

计算机视觉 (Chapter 1 作业)

智科三班 严中圣 222020335220177

2022 年 10 月 5 日

1.3 利用网上的数据库, 统计调查一下计算机视觉近年来的发展趋势和特点, 比较重要的事件有哪些。

解.

计算机视觉近四十年来的发展大致经历了四个阶段:

- (1) 第一阶段是马尔计算视觉。1982 年大卫马尔 (David Marr) 的《视觉》一书在计算机视觉领域中起到了关键性的作用, 它标志着计算机视觉正式成为一门独立的学科。马尔的计算视觉分为三个层次: 计算理论、表达和算法以及算法实现。马尔认为, 大脑的神经计算和计算机的数值计算没有本质区别, 而从现在神经科学的进展看, “神经计算” 与数值计算在有些情况下还是会产生本质区别。
- (2) 主动视觉与目的视觉是第二阶段, 学术界几位教授对马尔视觉计算理论提出了反对意见, 认为缺乏主动性、目的性和应用性。但由于这段时期没有过多进展, 对后续计算机视觉的发展影响不大, 因此很多时候没有把这一阶段单独列出介绍。
- (3) 多视几何和分层三维重建是计算机视觉发展的第三阶段, 其中代表人物包括法国的 O · Faugeras, 澳大利亚国立大学的 R · Hartely 和英国牛津大学的 A · Zisserman, 在这方面的研究重点是如何快速、鲁棒地重建大场景。
- (4) 最后来到了当代计算机视觉的阶段, 基于深度学习的视觉。在此阶段中, 文献大体上分为两个个阶段: 以流形学习为代表的子空间法和目前以深度神经网络和深度学习为代表的视觉方法。

从长远发展趋势来看, 目前计算机视觉还未进入技术发展的下半场, 未来随着计算机视觉技术在人脸识别上性能的不断突破, 限定场景识别准确率将不断提升, 愈来愈多的对象识别、分类问题将会逐步实现工业化, 渗透进更多的行业应用。

整体而言, 计算机视觉技术要想更好的实现应用落地, 需要在对具体业务场景的理解之上进行针对性开发, 以提供更加全面、及时的服务, 这要求计算机视觉企业未来在重视前沿算法研发的同时, 需要进一步加强算法和商业应用的融合。

1959 年, 神经生理学家 David Hubel 和 Torsten Wiesel 通过猫的视觉实验, 首次发现了视觉初级皮层神经元对于移动边缘刺激敏感, 发现了视功能柱结构, 为视觉神经研究奠定了基础——促成了计算机视觉技术 40 年后的突破性发展, 奠定了深度学习之后的核心准则。

1965 年, Lawrence Roberts 《三维固体的机器感知》描述了从二维图片中推导三维信息的过程。——现代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。

1969 年秋天，贝尔实验室的两位科学家 Willard S. Boyle 和 George E. Smith 正忙于电荷耦合器件（CCD）的研发。它是一种将光子转化为电脉冲的器件，很快成为了高质量数字图像采集任务的新宠，逐渐应用于工业相机传感器，标志着计算机视觉走上应用舞台，投入到工业机器视觉中。

1997 年，伯克利教授 Jitendra Malik（以及他的学生 Jianbo Shi）发表了一篇论文，描述了他试图解决感性分组的问题。研究人员试图让机器使用图论算法将图像分割成合理的部分（自动确定图像上的哪些像素属于一起，并将物体与周围环境区分开来）

1999 年，David Lowe 发表《基于局部尺度不变特征（SIFT 特征）的物体识别》，标志着研究人员开始停止通过创建三维模型重建对象，而转向基于特征的对象识别。

2009 年，李飞飞教授等在 CVPR2009 上发表了一篇名为《ImageNet: A Large-Scale Hierarchical Image Database》的论文，发布了 ImageNet 数据集，这是为了检测计算机视觉能否识别自然万物，回归机器学习，克服过拟合问题，经过三年多在筹划组建完成的一个大的数据集。李飞飞说“ImageNet 改变了 AI 领域人们对数据集的认识，人们真正开始意识到它在研究中的地位，就像算法一样重要”。ImageNet 是计算机视觉发展的重要推动者，和深度学习热潮的关键推动者，将目标检测算法推向了新的高度。

2012 年，Alex Krizhevsky、Ilya Sutskever 和 Geoffrey Hinton 创造了一个“大型的深度卷积神经网络”，也即现在众所周知的 AlexNet，赢得了当年的 ILSVRC。这是史上第一次有模型在 ImageNet 数据集表现如此出色。

1.5 视频序列包括一组 2-D 图像，如何用矩阵形式表示视频序列，请具体描述一下。

解.

对于单个 2-D 图像 F_i ，假设其尺寸为 $M * N * C$ ，可用矢量表示为

$$F_i = \begin{bmatrix} f_1 & f_2 & \dots & f_N \end{bmatrix} \quad (1)$$

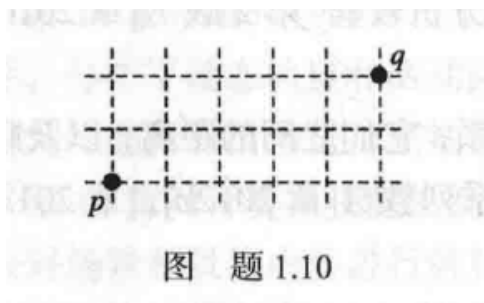
其中

$$f_j = \begin{bmatrix} f_{1j} & f_{2j} & \dots & f_{Mj} \end{bmatrix}^T, \quad j = 1, 2, \dots, N \quad (2)$$

故假设某一视频序列 V 包含图像 F_1, F_2, \dots, F_n ，则用矩阵形式表示为：

$$V = \begin{bmatrix} F_1 & F_2 & \dots & F_n \end{bmatrix}^T \quad (3)$$

***1.10** 试计算如图题 1.10 中两个像素 p 和 q 之间的 D_E 距离、 D_4 距离和 D_8 距离。



解.

$$D_E(p, q) = \sqrt{5^2 + 4} = \sqrt{29}$$

$$D_4(p, q) = 5 + 2 = 7$$

$$D_8(p, q) = \max(5, 2) = 5$$