

3D-human-pose 项目开题报告

智能 16 唐超

本项目主要对野外单幅图片的 3D 人体姿态估计问题进行研究，试图复现这一问题以往的研究成果并尝试做出一些改进。

1 项目描述

传统的人体姿态估计问题，往往基于深度图或基于多目图像识别 3D 人体姿态，本项目试图通过深度学习等方法，训练出一个可以识别 3D 人体姿态的模型，仅输入一副含有单人的野外图像给该模型，即可自动输出图像中人体的姿态估计结果。

2 研究思路

目前现存的使用深度学习网络解决这个问题主要有两个方向：一是由室内图片直接学习 3D 人体姿态再通过别的方法拓展到室外图片，二是先由图片学习得到对应的 2D 人体关节点，再根据各种方法学习 2D 关节点向 3D 人体姿态的转换。在第二种方法中，由单幅图片到 2D 关节点这一步已经有许多现存的做的比较好的方法，发展比较成熟，而且从 2D 关节点到 3D 人体姿态估计这一过程，可以选择的切入点比较多，比如距离矩阵，几何特征等等。因此本项目采取第二种方法，即：单幅图片 -> 2D 关节点 -> 3D 人体姿态。

3 相关材料

3.1 数据集

- MPII: 2D 的野外图像对应人体关节点数据库，包含了已做标注

的 25K 张图像，涵盖了 410 种人体姿态；

- Human3.6M: 3D 的室内实验室人体姿态图像数据集，总共含有 360 万张 3D 人体姿势对应的图像；

3.2 配置环境

- cuDNN: CUDA Deep Neural Network 的缩写，在配置 cuDNN 前需要先安装 cuda，cuda (Compute Unified Device Architecture) 是一种由 NVIDIA 推出的通用并行计算架构，该架构使 GPU 能够解决复杂的计算问题；
- Torch7: 一种科学计算框架；

3.3 论文及开源代码

- **Stacked Hourglass Networks for Human Pose Estimation**
https://link.springer.com/chapter/10.1007/978-3-319-46484-8_29
- **Structured Feature Learning for Pose Estimation**
http://www.ee.cuhk.edu.hk/~xgwang/projectpage_structured_feature_pose.html
- **VALSE2017 系列之三：人体姿态识别领域年度进展报告**
<http://www.10tiao.com/html/580/201705/2650326421/1.html>
- **单幅图片生成 2D 人体关节点相关论文：**
<https://github.com/shihenw/convolutional-pose-machines-release>
- **单幅图片生成 3D 人体姿态相关论文：**
<https://github.com/xingyizhou/pose-hg-3d>
<https://arxiv.org/pdf/1611.09010.pdf>
- **【翻译论文】 Learning to Estimate 3D Human Pose and Shape from a Single Color Image (CVPR 2018)**
- https://blog.csdn.net/Kansas_Jason/article/details/81021886

3.4 相关知识点

- 卷积网络
- 堆积沙漏网络

4 研究内容

- 了解 2D 的野外（MPII）或者是 3D 的室内实验室（Human3.6M）数据集的基本使用方法；
- 完成由单幅野外图片学习 2D 人体关节点的任务，可以基于开源代码对相关方法进行测试；
- 由上一步 2D 人体关节点，学习对应的 3D 人体姿态的估计，可以选择不同的方法完成任务；
- 将以上两步整合成一个整体的框架，尝试有什么改进的方法；

5 时间安排

