

- [26] Agrup, G. et al.: *ibid.*, 53, 453, 1973.
[27] ———.: *ibid.*, 55, 7, 1975.
[28] ———.: *ibid.*, 59, 381, 1979.

- [29] Rorsman, H. et al.: *ibid.*, 53, 248, 1973.

[本文于1980年3月25日收到]

视觉系统的工程模拟与视觉理论的新进展

陆惠民 王今著

(中国科学院生物物理研究所)

利用模型与模拟方法研究与解释物理、化学等复杂的自然现象是人们常用的一种方法。工程模型的特点是抓住事物的主要行为和性质,首先将复杂的事物大大简化,以便对事物的研究与模拟成为可能,然后根据条件的许可与需要补充细节,使之逐渐接近于事物的真实情况。

对于视觉这样一个复杂的系统,正如美国MIT的心理学家David Marr所说,当代神经生理学与心理物理学的最显著特点,是只描述各种细胞及其它实验对象的行为,而不对这些行为进行解释。为了探索视觉系统中信息处理的机理,以及结构与功能的关系,我们认为除了进行各种水平的生物学研究以外,还应借鉴工程上常用的模型与模拟方法,从系统的输入输出关系来进行研究,把它当成一个“黑箱”。由于电子计算机技术的飞速发展,以及控制论、信息论的应用日益广泛,为模拟视觉系统提供了实验技术和理论基础。国民经济与国防上的需要,又促进了视觉系统工程模拟的开展。工程模拟,不仅要提出理论模型,还要建立视觉系统的工程模型;有些工程模型就是所谓“机器视觉”装置,它可以作为“机械人”的眼睛,或做其它图象识别工作。

一、工程技术与视觉理论的新进展

物理学新理论与工程技术渗透到视觉系统的研究中,促进了我们对视觉系统的了解。约300年前,牛顿的光学理论为理解视觉的色觉

现象铺平了道路。随后,Thomas Young(1773—1829)和Helmholtz提出了色觉的三色素理论,认为在人的视觉系统中存在着三种视锥细胞,分别对红、绿、蓝三种颜色敏感,我们看到的各种颜色就是这三种视锥细胞产生的讯号的混合。这个理论至今已被证实。

近年来,信息的傅里叶分解与合成的理论引用到视觉系统中来,引起了视觉研究的革命性变化。

视觉信息是如何在人的视觉系统中进行编码的,或者说,物体的空间特征例如大小和形状,在视觉系统中是如何表达的呢?

一幅照片,可以看作是由不同灰度的象素组成的。将合适灰度的一个个象素,按照一定的次序可组合成任意的景物照片,电视与照片传真技术就是根据这个道理工作的。还有一种完全不同的方式可以同样分解与综合一幅复杂的视觉图象,即认为它是由各种适当的空间频率、取向和振幅的正弦波条纹图重迭而成的。数学上根据傅里叶变换可以证明这一点的。近10余年不少实验说明,视觉系统可能是按后一种方式加工空间信息的。目前比较流行的视觉理论认为,空间信息在人和高等动物的视觉系统中传递和加工时,不同精细度的空间特征(即不同的空间频率)是由分别的机构或多通道进行处理的。这个假设为不少实验所支持,文献[5]中举了三个方面的例证:

1. 关于猫视网膜单个神经节细胞的反应

电生理实验结果表明,有的神经节细胞对0.8

周/度的空间频率条纹最敏感,或者说它们对这个频率调谐。而另一些神经节细胞则对其它的空间频率条纹有最佳反应。可以设想从猫眼输入的景物中包含有各种空间频率条纹,它们使不同的神经节细胞起反应。

2. 人视觉系统的适应效应 让被试者预先注视 3.3 周/度,即比较高的空间频率适应条纹一定时间,然后再让他看类似的高空间频率的试验条纹,发现需要较高的反差才能看见,即对高频率的试验条纹适应效应大。而看低空间频率的试验条纹则影响不大。反之让被试者预先注视 0.7 周/度的空间频率适应条纹一定时间,然后让他再看类似的低空间频率的试验条纹,适应效应也大,而对高空间频率的试验条纹则影响不大。由此可以推测,视觉系统中处理这两种空间频率的机构不可能是相同的,它们适应不同空间频率后的反应说明,它们所调谐的空间频率是不同的。

3. Blakmore 等设计了一个简单易行的实验来支持上述理论。图 1 的左右两边具有相同的

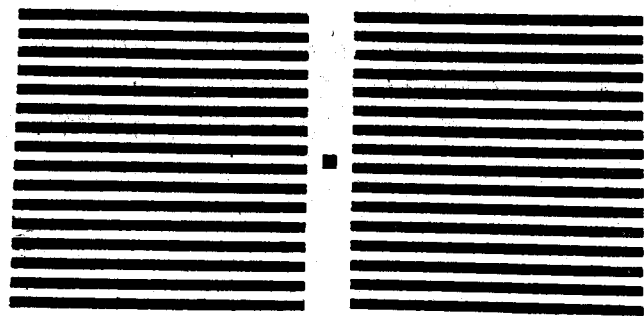


图 1 试验条纹图对
(引自 Blakmore 等)

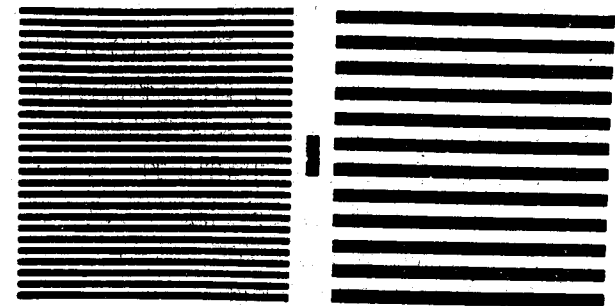


图 2 适应条纹图对
(引自 Blakmore 等)

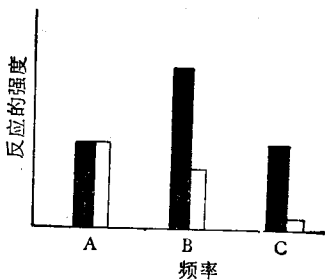


图 3 假设的分析器
(引自文献[1])

■ 适应前 □ 适应后

空间频率或者精细度,你可以注视条纹之间的小方块来确信是这样的情况。然后看图 2 60 秒钟,凝视点在中间的小长方形块上下移动,然后再回到图 1 注视中间的小方块,你会发现左右两边的条纹不再相同了,左边的条纹显得比右边的稀疏。这个实验可以用图 3 来解释。

假设在视觉系统中存在着空间频率分析器,分别由 A、B、C 表示,设 B 对某一频率 f 调谐, A 对低于 f 的某一频率调谐, C 对高于 f 的某一频率调谐。假如我们看一个空间频率为 f 的条纹,那末 B 分析器的反应就最强,而 A 和 C 的反应就较弱,这些分析器反应性的分布与感觉到的试验图形的精细度相对应。如果事先某些分析器(如 B 和 C)由于系统适应了稍大于 f 的某一空间频率条纹因而灵敏度降低了,再看 f 频率的试验条纹时,分析器反应强弱的分布发生了移动, A 分析器的反应变成了最大,所以看起来就觉得条纹 f 比实际上应该有的感觉更稀了。反之亦然。

L. Maffei 等的试验更进了一步^[2],他们报道说,人的视觉系统能够将一个图象分解成的傅里叶成分再合成,以构成原视觉图象。试验时将方波条纹的第一、二次谐波分别显示在两个相同的示波器屏上,受试者的两只眼分别看这两个图形。图形的空间频率、反差和相位可以由电位器调节,当调得合适时,感察者感知到一个非常接近于方波的条纹图。

工程技术促进视觉研究的另一个例

子是电子计算机发生的随机点的图(所谓 Julesz 图),它为研究双眼立体视觉以及研究视觉对图案质地的辨认提供了新的方法和途径。

二、由生物视觉到“机器视觉”

生物的启示常是工程技术发明创造的先导。让机器也长上“眼睛”已成为一些工程师的心愿。视觉信息直接输入计算机处理有很大的意义,因为视觉是人和动物接受外界信息最快的一种感觉。据日本京都大学坂井利之教授估计,人与外界环境交换信息的速度为:视觉~ 10^6 byte/秒;听觉~ 10^3 byte/秒;口讲< 24 byte/秒;手打字< 9 byte/秒。我们常用的光电纸带输入机,其输入速度~ 10^3 byte/秒。机器直接输入和处理图象是计算机的重要应用方向。

现在一个新兴的科学技术部门即“图象识别与人工智能”正在世界上蓬勃发展,比如“人工智能”在不长的时间里已开了五次国际会议,而“机器视觉”是“人工智能”中的基本智能之一。在1977年8月召开的第五届国际人工智能会议的论文集中,关于“机器视觉”的文章就有四十一篇。“机器视觉”的研究范围十分广泛,从各个方面模仿视觉系统的功能:从自动识别数字、符号、图片、简单的三维物体,到带有曲面的较复杂的物体、运动的物体,甚至让机器看懂简单的动画片。但这个学科还很很成熟,多数现代的图象识别系统在许多方面(速度可能例外)都不及视觉系统优越。例如阅读手写文字,对于图象识别系统就是一个至今未解决的难题;人眼可以容易地将复杂背景中的感兴趣的目标挑出来,可以在物体的大小、位置、取向发生某些变化后仍旧识别它们,而机器就难于做到这一点;视觉系统可以适应多种景物与图形,而一般的图象识别装置适应性窄得多,往往对某种图形的处理与识别是成功的,而换了另一种图形就不行了。这就说明为什么生物视觉对工程技术有那样大的吸引力,从图象信息处理技术发展的历史也说明了这一点。

美国 MIT 的 Louis., L. Sutro 报道^[3],1964 年美国制订了探索火星的计划,准备发送一个

自动着陆装置登上火星,它将拍摄火星照片并发回地球,于是提出了如何压缩每幅图片的信息量的任务,以便使该装置在火星表面恶劣的条件下可能正常工作的有限时间内发回更多的信息。当时,控制论的奠基人之一,Warren McCulloch 建议先从模拟青蛙眼开始,因为青蛙的视网膜和视觉系统的高级部分就进行着这种信息简化工作,蛙眼能够检测出昆虫,并把经过简化的主要信息报告给蛙脑。他们在前人生物学研究的基础上,模拟了蛙眼视网膜最复杂的一种神经节细胞的信息简化功能,建立了“昆虫检测器”模型。后来他们又提出了脊椎动物中枢神经系统的模型,并进行了工程模拟,为设计装置打下了基础。

意大利那不勒斯控制论研究室报道了崭新的图象平行加工技术的发展过程。我们一般所说的电子计算机及图象识别系统,对信息都是顺序进行加工的。人与动物的视觉系统与之不

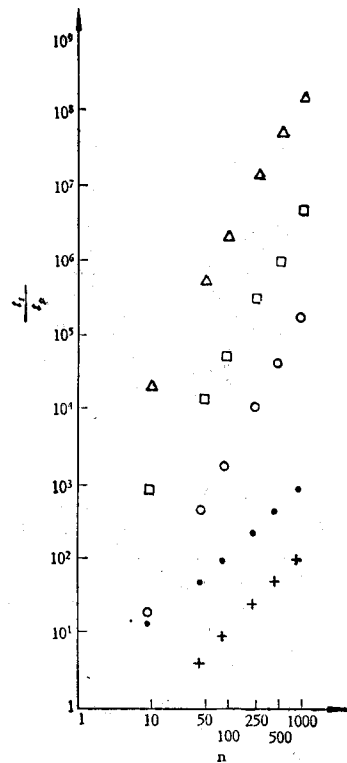


图 4

(引自[4]) t_s : 顺序加工时间; t_p : 平行加工时间; n : 像素数目; Δ : 细胞; $\bullet$: 确定周长; $\circ$: 平滑化; $\square$: 抽提轮廓; $+$: 取阈值

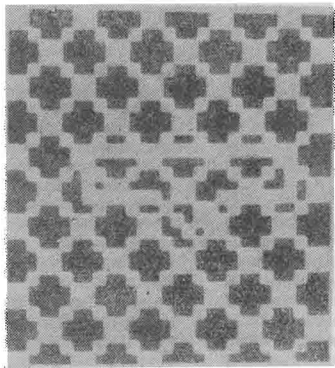
同,例如人眼的视网膜就是信息平行加工装置,视网膜有约一亿二千五百万个光感受器,却只有约一百万根视神经纤维传向大脑,其中约十万根与眼睛黄斑区的光感受器一一对应,所以其余九十万根对应于约一亿二千五百万个光感受器,也就是说视网膜对周边感光信息平行地进行了简化,简化比约 140:1。视觉图象是由视神经纤维并行传递的,这与电视摄影机不同。该研究室以及英国伦敦大学学院等已经提出了平行计算机逻辑元件设计以及图象平行加工算法。这种算法可以大大提高图象处理速度(图 4),它作为一种新技术编入了北大西洋公约组织先进科研项目丛书^[4]。这种平行加工技术有助于使像癌细胞识别一类的复杂的图象识别技术达到实用的速度,这种快速信息处理技术在军事上亦有重要意义。

三、视觉系统模型及其意义

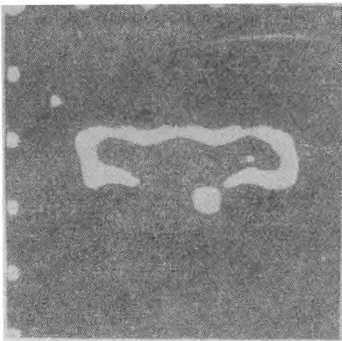
有关视觉系统的模型及模拟工作报道很

多,如日本东京广播科学研究实验室的福岛邦彦长期从事视觉系统的模拟工作,先后构成了模拟哺乳动物视觉系统的“曲线图形特征抽提器”、“自组织多层神经网络”。他还对双眼立体视觉、联想记忆等进行了模拟。这些工作对视觉研究与图象识别技术都有意义。他最近简要地介绍了视觉系统模拟工作的最新进展^[5]。

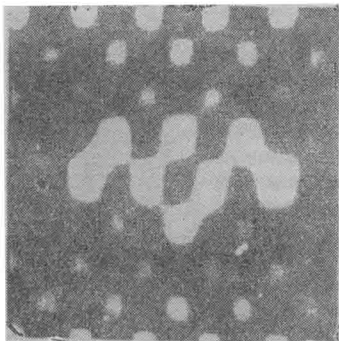
美国俄亥俄州空军研究所的 M. Kabrisky 从六十年代中期起就研究人视觉系统的工程模拟^[6],他总结有关的视觉系统资料,提出三点作为模型的依据:(1) 人大脑皮层的活动是以一个个很小的圆柱状体为单位进行的,这些小圆柱状体他称之为大脑皮层的基本计算单元;(2) 人的所有感觉资料(嗅觉除外)都在脑的中心即丘脑处中继一下,但是丘脑不大可能对经过它的资料进行处理,故只能认为皮层才是对视觉资料进行计算的地方;(3) 已经知道视皮层的 17 区与包围着它的垫圈形的 18 区之间有着麦穗状的联系。由此他提出视知觉过程主要就是



(a) 原图



(b) 经 MTF 空间滤波后



(c) 经低通空间滤波后(直流到 7 周/度)



(d) 经低通空间滤波后(直流到 3 周/度)

图 5 MTF 空间滤波器与普通空间滤波器效果比较

皮层中贮存资料与输入的资料进行二维互相关的运算过程。由此建立的视觉模型在 IBM 1620 计算机上进行了模拟,计算了一些简单几何图形在位置、大小和取向变化时与原图形的相关系数,模拟了皮层中建立短时记忆与长期记忆的过程,并对视觉不变性提出了可能的一种解释。

随后 Tallman^[7] 以及 M. Kabrisky 及其小组根据视觉研究的进展又提出了人视觉系统的模型。他们认为单眼、静态的黄斑视觉可能是皮层 17 区的基本计算单元阵列与 18 区的基本计算单元阵列进行离散的傅里叶变换与空间滤波运算的结果。人的视觉系统可能是低通滤波器,它的二维调制传递函数已经测出^[8],这个模型的模拟结果与生物学实验基本相符。

以上视觉模型促进了美国空间图象处理技术的发展,例如:

1. 模拟人眼二维调制传递函数 (MTF) 的滤波器对消除图象的背景干扰比常用的低通空间滤波器的性能要好 (图 5)。消除图象的背景干扰是视觉图象信息预加工的重要一环。

2. 解决了实时高分辨率雷达目标的鉴别任务。在七十年代初,美国大部分高分辨率图象雷达已达到了足够的分辨率来鉴别中等大小的目标,如飞机或建筑物。对于这些目标的自动识别,困难在于如何尽可能简单地解决识别目标的旋转、大小和位移的视觉不变性。由人视觉系统傅里叶变换的模型,得到了对于简单图形的 $\pm 30^\circ$ 的旋转不变性,而变换的幅度值具有位移不变性,由此提出了解决这个难题的一个办法^[8]。

3. 模拟人视觉系统空间滤波器的原理,极大地改进了从航测照片中检测与鉴定人造物体

的效果。应用视觉系统的傅里叶模型有助于从杂乱的背景中分离出感兴趣的目标^[9]。

四、结 语

视觉的生物学研究与工程技术是互相渗透、互相促进的,而视觉的工程模拟就是这二者之间的桥梁。现在视觉研究取得了不少进展,但是对视觉的脑机制了解得还很少,比较新的总结见文献[10]。我们正处于揭示视觉的奥秘的前夕。学习与记忆是怎么回事? 模棱视觉,视错觉、视觉的后效应、立体视觉是如何产生的? 大脑中究竟对视觉信息进行了什么运算? 为什么人脑进行算术运算与逻辑运算的速度和精度都远远比不上电子计算机,但视觉系统识别图象却这样敏捷、准确、灵活? 能否使未来的“机器视觉”也具有生物视觉系统的优良性能? 这些是生物学工作者与工程技术人员都关心的问题,而这些问题的解决有赖于生物学工作者、工程技术人员与数理工作者的共同努力。

参 考 文 献

- [1] Allan, Pantle: *IEEE NAECON*, 73, Record, 304—308, 1973.
- [2] Maffei, L. et al.: *Nature*, 240, 479, 1972.
- [3] Herbert, W. Robinson et al.: *Cybernetics, Artificial Intelligence and Ecology*, p. 225, 1972.
- [4] Caianiello, E. R.: *New concepts and technologies in Parallel Information Processing* 1975.
- [5] 福岛邦彦:《视觉生理与仿生学》,1980,(中译本)。
- [6] Kabrisky, M. A.: *Proposed model for Visual information Processing in the human brain*, 1968.
- [7] Tallman, O.: *IEEE NAECON*, 69, Record 261, 1969.
- [8] Arthur, P. Ginsburg: *IEEE NAECON* 73, Record, 309, 1973.
- [9] Roger A. Gagnon: *IEEE NAECON*, 73, Record, 317, 1973.
- [10] Hubel, D. H. 和 Wiesel, T. N.: 视觉的脑机制,《科学》,1980,1。

[本文于 1980 年 7 月 10 日收到]

科技消息

人线粒体基因组的结构

人线粒体基因组 DNA 是一个自主复制的独立单位,其全序列已测出,全长 16,569 个碱基对,并已突出其中为 12SrRNA、16SrRNA、22 种 tRNA、细胞色素 C 氧化酶的亚基 I、II、III; ATPase 亚基 6、细胞色素 b、以及其他 8 种蛋白编码的基因的位置。这些基因在

DNA 分子上排列极为紧凑,相磷基因之间根本没有,或者只有很少几个无编码的碱基隔开。许多场合,终链密码子并不由 DNA 序列编码,而是转录后加工过程中,在 mRNA 上接聚 A 时形成。

(刘蓉摘自“*Nature*”, 290, 457, 1981)