第二章

知觉与模式识别

肖承丽

1

第一节　知觉

1. 计算式大脑
2. 感觉与知觉
3. 知觉广度与感觉存储

2

计算式大脑

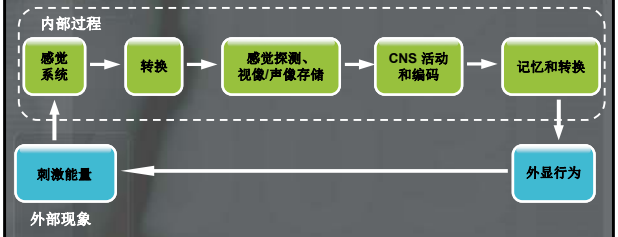
- 早上醒来 ≈ 启动你的认知计算机

- 被闹钟吵醒
- 睁开双眼
- 把水泼到脸上
- 闻到早餐的香味
- 吃一口新鲜的面包

- 感知周围环境的信息
- 注意周围的世界
- 在初始阶段加工信息

3

- 心灵对信息进行加工
- 涉及多个计算阶段



4

第一节　知觉

1. 计算式大脑
2. 感觉与知觉
3. 知觉广度与感觉存储

5

你能看多远?

- 你能看到的最远距离与很多因素相关，如物体大小、亮度等

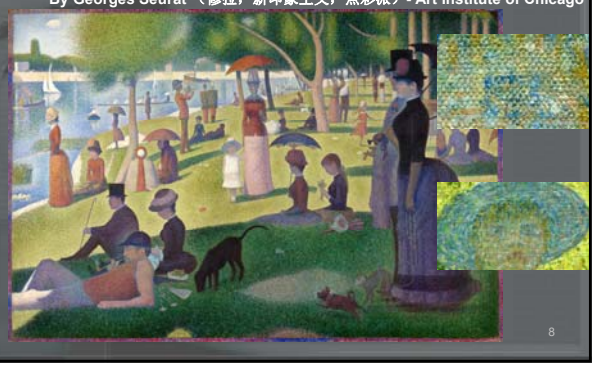
6

7

大碗岛星期天的下午

A Sunday Afternoon on the Island of La Grande Jatte

By Georges Seurat （修拉，新印象主义，点彩派）- Art Institute of Chicago



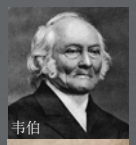
8

- 心理物理学
  - 研究外部世界物理属性的改变与相应的心理体验之间的关系的学科

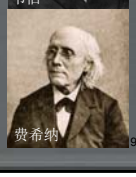
S=KlogR

S—感觉强度；K—常数；R—刺激强度

要使感觉呈算术级增强，那么物理刺激量则必须呈几何级数增加



韦伯

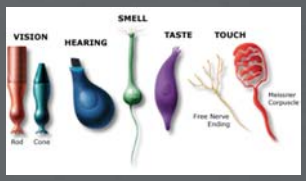


费希纳

9

感觉

- 对物理世界的能量的初次探测

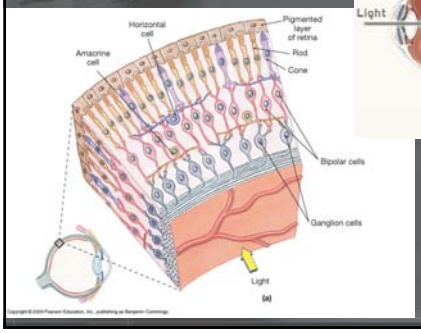


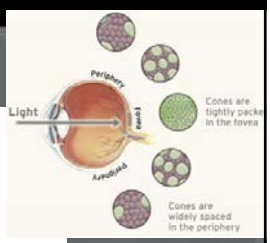
感觉	结构	刺激	感受器
视觉	眼睛	光波	视杆和视锥细胞
听觉	耳朵	声波	毛细胞
味觉	舌头	化学物质	味蕾
嗅觉	鼻子	化学物质	嗅觉受体神经元
触觉	皮肤	压力	神经细胞

10

视觉

- 视杆和视锥细胞

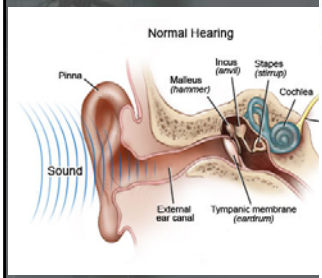


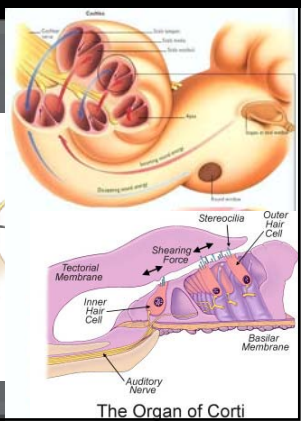


11

听觉

- 毛细胞

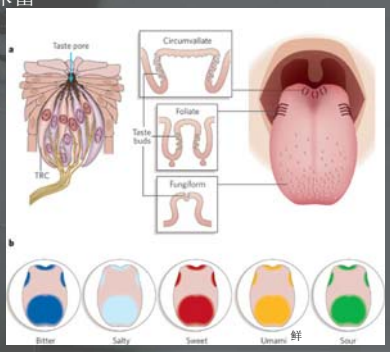




12

味觉

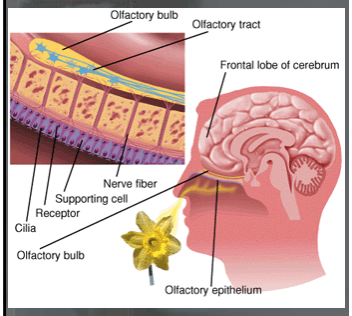
- 味蕾

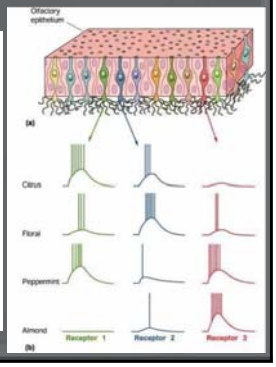


13

嗅觉

- 嗅觉受体神经元

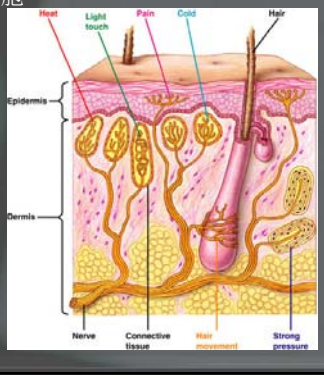




14

触觉

- 神经细胞



15

知觉

- 涉及解释感觉信息的高级认知过程

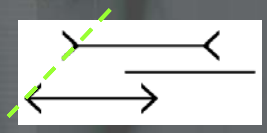
- 读书
- 听音乐会
- 享受按摩
- 闻香水
- 品尝佳肴

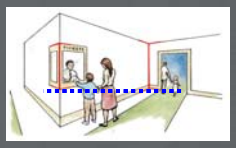
关于世界的知识 / 先前的经验

16

错觉

- 心智对现实的扭曲
  - 帮助我们认知我们的视觉和认知系统
- 缪勒-莱尔错觉（Muller-Lyer Illusion）





- 过去经验：让我们学会预测某些图形距离较远，某些图形距离较近
- 深深植根于脑中的始终如一的构造

17

"Waterfall"
by M. C. Escher

18

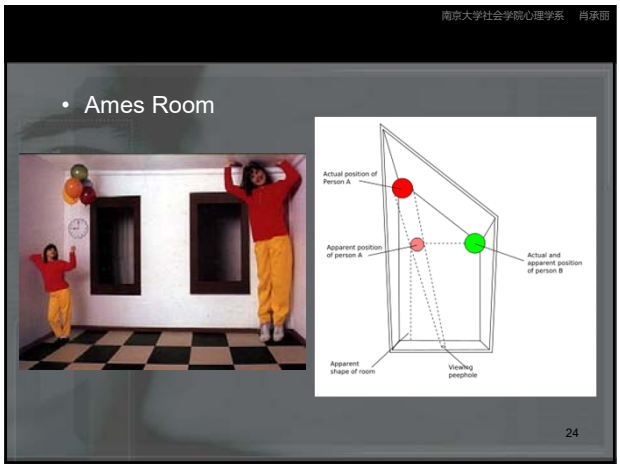
19

20

21

22

23



24

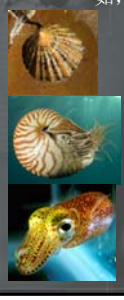
先备知识

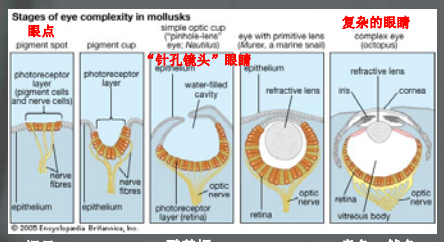
- 赋予刺激的初始感觉体验丰富多彩的意义

25

感觉——脑的先天禀赋

- 认识论
  - 心灵的窗口，即人类的感知系统是随着我们这个星球上的物质变化而出现的
  - 如，眼睛的进化





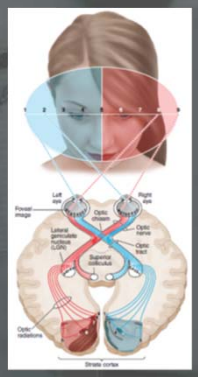
帽贝　　鹦鹉螺　　章鱼，鱿鱼

26

- 对感觉信号的探测和解释取决于：
  - 感觉系统和脑所感知到的刺激能量
  - 先前存储在记忆中的知识

27

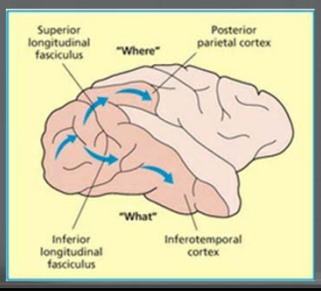
Two Visual Systems: Perception & Action



- 根据外侧膝状体细胞类型的不同，视觉神经通路可以分为：
- P通路（parvocellular pathway，又称为小细胞通路）
  - 主要接收来自视锥细胞的信号（对颜色和细节信息敏感），
- M通路（magnocellular pathway，又称为大细胞通路）。
  - 主要接收来自视杆细胞的信号（对运动信息敏感）

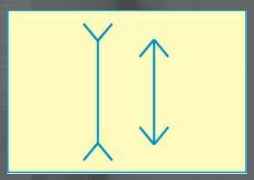
28

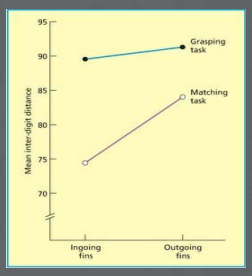
- P通路⇨下颞叶，加工视觉的特征属性信息，又被称为“what”通路
- M通路⇨顶叶，加工视觉的运动信息，又被称为“where”通路



29

- Haart, Carey, and Milne (1999)

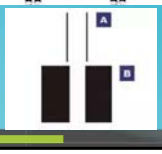


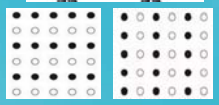


30

怎样穿最显瘦,听听心理学家怎么说

Cognition in Everyday Life





31

第一节 知觉

1. 计算式大脑
2. 感觉与知觉
3. 知觉广度与感觉存储

32

知觉广度

- 在一次短暂呈现中能够感受到的信息量

区分两个假设的结构：
- 知觉的感觉存储
- 短时记忆

33

- 第一项关于知觉广度的实验
  - 阅读并不是平滑地沿着文本线路进行扫描，而是以一种从一个注视点跳跃到另一个注视点的方式进行 (Javal, 1878)
  - 阅读（对文本材料信息的收集）发生在注视期间，而不是在跳跃(saccades眼跳)或扫描期间 (Cattell, 1886a, 1886b; Erdmann & Dodge, 1898)

一次短暂呈现中最多能够摄取4 – 5个无关联字母的信息

34

请尽可能多地记住下面将要呈现的字母：

X	A	Z
N	T	S
Q	R	C

35

一次短暂呈现中最多能够摄取4 – 5个无关联字母的信息

- 这一“铁律”持续了60年之久，直到斯伯林(George Sperling,1960) 证明它是错误的
  - 早期实验：记住多少报告多少 —— 测量被试的记忆
  - 部分报告法

36

请尽可能多地记住下面将要呈现的字母：

X A Z

N T S

Q R C

37

请尽可能多地记住下面将要呈现的字母：

R G C

L X N

S B J

38

请尽可能多地记住下面将要呈现的字母：

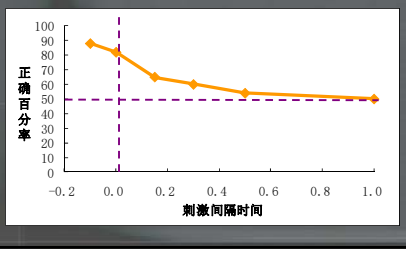
F S T

A K X

U Z W

39

- 改变字母呈现和声音信号呈现之间的时间间隔，从而测量视像存储的时间
  - 时间间隔增加，回忆正确率下降
  - 时间间隔增至1秒，回忆数量约为4.5个，和先前的实验结果类似
  - ⇨ 视像存储的时间约为250毫秒
    - 约等于阅读过程中的注视时间



40

- 奈瑟(Neisser, 1967)
  - 视像记忆 (Iconic memory)
    - 视觉印象的持续和一段时间内可供进一步加工
  - 声像记忆 (Echoic memory)
    - 听觉的感觉记忆

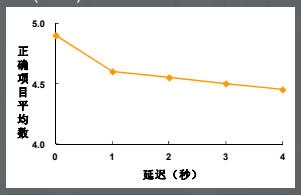
声像存储：
- 250毫秒~4秒
- 原始形式

短时记忆(STM):
- 10~30秒
- 部分受到加工

41

- 对四耳人的实验
  - Moray, Bates, & Barnett (1965)
- 对三耳人的实验
  - Darwin, Turvey, & Crowder (1972)

左耳	双耳	右耳
B	8	F
2	6	R
L	U	10



声像存储可以持续长达4秒，但是在听觉刺激呈现后1秒之内最为清晰

42

感觉存储的功能

- 视像存储、声像存储和其他的感觉信息的存储，使我们只需抽取有待进一步加工的信息
- 对来自外部物理世界的信息提取遵循节省法则

43

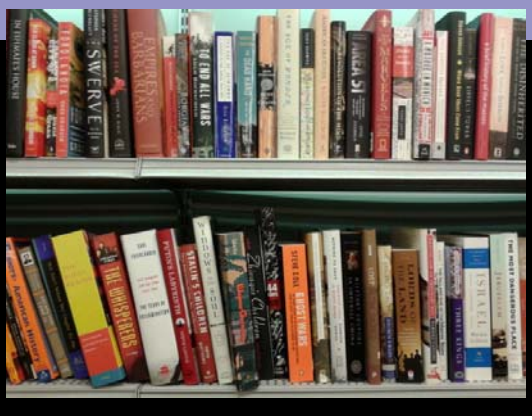
第二节　模式识别

南京大学社会学院心理学系
肖承丽

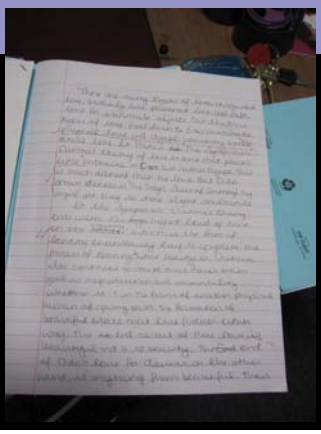
44

45

46

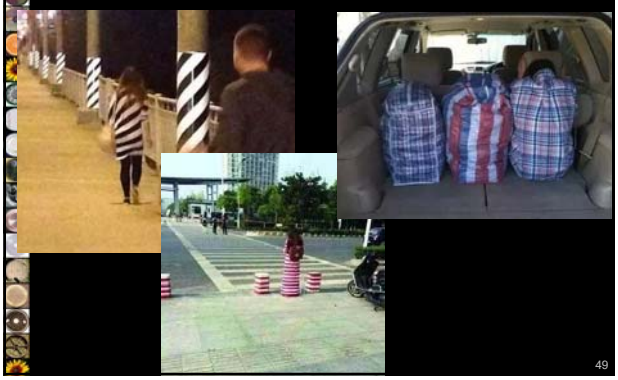


47



48

只是因为在人群中多看了你一眼，再也没能忘掉你容颜…



49

50

荷兰领馆哭了：别砸了，俄罗斯领馆在隔壁……

故事是这样的：

•土耳其击落俄罗斯战机

•俄罗斯人民很生气怒砸土耳其驻俄使馆

•于是乎，积怨已久的土耳其人纷纷涌向了……涌向了……俄罗斯使馆……的……隔壁：荷。。兰。。使。。馆。。。

•冒着“蛋雨”，欲哭无泪的荷兰总领事发了这么条推特：

•“老是不爽的土耳其示威者错把我们当成邻居俄罗斯领事馆了。。正在扔鸡蛋！！！”

51

- 法国使馆对土耳其人民能分清横竖表示了感谢

- 据说这事把卢森堡使馆惊出了一身冷汗

- 德国使馆则认为：此次乌龙虽然有损于土耳其人民的智商但至少能证明一点：土耳其人不是色盲！！！

52

模式识别

- 包含了感觉、知觉、短时记忆、长时记忆与旨在辨别刺激的认知搜索之间复杂的相互作用
- 特点：
  - 以较高的准确性迅速识别出熟悉的模式
  - 评价和理解不熟悉的模式
  - 准确知觉到以不同角度放置或旋转了的物体
  - 辨别部分不可见，被遮蔽，或处于某种形式的“噪音”掩蔽中的物体
  - 快速、自动、主观感觉轻松地执行模式识别

53

- 以较高的准确性迅速识别出熟悉的模式

54

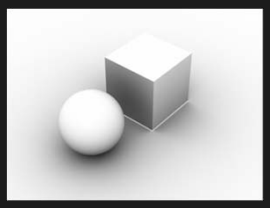
- 评价和理解不熟悉的模式

55

- 准确知觉到以不同角度放置或旋转了的物体

56

- 辨别部分不可见，被遮蔽，或处于某种形式的“噪音”掩蔽中的物体

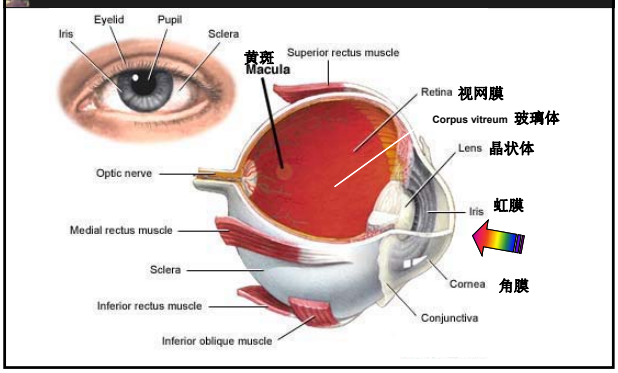


57

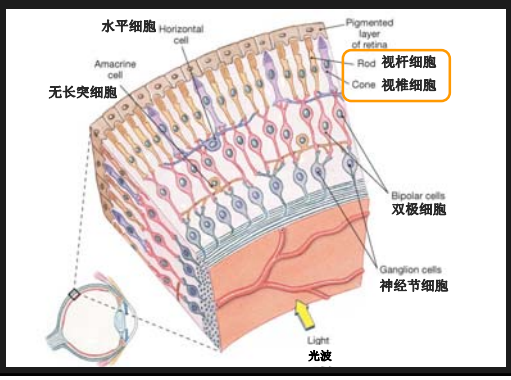
- 快速、自动、主观感觉轻松地执行模式识别

58

视觉

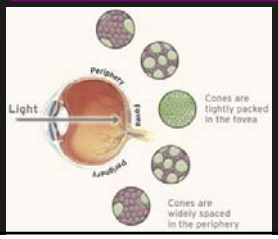


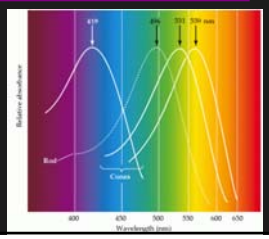
59



60

	视锥细胞	视杆细胞
数量	700万个	1.25亿个
对…敏感	光亮度好的刺激 颜色	光亮度差的刺激
分布	集中在中央凹	中央凹以外区域

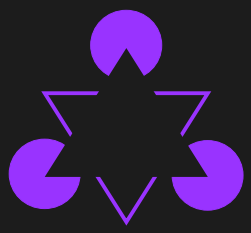




61

主观组织

- 错觉
  - 错觉轮廓
    - 人们看到物理上并不存在的轮廓
  - 侧抑制
    - 视网膜中相邻神经元倾向于抑制周围细胞，从而加强了轮廓



62

- 边缘的知觉清晰度与图形的亮度是诱导特征密度函数

63

- 解释
  - 进化观的解释
    - 看到形状、边缘和运动（有人或许会加上人脸）的需要是生存所必需的
    - 消除伪装作用 (Ramachandran, 1987)
  - 神经证据
    - 错觉轮廓激活了视皮层（V1，V2）
  - 格式塔心理学
    - 我们倾向于从所在环境中看到简单的、熟悉的或“良好的”图形

64

视觉模式识别

1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

65

格式塔理论

- 模式组织
  - 所有刺激共同作用而形成一个印象，它超越了所有感觉的总和

66

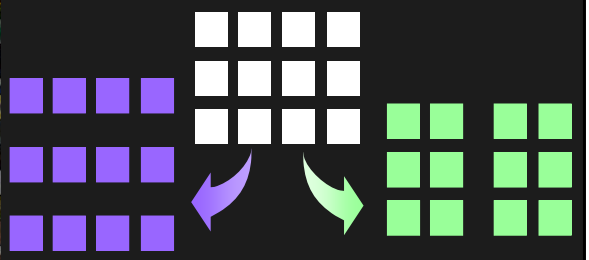
封闭律

- 彼此相属、构成封闭实体的各部分趋于组成整体

67

接近律

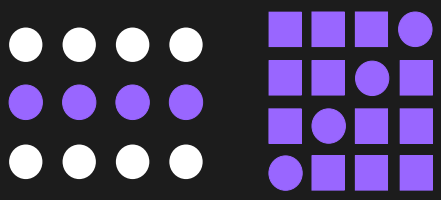
- 距离相近的各部分趋于组成整体



68

相似律

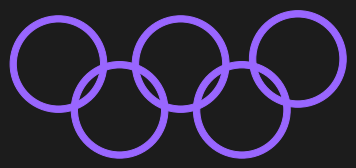
- 相似的各部分趋于组成整体，包括形状、颜色、大小、亮度等



69

Laws of simplicity简单律

- We see five circles in the picture, not the more complex array of nine objects.



70

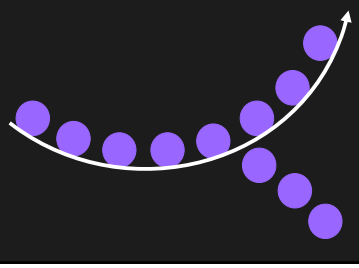
Law of Symmetry对称律

- (Figure ground relationships)— Symmetrical images are perceived collectively, even in spite of distance.

71

Law of Continuity连续律

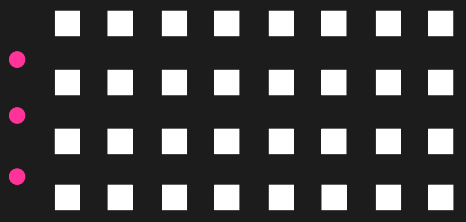
- The mind continues visual, auditory, and kinetic patterns.
  - Good continuation. The top branch is seen as continuing the first segment of the line. This allow us to see things as flowing smoothly without breaking lines up into multiple parts.



72

Law of Common Fate共同命运律

- Elements with the same moving direction are perceived as a collective or unit.



73

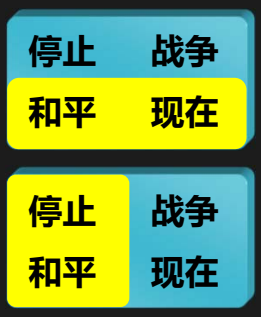
- Palmer (1977)

- 请快速判断下列图形是否是上面图形的一部分

快 快 慢 慢

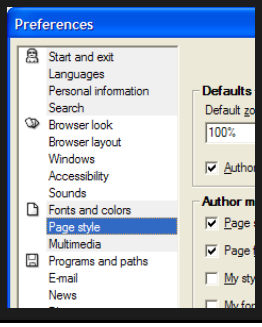
74

日常生活中的认知



75

- 相似律应用
  - 室内设计
  - 界面设计



76

- 接近律应用
  - 海报设计
    - 音乐电视大奖 (MTV Music Awards 2002)
    - 徽标（Logo）和赞助商分别聚集在两个地方

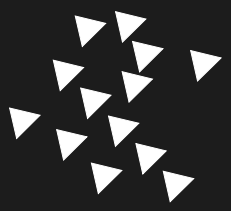
77

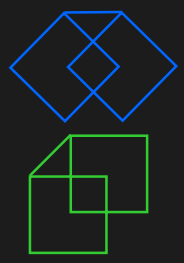
分组讨论：日常生活中应用格式塔原理的例子

78

高级心理过程

- 自主控制
- 过去经验





79

视觉模式识别

1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

80

典型表象

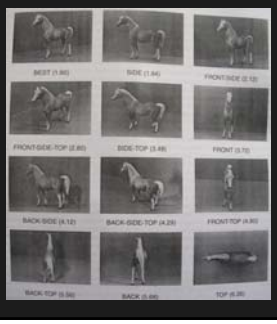
- 能对物体做出最佳表征的形象，或回忆一种形状时最先想起的表象
  - 请画一只咖啡杯和杯托

81

- 我们对于物体的日常经验可以发展出对该物体最具代表性的形象的永久记忆，以及涵盖其最多信息的形象的永久记忆
- 典型表象的研究，有助于我们了解
  - 原型形成
  - 思维经济性
  - 交流上的效率

82

- Palmer, Rosch, & Chase (1981)
  1. 被试评价拍摄角度的典型度和熟悉度
  2. 尽可能快地命名物体



83

- 越不典型的图片需要的反应时间越长
  - 该物体能够被识别的部分较少
  - 最好的（典型的）形象是最经常经验到的形象
  - 典型表象是对物体理想化的，或最好的形象

84

赫亚尔肖像
采自赫亚尔墓室的一扇木门
约前2778-前2723年

埃及人作画的方法始终如一：无论哪一个事物，他们都得从它最具特性的角度去表现：

•头部在侧面图中最容易看清楚，他们就从侧面画。但是，如果我们想起人的眼睛来，就想到它从正面看见的样子。因此，一只正面的眼睛就被放到侧面的脸上。

•躯体的上半部是肩膀和胸膛，从前面看最好看，那样我们就能看见胳膊是怎样和躯体结合。而且一旦活动起来，胳膊和腿从侧面看要清楚得多。

•他们喜欢从大脚趾向上翘的清晰的轮廓线。这样，两只脚都是从内踝那一面看，浮雕上的人看起来就仿佛有两只左脚

——【英】贡布里希《艺术的故事》

85

内巴蒙花园
陵墓壁画
埃及底比斯内巴蒙墓出土
约前1350年
64×74.2cm
英国，伦敦，大英博物馆

86

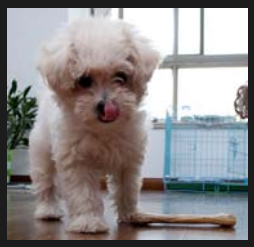
视觉模式识别

1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

87

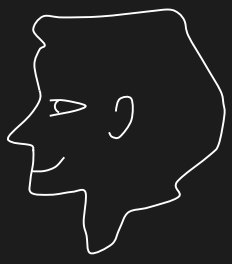
自下而上加工与自上而下加工

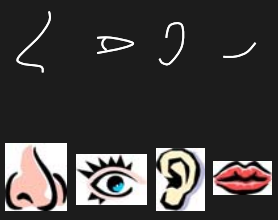
- 自下而上
  - 识别模式的各个部分
    ⇨ 识别整个模式
- 自上而下
  - 关于事物整体的假设：导致整体识别与之后的部分识别



88

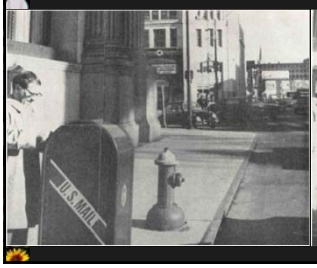
- 在多数情况下，对部分和整体的解释在自上而下与自下而上两个方向上同步发生 (Palmer, 1975)

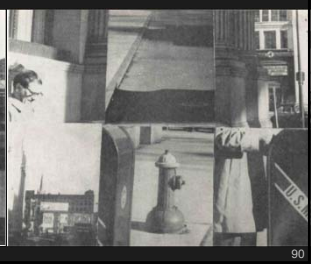




89

- 世俗知识能帮助我们在熟悉的情境中辨别物体，在不合适的情境中阻碍我们的辨认
  - 情境效应: 一个人在现实情境中寻找物体时，其识别准确性和辨别所需时间与该物体在位于该场景的适宜性有关
  - Biederman et al. (1973)
    - 寻找消防栓





90

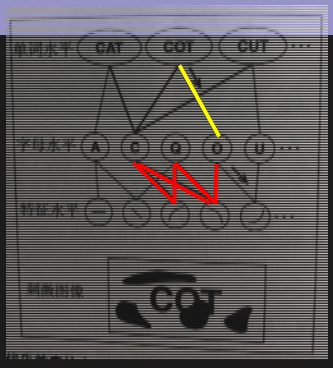
自上而下的加工模型

- 网络反馈模型
- 贝叶斯方法

91

- 网络反馈模型

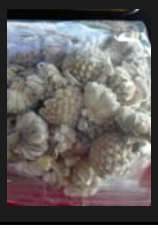
当特征信息与自上而下的影响相结合，字母O得到最强的激活



Rumelhart & McClelland (1987)

92

- 贝叶斯方法
  - 记忆中存储信息的影响具有概率性
  - E.g.，一个黄色弯曲细长的东西是?
    - 香蕉
    - 气球
    - 黄辣椒
    - ……



93

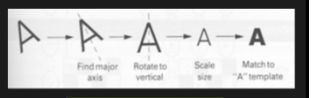
视觉模式识别

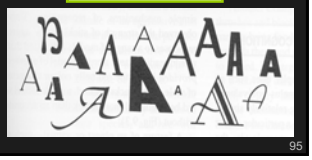
1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

94

模板匹配

- 模板指的是一种内部结构，当感觉刺激与之匹配后，便产生了对物体的识别
- 具有一定的理论和实际意义
- 但是
  - 需要存储亿万个模板
  - 搜索模板费时
  - 无法识别不熟悉的物体





95

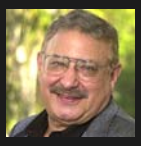
机器vs.人：模式识别之验证码

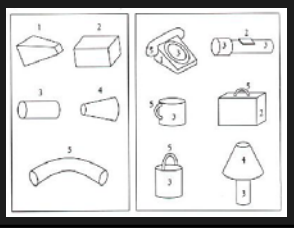
Cognition in Everyday Life

96

几何离子理论

- Irving Biederman
- 几何离子（Geon, geometrical ions, 几何学上的离子）
  - 所有复杂形状都是由几何离子组成的
- 成分识别(*recognition by components*, RBC)
  - 是物体识别的基础，在复杂形状中找出简单形状



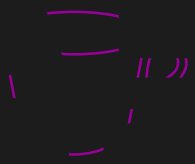


97

- 检验几何子理论：弱化形状 (Biederman, 1985)
  - 一个常见物体65%的轮廓被去除
  - 呈现时间：100毫秒

中间部分的线条被去除，观看者仍能看到基础部分

最顶端的线条被去除，包括重要的连接不同部分的边角



正确率：70%

正确率：50%

98

视觉模式识别

1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

99

特征分析

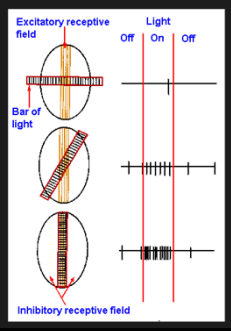
- 模式识别是一种高级信息加工，在此之前，复杂的输入刺激首先要根据其自身的简单特征得到识别
- 类似于自下而上加工
- 两个研究线路的证据
  - 神经的
    - 视皮层的神经细胞对光线的朝向敏感
  - 行为的
    - 眼动研究

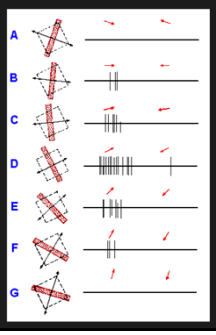
A

/ - \ ^

100

- Hubel & Wiesel
  - 不同细胞对直线，夹角，朝向，运动敏感





101

眼动和模式识别

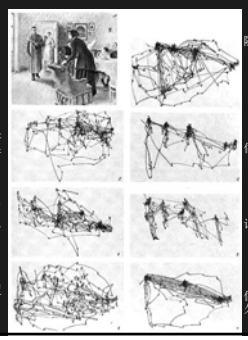
- 眼动研究的原理
  - 如果你在相对较长的时间内凝视着模式中某一特征，那么从中你提取到的信息要多于草草看一眼的特征

102

Yarbus (1967)

- 特征承载的信息越多，双眼停留其上的时间越长
- 注视点的分布与被试的意图有关

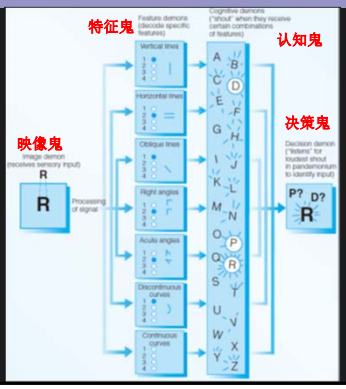
I. E. Repin
"An Unexpected Visitor"
列宾 《不期而至》



103

- Selfridge's Feature-Matching Model

魔域模型
The Pandemonium Model



104

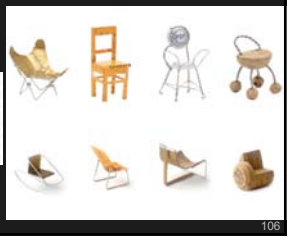
视觉模式识别

1. 格式塔理论
2. 典型表象
3. 自下而上加工与自上而下加工
4. 模板匹配
5. 特征分析
6. 原型匹配

105

原型匹配

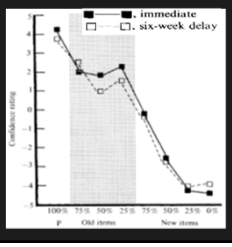
- 长时记忆中储存着某种抽象的模式作为原型
- 原型不仅仅是一组刺激的抽象物，而且更是模式的一个典型的或“最佳”的表征
- 寻找相似性，而非完全匹配

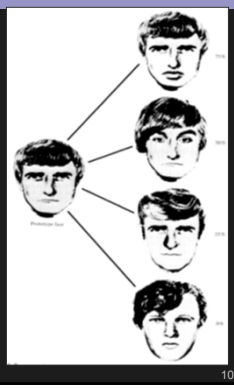


106

伪记忆

- Solso & McCarthy (1981)
  - 被试误将原型认作先前见过的图形
  - 原型形成的基础是频繁经验过的特征





107

- 获得有关模式的知识的过程包括两个阶段:
  - 获得模式中有关特征的信息 (快)
  - 获得特征之间关系的信息 (慢)

108

- 视觉原型形成及其作用的规律:
  1. 在对样例特征的平均的基础上形成原型;
  2. 只接触样例仍可获得原型信息;
  3. 熟悉原型比不熟悉（或新学会）原型变化范围更小;
  4. 以与原型的转换距离来对样例做判断;
  5. 通过对样例抽象形成原型；通过各种形式的原型之间以及原型与样例之间的距离来评估它们的关系.

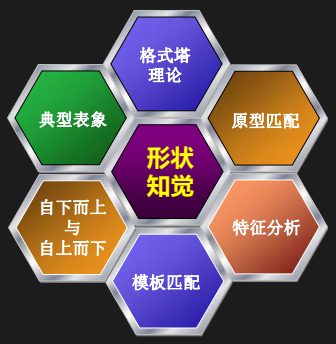
109

原型理论

- 趋中模型
  - 原型表征的是一组样例的平均数或均值
  - Posner & his colleagues
- 特征-频率模型
  - 原型表征的是众数或最常见的特征组合
  - Solso et al.

110

形状知觉：一条整合的途径



- 这些理论非“对”非“错”
- 互补的

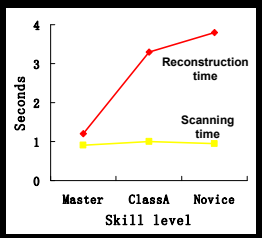
111

Pattern recognition among experts

- de Groot (1965,1966)
  - master and weaker players thought ahead and considered about the same number of moves, and had a similar search for patterns of moves.

112

- Chase & Simon (1973)
  - Masters have greater capacity to reproduce the pattern because they are able to encode the bits and pieces into chess schemata.



113

面孔识别（补充）

“每只蚂蚁，都有眼睛鼻子，它美不美丽，偏差有没有一毫厘，有何关系……”

——林夕 《开到荼蘼》

114

- 面孔识别的本族效应
- 虚构“红/绿人格”，对同类型人格的面孔识别也会更好

115

Holistic processing

- 面孔识别和物体识别的最大差异之一在于面孔识别是整体加工的（holistic processing）
- 我们不太可能依赖面部的局部信息来识别面部，因为：
  1. 不同的人有相同的面部信息（比如，瞳孔的颜色）
  2. 个体的面部信息可能改变（比如阴影、嘴巴形状）

116

- 支持面孔识别整体加工的证据（McKone, Kanwisher, & Duchaine, 2007）：
  - 颠倒效应（inversion effect）
  - 部分-整理效应（part-whole effect）
  - 合成效应（composite effect）

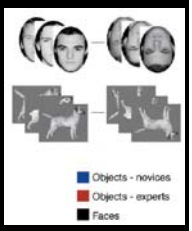
117

颠倒效应（inversion effect）

118

McKone (2004)

- 首先训练被试从6个角度学习两张面孔(0°，45°，67.5°，90°，112.5°，和 180°)。
- 测试时，快速（150毫秒）向被试呈现上述训练图片，要求被试判断图片是哪张面孔。
- 结果：
  - 被试对正立的面孔的识别正确率显著高于倒立面孔
  - 对于非面孔物体的识别，颠倒效应非常小，且随着练习快速消失
- 解释：人们对正立面孔的识别为整体加工，对颠倒面孔（或其他方向呈现的面孔）的识别为部分加工

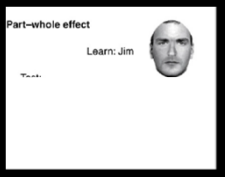


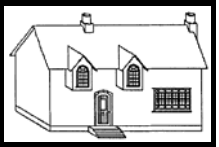
119

部分-整理效应（part-whole effect）

Tanaka & Farah (1993)

- 被试学习一组面孔并记住面孔对应的名字
- 测试：
  - 全脸测试：“哪个是Jim？”
    - 两张面孔仅有一处特征不同，如眼睛
  - 单独特征测试：“哪个是Jim的眼睛？”
    - 两个特征（如眼睛）
- 结果：
  - 正立：全脸测试正确率高于单独特征测试
  - 颠倒：两种条件测试成绩没有差异
  - 当学习和测试的图片为一般物体（如房屋），也没有部分整体效应出现





120

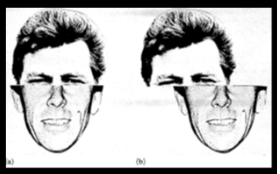
合成效应（composite effect）

Young, Hellawell, & Hay（1987）

•5张名人面孔混合拼接成合成面孔

•被试判断合成面孔的上半部分或下半部分属于哪一位名人

- 结果：
  - 正置：对齐（图a）慢于错开（图b）
  - 倒置：无差异
  - 其他物体识别：无上述合成效应

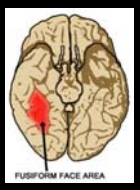


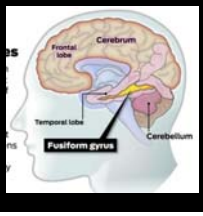
即使是合成面孔，被试对其的加工也是整体性的，且这种整体加工方式阻碍了被试对局部特征的判断

121

梭状回面孔区（fusiform face area）

- fMRI研究显示，相比于，人们观看面孔图片时，该区域的神经兴奋程度是观看其他类型图片时的神经兴奋程度的2倍（McKone, et al.，2007）。
- 面孔失认症病人的受损脑区普遍包括该区域（Barton, Press, Keenan, & O'Connor, 2002）





122

123