# 计算机视觉与模式识别

苏 远 歧

新型计算机研究所

## 第1章 绪论

计算机视觉是一门研究如何使机器“看”的科学，更进一步的说，就是是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。

1. 课程简介
   - 1.1 计算机视觉的定义
   - 1.2 模式识别的定义
   - 1.3 深度学习的定义
   - 1.4 计算机视觉的应用
2. 课程考核方式
   - 2.1 课程的模块
   - 2.2 考核方法
   - 2.3 参考书目
   - 2.4 答疑与联系方式

从一个失败的案例开始!

## 特斯拉（Tesla）的事故


- 2016年5月，在美国佛罗里达州US Highway 27A公路，一辆Tesla Model S型轿车直接与大货车侧面底部相撞，导致驾驶员死亡

• 雷达穿过厢式货车的地盘，未识别障碍物！


• 视觉系统去识别厢式货车的侧面


## 特斯拉的事故和计算机视觉

• 特斯拉采用的Mobileye的视觉系统
• ADAS: Advanced Driver Assistance Systems

## 特斯拉的事故和计算机视觉


- 特斯拉采用的Mobileye的视觉系统
- ADAS: Advanced Driver Assistance Systems

*Mobileye的人接受采访：*

我们曾经警告过特斯拉，不要过于激进地推进无人驾驶，不要过于依赖视觉系统的结果。

这个失败的案例告诉我们什么？

*尽管不够准确*

计算机视觉是可以帮助我们去认识这个世界！

*计算机视觉是可以帮助我们去认识这个世界！*

但是它不够准确

*尽管不够准确*

计算机视觉是可以帮助我们去认识这个世界！

一、课程简介

## 1 什么是计算机视觉呢?

**Computer vision** is an interdisciplinary field that deals with how computers can be made for gaining high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.

*https://en.wikipedia.org/wiki/Computer_vision*

*通俗地讲:*

- 计算机视觉是一门研究如何使机器“看”的科学，
- 更进一步的说，就是是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。

*https://baike.baidu.com/item/计算机视觉/2803351?fr=aladdin*

计算机视觉与模式识别

计算机视觉与模式识别

计算机视觉与模式识别

## 2 模式识别是什么?

**Pattern recognition** is a branch of machine learning that focuses on the recognition of patterns and regularities in data, although it is in some cases considered to be nearly synonymous with machine learning.

*https://en.wikipedia.org/wiki/Pattern_recognition*

## 模式识别干什么?

(1) 分类 (2) 回归 (3) 聚类

## 2.1 分类

图中是否存在猫?

## 分类

图中是否存在猫? → 有猫 → 1

## 分类

图中是否存在猫? → 有猫 → 1

图中是否存在猫? → 没猫 → 0

分类

f(  ) → 1

分类

f(  ) → 0

二类分类

f(  ) ∈ {0,1}

多类分类

图中是猫、狗、汽车、还是行人?

## 多类分类

f(  ) ∈ {猫, 狗, 汽车, 行人}

## 多类分类

图中是猫、狗、汽车、还是行人?

| | |
|---|---|
| 猫 | 1 |
| 狗 | 2 |
| 汽车 | 3 |
| 行人 | 4 |
| ..... | ..... |

## 多类分类

f(  ) ∈ {1,2,3,4}

## 2.2 回归

图中的猫多重?

1斤
1.5斤
2.1斤
3.1斤
5.5斤
7.4斤

回归

$f($  $) \in \{猫, 狗, 汽车, 行人\}$

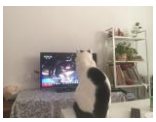
回归

$f($  $) \in \underbrace{(0, \infty)}_{\text{连续的值}}$

回归

$f($  $) \in \underbrace{8.5}_{\text{连续的值}}$

回归

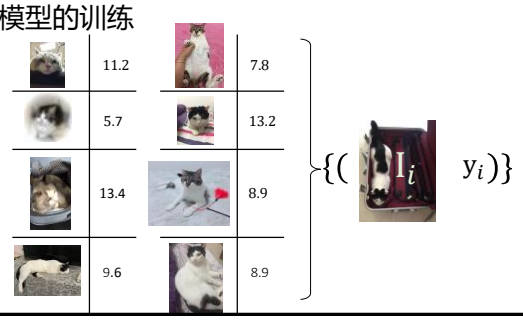
$f($  $) \in \underbrace{9.3}_{\text{连续的值}}$

如何获取分类和回归的函数

?

f(  )

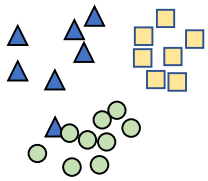
模型的训练




模型的训练

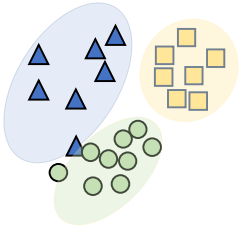
f( ) ← {(  $y_i$)}

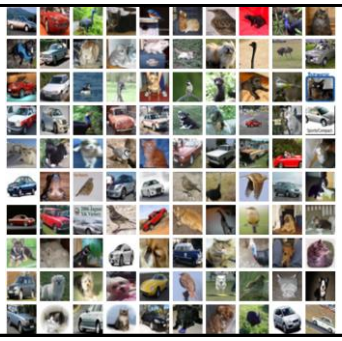
2.3 聚类 给定一组无标记的数据，我们要把它分成若干组



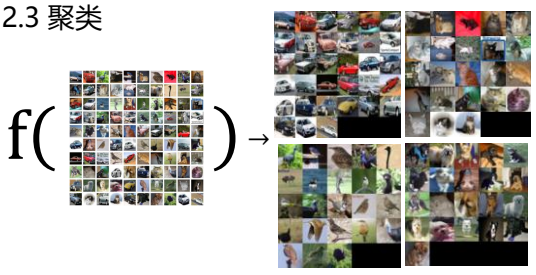
## 2.3 聚类

给定一组无标记的数据，我们要把它分成若干组



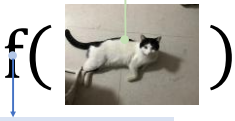


## 2.3 聚类



## 这门课中，模式识别与计算机视觉的关系

计算机视觉是我们的对象



模式识别是我们的工具

- 我们利用模式识别来研究计算机视觉

## 3 深度学习

**Deep learning** (also known as **deep structured learning** or **hierarchical learning**) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms.

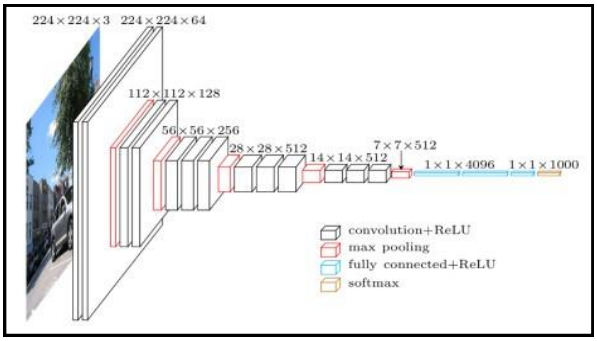
*https://en.wikipedia.org/wiki/Deep_learning*

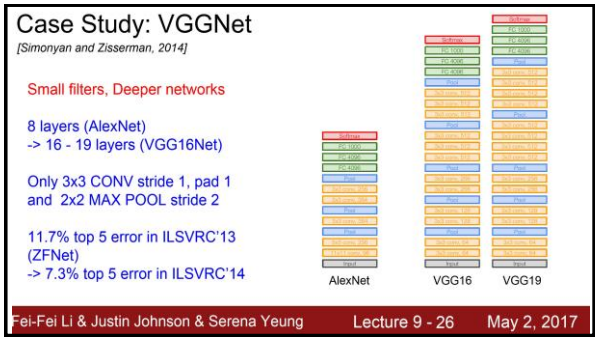
## 3 深度学习

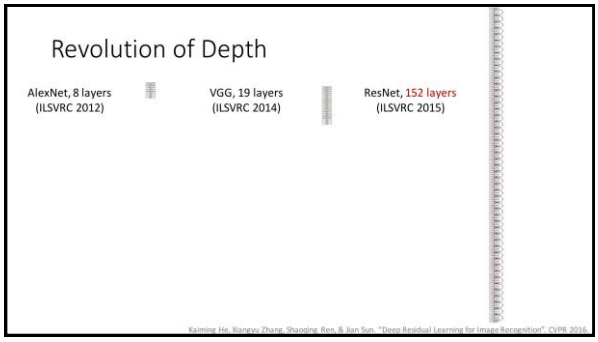
**Deep learning** (also known as **deep structured learning** or **hierarchical learning**) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms.

*https://en.wikipedia.org/wiki/Deep_learning*

A **deep neural network (DNN)** is an artificial neural network (ANN) with multiple hidden layers between the input and output layers.[1]



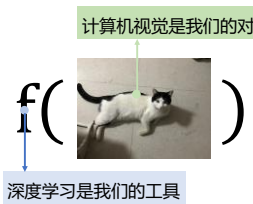



## 3 深度学习

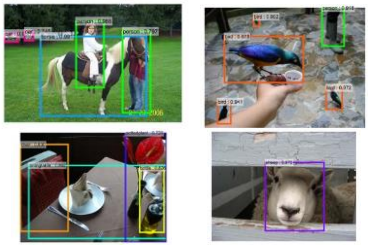
- ICCV2015深度学习超过45%
- 目前CVPR2018年估计有超过70%的论文都是基于深度学习的

http://www.58pic.com/sucai/19015904.html

## 这门课中，深度学习与计算机视觉的关系



## 4 计算机视觉能干什么?

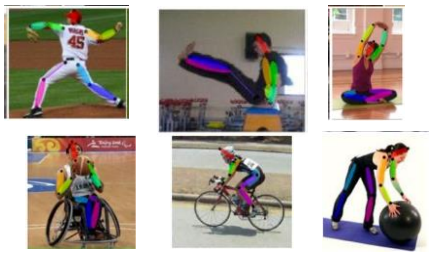


物体检测

- Faster R-CNN [Ren et al. 2015]

语义分割

- CRF-as-RNN [Zheng et al. 2015]

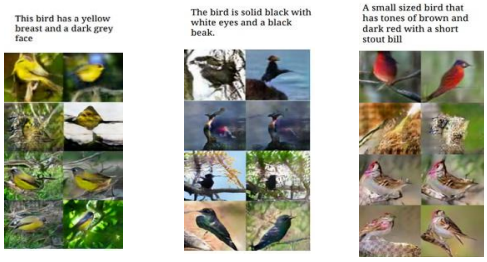


姿态识别

Convolutional Pose Machines [Wei et al. 2016]


Image Captioning

Neural Image Caption Generator [Vinyals et al. 2015]




Text to Image generation · Learning What and Where to Draw [Reed et al. 2016]


玩游戏 Google Deep Mind, Alpha Go, Alpha Zero,

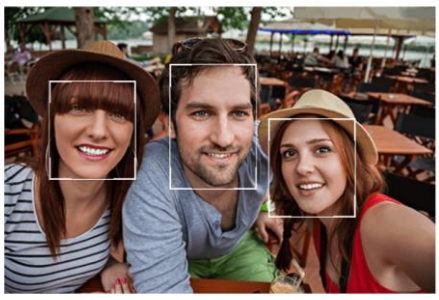
看看我们中国的公司

## 4 计算机视觉能干什么?

- 旷视科技
- 格林深瞳
- 商汤科技
- Momenta
- DeepAi


## 旷世科技

- 人脸识别
- 人体识别
- 文字识别
- 图像识别



能力介

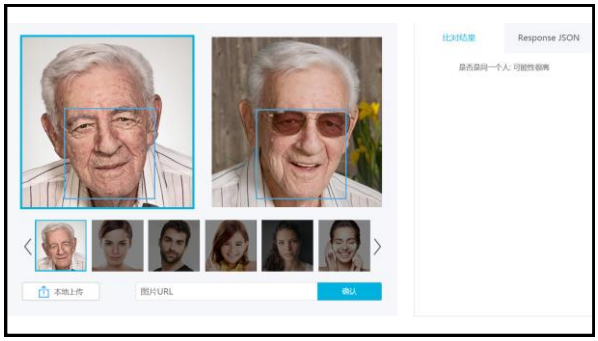
检测并定位图

人脸检测是对
便日后使用，


## 能力介绍

分析两张人脸属于同一个人的可能性大小。使用Face++的人脸比对技术，您会得到一个置信度分数和相应的阈值，以便评估相似度。

人脸比对技术被广泛用于基于人脸的用户认证以及人证合一。基于世界领先的深度学习算法，以及多年的行业落地应用打磨，Face++为各类场景提供精准、可靠的解决方案。目前已为300多家金融机构提供服务。




## 商汤科技

- 人脸与人体分析技术
- 通用与专业图像识别
- 海量视频理解与挖掘
- 图像视频处理增强

SLAM与3D视觉
机器人传感与控制
无人驾驶
深度学习平台

人脸检测与跟踪

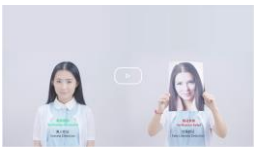
人脸特征点定位

人脸身份认证

人脸属性

人脸聚类

真人检测

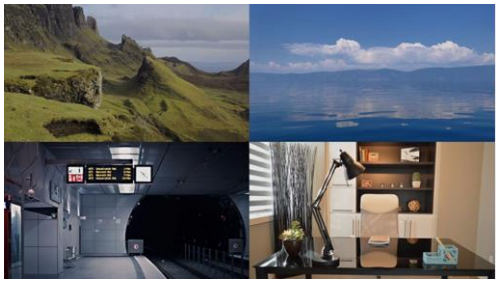


身份认证

美颜/美妆

身体关键点


一般物体的识别



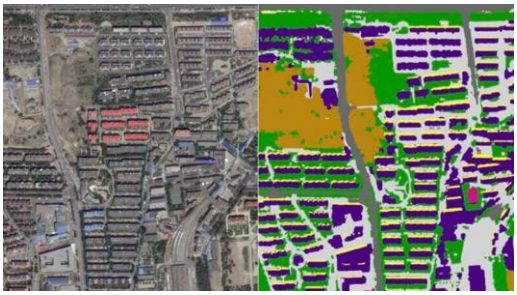
场景识别

车型识别


文字识别

文字识别


车牌识别



遥感图像中的实体识别

视频内容的审核

短视频的标签

视频内容的结构化

视频内容的分析

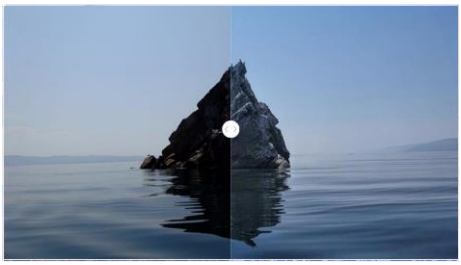
去雾

超分辨率

修复

暗光增强

失焦修复

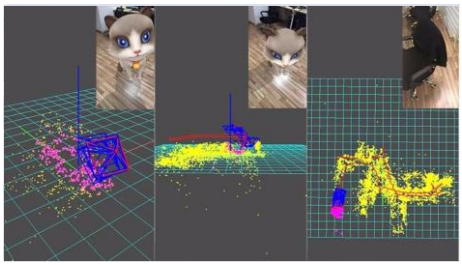


单图HDR

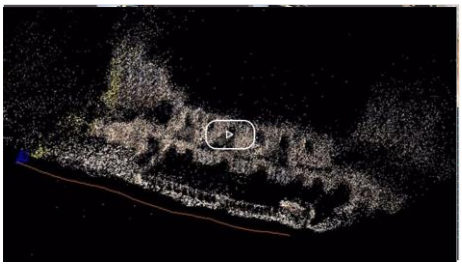
再聚焦

滤镜

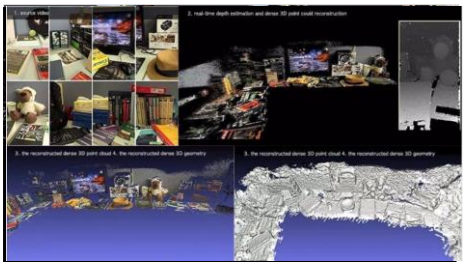
平面/三维物体的识别



SLAM

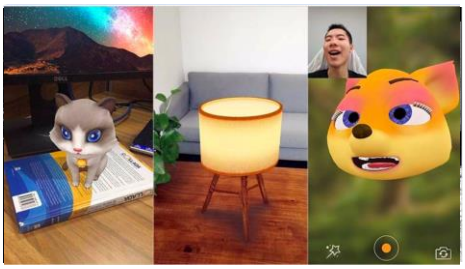


SfM



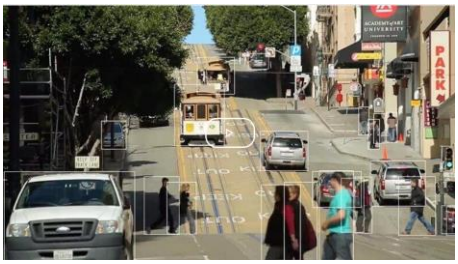
深度估计与三维重构

光照估计



轻量级的AR/VR

机器人抓取物体

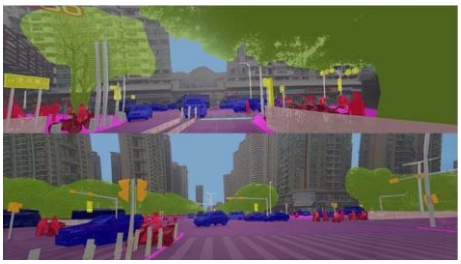



行人车辆检测

车辆属性检测

行人运动分析

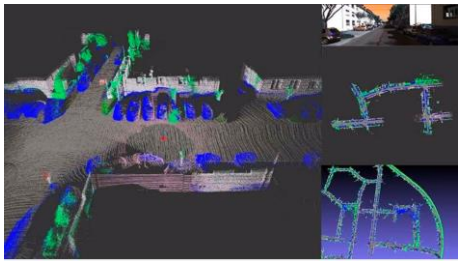
车道线检测



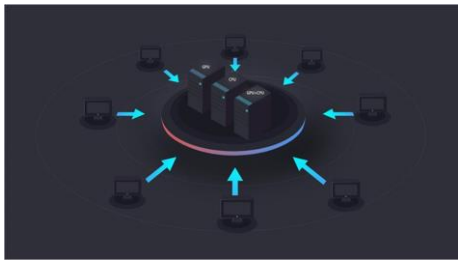
场景理解

交通指示牌识别

大规模场景的实时定位



高精度三维场景重建



深度学习的平台

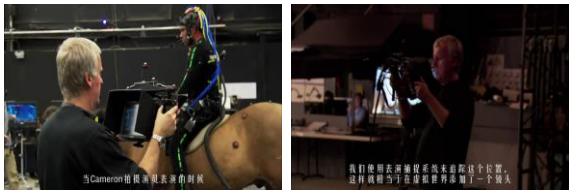
## 4 计算机视觉能干什么?

- 旷视科技
- 格林深瞳
- 商汤科技
- **Momenta**
- DeepAi

影视制作

虚拟中添加真实


真实中添加虚拟

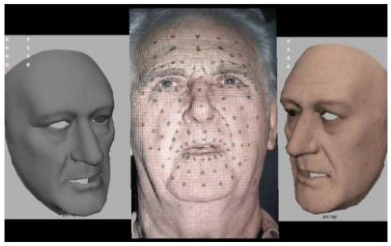


3D摄像机

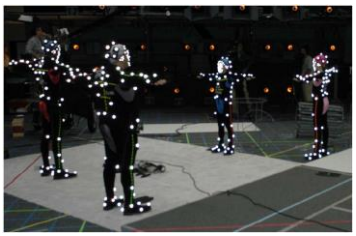
Blending

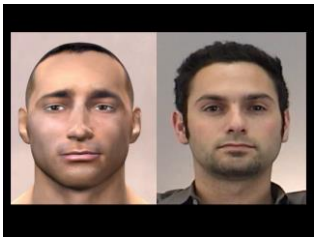
Morphing

Motion Capture



Monster House

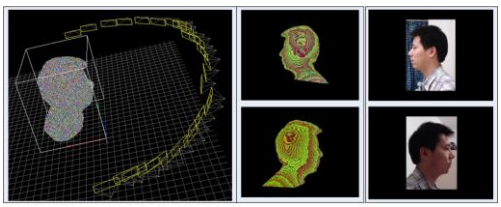




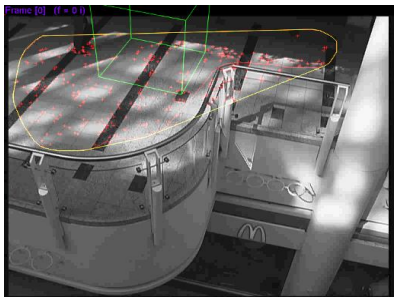
[George Borshukov SIGGRAPH 2004]

骇客帝国III（The Matrix Revolutions）2003年11月5日





目标的稠密重建及虚拟摄像机

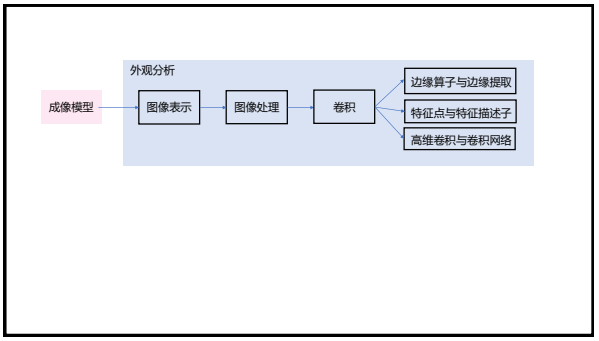


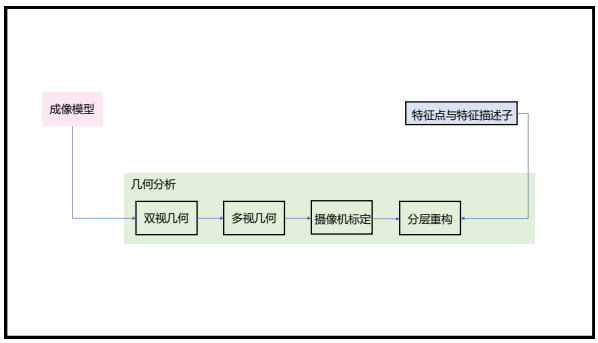
## 二、 考核方式

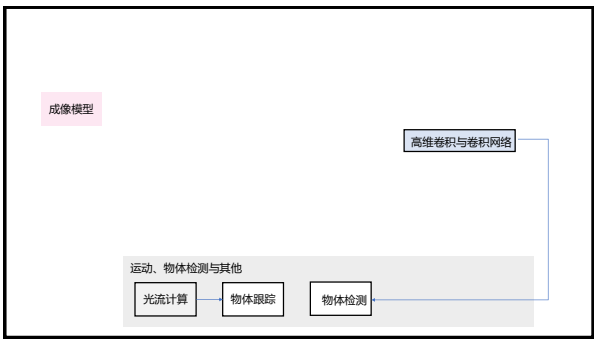
## 1 课程的模块

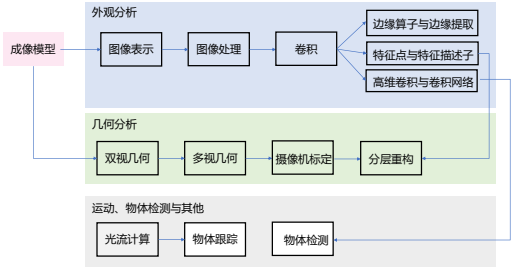
| 课内讲授36学时 | 实验8学时 |
| --- | --- |
| 外观分析 | 物体检测 |
| 几何分析 | 视觉SLAM |
| 运动、物体检测与其他 | 物体跟踪，三维物体的跟踪 |


外观分析
成像模型
几何分析
运动、物体检测与其他

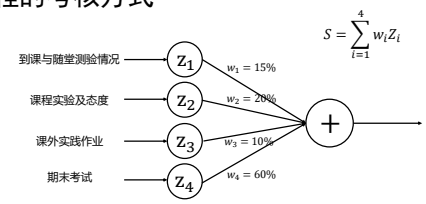

外观分析
成像模型
图像表示
图像处理
卷积
边缘算子与边缘提取
特征点与特征描述子
高维卷积与卷积网络


成像模型
特征点与特征描述子
几何分析
双视几何
多视几何
摄像机标定
分层重构


成像模型
高维卷积与卷积网络
运动、物体检测与其他
光流计算
物体跟踪
物体检测




## 2 课程的考核方式




## 作业内容

- 平时作业
  - 课堂上讲授的基本概念和原理
- 课外实践作业
  - 具体内容待定，属于实验的扩展内容
  - 题材不限
    - 室内的视觉导航、、
    - 基于深度学习的应用；

147

## 作业提交方式

- 上机作业
  - 每个班2次上机，每次4小时
  - 必须在机房提交，同时给出成绩
- 平时作业
  - 课堂内完成（10—15分钟）
- 课外实践作业
  - 利用课余时间完成，在指定的时间前提交到邮件

## 3 参考书目

[1].**计算机视觉：算法与应用**（Computer Vision: Algorithms and Applications）；[美] 塞利斯基（Richard Szeliski） 著；艾海舟，兴军亮等译；出版社：清华大学出版社，2012。

[2]. **计算机视觉——一种现代方法（第二版）** [Computer Vision: A Modern Approach, Second Edition]；作者：David, A. Forsyth, Jean Ponce；译者：高永强等；出版社：电子工业出版社；2017。

[3].**视觉SLAM十四讲：从理论到实践**；高翔，张涛等著；出版社：电子工业出版社，2017。

[4].**深度学习** (deep learning)；作者：[美] Ian，Goodfellow，[加] Yoshua，Bengio，[加] Aaron, Courville；译者：赵申剑等；出版社：人民邮电出版社，2017。

[5]. http://www.arxiv.org/. Topic: Computer Vision and Pattern Recognition

[6]. Matlab 2018a, Computer Vision Help Files.

[7] OpenCV

## 4 课程联系方式

- 姓　名：苏远歧
- 办公室： 西一楼403
- 手　机：13700230551
- Email:　yuanqisu@qq.com
  370021103@qq.com