



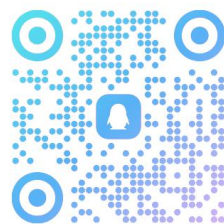
# 计算机视觉

陈飞

福州大学计算机与大数据学院

chenfei314@fzu.edu.cn

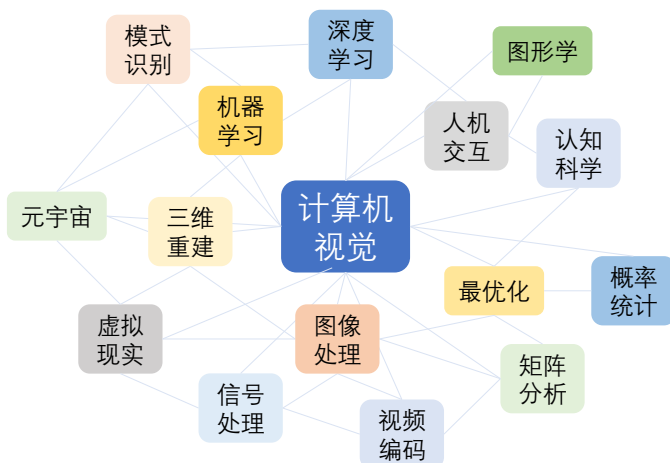
计算机视觉202502-陈...  
群号: 1049801850



## 课程内容



- 计算机视觉概述
- 视觉特征表达
- 视觉特征提取
- 视觉特征学习
- 图像分类
- 图像分割
- 目标检测
- 物体跟踪
- 多图拼接
- 三维重建



# 计算机视觉概述



## 1.1 计算机视觉简介

### 1.1.1 计算机视觉的发展历史

### 1.1.2 主流学术会议和期刊

## 1.2 计算机视觉的应用

### 1.2.1 智能监控

### 1.2.2 医学影像智能分析

### 1.2.3 智能机器人

### 1.2.4 车载视觉系统

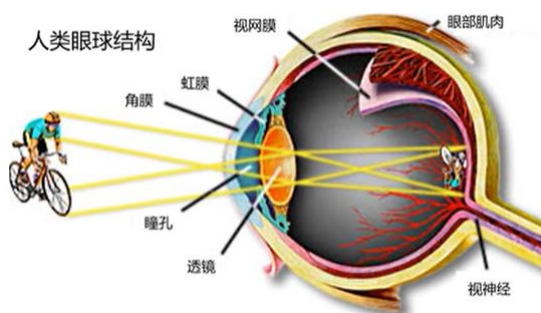
## 1.3 计算机视觉面临的挑战

3

# 1.1 计算机视觉简介



## • 人类视觉 (Human Vision)



1.光线进入：光线先通过角膜。

2.调节光量：通过瞳孔（由虹膜控制大小）。

3.初步聚焦：通过晶状体（由睫状肌调节厚薄）。

4.透射支撑：穿过玻璃体。

5.成像感光：光线最终落在视网膜上，形成倒立缩小的实像。

6.信号传导：视神经将信号传给大脑，大脑将其“转正”并形成意识。

4

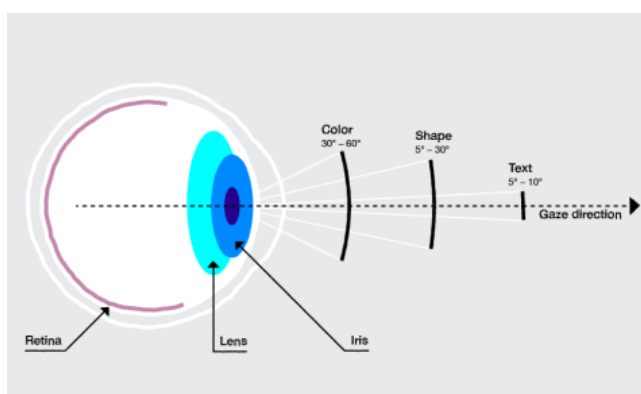
# 人眼如何观看事物？



- 可能性1:  
眼睛快速移动，不断扫描环视，补齐不清晰和看不到的区域。
- 可能性2:  
大脑对不清晰的画面进行“脑补”补齐不清晰或看不到的区域

5

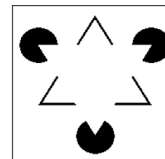
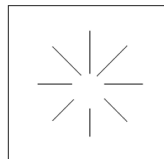
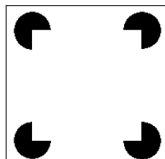
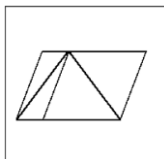
# 人类视觉 (Human Vision)



• **中心 vs. 周边**：人眼并不是整个视网膜都能看到相同的东西。只有视网膜中心（黄斑/中央凹）负责高分辨率、彩色的精细视觉（如阅读、认脸）。而视网膜周边主要负责感知运动、明暗和方位。

6

## 人眼如何看到事物？



- 可能性1:  
眼睛快速移动，不断扫描环视，补齐不清晰和看不到的区域。
- 可能性2:  
大脑对不清晰的画面进行“脑补”补齐不清晰或看不到的区域

7

## 人类视觉



- 相比看见，人类视觉系统更善于发现画面背后的故事

*“Vision can be understood as an information processing task which converts a numerical image representation into a symbolic shape-oriented representation”*

*- David Marr, 1982*

“视觉可以被理解为一个信息处理任务，将数值图像转变为一种面向形状的符号表示”

图像



意义

8

# 人眼 vs. 照相机



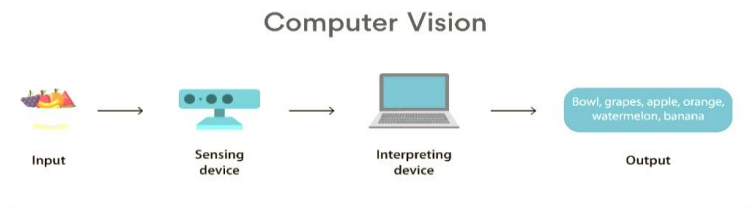
特征	人眼	照相机
镜头	可变形的晶状体（物理变焦）	刚性的玻璃/塑料镜片（位移变焦）
光圈	瞳孔（由虹膜控制，自动调节）	光圈叶片（可由相机或手动控制）
对焦点	单次只能对准一个平面，但扫视极快	可全画面均匀（取决于镜头设计）
感光元件	视网膜（中央凹高精度，周边低精度）	传感器（全局均匀像素）
动态范围	极高，能适应极大亮度跨度	有限，宽容度逐年提升但仍有差距
图像处理	大脑实时降噪、补盲点、自动白平衡	处理器运算
电源	生化能量（葡萄糖和氧气）	电池

9

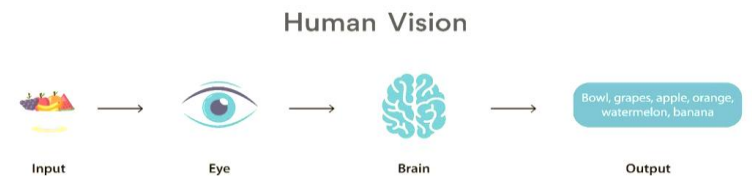
# 计算机视觉 (Computer Vision)



- 通过模拟人类的视觉系统，计算机视觉技术能够从图像或视频中提取信息、识别物体、理解场景，并做出相应的决策。



- 计算机视觉的应用非常广泛，包括自动驾驶、医疗影像分析、人脸识别、增强现实、安防监控等。



10



## 1.1.1 计算机视觉的发展历史

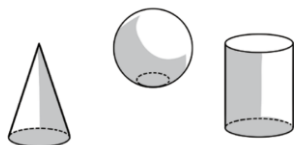
1. 早期阶段 (1960s-1970s)
  - 劳伦斯·罗伯茨提出让计算机从一张二维照片中提取出简单的三维结构信息
  - 简单的图像处理和分析上，如边缘检测、形状识别等
2. 理论发展阶段 (1980s) :
  - 1980年代，理论基础逐渐形成(几何学和统计学方法的应用)
  - 图像分割、特征提取、立体视觉等技术得到了发展
3. 机器学习与统计方法 (1990s-2000s) :
  - 机器学习方法开始应用于计算机视觉
  - 特征提取方法 (如SIFT、HOG) 成为主流，人脸检测的快速发展
4. 深度学习时代 (2010s至今) :
  - 2012年，AlexNet在ImageNet竞赛中取得突破性成绩
  - 卷积神经网络推动了图像分类、目标检测等任务的快速发展
  - 生成对抗网络 (GAN)、Transformer等新技术

11

## 早期阶段 (积木时代)



- 如何让计算机从二维图像中理解简单的三维世界，并**初步建立了学科的视觉理论框架**。



罗伯茨成功地让计算机识别出二维图像中的三维结构和空间布局

**1966年：**麻省理工学院的马文·明斯基 (Marvin Minsky) 教授安排了一个暑期项目，让本科生将摄像头连接到计算机上，试图让计算机直接理解它所看到的场景。



劳伦斯·吉尔曼·罗伯茨 (Lawrence Gilman Roberts)

罗伯茨 (1937—2018)，美国工程师，2001年因对互联网发展的贡献而荣获德拉普尔奖 (美国工程学界最高奖项之一)，五角大楼高级研究计划局 (ARPA) 的主管，首位设计和管理第一个分组网络 ARPANET 的互联网先驱。1965年，他在《三维固体的机器感知》一书中描述了从二维图片中推导三维信息的过程，成为计算机视觉的前导之一，开创了理解三维场景为目的的计算机视觉研究。

12

## 理论发展阶段（视觉计算理论）



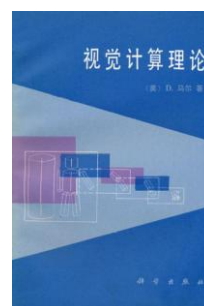
- David Courtnay Marr (1945-1980) 是计算神经科学领域的先驱之一。1973年，他应Marvin Minsky与Seymour Papert的邀请加入MIT人工智能实验室，从事视觉相关研究，并**开创性地提出了三维表达思想，使计算机视觉进入了一个新的时代**。MIT出版社在1982年出版了他的遗作《Vision: A Computational Investigation into the Human Representation and Processing of Visual Information》。



### Artificial Intelligence—A Personal View

D. Marr  
The Artificial Intelligence Laboratory,  
Massachusetts Institute of Technology,  
345, Technology Square, Cambridge, MA 02139, U.S.A.

Recommended by B. Meltzer

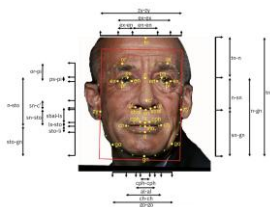
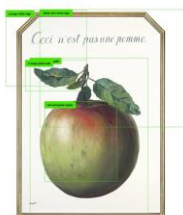


13

## 机器学习与统计方法



- 观念转变：从“重建”到“识别与学习”——让计算机能够识别和理解图像中的内容。
  - 确立了“视觉即学习”的范式：**将视觉问题从“设计特定算法”转变为“从数据中学习模型”**。
  - 引入了强大的数学工具：概率图模型（MRF, CRF）、核方法（SVM）、鲁棒统计和贝叶斯推理成为视觉研究者的标准工具箱。

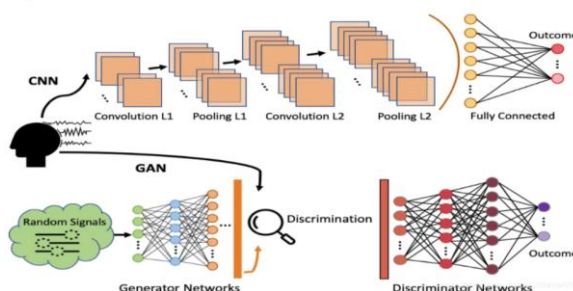


14

# 深度学习时代



- 这个时代彻底改变了视觉问题的解决方式，**将视觉识别的能力推向了前所未有的高度**。从2012年AlexNet的横空出世，到ResNet等架构的不断深化，再到如今Transformer、多模态、世界模型等新范式的百花齐放，视觉智能正以前所未有的速度演进。



15

## 1.1.2 主流学术会议和期刊



国际会议：

- IEEE International Conference on Computer Vision and Pattern Recognition (**CVPR**)
- International Conference on Computer Vision (**ICCV**)
- European Conference on Computer Vision (**ECCV**)
- IEEE International Conference on Robotics and Automation (ICRA)
- IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
- IEEE International Conference on Pattern Recognition (ICPR)



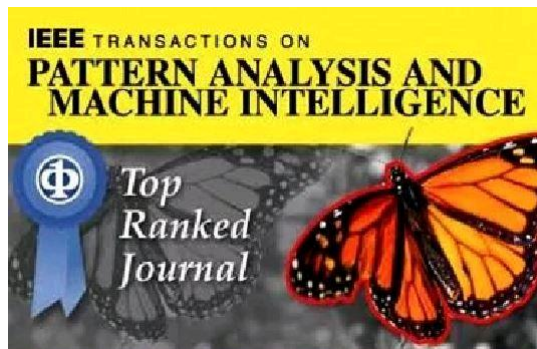
16

## 1.1.2 主流学术会议和期刊



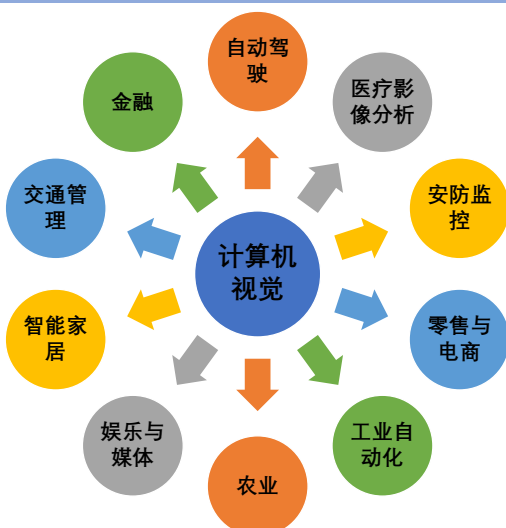
### 国际期刊:

- IEEE Transactions on Pattern Analysis and Machine Intelligence (**PAMI**)
- IEEE Transactions on Image Processing (**TIP**)
- International Journal of Computer Vision
- IEEE Transactions on Circuits and Systems for Video Technology
- IEEE Transactions on System, Man and Cybernetics, Part B
- IEEE Transactions on Autonomous Mental Development
- IEEE Transactions on System, Man and Cybernetics, Part A
- Pattern Recognition



17

## 1.2 计算机视觉的应用

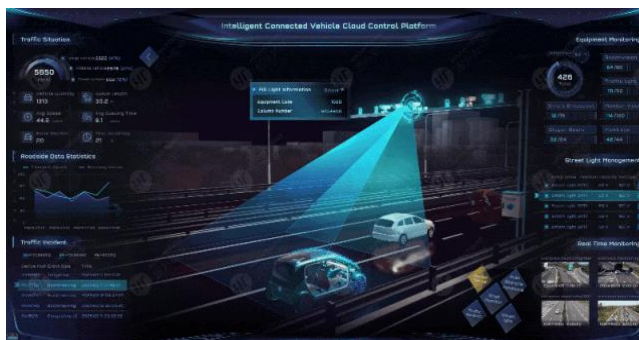


### 中国计算机视觉应用市场规模趋势

年份	市场规模 (亿元人民币)	年均复合增长率 (CAGR)	主要驱动因素
2022	500	-	政策支持、安防需求、技术进步
2023	650	30%	金融科技、医疗影像需求增长
2024	850	30.8%	零售智能化、自动驾驶发展
2025	1,500	32.5%	5G普及、边缘计算应用
2026	1,950	30%	多模态AI技术融合
2027	2,500	28.2%	农业、教育领域渗透
2028	3,000	20%	行业标准化、数据隐私规范
2029	3,500	16.7%	技术成熟、应用场景扩展
2030	4,000	14.3%	全面智能化、新兴领域需求

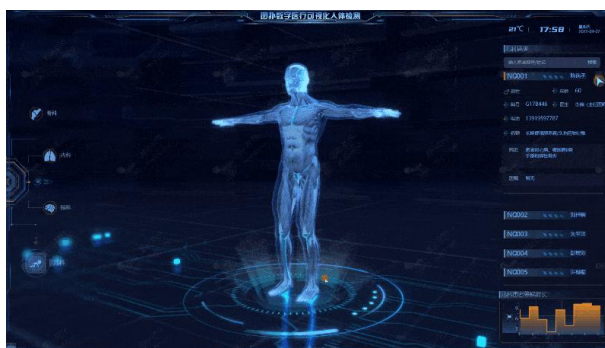
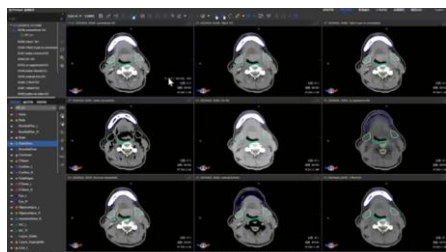
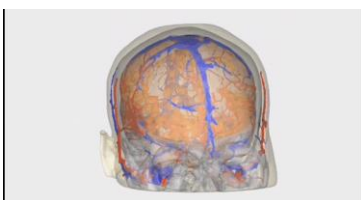
18

## 1.2.1 智能监控



19

## 1.2.2 医学影像智能分析



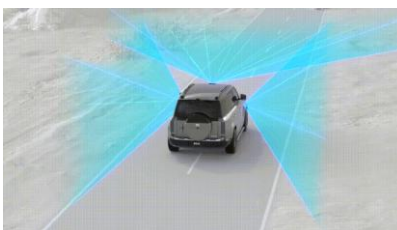
20

## 1.2.3 智能机器人



21

## 1.2.4 车载视觉系统



22

## 1.3 计算机视觉面临的挑战



### 数据质量/数量

- 标注成本高
- 多样性不足
- 数据噪声

### 模型泛化能力

- 过拟合
- 跨域适应

### 计算资源

- 高计算需求
- 实时性要求

### 复杂场景理解

- 遮挡与变形
- 光照变化
- 背景干扰

### 语义理解

- 高层次理解
- 多模态融合

### 鲁棒性/安全性

- 对抗攻击
- 模型解释性

### 伦理与隐私

- 隐私保护
- 偏见/公平性

### 实时性与效率

- 实时处理
- 边缘计算

### 3D视觉

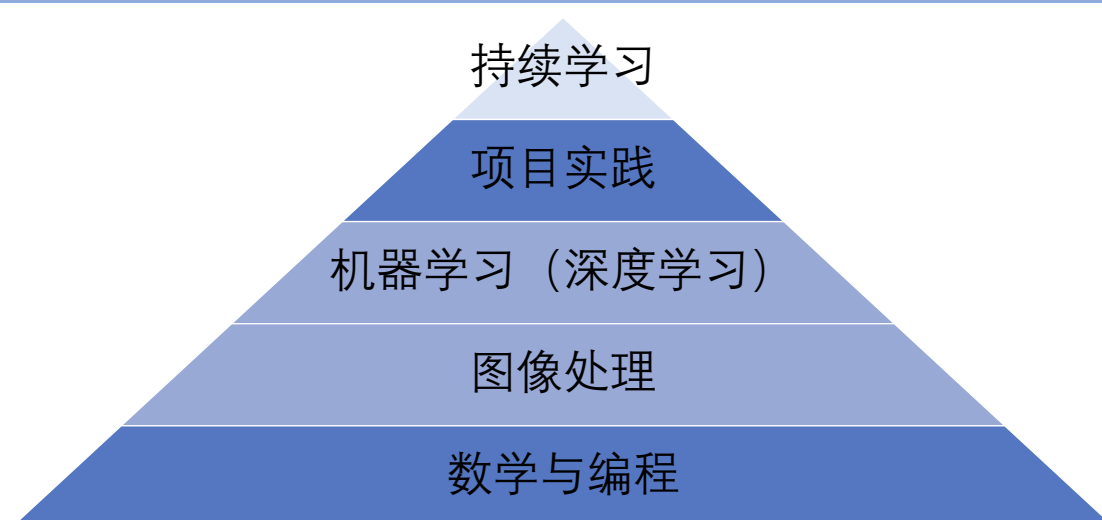
- 深度估计
- 点云处理

### 多任务学习

- 任务冲突
- 资源共享

23

## 1.4 如何学习计算机视觉？



24



Thank you!

E-mail: [chenfei314@fzu.edu.cn](mailto:chenfei314@fzu.edu.cn)