






刘大壮_

关注

11



64

3



专栏目录

一文带你读懂深度估计算法

原创 刘大壮_ 于 2021-11-20 15:35:58 发布 阅读量8.5k 收藏 64 点赞数 11

版权

分类专栏： 深度学习 文章标签： 算法 计算机视觉 深度学习



深度学习 专栏收录该内容

0 订阅 4 篇文章

订阅专栏

深度估计算法调研

1.常见的深度估计算法

寻找RGB图像与深度图之间存在着的某种映射关系

1) 几种常见算法：

- ① 基于几何的方法：从一系列二维图像序列中估计三维结构，SfM，稀疏特征的深度由SfM通过图像序列之间的特征对应和几何约束来计算。存在单目尺度模糊的问题，依赖于高精度的图像匹配或高质量的图像序列。
- ② 基于传感器的方法：利用深度传感器，可以直接获得相应图像的深度信息，测量范围有限，常见 Kinect 的有效测距范围仅为0.8m~4m。
- ③ 基于深度学习 的方法：CNN RNN VAE GAN

2) 常用数据集：

KITTI, NYU Depth V2, Make3D

3) 深度估计中一些评价指标：

- $RMSE = \sqrt{\frac{1}{|N|} \sum_{i \in N} \|d_i - d_i^*\|^2}$,
- $RMSE \log = \sqrt{\frac{1}{|N|} \sum_{i \in N} \|\log(d_i) - \log(d_i^*)\|^2}$,
- $Abs \text{ Rel} = \frac{1}{|N|} \sum_{i \in N} \frac{|d_i - d_i^*|}{d_i^*}$,
- $Sq \text{ Rel} = \frac{1}{|N|} \sum_{i \in N} \frac{\|d_i - d_i^*\|^2}{d_i^*}$,
- **Accuracies:** % of d_i s.t. $\max(\frac{d_i}{d_i^*}, \frac{d_i^*}{d_i}) = \delta < thr$,

4) 基于深度学习的单目深度估计：

① 监督或半监督学习：

CNN：

1.使用两个尺度的CNN对单张图片的深度进行估计：粗尺度网络预测图片的全局深度，细尺度网络优化局部细节。随后又提出了使用了VGG网络，利用第3个细尺度的网络进一步增添细节信息。

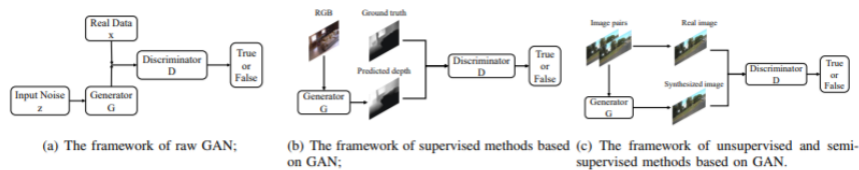
2.还有一些编码器解码器结构，其中比较有名的：

FastDepth: Fast Monocular Depth Estimation on Embedded Systems: encoder部分MobileNet 模型提取到7x7x1024的特征；Decoder部分采用了5次上采样，中间三次上采样结果通过Skip Connections的方法分别与Encoder部分的特征进行了特征融合

基于条件随即场：

Liu等将深度卷积神经网络 与连续条件随机场结合，提出深度卷积神经网络在统一的神经网络框架中学习连续CRF的一元势能项和成对势能项

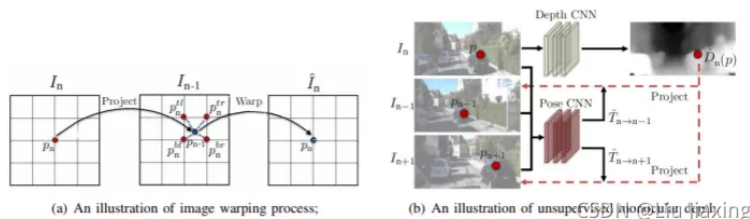
基于GAN：



② 无监督学习：

将帧间的几何约束作为监督信号，而不是用真实背景制作为标签

无监督方法的基本模型：无监督方法由单眼图像序列训练，几何约束建立在相邻帧之间的投影上



可传输性（泛化性）：可传输性是指同一网络在不同摄像机、不同场景和不同数据集上的性能
实时性：减少计算时间，提高运行速度

2.深度图转换成实际距离

1) 绝对深度：利用激光雷达等设备获取的相机与物体之间的距离

如KITTI深度图是利用360度雷达设备获取稀疏的深度图，是以uint16的格式存储的，要从深度图中读取真实距离，除以256之后以米为单位

将Kinect2原始深度数据转换为以米为单位的距离。Kinect实际上通过USB发送类似于视差图像的内容。OpenNI和libfreenect都能够使用设备报告参数（基线，焦距和到参考平面的距离IIRC）将其转换为深度图像。深度=基线*：



刘大壮_

关注

11

4

64

3



专栏目录



Beta



举报

2) 相对深度:

用来获得ground truth depth的硬件设备具有一定误差和环境限制，因此获得的深度图具有一定误差。在误差的基础上进行建模，使得后续模型拟合的难度加大了。因此提出，其实人类对深度的实际数值并不敏感，反而是对相对深度即物体的前后关系更加敏感。不能准确标注出多少米，其与真实深度之间差了一个基准值，这个值我们是不知道的

3.深度估计算法流程

1) 有监督学习:

① 训练集RGBD格式

RGB+Depth Map, Depth Map 类似于灰度图像, 只是它的每个像素值是距离物体的实际距离

groundtruth:常用的KITTI数据集和NYU-V2数据集中都主要包含两部分:

1. label datasets : 就是类别+相关信息
2. raw datasets: 产生RGB和depth图像

网络输入是raw dataset的RGB图像，输出是深度图，与depth图像作距离损失优化网络参数。训练好后，产生的深度图（灰度）中像素点的值体现视点距离

② 深度信息格式

许多研究图像深度的论文里有提到用kitti数据集中LIDAR scanner得到的深度信息作为ground-truth，其中matrix每行有x, y, z, intensity，意思是前后距离（深度），左右偏移，上下偏移，激光反射强度。

③ 格式RGB+相对深度

Single-Image Depth Perception in the Wild (经典论文):

DIW数据集：户外图像和相对深度注释。该数据集每幅图片选取一对像素点进行一次询问，标注。每张图片上询问一对采样点能最大限度利用采样点的相对深度信息。在一张图中询问多对采样点不如在多张图中每张图询问一对采样点。采样点的选择：采样的两点关于中心对称，这样实际概率上各占50%。

网络结构：网络是image2depth的

相对距离构成的损失函数：对dpeth图中at标注两个点之间根据注释rk算loss

$$L(I, R, z) = \sum_{k=1}^K \psi_k(I, i_k, j_k, r, z),$$

$$\psi_k(I, i_k, j_k, z) = \begin{cases} \log(1 + \exp(-z_{i_k} + z_{j_k})), & r_k = +1 \\ \log(1 + \exp(z_{i_k} - z_{j_k})), & r_k = -1 \\ (z_{i_k} - z_{j_k})^2, & r_k = 0. \end{cases}$$

2) 无监督学习:

数据集采用stereo images，两个相机（或者双目相机）在同一水平线上左右相距一定位置得到的两幅图片。这种图片获取的代价要比深度的ground truth 低一些，利用了深度信息和场景之间的一些物理规律来约束。

TABLE I
A SUMMARY OF DEEP LEARNING-BASED MONOCULAR DEPTH ESTIMATION. "MONO." REFERS TO "MONOCULAR", AND "MULTI-TASKS" MEANS THAT IN ADDITION TO POSE AND DEPTH ESTIMATION, THERE ARE OTHER TASKS THAT ARE JOINTLY TRAINED IN THE FRAMEWORK, SUCH AS SEMANTIC SEGMENTATION, MOTION SEGMENTATION, OPTICAL FLOW, ETC.

Methods	Years	Training set	Supervised (Sup) manner			Main contributions
			Sup	Semi-sup	Unsup	
Eigen <i>et al.</i> [35]	2014	RGB + Depth	✓			CNNs
Li <i>et al.</i> [66]	2015	RGB + Depth	✓			hierarchical CRFs
Liu <i>et al.</i> [69]	2015	RGB + Depth	✓			continuous CRF
Wang <i>et al.</i> [70]	2015	RGB + Depth	✓			Semantic labels, hierarchical CRFs
Shelhamer <i>et al.</i> [52]	2015	RGB + Depth	✓			Fully CNNs
Eigen <i>et al.</i> [51]	2015	RGB + Depth	✓			Multi-task
Szegedy <i>et al.</i> [59]	2015	RGB + Depth	✓			Inception Module
Mousavian <i>et al.</i> [71]	2016	RGB + Depth	✓			Multi-task
Roy <i>et al.</i> [73]	2016	RGB + Depth	✓			RFs
Mayer <i>et al.</i> [33]	2016	RGB + Disparity	✓			Multi-task
Laina <i>et al.</i> [54]	2016	RGB + Depth	✓			Residual learning
Jung <i>et al.</i> [63]	2017	RGB + Depth	✓			Adversarial learning
Kendall <i>et al.</i> [49]	2017	Stereo images + Disparity	✓			Disparity Loss
Zhang <i>et al.</i> [56]	2018	RGB + Depth	✓			Task-attentional, BerHu loss
Xu <i>et al.</i> [72]	2018	RGB + Depth	✓			Continuous CRF, structured attention
Gwn <i>et al.</i> [79]	2018	RGB + Depth	✓			Conditional GAN
Fu <i>et al.</i> [60]	2018	RGB + Depth	✓			Ordinal regression
Facil <i>et al.</i> [27]	2019	RGB + Depth	✓			Camera model
Wolt <i>et al.</i> [61]	2019	RGB + Depth	✓			Lightweight network
Garg <i>et al.</i> [28]	2016	Stereo images		✓		Stereo framework
Chen <i>et al.</i> [58]	2016	RGB + Relative depth annotations		✓		The wild scene
Gordard <i>et al.</i> [153]	2017	Stereo images		✓		Left-right consistency loss
Kaznietsov <i>et al.</i> [48]	2017	Stereo images + LiDAR		✓		Direct image alignment loss
Poggi <i>et al.</i> [96]	2018	Stereo images		✓		Trinocular assumption
Ramirez <i>et al.</i> [97]	2018	Stereo images + Semantic Label		✓		Semantic prediction
Aleotti <i>et al.</i> [31]	2018	Stereo images		✓		GAN
Pilzer <i>et al.</i> [102]	2018	Stereo images		✓		Cycled generative network
Luo <i>et al.</i> [99]	2018	Stereo images		✓		Stereo matching, view prediction
He <i>et al.</i> [106]	2018	Stereo images + LiDAR		✓		Weak-supervised framework
Pilzer <i>et al.</i> [103]	2019	Stereo images		✓		Knowledge distillation
Tosi <i>et al.</i> [101]	2019	Stereo images		✓		Stereo matching
Chen <i>et al.</i> [98]	2019	Stereo images		✓		Multi-task
Fei <i>et al.</i> [17]	2019	Stereo images + IMU + Semantic Label		✓		Multi-task, physical information
Feng <i>et al.</i> [65]	2019	Stereo images		✓		Stacked-GAN
Wang <i>et al.</i> [16]	2018	Mono. sequences		✓		Direct VO
Zhan <i>et al.</i> [14]	2018	Stereo sequences		✓		Deep feature reconstruction
Li <i>et al.</i> [112]	2018	Stereo sequences		✓		Absolute scale recovery
Wang <i>et al.</i> [113]	2019	Stereo sequences		✓		Multi-task
Zhao <i>et al.</i> [104]	2019	Stereo images + Synthesized GT		✓		Domain adaptation, cycle GAN
Wu <i>et al.</i> [105]	2019	Mono. sequences + LiDAR		✓		Attention mechanism, GAN
Zhou <i>et al.</i> [43]	2017	Mono. sequences			✓	Monocular framework, mask network
Vijayanarasimhan <i>et al.</i> [85]	2017	Mono. sequences			✓	Multi-task
Yang <i>et al.</i> [86]	2017	Mono. sequences			✓	Surface normal
Mahjourian <i>et al.</i> [50]	2018	Mono. sequences			✓	ICP loss
Yin <i>et al.</i> [15]	2018	Mono. sequences			✓	Multi-task
Zou <i>et al.</i> [89]	2018	Mono. sequences			✓	Multi-task
Kumar <i>et al.</i> [64]	2018	Mono. sequences			✓	GAN
Sun <i>et al.</i> [88]	2019	Mono. sequences			✓	Cycle-consistent loss
Wang <i>et al.</i> [87]	2019	Mono. sequences			✓	Geometry mask
Bian <i>et al.</i> [44]	2019	Mono. sequences			✓	Scale-consistency
Casser <i>et al.</i> [82]	2019	Mono. sequences			✓	Object motion prediction
Ranjian <i>et al.</i> [90]	2019	Mono. sequences			✓	Multi-task
Chen <i>et al.</i> [91]	2019	Mono. sequences			✓	Camera intrinsic prediction
Gordon <i>et al.</i> [92]	2019	Mono. sequences			✓	Camera intrinsic prediction
Li <i>et al.</i> [93]	2019	Mono. sequences			✓	GAN, LSTM, mask
Almaloglu <i>et al.</i> [94]	2019	Mono. sequences			✓	CCGAN, LSTM

CSDN @ Liu Jiaxing

文章知识点与官方知识档案匹配，可进一步学习相关知识

算法技能树 首页 概览 64980 人正在系统学习中

深度估计算法原理与论文解读



刘大壮_

关注

11

☆ 64

3

专栏目录

52053775的博客

3072

深度估计，即通过输入的彩色图像，获得每个像素点离相机距离的远近（热度图），热度图的深浅表示距离的远近。可以应用于AR、VR以及自动驾驶/辅助驾驶任务中，通常是作为上游任务，与通过传感...

一文教你**读懂**3D目标检测

Z's Palace 2万+

本文共4593字，预计需要20分钟，可以先收藏再看哦！简介目标检测是**计算机视觉**领域的传统任务，与图像识别不同，目标检测不仅需要识别出图像上存在的物体，给出对应的类别，还需要将该物体的位...

3条评论



我是一个对称矩阵 **热评** 博主您好，我最近在看深度估计的论文，现在有个疑惑想请问下：我看到在一个深度估计模型的输出中，深度值范围在0~NaN（最大...

写评论

单目测距——基于YOLOv5的单目**深度估计算法**的实现_单目**深度估计**测距-C...

11-5

基于YOLOv5实现的**深度估计**示例: 结果示例 一、YOLOv5的简单介绍: YOLOv5是一个在COCO数据集上预训练的物体检测架构和模型系列,由Ultralytics团队对未来视觉AI方法的开源研究。YOLOv5是YOL...

基于OpenCV的双目**深度估计**实现与改进_双目 **深度** opencv

11-5

双目**深度估计** 一、传统方法 常用的方法有SAD匹配**算法**,BM**算法**,SGBM**算法**,GC**算法** 1.1、**SAD算法** SAD(Sum of absolute differences)是一种图像匹配**算法**,基本思想是:差的绝对值之和。此**算法**常用于...

计算机视觉算法中的 **深度估计** (Depth Estimation)

牛肉胡辣汤 1524

需要注意的是，这个示例只是一个简单的演示代码，实际的**深度估计算法**可能需要更复杂的网络结构和训练过程。在传统的双目视觉系统中，通过计算两个摄像头之间的视差（disparity），可以推断出物体...

10 种机器学习**算法**的要点（附 Python 和 R 代码）

onlyForCloud的专栏 5515

http://blog.jobbole.com/92021/ 本文由 伯乐在线 - Agatha 翻译，唐尤华 校稿。未经许可，禁止转载！英文出处：SUNIL RAY。欢迎加入翻译组。前言 谷歌董事长施密特曾说过：虽然谷歌的无人驾驶汽车...

十种**深度学习算法**要点及代码解析_深度学习实践运用

10-28

Gradient Boost 和 Adaboost **算法** 1、线性回归 线性回归通常用于根据连续变量**估计**实际数值(房价、呼叫次数、总销售额等)。我们通过拟合最佳直线来建立自变量和因变量的关系。这条最佳直线叫做回归...

深度学习常用评价指标_深度学习评价指标

11-5

深度学习常用评价指标 评价指标: 准确率 (Accuracy),混淆矩阵(Confusion Matrix),精确率(Precision),召回率(Recall),平均正确率(AP),mean Average Precision(mAP),交并比(IoU),ROC + AUC,非极大值抑制(...

深度图可视化显示(kitti) 最新发布

weixin_38252409的博客 611

kitti数据是一个通用数据，有关kitti的**深度**图像内容我已有博客介绍。这里，我将给出一个工具，用于显示**深度**值对应像素坐标再图像上，也给出**深度**值可视化显示内容。

深度估计（一）

qq_41769706的博客 2448

深度估计学习笔记

数据挖掘十大经典**算法**(详解)_挖掘**算法**有哪些

11-6

ID3**算法** 1.概念提取**算法**CLS 1) 初始化参数C={E},E包括所有的例子,为根. 2) IF C中的任一元素e同属于同一个决策类则创建一个叶子 节点YES终止. ELSE 依启发式标准,选择特征Fi=[V1,V2,V3,...Vn]并创...

姿态估计 从原理到实践_姿态估计**算法**

10-15

解决密集人群的姿态**估计**问题 3.PoseFlow 设计了一种在线优化框架来建立跨帧之间的pose的联系解决连续姿态**估计**问题 使用**深度学习**和OpenCV进行手部关键点检测:Hand Keypoint Detection using Dee...

深度估计(Depth Estimation)

hxqxw的博客 1万+

深度估计，顾名思义，就是利用一张或者唯一/多个视角下的RGB图像，**估计**图像中每个像素相对拍摄源的距离。**深度估计**是从2D图像**估计深度**是场景重建和理解任务的关键步骤，在**计算机视觉**领域属于3...

计算机视觉基础（13）——**深度估计**

猪猪和超超的博客 8390

本节是**计算机视觉**的最后一节，我们将学习**深度估计**。从**深度**的概念和度量入手，依次学习单目**深度估计**和双目/多目**深度估计**，需要知道**深度估计**的经典方法，掌握**深度估计**的评价标准，注意结合对极几...

最优化方法:**深度学习**最优化方法_decoupled weight decay regularization...

11-1

[**深度**解读最流行的优化**算法**:梯度下降] 特点: 前期g_t较小时,regularizer较大.能够放大梯度;后期g_t较大的时候,regularizer较小,能够约束梯度; 主要优势之一,是它不需要对每个学习率手工地调节。自...

深度估计

庾信平生最萧瑟 579

深度估计 1、与**深度**相关通道的概念 **深度图**: 灰色图，该图像的每个像素值都是摄像头到位表面之间距离的**估计**值，比如：CAP_OPENNI_DEPTH_MAP通道的图像给出了基于浮点数的距离，该距离以毫...

一文读懂Tesla数据标注系统

04-02

一文读懂Tesla数据标注系统 #### 一、引言 随着自动驾驶技术的发展，数据标注作为构建高质量训练数据集的重要环节，对于自动驾驶系统来说至关重要。Tesla作为自动驾驶领域的领头羊，其数据...

一文读懂 深度强化学习**算法** A3C （Actor-Critic Algorithm）

a1424262219的博客 2586

一文读懂 深度强化学习**算法** A3C （Actor-Critic Algorithm） 2017-12-25 16:29:19 对于 A3C **算法**感觉自己总是一知半解，现将其梳理一下，记录在此，也给想学习的小伙伴一个参考。 想要认识...

深度估计参考软件

05-09

提供了一种**深度估计**得标准参考软件，可以得到较好的**深度估计**图，对做立体匹配，及**深度图**提取有很大的帮助。

一文读懂机器学习需要哪些数学知识

11-14

一文读懂机器学习需要哪些数学知识 #### 前言 随着人工智能技术的快速发展，尤其是机器学习领域的不断进步，对于希望进入这一领域的学习者来说，了解所需的数学基础知识至关重要。本文由留...

目标检测 | 一文读懂YOLO系列**算法**（从YOLOv1到YOLOv8）

突然好想你 1656

YOLO (You Only Look Once) 系列**算法**是一种目标检测**算法**，主要用于实时物体检测。相较于传统的目标检测**算法**，YOLO具有更快的检测速度和更高的准确率。本节课就给大家重点介绍下YOLO系列**算...**

单目**深度估计**–深度学习 篇 热门推荐

weixin_43702653的博客 3万+

文章目录一：**深度估计**应用背景1.**深度估计**的定义2.**深度估计**的应用场景3.几种**深度估计**的方法4.使用**深度学习估计**的优缺点二：单目**深度估计**模型1.使用的数据集2.整体网络架构3.分模块解析 I：层级 II...

[论文阅读] BoostingMonocularDepth **深度估计算法**

ljsx123的博客 2823

1.**算法** 1) 分辨率和**深度估计**结果： 低分辨率——>良好的**深度**一致性，**深度**细节差 高分辨率——>良好的**深度**细节，**深度**一致性差 出现这种结果的原因： 1.CNN网络感受野有限，高分辨率下一些像素不...

深度估计方法monodepth2 PSMnet

sizlincnt的博客 814

monodepth2 批量处理 PSMnet python3.6 torch1.4 torchvision0.5.0 运行记录 其他 AD-census rSGM SGBM

一文读懂MADDPG**算法**

09-29

一文读懂MADDPG**算法**的文章介绍了MADDPG**算法**的基本原理和应用。MADDPG代表多智能体**深度**确定性策略梯度**算法**，是一种用于协同多智能体决策的强化学习**算法**。该**算法**通过一个actor-critic...



刘大壮_ 码龄5年 暂无认证

6 112万+ 88万+ 3万+ 原创 周排名 总排名 访问 等级



刘大壮_ 关注

11 64 3

专栏目录

99
积分

10
粉丝

27
获赞

6
评论

183
收藏

私信

关注

大额流量券送不停

多发多得，流量翻倍！

去看看

热门文章

- 一文带你读懂深度估计算法 8518
- 知识蒸馏基础篇（经典论文） 7352
- yolov3和yolov5 android端实时部署(ncnn) 6029
- yolov3 MNN框架部署C++版 3775
- [论文阅读] BoostingMonocularDepth 深度估计算法 2823

分类专栏

- 工程化部署 2篇
- 深度学习 4篇
- 目标检测
- 模型压缩
- C++

最新评论

- yolov3和yolov5 android端实时部署(ncnn)
清风 : 弄出来了吗
- yolov3和yolov5 android端实时部署(ncnn)
一个渣渣sql_boy: 这个安卓工程如何让在界面上显示YOLOv5的各个检测框的数目， ...
- 一文带你读懂深度估计算法
wlgjy: 我也在看深度估计，能加个v交流一下吗
- 一文带你读懂深度估计算法
我是一个对称矩阵: 谢谢回复，我的问题是基于单目深度估计的，我的demo是来自 ...
- 一文带你读懂深度估计算法
我是一个对称矩阵: 博主您好，我最近在看深度估计的论文，现在有个疑惑想请问一下 ...

最新文章

- yolov3和yolov5 android端实时部署(ncnn)
- yolov3 MNN框架部署C++版
- [论文阅读] BoostingMonocularDepth 深度估计算法
- 2022年 1篇 2021年 5篇

Beta

举报

刘大壮_

关注

11

64

3

专栏目录

广告



地平线
Horizon Robotics

智能驾驶开发者
认证培训

全面覆盖智驾硬件、
软件、算法、工具链等领域

② 从入门到精通！助力 360 度全维提升

③ 一次掌握智驾全栈知识

④ 线上理论课程+线下实践课程

⑤ 从开发板点亮，到嵌入式应用开发

⑥ 初、中、高级系列课程，适配不同需求

立即报名

目录

深度估计算法调研

1.常见的深度估计算法

2.深度图转换成实际距离

3.深度估计算法流程



Beta



举报

刘大壮_

关注

11

64

3

专栏目录

https://blog.csdn.net/ljx123_/article/details/121440256

6/6