

实时人脸贴纸：基于3D人脸模型

唐雯豪(1800013088), 韩宇栋(1800013097), 崔轩宁(1800013083)

联系邮箱: tangwh@pku.edu.cn, hanyd@pku.edu.cn, cxn@pku.edu.cn

指导教师: 刘家瑛

北京大学, 人工智能引论课程
2018-2019, 春季学期

引言

Faceu和B612等相机APP深受青少年的欢迎, 原因之一便是其中丰富的贴纸选择。脸部贴纸 (facial filters), 是指通过计算机视觉方法将图案贴到人脸图片上。提出了一个基于三维人脸模型的人脸贴纸添加方案:

- ① 使用MTCNN进行人脸检测
- ② 使用3DDFA进行三维人脸关键点标注
- ③ 利用三维人脸信息添加贴纸, 并射影变换到原图片

此方案的优势在于:

- ① 使用三维信息, 贴纸添加在大角度侧脸上也有很好的表现。
- ② 拥有优秀的实时性能, 在GPU上可达20fps甚至更高, 在CPU上可达10fps。同时, 实现了一个带有GUI界面的脸部贴纸添加Demo, 支持实时视频、拍照以及拖动图片, 并且用户可以自由搭配不同贴纸。

关键词: Face Detection, Facial Landmark Detection, 3DMM

方法 Part I

Face Detection:

- 人脸检测环节是比较耗时的一部分, 经过对多种方法的测试, 选取了MTCNN作为人脸检测方法。

检测器	FPS
MTCNN	25.694
YOLO V2	4.504
RFCN-ResNet101	4.564
TinyYOLO	16.851
dlib*	6.067

不同方法fps对比 (测试环境: GTX1050Ti with Max-Q Design)
*: 基于HOG的传统方法测试在CPU上进行, 且准确率较差

- MTCNN采取了有三个子网络的层叠结构:

P-Net: 产生人脸候选窗口

R-Net: 拒绝大量错误窗口

O-Net: 进一步筛选

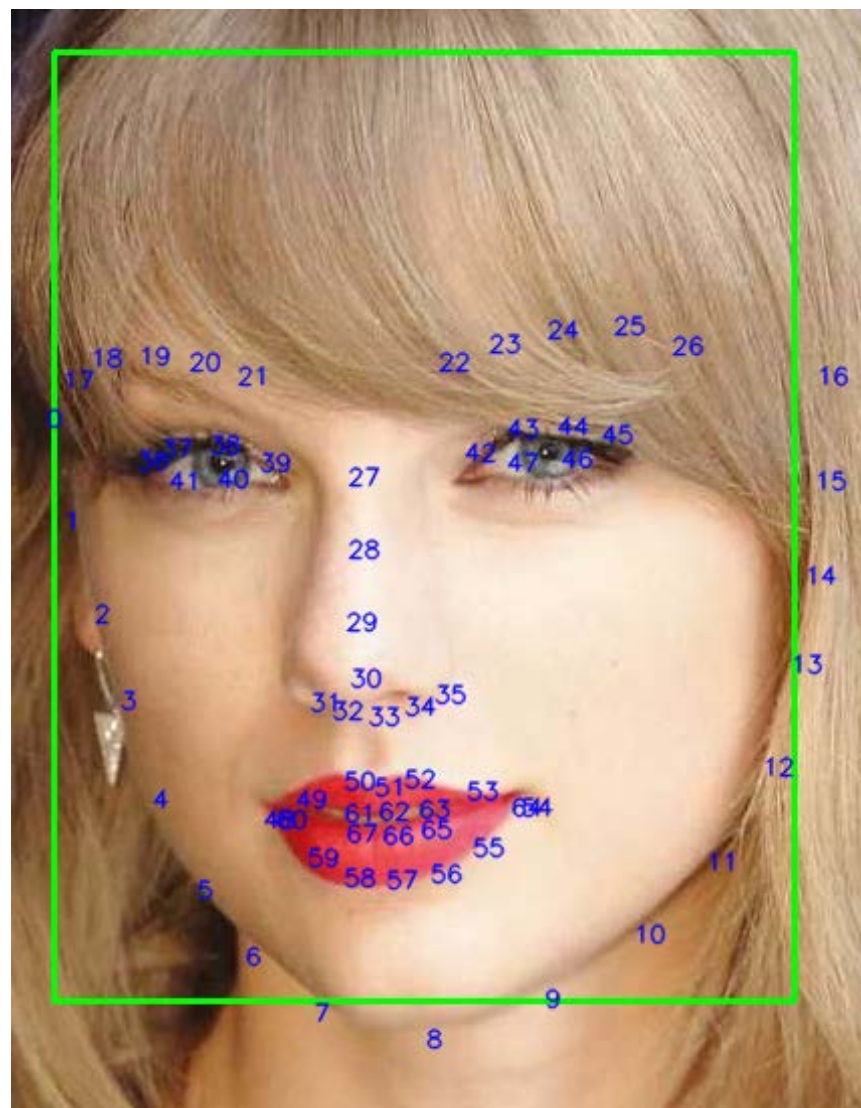
方法 Part II

Facial Landmark Detection:

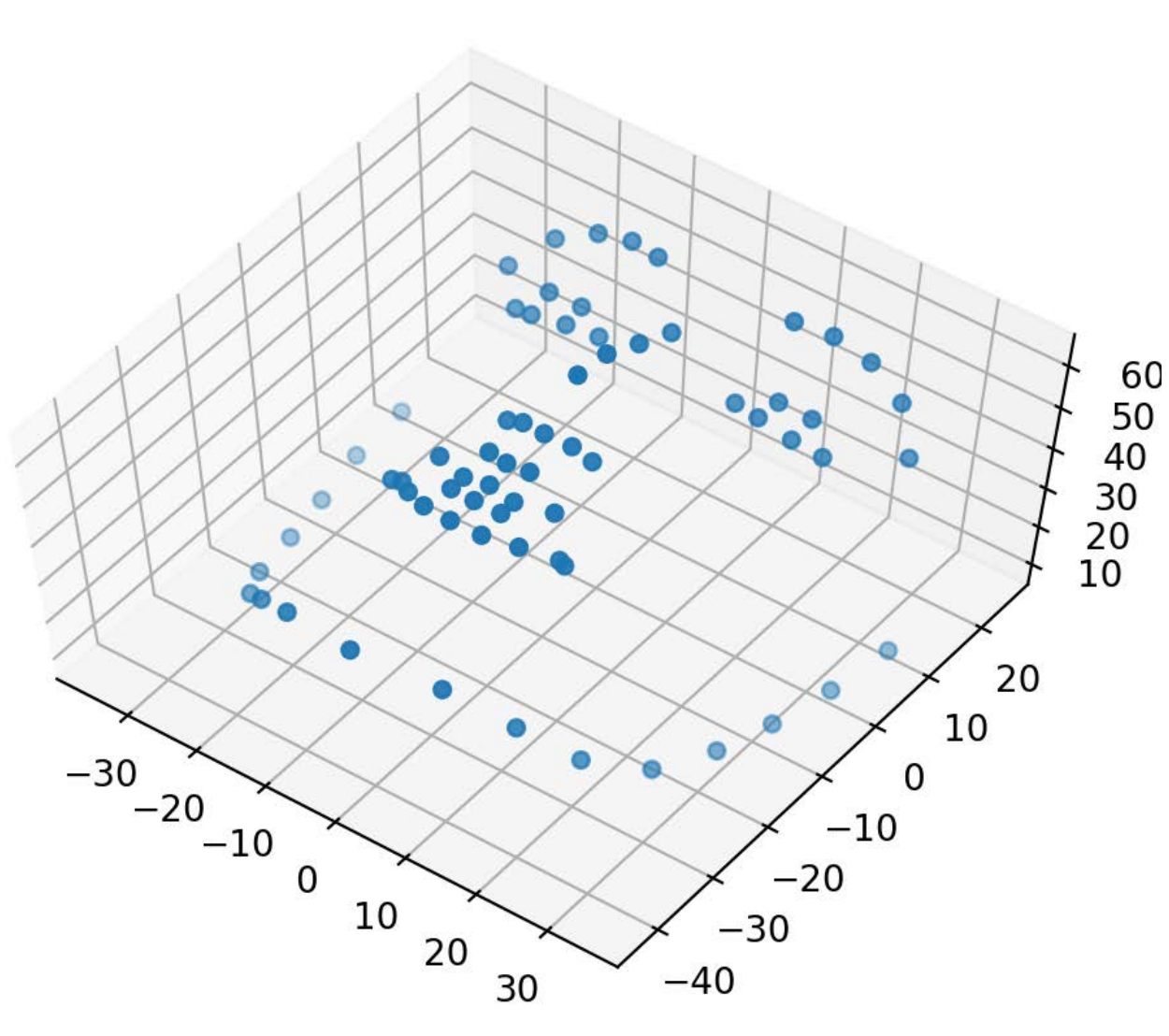
- Motivation: 常见的2维平面关键点检测算法对大角度侧脸的鲁棒性不好
- 使用了3D Dense Face Alignment (3DDFA)来实现人脸特征点标注
- 通过Cascaded Convolutional Neural Networks实现用Dense 3D Morphable Model (3DMM)人脸模型拟合人脸
- 3DMM人脸模型的公式描述:
$$\mathbf{S} = \bar{\mathbf{S}} + \mathbf{A}_{id}\alpha_{id} + \mathbf{A}_{exp}\alpha_{exp}$$
- 拟合过程使用了Cascaded Regression的方法, 其中每个Regressor是一个CNN, 使用了相同的网络结构, 回归过程公式如下:
$$\Delta \mathbf{p}^k = \text{Net}^k(\text{PAF}(\mathbf{p}^k, \mathbf{I}), \text{PNCC}(\mathbf{p}^k, \mathbf{I}))$$
- 网络的损失函数使用了OW-PDC

Composing Filters with Images:

- 得到3DMM人脸模型之后, 建立一个以人脸为中心的三维坐标系, 并得到68个人脸关键点在三维空间中的坐标



68个关键点的编号和位置



以人脸为中心三维坐标系中的68个关键点

- 在三维坐标中通过固定一个锚点和一个放缩尺度的方式确定贴纸中4个点的位置, 利用摄像机矩阵将这四个点变到摄像机坐标系中, 最后通过射影变换将贴纸图片变到原图像平面上。公式如下:
$$\mathbf{V}(\mathbf{X}) = \mathbf{f} * \mathbf{Pr} * \mathbf{R} * \mathbf{X} + \mathbf{t}$$
- 对于加贴纸过程耗时问题, 采取了一些代码层级的优化措施, 比如: 分离alpha通道, 只对非透明区域进行处理等。

结果



总结

提出了一个使用MTCNN检测人脸、使用3DDFA得到三维人脸模型并添加人脸贴纸的方案流程, 并且实现了一个带有GUI界面的脸部贴纸添加Demo, 支持实时视频、拍照以及拖动图片添加贴纸, 用户可以自由搭配不同贴纸。该方法有着很好的实时性能, 并且对大角度的侧脸有很好的鲁棒性, 最终的效果可以和Faceu, B612等APP相媲美。

在这次项目中, 我们学会了许多人脸检测和人脸关键点检测的基于深度学习的现代方法与空间射影几何的基础理论, 同时编写代码和GUI界面的能力有了很大的提升。

下一步的改进有三个方向: 视频帧间连续性的优化、性能提升和三维贴纸的添加。

参考文献

- [1] Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE Signal Processing Letters 23.10 (2016): 1499-1503.
- [2] Zhu, Xiangyu, et al. "Face alignment in full pose range: A 3D total solution." IEEE transactions on pattern analysis and machine intelligence 41.1 (2019): 78-92.

