

CVTE 工作总结汇报

周佳楸

2019 年 11 月 3 日

1 Abstract

这份汇报主要总结了从 2018 年 12 月到 2019 年 9 月的工作经历, 其中主要负责的是两个项目, 一个是 LED 色度校正项目, 主要完成了亮度校正算法以及色度校正方案的调研; 另一个是 MAXHUB 视线校正项目, 此项目还在进展过程中。此外, 在工作期间完成了单目图像深度估计的毕业设计论文, 对单目图像深度的估计也有一定的了解。

2 MAXHUB 视线校正项目

2.1 背景描述

2.1.1 需求描述

为了有更好的视频会议体验, MAXHUB 想使得人们在开会议的时候可以进行眼神的交流, 如下图所示:



图 1: 眼神交流示意图

2.1.2 问题解析

MAXHUB 产品摄像头布置采用的是广角镜头, 显示屏顶端与水平方向成 23 度俯视场景。相机的布置方案导致相机从上往下俯视, 人之间眼神的交流需要抬头直视摄像头才能进行。

拟解决方案:

- 通过 2D 图生成三维人脸, 旋转人脸模拟直视相机重投影回原图
- 通过相机标定, 制造一个虚拟相机在屏幕中间去拍人脸, 模拟眼神交流

2.2 模拟直视相机

2.2.1 具体实施方案

1. 利用神经网络从一张 2D 图生成 3D 图, 其中生成 3D 人脸的模型有两个, 一个是带脖子的模型 3DDFA, 一个是没有带脖子的模型 PRNet, 分别对两个模型进行实验, 其生成的 3D 人脸图如下图所示:

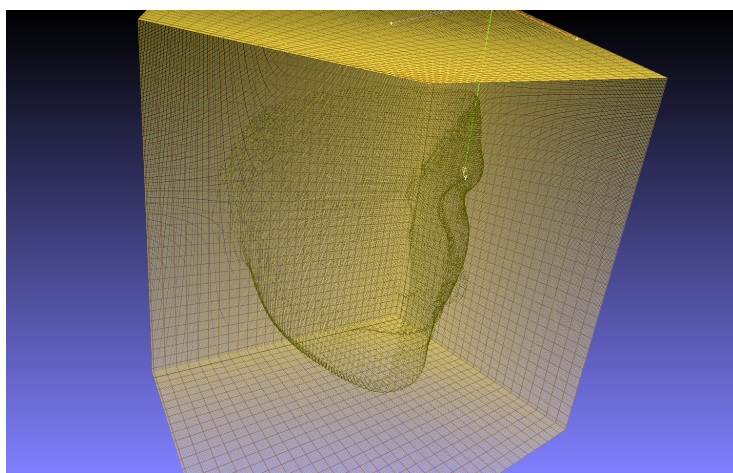


图 2: 三维人脸图

2. 通过估计得到的 3Dpose 计算人脸与相机所成的角度, 其中还做了人的眼神估计的角度, 使用的是 RecurrentGaze 的模型, 其效果如下图所示



图 3: 眼神估计图

3. 通过预估得到的角度, 对生成的 3D 人脸进行旋转, 将原图得到的纹理重新映射到转后的三维模型上, 使得人脸好像在直视相机

4. 最后重新投影回到原来的图片, 进行图像融合, 其效果如下所示



图 4: 校正效果图

2.2.2 存在问题

- 人脸与原图融合问题
- 3D 人脸模型重新映射出现的人脸变形问题

2.2.3 问题解析

根本问题在于生成的 3D 模型只针对于人脸, 使得其他部分与人脸的融合都会有一定的不自然问题, 而且对于在原图中的纹理人脸是向下看的, 其根本属性导致了人脸的变形问题

2.3 虚拟相机拍摄

2.3.1 简介

考虑到之前的方案存在不自然的问题, 是由于 3D 人脸的局限性, 因此考虑通过相机标定来制造虚拟相机从中间拍摄图像得到眼神交流的图像, 其简单的小实验如下所示。

拍摄一张相机直视的图片:



图 5: 利用相机直接拍摄得到的图片

和一张相机旋转的图片:



图 6: 利用相机旋转直接拍摄得到的图片

通过模拟虚拟相机, 旋转相机 15 度用仿射变换得到的图片:



图 7: 算法模拟相机旋转得到的图片

2.3.2 具体实现方案

1. 利用 google 单目深度估计网络估计人体的绝对深度, 然后利用之前所使用的三维人脸模型 PRnet 进行深度调整得到整体的人体深度

2. 利用标定板对相机进行标定, 得到内参矩阵 K 通过针孔相机建立的模型, 即以下的公式便可以从图像坐标转换三维坐标, 如下所示:

$$Z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = K \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

3. 通过得到的三维坐标对相机的坐标系进行旋转, 由于只有 X 轴上的旋转, 所以公式如下:

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x & 0 \\ 0 & \sin\theta_x & \cos\theta_x & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

4. 得到相机旋转后的三维坐标后, 再将其投影回到原来的 2D 图像中, 模拟相机的采集过程, 然后再将原来人体的纹理贴回到图像坐标中, 与原来图片进行融合, 达到视频眼神交流的效果。

2.3.3 存在问题

- 人体边缘与图像融合存在一定问题
- 绝对深度以及内参矩阵的估计不准确性带来了生成图像的误差

2.3.4 问题解析

根本问题在于相机从中间拍摄会比从上面拍摄的人脸图像要大, 导致融合的不自然问题。另一方面, 相机为广角镜头, 标定会有一定的难度, 单目生成的深度也没有那么准确。

2.4 项目收获

2.4.1 经验积累

- 有的时候, 发现问题比不断的去调整表象更加重要
- 解决问题的思路在于不断的剖析问题, 从问题的本质出发

2.4.2 技术积累

- 相机的内参标定以及三维坐标系旋转
- 三维 mesh 的纹理映射, 以及对人脸进行三角剖分映射
- 利用 OPENCV 实现图像融合算法
- 三维人脸以及眼神估计的模型了解与应用

3 LED 色度校正项目

3.1 背景描述

3.1.1 问题描述

LED 显示屏在使用时间变长之后, 会出现花屏的现象, 如下图所示:

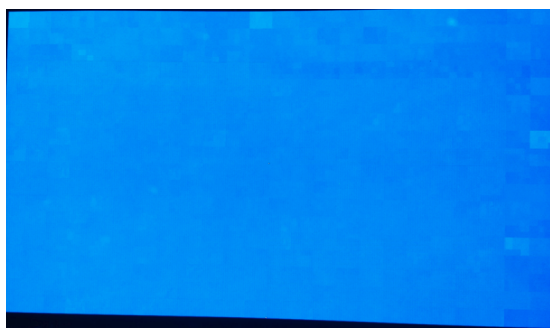


图 8: 发光不均匀的 LED 显示屏

对于 LED 色度校正项目所要达到的效果, 如下图所示:



图 9: 校正效果图演示

3.1.2 问题解析

花屏现象主要是由于 LED 灯板出厂后亮度衰减程度不一样, 导致发光亮度不一致而产生花屏。从使用场所出发, 一般为大型广场的广告屏, 整批更换不太现实; 进一步地, 进行部分更换的话, 由于使用时间的长短不一, 同样会出现不均匀性的问题。因此, 通过软件算法来控制 LED 灯板的亮度一致性成为经济有效并且主流的方式。

3.2 解决方案-LED 亮度校正

3.2.1 分解项目任务

- 对每一个 LED 灯点进行定位
- 提取每一个 LED 灯点的亮度
- 选取合适的亮度, 输出校正文件

3.2.2 具体流程

LED 灯板照片可视化的效果如下:

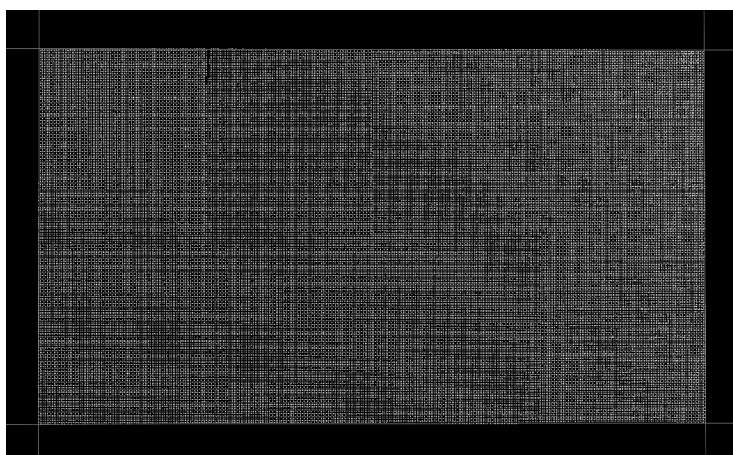


图 10: LED 灯板演示图

1. 灯点定位: 由于 LED 灯板的规则性, 因此可以利用其进行相机标定。其中实现的关键难点也就在于相机的标定, 需要对 LED 灯板四个边角的灯点进行定位, 如图 10 所示的边线即是四个边角的连线, 确定了四个边角之后再进一步确定一层一层剥离 LED 灯板得到里层的边角, 便可以模拟提取的角点进行相机的标定。这样, 得到相机的内参矩阵以及畸变矩阵, 就可以对 LED 灯板图像进行畸变校正, 使得坐标与 (0,0)-(216,384) 一一对应起来, 实现灯板的定位。

2. 提取亮度: 通过定位后, 计算每个区域的亮度即像素值进行叠加, 得到每个区域的亮度。

3. 输出校正文件: 得到每个区域的亮度, 便可以定义灯板整个区域所能达到的最低亮度, 然后计算比例便可以输出校正文件。

3.3 LED 色度校正调研

色度校正部分只完成了前面的调研方案部分, 其核心在于相机 RGB 空间与标准色彩空间 CIE-XYZ 之间的转换。

3.4 项目收获

3.4.1 经验积累

- 实验算法的验证从底层照片的获取到代码的编写, 每一步都需要严谨
- 对于算法不 work, 一定要沉下心去寻找摸索每一个可能出错的点
- 验证算法必须有一个统一的标准, 否则无法知道自己的算法完成进度

3.4.2 技术积累

- 单反相机快门速度, 曝光时间等参数的使用
- 基于针孔模型的相机内参标定以及畸变校正
- Opencv 图像处理中包括寻找凸包, 腐蚀膨胀等几何学操作

4 实习总结

4.1 思想总结

- 在算法出现问题的时候，先确保理论正确，然后再确保编码正确。
- 在进行每一步的算法开发的时候，必须明确每一步的目的，在算法效果不 work 的时候，一定要知道为什么不 work，否则就是浪费时间。
- 对所实现的东西问题的定义非常重要，只有正确定义问题，才能确保所行进的方向没有出现错误。

4.2 技术积累

- 单反相机各参数作用以及相机镜头的内参标定和畸变校正
- Opencv 图像处理和融合算法以及几何学操作等算法
- 三维物体的纹理映射，人脸三角剖分等算法原理
- 眼神估计，单目深度估计，三维人脸的模型了解与使用