

基于改进立体网络的深度估计

徐万朋¹, 邹玲³, 吴玲达¹, 齐越², 钱昭勇¹

¹航天工程大学复杂电子系统仿真科学与技术实验室, 中国北京市, 101416

²鹏城实验室, 中国深圳市, 518055

³北京电影学院数字媒体学院, 中国北京市, 100088

摘要: 自监督深度估计方法通过在训练数据中利用目标图像和参考图像之间的视角合成, 呈现了可以与全监督方法相媲美的结果。然而, 作为主干网络的ResNet最初是为了应对分类问题而设计的, 在应用于下游领域时存在一些结构上的缺陷。图像中的低纹理区域也使深度估计的效果受到很大影响。为了解决这些问题, 本文提出一系列改进, 以实现更加有效的深度预测。首先, 我们通过改进网络结构来促进网络中的信息流通, 并提高学习空间结构的能力。其次, 使用二值蒙版去除目标图像和参考图像之间低纹理区域中的像素, 以更准确地重建图像。最后, 我们随机输入目标图像和参考图像对数据集进行扩充, 并在ImageNet上进行预训练, 从而使模型获得良好的通用特征表示。我们使用立体图像对作为输入, 在KITTI自动驾驶数据集的特征分割上验证了最先进的性能。

关键词: 单目深度估计; 自监督; 图像重建

<https://doi.org/10.1631/FITEE.2000676>