

# 读心术——视听信息的脑神经编码

---

---

胡新韬

自动化学院

2020-5-20

# 读心术——伪科学还是黑科技？



古今中外，“读心术”层出不穷。无一例外，被斥为伪科学，如过街老鼠。

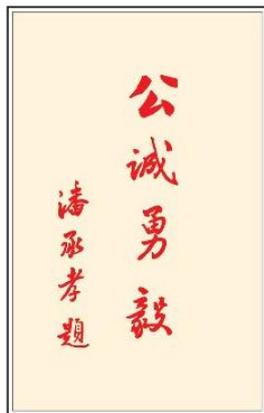
# 读心术——伪科学还是黑科技？



古今中外，“读心术”层出不穷。无一例外，被斥为伪科学，如过街老鼠。

西北工业大学 校训

西北工业大学 校风



“三实一新”

基础扎实

工作踏实

作风朴实

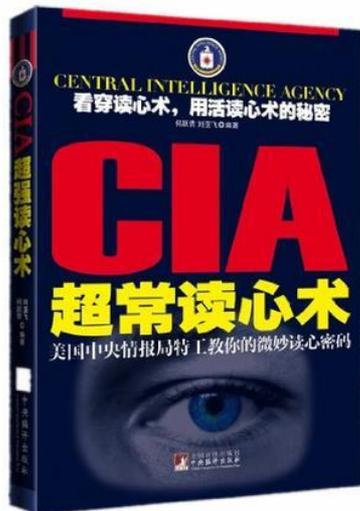
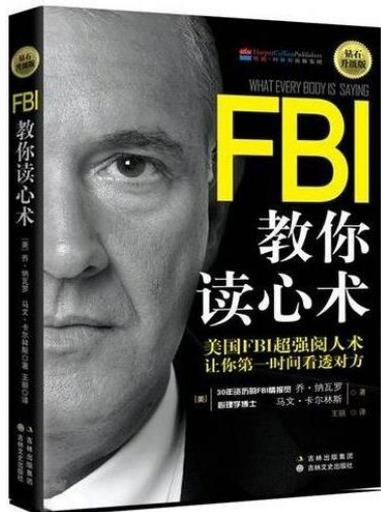
开拓创新



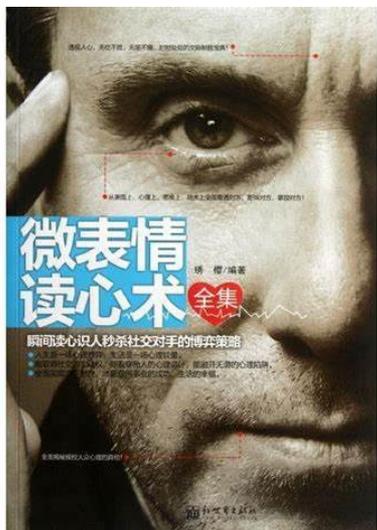
我读书不少，你不要骗我！

老师宣扬伪科学，我要举报！！！！

# 读心术——伪科学还是黑科技？



根据心理学原理，利用微表情、微动作等“读心”



老师不务正业，我要举报！！！！

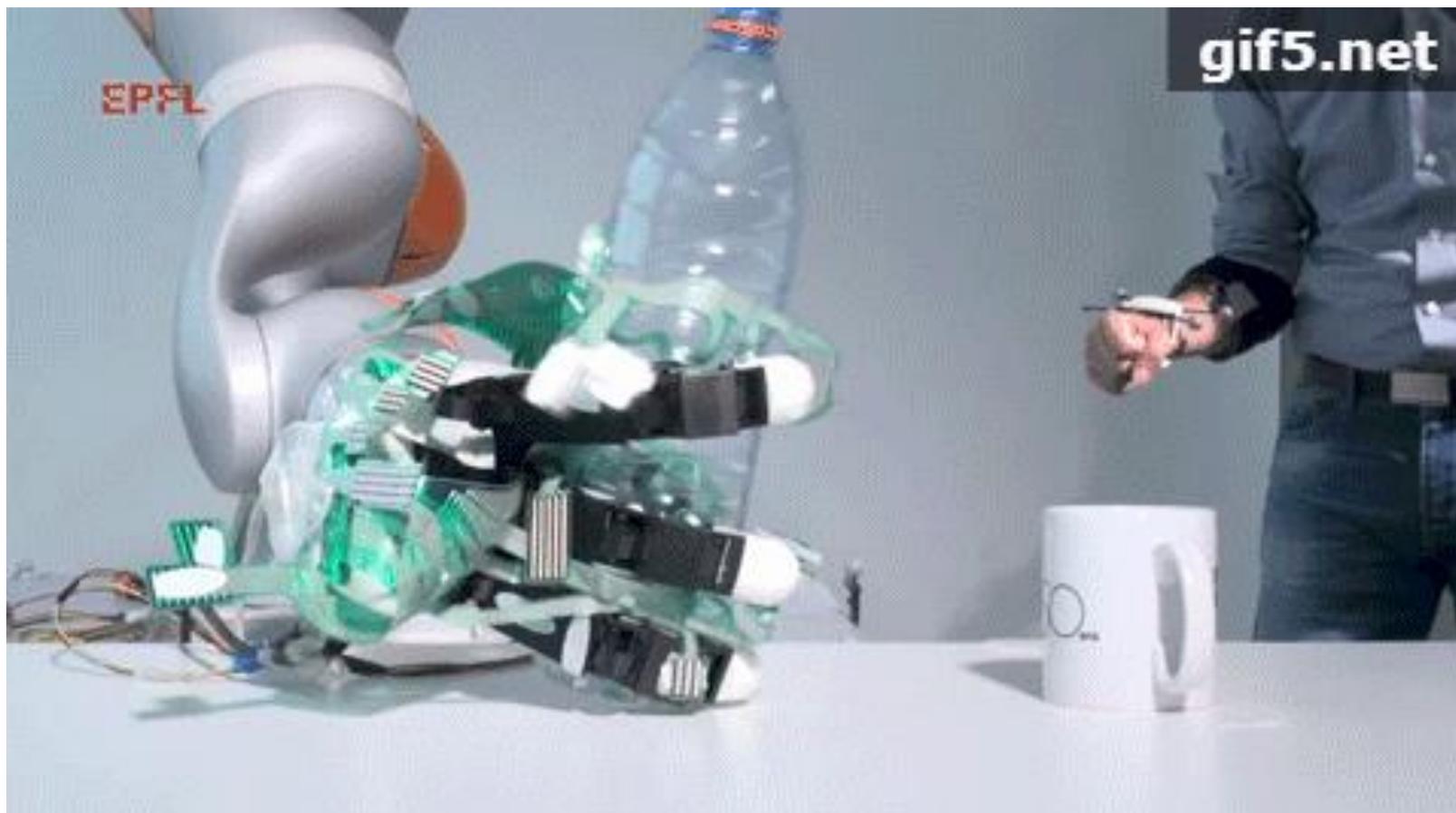
# 读心术——伪科学还是黑科技？



我们在一只名叫Aurora的猴子身上开始试验

猴子利用BCI，通过“意念”控制机械臂  
脑机接口（Brain-computer Interface, **BCI**）

# 读心术——伪科学还是黑科技？



**BCI** (Brain-computer Interface) + **AI** (Artificial Intelligence)

EPFL: 人负责有意识动作, AI负责无意识动作。

# 读心术——伪科学还是黑科技？



在最好的情况下，单词错误率为5%

【“毛骨悚然”的里程碑】

**BCI** (Brain-computer Interface) + **AI** (Artificial Intelligence)

想法转换成文本

nature  
neuroscience

Technical Report | Published: 30 March 2020

## Machine translation of cortical activity to text with an encoder–decoder framework

Joseph G. Makin , David A. Moses & Edward F. Chang 

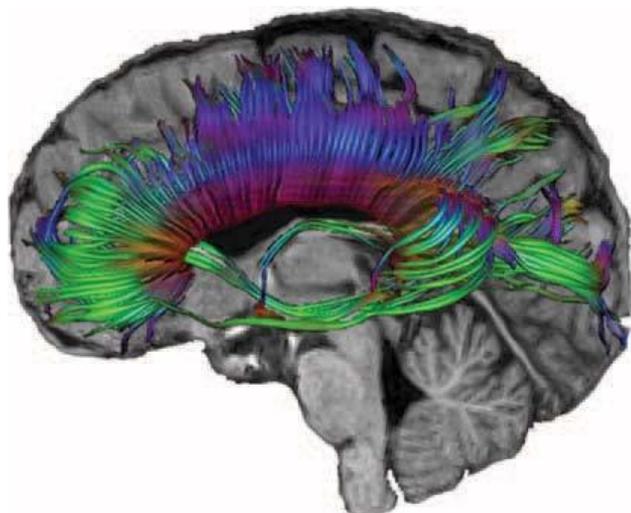
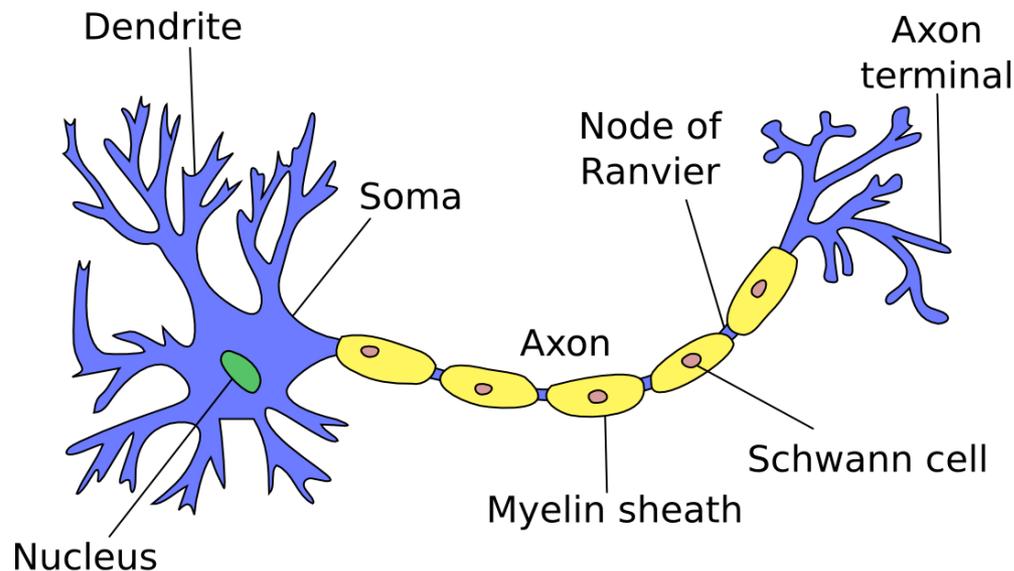
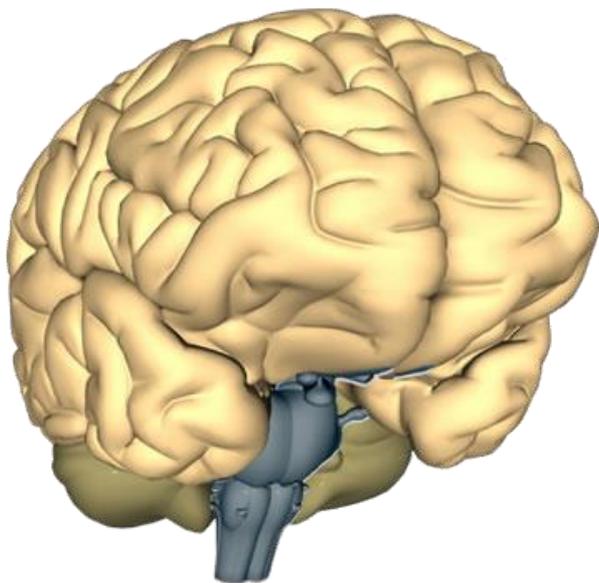
*Nature Neuroscience* (2020) | Cite this article

3849 Accesses | 1 Citations | 818 Altmetric | Metrics



华裔科学家、  
加州大学旧金山分校教授  
Edward  
Chang

# 读“心”术——It's all about the brain



体积：~1.3升

重量：~1.5千克

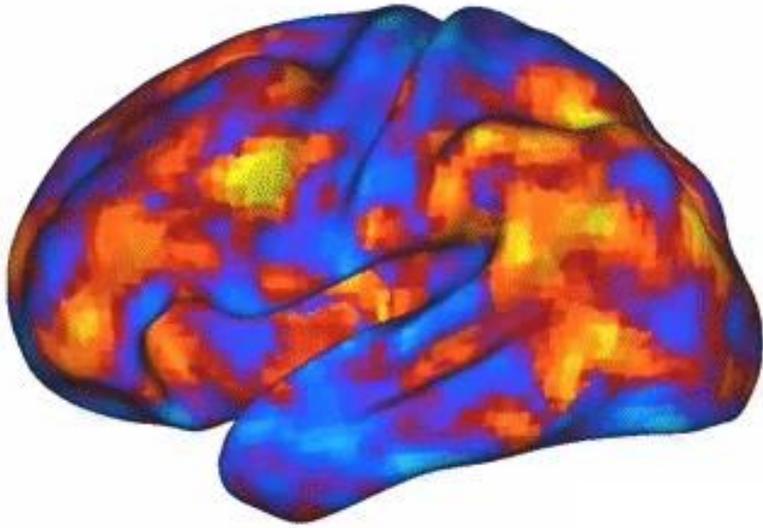
皮层表面积：~1.6m<sup>2</sup>

神经元：800~1000亿

神经连接：100兆（10W公里轴突）

# 读心术——基本要素

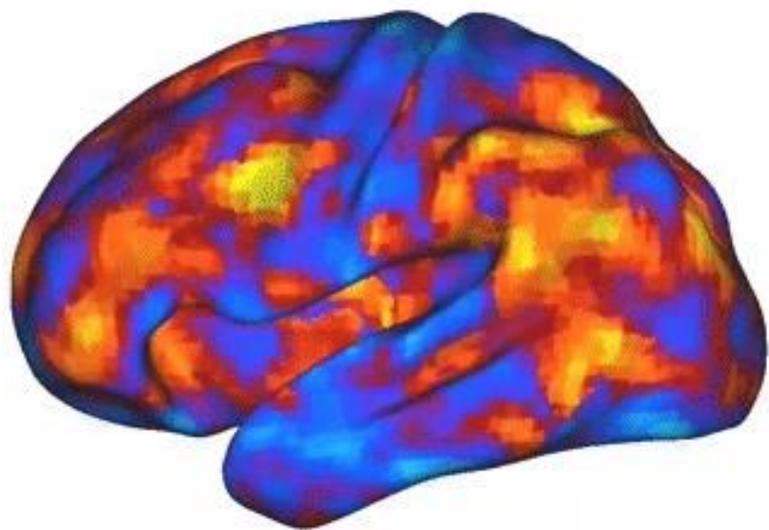
---



脑活动的观测、记录与分析

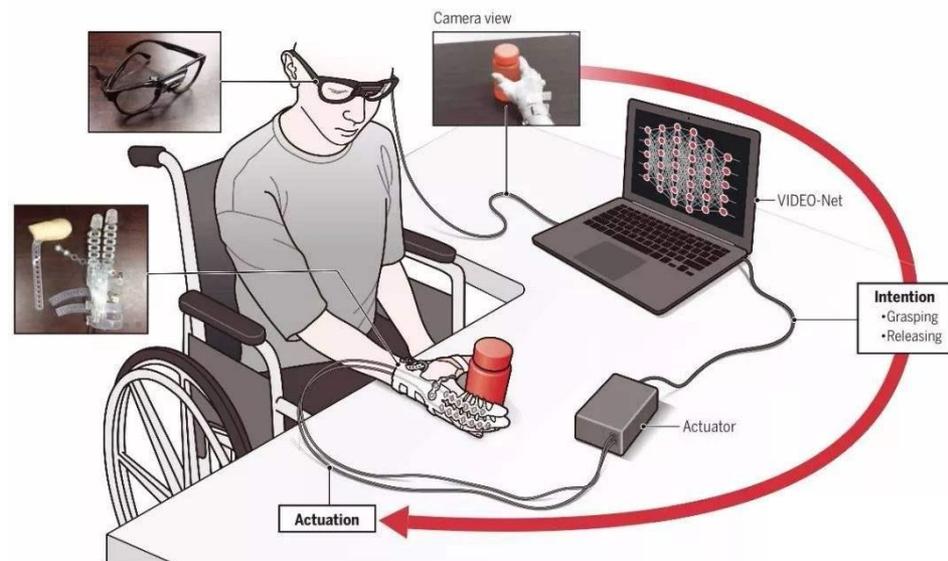
信号处理问题

# 读心术——基本要素



脑活动的观测、记录与分析

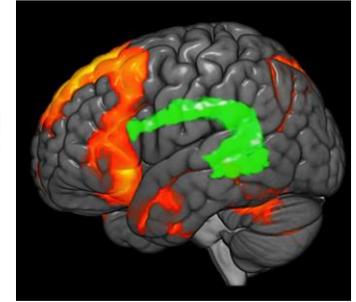
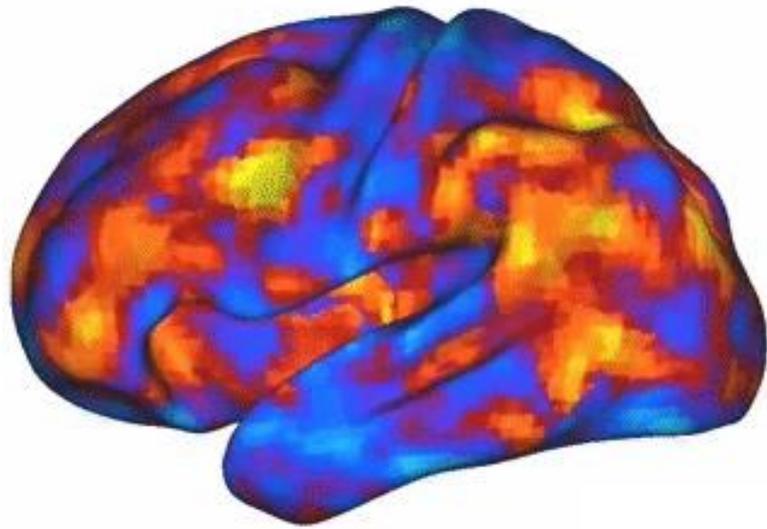
信号处理问题



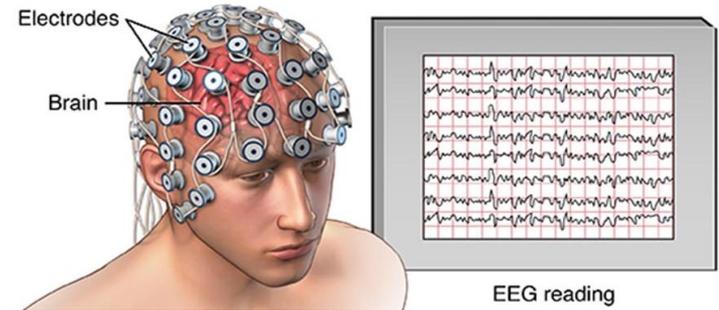
脑状态识别

模式识别问题

# 读心术——基本要素



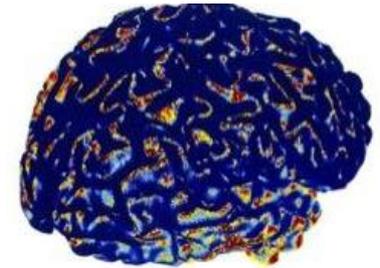
功能磁共振 (fMRI)



脑电图 (EEG)

脑活动的观测、记录与分析

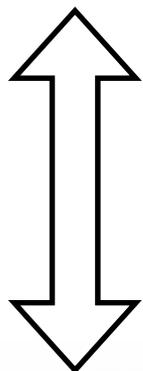
信号处理问题



脑磁图 (MEG)

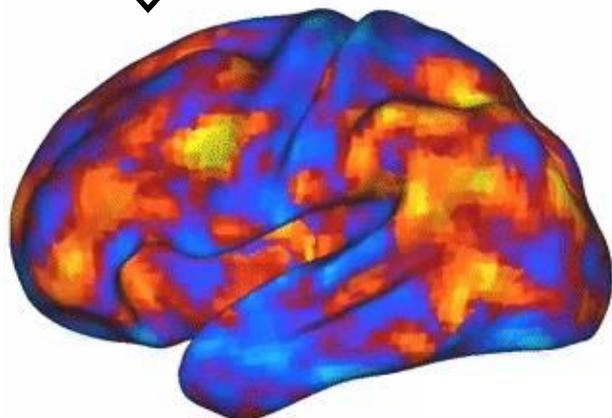
# 读心术——基本要素

看到的、听到的、想到的...

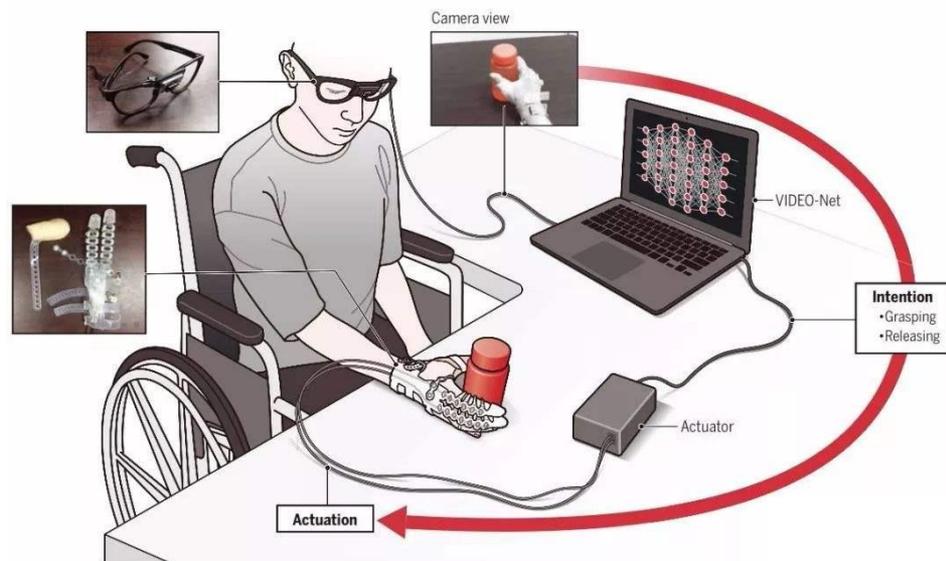


□ 映射

- 回归
- 分类



脑活动



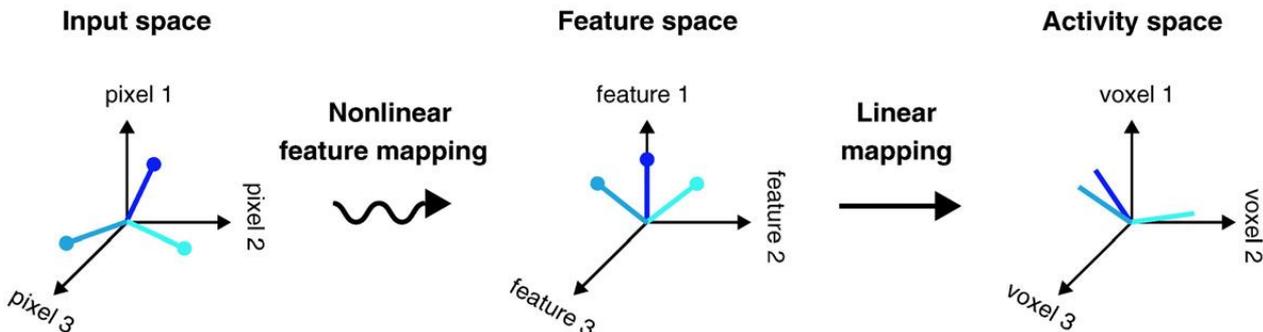
脑状态识别

模式识别问题

# 读心术——基本要素（一般化模型）

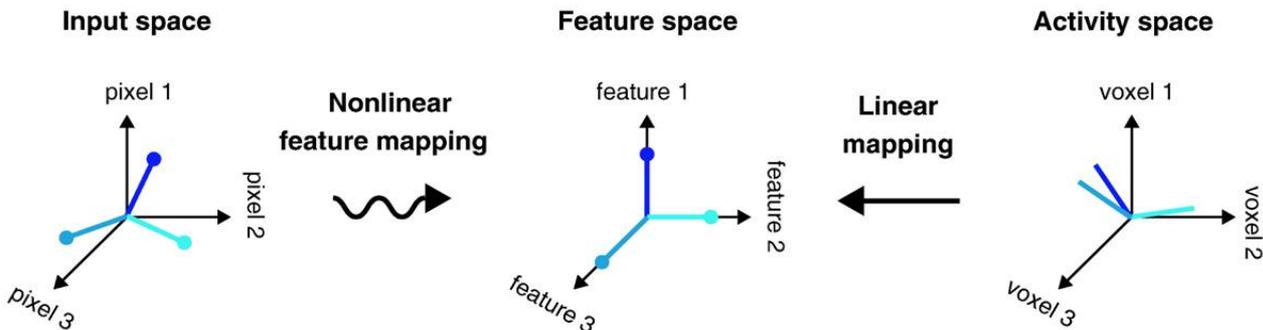


Linearizing encoding model



神经编码模型  
外界输入→脑活动

Linear classifier



神经解码模型  
脑活动→外界输入

# 视觉信息神经编解码——场景分类 (Classification)

Beaches



Buildings



Forrest



Highways



Industry



Mountains



输入空间: 图片

特征空间: 类标记

脑活动空间: fMRI体数据

映射方法: 支持向量机SVM

# 视觉信息神经编解码——场景分类 (Classification)

Beaches



Buildings



Forrest



Highways



Industry



Mountains



输入空间: 图片

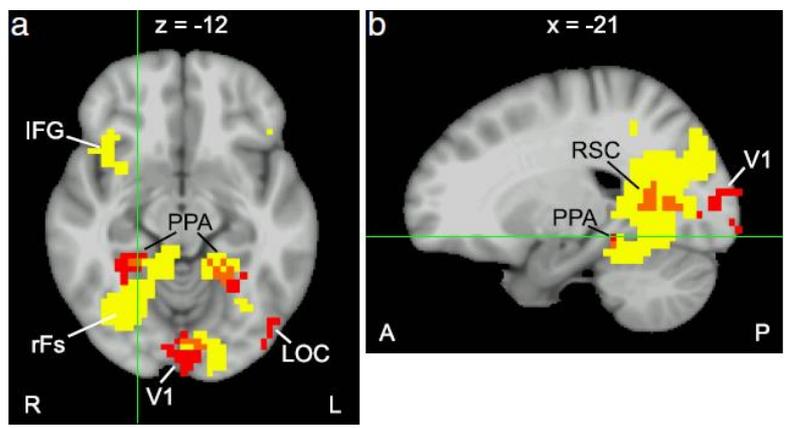
特征空间: 类标记

脑活动空间: fMRI体数据

映射方法: 支持向量机SVM

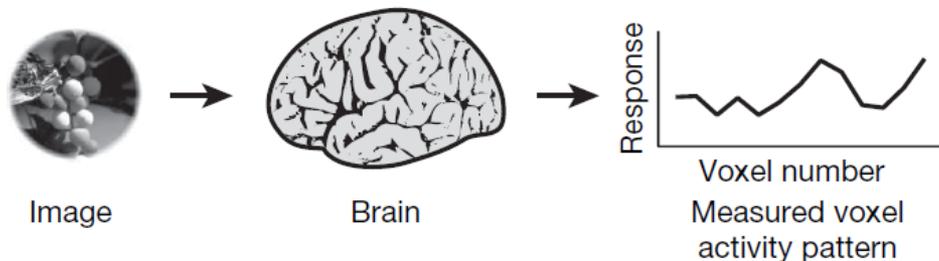
ROI	Decoding accuracy
V1	26%*
FFA	22%
LOC	24%*
RSC	27%*
PPA	31%**

**What?**

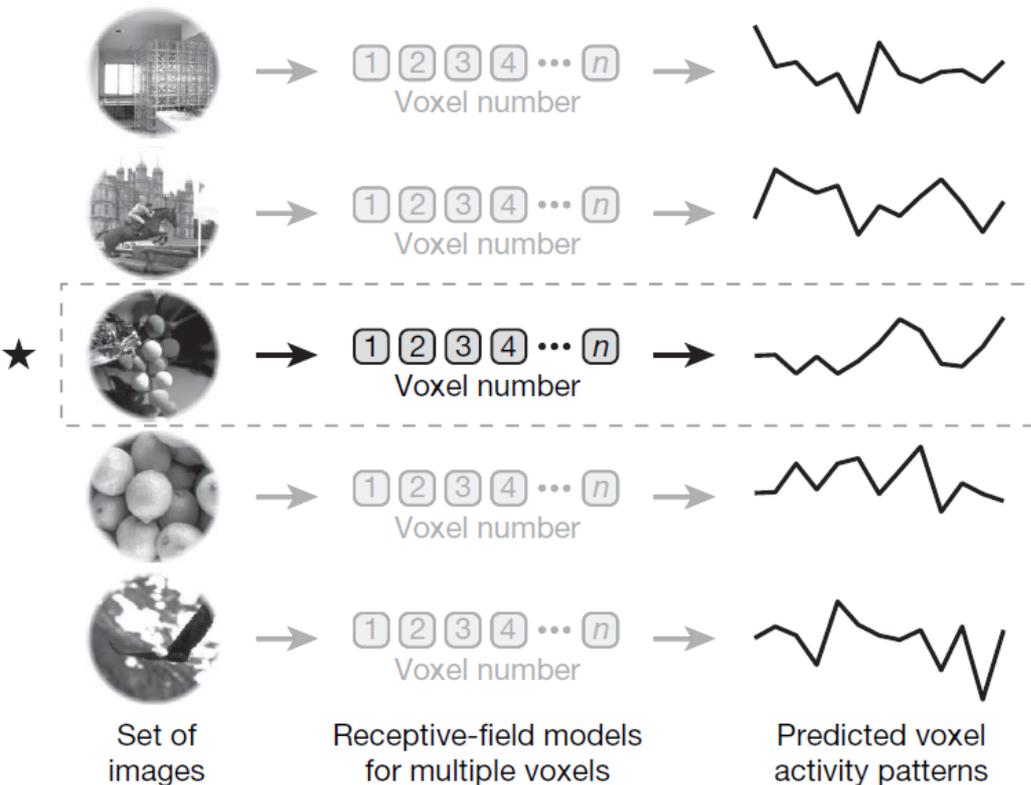


# 视觉信息神经编解码——图像辨认 (Identification)

(1) Measure brain activity for an image



(2) Predict brain activity for a set of images using receptive-field models



输入空间：图片

特征空间：Gabor滤波器组

脑活动空间：fMRI体数据

映射方法：支持向量机回归

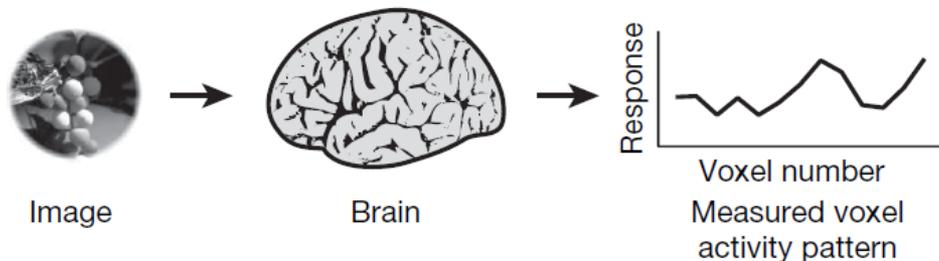
准确率：

120张中找1张：92%

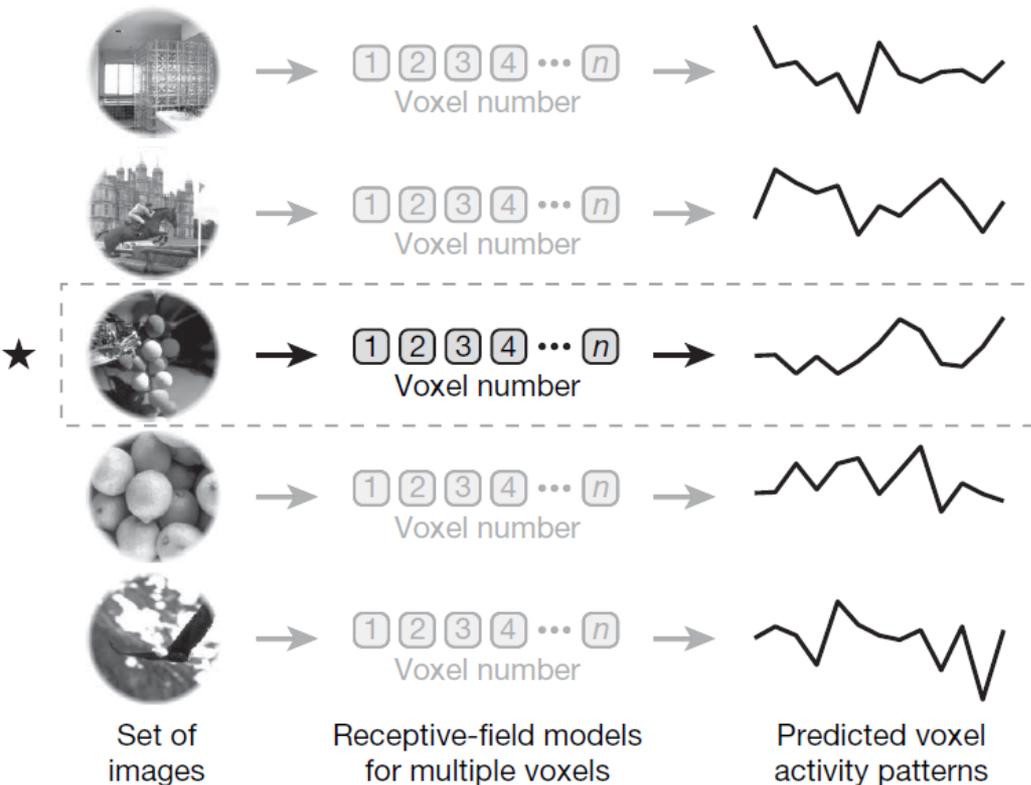
1000张中找1张：82%

# 视觉信息神经编解码——图像辨认 (Identification)

(1) Measure brain activity for an image



(2) Predict brain activity for a set of images using receptive-field models



输入空间：图片

特征空间：Gabor滤波器组

脑活动空间：fMRI体数据

映射方法：支持向量机回归

准确率：

120张中找1张：92%

1000张中找1张：82%

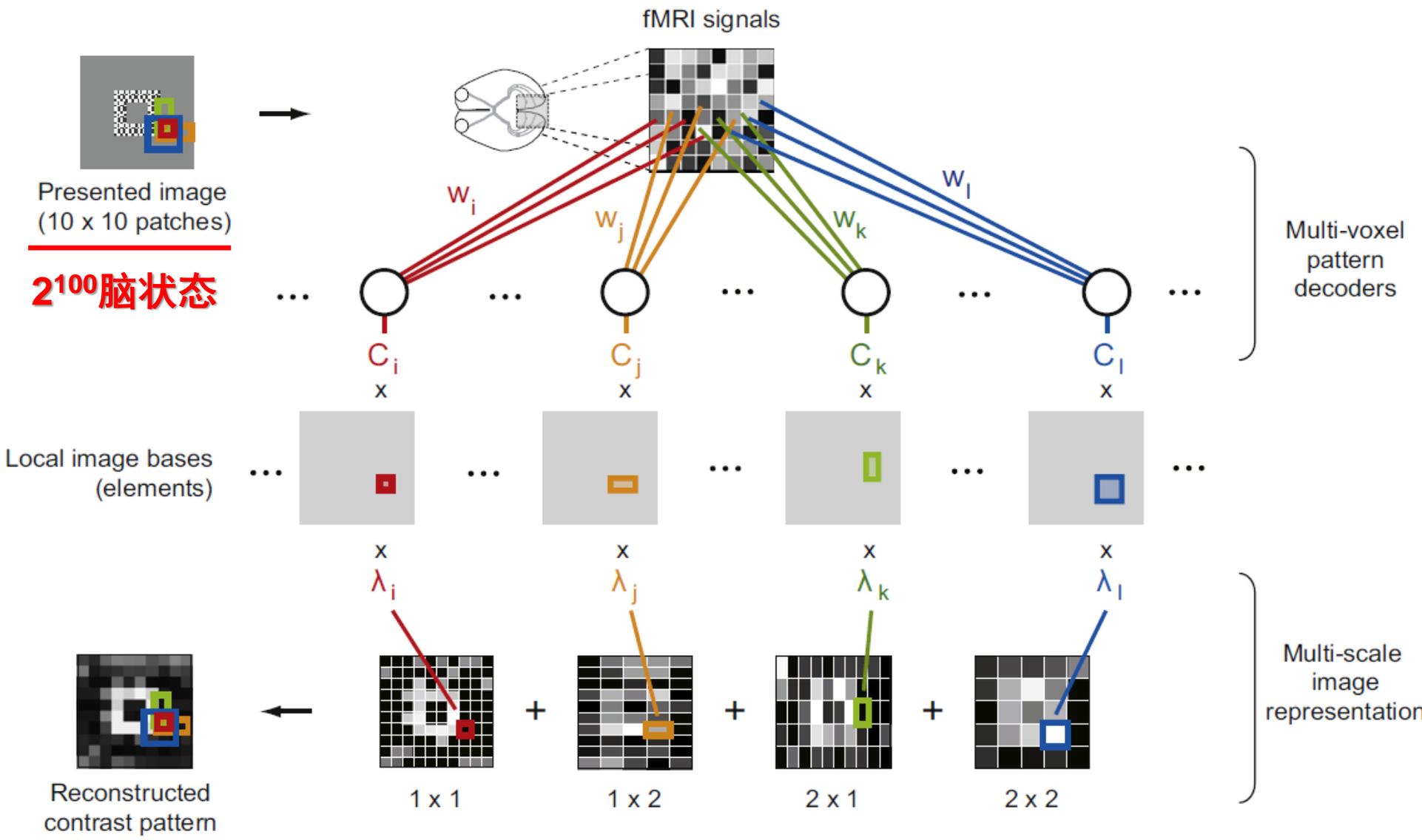


?

不过如此

**重点：**系统的验证了主视觉皮层（V1）对边缘朝向（Gabor滤波器组）的感知！

# 视觉信息神经编解码——图像重构 (Reconstruction)

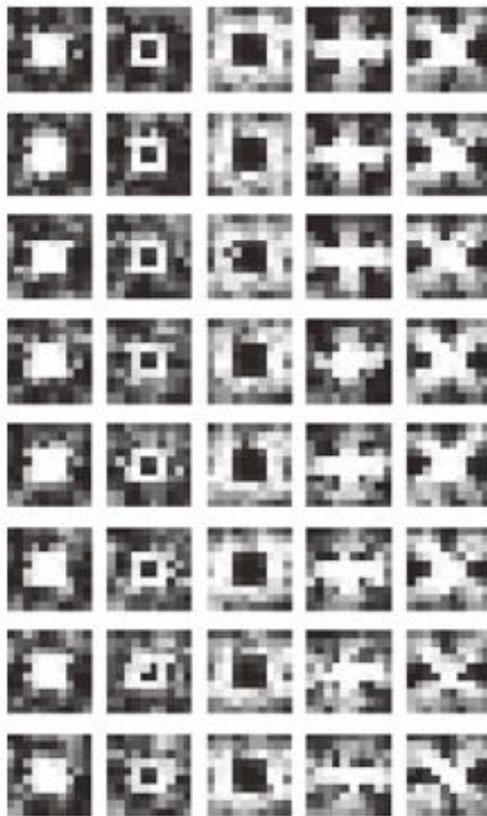


# 视觉信息神经编解码——图像重构 (Reconstruction)

Presented contrast pattern

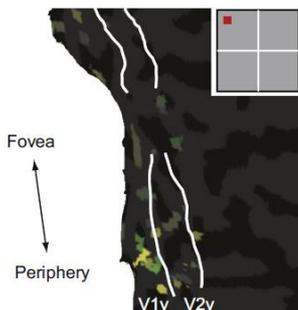
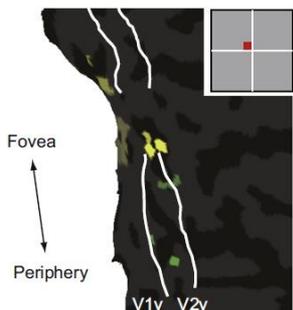


Reconstructed contrast pattern



Foveal patch decoder

Peripheral patch decoder

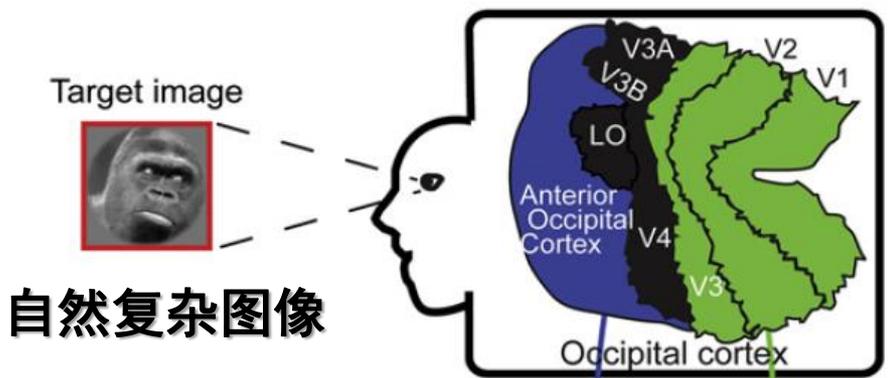


**重点:** 视觉信息在大脑中的多尺度表达

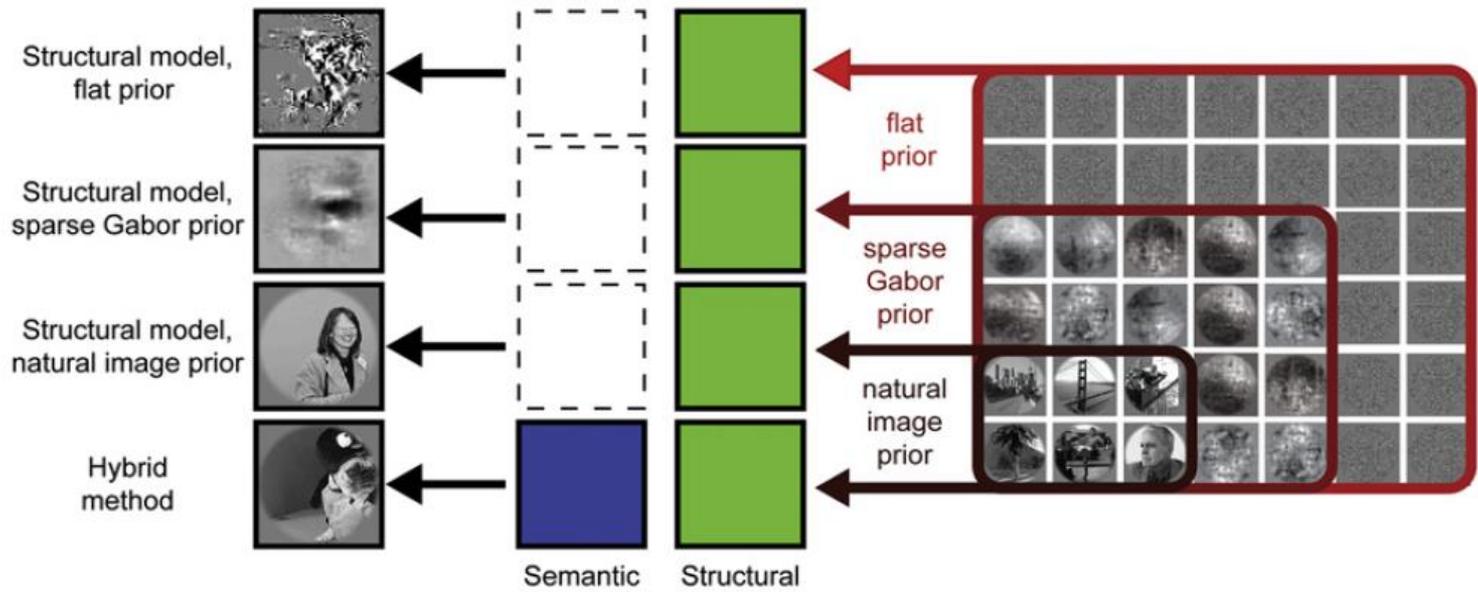
Mean of reconstructed contrast pattern



# 视觉信息神经编解码——图像重构 (Reconstruction)



- Flat prior: 无先验信息限制
- Gabor prior: 低阶统计先验
- Natural image prior: 自然图像先验
- Semantic prior: 语义信息先验



Reconstructions

Encoding Models

Image Priors

# 视觉信息神经编解码——图像重构 (Reconstruction)

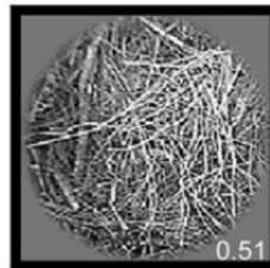
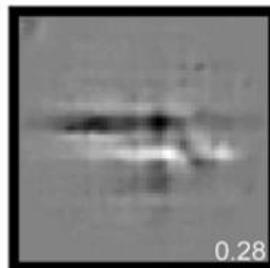
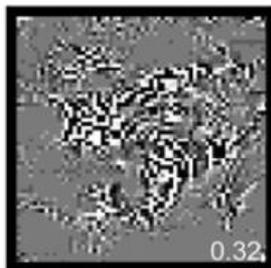
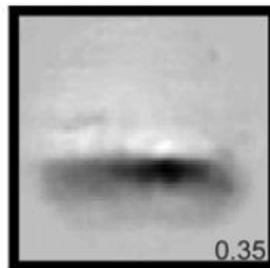
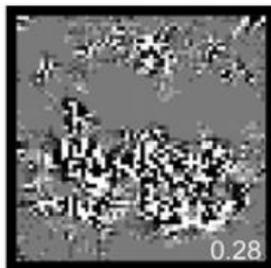
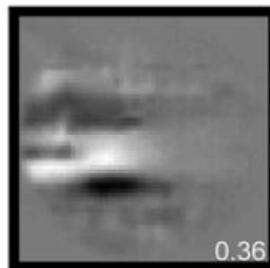
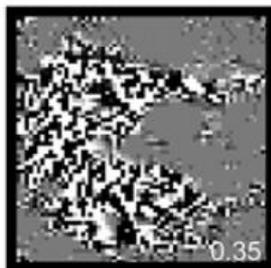
Target image

Reconstructions with structural encoding model

Flat prior

Sparse Gabor prior

Natural image prior



结构先验重构效果

# 视觉信息神经编解码——图像重构 (Reconstruction)

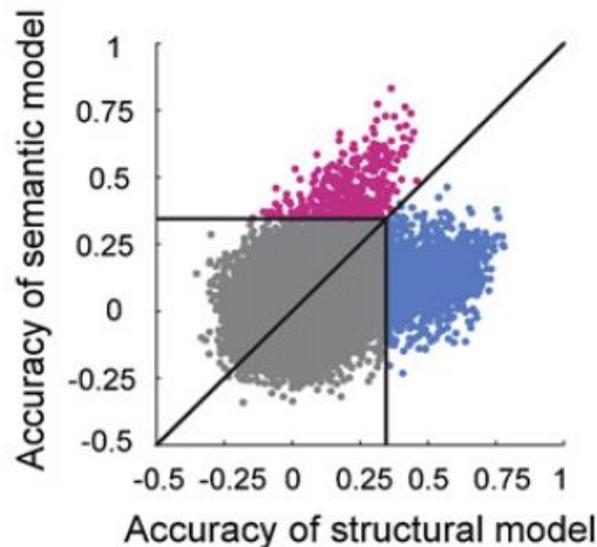
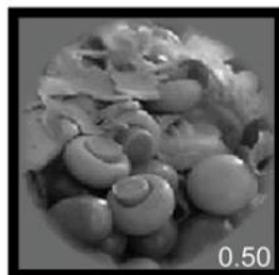
结构先验 + 语义先验

Target image

Reconstructions with natural image prior

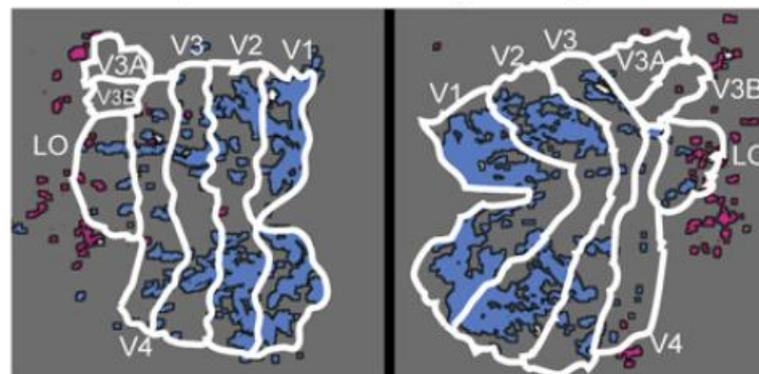
Structural model only

Structural and semantic models (hybrid method)



Left occipital cortex

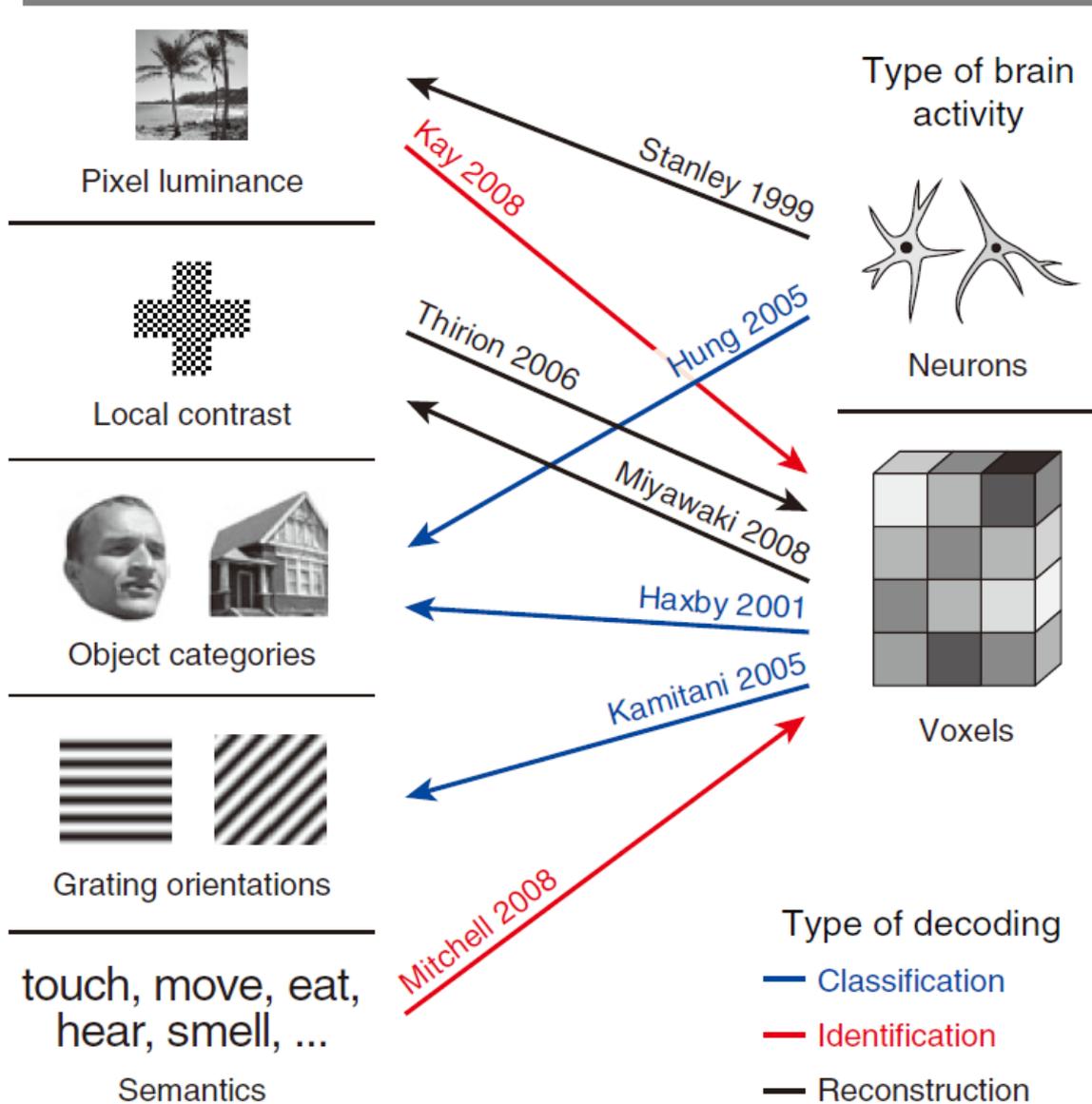
Right occipital cortex



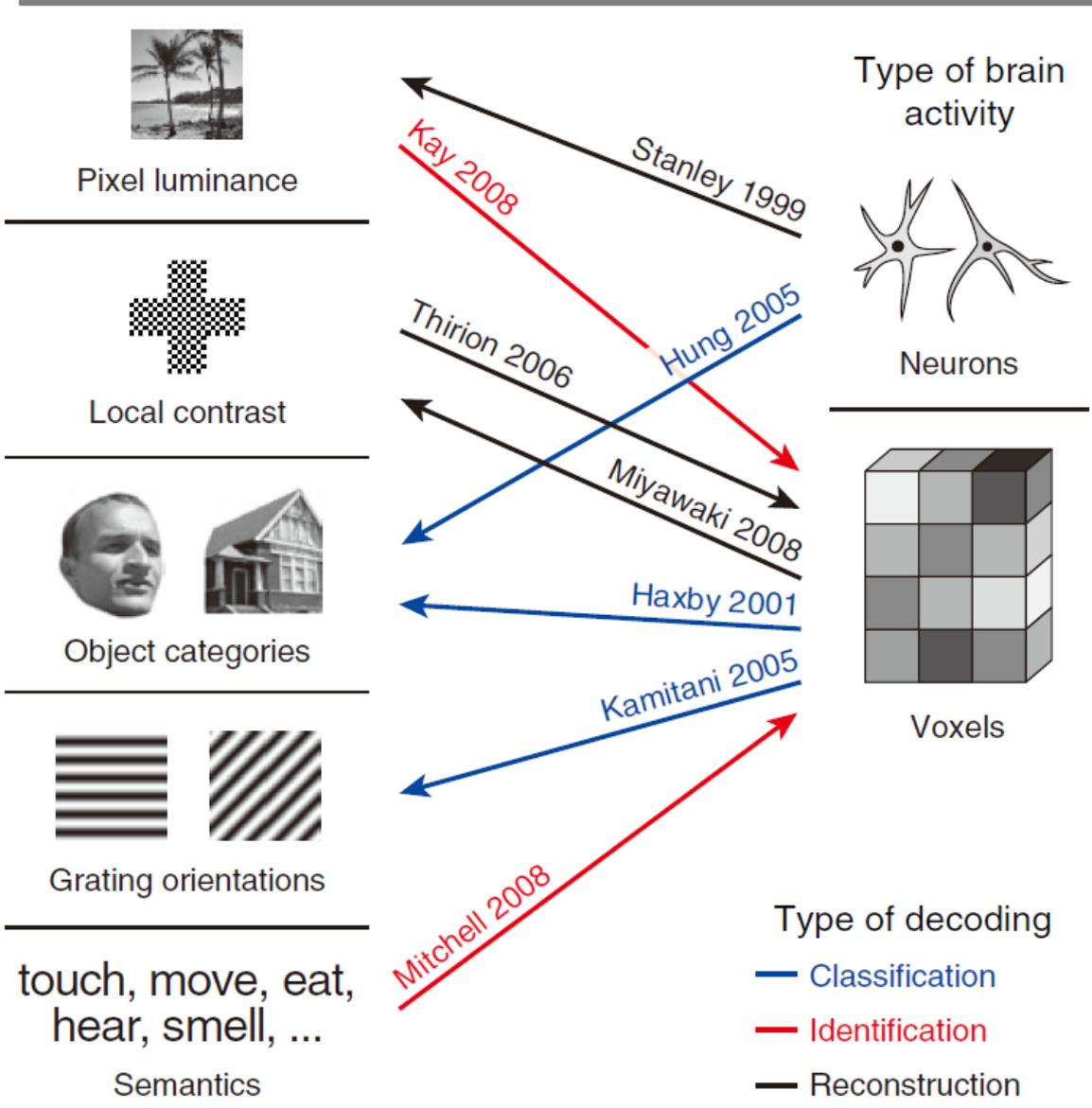
结构信息: V1-V3

语义信息: Anterior visual area

# 视觉信息神经编解码

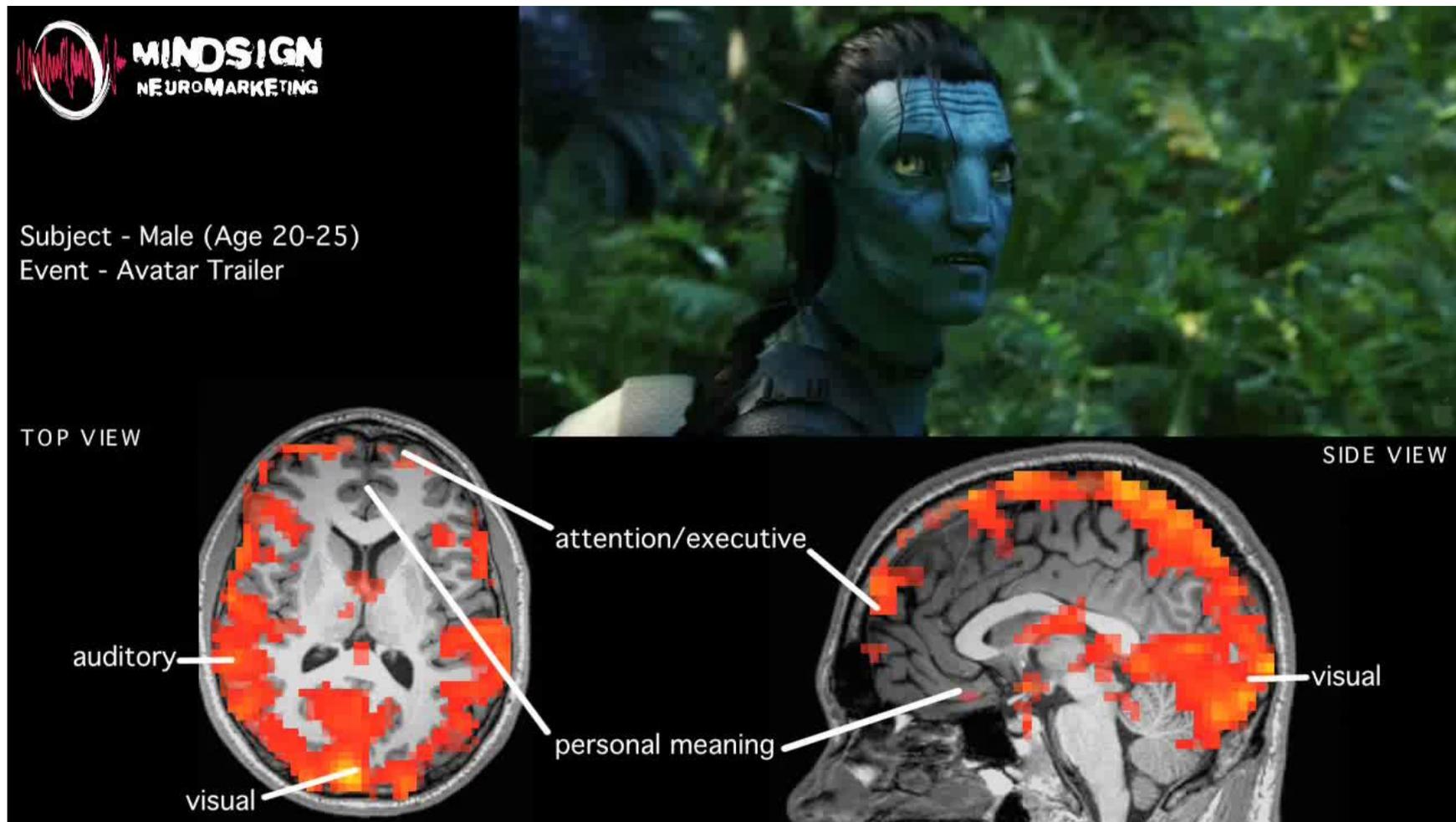


# 视觉信息神经编解码



- 神经编解码模型是分析手段，重要的是如何利用神经编解码模型，回答视听觉认知的相关科学问题。
- 催生了很多重要的科学发现！
- 能服务于（计算机）视觉信息分析么？

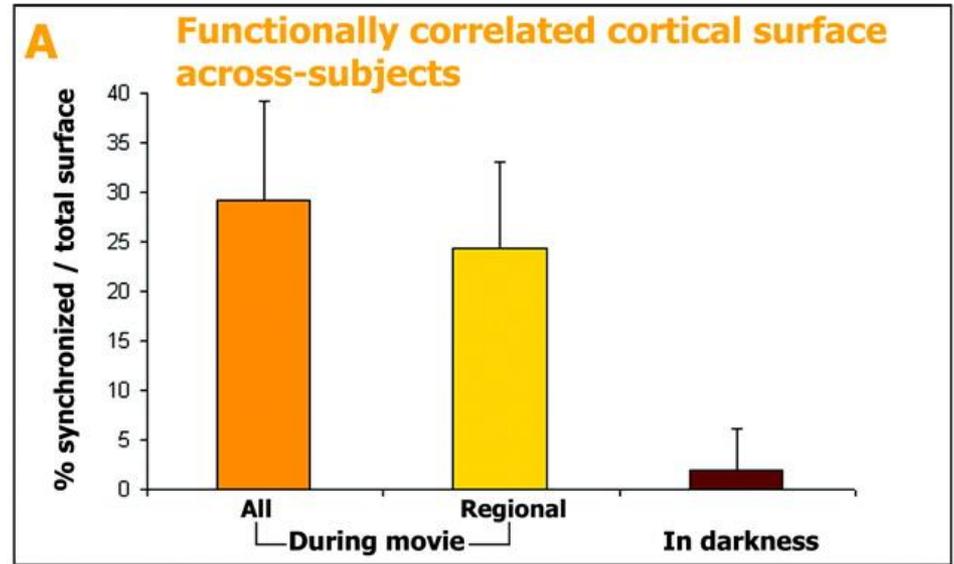
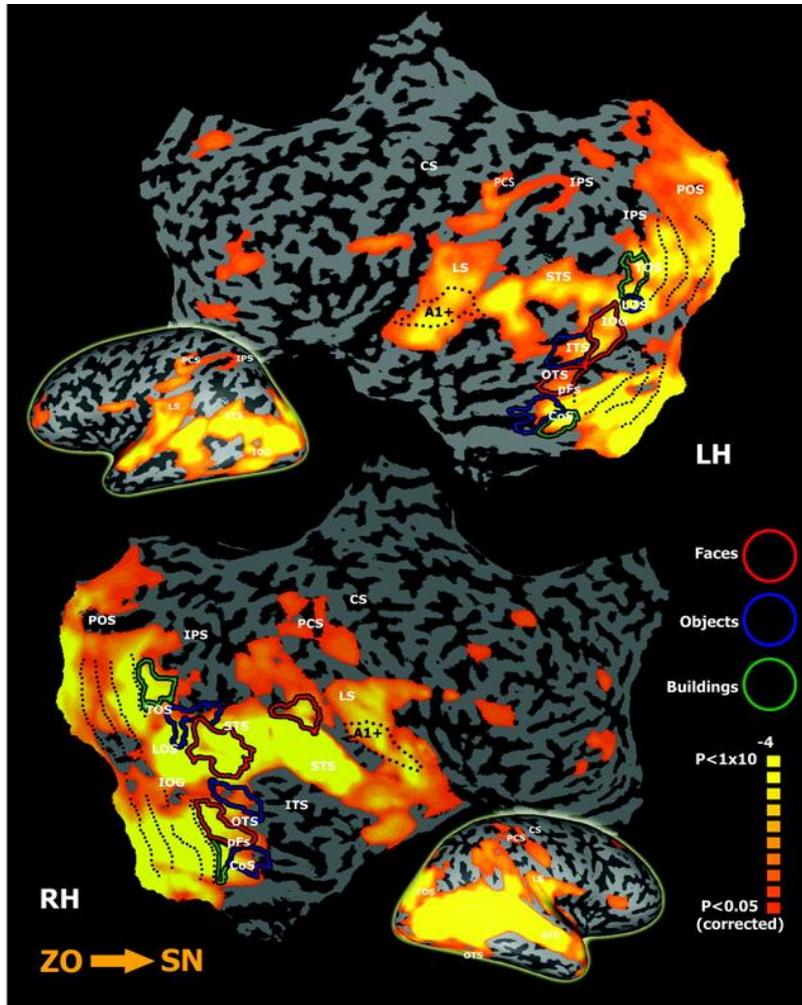
# 脑功能引导的视觉信息分析



依据1：脑成像技术可有效量化视频理解过程中的脑功能响应

Hasson, Science, 2010.

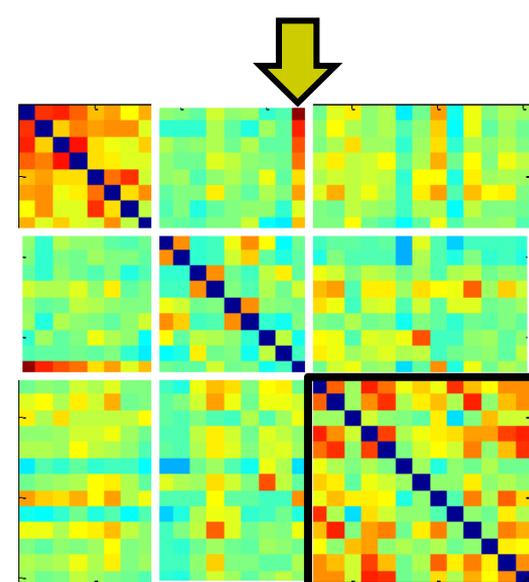
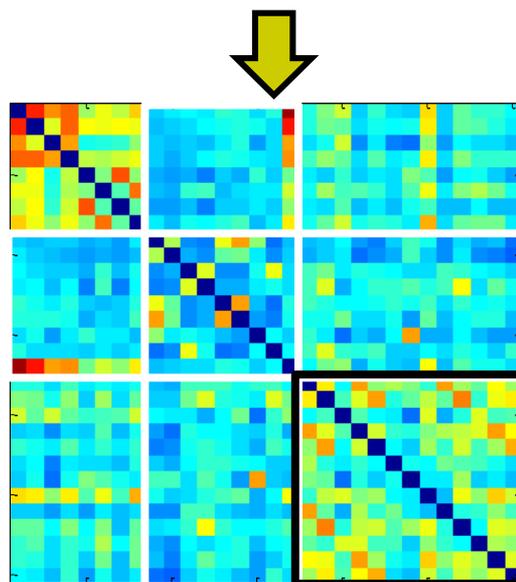
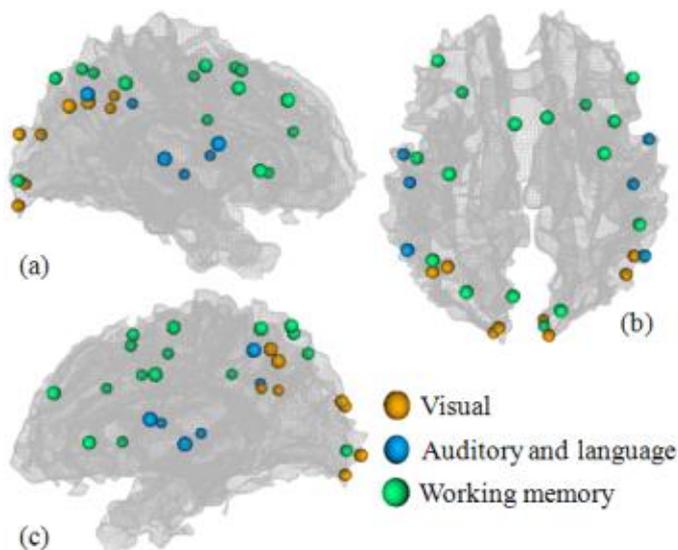
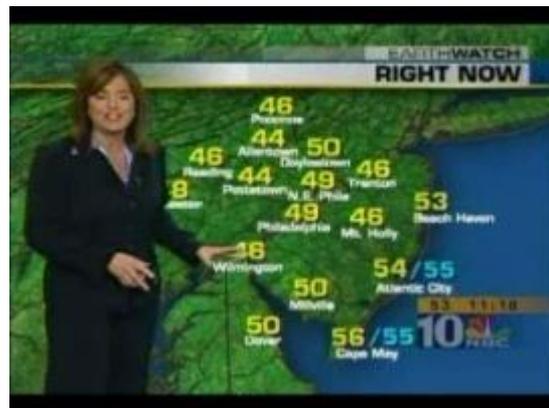
# 脑功能引导的视觉信息分析



同一视频刺激下，不同被试在某些脑区，具有类似的脑功能活动（可重复性高）  
(inter-subject correlation)

依据2：视频刺激下的脑功能活动具有可重复性

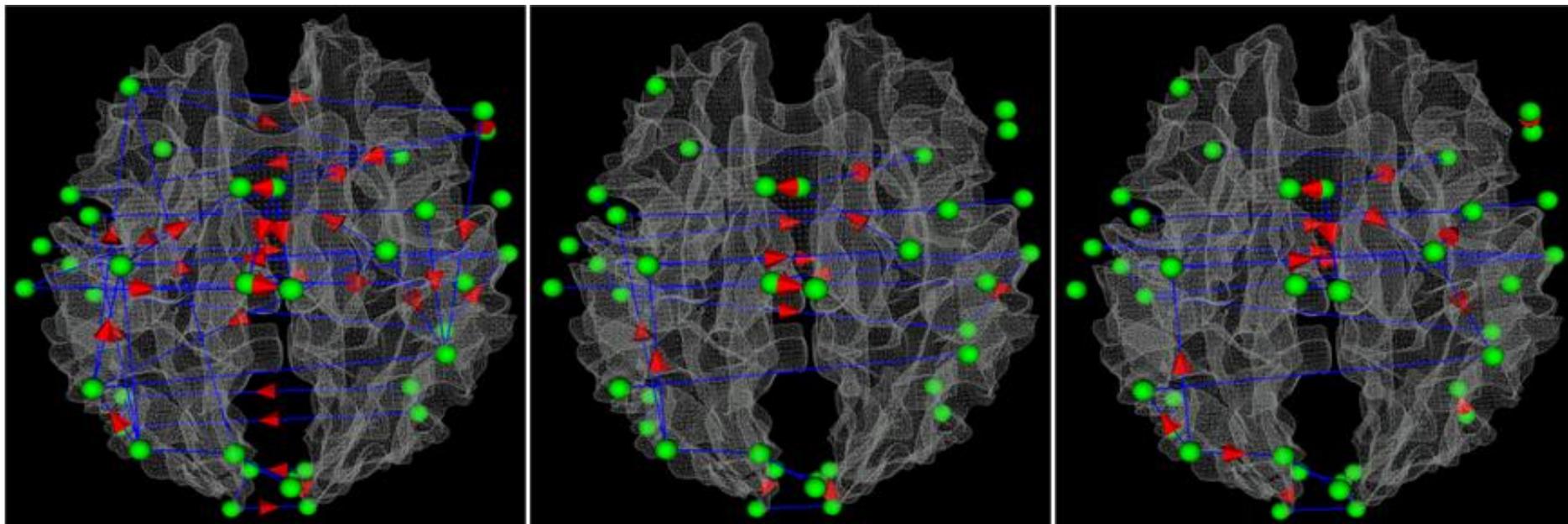
# 脑功能引导的视觉信息分析



**依据3：脑功能响应特征具有很高辨识度**

Hu X, IEEE TMM, 2012; ACM Multimedia, 2010. (最佳学生论文奖提名)

# 脑功能引导的视觉信息分析



广告

体育运动

天气预报

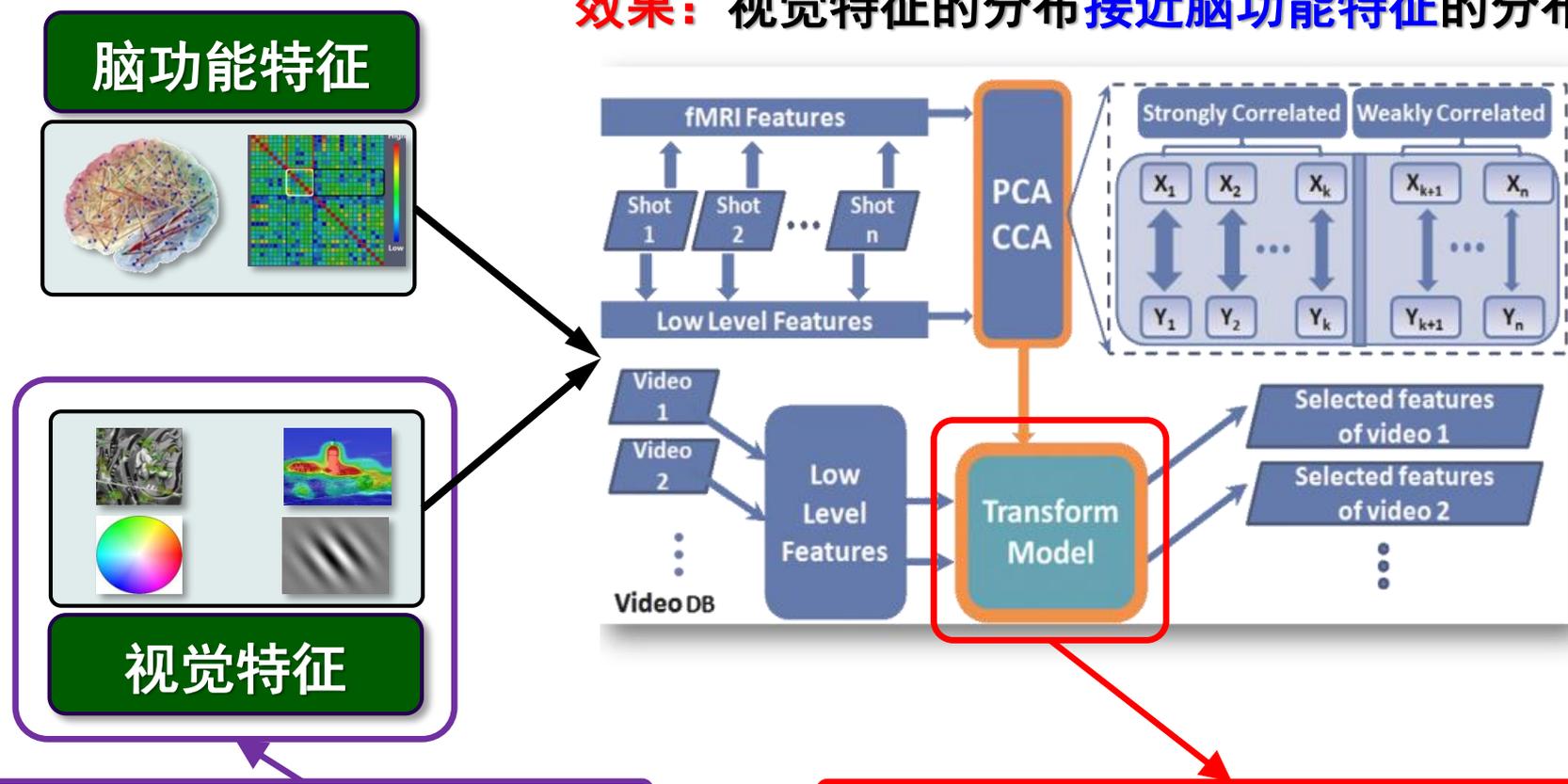
**依据4：理解不同类型视频内容的脑神经环路不同**

# 脑功能引导的视觉信息分析——视频分类

□ **创新方法：**建立了脑功能特征与视觉特征的**协同计算方法**

基于典型相关分析的融合学习

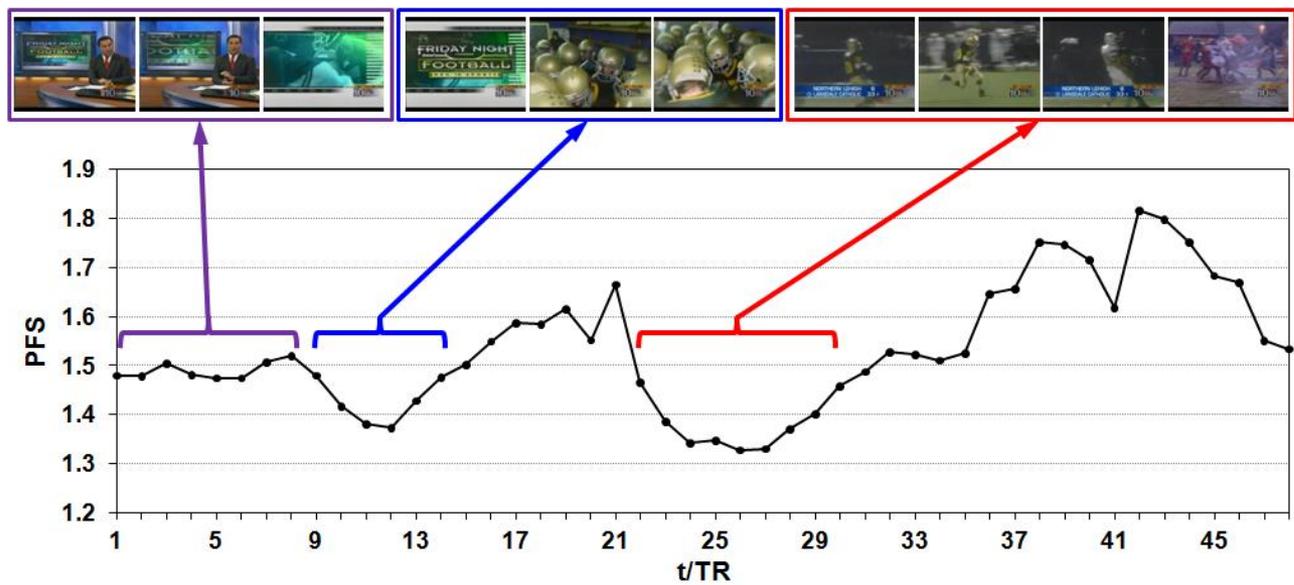
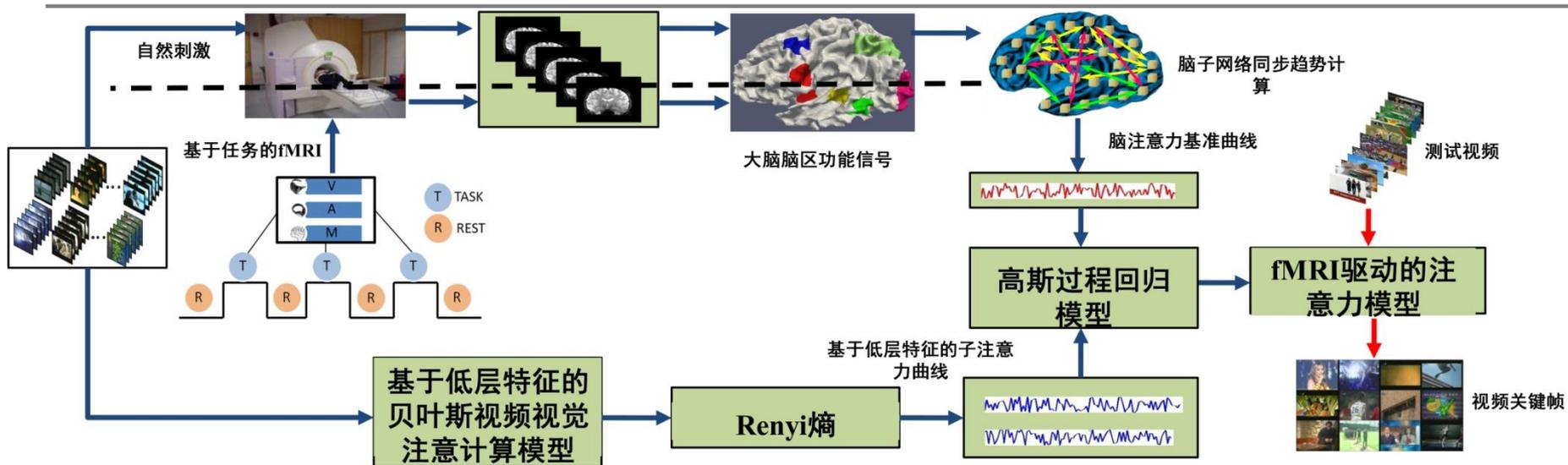
**效果：**视觉特征的分布**接近脑功能特征**的分布



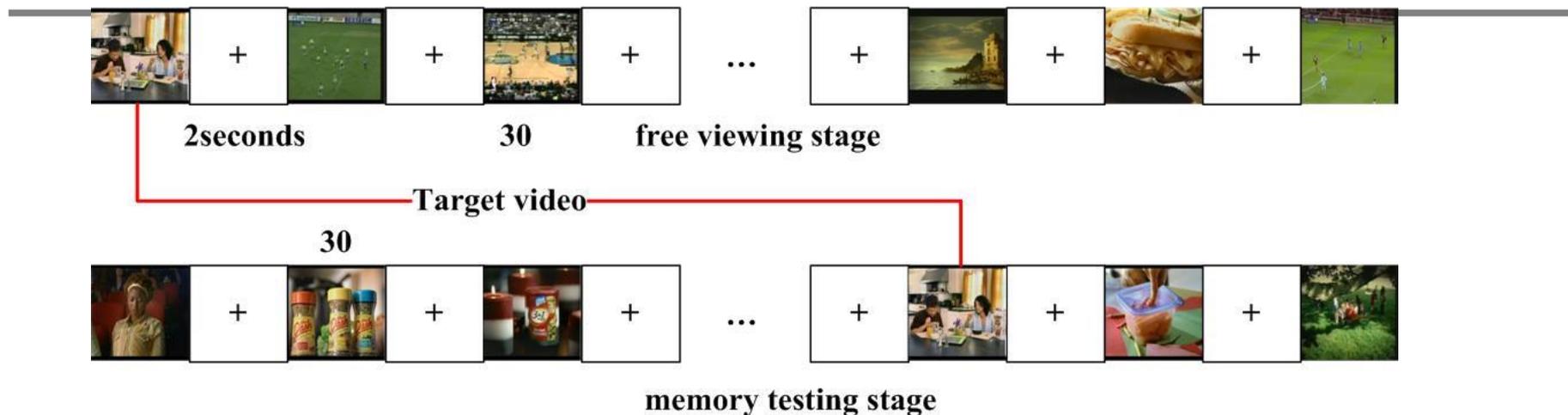
传统方法仅依靠视觉特征

学习到的模型可应用于新数据

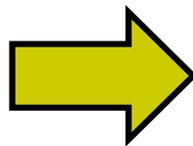
# 脑功能引导的视觉信息分析——视频摘要



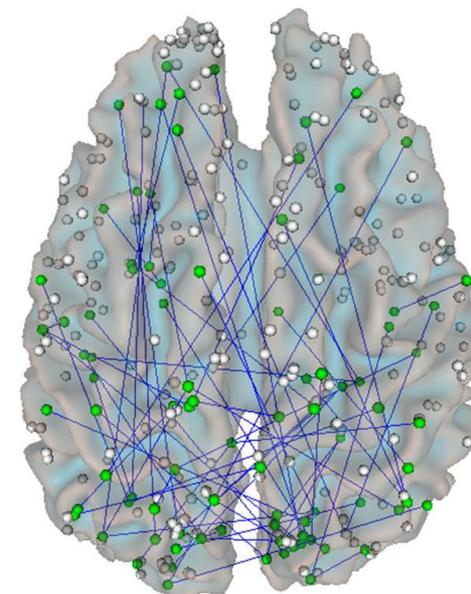
# 脑功能引导的视觉信息分析——视频记忆度预测



监督学习选择参与视频理解和记忆的脑区



ROIs	Percentage
attention	0.1010
l.speech	0.0575
l.semantics	0.0505
execution	0.0502
emotion	0.0467
inhibition	0.0432
uga.attention	0.0394
l.orthography	0.0324
uga.emotion	0.0321
uga.decision	0.0319



---

# 谢谢!

## 合作者

韩军伟，西北工业大学自动化学院教授，长江学者

郭 雷，西北工业大学自动化学院教授，杰青

刘天明，佐治亚大学计算机学院杰出教授

郭 聪，昆士兰医学研究所特聘研究员

张 京，耶鲁大学统计系助理教授

**THANKS** *for your time*