

# 多媒体大数据管理与分析

## 多媒体大数据的发展与相关研究问题

第三讲

刘家俊  
中国人民大学信息学院  
2017年7月

# 多媒体大数据的发展与挑战

- ❖ 多媒体大数据时代的到来
- ❖ 多媒体大数据的技术挑战
- ❖ 多媒体大数据研究与应用



# 相关计算机研究领域

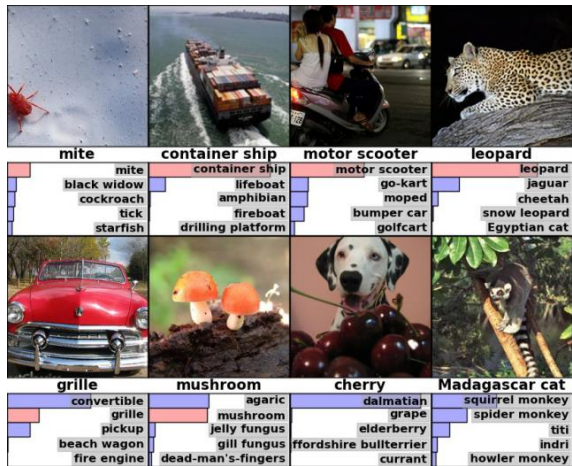


图片来源：Hubert Project

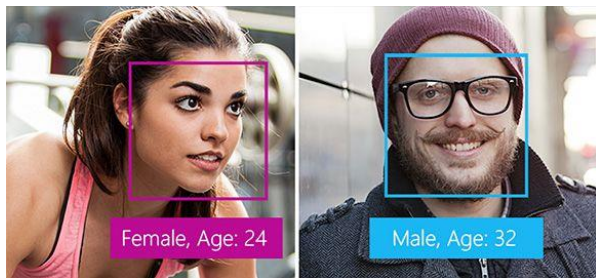
# 检测、分类与检索

## ❖ 是大多数多媒体应用的技术理论基础

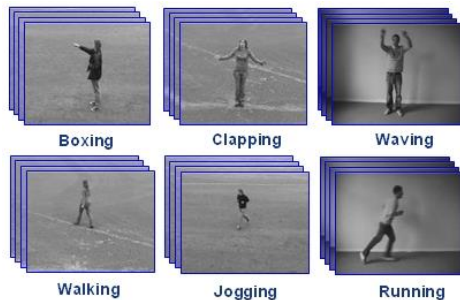
## ❖ 检测媒体中是否出现了1000类物件中的某些类



## ❖ 检测媒体中是否出现人脸，如是，给出其位置及人物的性别、年龄



## ❖ 检测视频中对象的姿势和动作类型



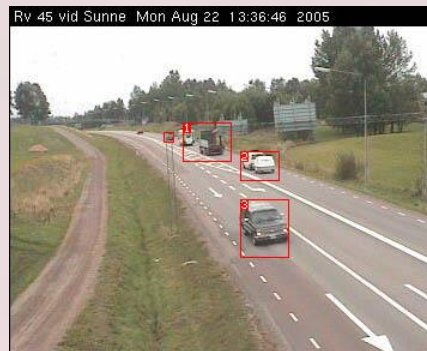
## ❖ 从数据库中检索出与查询记录内容近似的多媒体记录



图片来源:互联网

# 基于对象检测的智能视频监控

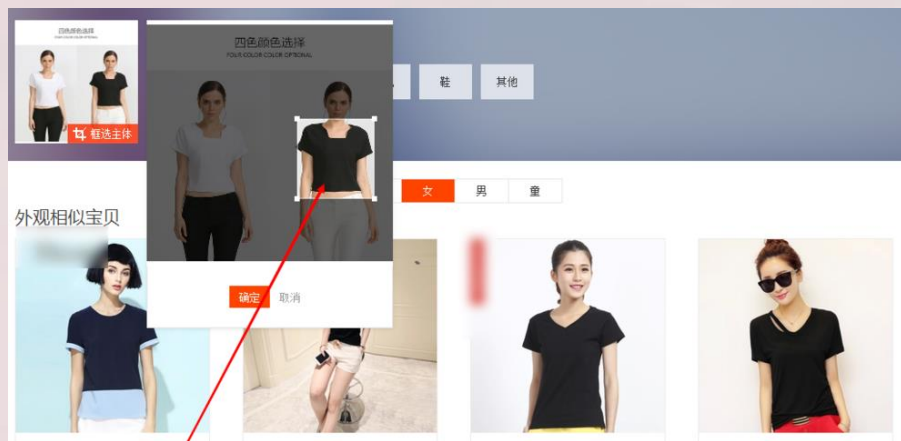
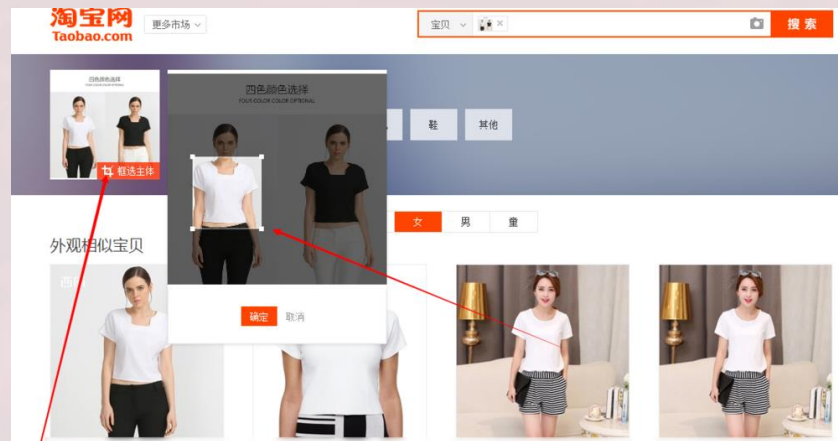
- ❖ 跨摄像头识别并追踪视频中的人、车辆、物体等并识别性别、衣着颜色、车牌、车型、等属性以及行进轨迹





# 拍照搜商品

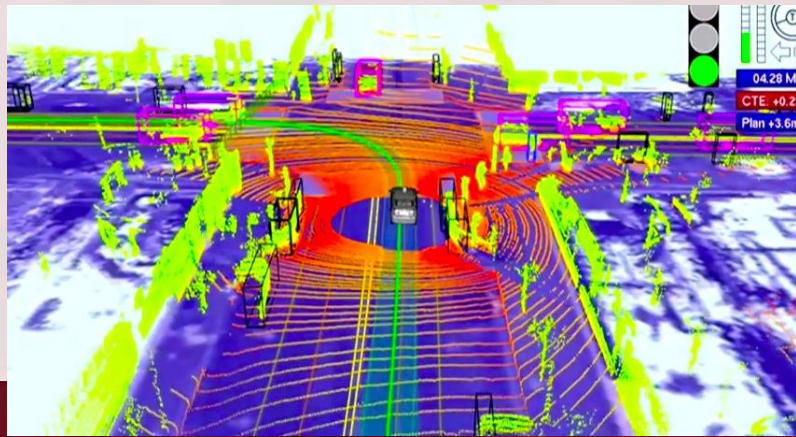
❖ 根据上传照片中框选部分的内容找到外观上相同或相似的商品



# 自动驾驶

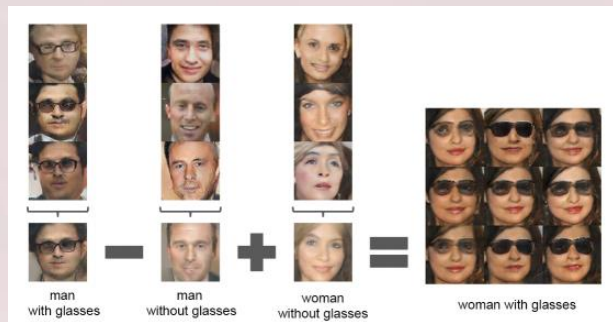
- ❖ 综合使用视觉、雷达、激光雷达等信号采集设备感知周围环境
- ❖ 利用车内传感器采集车速、行驶方向等自身状态
- ❖ 通过学习海量人类驾驶行为数据，训练自动驾驶模型
- ❖ 局限：受极端条件影响，仍不能完全替代人类驾驶员
  - 2016年5月7日，一辆特斯拉Model S车辆在自动驾驶模式下发生事故
  - 阳光照射挂车白色面板造成强烈的反光，使得自动驾驶系统和人都未能识别到前方的车辆

图片来源:互联网



# 图像风格化及生成

- ❖ 使用生成对抗网络生成类型图片



- ❖ 利用感知损失函数(perceptual loss function) 训练视觉风格转换神经网络



Radford et al., "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", Arxiv 2016

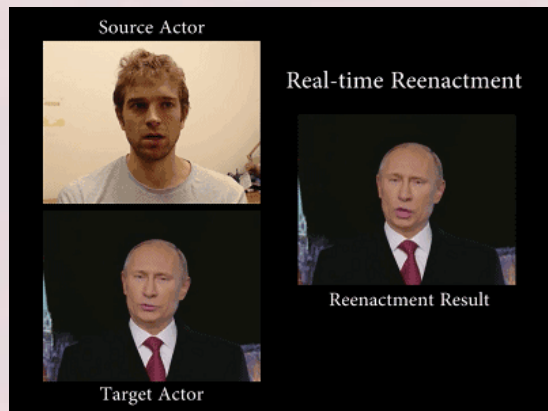
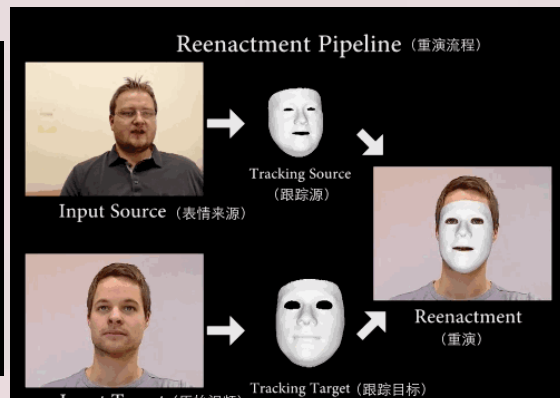
Johnson et al., "Perceptual losses for real-time style transfer and super-resolution." Arxiv 2016





# 音视频生成

- ❖ 通过学习目标的声音及脸部表情模型，实时以演员动作和神态为基础生成目标人物音视频



谢谢！

