

光学学报

基于编解码深度估计的雾天公路能见度检测方法

彭朋^{1,3*}, 董宇成¹, 李家春^{2*}, 姚怡涛¹

¹陕西科技大学电气与控制工程学院, 陕西 西安 710021;

²长安大学公路学院, 陕西 西安 710064;

³西北工业大学柔性电子研究院, 陕西 西安 710129

摘要 为解决传统能见度检测方法依赖人工标志物、效率低且误差较大的问题,提出了一种基于编解码深度估计的能见度检测方法。该方法以大气散射物理模型为基础,重点获取图像的大气透射系数与场景深度信息。首先,采用K均值聚类法分割公路监控视频图像中的有雾区域和可视道路区域,定位道路的最大场景深度区域;其次,结合区域熵值法改进暗通道先验算法,精确估算图像的大气透射系数;随后,利用基于编解码结构的深度估计网络获取图像深度信息,选用去除顶层分类任务的DenseNet-169网络作为编码器,引入Dense Block-B模块抑制网络输入中的特征冗余现象,提高特征提取准确性,并在编码器卷积模块中嵌入卷积模块的注意力机制模块(CBAM),以增强路面特征提取效果,减少无关特征干扰;最后,基于大气散射物理模型计算能见度。实验结果显示,该方法在准确率上达到89.83%,较BTS、LapDepth、MonoDepth、CADepth以及Lite-Mono方法分别提升5.90、5.00、4.90、4.23、3.20个百分点,表现出更高的准确性与鲁棒性。

关键词 能见度检测; 大气散射模型; 暗通道先验; 深度编解码网络; 图像分割

中图分类号 TP391.41 文献标志码 A

DOI: 10.3788/AOS250467

1 引言

能见度是指视力正常的人在所处当时的天气条件下,不借助外物便能够从天空背景中看到和辨认出目标物的最大水平距离^[1],是反映大气透明度的重要参数以及交通领域的重要指标。影响能见度的因素众多,其中雾霾天气对能见度的影响最大,雾霾天气空气中悬浮着的大量微小颗粒,在阻碍大气光传输的同时,会吸收来自物体表面反射的光线,使能见度降低。雾霾天气也是引发道路交通事故的主要因素之一^[2],据中国气象局统计,全国高速公路团雾易发路段3188段,涉及路网里程21000 km以上,每年因雾霾等恶劣天气造成的交通事故约占所有交通事故的24%,对社会经济和出行人的生命安全造成重大影响。因此,准确、高效地获取能见度数据,为交通部门提供数据支撑以及时做出相应的交通管控是近年来地面气象领域的重要研究内容之一。

当前,能见度获取的主要途径包括人工标识观测、仪器检测和基于视频图像的检测方法^[3]。人工标识观测通常由专业人员依据经验来观测放置的标志物,进而估算能见度,但该方法受限于观测者的主观判断和环境条件,效率低且误差较大。能见度仪器检测法依

赖于散射式或透射式能见度监测仪进行测量,虽然避免了人为因素的影响且精确度显著提升,但其测量范围较小,通常仅能反映局部区域的能见度^[4]。此外,该方法的成本较高、安装场景有限,难以满足公路全覆盖检测需求。随着全国高速公路监控网的逐步完善和机器视觉技术的快速发展,基于视频图像的检测方法逐渐成为公路能见度检测的研究热点,该方法主要包括基于大气散射物理模型的检测方法和基于样本数据驱动的检测方法^[5]。

基于大气散射物理模型的能见度检测方法通过大气散射模型研究图像特征信息与能见度之间的关系,推算能见度信息。杨天麟等^[6]基于矩形区域测距法和实际场景的物体大小,实现高速公路二维到三维的场景重构,并基于暗通道先验方法计算能见度。廖苗等^[7]提出利用雾线先验和霍夫投票估计大气光的想法,并基于图像颜色特征与大气透射系数的关系建立非线性模型。宋海声等^[8]引入暗通道先验理论,通过自适应去雾权值优化大气光值,结合自适应滤波窗口获取大气透射系数,并通过车道线首尾端点的透射系数计算大气能见度。此外,Li等^[9]通过比较输入图像与参考图像的透射系数,反演计算大气消光系数以实现能见度估算。这些研究均表明,基于物理

收稿日期: 2025-01-10; 修回日期: 2025-02-06; 录用日期: 2025-02-27; 网络首发日期: 2025-03-05

基金项目: “十四五”重点研发计划(2022YFD2100600)、陕西省自然科学基础研究计划(2024JC-YBQN-0400)

通信作者: *zhs@chd.edu.cn; **pengpeng@sust.edu.cn

模型的方法可以有效利用图像自身特征与大气散射模型的关系实现能见度检测。然而,暗通道先验算法在复杂环境中可能存在误差,需要进一步优化以提高适用性。

另一方面,基于样本数据驱动的方法依赖深度学习技术,通过对标注图像的大规模训练,学习图像的像素特征与能见度之间的非线性映射关系。例如:Outay等^[10]使用预训练的AlexNet结合非线性支持向量机(SVR)识别能见度;唐绍恩等^[11]通过预训练的VGG-16对图像子区域进行编码,并使用支持向量回归融合权重计算能见度;You等^[12]结合卷积神经网络(CNN)和递归神经网络(RNN),利用全局特征和局部细节从多尺度视角估计能见度。尽管数据驱动方法在特征提取与建模能力上具有优势,但其训练依赖于大规模标注数据集,并可能在实际应用中对样本域外的数据表现不足。

基于以上分析,本文主要针对能见度检测方法进行改进,首先基于大气散射物理模型提出了优化的暗通道先验算法,随后结合深度估计网络,提升复杂场景下的能见度检测性能。具体而言,本文以大气散射物理模型为基础,利用图像分割技术提取场景最大深度区域,通过改进暗通道先验算法获取精确的大气透射系数,并引入深度估计网络以提取图像的深度信息,从而实现能见度的准确检测。

2 能见度检测基础理论

在高速公路场景中,能见度主要受光在大气中传播时介质作用导致的衰减影响。基于此原理,Koschmieder定律成为计算白天目标能见度的重要理论基础,其将大气消光系数与能见度联系起来,数学表达式为

$$C = C_0 e^{-\sigma d} \quad (1)$$

式中: C 为目标观测物的观测亮度值; C_0 为目标观测物的本身自有亮度值; σ 为大气消光系数; d 为图像中观测点到观测目标的距离。目标观测物的观测亮度值和固有亮度值之比为视觉对比阈值 ϵ ,可表示为

$$\epsilon = \frac{C}{C_0} \quad (2)$$

国际照明委员会(CIE)规定:当视觉对比阈值 ϵ 为0.05时,人眼最大可视距离为能见度距离 V ,将式(2)代入式(1)中,可得

$$V = d = -\frac{\ln \epsilon}{\sigma} = \frac{2.996}{\sigma} \quad (3)$$

式(3)中,当视觉对比阈值 ϵ 设置0.05时,图像中观测点到观测目标的距离即能见度距离 V ,且能见度距离 V 与大气消光系数 σ 成反比。综上可得,通过计算大气消光系数 σ 可反演能见度距离 V ,本文基于大气散射模型求取大气消光系数 σ 。

基于光在大气中的传播过程受介质影响而逐渐衰

减的原理,大气散射模型应运而生,其物理意义在于:观测到的光信号是目标物体发出的光信号在经过大气吸收与散射衰减后,与大气自身辐射和散射光叠加的结果,其数学表达式如下:

$$I(u, v) = e^{-\sigma D(u, v)} J(u, v) + A [1 - e^{-\sigma D(u, v)}] \quad (4)$$

式中: I 为观察到的图像强度(有雾图像的强度); J 为场景辐射亮度(无雾的图像); $D(u, v)$ 为图像深度信息; A 为大气透光强度值; $e^{-\sigma D(u, v)}$ 表示大气透射系数,记作 $t(u, v)$,其值介于0和1之间:

$$t(u, v) = e^{-\sigma D(u, v)} \quad (5)$$

经变换可得

$$\sigma = -\frac{\ln t(u, v)}{D(u, v)} \quad (6)$$

由式(6)可得,大气消光系数 σ 可通过计算图像透射系数 $t(u, v)$ 和图像深度信息 $D(u, v)$ 获得。将式(6)代入式(3),可得

$$V = -\frac{2.996 D(u, v)}{\ln t(u, v)} \quad (7)$$

综上,通过大气消光系数 σ 可将能见度距离 V 与大气透射系数 $t(u, v)$ 和图像深度信息 $D(u, v)$ 关联起来,从而实现能见度的计算。

3 雾天能见度估计方法

由式(7)可知,能见度距离 V 可以通过分别求解大气透射系数 $t(u, v)$ 和图像深度信息 $D(u, v)$ 获得。本文提出的雾天能见度估计算法的整体框架如图1所示,该框架主要包括三个关键部分:1)进行图像场景分割,获取图像最大的景深区域;2)通过改进的暗通道先验方法计算图像最大景深区域的大气透射系数;3)基于编码-解码结构的深度估计网络,提取最大景深区域的深度信息,从而完成能见度的计算。

3.1 图像场景分割

常见的图像分割方法可分为基于深度学习的方法和传统图像分割方法。基于深度学习的方法通过训练深度网络模型,对大量已标注的图片进行特征提取学习,通过不断迭代实现图像的分割,常见的网络包括U-Net^[13]、Mask R-CNN^[14]、DeepLab^[15]系列等,虽然这些网络的泛化性较好,但需要在对海量的数据样本进行标注后,进行训练学习。传统的图像分割方法包括区域增长、阈值分割、边缘检测以及分水岭算法。它们都利用像素相似性、设置像素阈值、亮度突变、拓扑理论等实现图像分割,对于背景复杂度不高的图像有较好的分割效果。

K 均值聚类算法^[16]是一种经典的图像分割方法,其基于像素相似性原理,将数据中的样本划分为若干类别,使相似度高的样本归为同一类别,从而增大不同类别之间的差异性,其核心目标是将数据点划分为 K 个簇,并通过优化每个簇的中心位置,使类内差异最小

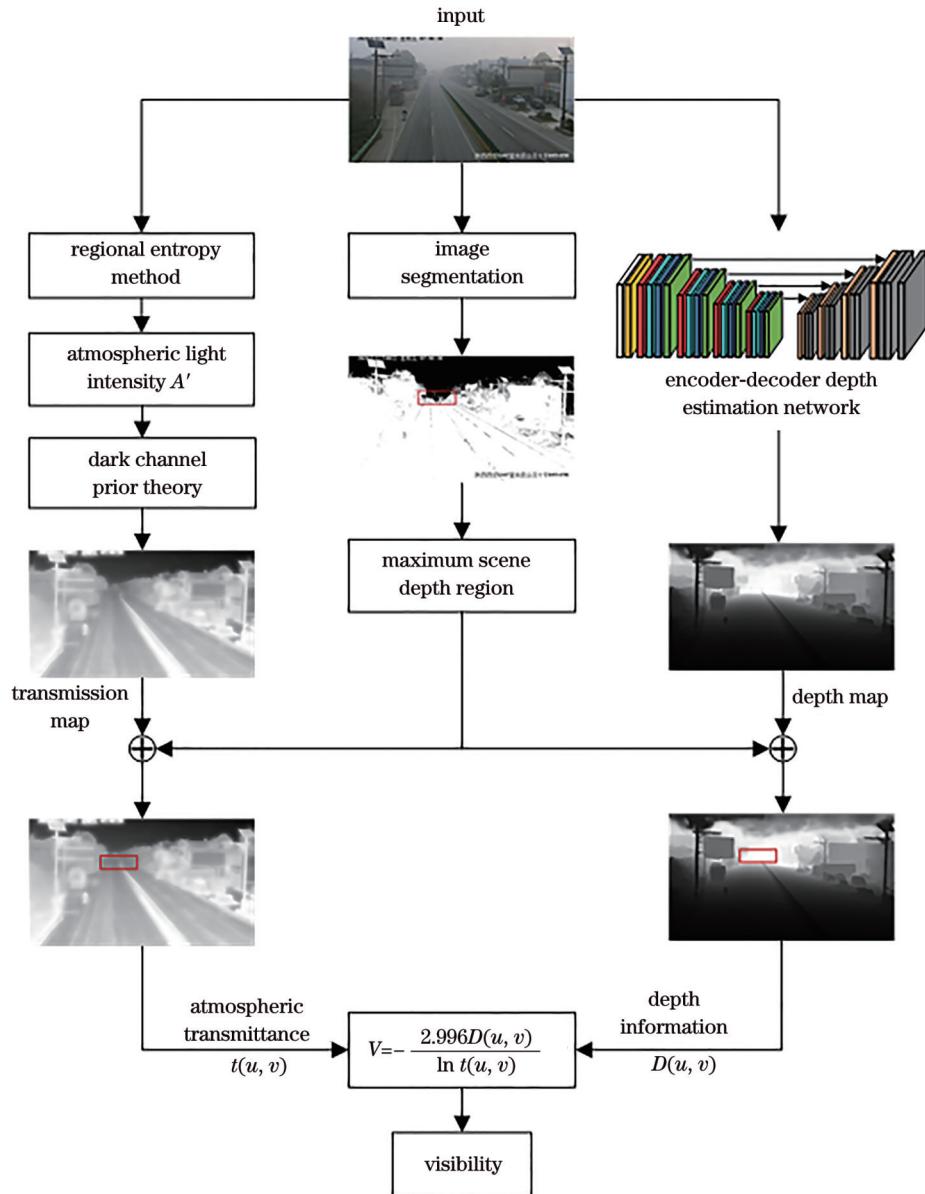


图 1 雾天能见度估计算法框架

Fig. 1 Framework of foggy visibility estimation algorithm

化,从而实现高效聚类和图像分割。在研究大量道路有雾图像的基础上,发现对于雾霾天气的图像,使用灰度图像进行处理更加合理。此类图像的主要特征是:灰度值大的区域通常对应天空或有雾区域,灰度值小的区域对应可视道路区域,在有雾区域与可视道路区域的交界处灰度值会出现显著突变,而其他区域的灰度值变化较为平缓。基于这一特点,本文采用 K 均值聚类算法对道路有雾图像进行分割。具体而言,本文将灰度图像中的数据点划分为两个簇:一个簇对应灰度值较大的有雾区域,如天空和雾霾;另一个簇对应灰度值较小的可视道路区域。基于相似性原则的聚类方法可将输入图像中的灰度值有效分离,实现对道路有雾图像的准确分割,为复杂气象条件下的道路分析提供技术支持。通过设置亮度阈值生成图像的亮区掩

模,帮助聚类算法更好地分辨不同的对象或特征,从而提高分类的准确性。具体而言,首先将彩色图像转换为灰度图像,接着设定亮度阈值为 150(灰度值),大于该阈值的像素为亮区,然后将灰度图像中大于设定阈值的部分标记为 1,其他部分标记为 0,从而生成一个二值掩模图像。此掩模图像中的 1 表示亮区,0 表示其他区域。图像分割结果见图 2。图 2(a)为输入的雾天道路原图,图 2(b)为输入原图的亮区掩模,图 2(c)为对道路图像进行分割的结果。

在雾天公路场景中,浓雾区域和道路区域为图像的主体部分,同时有雾区域与可视道路区域的灰度值差异较大。这种特性使得 K 均值聚类算法对天空和路面进行聚类更加适用,能够有效分割有雾区域和可视道路区域。相比于基于深度学习的图像分割方法,该

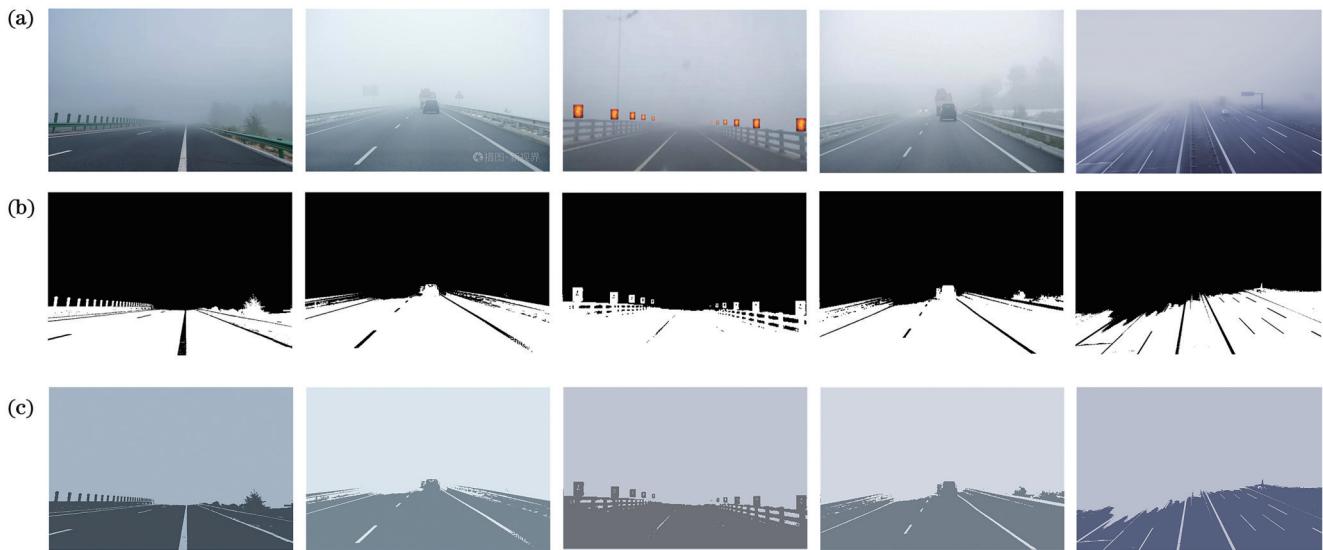


图2 图像分割结果图。(a)原图;(b)亮区掩模;(c)分割结果

Fig. 2 Image segmentation results. (a) Original images; (b) bright area masks; (c) segmentation results

算法不仅显著降低了工作量,同时也能保持良好的分割性能。在分割后的图像中,对逐个像素点与其相邻像素点进行比较,当检测到像素值发生突变时,则判定该像素点为最大景深区域内的边界像素点。对整个图像的所有像素点进行遍历,记录所有满足突变条件的像素点,最终确定最大景深区域的边界范围,并进一步提取完整最大景深区域。

3.2 改进暗通道先验算法获取透射系数

3.2.1 求图像大气光强度

暗通道先验理论由He等^[17]提出,在大多数非天空区域图像中,RGB三个通道中至少有一个通道在某些像素点上为接近于0的值,该通道被称为暗通道。对于一幅图像 J ,暗通道可以定义为

$$J^{\text{dark}}(u, v) = \min_{c \in \{r, g, b\}} \left[\min_{y \in \Omega(u, v)} J^c(u, v) \right] \quad (8)$$

式中: J^c 是 J 的颜色通道; $\Omega(u, v)$ 是以像素 (u, v) 为中心的区域; J^{dark} 是图像 J 的暗通道。根据暗通道先验知识,无雾图像的暗通道趋于0,即

$$J^{\text{dark}} \rightarrow 0 \quad (9)$$

在暗通道先验理论中,大气透光强度值是一种用来衡量光线透过大气后亮度变化的指标,其计算方法基于暗通道图像特性。首先,将原始图像转化为暗通道图像,以突出图像中较暗的区域,同时保留亮度极高的部分。随后,在暗通道图像中选取亮度最高的前0.1%的像素区域,这些区域通常对应场景中受强光影响的部分。接着,在原始图像中找到这些像素区域对应的亮度值,并从中挑选最高的亮度值作为大气透光强度的参考值。在筛选这些区域时,原始图像中可能存在背景干扰,例如白色区域、强反光区域或高亮度的灯光区域。这些干扰区域的亮度较高,可能与暗通道图像中的高亮度区域混淆,从而导致对大气透光强度

值的误判。这种误判会直接影响大气透光强度的精确计算,进而对后续的图像处理或相关参数估算造成不利影响。因此,本文对暗通道先验算法进行改进,将图像中的天空区域与道路区域进行分割,然后用区域熵值法求解大气透光强度,在此基础上基于暗通道先验求解图像透射系数。改进后的暗通道先验算法有助于更精准地筛选雾浓度最大位置的像素点作为大气光强度值,避免像素值分布不均匀导致的大气光强度值不准确。具体步骤如下:

1) 在天空有雾区域与可视道路区域的交界处使用图像分割算法将图像分为天空有雾区域和可视道路区域,对于道路有雾图像,除交界处灰度值出现突变外,其余区域的灰度值变化通常较为平缓,本文使用 K 均值聚类算法,根据相似性原则将图像中灰度值大的天空有雾区域划分为同一类簇,将灰度值小的道路区域划分为同一类簇。图3(a)为输入的原始视频图像,图3(b)为图像分割后的结果,其中矩形框像素位置为最大景深位置。

2) 对图像分割结果中路面与天空交界的最大景深位置区域引入区域熵值法进行大气光强度值的筛选。区域熵值定义为

$$h(u, v) = - \sum_{l=0}^{L-1} P_l \log_2 P_l \quad (10)$$

式中: l 为邻域大小 $m \times n$ 的方格区域内的灰度值; L 为像素值0到255的256个灰度级; P_l 为邻域大小 $m \times n$ 的方格区域内像素灰度值 l 的概率, P_l 的定义为

$$P_l = \frac{n_l}{m \times n} \quad (11)$$

式中: n_l 为在大小为 $m \times n$ 的局部邻域内,灰度值为 l 的像素数。

由式(10)和式(11)可知,对于公路而言,路面的像

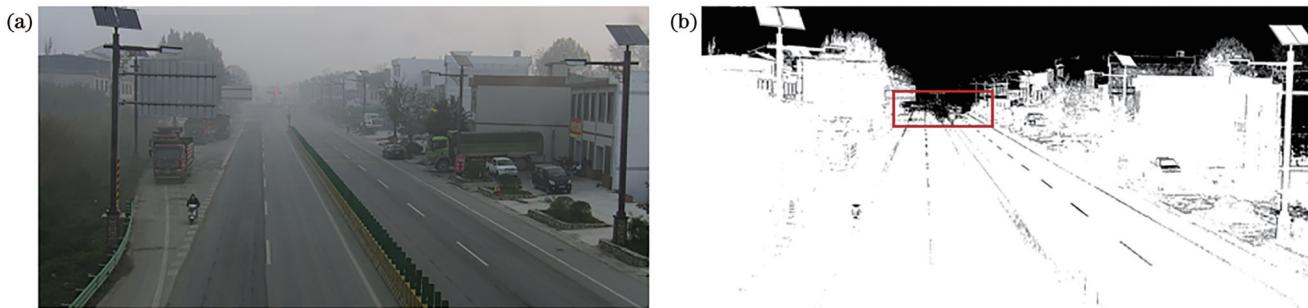


图3 视频图像分割。(a)视频图像;(b)分割结果

Fig. 3 Video image segmentation. (a) Video image; (b) segmentation result

素值相较于有雾区域更低:当雾浓度较低时,区域灰度值变化不稳定,区域熵值较大;当雾浓度较高时,区域的灰度值稳定则熵值较小,浓雾区域的准确性较高。针对原始图像,采用区域熵值法进行处理,结果如图4所示。图4中亮度较高的区域表示像素变化剧烈的区域,而亮度较低的区域则表示像素变化相对平缓的区域。通过区域熵值法在目标区域中筛选出熵值最小的像素位置,并将该位置的灰度值确定为大气光强度值。

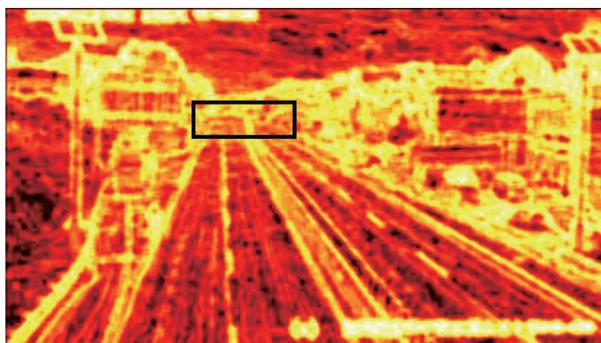


图4 区域熵值法结果图

Fig. 4 Image of result of local entropy method

3.2.2 求取透射系数

在暗通道理论的基础上,求取 x 邻域内的透射系数。首先假设以 x 为中心的邻域内透射系数相等,由式(8)联立式(4),并对两边求取最小值:

$$\frac{\min_{c \in \{r, g, b\}} \left\{ \min_{y \in \Omega(u, v)} [I^c(u, v)] \right\}}{A'} = t(u, v) \frac{\min_{c \in \{r, g, b\}} \left\{ \min_{y \in \Omega(u, v)} [I^c(u, v)] \right\}}{A'} + 1 - t(u, v) \quad (12)$$

式中: A' 为大气光强度。

根据暗通道先验知识,无雾图像的暗原色通道趋于0,大气环境光满足

$$\min_{c \in \{r, g, b\}} \left\{ \min_{y \in \Omega(u, v)} [J^c(u, v)] \right\} = 0 \quad (13)$$

将式(13)代入式(12),可以得到透射系数为

$$t(u, v) = 1 - \frac{\min_{c \in \{r, g, b\}} \left\{ \min_{y \in \Omega(u, v)} [I^c(u, v)] \right\}}{A'} \quad (14)$$

根据式(14)可得到图像像素点的大气透射系数,但实际大气中存在一定的雾霾颗粒,导致能见度降低。因此,需要在计算式(14)时在减数项前添加误差系数 ω , ω 通常取0.9^[18],得到的表达式为

$$t(u, v) = 1 - \omega \frac{\min_{c \in \{r, g, b\}} \left\{ \min_{y \in \Omega(u, v)} [I^c(u, v)] \right\}}{A'} \quad (15)$$

根据式(15)可得到图像中每个像素点的透射系数,所有像素点的透射系数组成透射图。图5为获取的透射图像。如图5(c)所示,获得的原始透射图受到噪声的影响,图像较为粗糙。为解决该问题,引入导向滤波的方法^[19],该方法在保护图像边缘信息的同时具有良好的滤波效果,如图5(d)所示。

3.3 基于DenseNet-169的编解码网络获取深度信息

本文使用单目视觉估计图像的场景深度信息。基于DenseNet-169^[20]的编码器解码器网络实现深度估计,网络结构如图6所示,图中BN代表批归一化。该网络基于DenseNet-169构建,采用编码-解码结构,其中编码器通过Dense Block的密集连接提取多层次的特征,解码器通过上采样逐步恢复分辨率。通过引入Dense Block-B模块抑制网络输入中的特征冗余现象,提高特征提取准确性,并针对原始图像背景中存在的干扰影响,在编码器卷积模块中嵌入卷积模块的注意力机制模块(CBAM),以增强路面特征的提取效果,减少无关特征干扰。通过 1×1 和 3×3 卷积操作实现特征压缩与空间特征提取,多尺度信息的融合进一步提升了网络的性能,最终输出图像场景的深度图,每个像素点的深度信息可在深度图中查询获得。

如图7所示,DenseNet编码器采用密集连接机制,将所有层级相互连接,每一层的输入由前面所有层的输出在通道维度上进行连接与合并,这种设计可实现特征复用,既可保留低维度的特征,又有效缓解梯度消失问题,保证网络的性能。然而,这种特征传输方式也带来了两个问题:

1) 特征冗余。多层级的特征融合可能导致冗余或无关特征被重复利用,增加了网络的计算负担。因此,需要进一步优化DenseNet的特征提取与传输机制,以更好地提高网络提取深度信息的准确性。

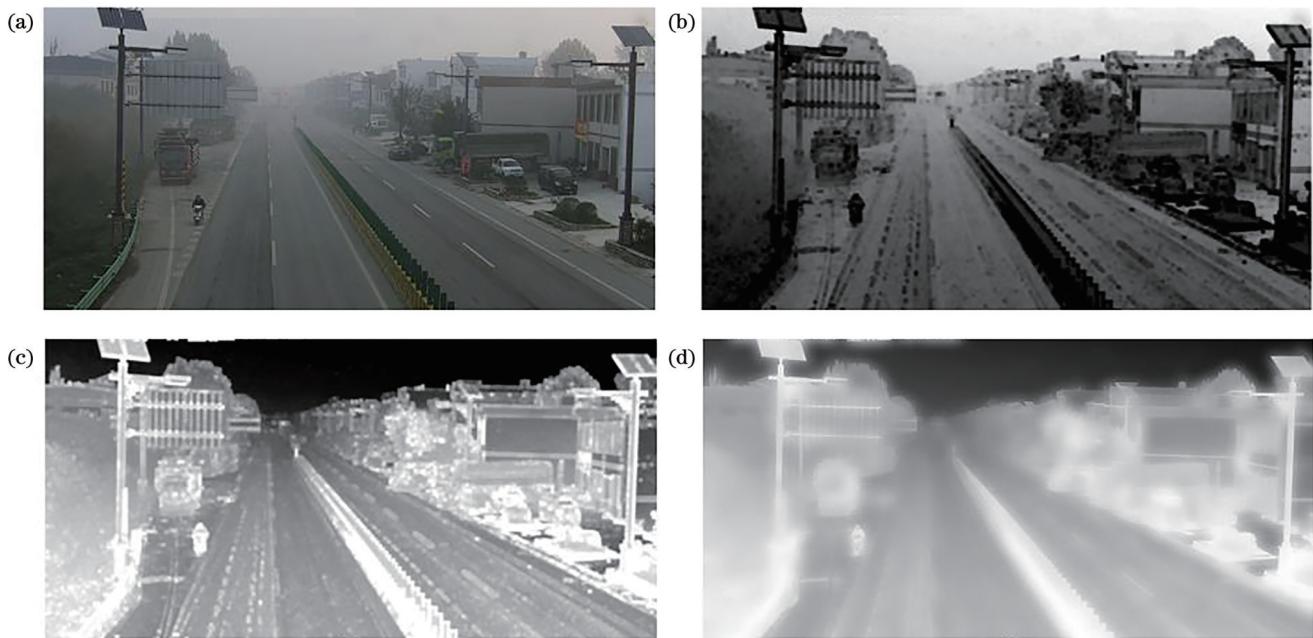


图 5 透射图像获取。(a)视频图像;(b)暗通道图像;(c)原始透射图;(d)优化透射图

Fig. 5 Acquisition of transmission maps. (a) Video image; (b) dark channel image; (c) original transmission map; (d) optimized transmission map

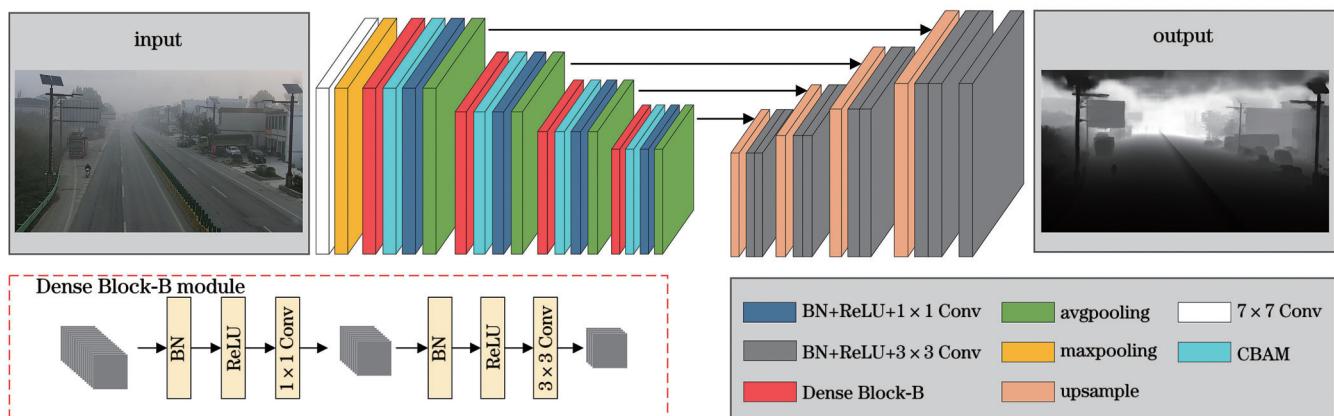


图 6 基于 DenseNet-169 的编解码网络
Fig. 6 Encoder-decoder network based on DenseNet-169

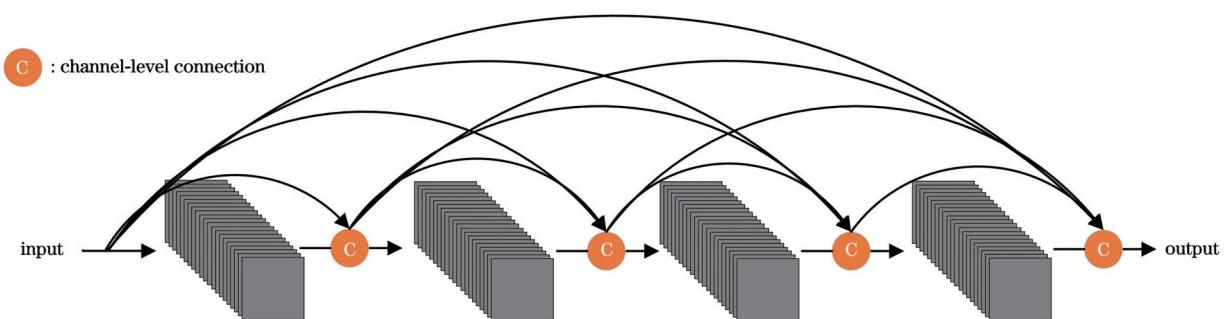


图 7 密集连接原理机制
Fig. 7 Mechanism of dense concat

2) 无关特征干扰。由于输入图像中包含大量背景信息(如树木、车辆、房屋等),网络在学习能见度特征时,易引入无关特征,影响识别准确性。

3.3.1 改进 Dense Block 密集块

为了解决 DenseNet 中因特征冗余而导致深度信息提取准确率较低的问题,对 Dense Block 结构进行了

改进,提出了 Dense Block-B。在原始的 Dense Block 中,每一层的特征图都与前面所有层的特征图进行拼接,实现特征复用。尽管这种方法能够保留更多的信息,但随着网络深度的增加,这种特征拼接会引入大量冗余信息,导致计算和存储的负担加重。为了解决这一问题,在 Dense Block-B 中引入了瓶颈层(Bottleneck Layer)。

具体来说,在 Dense Block-B 中(图 8),输入的特

征图(通道数为 $l \times k$)首先通过一个 1×1 卷积层(1×1 convolution),将通道数压缩至 4000。这一操作显著减少了特征数量,达到了降维的目的,提升了模型的计算效率。 1×1 卷积能够有选择性地保留重要的特征通道,同时抑制冗余或无关的特征,减少特征图中的冗余信息,并且在降维的过程中优化了特征表示,增强了有效特征的表达能力。通过这一改进,Dense Block-B 能有效提升深度信息提取的准确性。

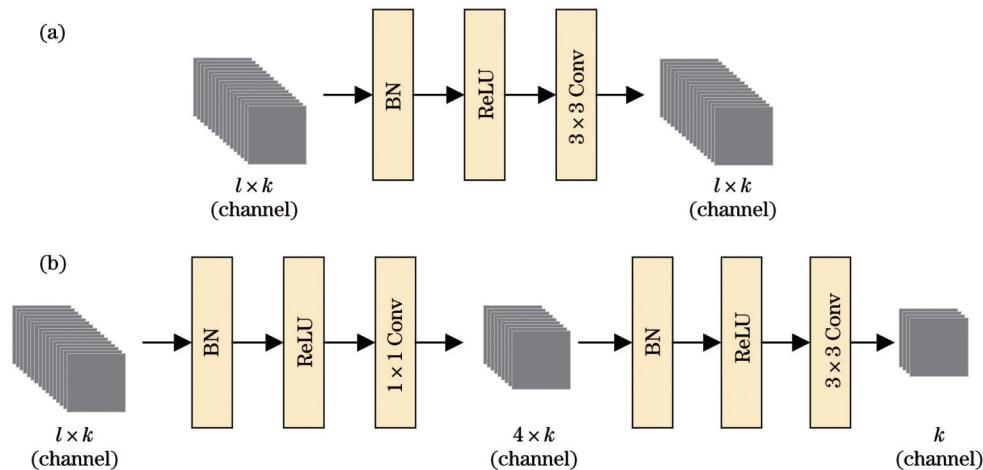


图 8 密集块结构改进。(a) Dense Block;(b) Dense Block-B

Fig. 8 Improvement of Dense Block. (a) Dense Block; (b) Dense Block-B

3.3.2 引入 CBAM

由于输入图像处于雾天环境,网络的卷积层和池化层在提取能见度特征时,往往会学习到大量与任务无关的特征(如树木、车辆、房屋等),从而降低网络的精度。为了解决这一问题,本文在网络编码器部分引入了 CBAM,通过在通道和空间两个维度上对特征图进行加权操作,CBAM 能够聚焦于有效特征并抑制无

关特征,从而帮助网络更准确地学习道路的深度信息,CBAM 结构如图 9 所示。

CBAM 由通道注意力模块和空间注意力模块两部分组成,将两者的输出特征逐元素相乘,最终得到精准的注意力特征。在通道注意力模块中,使输入特征图经过全局最大池化和全局平均池化操作,得到每个通道的最大特征值和平均特征值。这些值会

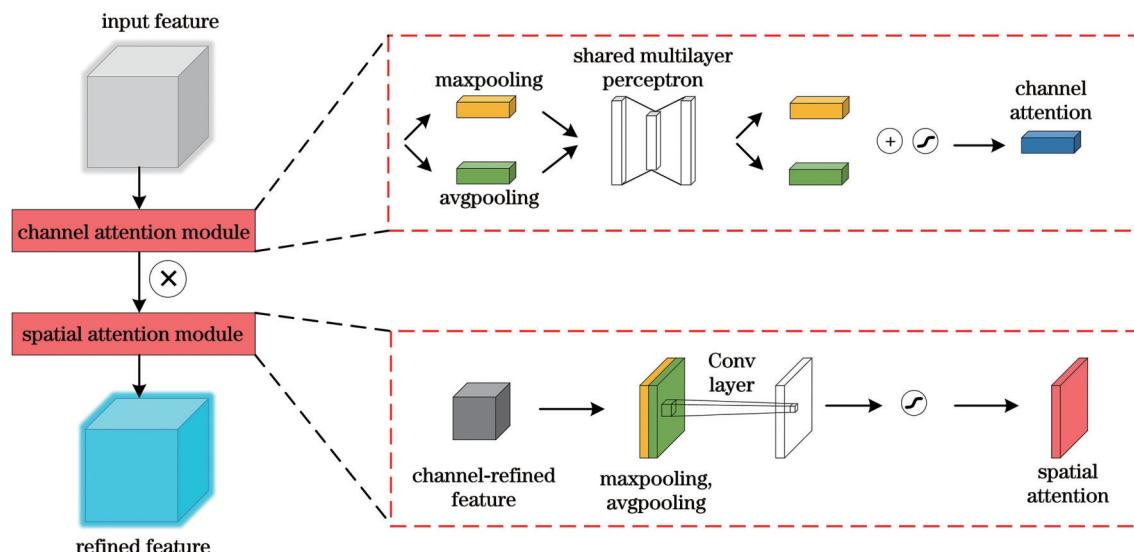


图 9 CBAM 结构

Fig. 9 Structure of CBAM

被输入到共享全连接层,用于学习每个通道的注意力权重,并通过 sigmoid 激活函数将权重限制在 0 到 1 之间,最终根据这些权重对通道特征进行加权输出。

在空间注意力模块中,对输入特征图在通道维度下进行最大池化和平均池化操作,以生成不同尺度的上下文信息。接着,将这些上下文信息拼接在一起,形成具有多尺度信息的特征图。通过卷积操作生成空间注意力权重,并对每个空间位置的特征进行加权,输出加权后的特征图。

解码器部分由双线性上采样和两个标准卷积层并结合跳跃连接组成。每个模块包含一个双线性上采样层和两个 3×3 卷积层,输出滤波器的数量设置为输入滤波器数量的一半。第一个卷积层接收前一层的输出和与之具有相同空间维度的编码器池化层的输出,并进行串联。通过跳跃连接,网络有效地将

编码器中的特征图细节传递到解码器,并与解码器内部的特征图融合,帮助解码器生成更加细致的输出图像。编码器的特征图与解码器的特征图通过加法方式融合,而非并列融合,这避免了增加解码器中的通道数,从而优化了计算资源的使用。

4 实验结果和分析

4.1 实验环境与数据来源

本文实验在台式机上完成,实验设备配置如下:操作系统为 Windows 11,处理器为 Intel i9-13900KF,显卡为 NVIDIA RTX A6000,显存为 48 GB,运行内存为 64 GB。所用网络模型基于 PyTorch 1.9.1 框架,CUDA 版本为 11.7,CUDNN 版本为 7.6。实验数据采集与西安市公路局合作完成,在陕西省西安市 S107 蓝田段公王十字 K49+890(图 10)建立了能见度监测点,安装了固定相机及配套能见度监测仪。

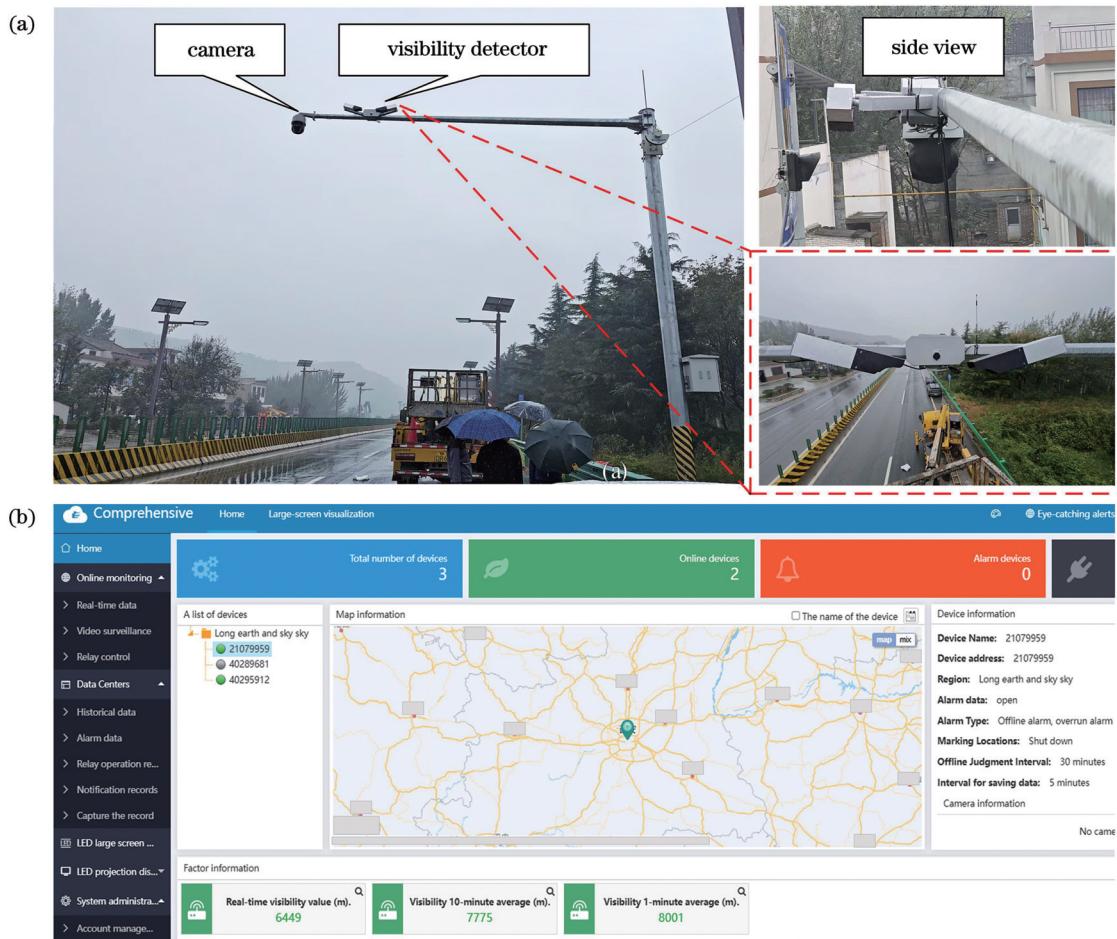


图 10 能见度监测点。(a)设备部署;(b)云平台监测

Fig. 10 Visibility monitoring station. (a) Device deployment; (b) cloud platform monitoring

如图 10 所示,监测点数据采集自 2024 年 9 月开始,每天从 07:00 至 21:00,每隔 5 min 采集一次,将采集数据上传至云平台保存,数据记录至今持续进行中,为实验提供了充足的样本支持。

4.2 评价指标

本文能见度估计模型基于单目深度估计网络,在单目深度估计任务中准确率(A)是常用的评价指标,它表示模型估计深度值和实际深度值之间的匹配程

度^[21]。准确率越高代表模型估计深度值与实际深度值的相对误差越小,模型的泛化性越好,其表达式为

$$A = \max\left(\frac{d_i}{\hat{d}_i}, \frac{\hat{d}_i}{d_i}\right) < \delta (\delta = 1.25, 1.25^2, 1.25^3) \quad (16)$$

式中: d_i 为实际深度值; \hat{d}_i 为估计深度值; δ 为像素阈值,表示所有估计深度值中小于所设阈值的像素点占总体像素的比例,不同阈值表示模型对误差的不同容忍程度。 δ 为1.25时,模型对误差的容忍度低,表示大多情况下模型可以估计出准确的数值,此时多用于评估模型近距离的深度误差; δ 为1.25²时,模型对误差的容忍度较低,准确率较高,多用于评估模型中等距离的深度误差; δ 为1.25³时,模型对误差的容忍度相对较高,准确率相对于前两个阈值较低,多用于评估模型远距离的深度误差。结合3种阈值可更加全面地评估模型性能。

绝对相对误差(A_R)代表每个像素的估计深度值和实际深度值之间的绝对差值占实际深度值的比例,通常以所有像素的平均值来表示,反映模型整体的估计性能,表达式如下

表1 不同深度估计网络算法的性能比较

Table 1 Comparison of performance of different depth estimation network algorithms

Model	A_R	RMSE /m	A /%		
			δ is <1.25	δ is <1.25 ²	δ is <1.25 ³
BTS	0.116	0.390	81.0	84.1	86.7
LapDepth	0.101	0.432	84.1	84.8	85.6
MonoDepth	0.109	0.440	85.4	85.3	84.1
CADepth	0.103	0.408	85.7	86.0	85.1
Lite-Mono	0.099	0.389	87.9	86.5	85.5
DenseNet-169	0.097	0.385	88.6	86.5	85.4

4.3.2 不同注意力机制下基于DenseNet-169的编解码网络训练结果

注意力机制可通过施加权重的方式对图像无关的外部特征进行过滤,关注有效特征,提升模型的鲁棒性。为提升网络性能、验证算法的有效性,本文在基于DenseNet-169的编解码网络中添加4种不同的

$$A_R = \frac{1}{N} \sum_{i=1}^N \frac{|d_i - \hat{d}_i|}{d_i} \quad (17)$$

式中: N 为图像像素点数。

均方根误差(RMSE)是估计深度值与实际深度值之间差的平方的平均值的平方根,可更全面地反映模型的深度估计整体精度,RMSE可表示为

$$R_{\text{MSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N |d_i - \hat{d}_i|^2} \quad (18)$$

4.3 实验结果和分析

4.3.1 不同深度估计网络算法的性能比较

实验选择VKITTI 2以及自建深度数据集进行训练,基于上述3个评价指标对BTS^[22]、LapDepth^[23]、MonoDepth^[24]、CADepth^[25]、Lite-Mono^[26]、DenseNet-169 6种不同的单目深度估计网络进行比较,对比结果见表1,其中RMSE、 A_R 越小则模型性能越好,准确率A越高则模型性能越好。

测试结果如表1所示,相比于BTS、LapDepth、MonoDepth、CADepth、Lite-Mono 5种深度估计网络,DenseNet-169的性能更好,在RMSE、 A 、 A_R 指标上均优于其他模型。

注意力机制进行实验。实验加入的4种注意力机制分别为CA^[27](Coordinate Attention)、CBAM^[28]、SSA^[29](Separable Self Attention)、ECA^[30](Efficient Channel Attention),将以上4种注意力加入基于DenseNet-169的编解码网络中,采用RMSE、 A 、 A_R 这三个评价指标进行对比分析,测试结果见表2,将

表2 不同注意力机制下的网络性能对比

Table 2 Comparison of network performance under different attention mechanisms

Model	A_R	RMSE /m	A /%		
			δ is <1.25	δ is <1.25 ²	δ is <1.25 ³
DenseNet-169	0.097	0.385	88.6	86.5	85.4
Dense+CA	0.088	0.323	87.4	90.1	89.2
Dense+ECA	0.098	0.379	92.9	86.8	87.3
Dense+SSA	0.081	0.310	86.3	88.9	90.3
Dense+CBAM	0.082	0.305	89.6	90.4	89.5

CBAM引入DenseNet-169后,与其他3种注意力机制相比,其RMSE和 A_r 最小,在阈值 δ 为1.25²时准确率最高。本文改进的Dense+CBAM模型的平均准确率为89.83%,BTS、LapDepth、MonoDepth、CADepth以及Lite-Mono方法的平均准确率为83.93%、84.83%、84.93%、85.60%以及86.63%(表1)。与BTS、LapDepth、MonoDepth、CADepth、Lite-Mono相比,本文方法在平均精度上分别提升了5.90、5.00、4.90、4.23、3.20百分点。

4.3.3 实验结果分析

中华人民共和国国家标准GB/T 33673—2017中规定水平能见度等级,如表3所示,水平能见度 $V \geq 10 \text{ km}$ 为优等级(1级), $2 \text{ km} \leq V < 10 \text{ km}$ 为良等级(2级), $1 \text{ km} \leq V < 2 \text{ km}$ 为一般等级(3级), $500 \text{ m} \leq V < 1 \text{ km}$ 为较差等级(4级), $50 \text{ m} \leq V < 500 \text{ m}$ 为差等级(5级), $V < 50 \text{ m}$ 为极差等级(6级)。实验从2024年9月至今在陕西省西安市S107蓝田段公王十字K49+890处建立能见度监测点,采集雾天视频图像及能见度数据。

截取监测点2024年11月24日14时25分到15时15分的监控视频图像,视频图像见图11,实验结果见表4,该表记录了14:25—15:20期间的能见度数据,时

表3 水平能见度等级

Table 3 Levels of visibility

Level	Qualitative descriptor	Visibility
1	Excellent	$V \geq 10 \text{ km}$
2	Good	$2 \text{ km} \leq V < 10 \text{ km}$
3	Average	$1 \text{ km} \leq V < 2 \text{ km}$
4	Poor	$500 \text{ m} \leq V < 1 \text{ km}$
5	Bad	$50 \text{ m} \leq V < 500 \text{ m}$
6	Terrible	$V < 50 \text{ m}$

间间隔为5 min。本文基于视频图像提取的能见度值与能见度监测仪的结果进行比较,在能见度等级检测上,本文方法获取的实验结果与能见度监测仪基本保持一致,仅在15:00出现差异。具体来说,能见度监测仪测得的能见度为512 m,等级为4级,而本文研究方法测得的能见度为459 m,等级为5级。这是因为15:00的能见度为512 m,达到4级,而本文方法由于存在53 m的误差,等级判定为5级。整体来看,本文研究方法的平均误差为73 m,平均准确率达到89.83%,具备较高的测量精度与可靠性。

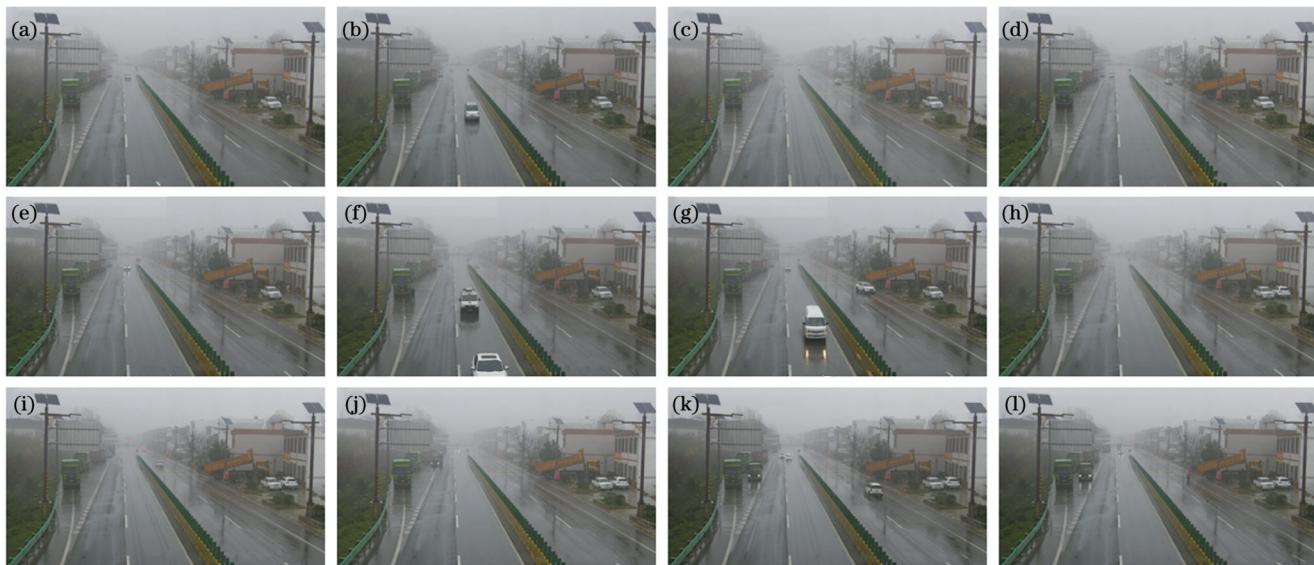


图11 以5 min间隔截取的监控视频图像。(a) 14:25;(b) 14:30;(c) 14:35;(d) 14:40;(e) 14:45;(f) 14:50;(g) 14:55;(h) 15:00;(i) 15:05;(j) 15:10;(k) 15:15;(l) 15:20

Fig. 11 Surveillance video frames captured at 5-minute interval. (a) 14:25; (b) 14:30; (c) 14:35; (d) 14:40; (e) 14:45; (f) 14:50; (g) 14:55; (h) 15:00; (i) 15:05; (j) 15:10; (k) 15:15; (l) 15:20

图12展示了能见度监测仪与基于DenseNet-169的5种网络方法在不同时刻的能见度检测结果和误差变化情况。Dense+SSA网络对远距离深度信息的关注度较高,在700 m以上范围的能见度预测中表现出较高的准确性,但对于近距离的深度信息关注不足,在低能见度条件下的预测准确率较低;Dense+ECA网络则对近距离深度信息更加敏感,这使得其在550 m

以内的能见度预测中表现较好,但随着能见度距离的增大,其预测准确率逐渐下降;未引入注意力机制的DenseNet-169网络相较于其他网络,虽然在15:00时刻的误差小于Dense+CA和Dense+SSA,但其他时刻估计的能见度值误差大;Dense+CA网络在14:30和14:55时刻的误差最小,在其他时刻的误差波动较大。相比之下,本文提出的方法在远距离和近距离的

表4 本文方法的能见度估计结果
Table 4 Visibility estimation results of proposed method

Time	Figure	Visibility detector		Proposed method		Error
		Visibility /m	Level	Visibility /m	Level	
14:25	Fig. 11(a)	632	4	573	4	59
14:30	Fig. 11(b)	697	4	615	4	82
14:35	Fig. 11(c)	595	4	646	4	51
14:40	Fig. 11(d)	681	4	716	4	35
14:45	Fig. 11(e)	704	4	629	4	75
14:50	Fig. 11(f)	670	4	587	4	83
14:55	Fig. 11(g)	688	4	615	4	73
15:00	Fig. 11(h)	512	4	459	5	53
15:05	Fig. 11(i)	793	4	683	4	110
15:10	Fig. 11(j)	653	4	721	4	68
15:15	Fig. 11(k)	834	4	736	4	98
15:20	Fig. 11(l)	860	4	771	4	89

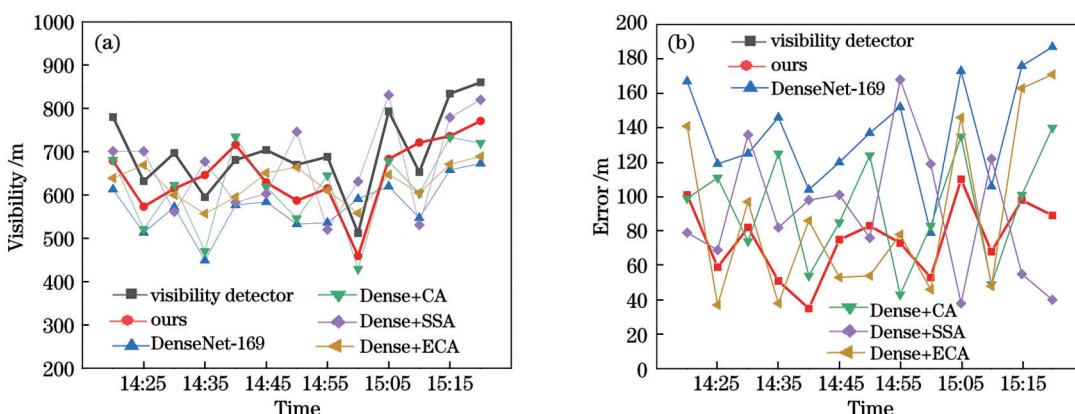


图12 基于不同模型方法的结果对比。(a)能见度曲线;(b)误差曲线

Fig. 12 Comparison of results based on different model methods. (a) Visibility curves; (b) error curves

能见度预测中均具备相对较高的准确性,整体误差较小,泛化能力更强。尤其是在14:25至15:20的整个时间段中,本文方法的检测结果与能见度监测仪提供的数据最为接近,表现出较高的鲁棒性和可靠性。与其他4种网络相比,本文方法能够有效平衡对远距离和近距离深度信息的关注,还实现了对能见度等级和具体数值的精确预测,整体性能更优。本文方法的准确率为89.83%。

5 结 论

针对传统能见度检测方法依赖人工标志物、效率低且误差较大的问题,提出一种基于编解码深度估计的能见度检测方法。该方法利用公路监控视频,结合大气散射模型,通过获取大气透射系数和场景深度信息计算能见度。为提高检测精度,提出一种编解码深度估计网络,以有效提取图像深度信息,所提网络具有较高的准确性。同时,利用浓雾区域灰度变化较平稳的特性,引入区域熵值法筛选精准的大气透光强度。

此外,引入最大景深区域思想,采用图像分割技术提取可视路面与天空有雾区域的交界区域,关注区域内像素信息,以减小无关特征的干扰。实验结果表明,本文方法相比于传统方法具有较高的准确性,准确率为89.83%,可实现高效、准确的能见度获取,为交通管控提供可靠的数据支撑,提高交通运行效率。未来工作将进一步考虑背景环境的影响,在图像特征提取和背景分割方面进行优化。

参 考 文 献

- [1] 国家质量监督检验检疫总局,中国国家标准化管理委员会.水平能见度等级:GB/T 33673—2017[S].北京:中国标准出版社,2017.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Grade of horizontal visibility: GB/T 33673—2017[S]. Beijing: Standards Press of China, 2017.
- [2] 王玉峰,华灯鑫.激光遥感斜程能见度探测技术研究进展(特邀)[J].光学学报,2024,44(6):0600001.
Wang Y F, Hua D X. Research progress of laser remote sensing

- [3] 马愈昭, 朱峻, 张宇航. 基于大气辐射传输和 δ -二流近似的斜程能见度反演方法[J]. 光学学报, 2023, 43(12): 1201004.
Ma Y Z, Zhu J, Zhang Y H. Slant-range visibility retrieval method based on atmospheric radiative transfer and δ -two-stream approximation[J]. Acta Optica Sinica, 2023, 43(12): 1201004.
- [4] 周涵, 王旭初, 袁越. 基于深度交互性多尺度感受野特征学习的三维点云配准网络[J]. 光学学报, 2024, 44(14): 1415001.
Zhou H, Wang X C, Yuan Y. Three-dimensional point cloud registration network based on deep interactive multi-scale receptive field feature learning[J]. Acta Optica Sinica, 2024, 44(14): 1415001.
- [5] 林曦, 郭阳, 赵永强, 等. 基于邻域像素注意力机制的光场深度估计方法[J]. 光学学报, 2023, 43(21): 2115003.
Lin X, Guo Y, Zhao Y Q, et al. Depth estimation method of light field based on attention mechanism of neighborhood pixel [J]. Acta Optica Sinica, 2023, 43(21): 2115003.
- [6] 杨天麟, 王卫杰, 康楠, 等. 采用改进暗通道先验算法的高速公路能见度检测[J]. 哈尔滨工业大学学报, 2023, 55(3): 100-108.
Yang T L, Wang W J, Kang N, et al. Highway visibility detection using improved dark channel prior algorithm[J]. Journal of Harbin Institute of Technology, 2023, 55(03): 100-108.
- [7] 廖苗, 陆颜, 张锦, 等. 基于雾线和颜色衰减先验的图像去雾方法[J]. 通信学报, 2023, 44(1): 211-222.
Liao M, Lu Y, Zhang J, et al. Image dehazing method based on haze-line and color attenuation prior[J]. Journal on Communications, 2023, 44(1): 211-222.
- [8] 宋海声, 刘宇平, 郑海亮, 等. 基于暗亮原色先验理论的道路能见度检测[J]. 激光与光电子学进展, 2021, 58(6): 0601001.
Song H S, Liu Y P, Zheng H L, et al. Road visibility detection based on prior theory of dark and bright primary colors[J]. Laser & Optoelectronics Progress, 2021, 58(6): 0601001.
- [9] Li Q, Li Y, Xie B. Single image-based scene visibility estimation[J]. IEEE Access, 2019, 7: 24430-24439.
- [10] Outay F, Taha B, Chaabani H, et al. Estimating ambient visibility in the presence of fog: a deep convolutional neural network approach[J]. Personal and Ubiquitous Computing, 2021, 25(1): 51-62.
- [11] 唐绍恩, 李塞, 胡磊, 等. 一种基于迁移学习的能见度检测方法[J]. 计算机工程, 2019, 45(9): 242-247.
Tang S E, Li Q, Hu L, et al. A visibility detection method based on transfer learning[J]. Computer Engineering, 2019, 45(9): 242-247.
- [12] You Y, Lu C W, Wang W M, et al. Relative CNN-RNN: learning relative atmospheric visibility from images[J]. IEEE Transactions on Image Processing, 2019, 28(1): 45-55.
- [13] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention—MICCAI 2015. Lecture notes in computer science. Cham: Springer International Publishing, 2015, 9351: 234-241.
- [14] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22–29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [15] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision—ECCV 2018. Lecture notes in computer science. Cham: Springer International Publishing, 2018, 11211: 833-851.
- [16] Ahmed M, Seraj R, Islam S M S. The k -means algorithm: a comprehensive survey and performance evaluation[J]. Electronics, 2020, 9(8): 1295.
- [17] He K M, Sun J, Tang X O. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [18] 蒲亨飞, 李震, 李良荣. 基于暗通道先验的去雾算法研究[J]. 激光与光电子学进展, 2021, 58(20): 2010011.
Pu H F, Li Z, Li L R. Research of dehazing algorithm based on dark channel prior[J]. Laser & Optoelectronics Progress, 2021, 58(20): 2010011.
- [19] He K M, Sun J, Tang X O. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1397-1409.
- [20] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21–26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [21] 肖磊, 胡鹏, 马俊杰. 局部注意力作用下基于全局信息关联的自监督单目深度估计模型[J/OL]. 激光与光电子学进展: 0815010[2025-01-10]. <https://www.opticsjournal.net/Articles/OJ2a2a073b3f2f9fb4/Abstract>.
Xiao L, Hu P, Ma J J. A self-supervised monocular depth estimation model based on global information correlation under the influence of local attention[J/OL]. Laser & Optoelectronics Progress: 0815010[2025-01-10]. <https://www.opticsjournal.net/Articles/OJ2a2a073b3f2f9fb4/Abstract>.
- [22] Lee J H, Han M K, Ko D W, et al. From big to small: multi-scale local planar guidance for monocular depth estimation [EB/OL]. (2019-07-24)[2024-12-12]. <https://arxiv.org/abs/1907.10326>.
- [23] Song M, Lim S, Kim W. Monocular depth estimation using Laplacian pyramid-based depth residuals[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(11): 4381-4393.
- [24] Godard C, Mac Aodha O, Brostow G J. Unsupervised monocular depth estimation with left-right consistency[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21–26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6602-6611.
- [25] Yan J X, Zhao H, Bu P H, et al. Channel-wise attention-based network for self-supervised monocular depth estimation[C]//2021 International Conference on 3D Vision (3DV), December 1–3, 2021, London, United Kingdom. New York: IEEE Press, 2021: 464-473.
- [26] Zhang N, Nex F, Vosselman G, et al. Lite-mono: a lightweight CNN and transformer architecture for self-supervised monocular depth estimation[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 17–24, 2023, Vancouver, BC, Canada. New York: IEEE Press, 2023: 18537-18546.
- [27] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20–25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [28] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision—ECCV 2018. Lecture notes in computer science. Cham: Springer International Publishing, 2018, 11211: 3-19.
- [29] Mehta S, Rastegari M. Separable self-attention for mobile vision transformers[EB/OL]. (2022-06-06)[2024-12-12]. <https://arxiv.org/abs/2206.02680>.
- [30] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13–19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.

Method for Measuring Visibility on Foggy Highways Based on Depth Estimation of Encoder and Decoder

Peng Peng^{1,3**}, Dong Yucheng¹, Li Jiachun^{2*}, Yao Yitao¹

¹School of Electrical and Control Engineering, Shaanxi University of Science & Technology, Xi'an 710021, Shaanxi, China;

²Highway School, Chang'an University, Xi'an 710064, Shaanxi, China;

³Institute of Flexible Electronics, Northwestern Polytechnical University, Xi'an 710129, Shaanxi, China

Abstract

Objective Visibility refers to the maximum horizontal distance at which an individual with normal vision can identify and distinguish an object against the sky background under prevailing weather conditions without external assistance. It is a critical parameter reflecting atmospheric transparency and serves as a key indicator in the transportation sector. Various factors influence visibility, with fog and haze having the most significant influence. In foggy and hazy weather, the fine particles suspended in the air hinder light transmission and absorb light reflected from object surfaces, significantly reducing visibility. On highways, reduced visibility due to fog and haze is a major cause of traffic accidents, posing severe risks to public safety and economic stability. The accurate and efficient acquisition of visibility data is essential for effective traffic management. Therefore, developing a high-precision visibility detection method that provides reliable data support for transportation authorities has become a key research focus in ground-based meteorology. To address this, we propose an advanced visibility detection method based on a depth estimation network and an atmospheric scattering model.

Methods In this paper, we propose a novel visibility detection method leveraging an encoder-decoder structured depth estimation network. Using highway surveillance videos, the method determines visibility by integrating principles of the atmospheric scattering model to extract atmospheric transmission coefficients and scene depth information. First, the K-means clustering algorithm is applied to segment video frames into foggy and visible road regions, enabling the identification of the maximum scene depth area of the road. Next, the dark channel prior algorithm is enhanced using a regional entropy method to refine the selection of atmospheric light intensity, ensuring a more precise estimation of atmospheric transmission coefficients. Subsequently, an encoder-decoder-based depth estimation network extracts depth information from the images. Finally, visibility is calculated based on the atmospheric scattering physical model.

Results and Discussions The depth information obtained using the DenseNet-169 network is compared with that derived from five other depth estimation networks: BTS, LapDepth, MonoDepth, CADept, and Lite-Mono (Table 1). The results show that DenseNet-169 outperforms these models in terms of root mean square error (RMSE), absolute relative error (A_R), and accuracy. This demonstrates that the depth information generated by DenseNet-169 is highly precise and effectively supports visibility estimation. To further enhance network performance, four attention modules—coordinate attention (CA), convolutional block attention module (CBAM), separable self-attention (SSA), and efficient channel attention (ECA)—are integrated into DenseNet-169 and compared against the original network without attention modules (Table 2). Among these, the integration of the CBAM module achieves the best performance, yielding the lowest RMSE and A_R values while attaining the highest accuracy at a threshold of 1.25². As a result, we adopt the Dense+CBAM encoder-decoder network for depth estimation. Experimental data are collected from surveillance video recorded on November 24, 2024, between 14:25 and 15:15. A sample video frame is shown in Fig. 11, while Table 4 and Fig. 12 present visibility detection results and error variations across different timestamps. The proposed method's visibility estimation results are compared with those of a visibility monitoring device and four alternative networks. The findings indicate that the proposed method closely aligns with the monitoring device, with only minor deviation observed at 15:00. Overall, the proposed method demonstrates high accuracy in both short-range and long-range visibility estimation, achieving an accuracy rate of 89.83% and an average error of approximately 73 m. Furthermore, the method exhibits strong generalization capabilities, highlighting its reliability and precision in visibility measurement.

Conclusions We propose a novel visibility detection method based on an encoder-decoder depth estimation network and an atmospheric scattering model. This method employs the atmospheric scattering physical model as its theoretical foundation to calculate visibility by deriving atmospheric transmittance coefficients and scene depth information from images. To enhance accuracy, a regional entropy method is introduced to refine atmospheric light intensity estimation,

leveraging stable gray-level variations in foggy regions. Image segmentation techniques are applied to locate the boundary between visible road surfaces and foggy sky regions, focusing on pixel information within the target area to minimize the interference from irrelevant features. In the depth estimation network, the encoder-decoder structure of DenseNet-169 is optimized by incorporating the Dense Block-B module, which suppresses redundant features in the input and enhances feature extraction efficiency. In addition, to mitigate interference from the original image background, the CBAM is embedded in the encoder's convolutional module, improving road surface feature extraction while reducing the influence of irrelevant features. Experimental results demonstrate that, compared to existing methods, the proposed approach achieves higher accuracy in both long-range and short-range visibility estimation, with an accuracy rate of 89.83%, reduced overall error, and enhanced generalization capability. This method enables efficient and precise visibility measurement, providing reliable data support for traffic management.

Key words visibility detection; atmospheric scattering model; dark channel prior; encoder-decoder network; image segmentation