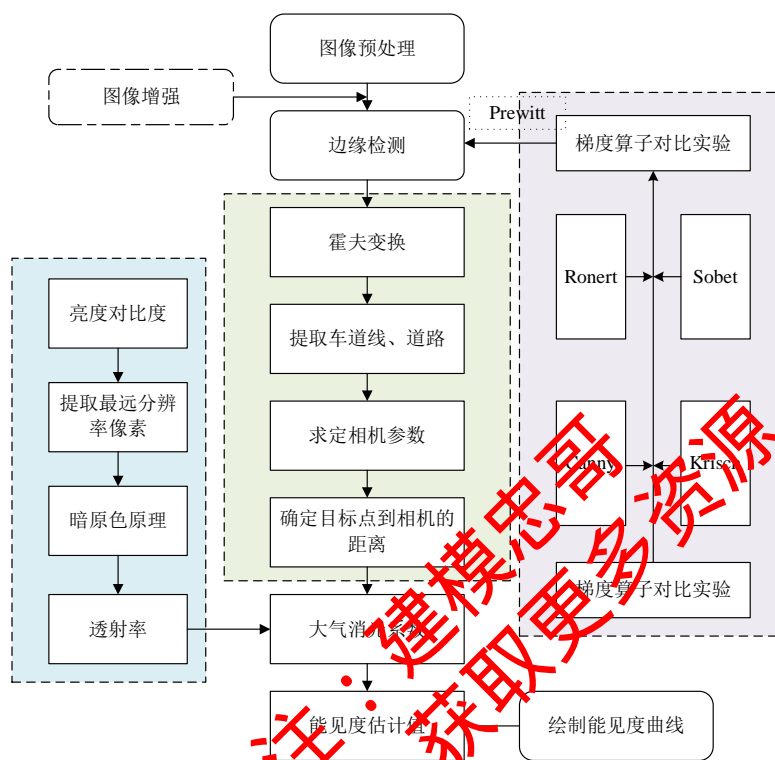


图 3-3 问题三技术流程图

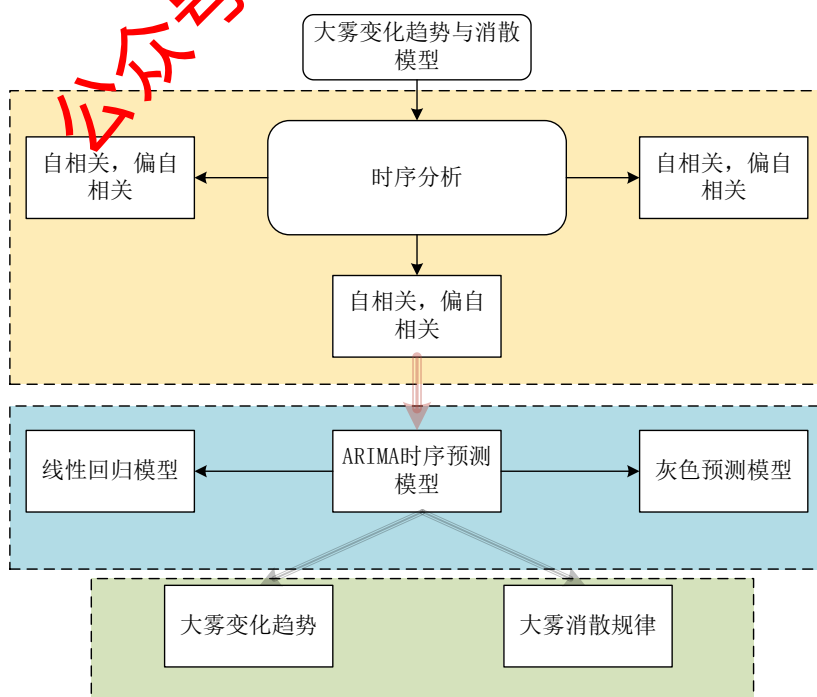


图 3-4 问题四技术流程图

## 4. 问题一建模与求解

### 4.1 问题分析与思路

在问题一中，要求根据题目提供的数据描述能见度与地面气象观测（温度、湿度和风速）的关系，并给出具体的关系式。影响能见度的重要因素是雾和霾，其形成与气象因子具有重要的相关关系，为从定量何定性两个角度去评估和分析二者之间的内在关联，本章尝试利用统计学方法分析某机场气候因子与能见度的特征变化，在此基础上结合数学回归模型构建二者之间的关系表达式，具体做法如下：

首先，对题目提供的 AMOS 观测数据进行预处理，按照合适的时间间隔关联能见度观测数据与气候数据，然后结合文献调查和相关专业知识对数据中的各项指标进行统计归纳整理，并利用描述性统计方法分析各项因子的特征和变化规律，以去除冗余变量和挖掘各因子对能见度的影响信息分析其间的内在联系，常用的方法有分类统计、倾向变化率、相关系数和可视化分析；然后基于以上特征信息利用非线性回归方程模拟单个气象因子与能见度之间的统计学关系，最后以该统计学关系为基础构建所有气象因子与能见度之间的多元回归模型，利用遗传算法进行最优参数求解从而确定能见度与各气象因子之间的非线性回归方程。

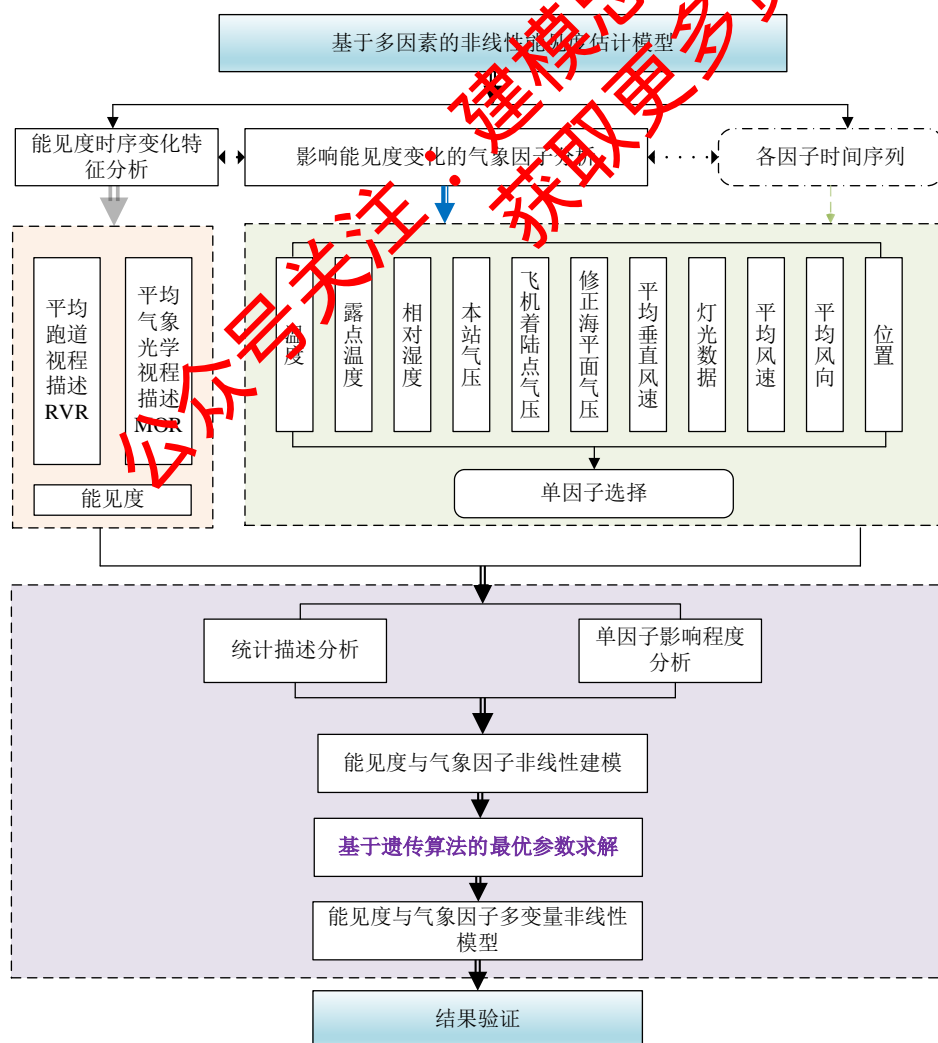


图 4-1 问题一技术流程图

## 4.2 数据预处理

### 4.2.1 信息剔除与数据化

由于气象观测信息很多，形式也多种多样，因而首先需要筛选题目数据中与雾的形成和消散有关的信息，将不必要的数据剔除。共提取出与能见度相关的气象因子如表 4-1 所示。

表 4-1 提取因子说明

变量	说明
RVR_1A	1 分钟平均跑道视程
MOR_1A	1 分钟平均气象光学视程，即能见度
PAINS(HPA)	本站气压
QFE 06	飞机着陆地区最高气压
QNH	修正海平面气压
TEMP	温度
RH	相对湿度
DEWPOINT	露点温度
WS2A	2 分钟平均风速
WD2A	2 分钟平均风向
CW2A	2 分钟平均垂直风速

### 4.2.2 数据标准化

数据中 PTU\_R06\_15.his 的时间间隔为 1 分钟，而 VIS\_R06\_15.his 和 WIND\_R06\_15.his 的时间间隔为 15 秒，因此需要统一时间尺度将气象因子与能见度数据进行合并。考虑到数据记录的时间为大约为第一天的 8 点到第二天 8 点，时间跨度较大，因而为较好的反应一天的能见度和各气象因子变化以及更方便有效地处理分析数据，最终选取以 15 分钟为间隔整合数据作为后续的实验样本。

表 4-2 处理后时间间隔为 15 分钟的部分数据

LOCALDATE (BEIJING)	RVR_ 1A	MOR _1A	WS2 A	WD2A	CW2A	PAINS (HPA)	QNH	QFE R06	TEMP	RH	DEWP OINT
2019... 8:00:00	3000	7000	3.78	86	1.78	1023.3	1024.8	1023.2	8.6	90	7.06
2019... 8:15:00	3000	8000	4.45	94	2.61	1023.3	1024.8	1023.2	9.5	86	7.29
2019... 8:30:00	3000	7000	4.55	89	2.34	1023.2	1024.7	1023.1	9.9	84	7.34
2019... 8:45:00	3000	7000	3.73	93	2.14	1023.3	1024.8	1023.2	10.6	82	7.68
2019... 9:00:00	3000	7000	4.12	108	3.15	1023.2	1024.7	1023.1	11.2	80	7.9
2019... 9:15:00	3000	6000	5.7	113	4.67	1023.3	1024.8	1023.2	12.1	78	8.41

### 4.3 能见度时序变化特征分析

将两个数据集 AMOS20200313 和 AMOS20191216 的能见度数据以 15 分钟为间隔进行可视化，同时计算其不同时段段的倾向变化率，如图 4-2 所示。

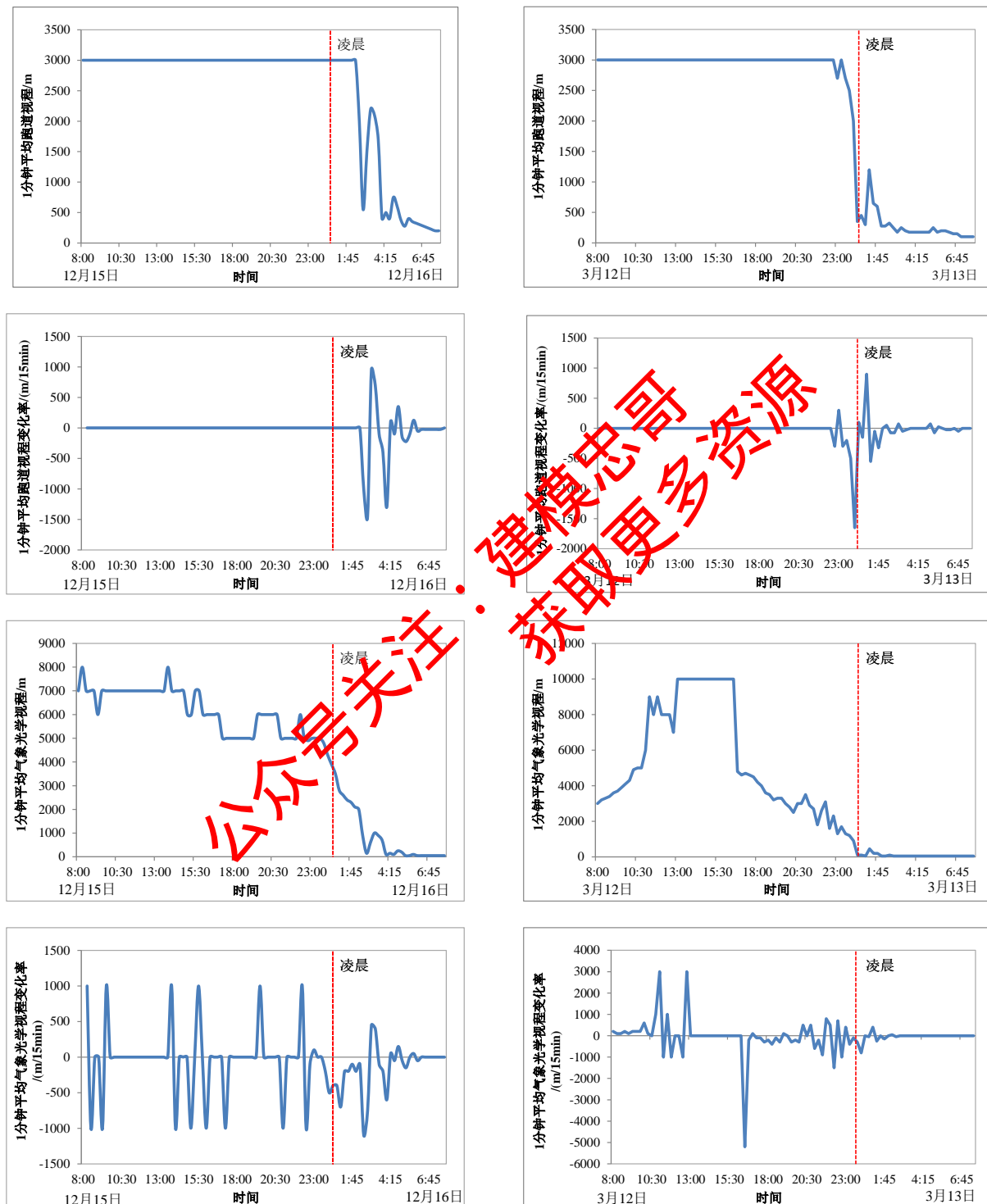


图 4-2 能见度时序变化特征图（左图为 AMOS20191216，右图 AMOS20200313）

从图 4-2 中可以看出，能见度的曲线变化一天的变化趋势主要体现在凌晨附近开始下降，早上 8 点开始逐渐上升。平均跑道视程曲线的突变时段主要体现在凌晨（00:00）到早上 7 点附近，气象光学视程整天都在波动变化的突变时段主要为凌晨和早上 8 点附近。



#### 4.4 气象因子和能见度的相关性分析

题目所给数据中包含了多种气象数据，为了更加科学地理解和阐述这些变量之间的关系，对能见度和所有气象因子进行 Pearson、Kendall 和 Spearman 检验，检验结果如下图 4-3 所示。

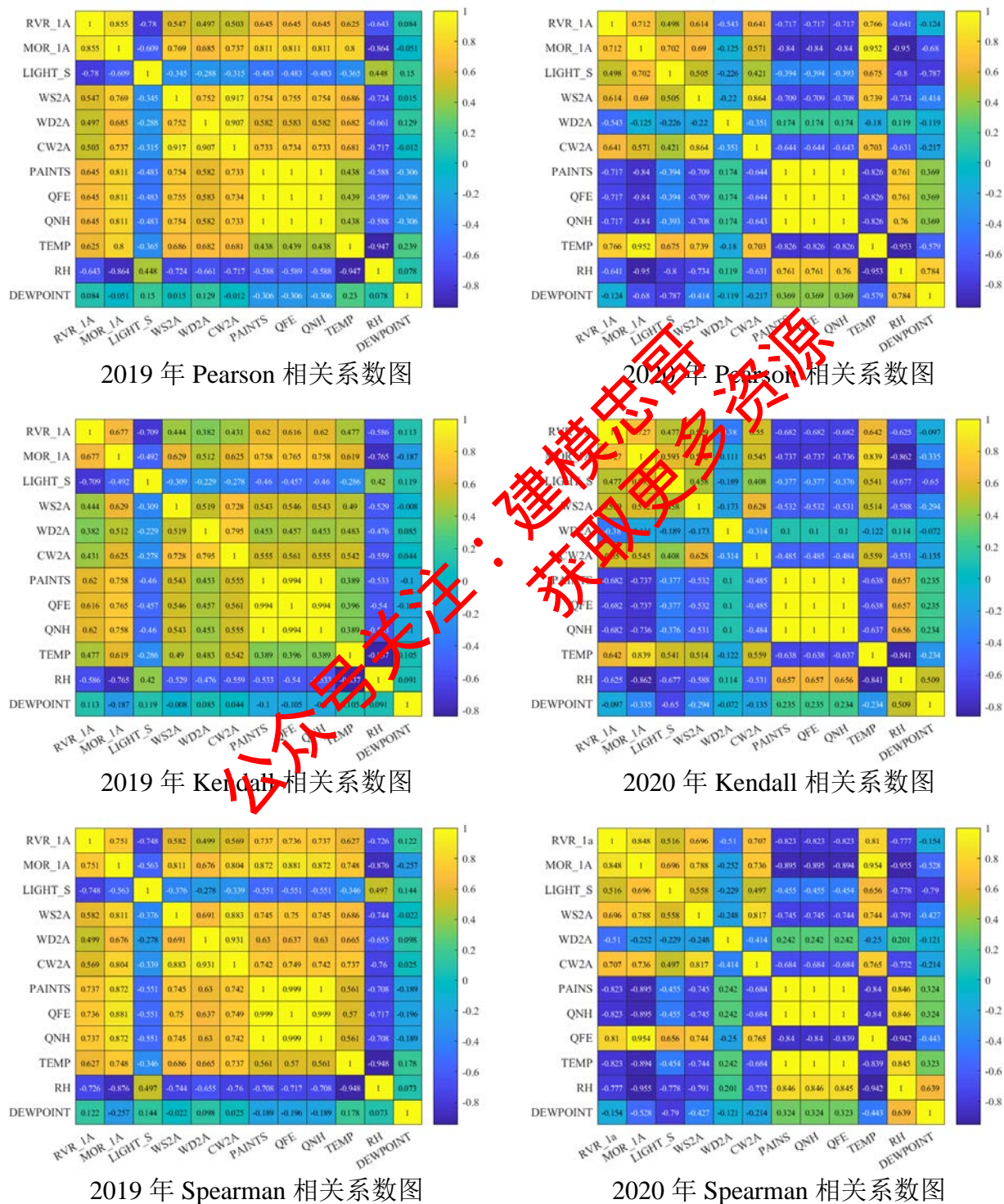


图 4-3 各变量的相关性分析

从图 4-3 中可以看出，气象光学视程和跑道视程之间具有很强的正相关性，查阅相关文献<sup>[3][4]</sup>可知，气象光学视程是气象学上的一个广义和普遍的概念，即我们常说的能见度，



而跑道视程是将能见度应用于航空上一个狭义的特殊概念。因而在后续的讨论中对能见度的建模与分析都是以气象光学视程为主。

诸如风速、风向、垂直风速、气压和温度均与能见度有较强的相关性，其中以风速、本站气压和温度对能见度的影响最强，但是气压和风向表现出季节差异性，即在冬季与能见度呈现正相关，而春季表现出负相关；湿度对能见度具有较强的负相关，露点温度对能见度的影响也为负相关但程度较小。在同一类型的气象因子中，如风速、风向与垂直风速，温度与露点温度，本站气压、修正海平面气压与飞机着陆地区最高气压之间具有非常强的相关性（>0.7），因而在后续的分析考虑中，仅选取同类别中与能见度相关性的因子作为最终变量，即风速、温度、本站气压与相对湿度。

4.5 能见度与各气象因子的非线性特征分析

4.5.1 能见度与风速的关系

风速是影响能见度的重要因子之一，本文数据中与风速相关的数据有 1 分钟平均风速（WS2A）、1 分钟平均风向（WD2A）和一分钟平均垂直风速（CW2A）。经图 4-3 中的相关性分析和先验知识，我们认为该三个变量间具有较强的共线关系，因而只选择对能见度相关性较大的平均风速作为风速因子。对本章数据中的风速进行统计如表 4-3 和图 4-4 所示。

表 4-3 不同风速区间能见度均值统计表

年份	风速 (m/s)	≤1	1-2	2-3	3-4	4-5	5-6	>6
2019.12	MOR(m)	-	-	-	3597.8	6812.5	6833.3	7000.0
2020.03	MOR(m)	518.3	3477.4	6485.7	7325.0	4800.0	-	-

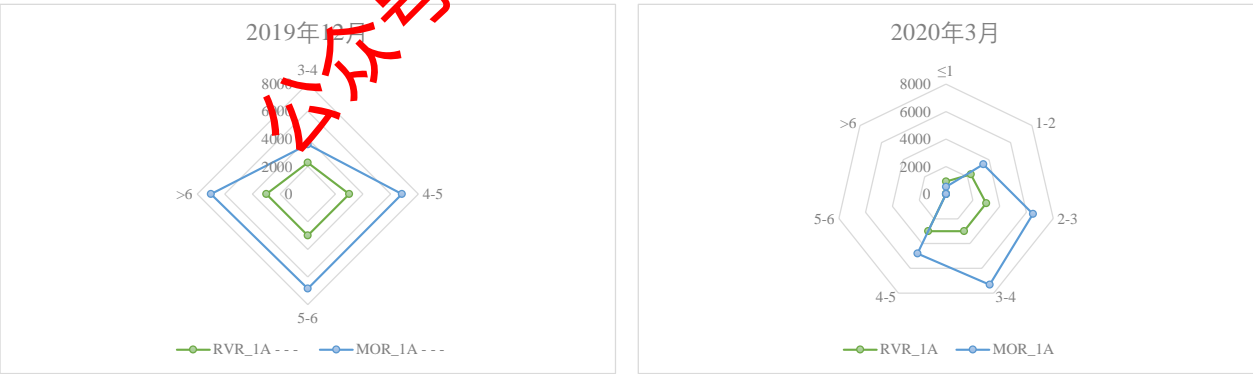


图 4-4 风速分布图

从图中可以看出 2020 年 3 月的风速等级明显弱于 2019 年 12 月，其对应的能见度也明显低于 2019 年；从风速上看，风速越大所对应的能见度也越高，具有一定的和正相关，这与上一小节中分析一致。图中不同时间风速的差异主要体现在对应季节的不同，风速对能见度的作用表现在对污染物的扩散、稀释和移动具有推动作用，通常情况下风速越大，作用应该越明显，而冬季相对于风冷风较多，春季多和煦微风。

利用散点图构建能见度与风速之间的统计学关系，从图 4-5 中我们可以发现，风速与能见度之间并非简单的线性关系，从结果上看二者呈现二次函数关系，拟合结果如下图所示，R2 系数分别为 0.6256 和 0.5207。

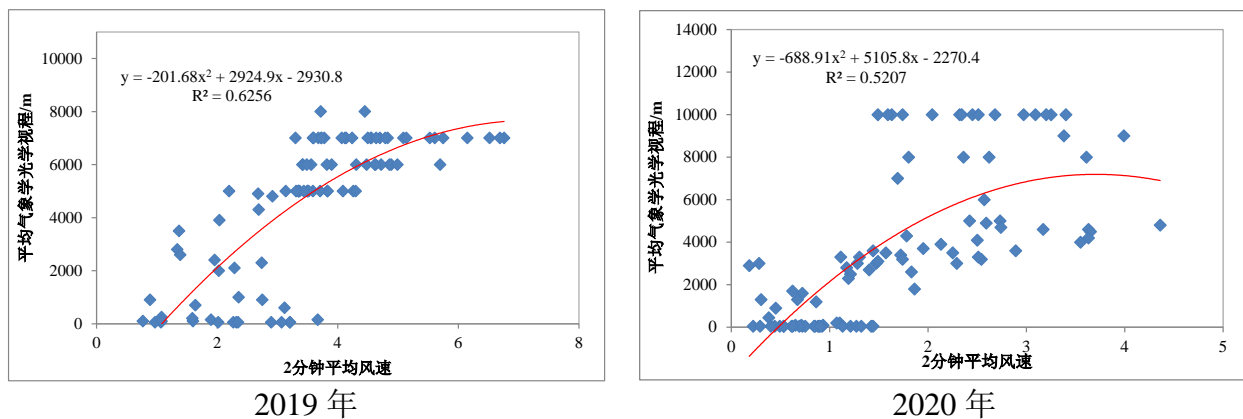


图 4-5 能见度与风速的非线性拟合

#### 4.5.2 能见度与本站气压的关系

气压作为重要的气象因子之一，对能见度也有较大的影响，本文数据中与气压相关的数据有本站气压、修正海平面气压和飞机着陆地区最高气压。经 4.4 小节中的相关性分析和先验知识，我们认为该三个变量间具有较强的共线关系，因而只选择对能见度相关性较大的本地压强作为气压因子。对本章数据中的气压进行统计，结果如表 4-4 和图 4-6 所示。

表 4-4 不同气压区间能见度均值统计表

时间	本站气压 (pa)	≤1017	1017-1018	1018-1019	1019-1020	1020-1021	1021-1022	>1022
2019.12	MOR (m)	73.08	670.00	4511.76	5768.57	7000.00	7000.00	7000.00
时间	本站气压 (pa)	≤1013	1013-1014	1014-1015	1015-1016	1016-1017	1017-1018	>1018
2020.03	MOR (m)	10000.00	7744.44	6325.00	6260.00	4043.75	1560.34	58.82

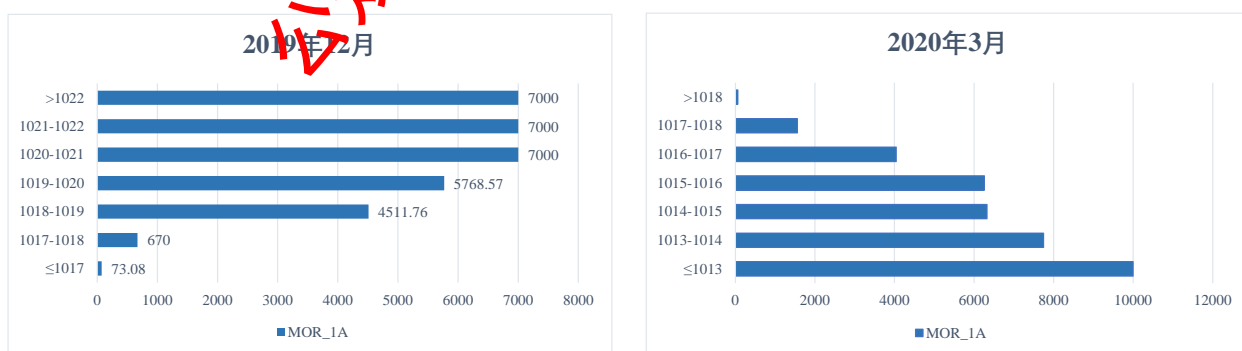


图 4-6 本站气压统计数据

从上表和图中介可以看出，不同的本站气压对能见度有较强的关联，同时春季的气压等级普遍低于冬季气压，但呈现出季节差异性，主要体现在冬季气压与能见度之间呈正相关，春季呈现负相关，这与 4.4 节中的分析一致；造成这种差异的主要原因在于气压对能见度的影响是间接的，它通过影响风速和温度，在冬季容易形成上升气流，导致地表气压局部低，地表气温低，相应能见度越大，而春季多为暖流，且压强本身较为稳定且强度低。

利用散点图构建能见度与气压之间的统计学关系，从图 4-7 中我们可以发现，气压与能见度之间并非简单的线性关系，从结果上看二者呈现三次函数关系，拟合结果如下图所

示，R2 系数分别为 0.9085 和 0.7083。

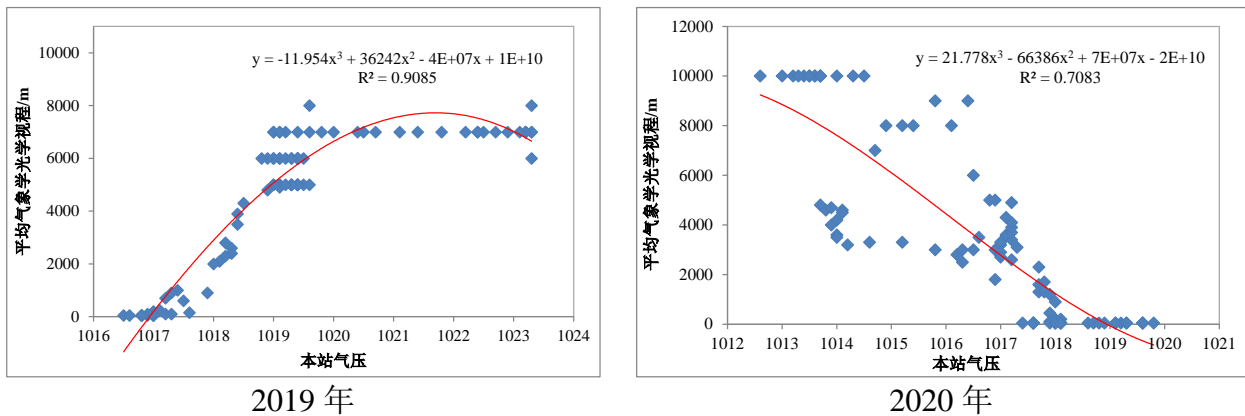


图 4-7 气压与能见度的非线性拟合

4.5.3 能见度与温度的关系

温度是影响能见度的一种重要气象因子，本文数据中与温度相关的数据有温度（TEMP）和露点温度（DWPOINT）。经 4.4 小节中的相关性分析和先验知识，我们认为二者间具有较强的共线关系，因而只选择与能见度相关性较大的温度（TEMP）作为温度因子，对本章数据中的温度进行统计，结果如表 4-5 和图 4-8 所示。

表 4-5 不同温度区间能见度均值统计表

时间	TEMP(°C)	≤8	8-10	10-12	12-14	14-16	16-20
2019.12	MOR_1A	175.00	2027.50	3553.13	5766.67	6382.35	7000.00
2020.03	MOR_1A	50.00	179.31	1876.93	3160	4085.71	8220.00

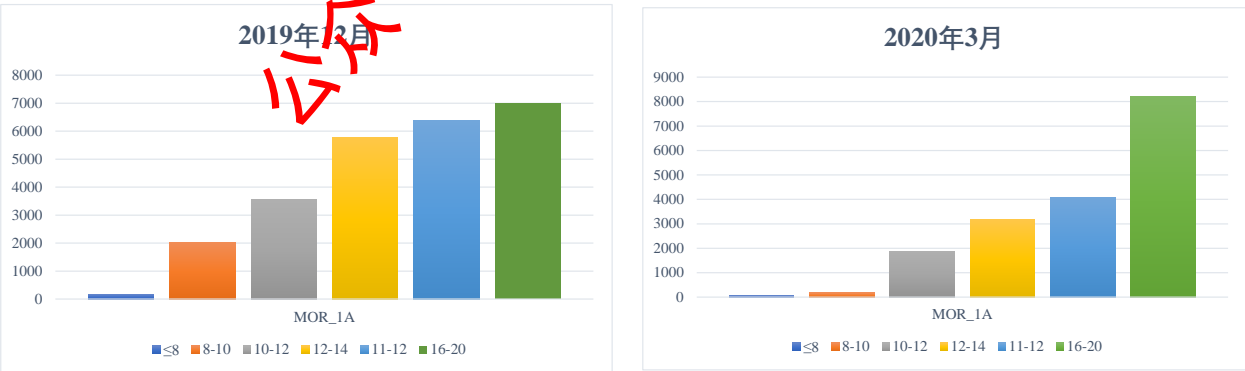


图 4-8 温度统计图

从上表和图中可以看出，温度与能见度之间呈现明显的正相关，且这种变化趋势在春季尤为明显，冬季随着温度的上升，能见度的上升趋势变缓，这与前面 4.4 小节分析的结论一致；温度对于能见度的影响主要作用于空气中污染物的扩散和移动作用，温度越高分子运动越快，同时温度的降低有利于雾气的形成。

利用散点图构建能见度与温度间的统计学关系，从图 4-9 中我们可以发现，气压与能见度之间并非简单的线性关系，从结果上看二者呈现四次函数关系，拟合结果如下图所示，R2 系数分别为 0.6649 和 0.9238。

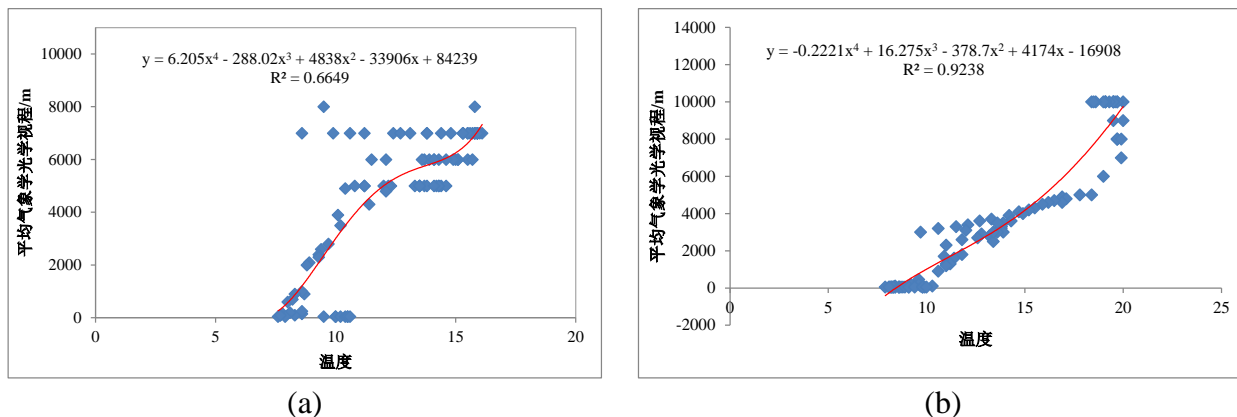


图 4-9 温度与能见度的非线性拟合(a):2019(b):2020

#### 4.5.4 能见度与相对湿度的关系

相对湿度也是影响能见度的重要因子之一，本章数据中的气压进行统计，结果如表 4-6 和图 4-10 所示。

表 4-6 不同湿度区间能见度均值统计表

时间	RH(%)	≤40	40-50	50-60	60-70	70-80	80-90	90-100
2019.12	MOR_1A	-	-	7000.00	6761.90	5761.90	5646.67	1142.19
2020.03	MOR_1A	9187.50	8875.00	4750.00	4212.50	3700.00	3250.00	805.43

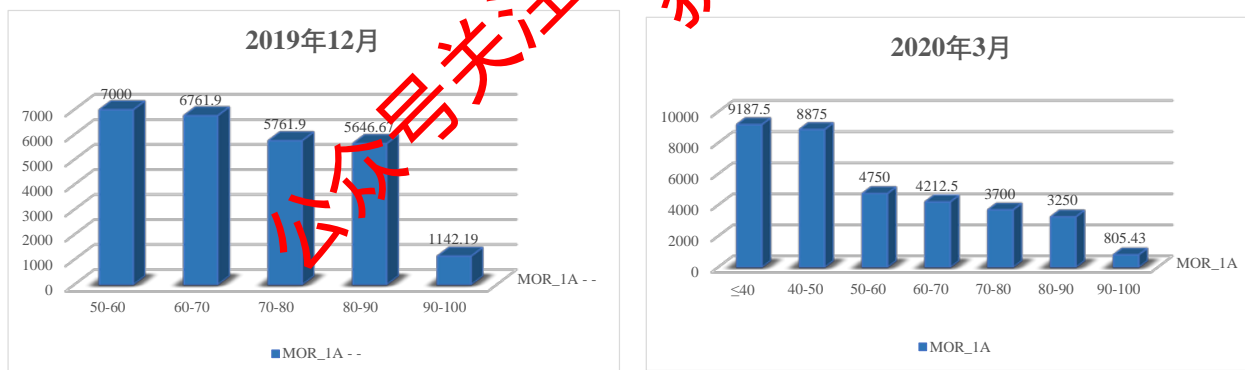


图 4-10 气压统计图

上表和图中可以看出，相对湿度与呈现较强的负相关关系，在季节上一致，这与 4.4 节中分析的结果一致；湿度对能见度的影响主要体现在对大气消光系数的影响，湿度增大时，气溶胶吸湿增长及消光系数明显增加，进而影响能见度，另一方面，在相对湿度大的环境下，水汽本身对光的吸收和散射会增大，空气中颗粒物容易被它们吸附，从而影响大气透明度。

利用散点图构建能见度与相对湿度的统计学关系，从图 4-10 中我们可以发现，气压与能见度之间并非简单的线性关系，从结果上看二者呈现二次函数关系，拟合结果如下图所示，R2 系数分别为 0.8635 和 0.9047。

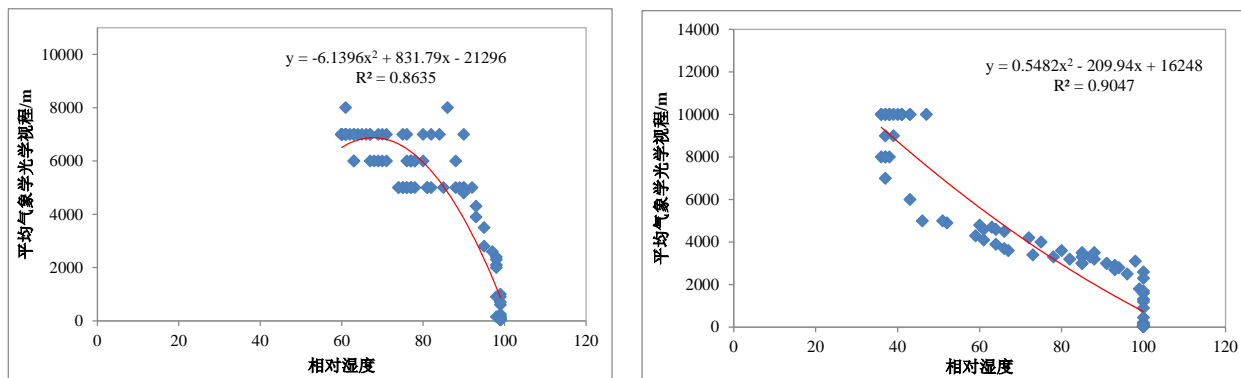


图 4-11 相对湿度与能见度的非线性拟合

## 4.6 能见度与气象因子的多变量非线性回归模型

### 4.6.1 能见度与气象因子的多元回归方程的确定

结合上述的因子分析，大气能见度与温度、气压、湿度和风速密切相关，为研究这些因子共同作用对大气能见度之间的影响，姑且假设这些因子对能见度之间的共同作用是加性的，因而其回归拟合公式为：

$$VIS = a_0 + a_1 * vs^2 + a_2 * vs + a_3 * pa^3 + a_4 * pa^2 + a_5 * pa + a_6 * t^4 + a_7 * t^3 + a_8 * t^2 + a_9 * t + a_{10} * u^2 + a_{11} * u \quad (4-1)$$

式中， $VIS$  表示大气能见度(m)， $vs$  表示风速(m/min)， $pa$  表示压强(pa)， $t$  为温度( $^{\circ}C$ )， $u$  为相对湿度(%),  $a_0 - a_{11}$  为各项系数。以此为目标方程，利用本章提供的样本数据，通过遗传算法进行最优系数求解。

### 4.6.2 遗传算法求解多元非线性方程的最优回归系数

遗传算法是一种借鉴生物界进化规律演化而来的随机搜索算法，主要用于解决 N-P 问题（枚举问题），通常用于求解非线性方程的全局最优解。本章是解决能见度和气候因子多元非线性方程的最优化问题，因而可以使用遗传算法进行估计。

遗传算法是随机生成一定数量的候选解（个体）的抽象表示（染色体）的种群向更好的种群进化。进化从完全随机个体的种群开始，之后一代代地进行遗传与进化。在每一代中，需要对整个种群中的每个个体进行适应度的评价，基于适应度值（优先选择适应度高的个体）从当前种群中随机地选择多个个体，通过自然选择和突变产生新的生命种群，该种群在算法的下次迭代中成为当前种群。基本的过程如下所示：

1) 从多元非线性方程的反演参数解空间随机生成一组参数  $p_i = [a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}]$ , ( $i = 1, 2, \dots, n$ )，构成遗传算法的初始种群， $n$  为种群的规模；

2) 计算每个个体 ( $p_i$ ) 对应的估计值，将该参数带入到多元非线性方程中计算各时刻能见度值；

3) 对比分析该个体  $p_i$  解析的结果与实际能见度值的拟合度指标  $R_i^2$ ，把它作为遗传算法的适应度值（评价函数）；

4) 选择操作：以每个个体  $p_i$  的适应度值  $R_i^2$  作为检验概率，采用蒙特卡洛模拟有放回的抽样方法从种群中选择  $n$  个新个体组成新的群落；

5) 交叉操作：从新群落中随机抽取两个个体，对组成个体的部分串进行交换，生成

两个新的个体，如果新的个体具有更好的适应度，则交换成功，否则将个体退回交换前的状态。如此重复多次，可以得到新的种群；

6) 变异操作：从新种群中随机选择部分个体，对个体中部分串进行随机扰动，是新的种群发生突变；

7) 对选择、交叉和变异后的新种群进行整体评价，若新生成的种群达到要求（新的群落比就群落没有明显改善或者其他限定条件），则结束遗传算法，最终生成的新种群中适应度最高的个体即为能见度与气象因子的多元非线性回归放方程的近似最优解；若没有达到幽囚，则返回 2)，知道新生成的种群达到要求为止。

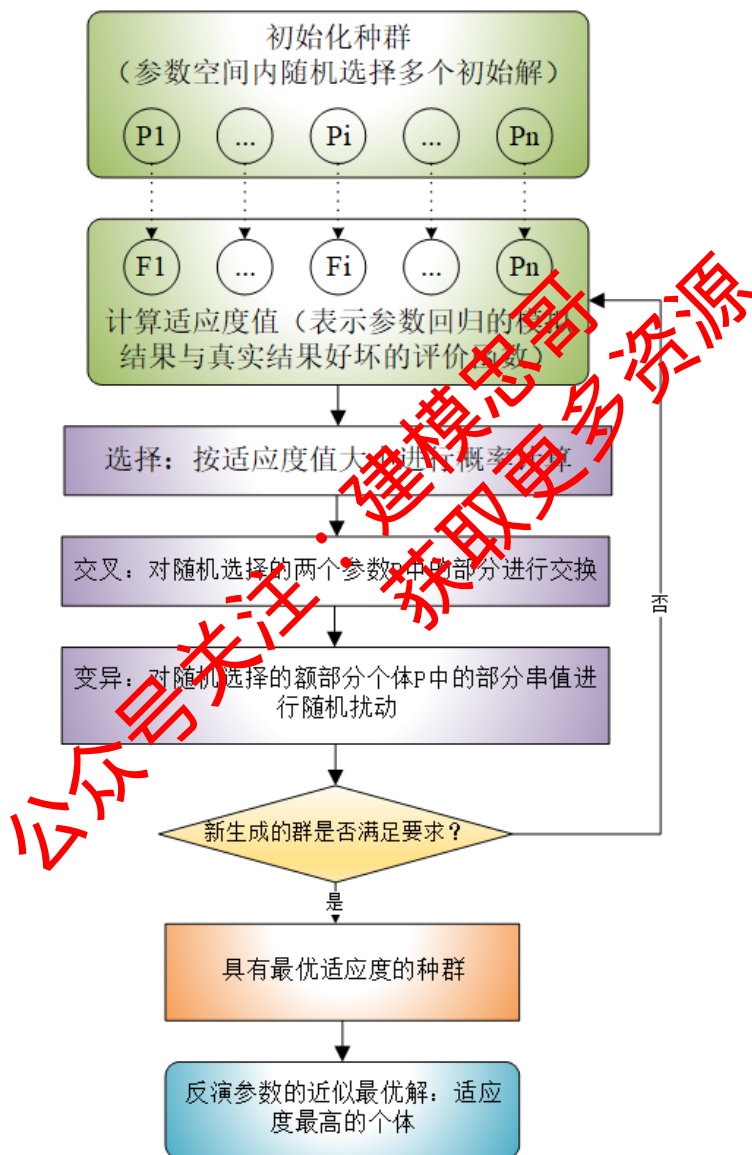


图 4-12 遗传算法流程图

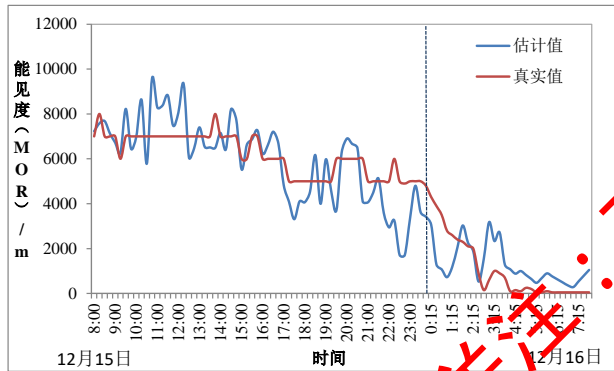
#### 4.6.3 结果分析

从上文的分析可知，多种气象因子与能见度呈现季节性差异，因而在结果中我们将本章数据分为春季（2020 年 3 月）和冬季（2019 年 12 月）两种情况分别进行最优参数求解。同时考虑到遗传算法的局部最优现象，我们假设遗传算法发生一个局部最优的可能性为  $p$ ，则  $n$  次重复试验中，发生局部最优的可能性为  $p^n$ ，所以下面进行重复 200 次遗传算法，选取  $R^2$  值最大的作为最终求解，结果如下所示。

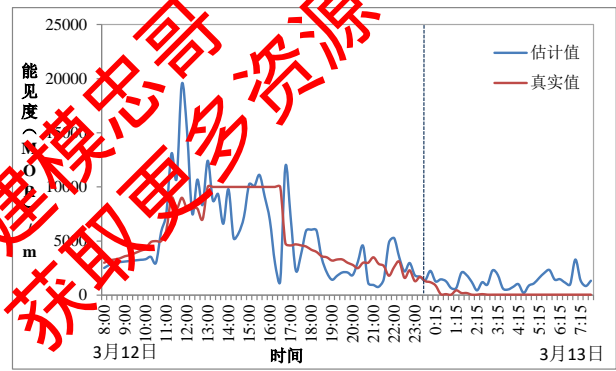


表 4-7 多次重复遗传算法的参数求解结果

时间	a0	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	R2
春季	487.42	590.75	1068.76	0.00023	-0.33	69.28	0.43	-7.36	9.04	760.36	2.83	-94.37	-0.04
	572.55	394.32	949.19	-0.00043	0.20	220.96	-0.36	16.08	-130.95	859.18	5.21	-320.23	-0.11
	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.
冬季	228.22	82.69	882.93	0.00029	-0.34	38.71	0.46	-7.25	28.33	298.11	1.69	-65.15	0.52
	468.26	122.46	1051.23	-0.0028	2.90	38.051	-0.60	1.28	-33.67	79.49	-6.58	-83.34	0.03
	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.



冬季



春季

图 4-13 非线性拟合结果

从图 4-13 中可以看出，利用多元非线性回归模型模拟的能见度变化与真实值较为接近，能较好的反应出能见度的变化趋势，R2 值分别为 0.5182 和-0.0381，由此证明该模型具有一定的科学性和合理性。

从上表中可以看出，多次执行遗传算法的 R2 值不相同，这是由于遗传算法求取是近似最优解，我们选择 R2 最大的一次遗传算法求解结果作为最终的模型参数。最终的关系式为：

$$\begin{cases} \text{春季: } VIS_{\text{春}} = 487.42 + 590.75 * vs^2 + 1068.76 * vs + 0.00023 * pa^3 - 0.33 * pa^2 + 69.28 * pa \\ \quad + 0.43 * t^4 - 7.36 * t^3 + 9.04 * t^2 + 760.36 * t + 2.83 * u^2 - 94.37 * u \\ \text{冬季: } VIS_{\text{冬}} = 228.22 + 82.69 * vs^2 + 882.93 * vs + 0.00029 * pa^3 - 0.34 * pa^2 + 38.71 * pa \\ \quad + 0.46 * t^4 - 7.25 * t^3 + 28.33 * t^2 + 298.11 * t + 1.69 * u^2 - 65.15 * u \end{cases}$$

(4-2)



## 5. 问题二建模与求解

### 5.1 问题分析

问题二要求根据题目提供的某机场视频数据和能见度数据，利用深度学习模型建立视频数据的能见度估计模型，并要求对估计结果进行精度评估。本题实际上是要寻找和挖掘影响能见度检测的视频图像特征信息，为雾天能见度反演模型的构建提供重要的理论支持。基于此目的，本章尝试通过大气光学分析与图像处理和人工智能技术相结合对能见度数据与视频图像数据进行多特征提取和分析，寻找并求证影响雾天能见度检测的影响因素及其相关性程度，以此建立视频图像与真实场景之间的关系。进而在此基础上充分考虑图像特征的连续变化信息构建基于视频图像数据的雾天能见度反演模型，并对反演结果进行精度评价，具体流程如下：

首先，需将视频转化为关键帧，并对关键帧图像进行选择建立与机场实测能见度数据的对应关系；

然后对图像质量较差的进行预处理，包括 ROI 选取、降噪处理、增强处理，剔除和替换等；下一步将对图像进行多特征计算，并基于分窗格处理提取图像兴趣窗口内的特征向量，通过相关性分析筛选最能影响雾天能见度反演的图像特征向量以及图像特征分窗的最佳窗口大小；

其次，通过分析每个兴趣窗口内图像特征与能见度的相关性大小提取各图像特征对应的最佳兴趣窗口，并计算最优兴趣窗口内的图像特征均值作为自变量矩阵，同时利用多元线性回归方程建立特征值与能见度之间的函数模型。

由于自然界中没有绝对的线性关系的存在，且考虑到 BP 神经网络同时兼顾线性和非线性的特点，利用其对于多元线性拟合值与实测能见度值之间的残差进行修正，从而构建基于深度学习修正线性回归残差的组合模型，即利用线性回归模型对能见度进行初步估计，并通过 BP 神经网络修正线性残差，从而实现对不同能见度的精准反演。

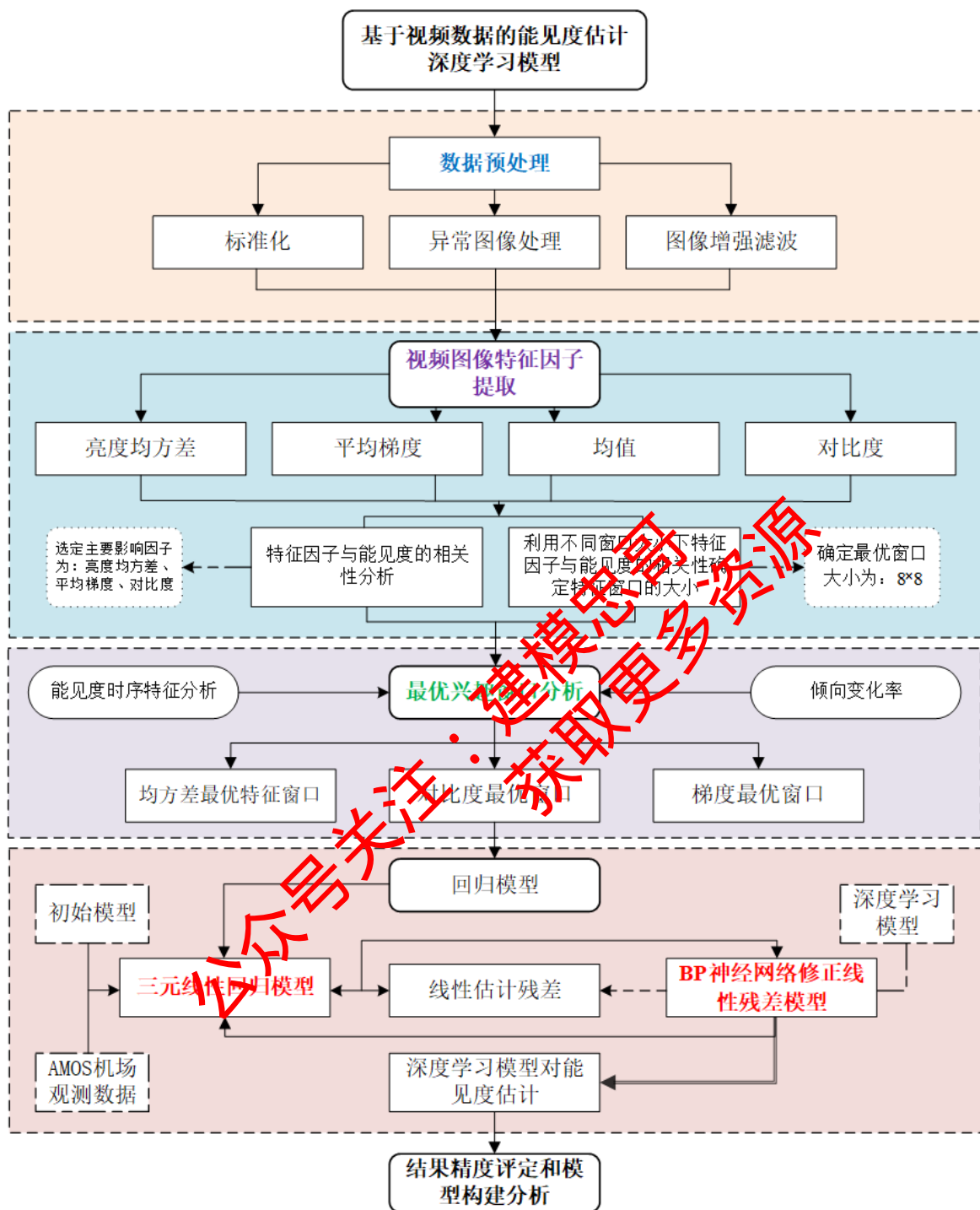


图 5-1 问题二的解决流程

## 5.2 视频数据预处理

### 5.2.1 视频转关键帧图像

本章数据中所给的视频数据是从 2020 年 3 月 13 日 00 时 00 分 26 秒开始到上午 11 时 47 分 48 秒，时间间隔为 1 秒，而机场 AMOS 观测数据中，实测能见度是从 2020 年 3 月 12 日 8 时开始到 2020 年 3 月 13 日 7 时 59 分 45 秒，时间间隔为 15 秒。为保证建立视频信息与实测能见度值的对应关系，我们以 1 分钟为时间间隔提取视频的关键帧照片，并将实测能见度也以 1 分钟为间隔建立与关键帧图像的对应关系。



图 5-2 一分钟为间隔的关键帧序列

### 5.2.2 异常图像处理

在实际的转化过程中，有一些图像出现虚晃、模糊和颜色变化等差异。对于单个图像出现异常问题的，我们选用其相邻 1 分钟间隔内图像进行替代，如下图 5-3 所示。对于连续图像序列出现问题的应该允与剔除，如下图 5-4 所示



图 5-3 单个异常图像



图 5-4 连续异常图像序列

### 5.2.3 ROI 选取

提取的图像中包含了部分与能见度无关的因素，例如顶部变化的文本信息，些因素在特征提取时不能正确反映能见度值，在后期模型训练中会形成误差。根据能见度的定义以及人眼在观察周围目标物时，亮度对比  $k = \frac{B_l - B_g}{B_g}$  若小于 0.02，人眼将无法判别目标物<sup>[8]</sup>。

因而，我们选取 ROI 的准则为：

- 1) ROI 内存在亮度对比度差异大的目标物；
- 2) ROI 内不包括文本信息等；
- 3) ROI 内无明显移动物体。最终我们截取中间部分，像素大小为 1000\*400，如图 5-5 所示。

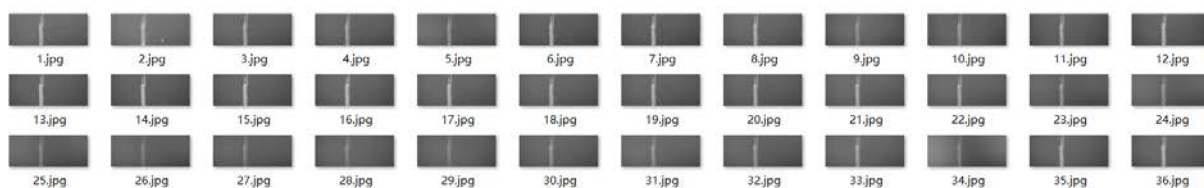


图 5-5 裁剪后关键帧序列

#### 5.2.4 降噪处理和增强处理

监测摄像头在正常工作情况下可能会受到多种因素的影响，在生成图像的过程中产生干扰噪声。同时，还需要对原始图像进行增强操作，消除由于环境光照不均而产生的图像质量不一问题，增强整体对比度，使模型能获取更加准确的图像特征。本文使用的图像增强方法为 **Retinex**，降噪方法选择中值滤波。

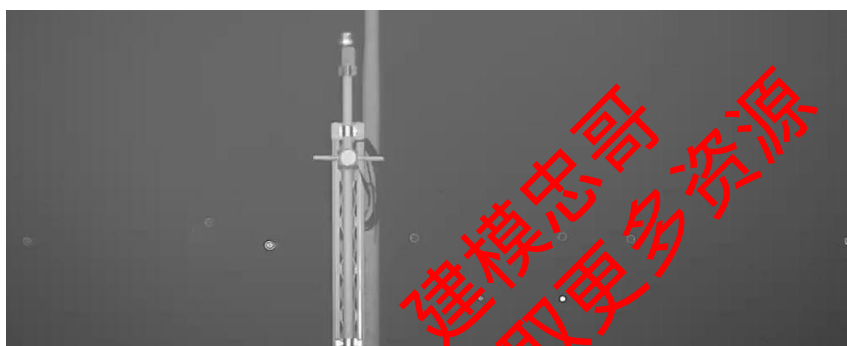


图 5-6 原始图像

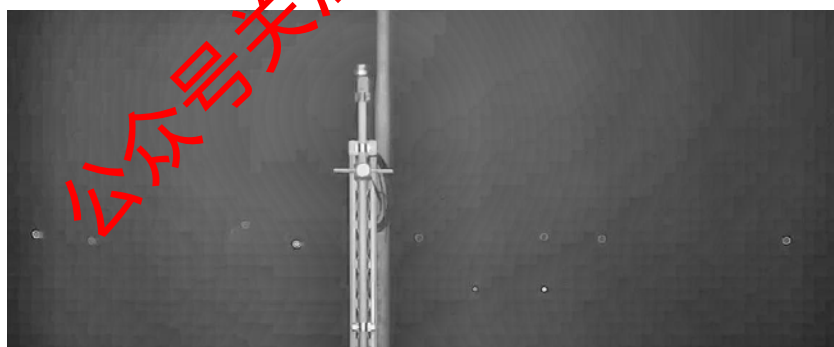


图 5-7 增强图像

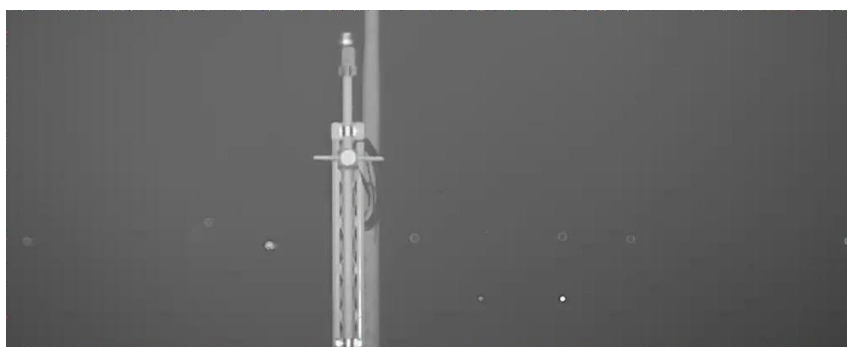


图 5-8 降噪后图像

### 5.3 能见度时序变化特征分析

将处理后的能见度序列数据进行可视化，并计算相邻时间的变化率，结果如下图 5-9 所示。

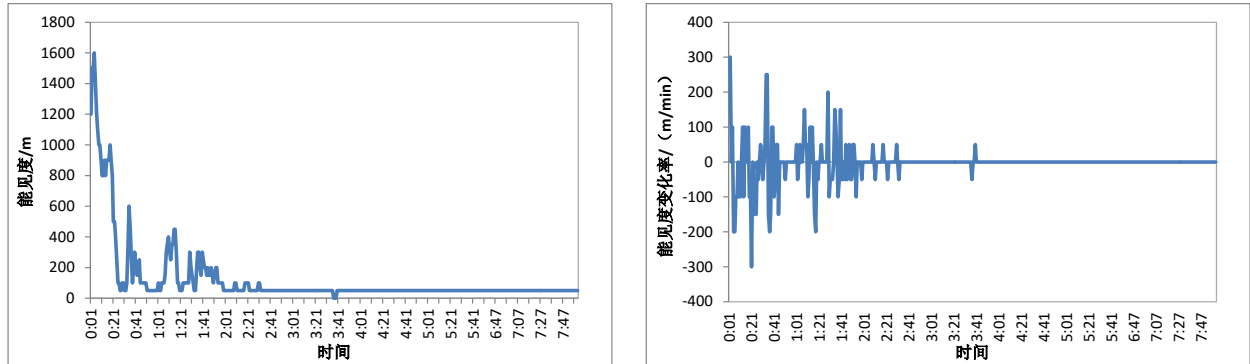


图 5-9 能见度时序变化特征图

从图 5-7 中可以看出 2020 年 3 月 13 日 0 时 1 分到早上 8 点的能见度变化整体呈下降趋势，并且其变化趋势呈现一定的周期性特征。

### 5.4 图像特征因子分析

#### 5.4.1 图像特征因子

基于视频数据的能见度估计，最重要的问题就是从图像中寻找与能见度相关的图像特征，图像特征包括颜色特征、纹理特征等，通过从图像中挖掘这些有效的特征信息，可以构建这些特征组成的特征向量与能见度之间的内在关系，从而计算能见度值。通过多方面的搜集资料和查阅文献<sup>[8][9][10][11]</sup>，我们选择图像的方差、图像均值、图像对比度和图像平均梯度作为特征因子。

##### 1) 图像均方差特征

光学条件下判别能见度的关键性因素是背景与目标物的对比，图像的均方差是用来衡量窗格内的像素特征值变化幅度，可以反映该窗格内的离散程度，我们通过视觉感知对比度来描述场景中目标物的相关信息，一般来说，均方差越大，图像越清晰，相应的能见度也越大，反之图像越模糊，能见度越小，因此，我们需要将彩色图像 RGB 颜色空间转化为 HSI 颜色空间进行处理，提取亮度分量，计算公式如下式(5-1)所示：

$$I = \frac{1}{3}(R + G + B) \quad (5-1)$$

$$\sigma = \sqrt{\frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n (I(i, j) - u)^2} \quad (5-2)$$

式中， $R$ 、 $G$ 、 $B$ 分别为彩色空间中的红绿蓝三个波段， $I$ 为亮度特征值。 $\sigma$ 为均方差， $I$ 为像素函数， $m$ ， $n$ 为窗格的长和宽， $i$ ， $j$ 为窗格内的像素坐标。

##### 2) 图像平均梯度特征

边缘信息是图像特征中重要的一种，人们通常比较关注物体的边缘信息，边缘信息可以获取物体的形状大小等信息。由于目标物距离比较远或者能见度较低，人通常无法分辨物体的边缘信息，因而边缘信息与能见度具有高度相关性。我们通过计算平均梯度的大小反正图像边缘细节的变化率，即表示图像清晰度，平均梯度越大，图像越清晰，对应的能见度也越大，反之图像越模糊，能见度越低，其计算公式如式(5-3)所示：



$$g = \frac{1}{(m-1) \times (n-1)} \times \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sqrt{\frac{(f(i, j) - f(i+1, j))^2 + (f(i, j) - f(i, j+1))^2}{2}} \quad (5-3)$$

式中， $g$  为平均梯度， $f$  为像素函数， $m, n$  为窗格的长和宽， $i, j$  为窗格内的像素坐标。

### 3) 图像对比度特征

人眼对高对比度的图像目标轮廓区域非常敏感，图像对比度在一定程度上也能反映目标物的轮廓以及物体内部纹理信息，其基本思想是去除灰度图像的边缘像素点，其他任意像素  $(i, j)$  相邻的像素为  $(i', j')$ ，则像素  $(i, j)$  可通过式(5-4)计算的到 8 个对比度特征是  $C$ ，选择 8 个对比度特征值中最大值最为像素  $(i, j)$  最终的对比度特征值，并计算该窗格内对比度的均值作为该窗格的特征。

$$C = \frac{|f(i, j) - f(i', j')|}{\max(f(i, j), f(i', j'))} \quad (5-4)$$

### 4) 图像均值特征

人的视觉除了对图像对比度和边缘信息较为敏感外，还对图像的颜色深浅和明暗程度较为敏感。通过计算兴趣窗口内像素均值来表征图像明暗程度、颜色深浅，其计算公式如下式：

$$\mu = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n f(i, j) \quad (5-5)$$

式中， $\mu$  为均值， $f$  为像素函数， $m, n$  为窗格的长和宽， $i, j$  为窗格内的像素坐标。

#### 5.4.2 图像子窗格划分

为了更好地捕获兴趣区域的像素特征，减少噪音影响。我们需要对图像进行分窗格处理，即按照  $n \times n$  的大小，对图像进行区域划分，结果如下图所示，红色方格即为一个区域。下面特征因子的提取均是以一个窗格为基本单元统计计算得到。

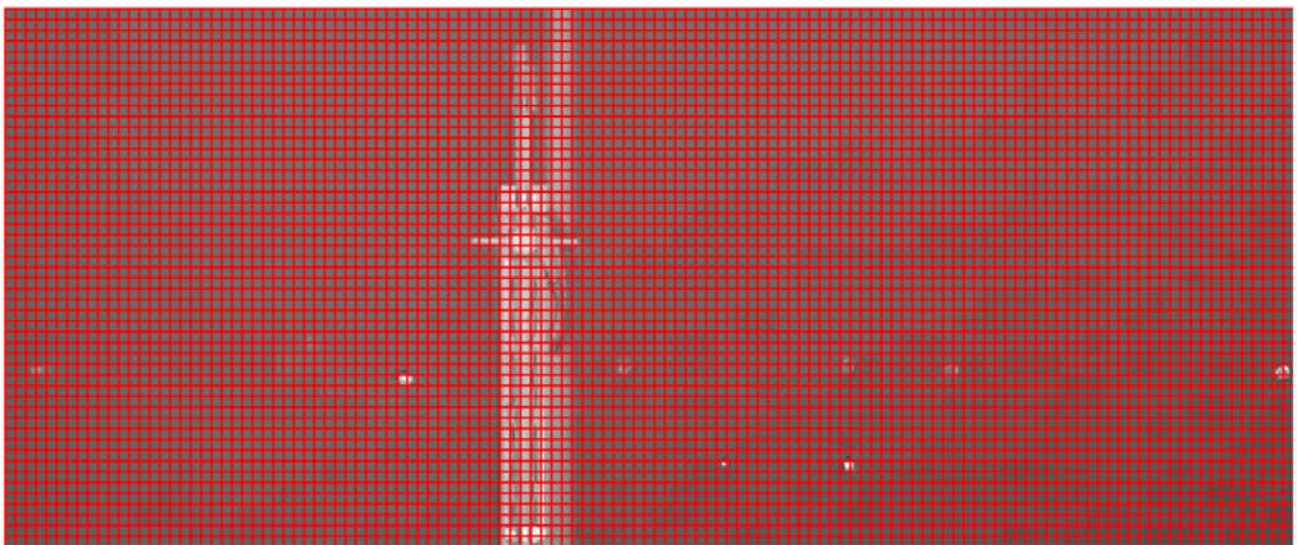


图 5-10 图像子窗格分布

### 5.4.3 特征因子与能见度的相关性分析

基于视频的能见度反演模型利用的是大气光学分析与数字图像图像处理及人工智能技术结合提取与能见度高度相关的图像特征，进而估计能见度信息。因此首要解决的问题就是寻找和求证与能见度高度相关的图像特征。

1) 我们先以  $10 \times 10$  的兴趣窗格大小计算各窗格内的图像平均梯度、图像亮度均方差、图像均值和图像对比度特征  $(\sigma_{ij}, g_{ij}, c_{ij}, \mu_{ij})$ ， $i$  表示影像编号， $j$  表示一幅图像中的窗口编号；

2) 求取这些特征值的均值，即计算  $\sigma_i = \sum_{j=1}^{end} \sigma_{ij}$ 、 $g_i = \sum_{j=1}^{end} g_{ij}$ 、 $c_i = \sum_{j=1}^{end} c_{ij}$ 、 $\mu_i = \sum_{j=1}^{end} \mu_{ij}$ ，得到对应影像的特征向量  $(\sigma_i, g_i, c_i, \mu_i)$ ，并与对应能见度建立联系形成相关因子矩阵  $(\sigma_i, g_i, c_i, \mu_i, VIS_i)$ ；

3) 以此循环所有图像，即可得到一个  $n \times n$  维的相关因子矩阵， $n$  为关键帧图像数量；

4) 分别进行 Kendall、Pearson 和 Spearman 相关性分析。

结果如下图 5-11 所示。

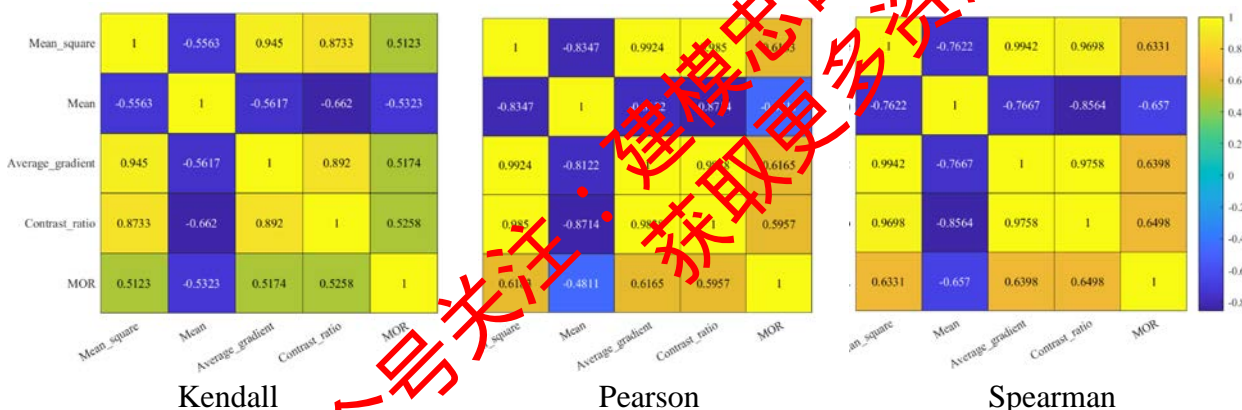


图 5-11 各图像特征因子与能见度的相关系数图

从图 5-11 中可以看出三种方法计算相关系数具有较高的一致性，其中图像亮度均方差 (Mean\_square)、平均梯度 (Average\_gradient) 和对比度 (Contrast\_ratio) 与能见度 (MOR) 均具有加强的相关性，但是图像均值与能见度的相关系数不稳定且较其他因子明显偏低，导致这种现象的原因可能是均值描述的窗格的敏感程度，易收到光线强弱的影响，从视频中可以看出检测时间基本为夜间，所以均值对能见度影响难以评估。因而我们最终选择图像亮度均方差、平均梯度和对比度作为最终的图像特征因子，构建基于视频数据的大雾天气下能见度反演模型。

### 5.4.4 不同窗格下图像特征因子与能见度的相关性分析

基于图像对能见度进行估计，首先就需要对图像进行分窗处理，在提取兴趣区域内的特征因子，因而不同的窗格大小可能会对图像特征提取的精度有影响。因而本节分别进行不同大小的窗格下的特征因子与能见度的相关性分析，这里的特征因子只选择上节中最后选出的三种因子，窗口大小分别设为  $3 \times 3$ 、 $5 \times 5$ 、 $8 \times 8$ 、 $10 \times 10$ 、 $12 \times 12$ 、 $20 \times 20$ 、 $25 \times 25$  和  $30 \times 30$ ，结果如图(5-12)所示。



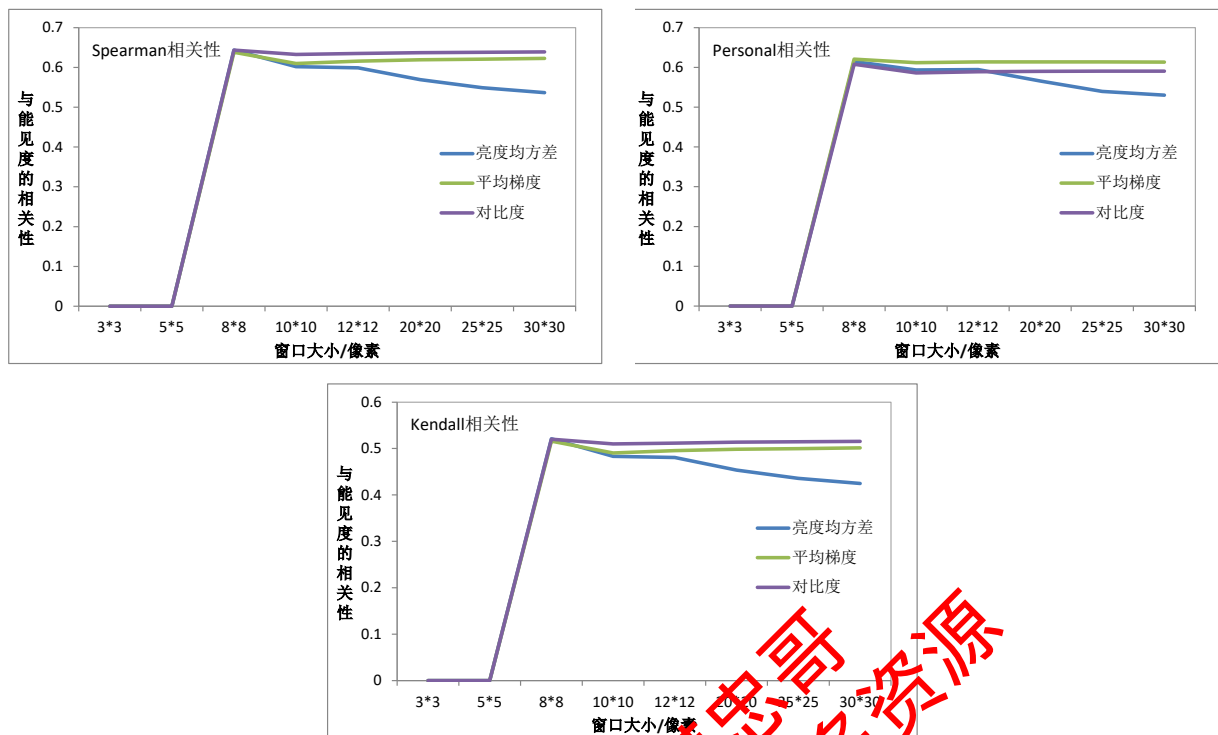


图 5-12 不同窗口下图像特征因子与能见度的相关性分析

从图 5-12 中可以看出随着窗口增加, 各特征因子与能见度的相关性呈现先增加后减小的趋势, 其中在  $8 \times 8$  的窗口大小达到最大值, 因此后续图像特征因子的窗口大小取  $8 \times 8$  个像素。

## 5.5 兴趣窗格的选择

图像中仅有部分区域是能见度对特征因子最敏感的区域, 这些区域也称为兴趣窗格。为得到不同特征对应的最优兴趣窗格, 即该窗格内该特征与能见度的相关性最大, 需要对不同特征因子下每个窗格与能见度的相关性, 以此选出最优兴趣窗格, 并将该窗格内某特征因子的平均值作为整幅图像该特征因子的实际数值, 进而与该幅图像的能见度构建变量矩阵, 用于反演模型构建。具体做法如下:

- 1) 我们先以  $8 \times 8$  的兴趣窗格大小计算各窗格内的图像平均梯度、图像亮度均方差、图像均值和图像对比度特征  $(\sigma_{ij}, g_{ij}, c_{ij})$ ,  $i$  表示影像编号,  $j$  表示一幅图像中的窗口编号;
- 2) 记录一幅图像的窗格数量为  $m$ , 构建每个窗格与该幅图像的能见度对应关系  $(reg_{i1}, reg_{i2}, \dots, reg_{im}, VIS_i)$ ;
- 3) 分别令  $reg_{ij} = \sigma_{ij}, g_{ij}, c_{ij}$ , 统计所有图像的上述相关因子矩阵, 进行分别进行相关性分析;
- 4) 选取相关因子排名前 10 的窗格作为最优的兴趣窗格, 进而构建反演模型的因变量和自变量矩阵  $(VIS_i, \sigma'_i, g'_i, c'_i)$ , 其中  $\sigma'_i, g'_i, c'_i$  的计算是基于选取的最优兴趣窗格而非所有窗格。

结果如下图 5-13 和图 5-14 所示, 图 5-13 中的折线图点编号为该兴趣窗口的编号。

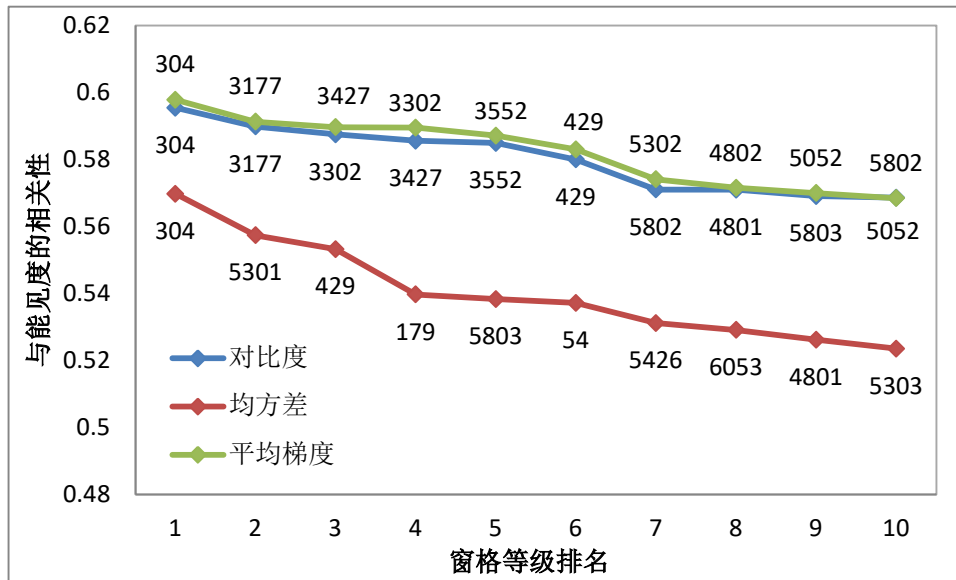
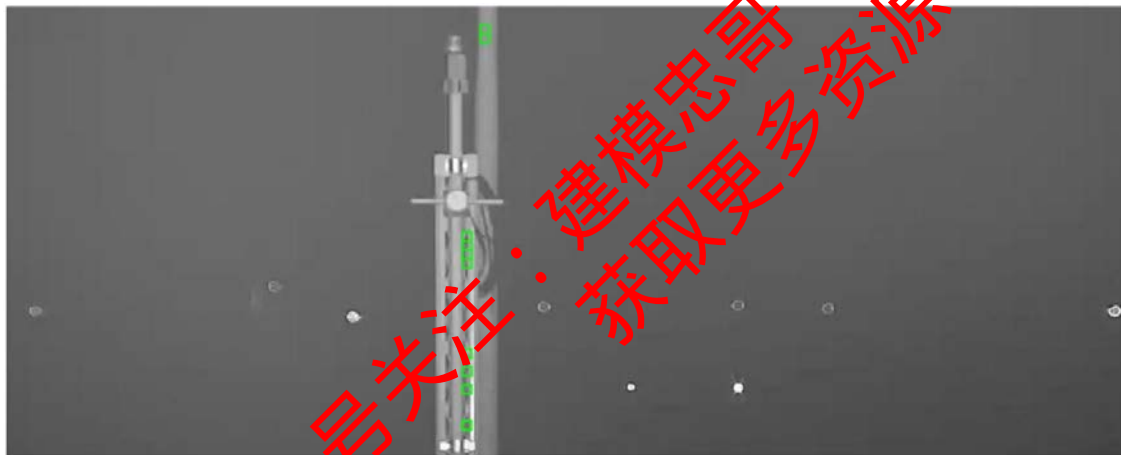
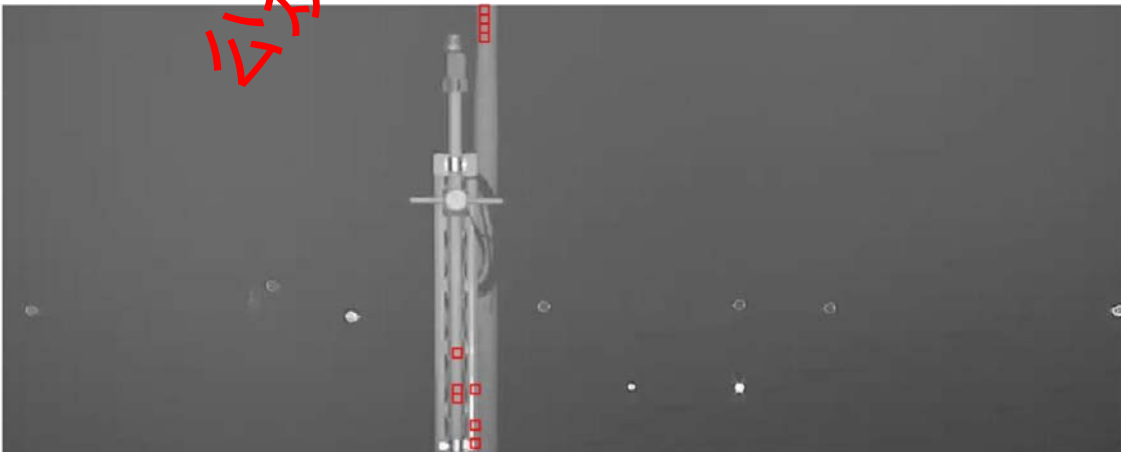


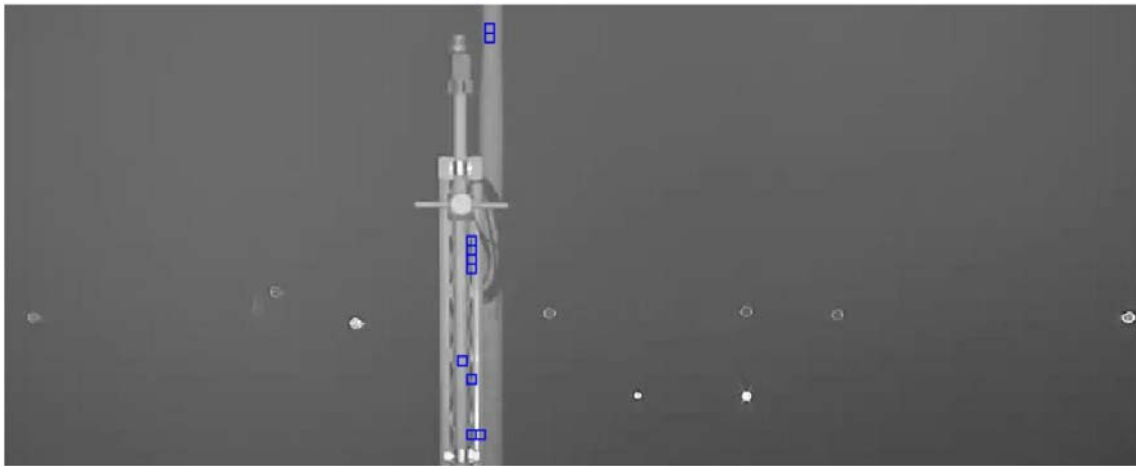
图 5-13 特征因子与能见度相关系数折线图



(a) 平均梯度因子对应最优兴趣窗口分布



(b) 均方差因子对应最优兴趣窗口分布



(c) 对比度因子对应最优兴趣窗口分布

图 5-14 各图像特征因子对应的最优兴趣窗口分布

从图 5-13 中可以看出排名前 10 的兴趣窗格与能见度的相关性保持较高的水平，均能达到 0.58 以上；不同排名的窗格之间相关性差异较小，整体较为稳定，其中对比度与平均梯度要明显高于均方差特征，这主要是因为该组图像拍摄于夜间，图像整体的亮度变化水平不高。

从图 5-14 总可以看出，这些特征因子对应的最优兴趣窗格主要分布在由明显建筑物轮廓的地方，如图中建筑物白色边缘与圆柱形表面，最优兴趣窗格的分布整体上较为集中，同时不同因子对应的最优兴趣窗格的分布具有高度的一致性，这主要是夜间拍摄的图像整体光线较暗，同时能见度较低，远处的景物基本模糊，图像中仅有一处近景的建筑物比较清晰，大部分的图像特征因子也分布在这里，因此最优兴趣窗口的分布一般集中在具有摄像头附近区域具有明显轮廓的目标物区域。

## 5.6 能见度反演模型的构建

线性回归模型是众多领域中常见的反演模型，但由于复杂的环境因素具有干扰性，对模型的反演效果造成很大影响，同时线性回归也较为单一，模型的精度有待提高。因而，我们构建了基于 BP 神经网络修正线性回归残差的组合模型，基本思想是首先利用线性回归对模型能见度进行初步粗糙的估计，然后利用 BP 神经网络修正线性残差，不仅仅发挥了线性回归模型的运算简单特点，同时又充分利用了 BP 神经网络兼顾线性和非线性的特点。

### 5.6.1 三元线性回归建模

实际中，能见度往往同时受到复杂环境下的多个因素影响，单一考虑一个因素显然时不够的，从大气光学和图像底层特征分析可知图像的对比度特征和边缘梯度特征等能描述能见度信息。因而我们尝试使用三元一次线性回归模型对能见度进行初步估计，模型的形式如下式所示。

$$VIS = \alpha_1 \times grad + \alpha_2 \times meas + \alpha_3 \times cons \quad (5-6)$$

式中， $VIS$  为能见度， $grad$  为梯度因子， $meas$  为均方差因子， $cons$  为对比度因子  $\alpha_1$ 、 $\alpha_2$ 、 $\alpha_3$  分别为回归系数，结果如下图\*所示，三元线性拟合的结果与实际数据在趋势上具有较好的一致性， $R^2$  系数为 0.3266。

表 5-1 三元线性回归模型参数

$\alpha_1$	$\alpha_2$	$\alpha_3$	$R^2$
-20.4735	23.9235	1969.8862	0.3266

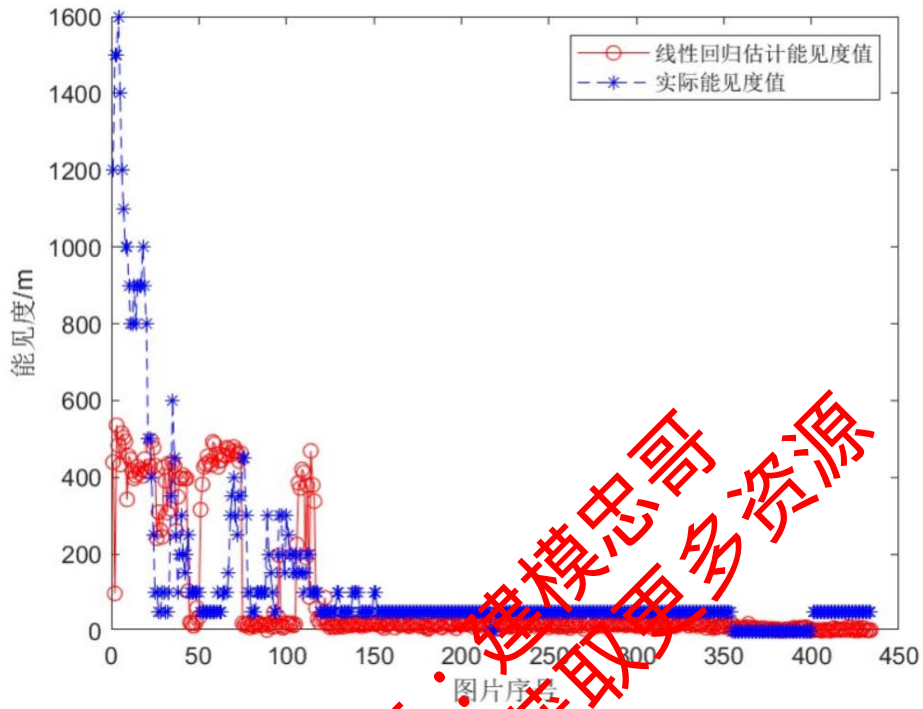


图 5-13 三元线性回归拟合结果

### 5.6.2 BP 神经网络

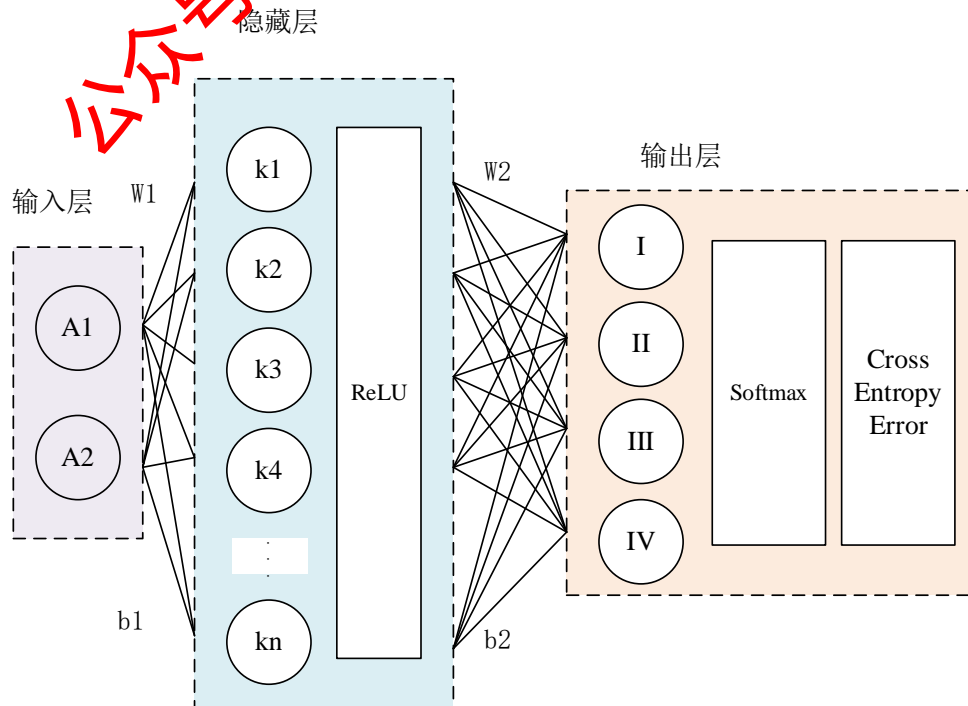


图 5-14 BP 神经网络结构

### 5.5.3 BP 神经网络修正线性残差模型的构建

残差修正模型目前已经应用于各个领域，利用 BP 神经网络来建立修正残差模型的主要步骤如下：

1) 利用三元线性回归模型对图像的特征序列  $(\sigma'_i, g'_i, c'_i)$  和能见度观测序列  $VIS_i$  进行拟合，得到能见度估计值  $\hat{VIS}_i$ ，计算残差  $e_i = VIS_i - \hat{VIS}_i$ ；

2) 建立残差序列与图像特征因子  $(e_i, \sigma'_i, g'_i, c'_i)$ ，进一步利用 BP 神经网络进行拟合，建立基于 BP 神经网络的残差修正模型，利用该模型可以预测残差序列  $\hat{e}_i$ ；

3) 将预测的残差序列  $\hat{e}_i$  与三元线性回归估计的  $\hat{VIS}_i$  进行相加即可得到基于 BP 神经网络修正线性残差组合模型的反演结果。

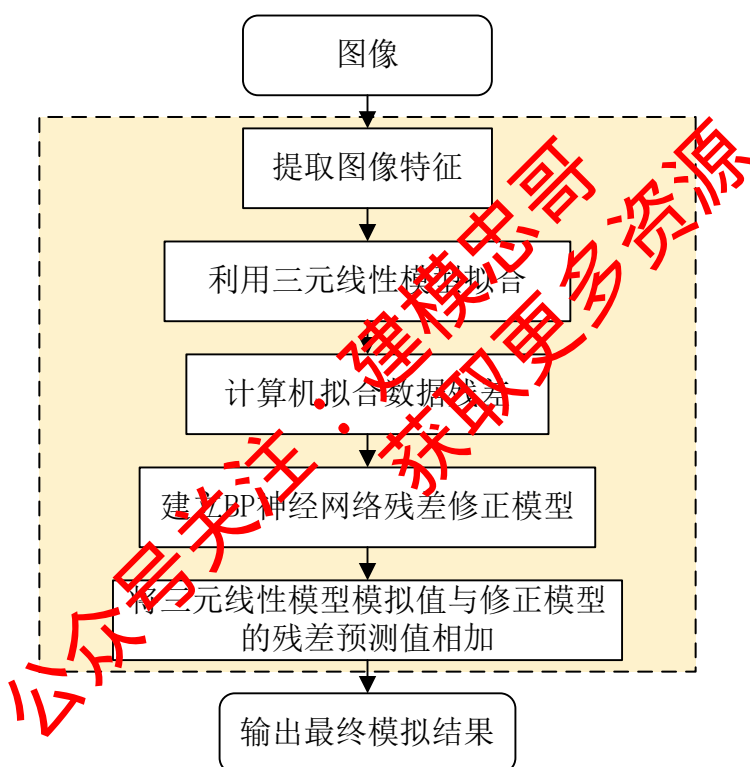


图 5-15 BP 神经网络修正线性残差组合模型

### 5.5.4 精度分析

我们选用 BP 神经网络的以  $e_i$  序列作为  $3 \times 15 \times 1$  的 BP 神经网络输出想，输入层和隐藏层均采用  $\tansing$  函数作为传输函数， $\text{purelin}$  函数作为输出层的传输函数，学习速率为 0.01，收敛误差为 0.01，最大迭代次数为 10000，结果如下图 5-16 所示。

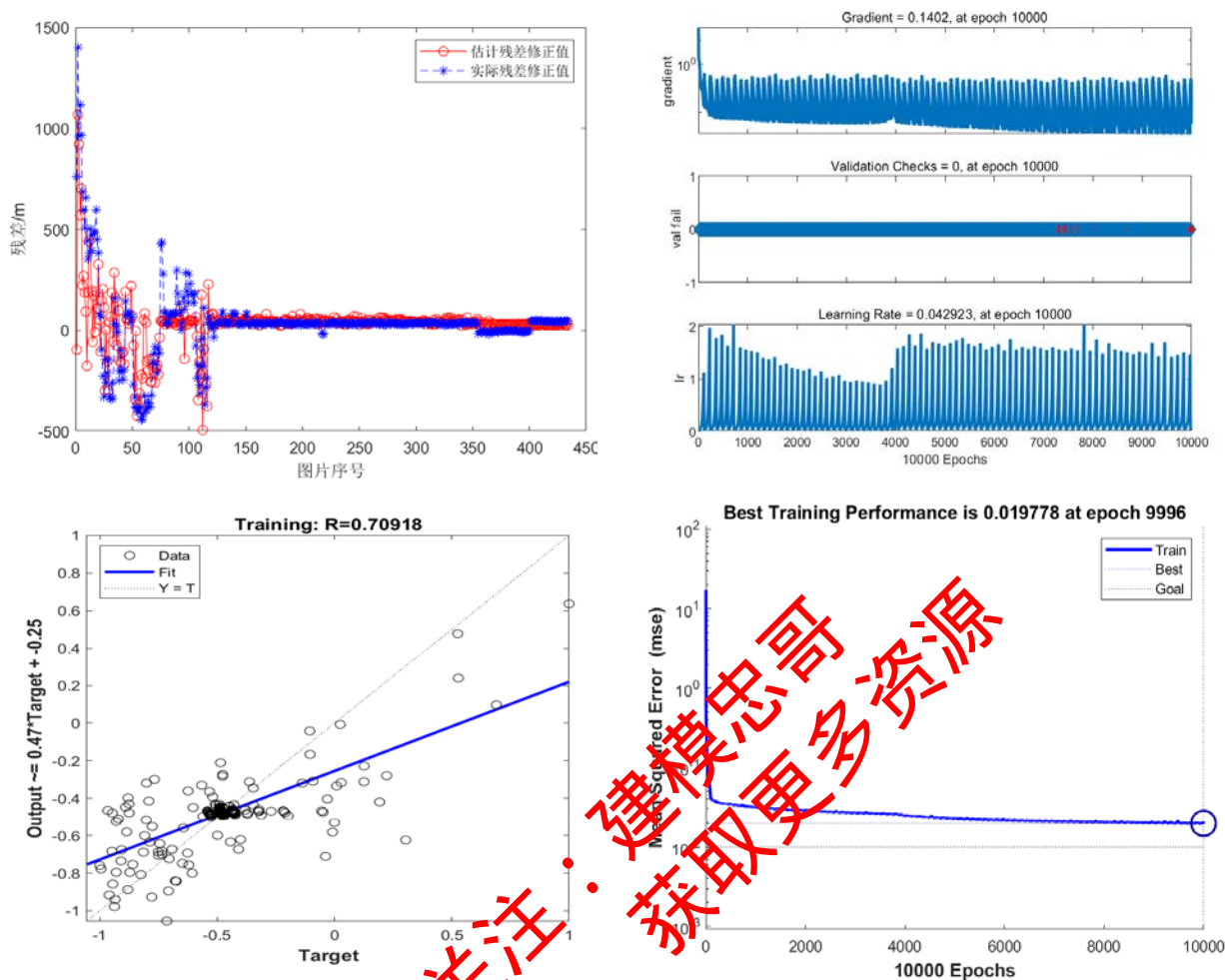


图 5-16 BP 神经网络拟合残差精度分析图

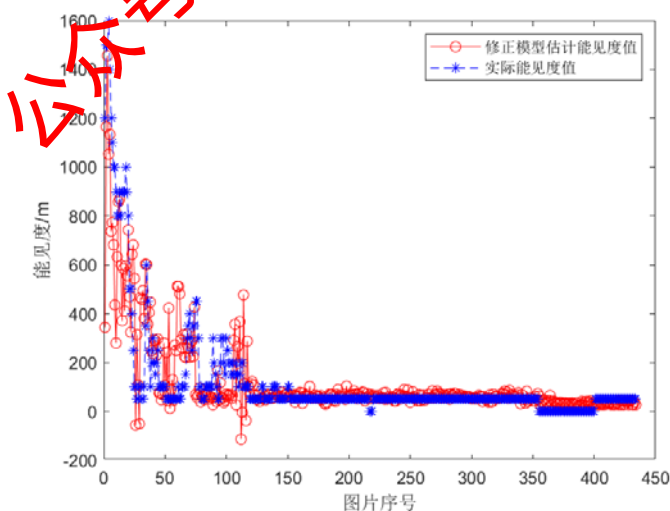


图 5-17 BP 神经网络修正线性残差后的能见度对比

从图 5-16 中可以看出 BP 神经网络预测的残差值与实际残差值具有高度的一致性，神经网络的训练结果逐渐平稳，不断向收敛误差逼近，最终的误差结语 0.01 和 0.1 之间，拟合的  $R^2$  系数为 0.70918。将残差带回到三元线性估计的能见度后得到修正了线性残差的最终能见度值，如图 5-17，其拟合结合较之前单纯采用线性拟合有较大的提升，整体趋势基本一致，计算  $R^2$  系数为 0.8342，这也充分证明了该方法的可靠性和正确性。



## 6. 问题三建模与求解

### 6.1 问题分析

问题三要求根据高速公路某路段的检测数据，利用图片中物体的不同景深，以及亮度差异，建立不依赖直接测量能见度数据的能见度估计算法。本题实际上是想建立获得的图片信息与能见度的关系，首先，能见度与大气光学传输作用息息相关，能见度为人眼可分辨出目标的最大水平距离，能否识别目标的关键点就在于该像素与背景的亮度差异是否明显，若大于人眼可分辨的阈值则可见，因此，基于亮度对比度的思想，找到距离相机最远可分辨的行像素，通过暗原色先验原理，计算改行目标像素的透射率，从而得到大气消光系数，获得能见度。

首先，需对图片进行预处理，去除无效信息的干扰，然后为了更好的识别图像的边缘特征，需进行图像增强，采用非线性图像增强的方法，抑制噪声的影响，再进行边缘检测，因为图片中有大面积的团雾，为了获的更好的效果，先后进行了 5 种边缘检测，包括梯度法边缘检测、sobel 算子边缘检测、Roberts 边缘算子检测、prewitt 边缘算子检测，进行对比，采用最好效果的算子进行检测，再经过霍夫变换检测车道线，并得灭点，即可对摄像机进行标定，求出路面上一点至摄像机的垂直距离，通过亮度对比度，寻找距离相机最远的可分辨像素的行坐标，即‘消失线’，选取该线上的 5 个像素点作为暗原色先验原理的分析目标，获得目标的透射率，得出大气消光系数，由 Koschmieder 定律，求得能见度。

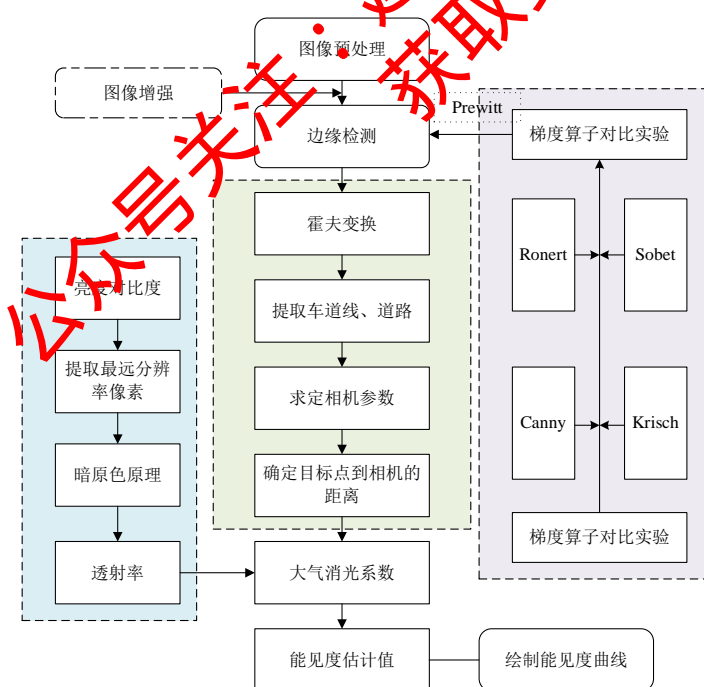


图 6-1 问题三技术流程图

### 6.2 数据预处理

#### 6.2.1 剔除无关信息

本章所给数据共计 100 张图片，时间范围是从 2016 年 4 月 14 日 6 点 30 分 26 秒开始到 2016 年 4 月 14 日 7 点 39 分 11 秒结束，路段标识为“K113+000”。图片中主要是以此



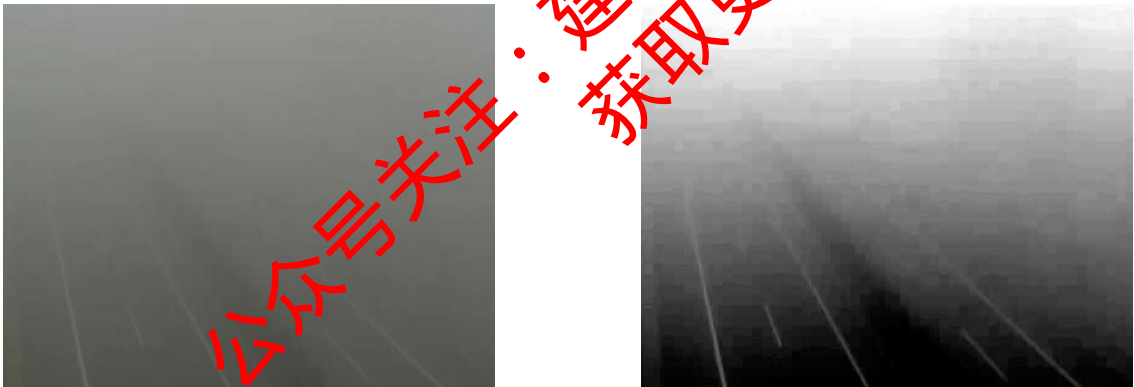
高速道路信息为主，距离摄像机越近的地方边缘信息越明显。图像中除了主要道路信息，还包含了监控日期、时间、路段标识等无关信息，会影响后续的检测效果，因此我们对图像进行裁剪，剔除 无关信息。



图 6-3 裁剪处理后的结果

6.2.2 图像增强

实际环境中存在着多种因素会影响相机拍摄的图像，为了增强图像的对比度，便于后续提取有效的图像特征信息，需要采用直方图均衡化增强图像的对比度。结果如下图 6-4 所示。从灰度直方图中可以看出，原图的灰度比较集中，经过直方图均衡化后，灰度分布广泛，从而达到增强对比度的效果，更容易区分不同灰度间的细节。



(a) 原始图像

(b) 直方图均衡化后

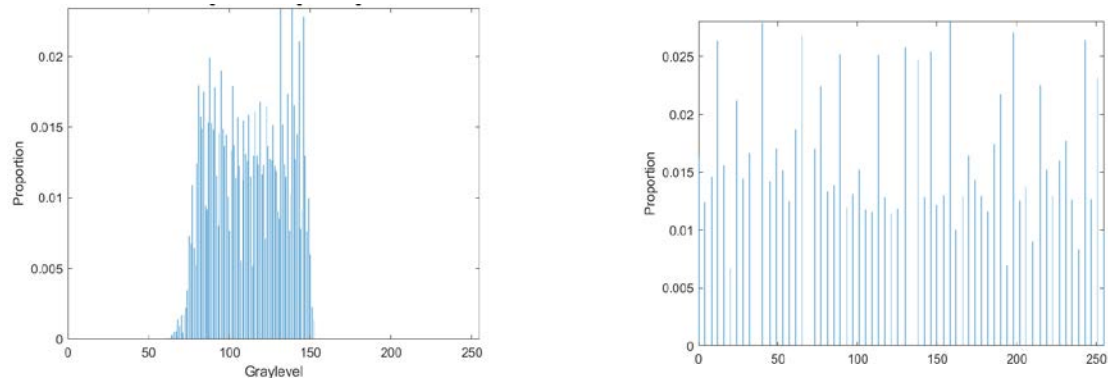


图 6-4 图像增强前后对比图

6.3 边缘检测

为了能提取出车道线，需对增强后得图像进行边缘检测，为了获得更好的效果，采用 5 中算子进行对比，分别是梯度法边缘检测、sobel 算子边缘检测、Roberts 边缘算子检测、

prewitt 算子边缘检测。结果如图 6-5 所示。其中，Prewitt 算子边缘检测的效果最好，Krisch 梯度算子与梯度法几乎没有检测出道路，Sobel 梯度算子边缘检测结果较好，Roberts 其次。因此本章采用基于 Prewitt 边缘检测的图像进行后续图像处理。

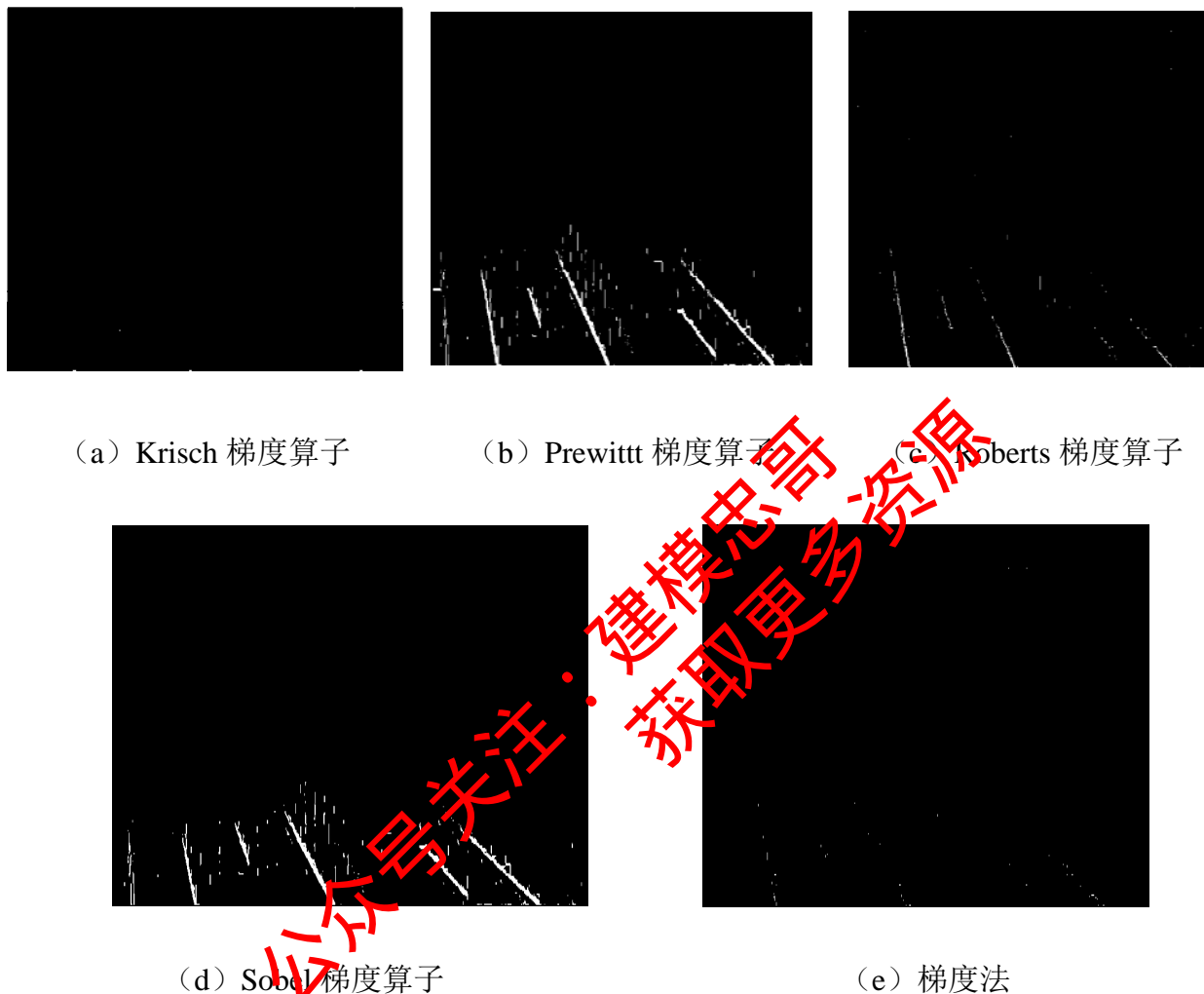


图 6-5 边缘检测图

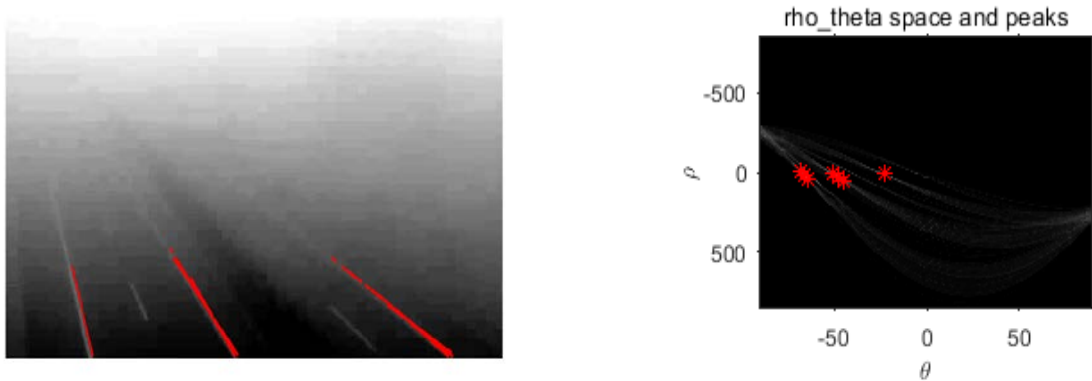
## 6.4 道路提取

### 6.4.1 霍夫变换

霍夫变换是一种特征检测(feature extraction),被广泛应用在图像分析(image analysis)、计算机视觉(computer vision)以及数位影像处理(digital image processing),主要是用来辨别找出物件中的特征,例如:线条。其实质是一种统计投票的决策算法,通过一点(x,y)的直线有  $n$  条,将这些直线对应至极坐标中,可以形成一条曲线,将所有存在直线的所有像素点映射值二维参数空间中去,并对参数频数进行统计叠加,频数越高说明该像素对应参数所生成的直线在原图相中用过的感兴趣点越多,即被检测出直线的概率越大。

将经过增强处理和边缘算子的图像带入霍夫变换算法,结果如图 6-6 所示。图中红色线条表示我们提取的主要车道线段,右图白线表示直线变换到极坐标的曲线,曲线上的像素点越明亮,说明该像素对应参数所生成的直线在原图像中通过的感兴趣像素点越多,红

点对应着道路上的点，每个点生成一条直线。其中，车道线的提取效果较好，部分短车道间距线未能提出，主要是因为车道线比道路线段短，在取频数最大时取不到该线段。



(a) 霍夫变换检测车道线 (b) 霍夫变换空间

图 6-6 霍夫变换提取道路线

6.5 道路兴趣区的提取

为了消除道路周边建筑物和树木等物体的干扰，后续的能见度检测过程将锁定在道路的区域，利用上述提取的车道线结果，并将其延长至交点，该区域称为道路兴趣区。具体的提取算法如下：

- 1) 初始化以该边缘点位中心的  $1 \times 2$  的窗口，确定矩形窗口的边长与白色像素相交的数量，若没有交点，则记录当前矩形窗口的大小和位置，若有交点，则进行第二步操作；
- 2) 以初始矩形窗的  $L$  倍进行外扩，若与边缘像素点没有交点则记录，反之，则进行第三步操作
- 3) 矩形窗口阈值大小设为  $T1$ ，若矩形窗口尺寸超过阈值大小，则终止，且不记录，跳转第一步对下一个边缘点开始检测，反之，则跳转第二步，直到图片中的所有像素全被遍历，若满足没有交点，矩形窗口记录为  $Rect_i(i = 1, 2, \dots)$ ，像素记录为  $P_i(1, 2, \dots)$ ，因为只需要车道线，设定矩形窗内白色像素点的最小阈值  $T2$ ，若小于这个值则舍去，矩形的端点坐标即为车道线的端点坐标。

结果如图 6-7 所示。图中左图即为我们提取的道路兴趣区。



图 6-7 道路兴趣区提取

## 6.6 相机标定

根据路况摄像机成像模型，建立路面坐标系、摄像机坐标系与成像坐标系的转换关系，推算出路面上一点到摄像机坐标的垂直距离与该点在像平面上对应像素点的关系<sup>[17]</sup>：

$$d = \begin{cases} \frac{\lambda}{S - S_v}, & S > S_v \\ \infty, & S \leq S_v \end{cases} \quad (6-1)$$

式中， $d$  为路面上一点到摄像机的垂直距离， $S$  表示路面上一点在图像中对应像素点的行坐标（自上向下）， $S_v$  表示灭点在像平面中的行坐标，即图 6-7 中的交点坐标，由霍夫变换可知道每条线的方程式，从而求出灭点坐标为（40，57），由于假设摄像机是一直不产生移动的，因此每副图的灭点是相同的。 $\lambda$  表示摄像机高度、焦距、摄像机光轴与路面夹角 3 个参数的关系式。

在算法理论分析的第二点中，我们知道高速公路上车道分界线的长度为 6m，因此对  $\lambda$  的求取只需要在图像上找到一条清晰的车道分界线，并获得该车道分界线的两端点行坐标即可求得  $\lambda$  值。假设车道分界线的头端到摄像机的距离为  $d_2$ ，在图像中的行坐标为  $S_2$ ；尾端到摄像机的距离为  $d_1$ ，在图像中的行坐标为  $S_1$ ，代入式 (6-2) 中得：

$$\lambda = \frac{d_2 - d_1}{\frac{1}{S_2 - S_v} - \frac{1}{S_1 - S_v}} \quad (6-2)$$

选取图 6-7 中的 A（203,234）、B（229,263）两点坐标，带入 6-1 和 6-2，解得  $d$  大约为 6.5。

## 6.7 亮度对比度提取道路最近可视点

图像道路兴趣域中，像素点  $x_i$  和相邻像素点  $x_1$  的对比度归一化定义为：

$$C(x, x_1) = \frac{|f(x) - f(x_1)|}{\max(f(x), f(x_1))} \quad (6-3)$$

其中， $f(x)$  表示图像在  $x$  像素点的亮度值， $f(x_1)$  表示图像在  $x_1$  像素点的亮度值。

在公式(6-10)的基础上，定义每个像素点的对比度值。取  $x$  的上、下、左、右，左上、右上、左下、右下 8 个邻域，记为  $\Omega(x)$ ，将对比度最大值作为  $x$  点的对比度值，表示如下：

$$C(x) = \max_{x_i \in \Omega(x)} (C(x, x_i)) \quad (6-4)$$

其中， $x_i$  是像素点周边的 8 个相邻像素点。

在大气光学中，定义亮度对比度为：

$$C = \begin{cases} \frac{|B'_0 - B_0|}{B'_0}, & B'_0 \geq B_0 \\ \frac{|B'_0 - B_0|}{B_0}, & B'_0 < B_0 \end{cases} \quad (6-5)$$

式中， $B_0$  为目标物的固有亮度， $B'_0$  为背景的固有亮度。

由于大气消光作用，亮度对比度  $C$  随着观测者与目标物距离的增大，对比度会越来越小，你看到的目标会越来越模糊，当  $C$  小于临界值  $\varepsilon$  后将会无法看清楚目标物。当  $C = \varepsilon$  时，对应的距离就是能见度值， $\varepsilon$  普遍采用 CIE（国际照明协会）在 1987 年发布的标准，规定其为 0.05，即可以令阈值为  $C(x) = 0.05$  提取出图像道路中的可分辨像素点，选取距离摄像机最远的可分辨像素点作为后续分析的目标点。在实际算法实现中，为了确保检测目标的稳定性，选取一行可分辨像素点在 5 个以上且距离摄像机最远的像素行作为后续暗原色先验原理的分析目标，结果如图 6-8 所示。

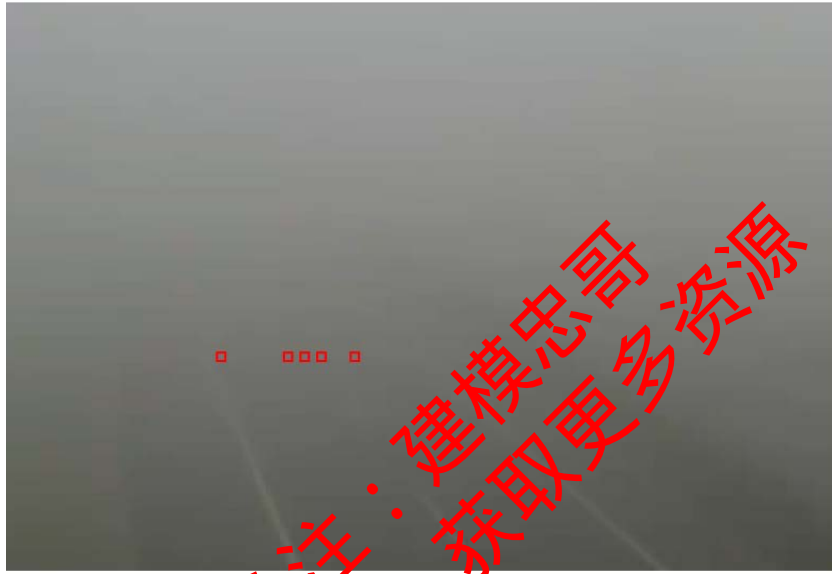


图 6-8 道路最远可视点分布

## 6.8 暗原色法计算消光系数

Koschmieder 定律已被广泛应用于计算机视觉和图形学领域，表示为如下关系：

$$I(x) = t(x)J(x) + A(1 - t(x)) \quad (6-6)$$

$x$  表示图像中的像素点， $I(x)$  为图像原图， $J(x)$  为无雾天气下的清晰图像， $A$  为大气光强度值， $t(x) \in [0, 1]$  表示大气透射率。大气透射率与像素点到摄像机的实际距离  $d(x)$  和大气消光系数  $k$  有关，表示为：

$$t(x) = e^{-kd(x)} \quad (6-7)$$

暗原色先验原理是基于大量户外无雾图像观察到的一条统计规律：在绝大多数户外无雾图像的每个局部区域至少存在某个颜色通道的强度值很低。即对于一幅清晰无雾图像  $J$ ，有：

$$J^{dark}(x) = \min_{c \in \{r, g, b\}} [\min_{y \in \Omega(x)} (J^c(y))] \quad (6-8)$$

式中， $J^c$  表示  $J$  的一个颜色通道， $\Omega(x)$  是以  $x$  为中心的邻域。除了天空区域外， $J^{dark}$  值趋于 0，称  $J^{dark}$  为  $J$  的暗通道。

在暗原色理论的基础上，求取  $x$  邻域内的透射率。首先假设以  $x$  为中心的邻域内透射率相等，然后对(6-9)式两边求取最小值：



$$\frac{\min_{c \in \{r, g, b\}} [\min_{y \in \Omega(x)} (I^c(y))]}{A} = t(x) \frac{\min_{c \in \{r, g, b\}} [\min_{y \in \Omega(x)} (I^c(y))]}{A} + (1 - t(x)) \quad (6-9)$$

根据暗通道趋于 0 的结论，有：

$$\min_{c \in \{r, g, b\}} [\min_{y \in \Omega(x)} (I^c(y))] = 0 \quad (6-10)$$

进而，将式(6-13)代入式(6-12)中求得大气透射率

$$t(x) = 1 - \frac{\min_{c \in \{r, g, b\}} [\min_{y \in \Omega(x)} (I^c(y))]}{A} \quad (6-11)$$

对于大气光强度值  $A$  的估计，采取道路灭点上方图像中大于最大暗原色值 0.9 倍的像素点中  $r, g, b$  三个颜色分量的最大值作为大气光强度的估计值。

结果如图 6-9 所示。

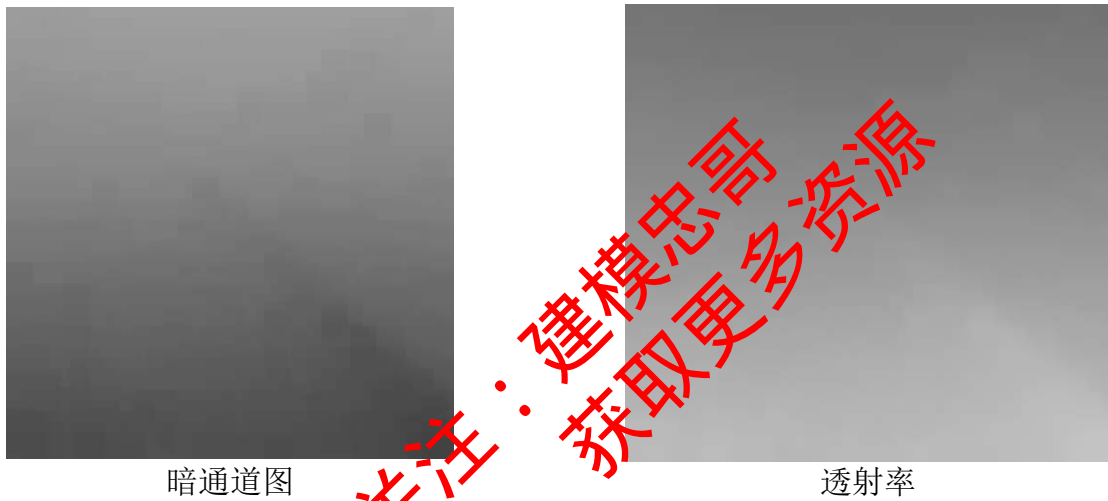


图 6-9 暗通道图像和透射率图像

## 6.9 能见度估计

根据之前理论分析可知，影响能见度的主要因素是物理光学传播状态。光在介质中传播时，受颗粒物吸收和散射的影响会一定的削弱和增强，这种关系表示为 Koschmieder 定律：

$$L = L_0 e^{-kd} + L_f (1 - e^{-kd}) \quad (6-12)$$

式中， $L$  为物体的视亮度， $L_0$  为物体本身亮度， $L_f$  为背景亮度， $d$  为物体本身亮度， $k$  为物体的消光系数。 $L_0 e^{-kd}$  描述了物体亮度的削弱， $L_f (1 - e^{-kd})$  描述了大气背景亮度的增强。由上式推导出亮度对比度计算公式如下：

$$C_d = \frac{|L - L_f|}{L_f} = \frac{|L_0 - L_f|}{L_f} e^{-kd} = C_0 e^{-kd} \quad (6-13)$$

式中， $C_d$  表示物体亮度对比度， $C_0$  表示物体固有亮度对比度。设， $C_d = 0.05$ ， $C_0 = 1$ ，则

由上式可得到大气能见度值  $V_{met}$  的计算公式：

$$V_{met} = -\frac{1}{k} \ln\left(\frac{C_d}{C_0}\right) \approx \frac{3}{k} \quad (6-14)$$

式中  $k$  为消光系数，该值可以有目标点的透射率、目标点距离摄像机的垂直距离分别由式（6-11）和式（6-1）求得。在实际算法中，透射率选取目标行中最远视点中的最小透射率，然后根据式（6-7）求得消光系数，即：

$$k = -\frac{\ln(t(x))}{d(x)} \quad (6-15)$$

最后将消光系数带入式（6-14）即可求得图像的能见度值。结果如图 6-10 所示，其中该时段能见度变化趋势较缓，集中在 47m 附近，波动范围  $\pm 10\text{m}$ 。



图 6-10 2016 年 4 月 14 日部分时段能见度变化曲线

## 7. 问题四建模与求解

### 7.1 问题分析

要求根据题目三求解的能见度变化规律，建立数学模型预测大雾的变化趋势，何时散去。本题实际上是要对题目三求解的能见度曲线进行时序特征分析，剖析曲线变化的变化规律，进而利用数学模型构建大雾变化预测模型。具体流程如下：

首先，对能见度曲线进行数据可视化，然后利用时序分析方法解析该曲线的变化特征，主要包括倾向变化率、平稳性、周期性、偏相关和自相关，进而通过线性回归模型、灰色预测模型以及 ARIMA 时序预测模型描述大雾的变化情况以及何时消散。

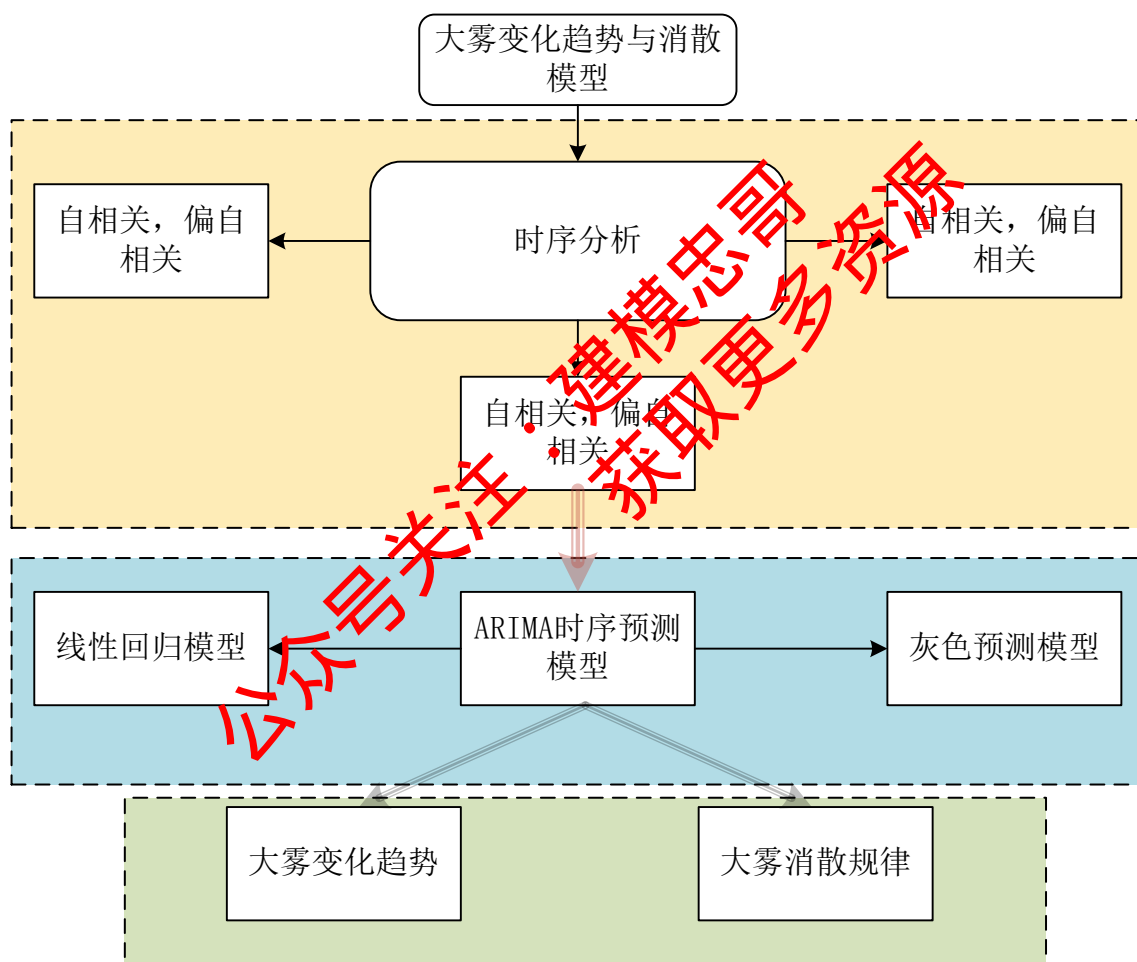


图 7-1 问题四的解决流程图

### 7.2 能见度变化时序特征分析

#### 7.2.1 能见度倾向率

能见度倾向率反应能见度要素的变化趋势，即能见度随时间的变化趋势，一般采用一次线性方程来定量描述，即

$$VIS = b + k * t \quad (7-1)$$

趋势变化率方程：

$$\frac{dVIS}{dt} = k \quad (7-2)$$

式中， $VIS$  为能见度，单位为米， $t$  为时间，单位为秒， $k$  为线性趋势项。

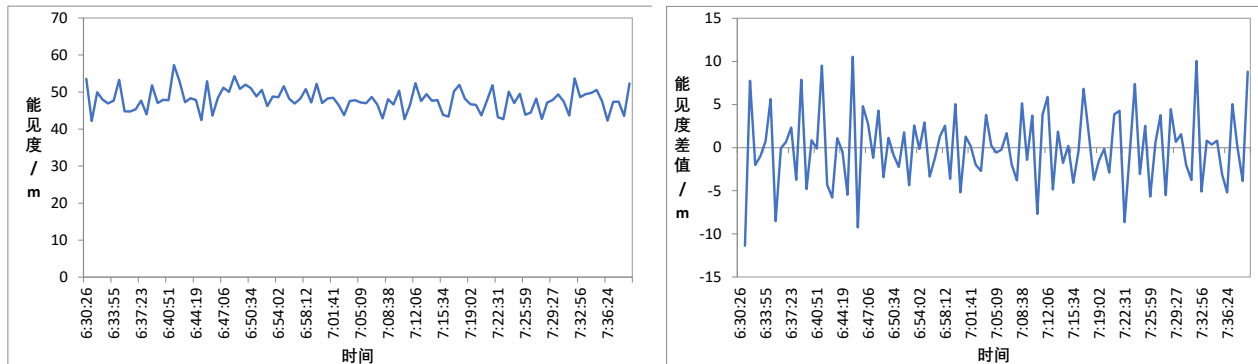


图 7-2 能见度倾向变化率

从图中可以看出，能见度的变化较为平缓，上下波段在 20 之间，并且能见度基本上保持恒定，大约在 47m 左右

### 7.2.2 时间序列的平稳性及周期性检验

时间序列是按照时间排序的一组随机变量，通常是在一定的时间间隔内以一定的采样率对某潜在过程进行观测或采样的结果。时间序列数据本质上反应了随机变量随时间的变化趋势。因此通过对能见度曲线的时间序列研究可以获取能见度随时间变化的规律

#### (1) 时间序列平稳性

某随即变量的时间序列具有平稳性表示该事件序列的本质特征能延续到未来，即该随机变量的历史和现状特征具有代表性和可延续性，因此可以基于历史和现状预测随机变量的未来趋势。

#### (2) 时间序列周期性

某随机变量由于外部影响随自然季节的交替变化出现波峰和波谷的规律。

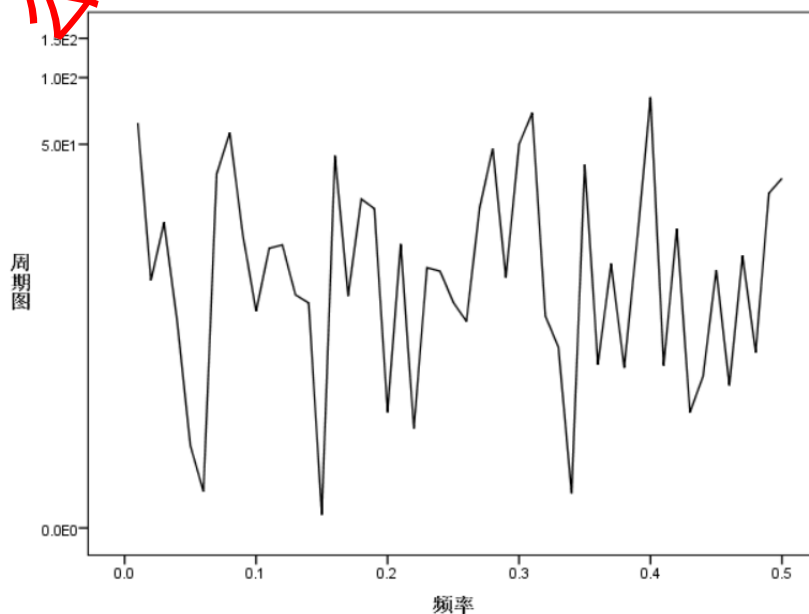


图 7-3 周期图

从图中可以看出，能见度随着时间的变化具有一定的周期，且符合平稳性检验。

### 7.2.3 自相关性与偏相关性

相关函数表现出时间序列中任意两个值之间的相关性是如何随着时间间隔而改变的，自相关函数刻画了时间序列相邻变量之间的相关性，偏相关函数则是排除了其他中间变量的影响，真实地反映了两个变量之间的相关性，并且二者紧密相连，同时两个相关图所反映的信息在时间序列分析各方面发挥着关键作用。

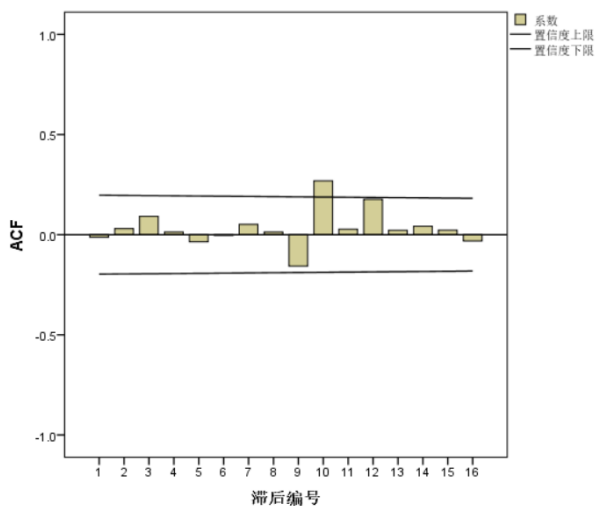


图 7-4 偏相关系数

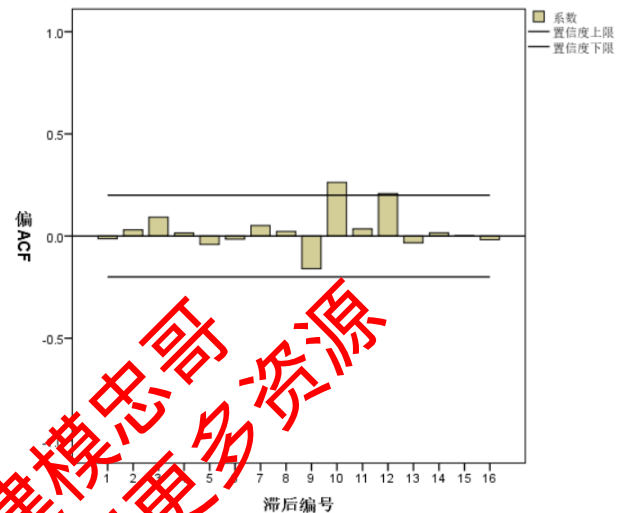


图 7-5 自相关系数

从上图中可以看出，均处在合理的区间范围。

## 7.3 ARIMA 时间序列模型

ARIMA 全程为自回归积分滑动平均模型（Autoregressive Integrated Moving Average Model，简称 ARIMA），ARIMA(p,d,q)模型是针对非平稳时间序列所建立的模型。ARIMA 的含义包含 3 各部分，即 AR、I、MA。

其中 AR 表示 auto regression，即自回归模型；

I 表示 integration，即单整阶数，时间序列模型必须是平稳性序列才能建立计量模型，ARIMA 模型作为时间序列模型也不例外，因此首先要对时间序列进行单位根检验如果是非平稳序列，就要通过差分来转化为平稳序列，就要通过差分来转化为平稳序列，经过几次差分转化为平稳序列，就成为几阶单整；

MA 表示 moving average，即平均移动模型。可见，ARIMA 模型实际上是 AR 模型和 MA 模型的结合。

相应的，有三个参数：p,d,q。其中：

p 代表预测模型中采用的时序数据本身的滞后数（lags），也叫作 AR/Auto-Regressive 项

d 代表时序数据需要进行几阶差分，才是稳定的，也叫 Integrated 项。

q 代表预测模型中采用的预测误差的滞后数（lags），也叫做 MA/Moving Average 项因而，用数学形式表示为（假设 p,q,d 已知）

$$\hat{y}_t = \mu + \Phi_1 * y_{t-1} + \dots + \Phi_p * y_{t-p} + \theta_1 * e_{t-1} + \dots + \theta_q * e_{t-q} \quad (7-3)$$

其中， $\Phi$ 表示AR的系数， $\theta$ 表示MA的系数。

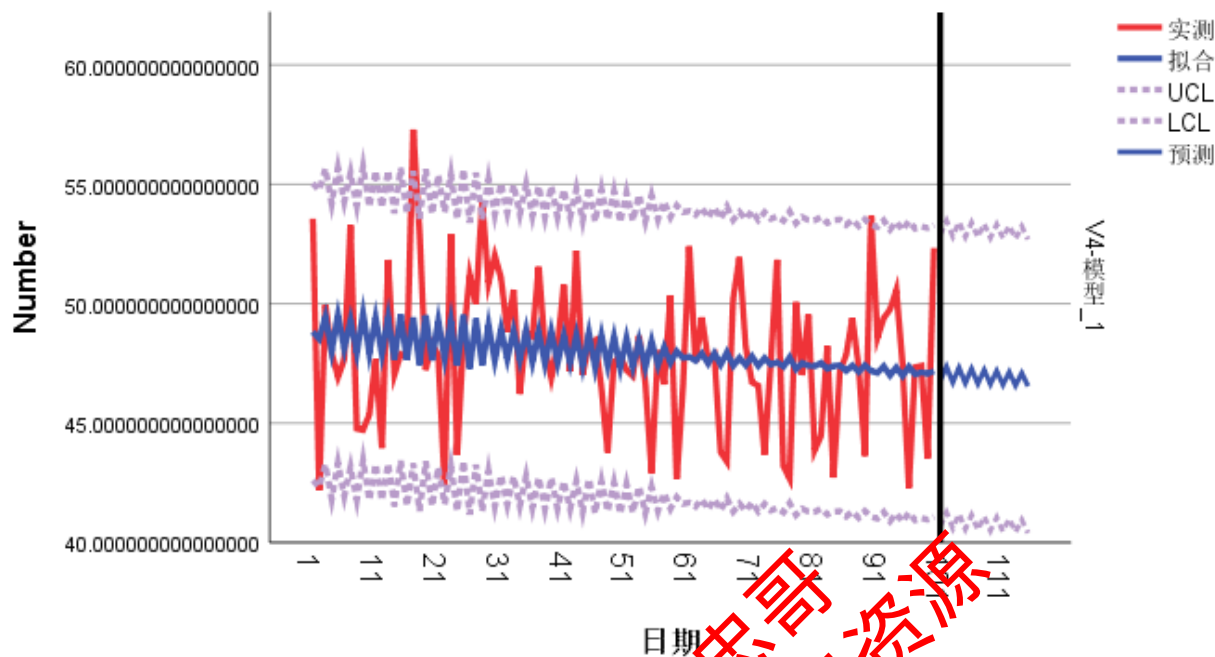


图 7-6 ARIMA 时间序列预测模型

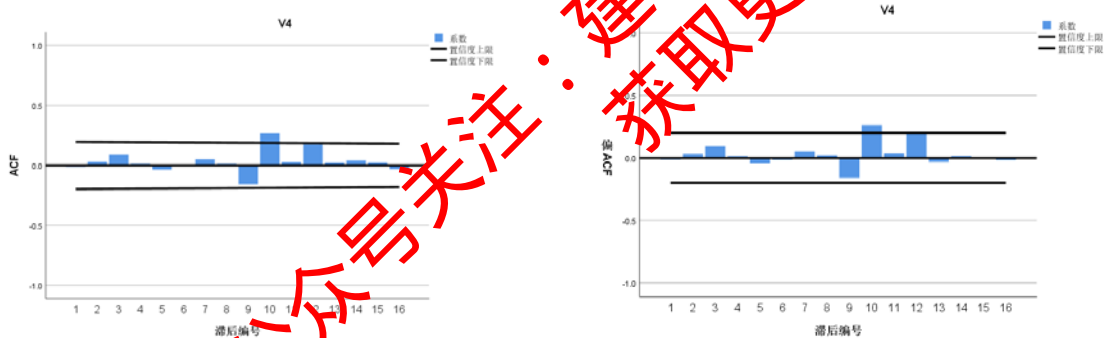


图7-7 ARIMA的自相关和偏自相关检验结果图

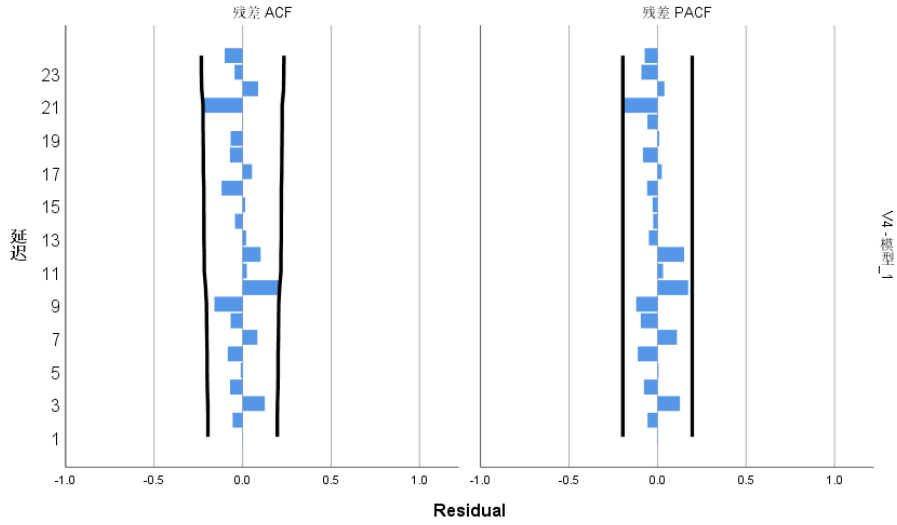


图7-8 ARIMA的预测残差ACF和残差PACF



从图 7-6 中的结果来看, ARIMA 时间序列预测模型估计的大雾变化趋势为加重, 即该时间段内, 能见度的变化趋势为逐渐下降, 但趋势变化较缓。

#### 7.4 GM(1,1)灰色预测模型

灰色系统是指既含有已知信息又含有未知信息的系统, 灰色模型即对灰色系统预测的模型, 表达式  $GM(n, x)$ , 是指用  $n$  阶微分方程对  $x$  个变量建立模型。通过把分散在时间轴上的离散数据看成一组连续变化的序列, 采用累加和累减的方式, 将灰色系统中的未知因素弱化, 强化已知因素的影响程度, 最后构建一个以时间为变量的连续微分方程, 通过数学方法确定方程中的参数, 从而实现预测目的。他无需大量数据样本, 短期预测效果好, 运算过程简单。但对非线性数据样本预测效果差。

(1) 设一组原始数据为  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ , 其中,  $n$  为数据个数对  $x^{(0)}$  累加以便弱化随机序列的波动性和随机性, 得到新的数列为:

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)), \text{其中, } x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), \quad k=1, 2, \dots, n$$

(2) 生成  $x^{(1)}$  的邻均值等权数列  $z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$ ,  $k=2, 3, \dots, n$  其中,  $z^{(1)}(k) = 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k)$

(3) 根据灰色理论对  $x^{(1)}$  建立关于  $t$  的白化形式的一阶一元微分方程 GM(1,1):

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u$$

其中,  $a, u$  为待解系数, 分别称为发展系数和灰色作用量,  $a$  的有效区间是  $(-2, 2)$ , 并记  $a, u$

构成的矩阵为灰参数  $\hat{a} = \begin{pmatrix} a \\ u \end{pmatrix}$ , 只要求出参数  $a, u$ , 就能求出  $x^{(1)}(t)$ , 进而求出  $x^{(0)}$  的预测值。

(4) 对累加生成数据做均值生成  $B$  与常数项向量  $Y_n$ :

$$B = \begin{bmatrix} -\frac{1}{2}Z^{(1)}(2) & 1 \\ -\frac{1}{2}Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}Z^{(1)}(n) & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

(5) 用最小二乘法求解灰参数, 则  $\hat{a} = (B^T B)^{-1} B^T Y_n$

(6) 将灰参数  $\hat{a}$  代入  $\frac{dx^{(1)}}{dt} + ax^{(1)} = u$ , 并对其进行求解, 得

$$\hat{x}^{(1)}(t+1) = (\hat{x}^{(1)}(1) - \frac{u}{a})e^{-a} + \frac{u}{a}$$

(7) 将上述结果累减还原, 即可得到预测值

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t)$$

(8) 利用模型进行预测:  $\hat{x}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n), \hat{x}^{(0)}(n+1), \dots, \hat{x}^{(0)}(n+m))$

灰色预测结果如下图 7-9 所示。

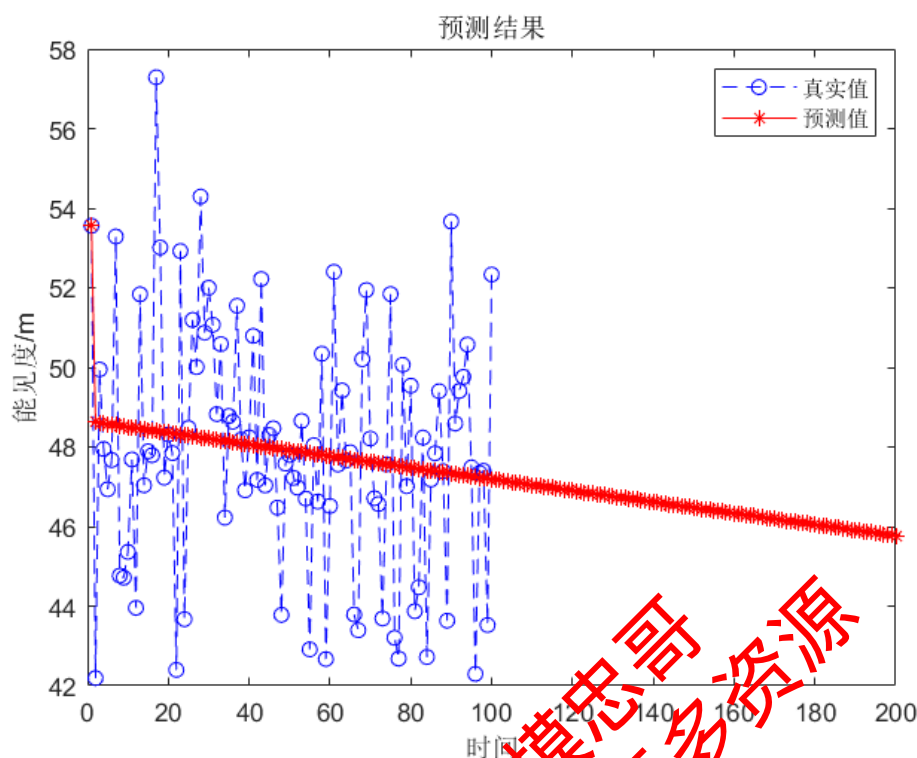


图 7-9 灰色预测结果

从图 7-9 中可以看出，灰色预测的预测值呈下降趋势，表明在 0-100 个时间间隔（42 秒）中，大雾的变化趋势时逐渐加重的，但变化速度较缓。此时大雾短时间内不会消散。

## 7.5 线性回归模型

对能见度曲线进行一元线性回归，从图中可以看出随着时间的增加，大雾逐渐减弱，当 MOR=150 时，时间为 8 点 50 分 56 秒，即为大雾消散时间。

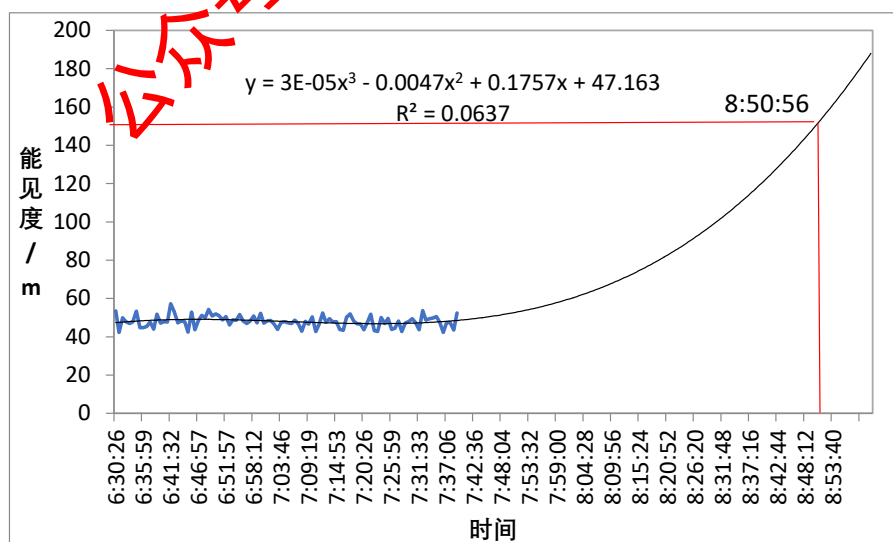


图 7-10 线性回顾预测大雾消散时间

## 参考文献

- [1] D. WU, X. TIE, C. LI, et al. An Extremely Low Visibility Event over the Guangzhou Region: A Case Study. *Atmos Environ.* 2005, 39(35): 6568-6577
- [2] Y. QI, A. MITCHELL MYSLIWIEC, MJ. KLEEMAN. Source Apportionment of Visibility Impairment Using a Three-Dimensional Source-Oriented Air Quality Model. *Environmental Science & Technology*, 2004, 38(4): 1089-1101
- [3] 龙凤翔,张瑀琳.桂林城区大气能见度与颗粒物浓度和气象因子关系研究[J].*气象与环境学报*,2020,36(01):21-27.
- [4] 郭智亮,高聪晖,谢文锋,陈淑敏,黎伟标.2005—2017 年白云机场能见度变化特征及其与影响因子关系研究[J].*气象科技进展*,2019,9(06):40-43.
- [5] LE. YU, ML. SHULMAN, R. KOPPERUD, et al. Characterization of Organic Compounds Collected During Southeastern Aerosol and Visibility Study: Water-Soluble Organic Species. *Environmental Science and Technology*, 2005, 39(3): 707-715
- [6] 韩明敏. 基于视频图像的高速公路能见度检测技术研究[北京交通大学硕士学位论文]. 北京: 北京交通大学, 2016
- [7] 龙凤翔,张瑀琳.桂林城区大气能见度与颗粒物浓度和气象因子关系研究[J].*气象与环境学报*,2020,36(01):21-27
- [8] 石玉立,王彬,卜帆.基于图像特征的大气能见度估算方法[J].*南京理工大学学报*,2018,42(05):552-559.
- [9] 邱新法,叶栋水,曾燕,叶秀枝,石一凡.基于视频的雾天能见度实时监测方法研究[J].*安全与环境学报*,2018,18(04):1413-1418.
- [10] 花毓幸. 基于图像兴趣窗格测算大气能见度的方法研究[D].南京信息工程大学,2016.
- [11] 项文书. 基于交通视频的能见度估计研究[D].上海交通大学,2014.
- [12] Baumer D, Versick S, Vogel B. Determination of the visibility using a digital panorama camera. *Atmospheric Environment*, 2008, 42(11): 2593-2602.
- [13] Hautiere N, Bigorne E, Aubert D. Daytime visibility range monitoring through use of a roadside camera. *Intelligent Vehicles Symposium*, 2008 IEEE. IEEE, 2008: 470-475.

## 附录

问题一

%% 数据与处理

clear;clc;

[datanum,txt] = xlsread("dataAMOS20191216");

datatxt = txt(3:end,2);

newtxt = {};

newNum = [];

for i = 1:size(datatxt,1)

time = datatxt(i);

timechar = cell2mat(time);

timeMilisecond = timechar(length(timechar)-1:length(timechar));% 1 秒间隔

standard = '00';

if timeMilisecond == standard % 1 秒间隔

newtxt(size(newtxt,1)+1,1) = datatxt(i);

newNum(size(newNum,1)+1,:) = datanum(i,:);

end

end

%% 处理 15 秒一个间隔

clear;clc;

[datanum,txt] = xlsread("dataAMOS20191216",'处理后 1 分钟间隔');

datatxt = txt(3:end,2);

newtxt = {};

newNum = [];

for i = 1:size(datatxt,1)

time = datatxt(i);

timechar = cell2mat(time);

timeMilisecond = timechar(length(timechar)-4:length(timechar)-3);

if mod(str2num(timeMilisecond),10) == 0 % 15 秒或 30 一个间隔

newtxt(size(newtxt,1)+1,1) = datatxt(i);

newNum(size(newNum,1)+1,:) = datanum(i,:);

end

end

%遗传算法

clear;clc;

% 目标函数  $y = a_0 + a_1 \cdot x_1^2 + a_2 \cdot x_1 + a_3 \cdot x_2^3 + a_4 \cdot x_2^2 + a_5 \cdot x_2 + a_6 \cdot x_3^4 + a_7 \cdot x_3^3 + a_8 \cdot x_3^2 + a_9 \cdot x_3 + a_{10} \cdot x_4^2 + a_{11} \cdot x_4$

%两边求对数得到化为线性:  $\ln y = \ln a_0 + \dots + \ln a_{11} + 3 \ln x_1 + 6 \ln x_2 + 10 \ln x_3 + 3 \ln x_4$

%在令 new = lny; new1 = lnx1; new2 = lnx2; new3 = lnx3; new4 = lnx4;

```

%得到方程 new = b0 + 3*new1 + 6*new2 + 10*new3 + 3*new4;
global y;
global x;
adress = {'dataAMOS20200313.xlsx','dataAMOS20191216.xlsx'};
for j = 1:2
data = xlsread(cell2mat(adress(j)),"处理后 15 分钟间隔");
y = data(:,4); %能见度
x1 = data(:,6); %风速
x2 = data(:,9); %本站气压
x3 = data(:,12); %温度
x4 = data(:,13); %相对湿度
x = [x1 x2 x3 x4]';

%定义待调用的两个子程序 seaircal()和 seair_constraint()
ObjectiveFunction = @seaircal;
nvars = 12;
%循环 ga()遗传算法
for i = 1:200
[d,fval] = ga(ObjectiveFunction,nvars,[],[],[],[],[],[],[],[],[],[]); %x 是待估变量，fval 是返回
的自适应值
result(i,:) = [d,-fval];
end
%将计算结果导出为*.xls 格式的Excel表格文档
result1 = sortrows(result,13,'descend');
xlswrite(cell2mat(adress(j)),result1,'遗传算法估计参数');
end
%% 程序功能：为遗传算法提供自适应函数（SEAIHDR 模型），返回 simerr 作为评价指标，
即每次进化会选出最小的 simerr 值
function simerr = seaircal(a)
global x;
global y;
for j = 1:size(x,2)
yichuan(j) = a(1) + a(2)*x(1,j)^2 + a(3)*x(1,j) + a(4)*x(2,j)^3 + a(5)*x(2,j)^2 + a(6)*x(2,j)
+ a(7)*x(3,j)^4 + a(8)*x(3,j)^3 + a(9)*x(3,j)^2 + a(10)*x(3,j) + a(11)*x(4,j)^2 + a(12)*x(4,j);
end
%% 计算 R2 值
%计算误差评定指数：模型参数条件下 SARS 传播过程与真实过程的拟合度值
simerr = -FITNESS(y',yichuan'); %因为 ga()函数默认输出

问题二
% 调试
% Input_path = 'F:\永无极限小组\E20104590100\问题二\图片\裁剪后\';
% namelist = dir(strcat(Input_path,'*.jpg'));
% name=namelist(4).name; %namelist(i).name; %这里获得的只是该路径下的文件名

```

```

% I=imread(strcat(Input_path, name)); % 图片完整的路径名
% [meansquare,meanG,tidu,contrast, regionI, regionG, const] = featureExtrude(I,8,8);
% function [meansquare,meanG,tidu,contrast, regionI, regionG, const] = featureExtrude(image,
ch, cw)
function feature = featureExtrude(image, ch, cw) %%ch 和 cw 是子窗体的尺寸
% 获取图像的 rgb 通道矩阵，并将其转化为 double 可计算类型
    R=image(:,:,1);
    G=image(:,:,2);
    B=image(:,:,3);
    R0=double(R);
    G0=double(G);
    B0=double(B);
    [rs,cs] = size(R0);% 行数列数

    %将 rgb 颜色空间转换到 HSI 颜色空间.I 为亮度特征值，为了计算均方差
    I = 1/3*(R0 + G0 + B0);
    %将 rgb 颜色空间转换到灰度图像，为了计算梯度
    G = double(rgb2gray(image)); %按照 0.3 0.69 0.11 (rgb)

% *****分块，即按照一定窗口大小计算各个特征*****
    % ch 为行间隔 cw 为列间隔，控制分块图像的大小，如现在的 8*8
    % numr 为间隔块个数 numc 为间隔块个数
    numr = round(rs/ch); % 计算总共多少行
    numc = round(cs/cw); % 计算总共多少列
    allregion = numr*numc; % 总共多少区域
    % 区域块分割
    t1 = (0:numr-1)*ch + 1; t2 = (1:numr)*ch;
    t3 = (0:numc-1)*cw + 1; t4 = (1:numc)*cw;
    count = 1; % 用来记录一幅影像生成了多少个窗口，最终结果应该等于 allregion+1
    % 获得一幅影像中的各窗口矩阵，即分块矩阵
    for i=1:numr
        for k=1:numc
            % 从 t1 行至 t2 的子块
            x=t1(i):t2(i);
            y=t3(k):t4(k);
            regionI(:, :, count) = I(x(1):x(length(x)),y(1):y(length(y)));
            regionG(:, :, count) = G(x(1):x(length(x)),y(1):y(length(y)));
            count = count + 1;
        end
        rectangle('Position', [t3(j) t1(i) length(x) length(y)], ...,
        'EdgeColor', 'r', 'LineWidth', 1); % 绘制矩形分割格子，红色
    end
end

% *****分块，即按照一定窗口大小计算各个特征*****
% 计算亮度均方差(是一个窗格的数值)

```



```

for j = 1: allregion
    f = regionI(:,j);
    u = mean(mean(f));
    s = 0; %每一个格子的方差
    for i=1:cw
        for k=1:ch
            s = s+(f(i,k)-u)^2;
        end
    end
    meansquare(j) = sqrt(s/(cw*ch)); %一幅影像一个子窗口的一个均方差特征
end

```

%计算灰度均值

```

for j = 1: allregion
    f = regionG(:,j);
    meanG(j) = mean(mean(f)); %一幅影像一个子窗口的一个均值特征
end

```

%计算平均梯度（每个像素的值是梯度，对于窗口还要求窗口内的所有像素均值,即我们做的这个  $\text{tidu\_j}/((\text{cw}-1)*(\text{ch}-1))$ ）

```

for j = 1: allregion
    f = regionG(:,j);
    tidu_j = 0; %每一个格子的平均梯度
    for i=1:cw-1
        for k=1:ch-1
            tidu_j=(tidu_j)+sqrt( ((f(i,k)-f(i+1,k))^2+(f(i,k)-f(i,k+1))^2)/2 );
        end
    end
    tidu(j) = tidu_j/((cw-1)*(ch-1)); %一幅影像一个子窗口的一个均方差特征
end

```

%计算灰度对比度（每个像素的值是对比度，对于窗口还要求窗口内的所有像素均值）

```

for j = 1: allregion
    f = regionG(:,j);
    gray_j = []; %每一个格子的对比度
    d = 1;
    for i=2:cw-1
        for k=2:ch-1
            gray_j(1) = abs(f(i,k)-f(i-1,k-1))/max(f(i,k), f(i-1,k-1));
            gray_j(2) = abs(f(i,k)-f(i-1,k))/max(f(i,k), f(i-1,k));
            gray_j(3) = abs(f(i,k)-f(i-1,k+1))/max(f(i,k), f(i-1,k+1));
            gray_j(4) = abs(f(i,k)-f(i,k-1))/max(f(i,k), f(i,k-1));
            gray_j(5) = abs(f(i,k)-f(i,k+1))/max(f(i,k), f(i,k+1));
            gray_j(6) = abs(f(i,k)-f(i+1,k-1))/max(f(i,k), f(i+1,k-1));
        end
    end
end

```

```

        gray_j(7) = abs(f(i,k)-f(i+1,k))/max(f(i,k), f(i+1,k));
        gray_j(8) = abs(f(i,k)-f(i+1,k+1))/max(f(i,k), f(i+1,k+1));
        const(d) = max(gray_j);
        d= d+1;
    end
end
contrast(j) = mean(const); %一幅影像一个子窗口的一个均方差特征
end
feature = [meansquare;meanG;tidu;contrast];

end

% I:原始图像
clear;clc;close all
Input_path = 'F:\永无极限小组\E20104590100\问题二\图片\裁剪后\';
namelist = dir(strcat(Input_path,*.jpg)); %寻找该文件夹中所有后缀为.jpg 的文件并且存入
其相关信息
files_name =sort_nat({namelist.name}); %cell2mat(files_name(1))可以变成字符串
data = xlsread('VIS_R06_12.xlsx','处理后数据');
len = length(namelist);

%%      ***** 将 能 见 度 人 为 分 级
***** %
% data50 = [];
% data200 = [];
% data500 = [];
% data1000 = [];
% data1600 = [];
% for i = 1: size(data,1)
%     if data(i,1)<=50
%         data50(1, size(data50,2)+1) = data(i,1); %能见度值
%         data50(2, size(data50,2)) = data(i,2); %对应序号也对应图片序号
%     elseif data(i,1)<=200
%         data200(1, size(data200,2)+1) = data(i,1);
%         data200(2, size(data200,2)) = data(i,2);
%     elseif data(i,1)<=500
%         data500(1, size(data500,2)+1) = data(i,1);
%         data500(2, size(data500,2)) = data(i,2);
%     elseif data(i,1)<=1000
%         data1000(1, size(data1000,2)+1) = data(i,1);
%         data1000(2, size(data1000,2)) = data(i,2);
%     else
%         data1600(1, size(data1600,2)+1) = data(i,1);
%         data1600(2, size(data1600,2)) = data(i,2);

```

```

%      end
% end
%% ***** 将 能 见 度 人 为 分 级
***** %%

%% ***** 统计影像特征因子和对应能见度的相关性
***** %%

% feacount = 4;
% feature = [];
% for k=1:len
%     name=namelist(k).name; %namelist(i).name; %这里获得的只是该路径下的文件名
%     I=imread(strcat(Input_path, name)); %图片完整的路径名
%     d = featureExtrude(I, 10, 10); %均方差, 均值, 梯度, 对比度
%     feature(size(feature,1)+1,1:feacount) = mean(d,2); %求所有列的均值,即特征值的均值
%     for j = 1:size(data,1) %遍历所有样本寻找与该图像特征对应的能见度
%         if str2double(name(1:length(name)-4)) == data(j,2)
%             feature(size(feature,1),feacount+1) = data(j,1);
%         end
%     end
% end
%% 求维度之间的相关系数
% rho = corr(feature, 'type','Spearman'); % Kendall Pearson Spearman

%% 绘制热图
% string_name={'Mean_square','Mean','Average_gradient','Contrast_ratio','MOR'};
% xvalues = string_name;
% yvalues = string_name;
% h = heatmap(xvalues,yvalues,rho, 'FontSize',10, 'FontName','Times New Roman');
% colormap(parula)
% saveas(gcf,sprintf('各特征因子与能见度的热力图(肯德尔).jpg'),'bmp'); %保存图片
%% ***** 统计影像特征因子和对应能见度的相关性
***** %%

%% ***** 统计窗口大小和对应能见度的相关性
***** %%

% feacount = 4;
% feature = [];
% relation = [];
% grid_size = [3 5 8 10 12 20 25 30]; %尺寸的选择注意, 不能使 7 15 等, 会导致分块的时候
% y=[1000 1001],主要问题是因为取整问题
% for i = 3 : size(grid_size,2)
%     for k=1:len
%         name=namelist(k).name; %namelist(i).name; %这里获得的只是该路径下的文

```

```

件名
% I=imread(strcat(Input_path, name)); % 图片完整的路径名
% d = featureExtrude(I, grid_size(i), grid_size(i)); % 均方差, 均值, 梯度, 对比度
% feature(size(feature,1)+1,1:feacount) = mean(d,2); % 求所有列的均值, 即特征值
的均值
% for j = 1:size(data,1) % 遍历所有样本寻找与该图像特征对应的能见度
%     if str2double(name(1:length(name)-4)) == data(j,2)
%         feature(size(feature,1), feacount+1) = data(j,1);
%     end
% end
% end
% % 求维度之间的相关系数
% rho = corr(feature, 'type', 'Kendall'); % Kendall Pearson Spearman
% relation(i,:) = rho(feacount+1,:);
% end
%% ***** 统计窗口大小和特征能见度之间的相关性
***** %%

%% ***** 统计每个窗口的最优兴趣窗口
***** %%

% feacount = 4;
% feature_meansquare = [];
% feature_tidu = [];
% feature_const = [];
% for k=1:len
%     name=cell2mat(files_name(k)); % namelist(i).name; % 这里获得的只是该路径下的文
件名
% % 遍历所有样本寻找与该图像特征对应的能见度
% for j = 1:size(data,1)
%     if str2double(name(1:length(name)-4)) == data(j,2)
%         feature_meansquare(k,1) = data(j,1);
%         feature_tidu(k,1) = data(j,1);
%         feature_const(k,1) = data(j,1);
%     end
% end
% I=imread(strcat(Input_path, name)); % 图片完整的路径名
% d = featureExtrude(I, 8, 8); % 均方差, 均值, 梯度, 对比度, 注意选择的最有窗口大
大小为 8*8
% feature_meansquare(k,2:length(d)+1) = d(1,:); % 均方差特征, 行是第几张图片, 列是
各个窗口
% feature_tidu(k,2:length(d)+1) = d(3,:); % 平均梯度特征, 行是第几张图片, 列是各个
窗口
% feature_const(k,2:length(d)+1) = d(4,:); % 平均梯度特征, 行是第几张图片, 列是各个
窗口

```

```

% end
%
% % 求维度之间的相关系数
% rho_const = corr(feature_const, 'type','Spearman'); % Kendall Pearson Spearman
% relation_const = rho_const(1,2:end); % 第一行即为能见度与每个网格之间的相关性（共
6250 个网格），同时去掉第一行第一列的 1
% [grid_relation_const,index] = sort(relation_const,'descend');
% zuiyou_grid_const = [index(1:10); grid_relation_const(1:10)]; %选取前 10 个作为最优窗口
%
% rho_tidu = corr(feature_tidu, 'type','Spearman'); % Kendall Pearson Spearman
% relation_tidu = rho_tidu(1,2:end); % 第一行即为能见度与每个网格之间的相关性（共 6250
个网格），同时去掉第一行第一列的 1
% [grid_relation_tidu,index] = sort(relation_tidu,'descend');
% zuiyou_grid_tidu = [index(1:10); grid_relation_tidu(1:10)]; %选取前 10 个作为最优窗口
%
% rho_meansquare = corr(feature_meansquare, 'type','Spearman'); % Kendall Pearson Spearman
% relation_meansquare = rho_meansquare(1,2:end); % 第一行即为能见度与每个网格之间的
相关性（共 6250 个网格），同时去掉第一行第一列的 1
% [grid_relation_meansquare,index] = sort(relation_meansquare,'descend');
% zuiyou_grid_meansquare = [index(1:10); grid_relation_meansquare(1:10)]; %选取前 10 个作
为最优窗口
% save('feature_grid','feature_const','feature_meansquare','feature_tidu')
% 最终的结果就是 zuiyou_grid_meansquare zuiyou_grid_tidu zuiyou_grid_const 三个变量
% 文件中的 feature_grid 就是每幅影像对应的一幅影像里每个窗口的一个因子值
%% ***** 统计 最 优 兴 趣 窗 口
***** %%

%% ***** 三元线性回归建模
***** %%

% load('feature_grid.mat'); %加载各窗口的各特征值与能见度的矩阵
% zuiyou_grid = [304,3177,3427,3302,3552,429,5302,4802,5052,5802;
%               304,5301,429,179,5803,54,5426,6053,4801,5303;
%               304,3177,3302,3427,3552,429,5802,4801,5803,5052];%顺序是梯度、均方差、对
比度对应最优窗口
% %将各个特征值对应最优窗口的属性值提取出来，并求取均值
% %提取最优窗口对应的梯度
% zuiyou_const = [];
% zuiyou_tidu = [];
% zuiyou_meansquare = [];
% for i = 1 : length(feature_const) %表示格网的编号，但是多了一个第一列是能见度，对
应的窗格编号应该是 i-1
%     for j = 1 : 10
%         if zuiyou_grid(1,j) == i-1
%             zuiyou_tidu(:,size(zuiyou_tidu,2)+1) = feature_tidu(:,i);

```

```

%           end
%           if zuiyou_grid(2,j) == i-1
%               zuiyou_meansquare(:,size(zuiyou_meansquare,2)+1) =
feature_meansquare(:,i);
%           end
%           if zuiyou_grid(3,j) == i-1
%               zuiyou_const(:,size(zuiyou_const,2)+1) = feature_const(:,i);
%           end
%       end
%   end
%
%   %确定最终需要进行线性回归的  $y = a*x1 + b*x2 + c*x3$ ,  $x1$  即为上面选取 10 个最右窗
口的值的均值,梯度, 均方差, 对比度
%   x1 = mean(zuiyou_tidu, 2); %对每一行所有列进行求和
%   x2 = mean(zuiyou_meansquare, 2);
%   x3 = mean(zuiyou_const, 2);
%   x = [x1 x2 x3];
%   y = feature_tidu(:,1); %三个特征第一列都是能见度店面才是各窗口对应的值, 所以取
一个即可
%   [b,bint,r,rint,stats] = regress(y,x); %b 是各个系数,bint 是系数的 95%的置信区间, r 返回
的是残差,rint 是异常值矩阵, stats 是 r2、F 检验、p 值以及误差方法的估计值
%
%   %计算残差 et = y - yg
%   yg = b(1)*x1 + b(2)*x2 + b(3)*x3;
%   e = y - yg;
%   save('liner_regress','y','x','yg','e')
%% ***** 三元线性回归建模 ***** %%
%
%% *****神经网络预测***** %%
% load('liner_regress.mat');
% p = x'; %输入矩阵
% t = e'; %目标矩阵
%
% %利用函数 premnmx 对数据进行归一化
% [pn, minp, maxp, tn, mint, maxt] = premnmx(p, t); %对于输入矩阵 p 和输出矩阵 t 进行归一
化处理
% dx = [-1, 1; -1, 1; -1, 1]; %归一化处理后最小值为-1, 最大值为 1
%
% %BP 神经网络训练
% %建立模型, 并用梯度下降算法训练,'tansig'函数为隐层的传输函数, purelin 函数为输出
层的传输函数
% net = newff(dx, [3, 15, 1], {'tansig','tansig','purelin'}, 'traingdx');
% net.trainParam.show = 1000; %1000 轮回显示一次结果

```



```

% net.trainParam.Lr = 0.01; %学习速率为 0.05
% net.trainParam.epochs = 10000; %最大训练轮回为 5000 次
% net.trainParam.goal = 0.01; %均方根误差
%
% net = train(net,pn, tn); %开始训练，其中 pn，tn 为归一化后的输入和输出样本
%
% %利用原始数据对 BP 网络仿真
% an = sim(net, pn); %用训练好的模型进行仿真
% a = postmnmx(an, mint, maxt); %把仿真的得到的数据还原为原始数据量级
%
% %本利因样本容量有限使用训练数据进行测试，通常必须用新鲜数据进行测试
% xx = 1:434;
% newt = a;
%
% figure(1);
% plot(xx, newt, 'r-o', xx, t, 'b--*'); %绘制公路客运量对比图
% legend('估计残差修正值','实际残差修正值');
% xlabel('图片序号');ylabel('残差/m');

%% ***** 神经网络预测 *****
%% ***** 神经网络预测 *****

%% ***** 神经网络预测 *****
%% ***** 神经网络预测 *****

% 最终绘图
% data = xlsread('VIS_R06_12.xlsx','线性拟合');
% y1 = data(:,1)'; %真实能见度
% y2 = data(:,5)'; %线性回归估计能见度值
% y3 = data(:,8)'; %修正估计能见度值
% y4 = data(:,6)'; %残差
% x = 1:434;
% newt = y3; %估计值
% t = y1; %真实值
%
% figure(1);
% plot(x, newt, 'r-o', x, t, 'b--*'); %绘制公路客运量对比图
% legend('修正模型估计能见度值','实际能见度值');
% xlabel('图片序号');ylabel('能见度/m');
%
% % saveas(gcf,sprintf('1.jpg'),'bmp'); %保存图片

% %计算 r2 值
% a = y1';
% b = y3';

```

```

% sizeA=size(a);
% sizeB=size(b);
% top = 0;
% for i=1:1:sizeA(1)
%     top = top + (a(i,1)-b(i,1))^2;
% end
% bot=0;
% ma=mean(a);
% for i=1:1:sizeA(1)
%     bot = bot + (a(i,1)-ma)^2;
% end
% sr = 1 - top/bot

```

### 问题三

%单幅图像的暗原色先验去雾改进算法，能够很好地改进天空或明亮部分色彩失真问题，matlab 代码注解详细，能很好地帮助理解算法过程。

```
clc;clear;
```

```

Input_path='cutting2\';
outpath ='toushe\';
picture_name_list= dir(strcat(Input_path,'*.bmp'));
len = length(picture_name_list);
toushe=[];
xy=[];
for mi = 1:len
img_name =picture_name_list(mi).name;
picture=imread(strcat(Input_path,img_name));
% 原始图像
G=liangdu(picture);
originImage=double(picture)/255; %读进来是 3 维矩阵 rgb，800*1200 像素
% 获取图像大小
[h,w,c]=size(originImage);%h 是高度，w 是宽度，c 是几个通道
win_size = 7; %这是设置邻域窗口，15*15，因第一个选定像素不能是第一个，应该是从
左到右的第 8 个，最后一个像素为从右向左第 8 个
img_size=w*h;
% figure, imshow(originImage);%显示原始图片
% title("原始图像");

win_dark=ones(h,w);
%计算分块 darkchannel
for j=1+win_size:w-win_size %保证不会选到边缘点，即第一个像素从左到右的第 8 个等
    for i=win_size+1:h-win_size
        m_pos_min = min(originImage(i,j,:)); %先选出该像素中这个三个通道中的最小值
        for n=j-win_size:j+win_size %选取该邻域内最小的像素值

```

```

        for m=i-win_size:i+win_size
            if(win_dark(m,n)>m_pos_min)
                win_dark(m,n)=m_pos_min;
            end
        end
    end
end
end
%选定精确 dark value 坐标
%win_b = zeros(img_size,1);
%win_dark 是暗通道
% figure, imshow(win_dark); %显示暗通道
% title("暗通道");
if mi==1
    imwrite(win_dark,'图\暗通道.bmp');
end
%需调节参数
w = 0.9;
win_t=1-w*win_dark; %这个地方的 w 就是权重参量，计算的结果就是透射率
% figure, imshow(win_t);%显示原始透射率图
% title("原始透射率");

% *****
% 导向滤波细化透射率
% - 引导图像: I (应为灰度/单通道图像)
% - 滤波输入图像: p (应为灰度/单通道图像)
% - 局部窗口半径: r
% - 正则化参数: eps
% *****

I=win_t;
p=win_t;
%需调节参数
r=4;
eps=0.001;

[hei, wid] = size(I);
N = boxfilter(ones(hei, wid), r);

mean_I = boxfilter(I, r) ./ N;
mean_p = boxfilter(p, r) ./ N;
mean_Ip = boxfilter(I.*p, r) ./ N;
% this is the covariance of (I, p) in each local patch.
cov_Ip = mean_Ip - mean_I .* mean_p;

```

```

mean_II = boxfilter(I.*I, r) ./ N;
var_I = mean_II - mean_I.* mean_I;

a = cov_Ip ./ (var_I + eps);
b = mean_p - a.* mean_I;

mean_a = boxfilter(a, r) ./ N;
mean_b = boxfilter(b, r) ./ N;

q = mean_a.* I + mean_b;
if mi==1
    imwrite(q,'图\透射率.bmp');
end
% figure, imshow(q);% 导向滤波细化后的透射率
% title("导向滤波细化后的透射率");
liangx=G(1,1);
liangy=G(2,1);
xy(size(xy,1)+1,:)=[liangx,liangy];
d=canshu(liangx);
toushe(size(toushe,1)+1,:)=[q(liangx,liangy),dl;%234,203,268,229,q(252,527),q(294,597)
end
% *****
%     自动获取大气光步骤，A 为最终大气光的值
% *****
% range=ceil(img_size*0.001);%取暗原色中最亮的%1 的点数，ceil 朝无穷大方向取整
% radi_pro=zeros(range,1); %用于记录最亮点内对应图片点象素的三个通道的颜色强度
%     for s=1:range
%         [a,b]=max(win_dark);
%         [c,d]=max(a);
%         b=b(d);
%         m=sparse(b,d,1,h,w);          %b,d 为最亮值的坐标
%         win_dark=win_dark-c.*m;        %消去选出的最大值
%         radi_pro(s)=sum(originImage(b,d,:)); %最大值对应象素三通道求和
%     end
% A=max(radi_pro)/3;%大气光的值，是用来求去雾图像的

% *****
%     算法改进步骤，可修正天空透射率以减小明亮部分的失真率
% *****
% inten=zeros(h,w);
%     for m=1:h
%         for n=1:w
%             inten(m,n)=mean(originImage(m,n,:));
%         end

```

```

%      end
% k=70;
% k=zeros(h,w)+k/255; % 容差
% % A=220/255;
% cha=abs(inten-A); % 差限
% alpha=min(max(k./cha,1).*max(alpha,0.1),1); % 算法改进关键部分
% figure,imshow(alpha);
% % *****
% alpha=repmat(alpha,[1,1,3]);
% dehaze=(originImage-A)./alpha+A;
% figure, imshow(dehaze);
% title("dehaze");
see=[];
for i=1:size(toushe,1)
    bf=toushe(i,1);
    bc=toushe(i,2);
    see(size(see,1)+1,:)=3/(-log(toushe(i,1))/toushe(i,2));
end

clear;
f=imread('非线性图像增强\1.bmp');%rgb2gray%cutting\1.bmp 均衡化 非线性图像增强
path1=('边缘检测\非线性的增强\梯度法.bmp');%均衡化后的增强，没有经过增强，非线性的增强
path2=('边缘检测\非线性的增强\roberts.bmp');
path3=('边缘检测\非线性的增强\prewitt.bmp');
path4=('边缘检测\非线性的增强\sobel.bmp');
path5=('边缘检测\非线性的增强\krisch.bmp');
path6=('边缘检测\非线性的增强\LoG.bmp');
f=rgb2gray(f);
T=20;%阈值
[m,n]=size(f);
%-----梯度法-----
f_g=zeros(m,n);
for i=2:m-1
    for j=2:n-1
        f_g(i,j)=abs(f(i+1,j)-f(i,j))+abs(f(i,j+1)-f(i,j));
        if f_g(i,j)<T
            f_g(i,j)=0;
        else
            f_g(i,j)=255;
        end
    end
end
end
figure(1);

```

```
subplot(2,3,1);imshow(uint8(f_g));title('梯度法');
imwrite(f_g,path1);
```

```
%-----roberts 算子-----
```

```
f_r=zeros(m,n);
for i=2:m-1
    for j=2:n-1
        f_r(i,j)=abs(f(i+1,j+1)-f(i,j))+abs(f(i,j+1)-f(i+1,j));
        if f_r(i,j)<T
            f_r(i,j)=0;
        else
            f_r(i,j)=255;
        end
    end
end
```

```
end
%f_r=imbinarize(imfilter(f_r,T));
imwrite(f_r,path2);
subplot(2,3,2);imshow(uint8(f_r));title('Roberts 算子');
```

```
%-----prewitt 算子-----
```

```
f_p=zeros(m,n);
for i=2:m-1
    for j=2:n-1

f_p(i,j)=abs(f(i-1,j-1)+f(i,j-1)+f(i+1,j-1)-f(i-1,j+1)-f(i,j+1)-f(i+1,j+1))+abs(f(i+1,j-1)+f(i+1,j)+f(i+1,j+1)-f(i-1,j-1)-f(i-1,j)-f(i-1,j+1));
        if f_p(i,j)<15
            f_p(i,j)=0;
        else
            f_p(i,j)=255;
        end
    end
end
imwrite(f_p,path3);
subplot(2,3,3);imshow(uint8(f_p));title('Prewitt 算子');
```

```
%-----sobel 算子-----
```

```
f_s=zeros(m,n);
for i=2:m-1
    for j=2:n-1

f_s(i,j)=abs(f(i-1,j-1)+2*f(i,j-1)+f(i+1,j-1)-f(i-1,j+1)-2*f(i,j+1)-f(i+1,j+1))+abs(f(i+1,j-1)+2*f(i+1,j)+f(i+1,j+1)-f(i-1,j-1)-2*f(i-1,j)-f(i-1,j+1));
        if f_s(i,j)<T
```



```

        f_s(i,j)=0;
    else
        f_s(i,j)=255;
    end
end
end
imwrite(f_s,path4);
subplot(2,3,4);imshow(uint8(f_s));title('Sobel 算子');

%-----krisch 算子-----
k(:,1)=[-3,-3,-3;
        -3,0,5;
        -3,5,5];
k(:,2)=[-3,-3,5;
        -3,0,5;
        -3,-3,5];
k(:,3)=[-3,5,5;
        -3,0,5;
        -3,-3,-3];
k(:,4)=[-3,-3,-3;
        -3,0,-3;
        5,5,5];
k(:,5)=[5,5,5;
        -3,0,-3;
        -3,-3,-3];
k(:,6)=[-3,-3,-3;
        5,0,-3;
        5,5,-3];
k(:,7)=[5,-3,-3;
        5,0,-3;
        5,-3,-3];
k(:,8)=[5,5,-3;
        5,0,-3;
        -3,-3,-3];
kk=zeros(size(f));
I=double(f);
for i=1:8
    f_k(:,i)=conv2(I,k(:,i),'same');
    kk=max(kk,f_k(:,i));
end
f_kk=imbinarize(kk,600);
imwrite(f_kk,path5);
subplot(2,3,5);imshow(f_kk);title('Krisch 算子');

```

公众号关注：建模忠哥  
获取更多资源

```
%-----LoG 算子-----  
log1=[0 0 -1 0 0;  
      0 -1 -2 -1 0;  
      -1 -2 16 -2 -1;  
      0 -1 -2 -1 0;  
      0 0 -1 0 0];  
  
f_l=conv2(f,log1,'same');  
f_ll=imbinarize(abs(f_l),300);  
imwrite(f_ll,path5);  
subplot(2,3,6);imshow(f_ll);title('LoG 算子');
```

公众号关注：建模忠哥  
获取更多资源

公众号关注：建模忠哥  
获取更多资源