

DOI:10.16136/j.joel.2023.07.0363

基于交叉贝叶斯融合全局和局部大气光的深度预测研究

段秀真¹, 夏晨星^{1*}, 罗双强², 葛斌¹, 高修菊³

(1. 安徽理工大学 计算机科学与工程学院, 安徽 淮南 232001; 2. 河南中烟工业有限责任公司安阳卷烟厂, 河南 安阳 455004; 3. 安徽理工大学 电气与信息工程学院, 安徽 淮南 232001)

摘要: 单目深度预测对于三维场景的理解和感知起着至关重要的作用。目前基于深度学习的方法虽然取得很好的效果, 但是其性能过于依赖于训练数据, 同时, 在复杂场景下, 基于深度线索的全局光不变假设效果欠佳。为此, 本文提出了一种基于交叉贝叶斯融合全局和局部大气光的深度预测方法, 具体地, 分别基于全局大气光和设计的局部大气光散射模型进行初始深度图预测, 为实现两者有效互补, 利用研制的交叉贝叶斯模型进行深度图的融合, 并利用边缘滤波机制进行优化得到最终深度图。通过一系列实验, 证明了该模型在深度预测数据集中能够取得很好的效果, 并且有助于提高显著性目标检测任务的准确性。

关键词: 单目深度预测; 全局大气光; 局部大气光; 交叉贝叶斯模型; 边缘滤波

中图分类号: TP391 **文献标识码:** A **文章编号:** 1005-0086(2023)07-0704-09

Cross Bayesian-based fusion of global and local atmospheric light for depth prediction research

DUAN Xiuzhen¹, XIA Chenxing^{1*}, LUO Shuangqiang², GE Bin¹, GAO Xiuju³

(1. College of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, Anhui 232001, China; 2. Anyang Cigarette Factory, China Tobacco Henan Industrial Co., Ltd., Anyang, Henan 455004, China; 3. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan, Anhui 232001, China)

Abstract: Monocular depth prediction plays a crucial role in 3D scenes understanding and perception. At present, deep learning methods have achieved great performance, however, their performance is dependent on training data. In complex scenes, the assumption of global light invariance based on depth cues is not effective. To this end, we propose a depth prediction method based on cross Bayesian fusion of global and local atmospheric light. Specifically, we use the global and the designed local atmospheric light scattering model to predict the initial depth map respectively. And then we construct a cross Bayesian model to fuse depth maps, in order to better realize the complementarity between the two models. Finally, we optimize the depth map by using the edge filtering mechanism. A series of experiments can prove that the model designed can achieve good results in the depth prediction datasets, and help to improve the accuracy of salient object detection task.

Key words: monocular depth prediction; global atmospheric light; local atmospheric light; cross Bayesian model; edge filtering

* E-mail: with_Sandra@163.com

收稿日期: 2022-05-17 修订日期: 2022-07-26

基金项目: 安徽理工大学研究生创新基金(2022CX2117)、国家自然科学基金(62102003)、安徽省自然科学基金(2108085QF258)、安徽省博士后基金(2022B623)和安徽理工大学青年科学研究基金一般项目(xjyb2020-04)资助项目

0 引言

深度预测旨在从一个或多个图像中预测场景的深度信息,是计算机视觉领域经典问题之一,其广泛应用于其他计算机视觉任务,如同步定位和映射^[1]、自动驾驶^[2]、三维重建^[3]、目标检测^[4]等领域^[5]。深度预测任务可以由专用硬件完成,如微软的测距仪和 Kinect 摄像头,但价格昂贵、场景有限,大多数图像都是由普通相机拍摄的,只含场景的颜色信息。为了重建深度信息,研究人员尝试直接从 RGB 图像中预测深度,这种方法与通过专业硬件设备获得深度信息相比,更具有普遍性。

目前的深度预测可以分为双目图像^[6]、多视点图像深度^[7]和单目图像^[8]预测。模仿人类的双目,利用双目视差一致性来预测深度的方法称为基于双目图像的深度预测方法。GODARD 等^[6]提出左右视差的一致性,从而保持单目深度预测的几何一致性。多视点图像深度预测以多个连续的图像作为输入,利用预测的深度、相对姿态和邻近图像重建目标图像。如 KHAN 等^[7]提出了一种可微分渲染和扩散稀疏深度点集的方法,直接优化密集深度图重建,最小化多视图 RGB 重投影损失。然而收集双目立体图像需要特殊的仪器,基于双目与多视角的深度预测主要是采用无监督的方法,而无监督学习的方法目前仍然存在尺度模糊性的问题,在模型性能与深度质量方面与有监督单目深度预测相比仍有差距。单目深度预测旨在从单一 RGB 图像中恢复深度信息,如 BHAT 等^[8]将深度范围划分为多个单元,图像最终深度值估计为单元中心的线性组合。随着深度学习的兴起与发展,及 RGB 图像的广泛可用性,从单目图像推断深度信息引起了越来越多的兴趣。然而,这是一个具有挑战性的不适宜问题,这意味着单张 RGB 图对应的场景可能有无数个。所以,目前单目图像深度预测是深度预测领域中最具挑战性和最重要的部分。

传统的单目图像深度预测方法主要是根据场景中的深度线索来预测深度。常用的单目图像深度线索包括线性透视、聚焦和散焦、阴影、闭塞等^[9-13]。MAXIMOV 等^[10]基于单个图像的焦点信息,提出了一种生成相对深度图的方法,利用图像分割算法将图像划分为统一的区域,将前景背景区域区分进行深度匹配,最后相对深度图被融合得到最终深度图。ZHOU 等^[12]利用光场的多方位对极几何,从光场的中心视图预测深度。基于深度学习的方法^[14-16]通常使用卷积神经网络(convolutional neural network, CNN)来预测场景

的深度信息。EIGEN 等^[14]提出了一个由两个 CNN 堆叠在一起的从粗到细的框架,粗尺度网络提取全局信息,精细尺度网络利用局部信息对粗深度预测结果进行细化。YUAN^[15]等利用全连接条件随机场(conditional random field, CRF)的潜力,结合 Transformer 将输入拆分成窗口,对每个窗口执行全连接 CRF 优化。

上述传统的方法都是基于有额外信息源的场景,而基于深度学习的方法的性能过于依赖训练数据,并且,传统的大气光散射模型在建立时会产生不可避免的误差。为解决上述问题的局限性,本文利用传统的方法避免对训练数据过度依赖,通过在全局大气光的基础上引入局部大气光信息来减少模型构建误差,同时利用两种大气光信息的互补性,构建贝叶斯模型融合全局和局部大气光信息,进一步减少了两种大气光单独预测时产生的构建误差,最后采用边缘保护滤波保留场景细节,进一步提升预测性能。

1 基于交叉贝叶斯的局部与全局混合大气散射深度预测模型

1.1 模型的提出

为减少传统大气光模型产生的构建误差,以及深度学习方法依赖数据集的问题,本文利用贝叶斯模型能更好处理不确定性、计算复杂度不会大幅度上升等优点构建出交叉贝叶斯模型,图 1 显示了本文的整体模型流程图。首先利用全局大气散射模型生成一个深度图,在此基础上引入局部大气光散射模型生成局部大气光深度图。其次,利用交叉贝叶斯模型融合两个模型的结果以获得融合深度图,最后进行融合深度图的优化。

1.2 全局大气光散射模型

清晰的图像是进行深度预测任务的前提,大气中许多混合介质,包括颗粒、水滴等,这些介质会吸收和散射部分可见光,在计算机视觉和图形学中,大气散射模型通常被用来描述薄雾等悬浮颗粒物干扰的图像,在对图像增强和复原的基础上还可以获取一些深度信息。如图 2 所示,全局大气散射模型定义如下:

$$I(x, y) = A \cdot \rho(x, y) \cdot t(x, y) + A \cdot (1 - t(x, y)), \quad (1)$$

式中, I 、 ρ 和 A 分别表示观测图像、场景背景光和大气光, $t(x, y)$ 表示反射光穿透介质能力的透视图。 $A \cdot \rho(x, y)$ 构成大气光的直接衰减项, $A \cdot (1 - t(x, y))$ 构成环境光模型。

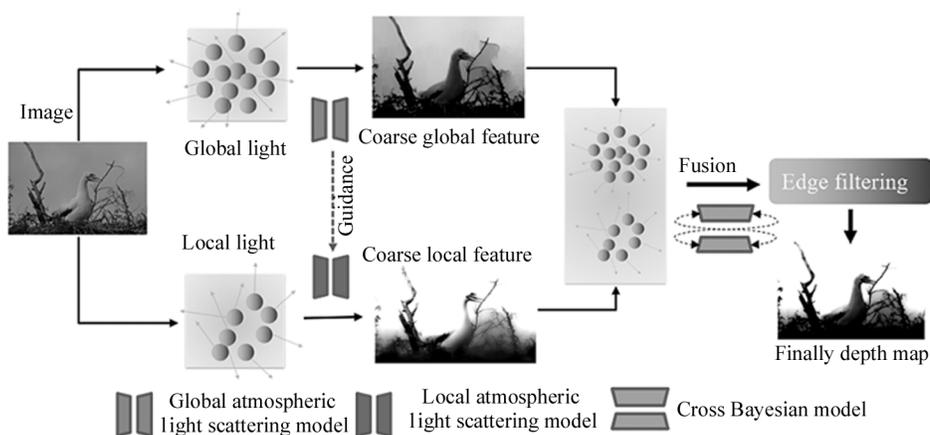


图 1 整体模型流程图

Fig. 1 Overall model flow chart

根据大气散射模型,设备捕获的总光能可分为两部分:场景的辐射光和环境光。在 HE^[17] 理论研究的基础上,通过获取场景的透射图,可以计算场景不同位置的深度信息。传输图可以表示为式(2):

$$t(x) = e^{-\beta t(x)}, \quad (2)$$

式中, β 表示散射系数, $d(x)$ 表示场景深度。根据该模型,可以根据场景的透射图和大气光信息恢复场景的深度信息。在无雾图像中,大多数像素区域中至少有一个颜色通道的强度值等于或接近 0,暗通道的先验值通常用于对图像进行消雾,表示为式(3):

$$I^{\text{dark}}(x, y) = J^{\text{dark}}(x, y)t(x, y) + A(1 - t(x, y)), \quad (3)$$

式中, I^{dark} 和 J^{dark} 分别表示获取图像和理想图像的暗通道。 $t(x, y)$ 通过从式(2)中求解以下方程计算得出:

$$t(x, y) = 1 - \frac{J^{\text{dark}}(x, y)}{A}. \quad (4)$$

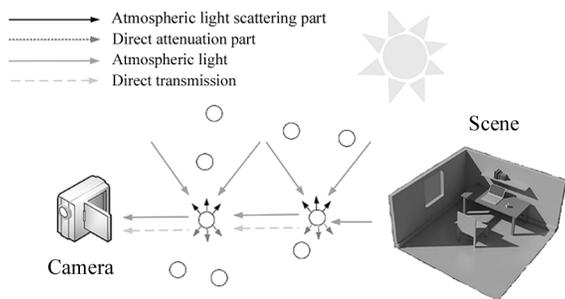


图 2 大气光散射模型

Fig. 2 Atmospheric light scattering model

在上述模型中, A 在单个图像中是常数。在实际场景中,传统的模型在某些特定区域会产生一定的误差。如图 3 所示,图(a)是 RGB 图,图(c)是

RGB 图(a)的深度图,图(b)是深度真值图,在许多区域,深度预测结果与深度真值图不一致。

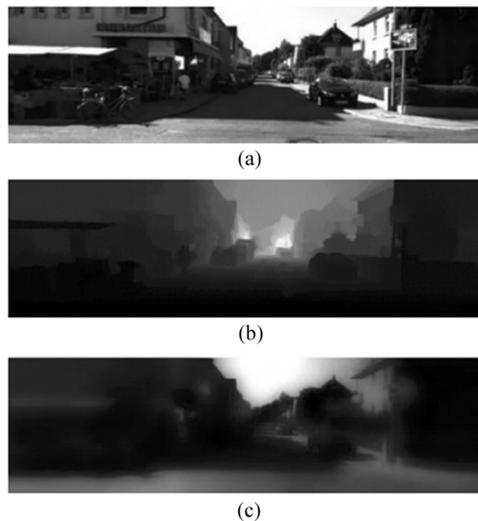


图 3 全球大气散射模型进行深度预测的例子

Fig. 3 An example of depth prediction using the global atmospheric scattering model

在传统的大气散射模型中,假设大气光值固定在某些场景中会产生不可避免的误差,例如图 4 中的阴影部分没有直接暴露在空气中,仅由少量间接辐射照亮,因此,传统模型假设的常数 A 值会在某些区域产生不可避免的误差。为了弥补这个缺点,一些方法假设大气光值是一个变量,但是,由于场景中的大部分悬浮粒子仍然直接被大气光照亮,这种基于像素位置的校正不可避免地会产生误差并且极大增加了计算复杂度。

根据上述分析,本文考虑在全局大气光模型中添加局部大气光信息,利用全局信息和局部信息的互补性,将直接衰减的大气光重新定义为场景入射

光,不同场景有独立的场景入射光,增强模型性能,使用第 1.3 节中介绍的局部大气光模型来模拟大气光。

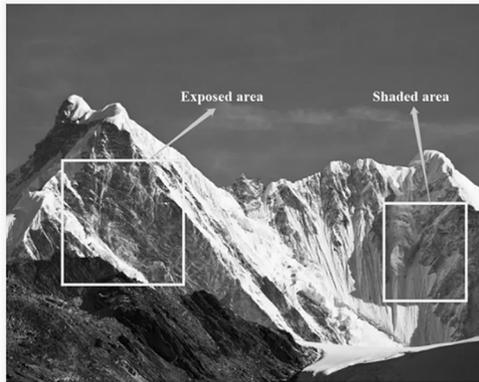


图 4 阴影区域和暴露区域的光照不均匀场景
Fig. 4 Scene with uneven illumination in shaded and exposed areas

1.3 局部大气光散射模型

为解决上述全局大气光模型所产生的校正误差和计算复杂度增大等不足,本文引入了局部大气光散射模型,首先要对场景进行分割,并对每个不同区域的入射光进行自适应评估,同一场景中的像素可能具有相似的深度,根据全局大气散射模型获得的初始深度图对场景进行分割,并在分割后对每个块进行分割。根据上述分析,重新定义的模型可以表示为式(5):

$$I(x, y) = L(i) \cdot \rho(x, y)T(i) + L(i) \cdot (1 - T(i))(x, y), \quad (5)$$

式中, I 、 $L(i)$ 和 $T(i)$ 分别表示 i 个场景、第 i 个场景中的恒定场景亮度和第 i 个场景中的透视率图。将由全局模型生成的原始深度图 D 划分为若干层以保持原始场景的边缘形状。为了避免分割后某些区域的像素聚集,所采用的场景深度层分割方法选择基于像素数的。每一层的长度为 $l = M \cdot N$, 其中 M 和 N 分别表示原始图像的长度和宽度。像素 (x, y) 的层计算如式(6)所示:

$$layer(x, y) = \begin{cases} i, D_r(\max(1, \lceil \frac{i-1}{n} \rceil)) < D_a(x, y) \\ < D_r(\lceil \frac{i}{n} \rceil) \\ 1, D_a(x, y) = D_r(1) \end{cases}, \quad (6)$$

式中, $\lceil \cdot \rceil$ 表示舍入, $D_r(j)$ 表示 D_r 的第 j 个元素。10 层场景分割的结果如图 5 所示,图 5(a)是样本图像,(b)是层数设置为 10 时的分层图图像,(c)和(d)

是原始图像和亮通道图像按照图(b)所示的分层图中的第一层亮通道图。

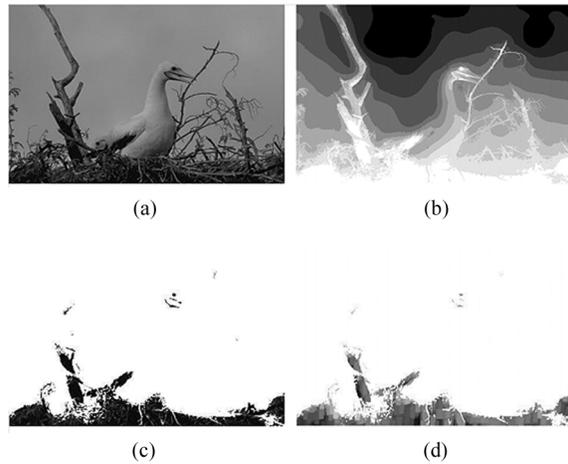


图 5 10 层场景分割的结果
Fig. 5 Results of scene segmentation in 10 layers

明亮通道图的演示如图 6 所示。场景分割后,局部大气光可能太暗,如果用传统的全局大气光获取方法来获取局部大气光,由于局部色差,其值可能低于真实值,所以本文的方法在局部大气光散射模型中引入了亮通道来获取局部大气光。通过评估原始输入图像的明亮通道信息的局部大气光,可以有效地减少由色差引起的局部大气光误差。参考 WANG^[18] 对大量的无霾图像进行的统计,即每个区域中都有一个像素,其 RGB 三个颜色通道中的一个通道值接近或等于 1。图像的明亮通道图可以定义为:

$$P^{\text{bright}} = \max_{c \in R, G, B} (\max_{(x', y') \in \omega(x, y)} (P^c(x', y'))) \rightarrow 1. \quad (7)$$



图 6 样本图和样本亮通道图

Fig. 6 Sample diagram and sample bright channel diagram

在这个模型中,本文对初始深度图的深度值进行分割,并对每个单独场景的明亮通道中相应像素亮度的前 10% 进行统计,取平均值作为单独场景的大气光值,建立局部区域的大气散射模型,通过式(5)得到不同区域的 $T(i)$ 。最后,将这些区域合并为局部大气散射模型。

局部大气光散射模型对不同区域采用了自适应的大气光策略,这对于减少因光线分布不均而产生的透视误差非常有效,与全局大气光策略相比,局部大气光散射模型可以更好地模拟复杂场景。实验发现全局大气光散射模型在平均透射率预测方面取得了更好的效果,而局部大气光散射模型在复杂场景中不同深度区域的透射率预测方面取得了更好的效果。基于以上分析,本文采用了基于交叉贝叶斯模型的融合策略,对两个模型产生的结果进行了融合,互充了两个模型的优势。

1.4 基于交叉贝叶斯的融合模型

近几年的研究中,贝叶斯公式(Bayes)被用来通过后验概率来衡量突出性,贝叶斯模型发源于古典数学理论,有稳定的分类效率,能个处理多分类任务,适合增量式训练,且对缺失数据不太敏感,算法也比较简单,本文利用贝叶斯模型的这种特性来解决全局大气光模型和局部大气光模型的融合问题,提出的深度图交叉贝叶斯模型如图7所示。Bayes公式具体如下:

$$p(F|H(z)) = \frac{p(F)P^F}{p(F)p(PF) + (1-p(F))P^B}, \quad (8)$$

式中, $p(F)$ 是先验概率, $H(z)$ 是像素 z 的特征向量, P^F 和 P^B 分别表示 $p(H(z)|F)$ 和 $p(H(z)|B)$ 。通过使用贝叶斯模型,将上述步骤得到的全局大气光模型图像和局部大气光模型图像融合成融合深度图。

给定两个深度图 G 和 L 为 D_1 和 D_2 ,取其中一个作为先验的 D_i ,用另一个 D_j 来计算似然。首先使用 D_i 的平均值将 D_i 划分为前景区和背景区,分别为 F_i 和 B_i 。在每个区域,通过比较 D_i 和 D_j 在像素 z 处的前景和背景区来计算似然性,似然性概率的计算方法是:

$$p(D_j|F_i) = \frac{N_{b_{F_i}(D_j(z))}}{N_{F_i}}, p(D_j|B_i) = \frac{N_{b_{B_i}(D_j(z))}}{N_{B_i}}, \quad (9)$$

式中, N_F 表示前景像素数, N_B 为背景像素数, $N_{b_{F_i}(D_j(z))}$ 为深度特征落入前景区 $b(D(z))$ 的像素数,其中包含特征 $D(z)$,背景的深度分布直方图也由 N_B 和 $N_{b_B(D(z))}$ 代表。

后验概率的计算是以 D_i 作为先验由以下式计算:

$$p(F_i|D_j(z)) = \frac{D_i(z)p'}{D_i(z)p' + (1-D_i(z))p'}, \quad (10)$$

式中, p' 表示 $p(D_j(z)|B_i)$ 。

同样,也可以计算 D_j 作为先验的后验概率。使用这两个后验概率来计算融合深度图: $D_B(D_1(z), D_2(z))$ 基于贝叶斯的融合:

$$D_B(D_1(z), D_2(z)) = p(F_1|D_2(z)) + p(F_2|D_1(z)), \quad (11)$$

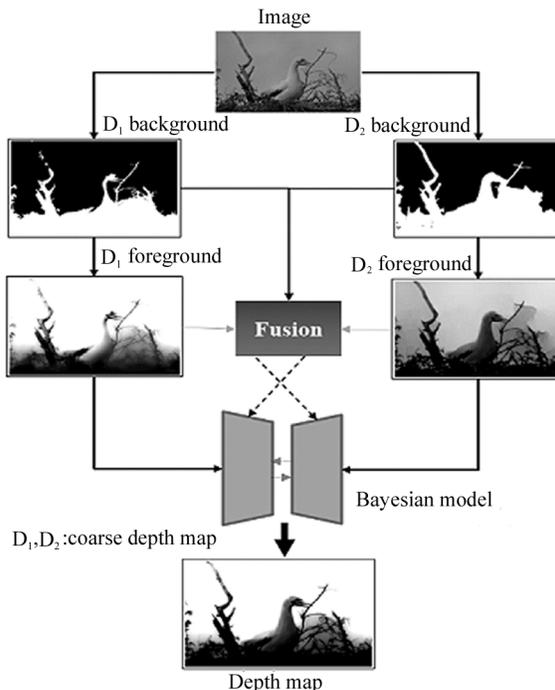


图7 深度图的贝叶斯融合模型

Fig.7 Bayesian fusion model of depth map

2 实验结果及分析

2.1 数据集

提出的算法在KITTI^[19]、SOD^[20]、ECSSD^[21]和ASD^[22]数据集中进行了测试。KITTI数据集是一个用于深度预测的大规模公共数据集,用来完成深度预测任务,其余数据集用来验证本文方法对显著性目标检测任务准确性的提升。其中SOD数据集有300张图像,是一组突出的物体边界,它有许多带有多个物体的图像。ECSSD^[21]数据集有1000张复杂场景图像,一部分来自Berkeley-300数据集。ASD数据集有1000张图像,其图像被手动分割成前景和背景区域。

2.2 评价指标

本文采用EIGEN^[13]等引入的误差和精度度量来验证本文所提出的方法的性能,并与其他单目深度预测模型进行了比较。误差度量中的绝对相对误差(Abs Rel)、平方相对误差(Sq Rel)、均方根误差(RMSE)表述如下:

$$Abs\ Rel = \frac{1}{T} \sum_{i=1}^T \frac{\|d_i^* - d_i\|}{d_i}, \quad (12)$$

$$Sq\ Rel = \frac{1}{T} \sum_{i=1}^T \frac{\|d_i^* - d_i\|^2}{d_i}, \quad (13)$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (d_i^* - d_i)^2}, \quad (14)$$

式中, d_i^* 和 d_i 分别为 i 像素的预测深度和真实深度, T 为像素总数。

2.3 定量和定性结果比较

本文的方法与其他几种方法的比较结果见表1, 其中 \downarrow 表示数值越低越好, 加黑字体表示最佳性能, 划线字体表示第二性能。结果显示, 本文的方法在各种评估中都优于其他基于深度学习的方法, 而且与其他方法相比, 本文的方法需要的运行时间更少, 且不需要训练样本。

表1 KITTI数据集的定量结果
Tab.1 Quantitative data set of KITTI

Method	$Abs\ Rel \downarrow$	$Sq\ Rel \downarrow$	$RMSE \downarrow$
EIGEN ^[13]	0.203	1.548	6.307
LIU ^[23]	<u>0.202</u>	1.614	6.523
ZHOU ^[24]	0.208	0.768	6.856
Make3D ^[25]	0.280	1.713	8.734
GARG ^[26]	0.203	1.955	<u>5.597</u>
Ours	0.193	<u>1.465</u>	5.595

为了测试和验证深度图的质量, 本文比较了不同深度图的主观效果。将本文的方法与基于深度学习的 EIGEN^[13]、LIU^[23]、ZHOU^[24]、Make3D^[25] 和 GARG^[26] 的方法进行比较, 与上述方法相比, 本文提出的方法所生成的深度图的分辨率更高, 如图8中框内的远程场景, 从左到右分别为输入图像、真值图、EIGEN 等的方法结果图和本文的结果。本文的方法保留了更丰富的边缘信息。本文方法的深度图像质量更高, 对细节和场景结构的表現更细致, 对不

同物体的分辨能力更强, 这主要得益于大气深度线索和边缘优化策略。

2.4 在视觉显著性检测中的应用

2.4.1 使用伪深度先验的显著性检测

深度信息已被证明是对视觉显著性检测的有用信息。然而, 现有的 RGB-D 显著性检测方法主要集中在简单和综合模型的设计上。在现有的 RGB-D 显著性检测方法中, 有两种主要方式来使用深度信息。一种是将深度信息直接明确地集成到特征库中, 作为颜色特征的补充; 另一种是通过设计的深度描述服从深度图中捕捉隐含属性, 从而进一步进行实验, 研究深度信息如何影响视觉显著性检测任务。

2.4.2 显著性检测的评价指标

本文的方法获得的深度信息被用来替换原始深度图, 并检测 PDP 中的显著性。该方法与一些先进的方法在 ASD、MSRA、SOD 和 ECSSD 数据集上的比较结果见表2。结果表明, 本文的方法在各种评价中都优于其他方法。与原始 PDP 算法的结果相比, 准确的深度信息在使用伪深度先验的显著性检测中起着关键作用。模型中使用的深度图越准确, 使用伪深度先验的显著性检测的改进就越明显。

表2 在 AUC 评价指标下的定量结果

Tab.2 Quantitative results under AUC evaluation index

Dataset	Ours	RC	LR	MR	HR
ASD	0.9752	<u>0.9435</u>	0.9575	0.9632	0.9639
MSRA	0.9368	0.8823	0.9121	0.9261	<u>0.9280</u>
SOD	<u>0.7594</u>	0.7870	0.7432	0.7534	0.8085
ECSSD	0.9009	—	<u>0.8619</u>	0.8827	0.8840

表2中, 加粗字体表示最佳性能, 划线字体表示第二性能。使用 ROC 曲线下的面积(AUC)得分评价标准来评估显著性检测方法的性能, AUC 分数越高, 说明性能越好。

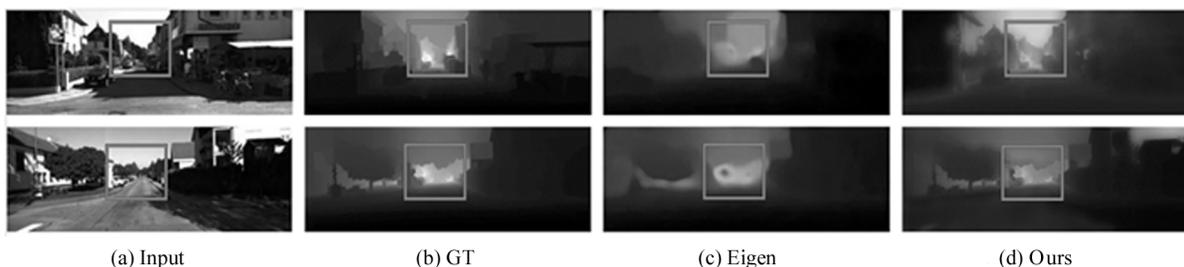


图8 KITTI数据集的定性结果

Fig.8 Qualitative results of KITTI dataset

本文的方法、MILPS 和 PDP 的定性结果显示在图 9 中。从左到右为:参考图像、实际真值图、MILP 结果图、PDP 结果图和本文的结果图。从图中可以看出,本文的方法表现更好,能更完整地保留了重要目

标的边缘信息。

图 10 显示了本方法与一些先进方法在 ASD 数据集上的 PR 曲线。结果表明,在视觉显著性检测中使用深度信息可以有效提高视觉显著性检测的

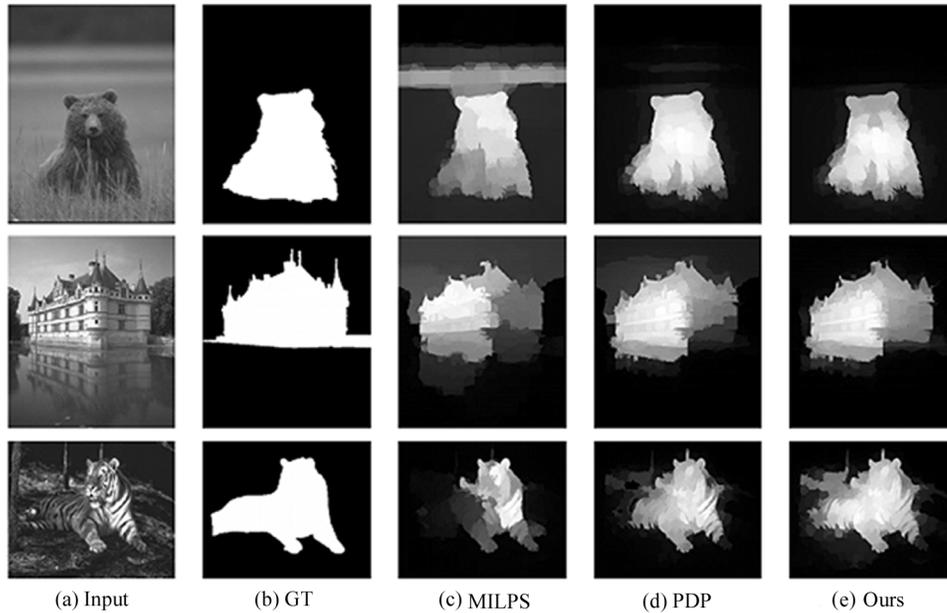


图 9 SOD 数据集的定性结果

Fig. 9 Qualitative results of SOD dataset

效果。

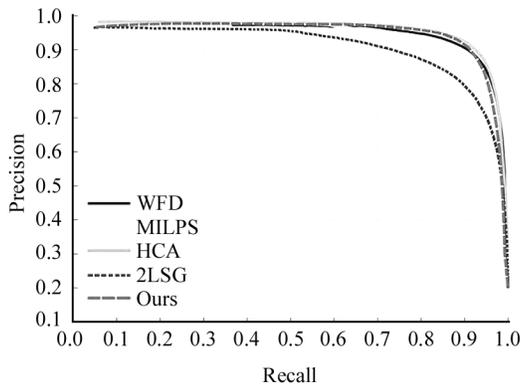


图 10 在 ASD 数据集上的 PR 曲线比较

Fig. 10 Comparison of precision-recall curves on ASD dataset

3 结 论

本文提出了一种基于大气光散射模型的深度预测方法。首先,利用全局大气光散射模型进行初始深度预测;其次,利用局部大气光模型进行局部深度预测;最后,全局大气光模型结果与基于贝叶斯的局部大气光模型结果相融合,并使用边缘保护滤波器

进行边缘优化以获得最准确的深度。最终的深度预测图在 KITTI 数据集上进行了测试,与现有的基于深度学习的方法相比,提出的方法取得了更准确的结果。此外,在 ECSSD、ASD 和 SOD 数据集上的实验结果表明,本文的方法所获得的深度信息能够有效地提高视觉显著性检测的准确性。所提出的方法对于有明显大气光散射现象的自然光源场景效果较好,对于复杂光源场景(如室内)场景有局限性。今后的工作重点是研究复杂光源场景的深度预测,采用图神经网络等方法,引入更多深度线索进行推理,重建场景深度。

参考文献:

[1] TIWARI L,JI P,TRAN Q H.et al. Pseudo RGB-D for self-improving monocular SLAM and depth prediction[C]// European Conference on Computer Vision, August 23-28, Virtual Event. Berlin, Heidelberg: Springer-Verlag, 2020: 437-455.

[2] YUAN Z, SONG X, BAI L, et al. Temporal-channel transformer for 3D lidar-based video object detection for autonomous driving[J]. IEEE Transactions on Circuits and

- Systems for Video Technology, 2021, 32(4):2068-2078.
- [3] LIU H, TANG X, SHEN S. Depth-map completion for large indoor scene reconstruction [J]. Pattern Recognition, 2020, 99:107112.
- [4] ZHANG S, WEN L, LEI Z, et al. RefineDet ++: Single-shot refinement neural network for object detection [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(2):674-687.
- [5] XU G Y, TANG W J. An image semantic segmentation method effectively fusing multi-scale features [J]. Journal of Optoelectronics • Laser, 2022, 33(3):264-271.
许光宇, 汤伟建. 一种有效融合多尺度特征的图像语义分割方法 [J]. 光电子 • 激光, 2022, 33(3):264-271.
- [6] GODARD C, AODHA O M, BROSTOW G J. Unsupervised monocular depth prediction with left-right consistency [C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017:270-279.
- [7] KHAN N, KIM M H, TOMPKIN J. Differentiable diffusion for dense depth estimation from multi-view images [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 19-25, 2021, Virtual Event. New York: IEEE, 2021:8912-8921.
- [8] BHAT S F, ALHASHIM I, WONKA P. Adabins: Depth estimation using adaptive bins [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 19-25, 2021, Virtual Event. New York: IEEE, 2021:4009-4018.
- [9] GUR S, WOLF L. Single image depth estimation trained via depth from defocus cues [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 16-20, 2019, Long Beach, USA. New York: IEEE, 2019:7683-7692.
- [10] MAXIMOV M, GALIM K, LEAL-TAIXE L. Focus on defocus: bridging the synthetic to real domain gap for depth estimation [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020, Seattle, USA. New York: IEEE, 2020:1071-1080.
- [11] MING A, WU T, MA J, et al. Monocular depth-ordering reasoning with occlusion edge detection and couple layers inference [J]. IEEE Intelligent Systems, 2015, 31(2):54-65.
- [12] ZHOU W, ZHOU E, LIU G, et al. Unsupervised monocular depth estimation from light field image [J]. IEEE Transactions on Image Processing, 2019, 29:1606-1617.
- [13] GUO F, TANG J, PENG H. Adaptive estimation of depth map for two-dimensional to three-dimensional stereoscopic conversion [J]. Optical Review, 2014, 21(1):60-73.
- [14] EIGEN D, PUHRSCH C, FERGUS R. Depth map prediction from a single image using a multi-scale deep network [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal, Canada. Cambridge, MA, United States: MIT Press, 2014:2366-2374.
- [15] YUAN W, GU X, DAI Z, et al. NeW CRFs: neural window fully-connected CRFs for monocular depth estimation [EB-OL]. (2022-03-03) [2022-05-17]. <https://arxiv.org/abs/2203.01502>.
- [16] ANISIMOVSKIY V, SHCHERBININ A, TURKO S, et al. Unsupervised monocular depth estimation CNN robust to training data diversity [C]//Canadian Conference on Artificial Intelligence, May 13-15, 2020, Ottawa, Canada. Berlin, Heidelberg: Springer-Verlag, 2020:36-48.
- [17] HE K, SUN J, TANG X. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(12):2341-2353.
- [18] WANG Y, ZHUO S, TAO D, et al. Automatic local exposure correction using bright channel prior for under-exposed images [J]. Signal Processing, 2013, 93(11):3227-3238.
- [19] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The KITTI dataset [J]. The International Journal of Robotics Research, 2013, 32(11):1231-1237.
- [20] MOVAHEDI V, ELDER J H. Design and perceptual validation of performance measures for salient object segmentation [C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, USA. New York: IEEE, 2010:49-56.
- [21] YANG C, ZHANG L, LU H, et al. Saliency detection via graph-based manifold ranking [C]//IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2010, San Francisco, USA. New York: IEEE, 2010:3166-3173.
- [22] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection [C]//2009 IEEE Confer-

- ence on Computer Vision and Pattern Recognition, June 20-29, 2009, Sparkle East, USA. New York: IEEE, 2009: 1597-1604.
- [23] LIU F, SHEN C, LIN G. Deep convolutional neural fields for depth estimation from a single image[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 5162-5170.
- [24] ZHOU T, BROWN M, SNAVELY N, et al. Unsupervised learning of depth and ego-motion from video[C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1851-1858.
- [25] SAXENA A, SUN M, NG A Y. Make3D: Learning 3D scene structure from a single still image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(5): 824-840.
- [26] GARG R, BG V K, CARNEIRO G, et al. Unsupervised cnn for single view depth estimation: Geometry to the rescue [C]//European Conference on Computer Vision, October 11-14, 2016, Amsterdam, The Netherlands. Berlin, Heidelberg: Springer-Verlag, 2016: 740-756.

作者简介:

夏晨星 (1991—), 男, 博士, 副教授, 硕士生导师, 主要从事图像与信息处理方面的研究。