



Computer Vision

第十四周 人脸识别 和手势检测

庞彦

yanpang@gzhu.edu.cn



01

Face Recognition

人脸识别全概述

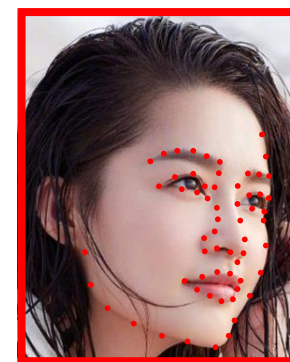
Schedule of Face Recognition



图像采集



人脸检测



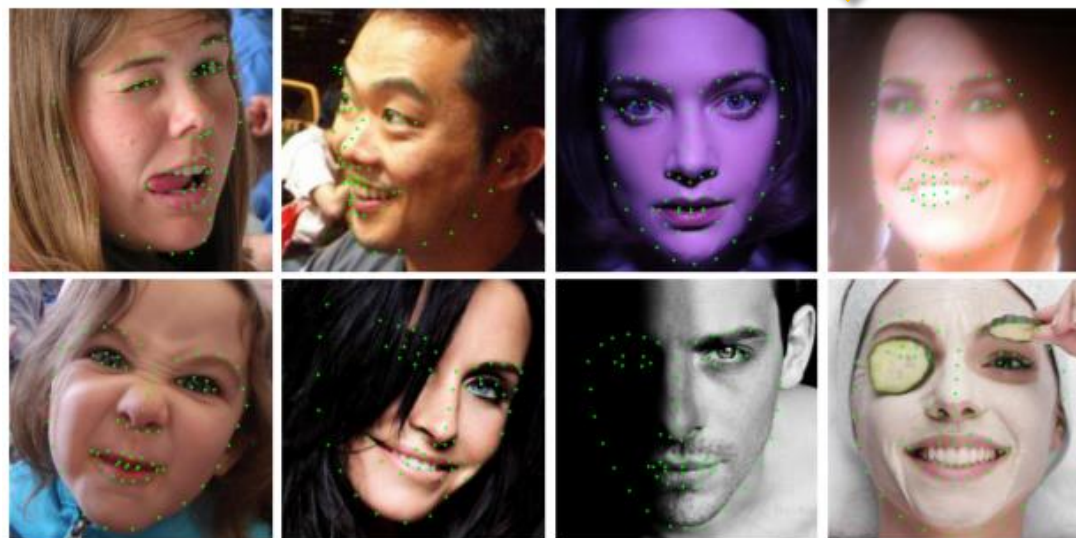
特征点检测



Y
N



人脸
比对



人脸识别应用模式



人脸识别应用模式

- 1:N识别
- 考勤信息管理
- 人脸识别嵌入式终端

- 动态人脸检测/跟踪技术
- 人脸属性分析终端
- 客流统计：数人头数、多维度查询报表、导出查询数据
- 客人统计：属性识别、运动轨迹分析、属性标签
- VIP客户管理：会员注册、精准营销、会员级别管理

人脸识别考勤系统



访客接待系统



智慧商业

- 动态人脸检测/跟踪技术
- 1:N识别
- 预约访客（VIP）接待提醒
- 新访客信息录入
- 人脸识别嵌入式终端



人脸识别客情统计系统



人脸识别门禁系统

- 动态人脸检测/跟踪技术
- 活体检测（安全要求高的区域）
- 1:N识别
- 可集成考勤功能
- 新访客信息录入
- 可集成访客接待功能

人脸识别应用模式

智慧生活

人脸识别物业门禁系统



- 人脸信息采集注册（App端或终端机）
- App刷脸登录
- 刷脸取件
- 活体检测防伪终端
- 私有云服务
- 人脸识别嵌入式终端

人脸识别智能快递投递系统



- 属性分析（帽子/墨镜/口罩等）
- 活体检测
- 现场照片与身份证比对认证
- 黑名单预警
- 人证合一验证终端

人脸识别应用模式

- 属性分析（帽子/墨镜/口罩等）
- 1:N识别
- 黑名单预警
- 动态人脸检测/跟踪技术

动态人证合一通关系统



静态人证合一通关系统



- 人证合一验证终端
- 属性分析（帽子/墨镜/口罩等）
- 活体检测
- 现场照片与身份证比对认证
- 黑名单预警

智慧安防

身份证照N:N查重
比对
人像碰撞比对，挖掘“漂白”身份逃犯等线索信息
现场照片身份证比对认证



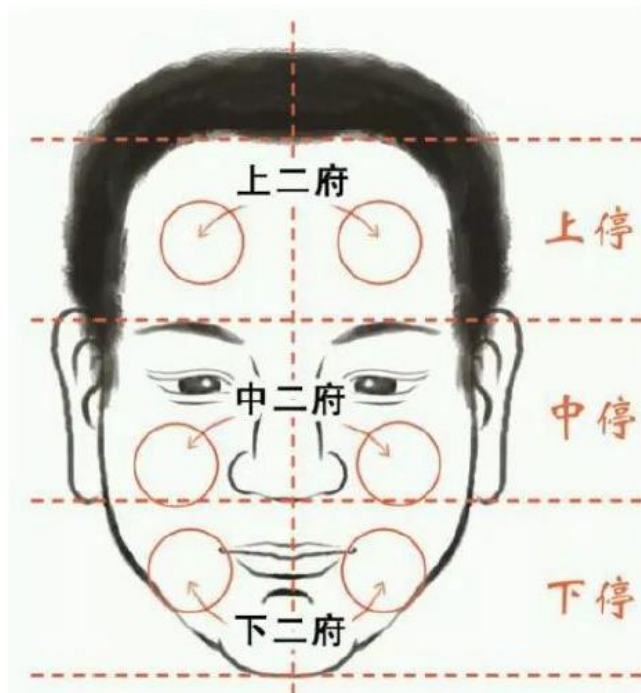
静态超大库人脸检索系统



动态人脸识别防控系统

- 动态人脸监控识别
- 现场照片与身份证比对认证
- 多种实时报警方式
- 动态人脸检测/跟踪技术

Facial Landmark Detection



Current

<http://vis-www.cs.umass.edu/lfw/results.html>

| Traditional Methods (~2013) | | Deep Learning (2013~) | |
|----------------------------------|---------------|-----------------------|---------------|
| Simile Classifiers | 0.8472±0.0041 | DeepID2+ | 0.9947±0.0012 |
| Attribute and Simile Classifiers | 0.8554±0.0035 | Betaface/com | 0.9953±0.0009 |
| Multiple LE + comp | 0.8445±0.0046 | DeepID3 | 0.9953±0.0010 |
| Associate-Predict | 0.9057±0.0056 | Insky.so | 0.9551±0.0013 |
| Tom-vs-Pete | 0.9310±0.0135 | FaceNet | 0.9963±0.0009 |
| Tom-vs-Pete + Attribute | 0.9330±0.0128 | Baidu | 0.9977±0.0006 |
| Combined Joint Bayesian | 0.9242±0.0108 | AuthenMetric | 0.9977±0.0009 |
| High-Dim LBP | 0.9517±0.0113 | YI+AI | 0.9983±0.0024 |
| DFD | 0.8402±0.0044 | Innovative Technology | 0.9988±0.0004 |
| TL Joint Bayesian | 0.9633±0.0108 | Oz Forensics | 0.9987±0.0018 |

Mean classification accuracy and standard error of the mean

Human Face Database

- [PubFig: Public Figures Face](#)
- [Large-scale CelebFaces Attributes \(CelebA\)](#)
- [Colorferet](#)
- [Multi-Task Facial Landmark \(MTFL\)](#)
- [Labeled Faces in the Wild Home \(LFW\)](#)
- [YouTube Faces](#)
- [IMDB-WIKI](#)
- [FDDB](#)
- [Caltech](#)
-

Challenge

干扰因素：

主动因素 + 固有因素；

推广性挑战：

海量比对 + 分布差异

Challenge

干扰因素之主动因素:



正常



姿态



遮挡



表情

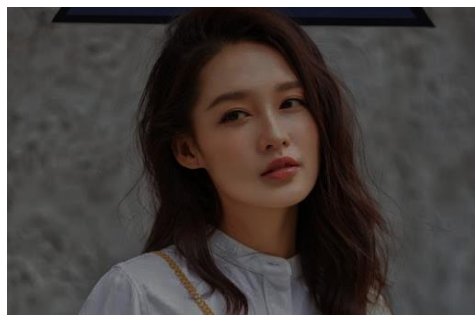


模糊

主动干扰因素可以通过**用户简单配合**来消除;
系统可以通过**拒绝识别**和**质量判断**来获取高质量图像用于识别;

Challenge

干扰因素之固有因素:



光照变化



年龄变化

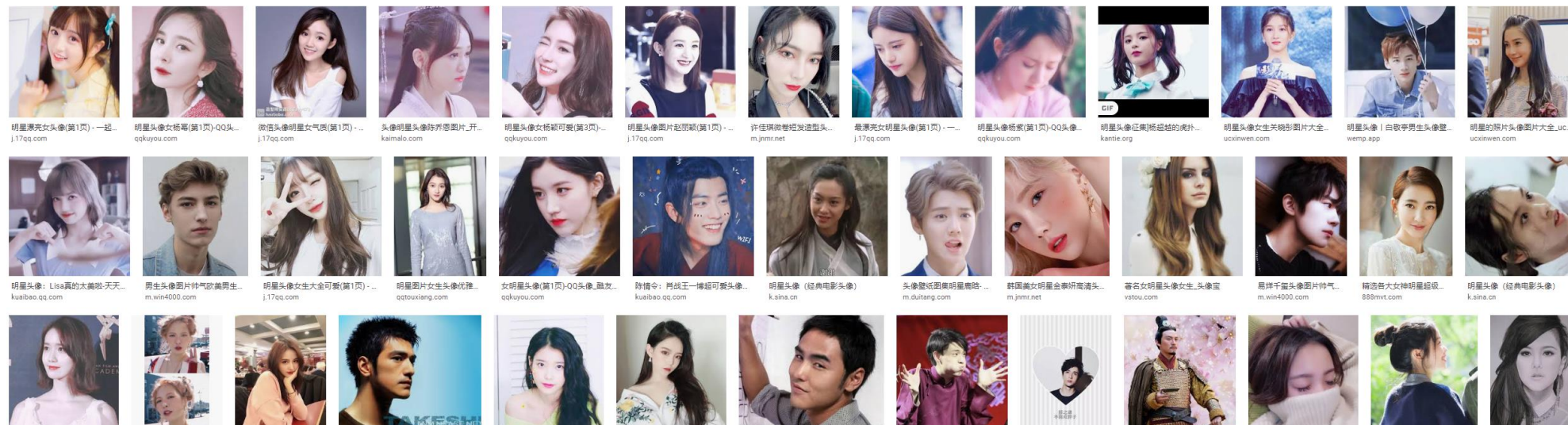


化妆

在现场很难去除的干扰因素，识别器需要对其鲁棒。

Challenge

推广性挑战：海量对比1：N（亿级）



当数据库足够大时，识别器将面临海量对比，总能遇到不同的人脸却看起来很相似的情形。

Challenge

推广性挑战：跨人种，跨年龄



中国人
正常年龄分布

Train

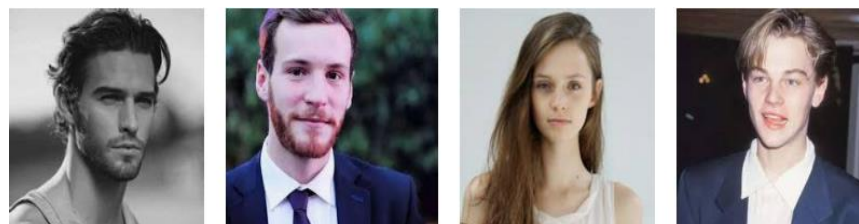
小孩



老人



白人



Test

训练时由于缺乏某个群体的训练集，测试时检测器会对该群体不敏感

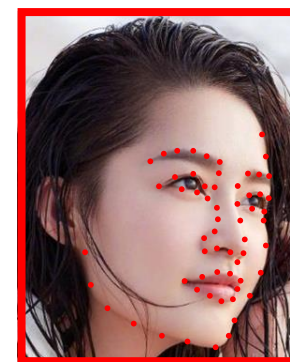
Schedule of Facial Recognition



图像采集



人脸检测



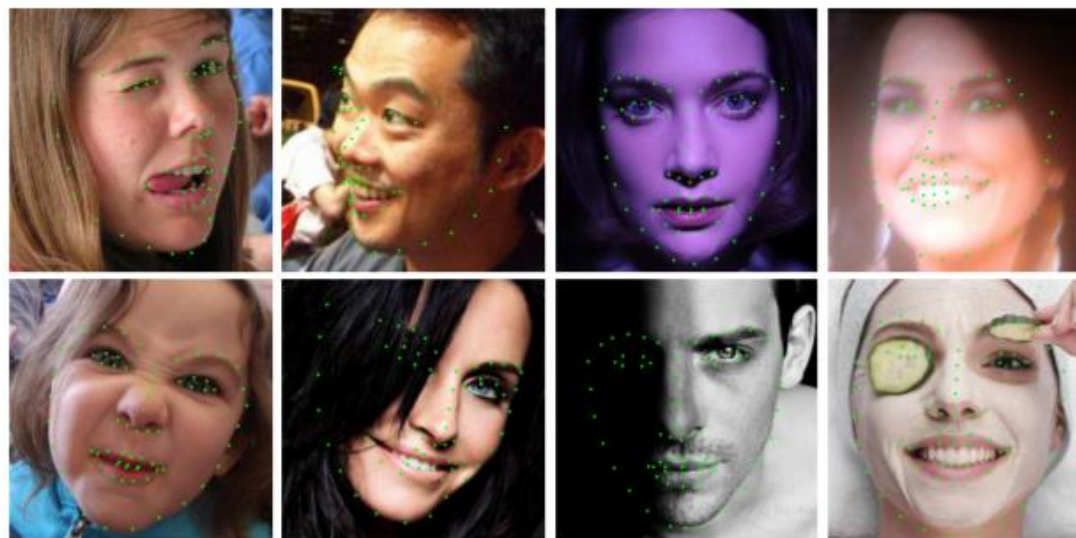
特征点检测



Y
N



人脸
比对



Orientation

更快的速度，需要满足嵌入式等设备；



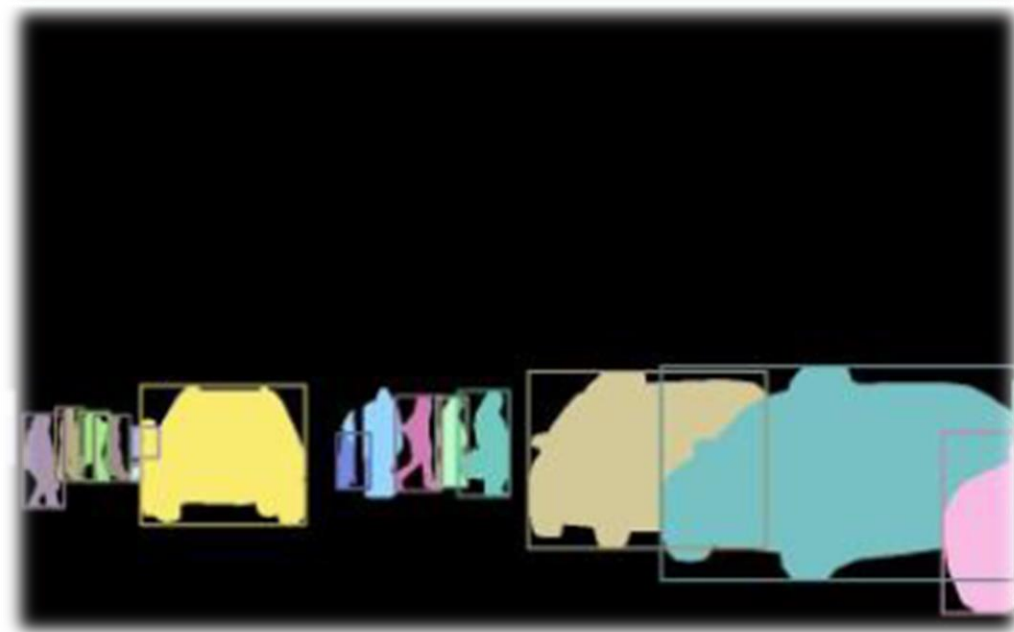
Orientation

更快的速度：需要满足嵌入式等设备；
更高的精度：解决小尺度、遮挡问题；



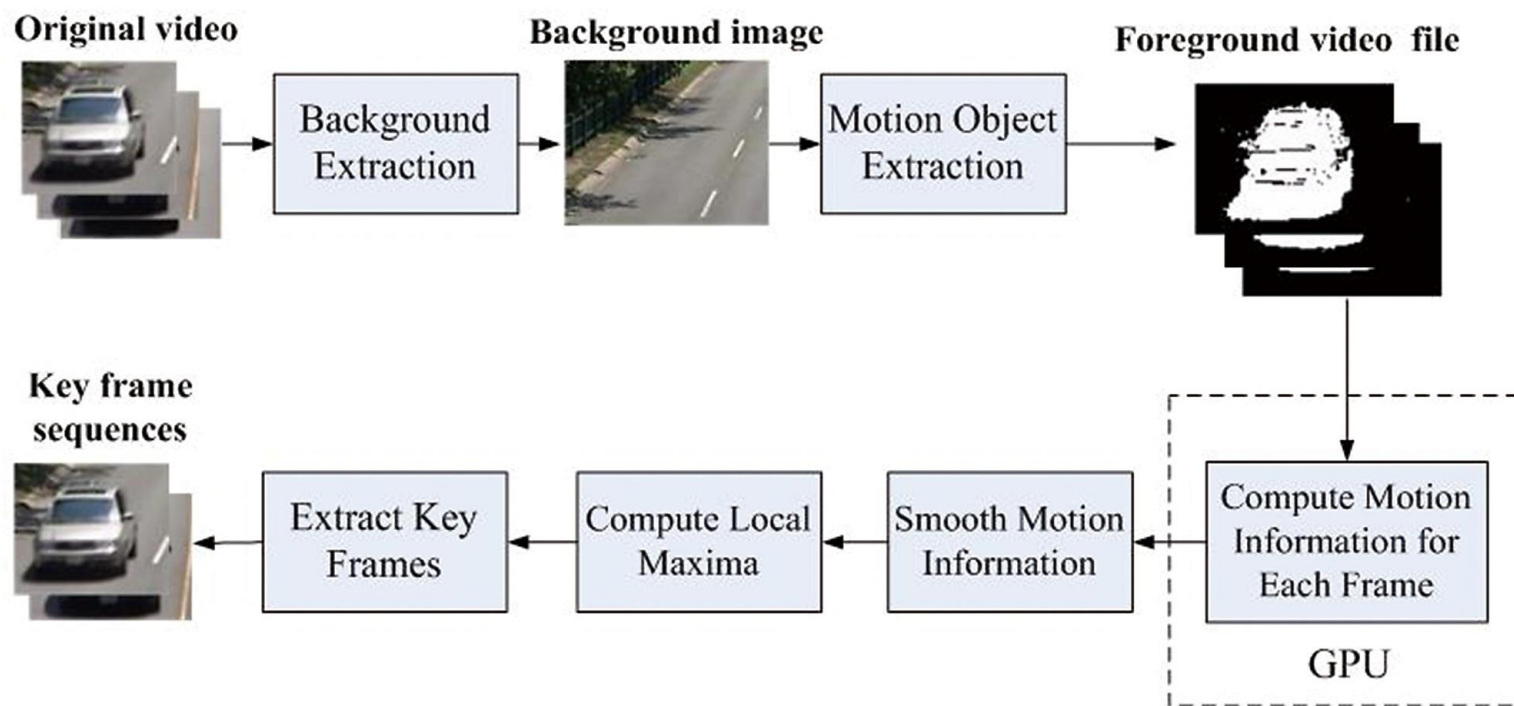
Orientation

更快的速度：需要满足嵌入式等设备；
更高的精度：解决小尺度、遮挡问题；
多任务联合：实例分割 = 检测 + 分割；



Orientation

更快的速度：需要满足嵌入式等设备；
更高的精度：解决小尺度、遮挡问题；
多任务联合：实例分割 = 检测 + 分割；
视频中检测：关键帧信息加速，提精度。



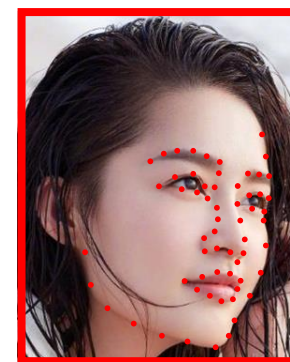
Schedule of Facial Recognition



图像采集



人脸检测



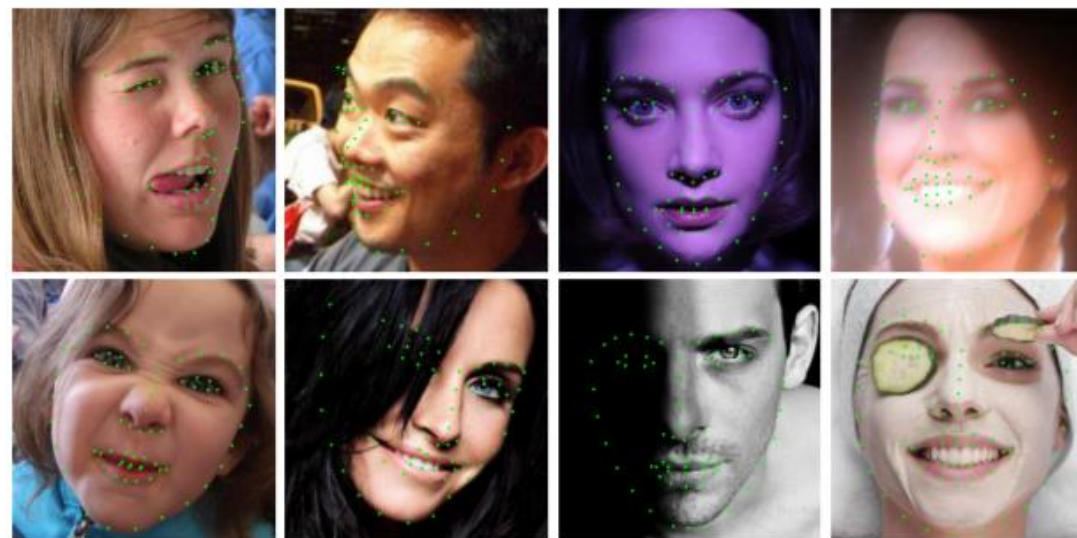
特征点检测



Y
N

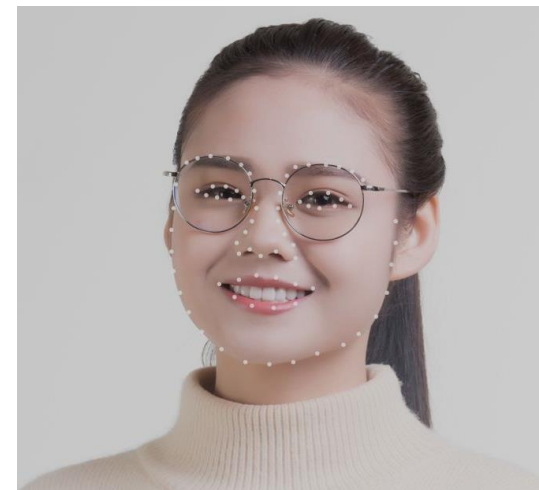
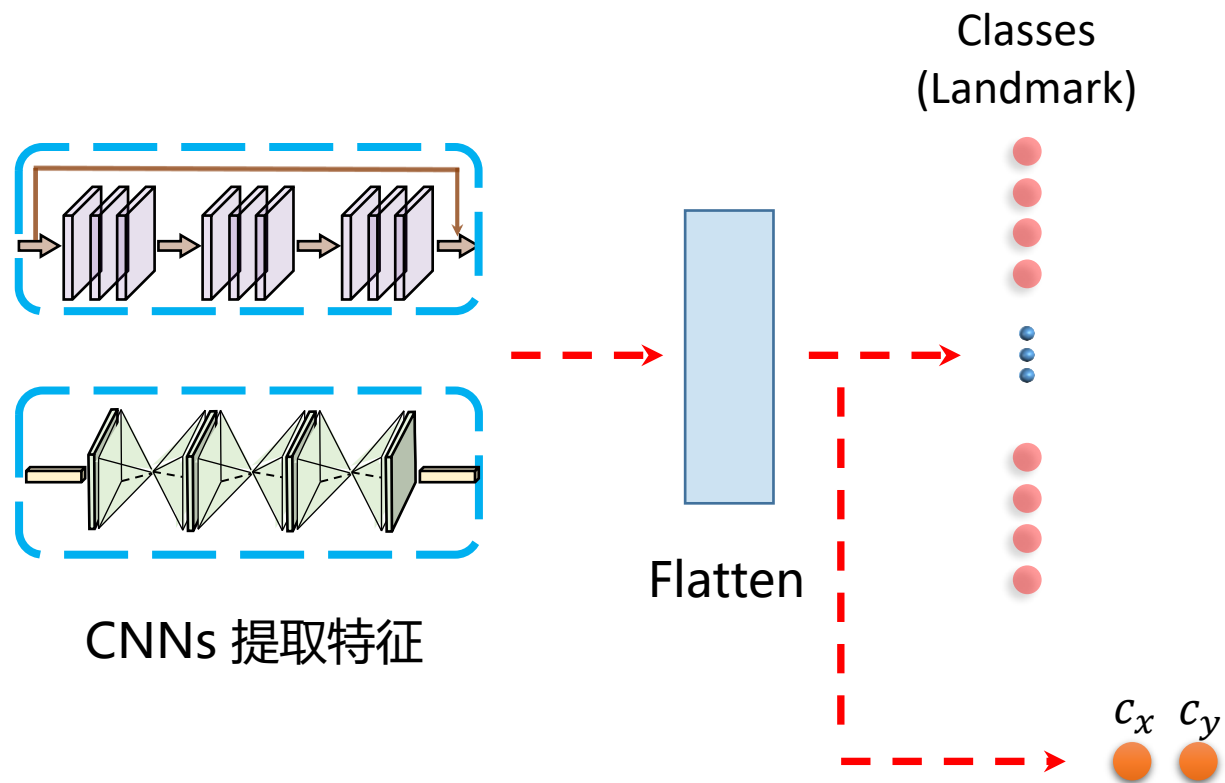


人脸
比对



Facial Landmark Detection

分类+定位问题



$$y = \begin{bmatrix} P_c \\ c_x \\ c_y \\ C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_n \end{bmatrix}$$

Facial Landmark Detection

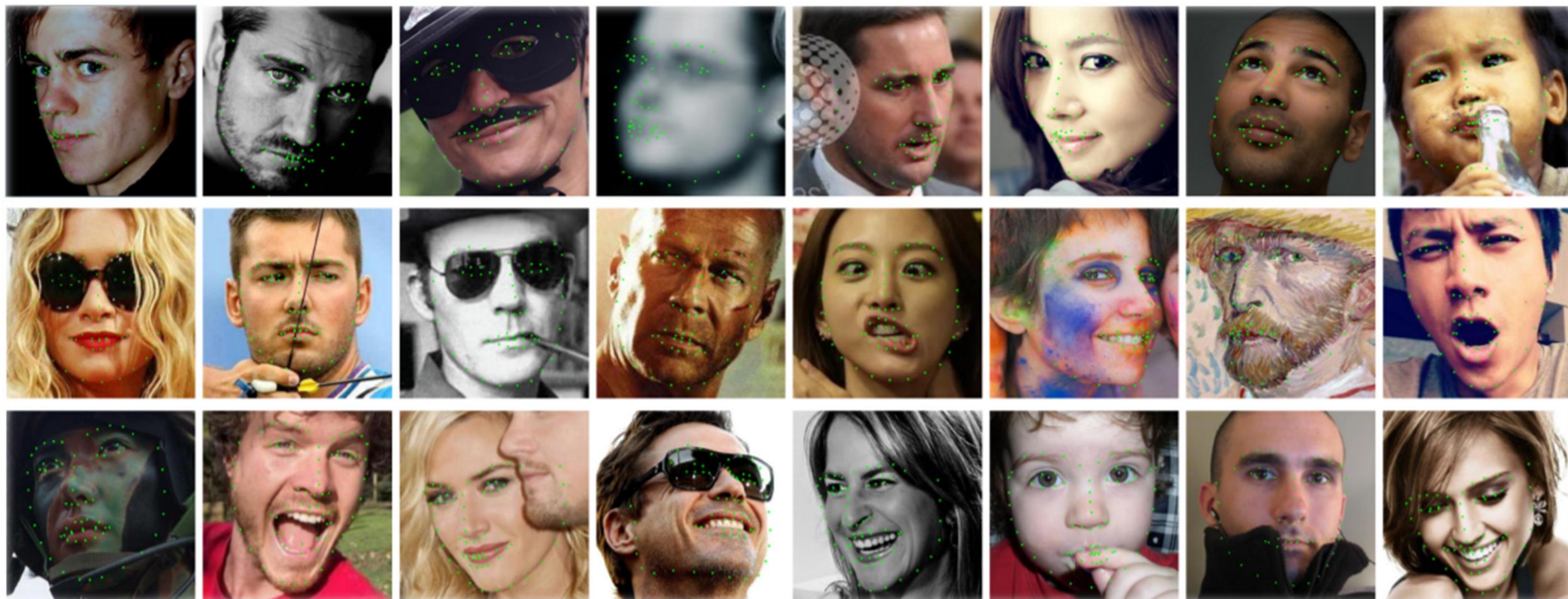
现有问题：

- 大量训练数据都是正脸，侧脸等从其他角度观看的图面较之而言非常少；
- 实际应用的中，对于一些极端情况，如遮挡（手部、眼镜）、光照（强光、弱光）、极端姿态（yaw、pitch、raw较大的时候）、极端面部表情等。

解决方案：

- 通过采用特征描述能力更强的backbone（如DenseNet、CSPDarkNet53等）；
- 增加极端情况下的训练数据、平衡各类情况下的训练数据的比例、控制数据数据的采样形式（非完全随机采样）。

PFLD: A Practical Facial Landmark Detector



LLCV: 8000

旷视研究院 LLCV 组的“8000点人脸关键点技术”；

单张彩色图像或单段视频作为输入，提取出面部稳定鲁棒的 8000 个 3D 关键点。

通过 10,000,000 张人脸数据的训练，该技术可以精细刻画全脸 3D 特征，支持全方位姿态、各种极端表情，并且能在移动端平台实时运行。



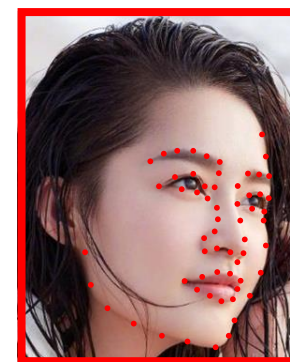
Schedule of Facial Recognition



图像采集



人脸检测



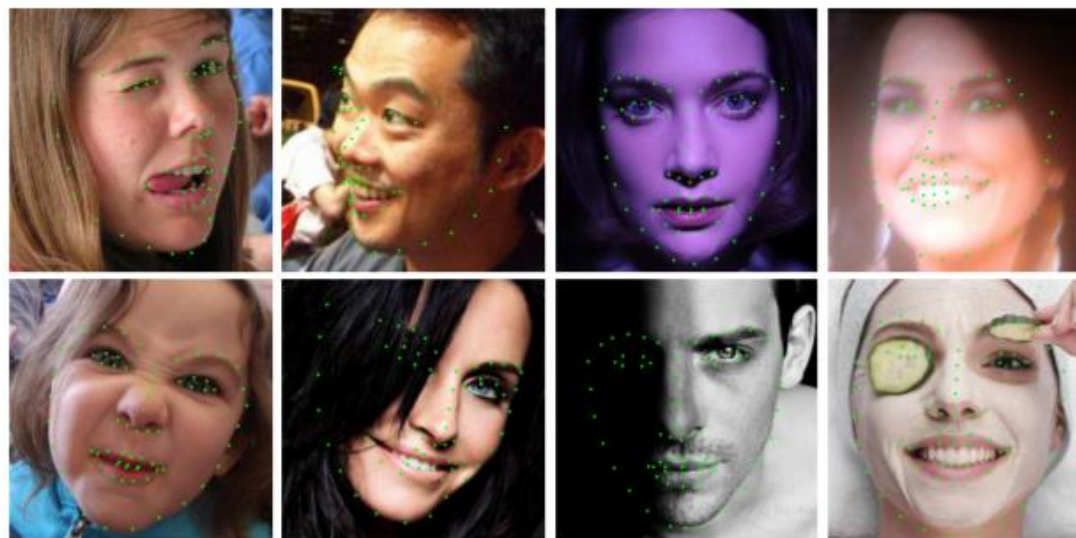
特征点检测



Y
N



人脸
比对



Facial Comparison

人脸比对类型可分为四种，分别是：人脸1:1、人脸1:n、人脸1:N、人脸M:N。

人脸1:1 静态人脸比对，证明你就是：



Facial Comparison

人脸比对类型可分为四种，分别是：人脸1:1、人脸1:n、人脸1:N、人脸M:N。

人脸1:n 动态人脸比对，我是谁？

将一张人脸照片与人脸库中的多张人脸逐一比对，将所有人脸全部比对一次。

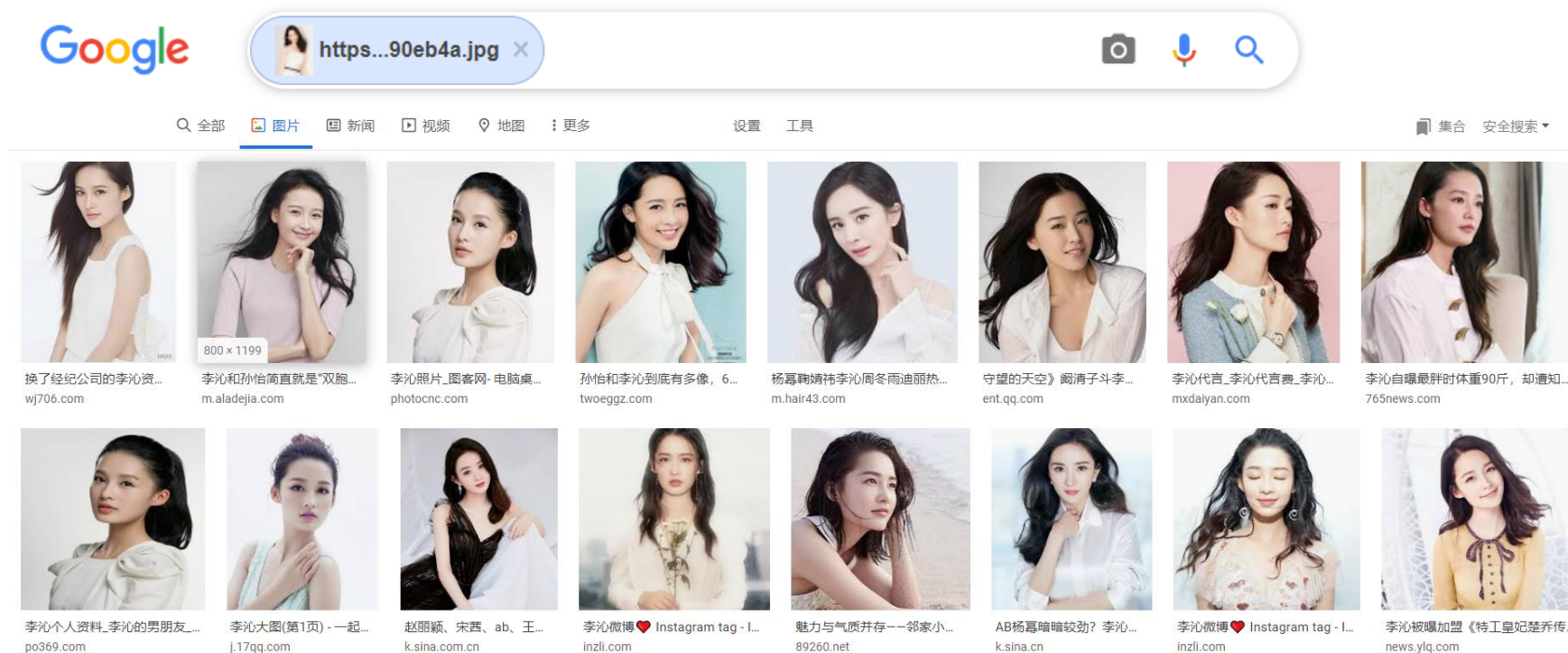


Facial Comparison

人脸比对类型可分为四种，分别是：人脸1:1、人脸1:n、人脸1:N、人脸M:N。

人脸1:N 静态大库检索，找相似。

图片检索，进行N次（亿级）人脸比对，并留下比分大于阈值的结果。



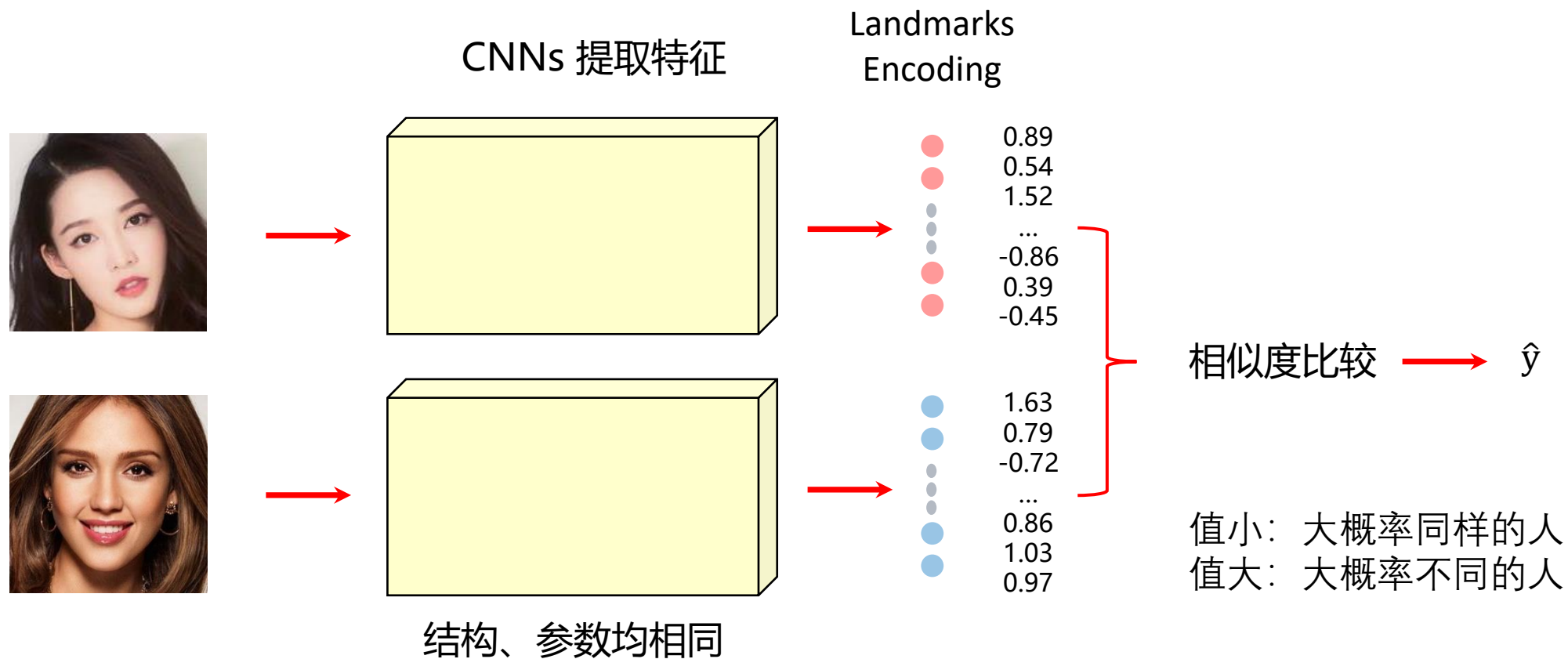
Facial Comparison

人脸比对类型可分为四种，分别是：人脸1:1、人脸1:n、人脸1:N、人脸M:N。

人脸M:N 不同人脸库进行比对。
相当于是M个人脸1:N相加的结果。



Siamese Network



Similarity

内积相似度:

$$\text{sim}(q, k) = q^T k$$

余弦相似度:

$$\text{sim}(q, k) = \frac{q^T k}{\|q\| \|k\|}$$

拼接相似度:

$$\text{sim}(q, k) = \omega^T k[q; k] = \omega_1^T q + \omega_2^T k$$

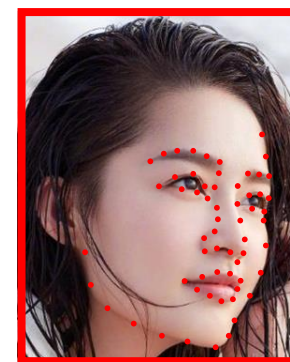
Schedule of Facial Recognition



图像采集



人脸检测



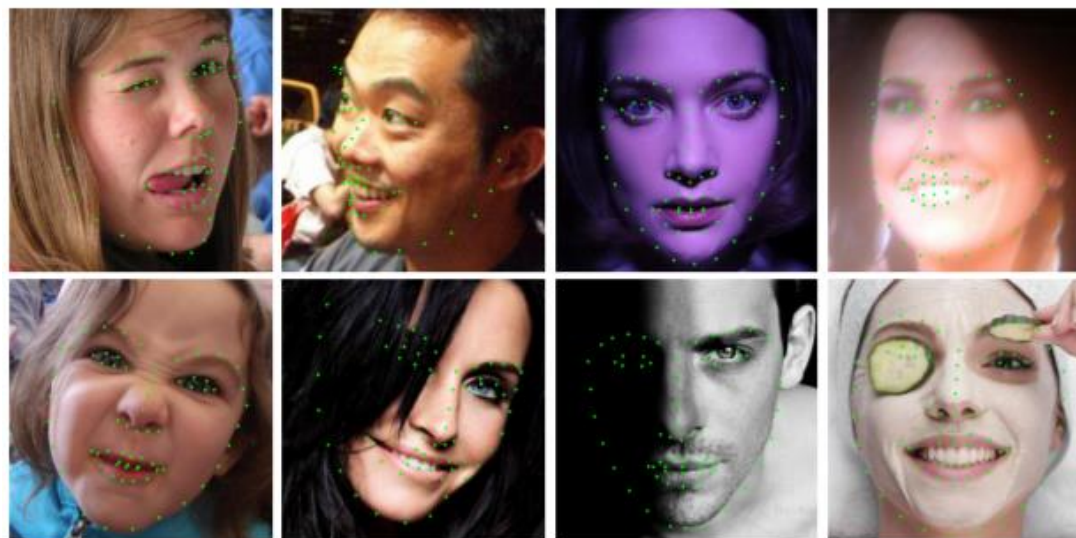
特征点检测



Y
N



人脸
比对





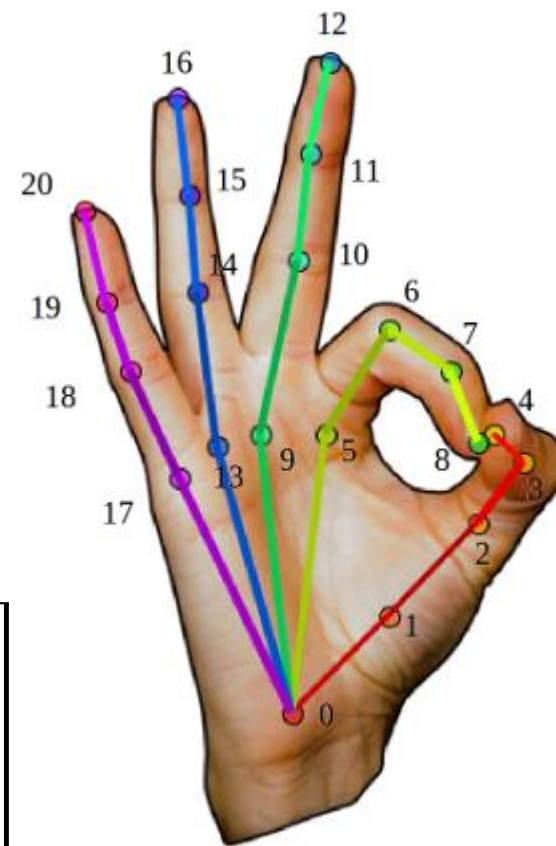
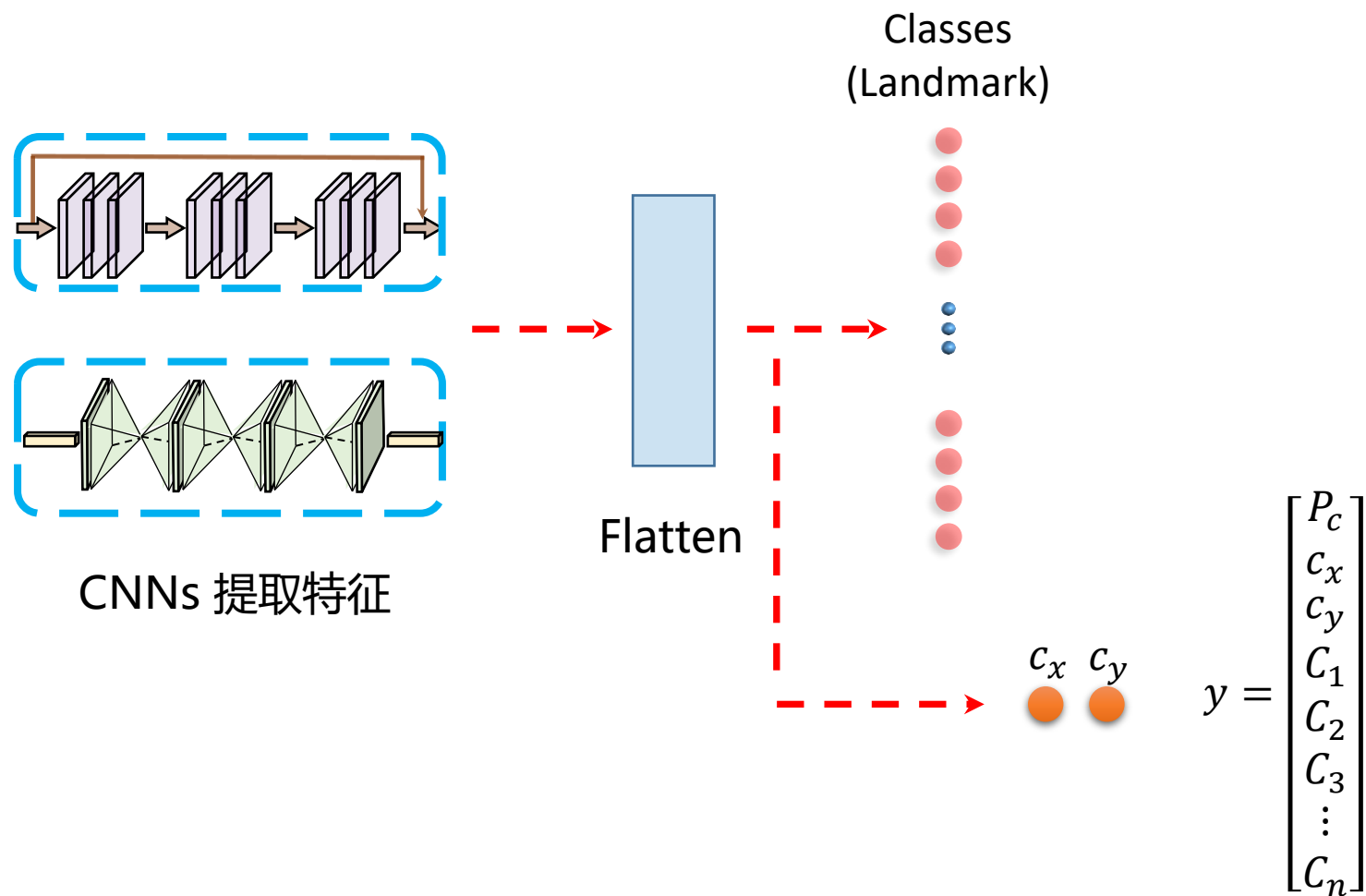
02

Facial Landmark Detection

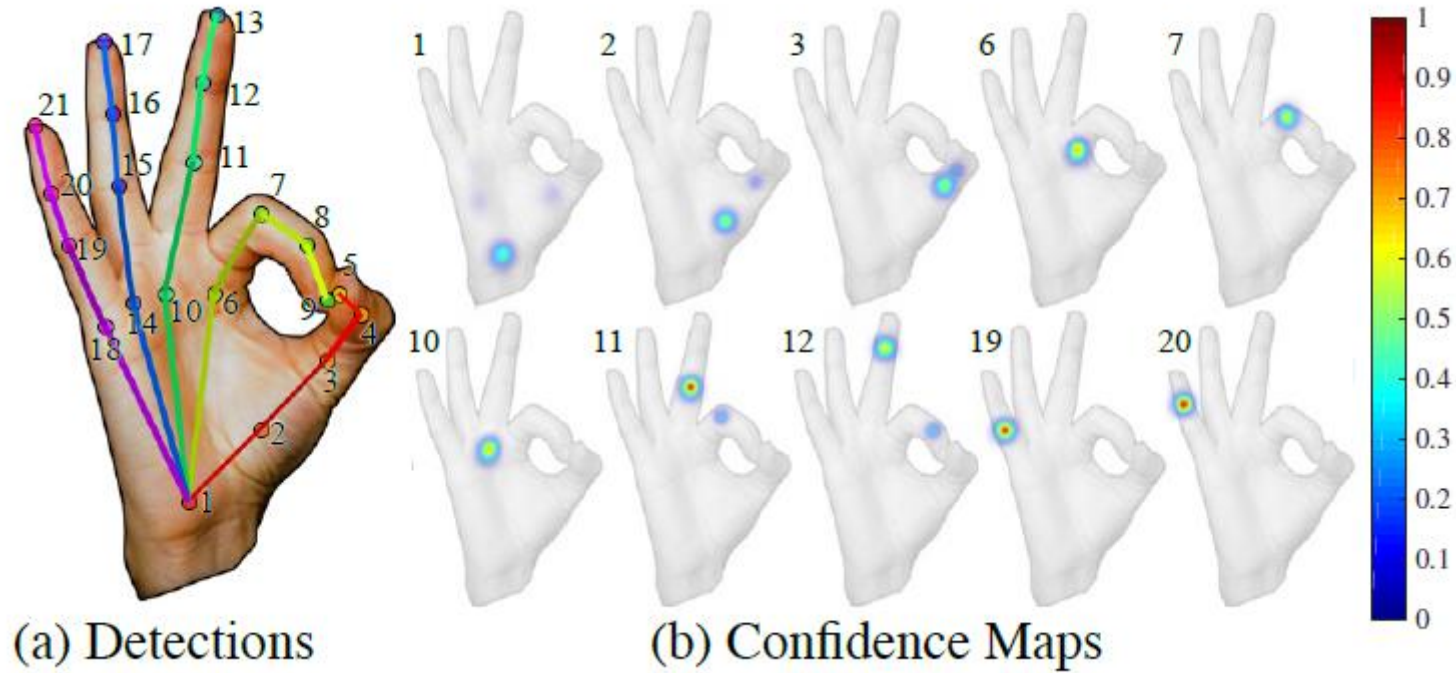
手势检测

Hand Key-Joints Detection

分类+定位问题



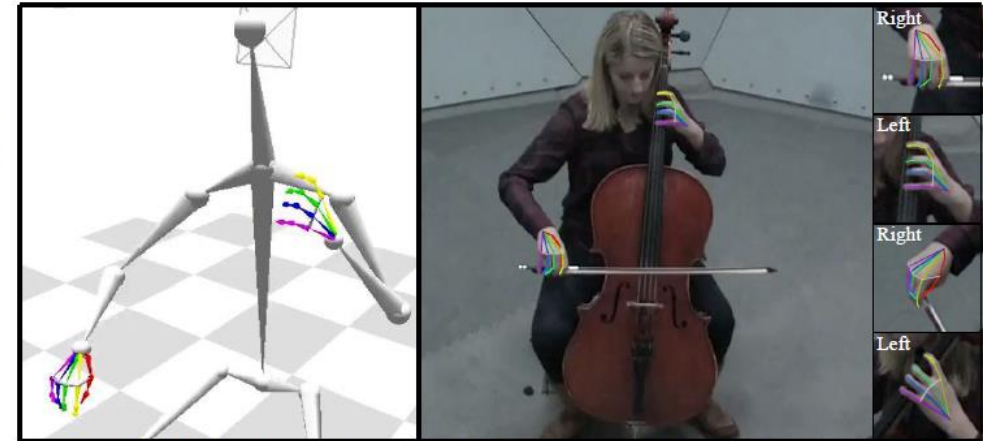
Hand Key-Joints Detection



2D-3D Hand Detection

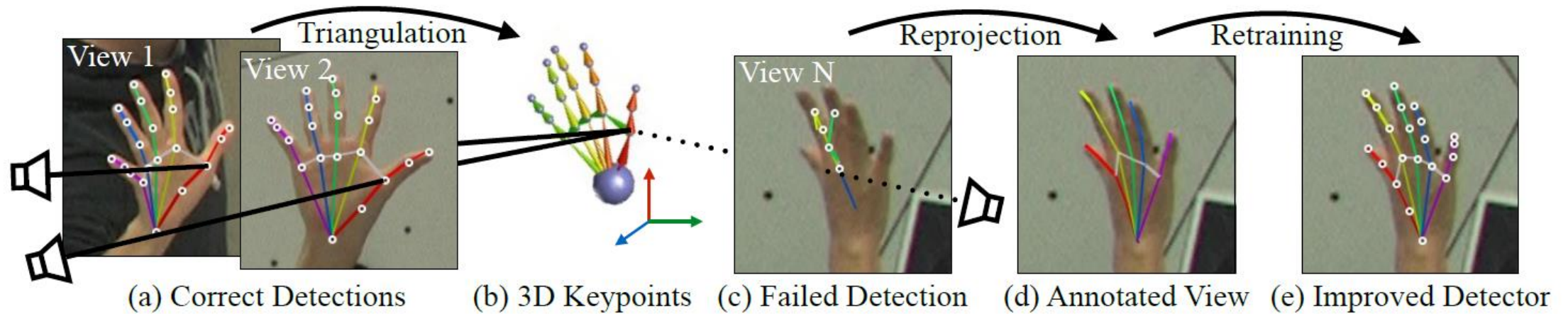


(a) Realtime 2D Hand Detection on YouTube and Webcam Videos

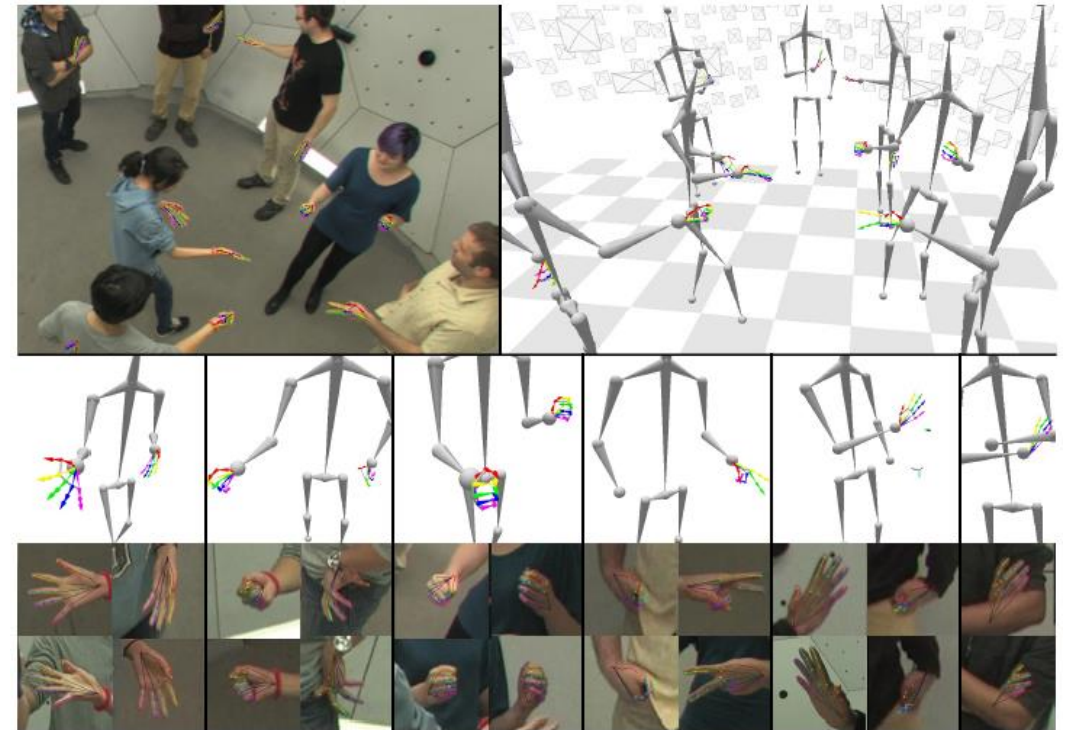
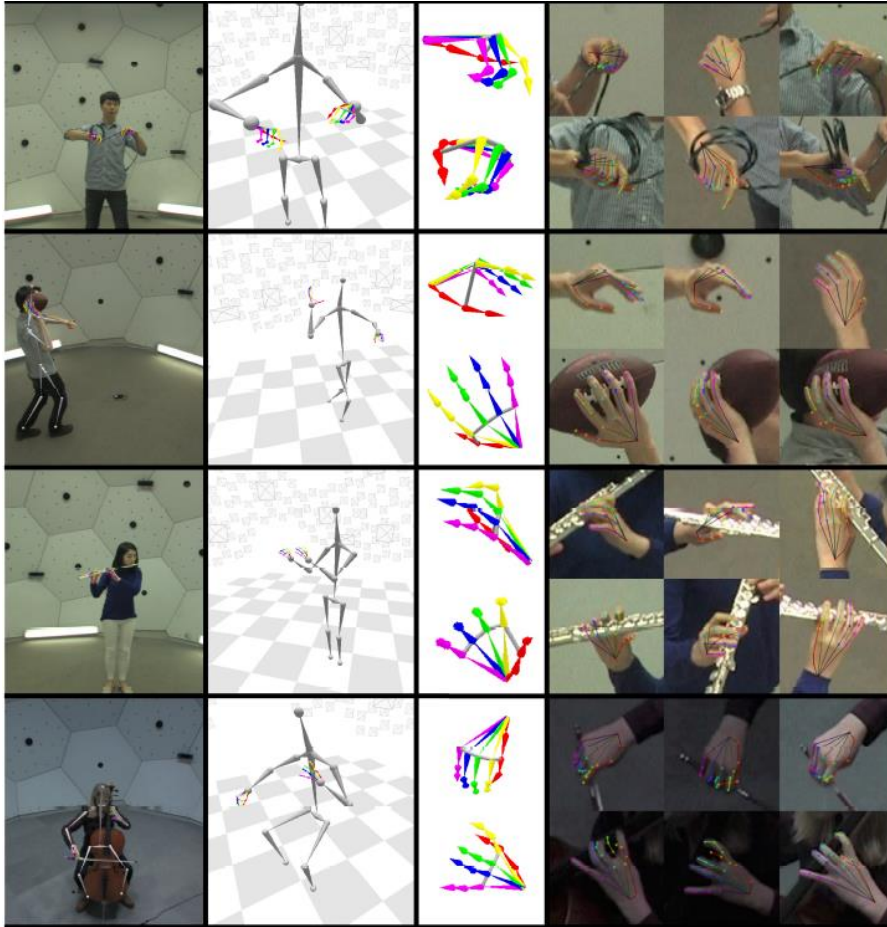


(b) 3D Hand Motion Capture by Triangulating Multiple 2D Detections

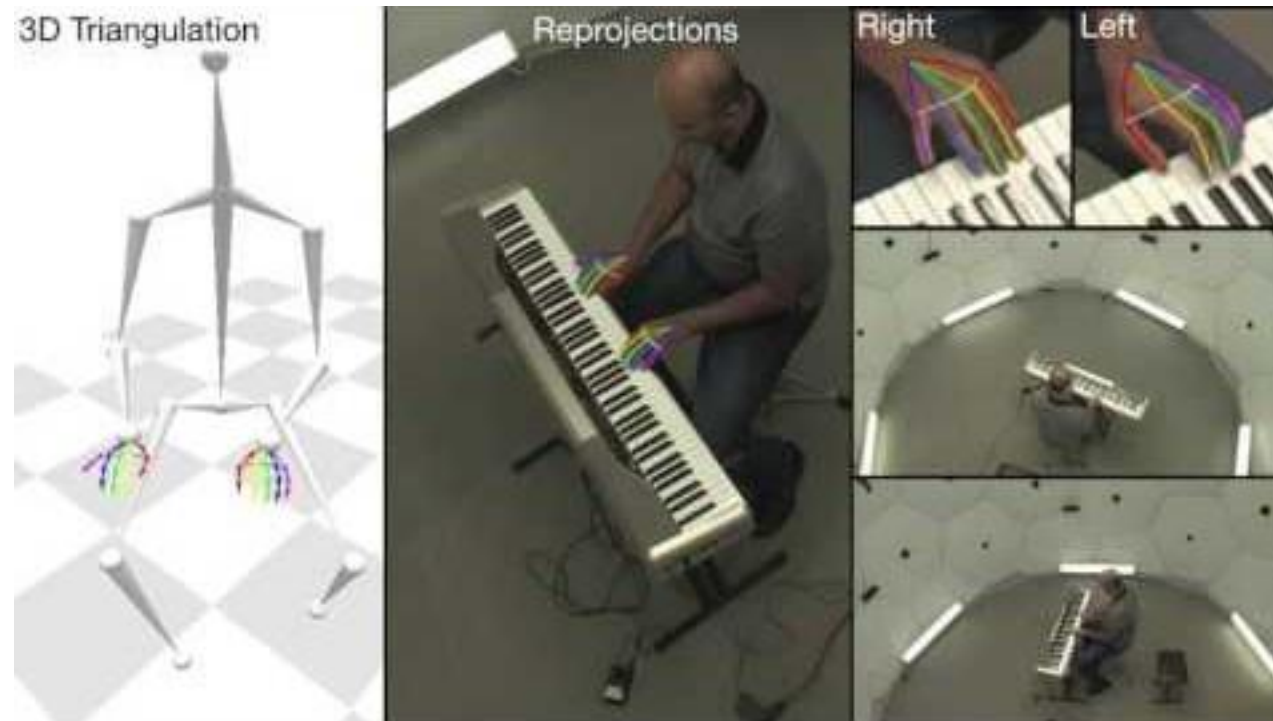
2D-3D Hand Detection



2D-3D Hand Detection

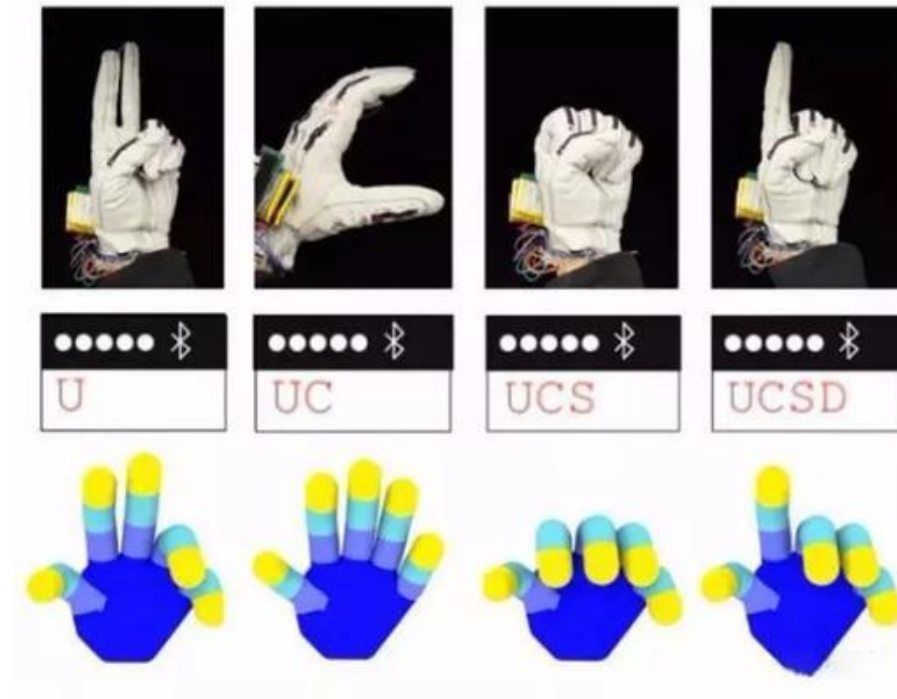


2D-3D Hand Detection



Application

Sign Language



Application

Music Education



Application

AR VR XR



(a)



(b)



(c)



(d)



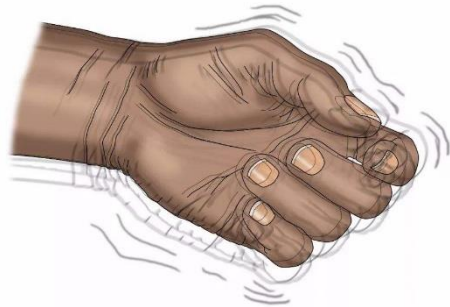
(e)



(f)

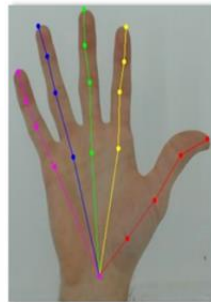
Application

Medial Area



Parkinson

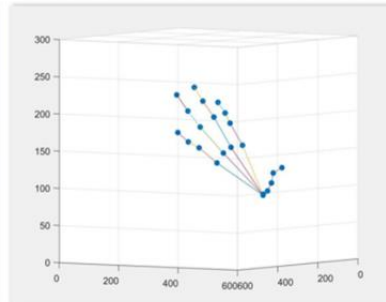
Item 21: Postural Tremor of Hands



Cam. 1

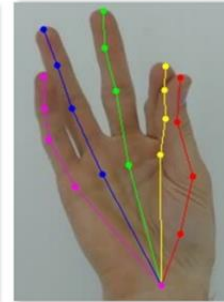


Cam. 2

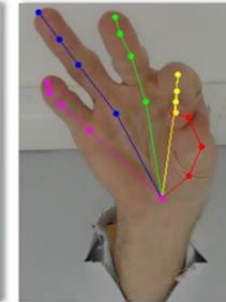


3D

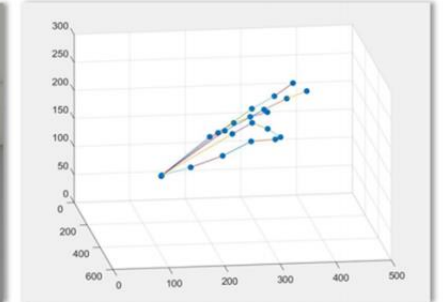
Item 23: Finger Taps



Cam. 1

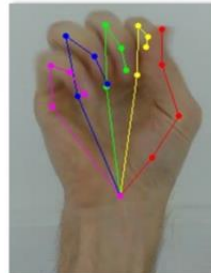


Cam. 2

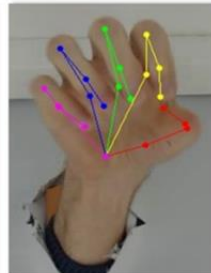


3D

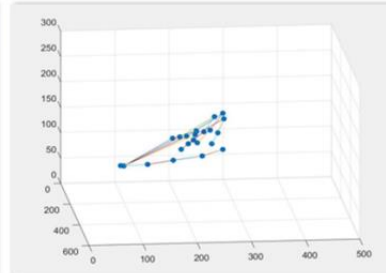
Item 24: Hand Movements



Cam. 1

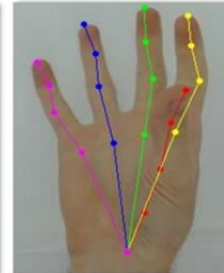


Cam. 2

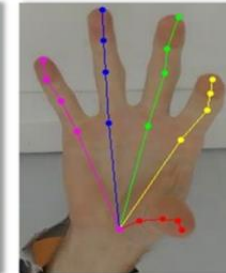


3D

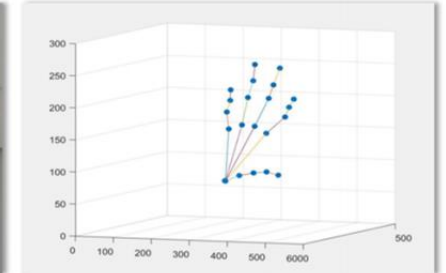
Item 25: Rapid Alternating Movements of Hands



Cam. 1



Cam. 2



3D

Q&A



Fall 2023