|  |  |
| --- | --- |
| 人脸对齐调研 | 文件类型：技术文档 |
| 文件编号： |
| 面向的部门： 研发中心 |
| 保密等级：高 |
| 作者： 刘凯 |
| 日期：2018-11-8 |
| 版本：1.0V |

人脸对齐调研



目 录

[1 人脸对齐简介 3](#_Toc529539999)

[2 级联回归人脸对齐 4](#_Toc529540000)

[3 深度学习人脸对齐 5](#_Toc529540001)

[参考文献 1](#_Toc529540002)

# 人脸对齐简介

人脸对齐(Face Alignment)，又称人脸特征点检测(Facial Landmark Detection)，即估计图片中人脸特征点的坐标，如眼睛、鼻子、嘴巴等。设常见的人脸特征点包括5点、19点和68点等，如图所示。



1. 5点人脸坐标，CelebA数据库



(b) 19点人脸坐标，AFLW数据库



(c) 68点人脸坐标，300W数据库

图1 不同数据库的人脸特征点

人脸对齐是很多应用中不可或缺的模块，如人脸识别(Face Recognition)，特征计算(Attribute Compute)，表情识别(Expression Recognition)等[1]。在实际应用中，人脸对齐面临着很多挑战，如姿势(Pose)、遮挡(Occlusion)、表情(Expression)和光照(Illumination)等，如图所示。



实现人脸对齐的方法大致可以分成两类[2]：基于人脸模型参数的算法和不基于人脸模型参数的算法(parametric shape model-based methods and nonparametric shape model-based methods)。基于人脸模型参数的算法是假设模型中的数据服从某些分布，如高斯分布、混合高斯分布等。不基于人脸模型参数的算法则不对模型中的数据没有先验分布的要求。基于人脸模型参数的算法可以细分成两个子类：Local Part-based Methods，如ASM(Active Shape Models)[3]算法和Holistic Methods，如AAM(Active Appearance Models)[4]算法。同样，不基于人脸模型参数的算法可以细分成四个子类：基于模板的算法(exemplar-based methods)、基于图模型的算法(graphical modelbased methods)、基于级联回归的算法(cascaded regression based methods)和基于深度学习的算法(deep learning based methods)。人脸对齐的算法分类如图所示。更多关于各类算法的代表算法和原理可参看文献[2]。因为级联回归和深度学习这两类算法性能较其他算法更好，因此本文主要研究这两类算法。



图：人脸对齐算法分类

# 级联回归人脸对齐

人脸特征点一般用一个矢量表示：表示，其中为第n个特征点的坐标，N为特征点的个数，一般为5或68等。**S**也被成为shape。

基于级联回归的人脸对齐算法一般是从一个初始的shape开始，通过一系列预先训练好的回归器不断地修正初始人脸，直至最后收敛：



其中*I*为图片，为第k个回归器。

级联回归人脸对齐的代表算法包括：LBF(Local Binary Features)[5], SDM (Supervised Descent Method)[6], CFSS(Coarse-to-Fine Shape Searching)[7], ERT (Ensemble of Regression Trees)[8]等。

# 深度学习人脸对齐

基于深度学习的人脸对齐算法是将人脸对齐问题转化为一个非线性回归问题，如TCDCN(Tasks-Constrained Deep Convolutional Network)[9]算法，将人脸对齐问题转化为多任务学习(MTL, Multi-Task Learning)问题：



其中为loss函数，为正确的人脸坐标，**W**为权值矩阵，为正则项。上述权值矩阵可以通过深度学习的方法得到。

基于深度学习的人脸对齐算法包括：TCDCN[9], CFAN(Coarse-to-Fine Auto-encoder Networks)[10], DRA(Deep Regression Architecture)[11], DSRN (Direct Shape Regression Networks)[12]。

参考文献

1. X. Jin, X. Tan. Face alignment in-the-wild: A Survey. Computer Vision and Image Understanding, 2017.
2. N. Wang, X. Gao, and D, Tao, et al. Facial feature point detection: A comprehensive survey. Neurocomputing, 275(31), 50-65, 2018.
3. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models-their training and application. Computer Vision and Image Understanding, 61(1):38–59, 1995.
4. T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In Proceedings of European Conference on Computer Vision, pages 484–498, 1998.
5. S. Ren, X. Cao, Y. Wei, and J. Sun. Face alignment at 3000 fps via regressing local binary features. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 1685–1692, 2014.
6. X. Xiong and F. Torre. Supervised descent method and its application to face alignment. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 532–539, 2013.
7. S. Zhu, C. Li, C. Loy, and X. Tang. Face alignment by coarse-to-fine shape searching. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 4998–5006, 2015.
8. Vahid Kazemi and Josephine Sullivan. One Millisecond Face Alignment with an Ensemble of Regression Trees. CVPR 2014.
9. Z. Zhang, P. Luo, C. Chen, and X. Tang. Facial landmark detection by deep multi-task learning. In Proceedings of European Conference on Computer Vision, pages 94–108, 2014. TCDCN
10. J. Zhang, S. Shan, M. Kan, and X. Chen. Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment. In Proceedings of European Conference on Computer Vision, pages 1–16, 2014. CFAN
11. J. lv, X. Shao, and J.Xing, et al. A Deep Regression Architecture with Two-Stage Re-initialization for High Performance Facial Landmark. CVPR, 2017.
12. X. Miao, X. Zhen, X. Liu and C. Deng. Direct Shape Regression Networks for End-to-End Face Alignment. CVPR, 2018.