

开题报告论文答辩PPT

基于几何结构的三维人体姿态估计



指导老师: 曾辽原



答辩人: 罗子建 2016010902012





01 选题依据及意义02 研究现状及发展态势03 课题研究内容1 主要技术路线



1、选题依据及意义

● 选题依据

● 研究意义

● 应用场景

选题 依据

深度学习被用于计算机视觉领域多个不同的研究方向,包括了图像分类、物体识别、目标检测、人脸识别、人体姿态估计与识别等。这一系列的研究方向中,与人相关的研究往往是最有价值的,而与人相关的分析中对人体姿态的估计与识别是当下的一个热点。

选题 依据

人体姿态估计是基于图像信息对人体的各个关节和刚性 部件进行准确检测和有效组合。其目的是获取人身体各个 关键点的位置,得到正确位置后,对关键点进行正确的连 接形成人体骨架信息,在后续研究中可以利用骨架信息对 人的动作和行为进行分析。



自动理解图像及视频序列中的人体姿态,一直都是计算机视觉的研究重点。

意义

第二点 意义

对于人体姿态估计,应用场景广泛,对应技术产品的落地机会大。

第三点 意义

有效的处理和理解数据中的人类活动,将会对人类社会的发展带来深远的影响。

应用一

在智能监控领域,利用人体姿态估计通过监控设备可 对人的某些反常行为进行辅助监控和预警。

应用二

在体育竞技领域中,可以获取人的实时位置和姿态, 增加竞赛成绩评价的手段。

应用三

在康复医疗领域,医生可以根据病人的康复训练完成 状况指定相应的训练计划。

应用四

在智能人机交互方面,允许用户通过某种动作与智能信息进行信息交互传递。



2、研究现状及发展态势

- 国内研究现状
- 国外研究现状

● 发展态势



1

2

2018.11, 电子科技大学CV_lab在<Fast Human Pose Estimation>的论文里,推出了用知识蒸馏的方法获得"模型小计算复杂度低"的易于在边缘设备部署的模型

2019.2,中国科学技术大学、微软亚洲研究院提出了,基于深度高分辨表示学习的姿态估计算法,并设计出一种不同分辨率子网络并行前进的网络结构。

2019.4, 华中科技大学刚刚开源的一款人体姿态迁移算法, 其基于GAN思想构建。给定一幅含有人物的图片和一个目标姿态, 将图片内人物转换成目标姿态的样子。

3

2019.5,浙江大学,德克萨斯大学奥斯汀分校,提出了一种多视图的快速鲁棒多人3D姿态估计算法

国外研究现状

上海交通大学推出的一款AlphaPose的框架,自顶向下方法,SSD-512检测人+stacked hourglass姿态估计



OpenPose人体姿态识别项目是 美国卡耐基梅隆大学(CMU) 基于卷积神经网络和监督学习 并以caffe为框架开发的开源库。 可以实现人体动作、面部表情、 手指运动等姿态估计。





3、课题研究内容

● 研究内容

● 关键问题

● 研究目标

研究内容分析

本课题是针对三维人 体姿态估计主流算法 的原理性验证及创新 探索 在深度学习的框架中,对设计的神经网络结构做改进,以提升识别率和降低平均误差(MPJPE)。

分析

该三维人体姿态估计算法,需完成特定场景的人体主要关节的估计和识别。

3 结合图像测试结果,做出对应的分析结论,总结创新点。



网络结构的设计

尤其是卷积神经网络结构,可以从图像中有效地获取深层次的特征信息,对于三维人体姿态估计这类对几何结构特征及其敏感的复杂任务,需要对网络结构进行改进,使其在更大的范围捕捉特征。

损失函数的设计

需要采用改造损失函数的方式,向卷积神经网络中引入几何约束信息,来实现更为准确的三维人体姿态估计。





如何提高泛化能力

图像存在复杂背景干扰和遮挡问题,还存在视角局限的问题,如何 在出现特殊视角(俯视,仰视)的时候还能准确地对人体姿态进行 估计,这也是研究考虑的一个重点。



设计一款用于三维人体姿态估计的神经网络结构, 并根据GPU上的图像测试结果,进行原理性验证。



通过训练来不断优化网络结构的参数,得出一个高准确率并且高泛化能力的3D人体姿态估计模型算法。



不断优化网络结构的参数,使算法的平均误差(MPJPE) 达到62mm以下



4、主要技术路线

- 主要技术路线
- 实施方案

● 工具

设计一个2D人体姿态估计模块 (获取二维人体关节位置),参考 OpenPose, AlphaPose框架

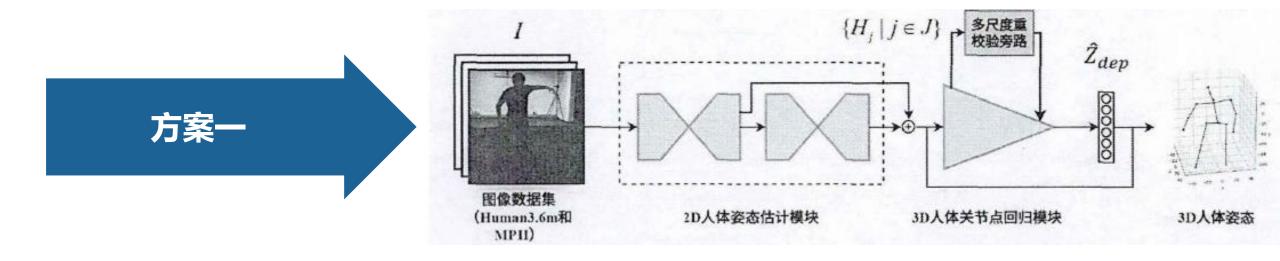
设计一个深度回归模块(从2D人体 关节的信息恢复3D人体姿态),并 通过引入的2D人体姿态进行几何 约束,最终得到对应的深度预测



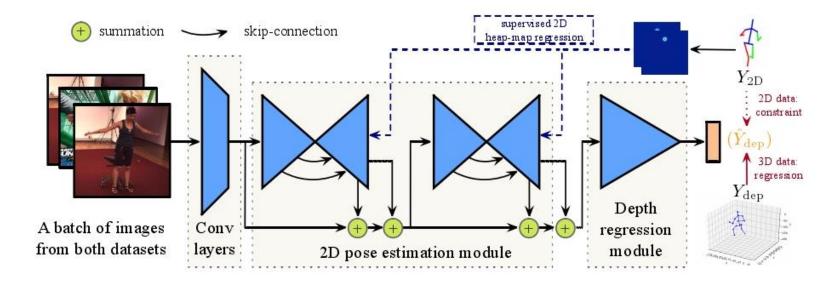
系统输入为包含人类的RGB图像

得到二维人体姿态估计对应的热力图(2D-HeatMap)和低分辨率的二维图像特征

系统输出为3D人体姿态估计数据









硬件

GPU(若干)

摄像头(若干)

数据集

Human3.6M,PoseTrack,
CrowdHuman,MPI-INF-3DHP



谢斯氏听

基于几何结构的三维人体姿态估计



指导老师: 曾辽原



答辩人: 罗子建 2016010902012