



中国科学院
CHINESE ACADEMY OF SCIENCES

认知学习模型的最新研究



中国科学院计算技术研究所

Institute of Computing Technology, Chinese Academy of Sciences



中国科学院
CHINESE ACADEMY OF SCIENCES

没有认知科学建立的"计算"、
"表达"等概念就不可能对认知和智
力开展真正科学的研究

——郑南宁



中国科学院计算技术研究所
Institute of Computing Technology, Chinese Academy of Sciences



有几只老虎?

中国科学院

Chinese Academy of Sciences

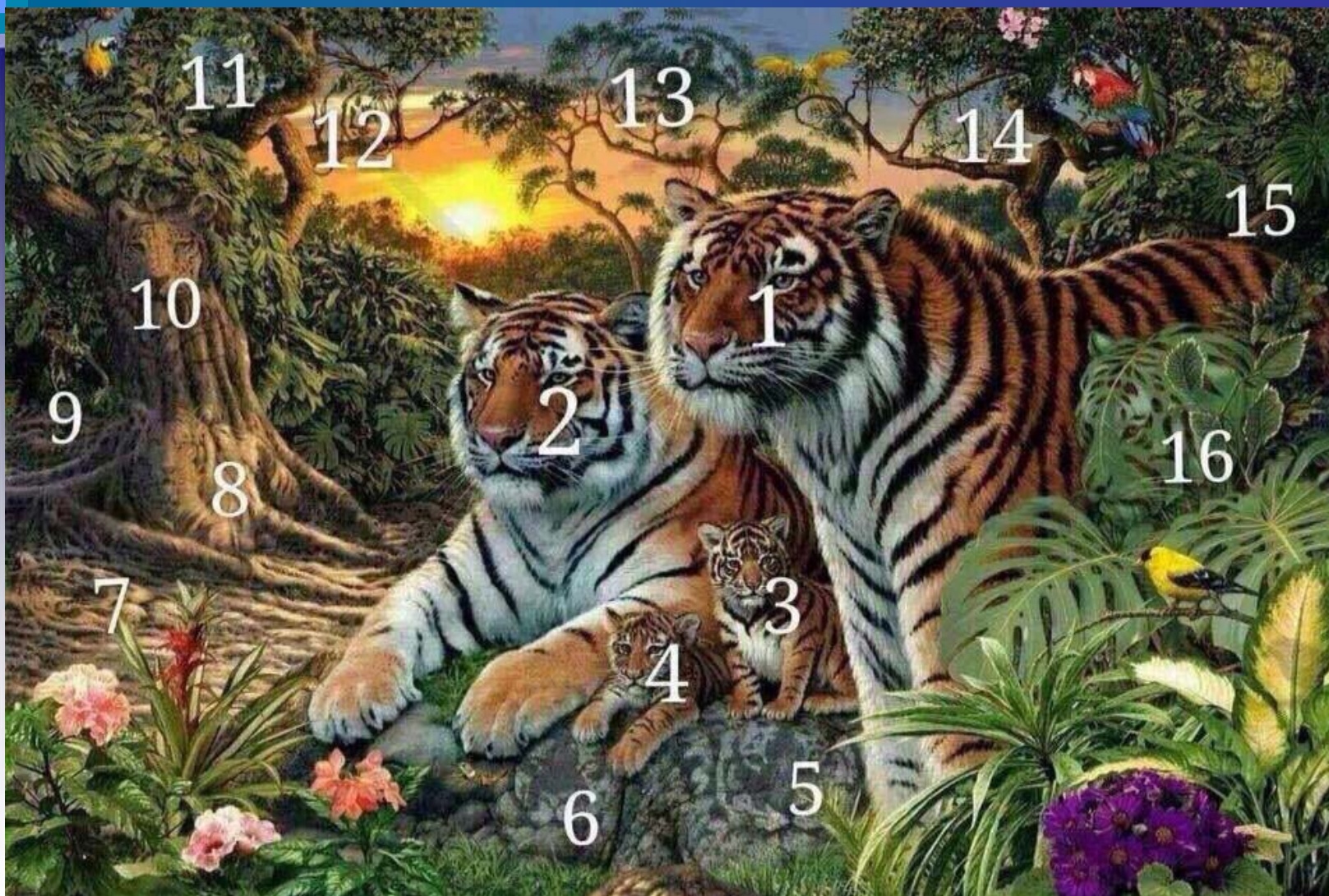




您看出来几只？还有吗？

中国科学院

Chinese Academy of Sciences





中国科学院
CHINESE ACADEMY OF SCIENCES

您能描述您的认知过程吗？



中国科学院计算技术研究所

Institute of Computing Technology, Chinese Academy of Sciences



认知科学

中国科学院

Chinese Academy of Sciences

- 探索物质的本质、宇宙的起源、生命的本质和智力的产生是人类科学事业面临的四大挑战
- 认知科学就是探索人类的智力如何由物质产生和人脑信息处理的过程
- 认知科学研究的范围包括知觉、注意、记忆、动作、语言、推理、思考、意识乃至情感在内的各个层面的认知活动





一.分类认知学习方法

- 构造性学习算法(覆盖算法)
- 基于规划的算法(SVM)
- 基于搜索机制的算法(BP)



覆盖型分类算法

中国科学院

Chinese Academy of Sciences

- 规则覆盖样本
 - 决策树算法
 - AQ系列算法
- 空间划分覆盖样本
 - 基于球邻域的空间划分法
(张铃教授、张钹院士)
 - 仿生模式识别BPR (王守觉院士)
 - 视觉分类方法VCA (徐宗本院士)
 - 分类超曲面算法HSC (何清)





规则覆盖算法—决策树算法

中国科学院

Chinese Academy of Sciences

- 以决策树的形式表示分类规则
- 采用分治法的思想，自顶向下归纳决策树
- 算法执行的每一阶段都寻找一个能将训练样例集按类别划分的最佳属性
- 然后对所得的样例子集进行递归处理
- 直至得到的样例子集只属于一个类别





规则覆盖算法—AQ系列算法

- 首先由Michalski提出，是一种基于规则的方法，把对样本集的学习问题转化为逻辑演算
- AQ 算法是搜索规则空间的算法，利用覆盖所有正例、排斥所有反例的思想来寻找规则
- 它相当于反复应用消除候选元素算法
- 它把学习鉴别规则问题转化为一系列学习单个概念问题
- 奠定了“覆盖”思想的基础



空间覆盖算法—基于球邻域的空间划分

中国科学院

Chinese Academy of Sciences

- 1999年，张铃与张钹教授提出二次规划优化函数的几何方法
- 非线性映射为球函数
- 在球面上，用球邻域概括相同类别的样本
- 计算分类超平面的问题转换为计算样本点两两之间距离所构成的距离空间上的覆盖问题





空间覆盖算法—基于球邻域的空间划分

n 维训练样本(归一化)



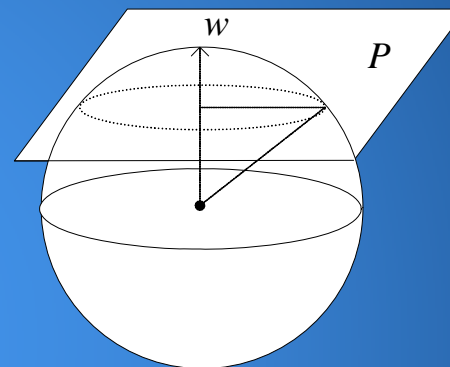
投影函数

$n+1$ 维单位超半球面

分类超平面
的计算



样本点两两之间距离所构成的
距离空间上的覆盖的计算





空间覆盖算法—基于球邻域的空间划分

■ 寻找覆盖的过程

在球面上，考虑一个被映射到球面上的样本，并建立以其坐标为中心的邻域，使得在这个邻域最大限度地包括相同类别的样本。然后，将这些样本删除，重复上述过程，直到覆盖所有的样本。

■ 评价

- 与Vapnik的思想类似，向高维空间映射样本
- 与SVM计算方法不同，引入邻域大大节省了计算资源



空间覆盖算法—仿生模式识别

中国科学院

Chinese Academy of Sciences

- Bionic (Topological) Pattern Recognition (BPR)
- 仿生模式识别是王守觉院士提出的一种新型模式识别理论，基于“认识”事物而不是“区分”事物
- 更接近于人类“认识”事物的特性，故称为“仿生模式识别”
- 它的数学方法在于研究特征空间中样本集合的拓扑性质，故亦称作“拓扑模式识别”





空间覆盖算法—仿生模式识别

中国科学院

Chinese Academy of Sciences

- 传统模式识别把不同类样本在特征空间中的最佳划分作为目标
- 仿生模式识别则以一类样本在特征空间的分布的最佳覆盖作为目标
- 对一类事物的“认识”，实质上就是对这类事物的全体在特征空间中形成的无穷点集合的“形状”的分析和“认识”





空间覆盖算法—仿生模式识别

中国科学院

Chinese Academy of Sciences

- 用“仿生模式识别”中的“仿生”二字,并非 Vapnik 所批评的“目标变成了构造利用大脑来推广的模型”,而只是为了从功能上更接近于人类对事物的识别过程,而从数学模型上加入了实际“被识别对象”的某些普遍存在的“特殊性”
- 这个所谓“特殊性”是针对目前传统的模式识别数学描述方法的“一般性与通用性”而言的,它在实际的模式识别问题中却仍然具有一般性与通用性





空间覆盖算法—仿生模式识别

$$S = \{x | x = S_i (i = 1, 2, \dots, \text{采集样本总数})\}$$

$$P_a = \{x | \rho(x, y) \leq k, y \in A, x \in R^n\}$$

$$A = \{x | x = x_i, i = (1, 2, \dots, n), n \in N, \rho(x_m, x_{m+1}) < \varepsilon,$$

$$\rho(x_1, x_n) < \varepsilon, \varepsilon > 0, n - 1 \geq m \geq 1, m \in N\}, A \subset R^n, S \subset A$$

为了实际用神经网络中若干神经元来实现近似覆盖 P_a , 我们用若干空间直线段来近似空间曲线 A , 每个神经元覆盖的是一个直线段与 n 维超球的拓扑乘积。

设采集原始样本集合为 S , 在其中取样本 j 个组成新的样本集 S' , 使

$$S' = \{x | x = S'_i (i = 0, 1, 2, \dots, j), \rho(s'_i, s'_{i+1}) \leq d <$$

$$\rho(s'_{i-1}, s'_{i+1}), d \text{ 为设定常数}, s'_0 = s'_j\}, S' \subset S$$

用 j 个神经元来近似地覆盖 P_a , 第 i 个神经元覆盖范围 P_i 为:

$$P_i = \{x | \rho(x, y) \leq k, y \in B_i, x \in R^n\}$$

$$B_i = \{x | x = \alpha S'_i + (1 - \alpha) S'_{i+1}, \alpha \in [0, 1]\}$$

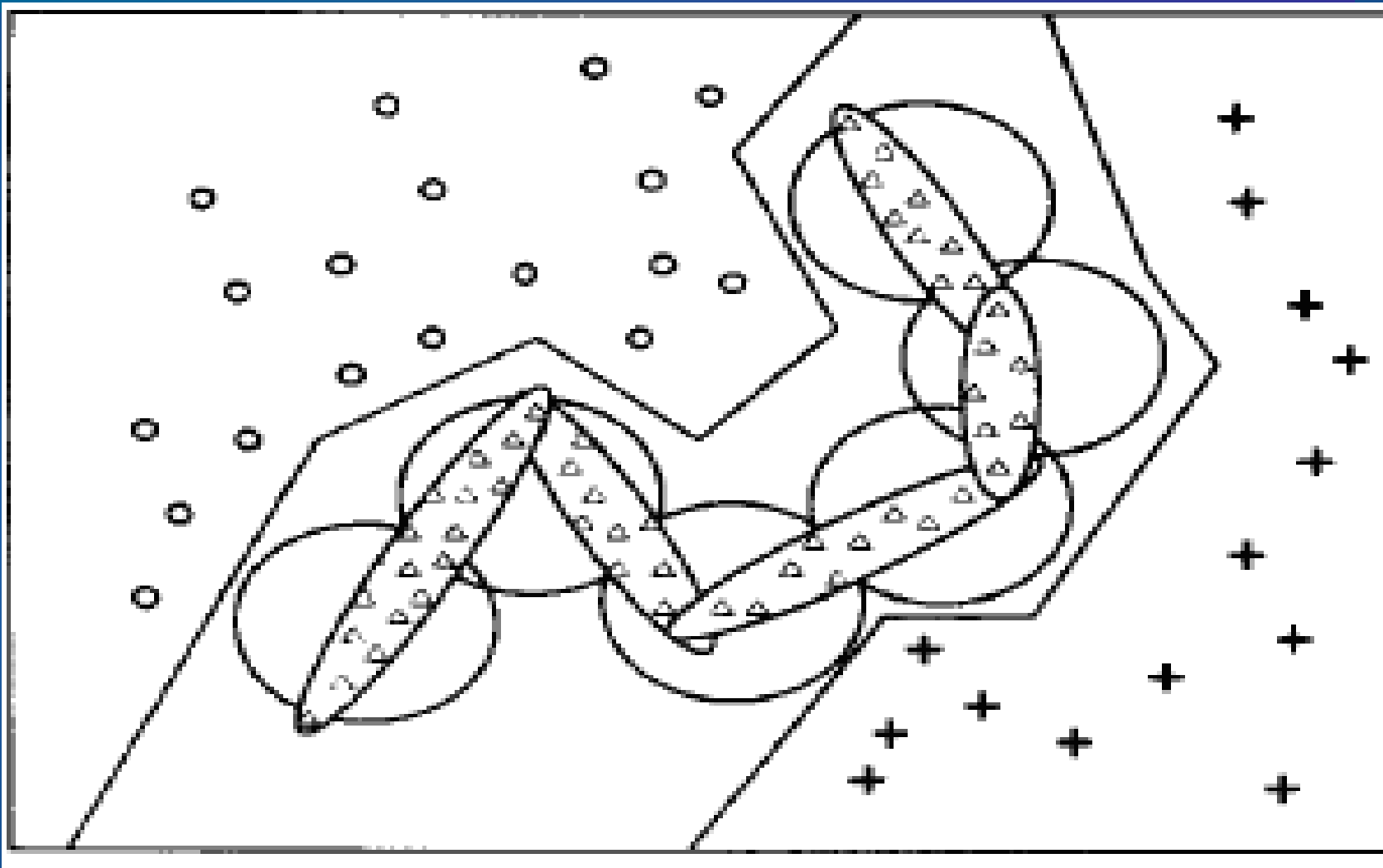
$$\text{全部 } j \text{ 个神经元覆盖范围为: } P'_a = \bigcup_{i=0}^{j-1} P_i$$

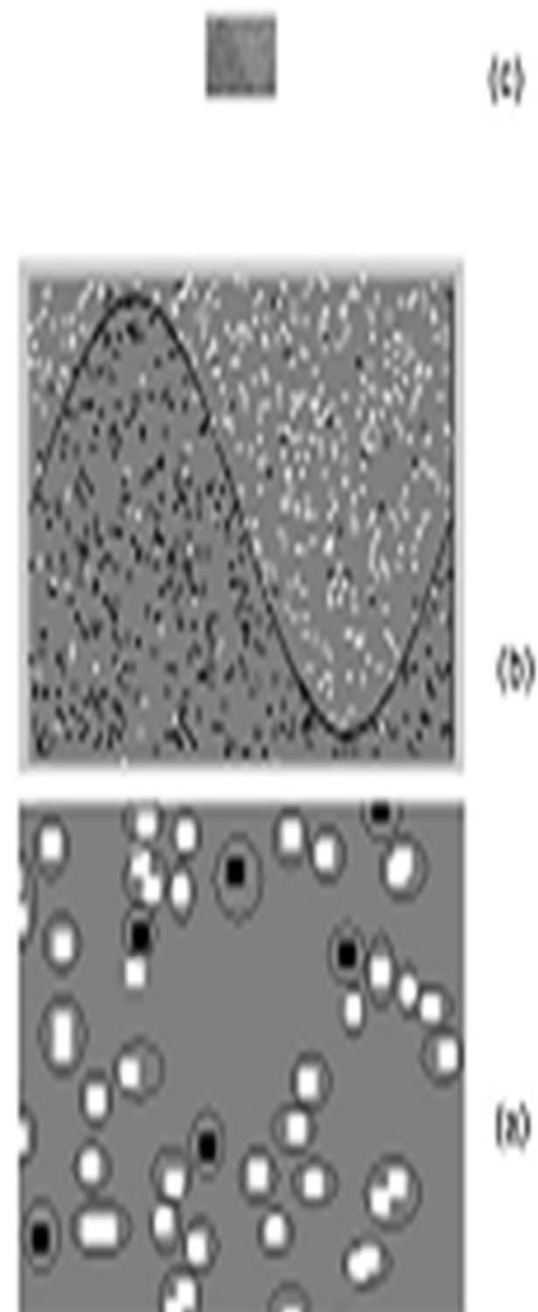
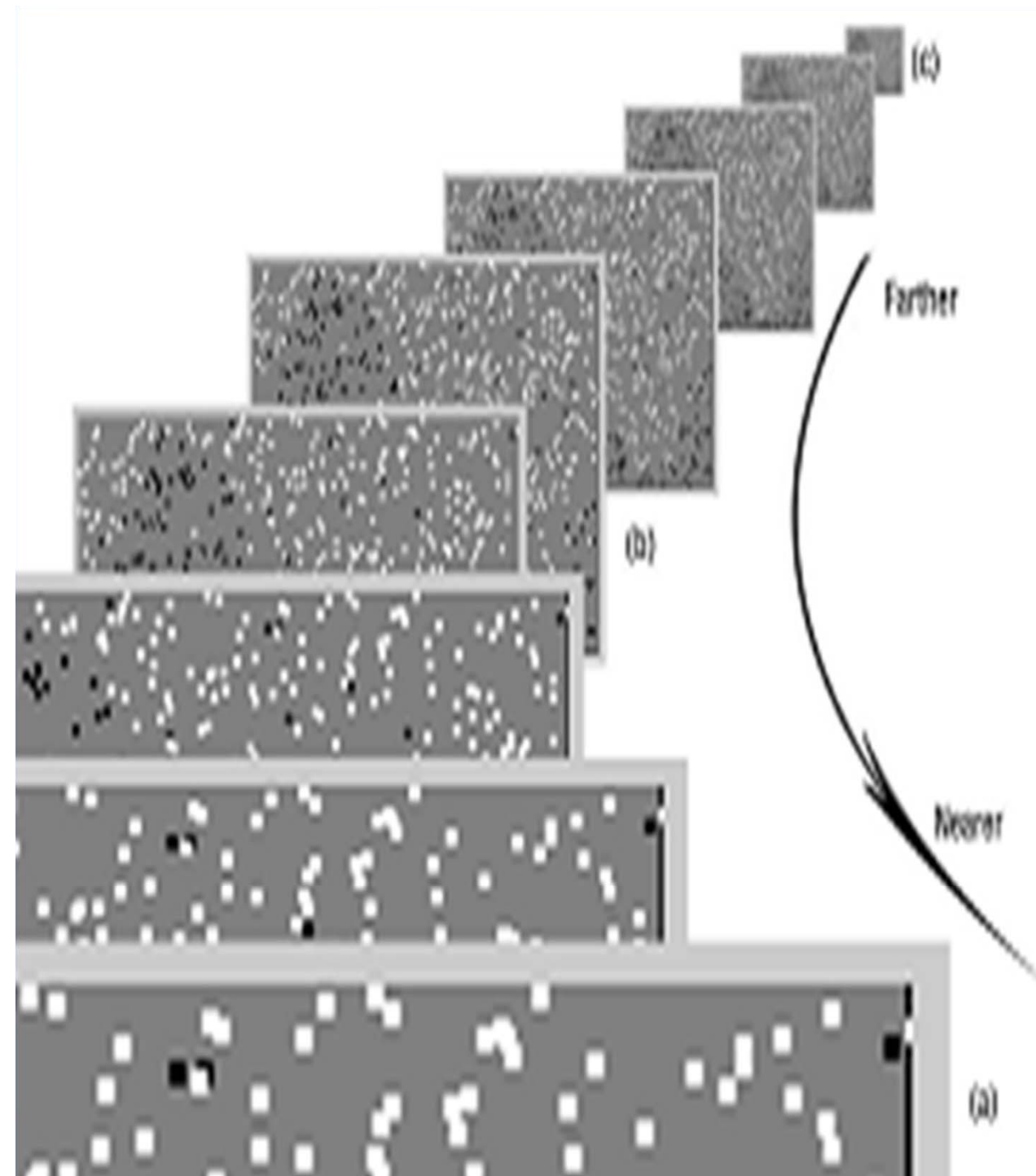


空间覆盖算法—仿生模式识别

中国科学院

Chinese Academy of Sciences







空间覆盖算法—视觉分类方法

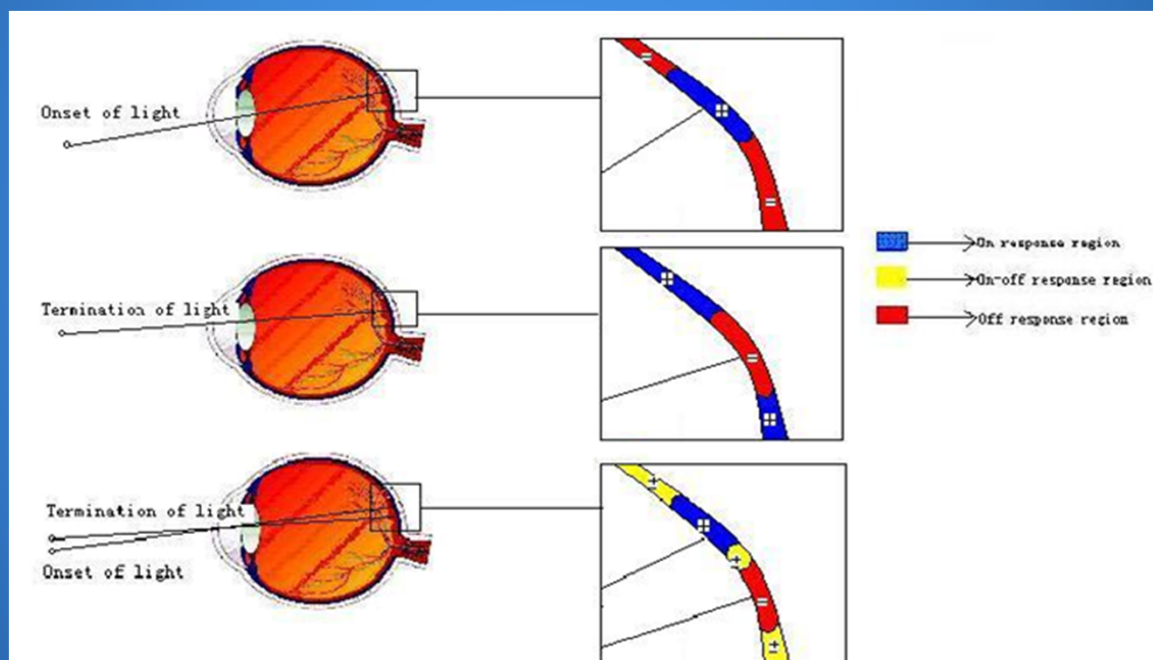
- 视网膜细胞对光线的变化有三种反应状态
 - ON: 光线刺激到达时
 - OFF: 光线刺激消失时
 - ON-OFF: 多束光线刺激
- 光源影响
 - 对单个固定点光源来说，它所引发的处于ON/OFF状态的视网膜细胞呈类高斯分布，其余处于OFF/ON状态的视网膜细胞也呈类高斯分布。
 - 对多个点光源来说，可能同时存在多个不同的ON、OFF、ON-OFF区域。特别地，ON-OFF区域构成ON区域与OFF区域的分界线



空间覆盖算法—视觉分类方法

中国科学院

Chinese Academy of Sciences

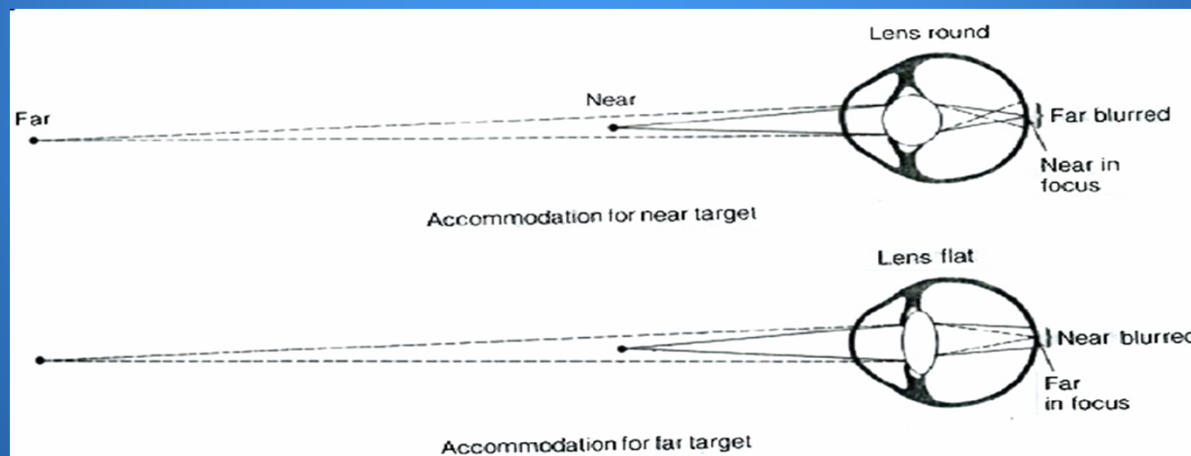




空间覆盖算法—视觉分类方法

■ 水晶体的曲度随着距离的变化自动调节

- 问题：同一个物体，随着距离(水晶体曲度)的变化，其视网膜成像是如何变化的？
- 当看近物时，水晶体曲度变大，比较凸起
- 当看远物时，水晶体曲度变小，比较扁平





覆盖算法—视觉分类方法

中国科学院

Chinese Academy of Sciences

■ 尺度空间理论 (Scale Space Theory)

- Let $p(x) : R^n \rightarrow R$ denote an image whose gray degree at x is $p(x)$. $P(x, \sigma)$ is the image appeared in the retina at the scale σ . The scale space theory then says that $P(x, \sigma)$ obeys to

$$\begin{cases} \frac{\partial P}{\partial \sigma} = \Delta_x P \\ P(x, 0) = p(x) \end{cases}$$

or equivalently,

$$P(x, \sigma) = p(x) * g(x, \sigma) = \int g(x-y) p(y, \sigma) dy,$$

where $*$ denotes the convolution operation and $g(x, \sigma)$ is Gaussian kernel defined by

$$g(x, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|x\|^2}{2\sigma^2}}.$$

■ Visual Classification Algorithm (VCA)

■ 基于视感知学习算法要解决的两大问题

- 寻找最佳尺度
- 构造分类模型





基于规划的算法-SVM

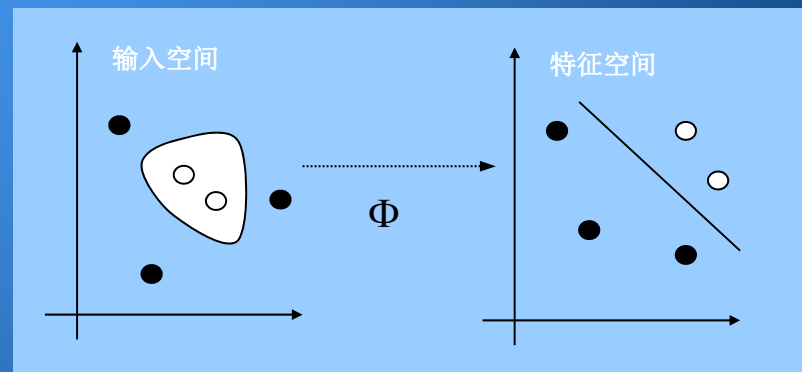
中国科学院

Chinese Academy of Sciences

■ 支持向量机

- 通过内积函数定义的非线性映射将样本空间映射到一个高维线性空间（特征空间），寻求支持向量和最优分类超平面
- 对小样本、非线性和高维数据具有很好分类性能
- 计算开销大

$$Q(\alpha_i) = \sum_{i=1}^n \alpha_i - \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j)$$





基于搜索机制的算法(BP)

- BP神经网络模型拓扑结构包括输入层（input）、隐层(hidden layer)和输出层(output layer)
- BP网络能学习和存贮大量的输入-输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程
- 它的学习规则是使用最速下降法，通过反向传播来不断调整网络的权值和阈值，使网络的误差平方和最小



王守觉院士关于SVM的论述

- Vapnik 在1974 年提出了“最优分类超平面”的概念 并在此基础上后来发展构造成了支持向量机(SVM)
- 数十年来人们在研究模式识别问题中, 考虑的出发点都是在若干类别的最佳分类划分上. 其根本原因或许在于用这样的数学描述与处理方法最具有—般性、通用性
- 但众所周知, 即使基于目前最先进的模式识别理论基础上的识别机, 其实际效果却仍然远不能令人满意



Vapnik 对人工智能学者的评价

- 在80 年代, 人工智能学者在计算学习领域中扮演了主要的角色
- 在这些人工智能研究者中间, 一些较极端的学者有很大的影响(正是他们强调了‘复杂的理论是没有用的, 有用的是简单的算法’)
- 这些人工智能学者对于处理学习问题有丰富的经验, 善于对一些理论上非常复杂的问题构造‘简单的算法’



参考文献

- Zongben Xu, Deyu Meng and Wenfeng Jing A New Approach for Classification: Visual Simulation Point of View in: J. Wang, X. Liao, and Z. Yi (Eds.): ISNN 2005, LNCS 3497, pp. 1–7, 2005.
- Ling Zhang, Bo Zhang, “A Geometrical Representation of McCulloch- Pitts Neural Model and Its Applications”, IEEE Transactions on Neural Networks, Vol.10, No.4, July 1999, pp. 925–929.
- 张铃, 张钹, 殷海风. 多层前向网络的交叉覆盖设计算法. 软件学报1999, 10(7): 737–742.



参考文献

中国科学院

Chinese Academy of Sciences

- 王守觉 仿生模式识别(拓扑模式识别) —— 一种模式识别新模型的理论与应用. 电子学报, 30(10) (2002) 1417-1420.
- 王守觉, 王柏南. 人工神经网络的多维空间几何分析及其理论. 电子学报, 2002, 30(1) : 1 - 4.
- Wang, S. J., Qu Y. F., Li W. J., Qin H.: Face Recognition: Biomimetic Pattern Recognition vs. Traditional Recognition. ACTA ELECTRONICA SINICA 32(7) (2004) 1057-1061

