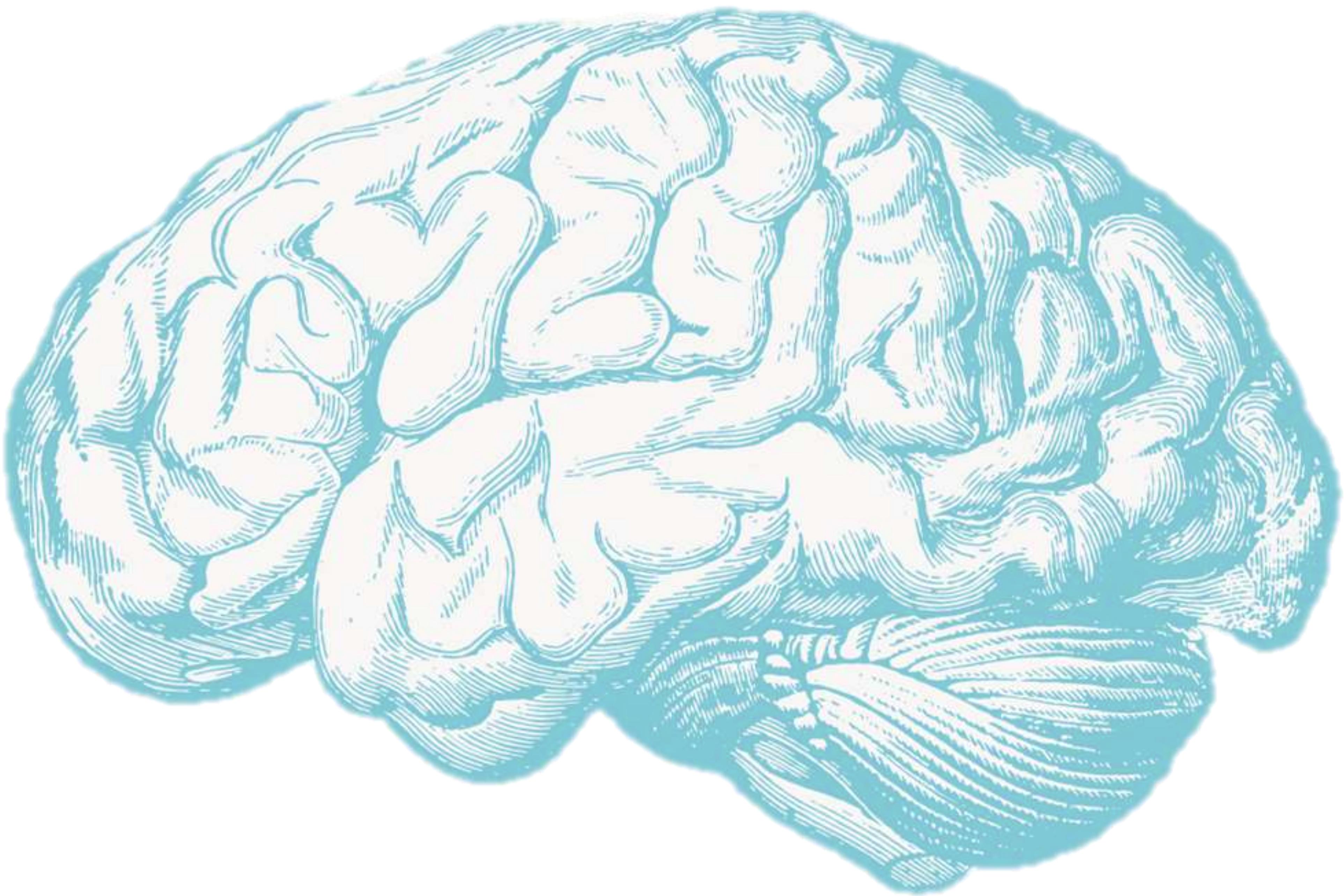
第五讲

脑与认知科学

脑与认知加工：视觉

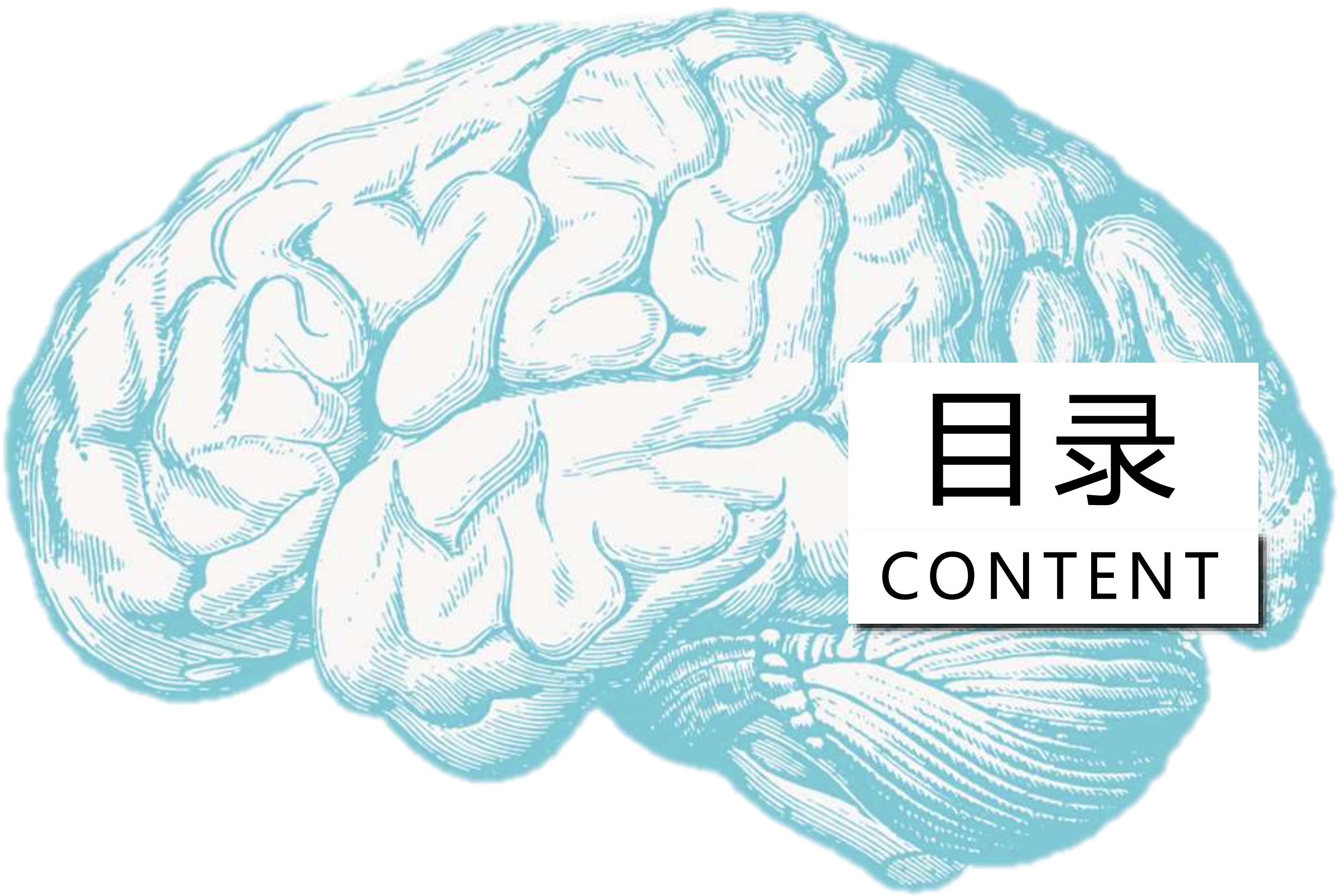
授课人： 甄宗雷 教授
北京师范大学 | 心理学部



教师介绍

甄宗雷，男，北京师范大学心理学部教授，博士生导师，中科院自动化所计算机应用技术博士，长期在认知神经科学方向开展工作采用磁共振/脑磁图，神经编解码模型和人工智能等多种方法研究人脑视觉和运动的认知神经机制及其发展发育，并基于人脑认知神经新发现开发类脑计算模型。研究结果已发表在Nature Neuroscience, Nature Communications, PLoS Biology, eLife, The Journal of Neuroscience等领域高影响力期刊。

Email: zhenzonglei@bnu.edu.cn



目录

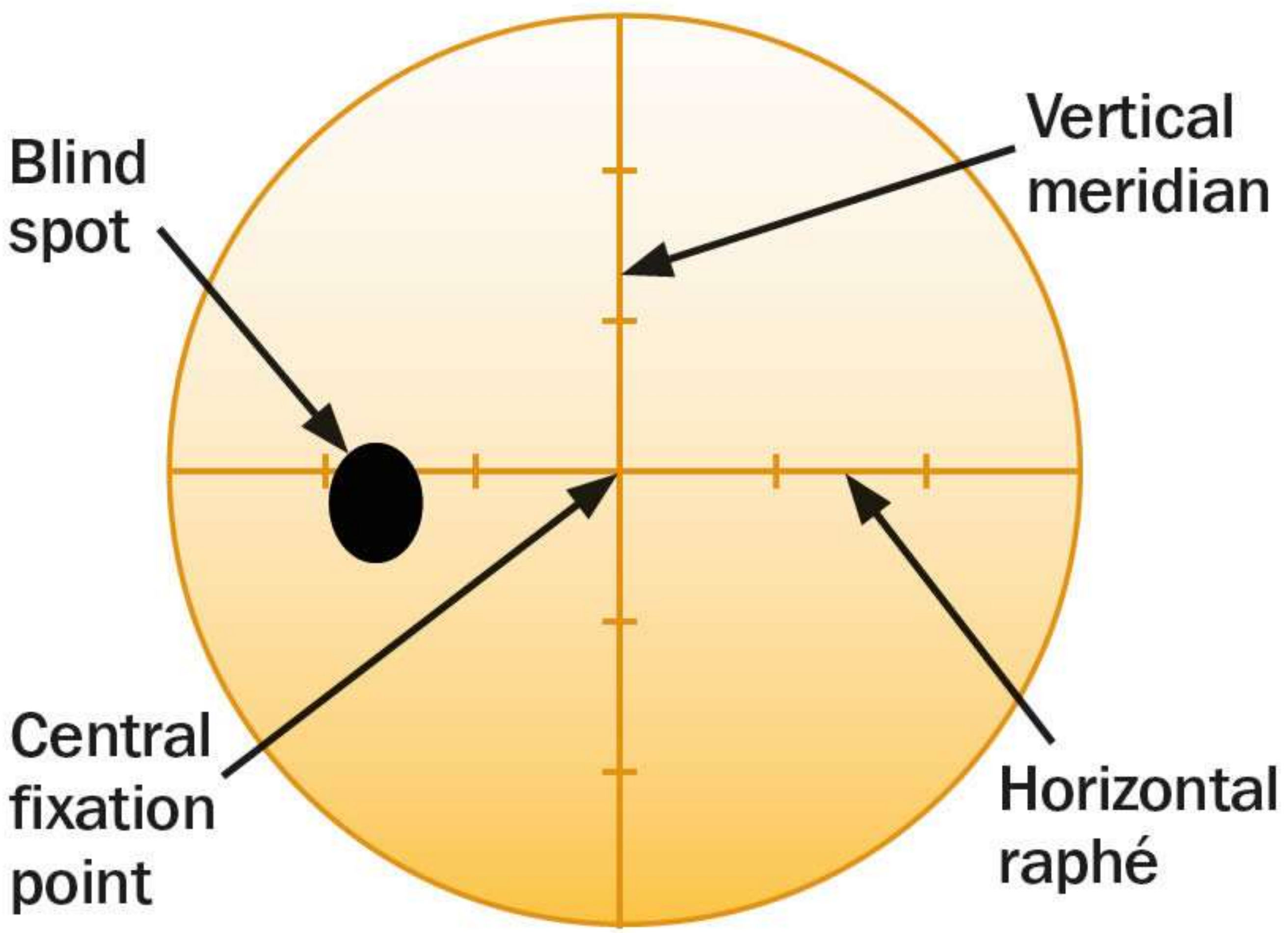
CONTENT

Visual field

Left VF **Right VF**

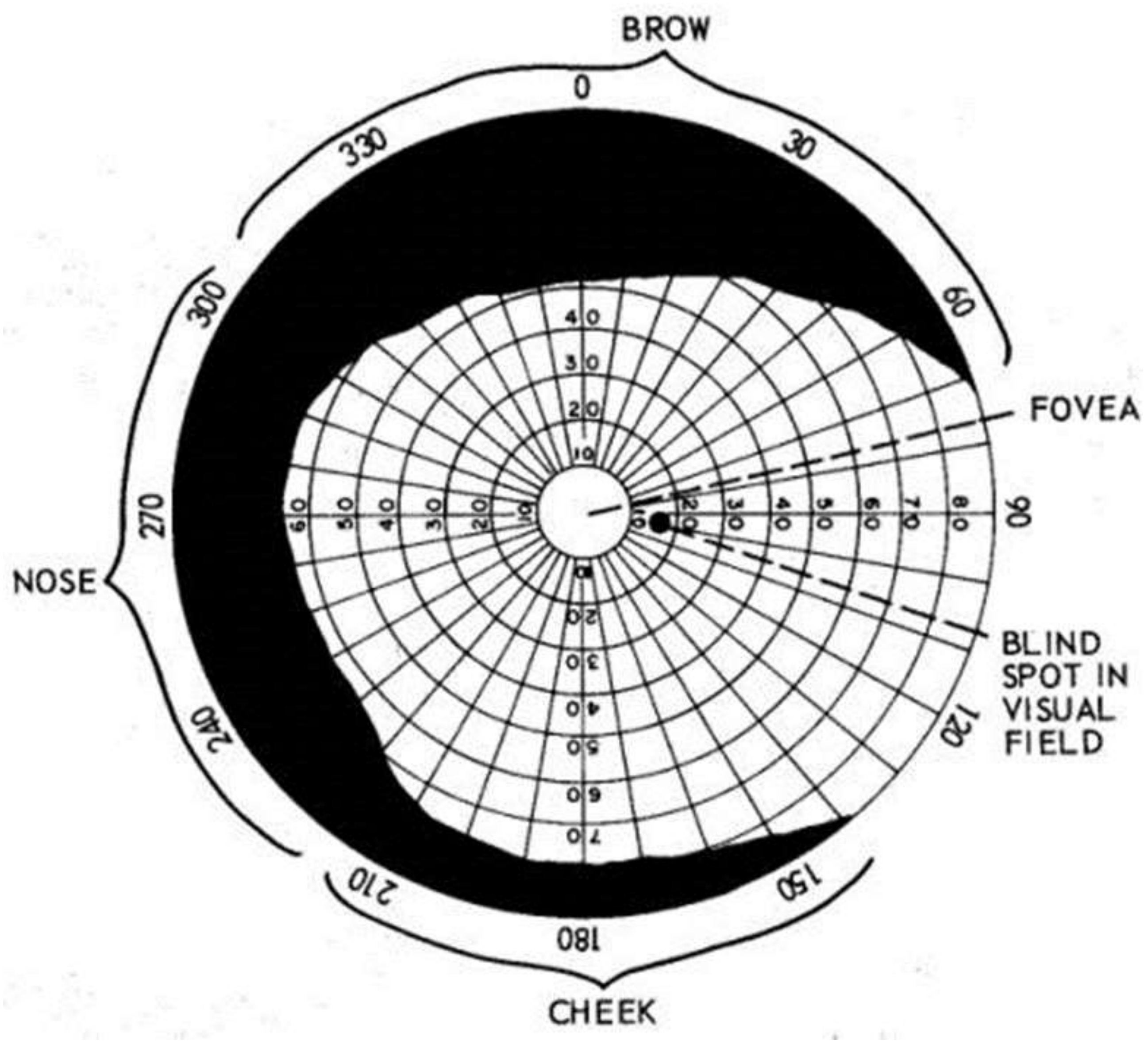
UVF

LVF



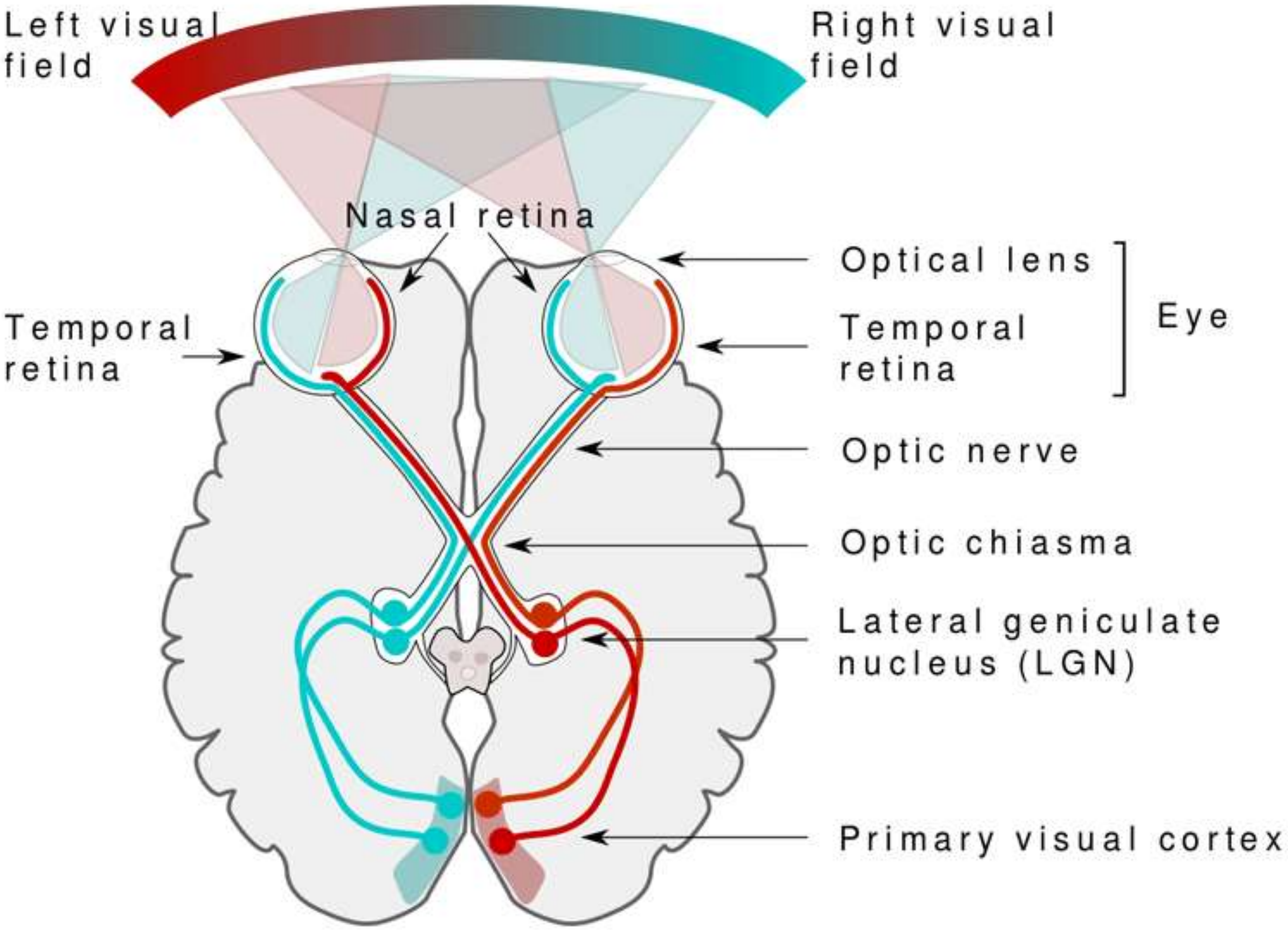
Left eye

Left VF **Right VF**



Right eye

Visual pathway



光学刺激

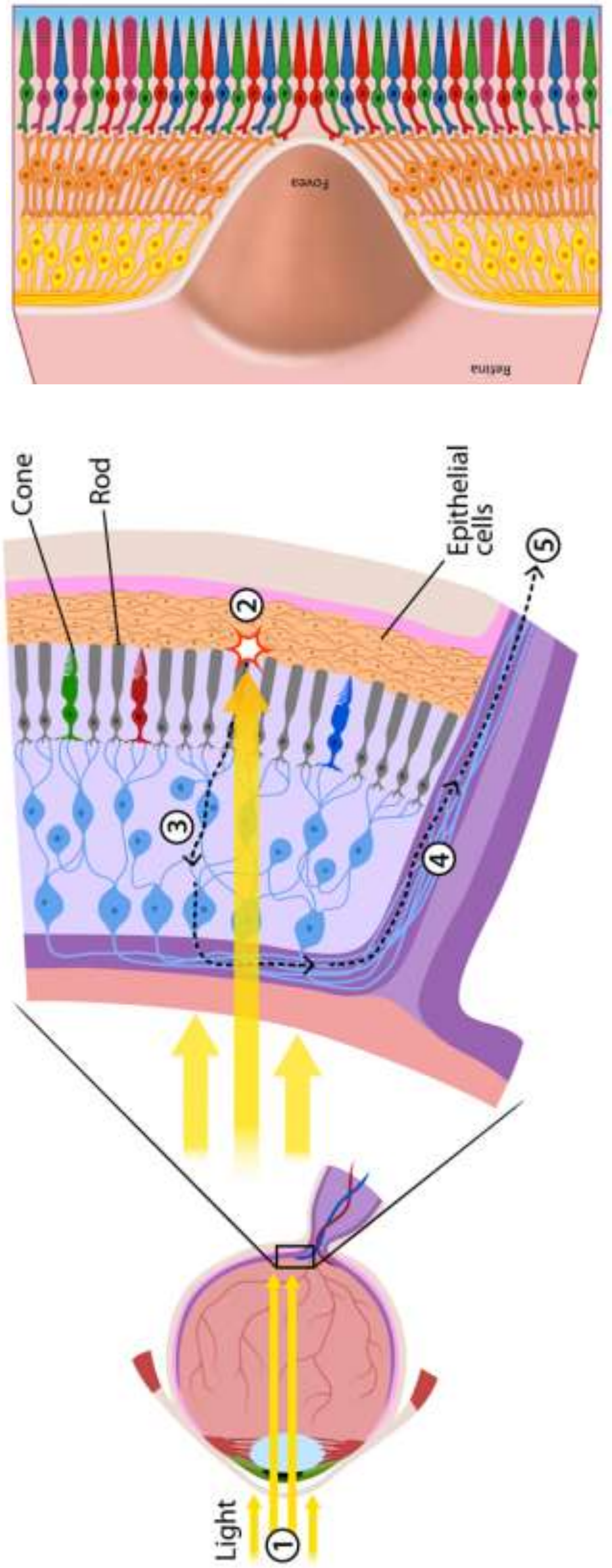
视网膜

外侧膝状体

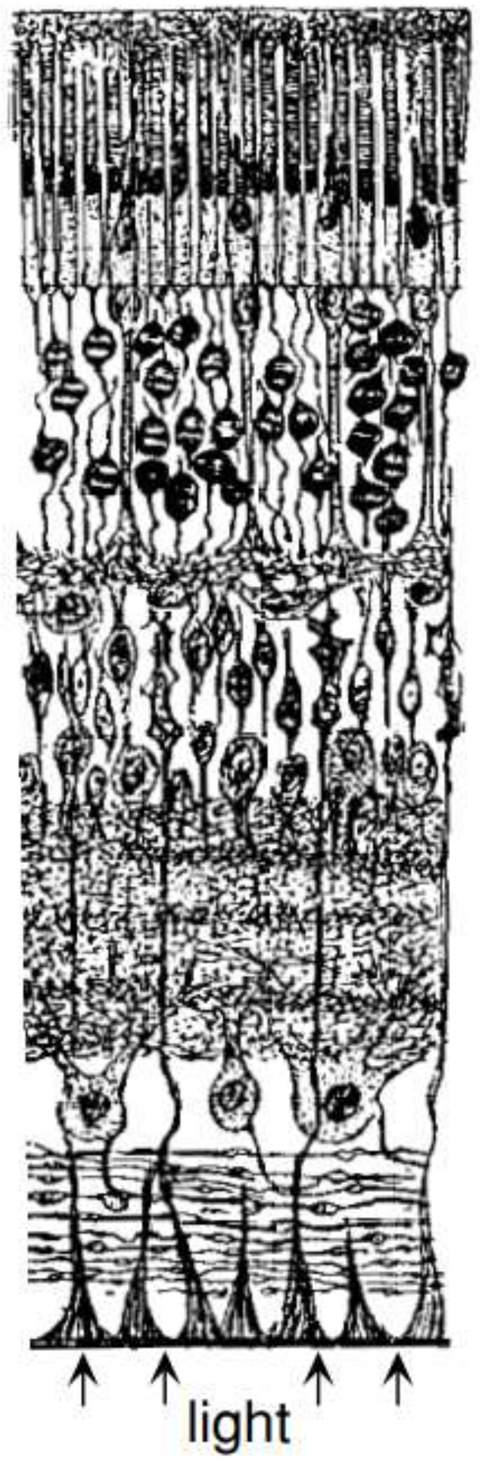
低级视觉皮层

高级视觉皮层

The anatomy and function of the retina



Circuitry



Cell types

rod and cone receptors (R)

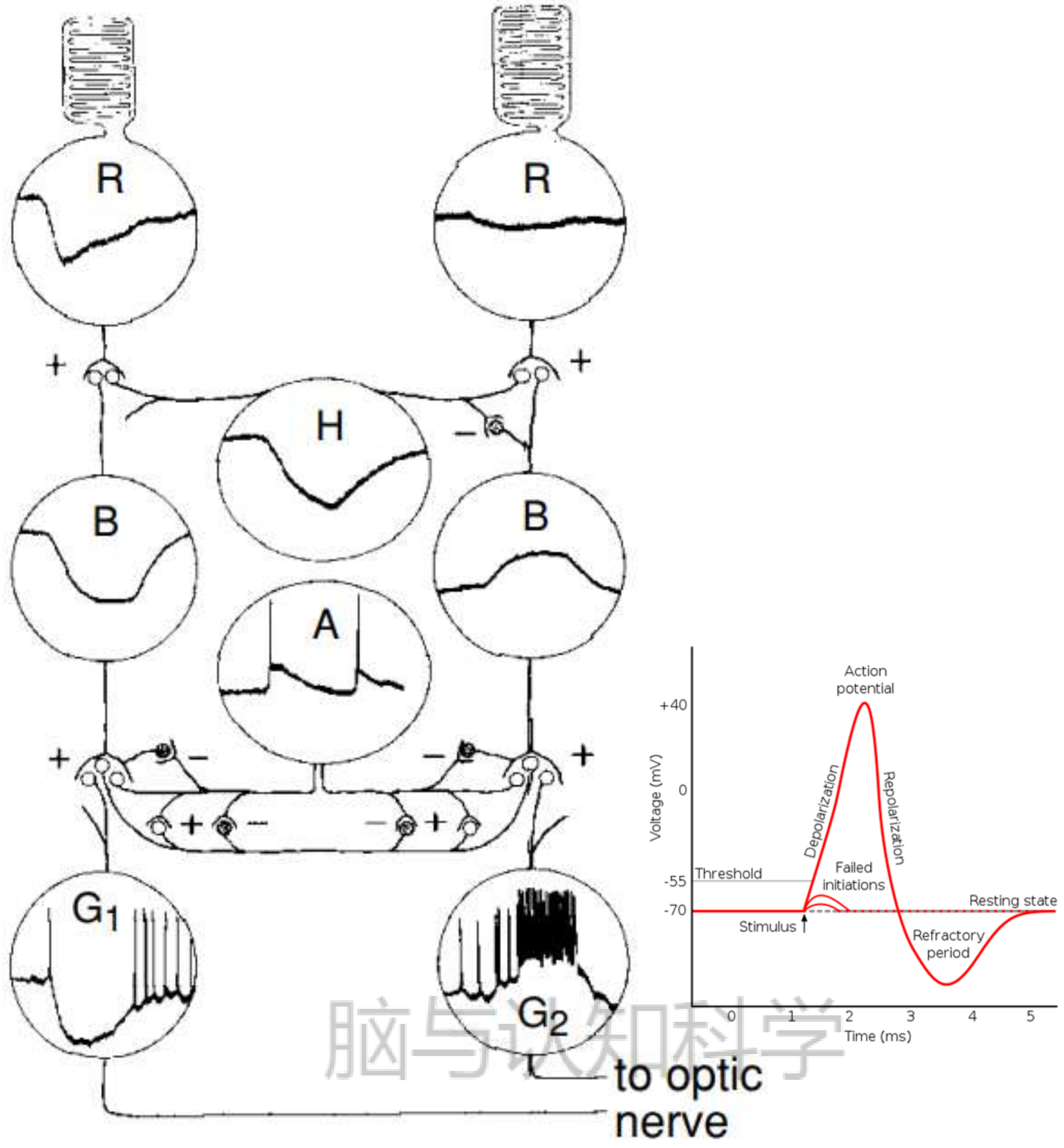
horizontal (H)

bipolar (B)

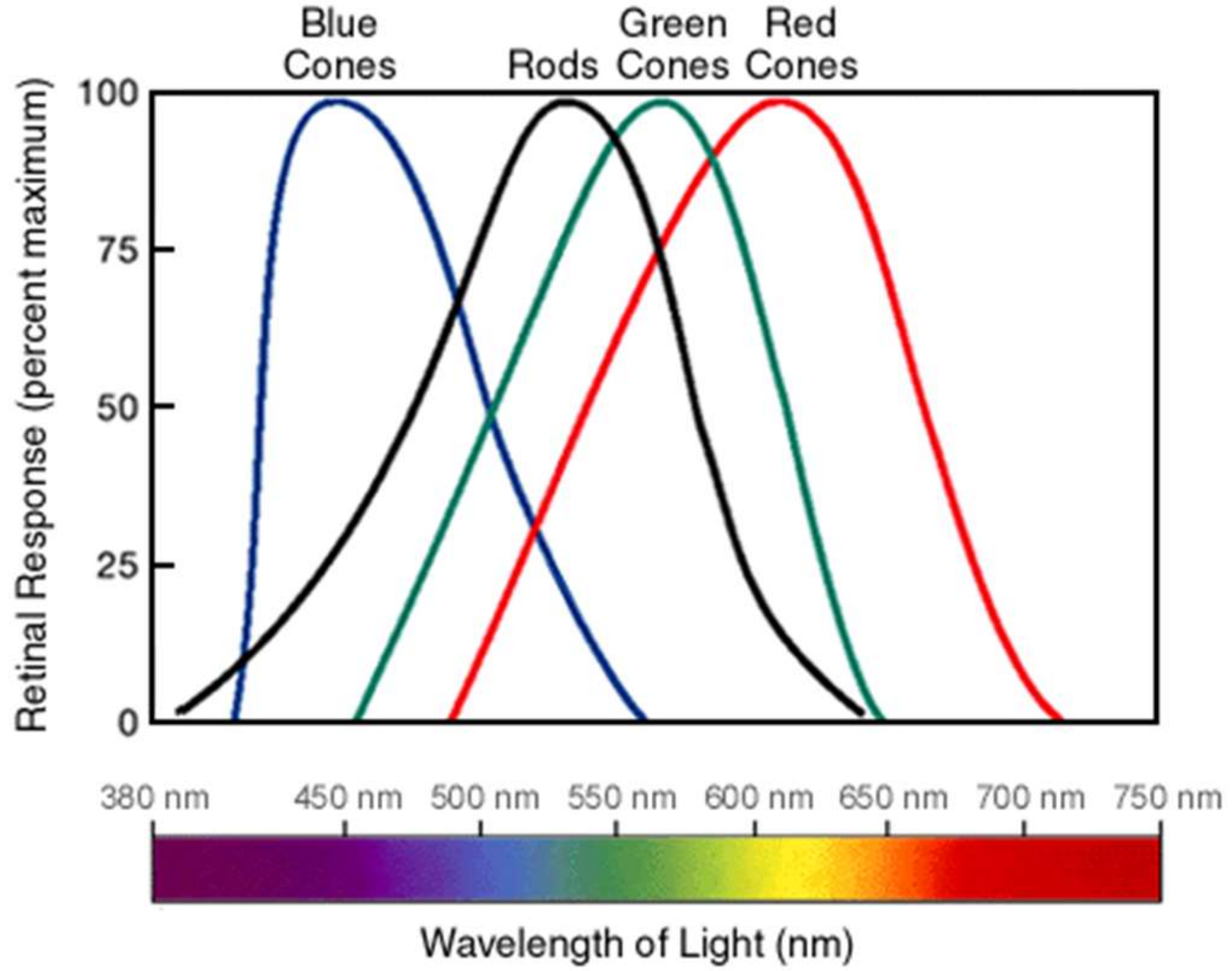
amacrine (A)

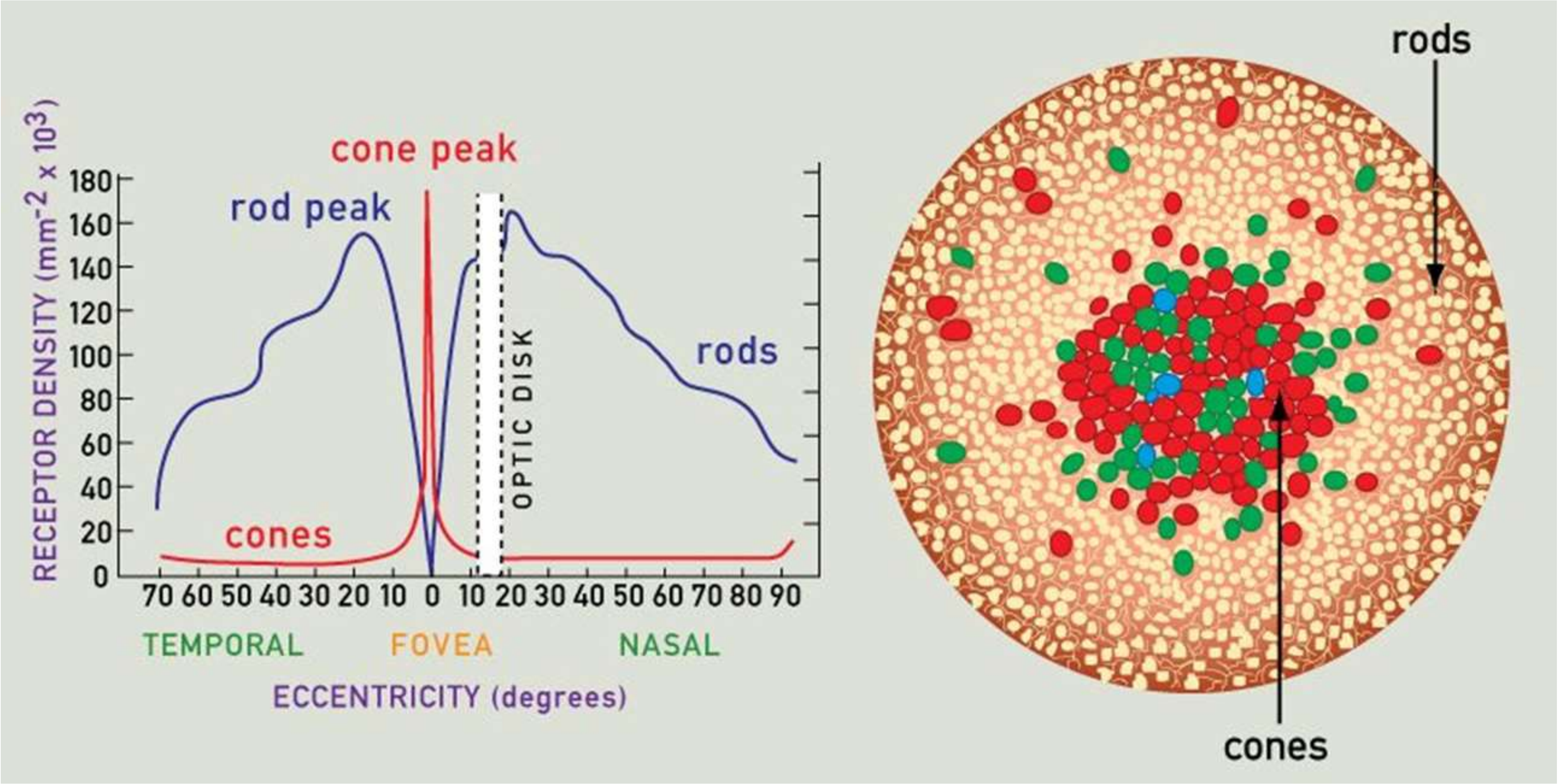
retinal ganglion (G)

Intracellular recordings

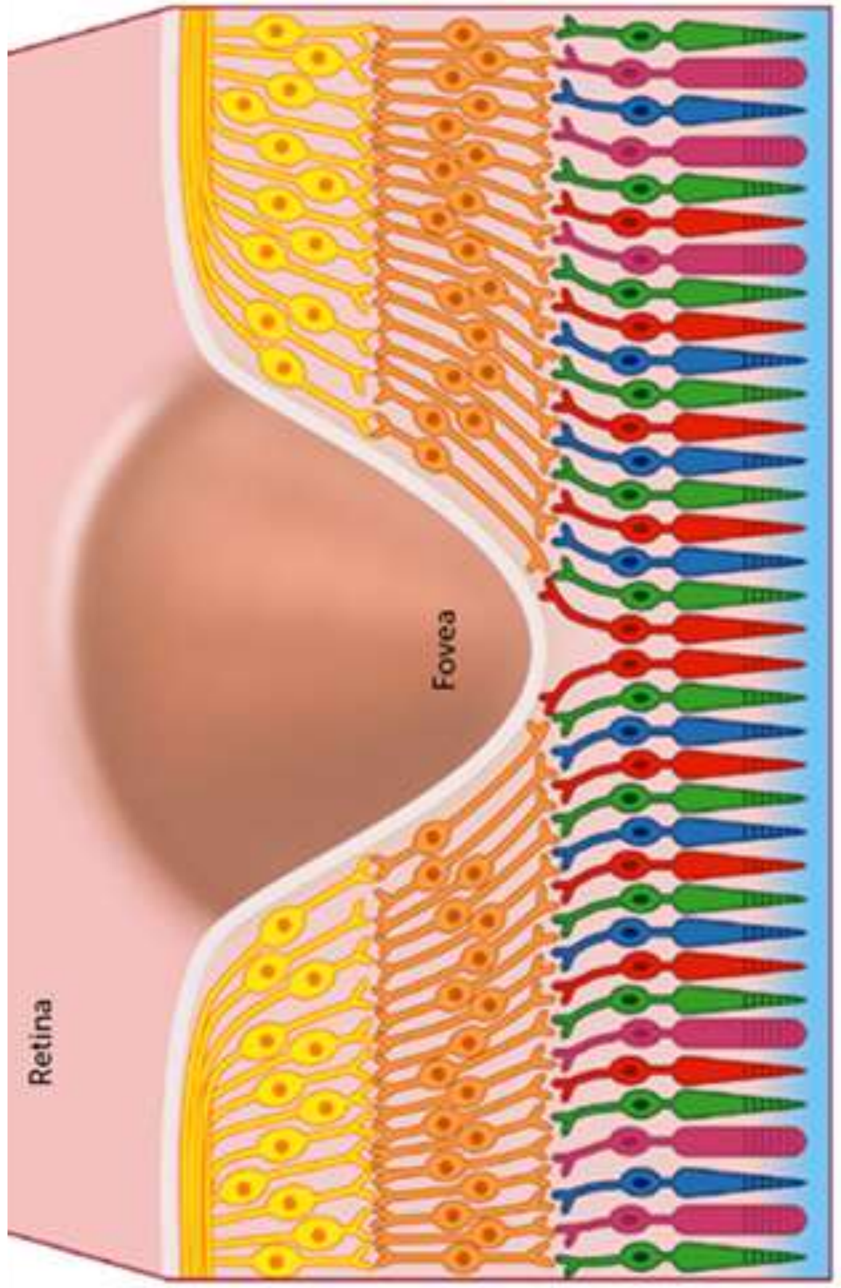


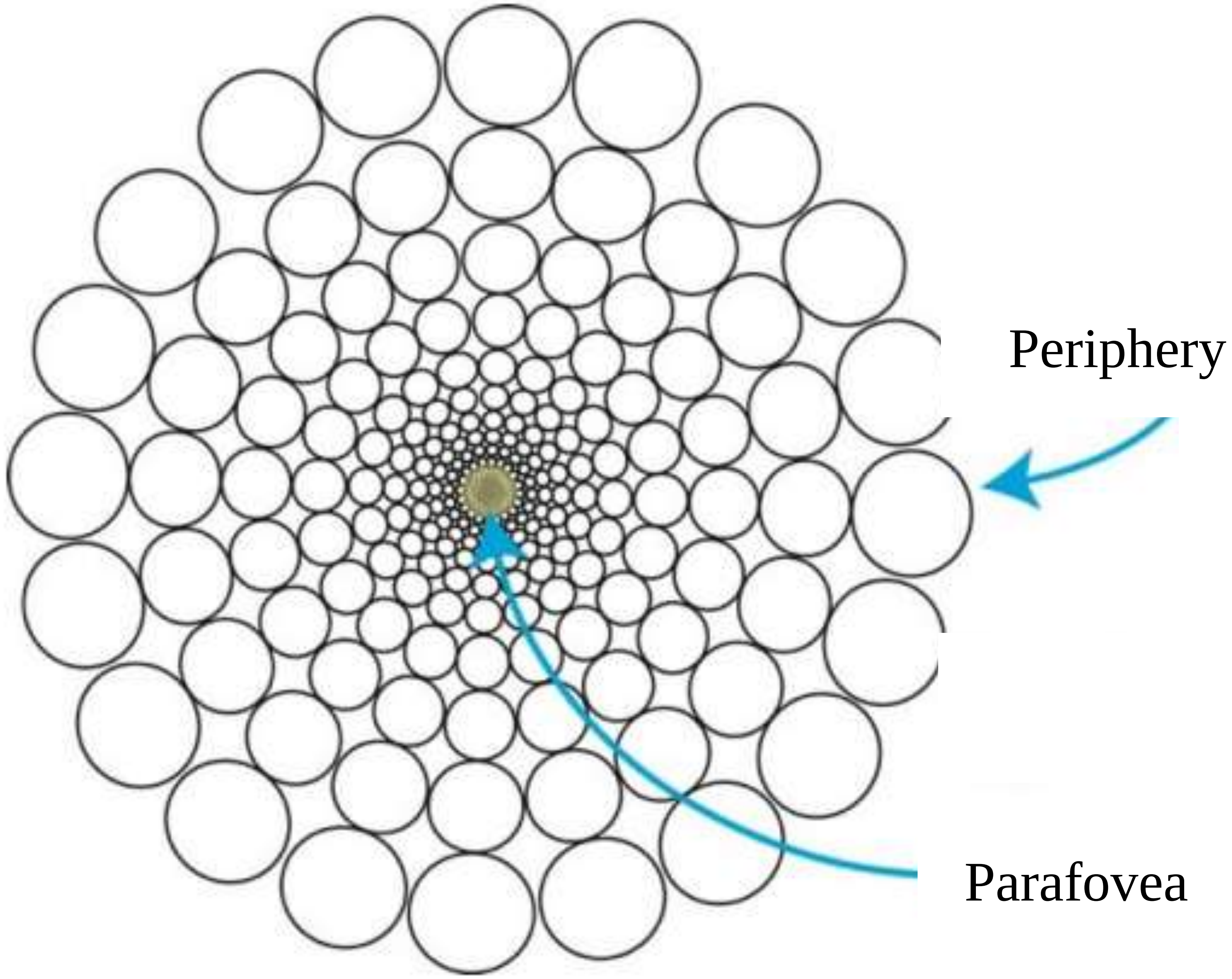
Spatial distribution of rods and cones





Receptive field for a neuron



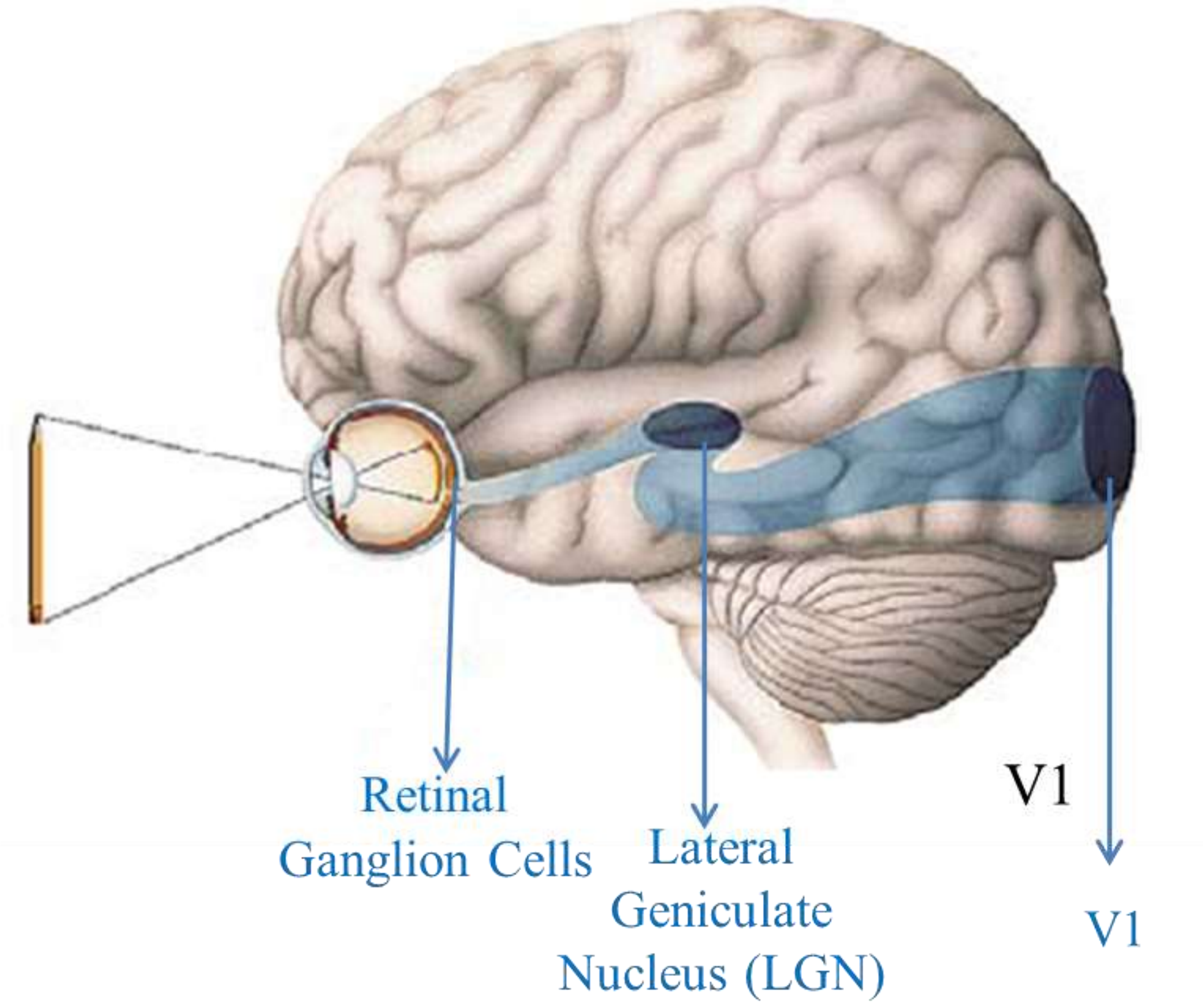


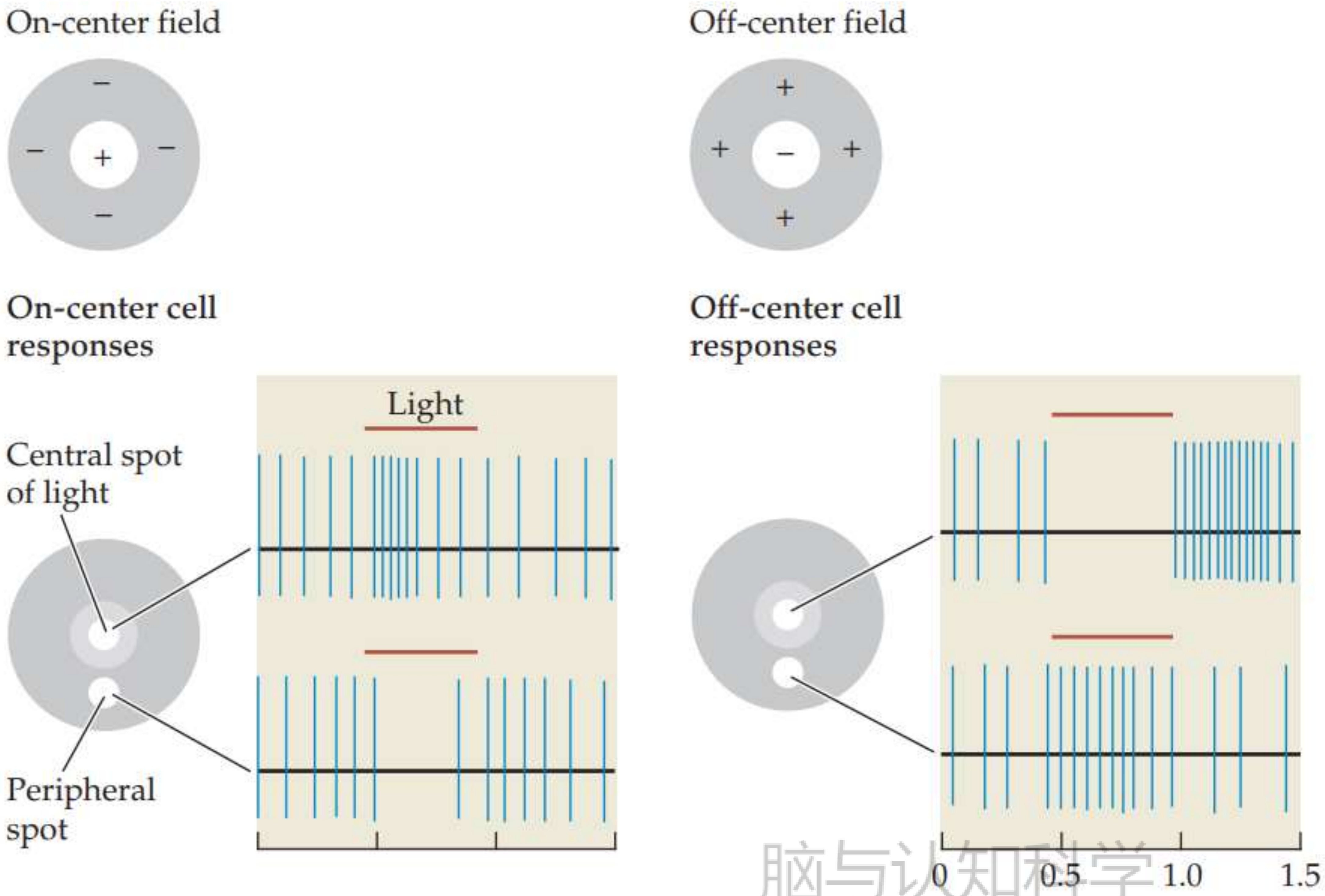
Receptive field(RF): The region of sensory space that activates a neuron.

Neurons in the retina, LGN, and primary visual cortex respond to light stimuli in restricted regions of the visual field.

Center-surround RF of ganglion cells

Receptive field(RF): specific properties of a sensory stimulus that generate a strong response from a cell.

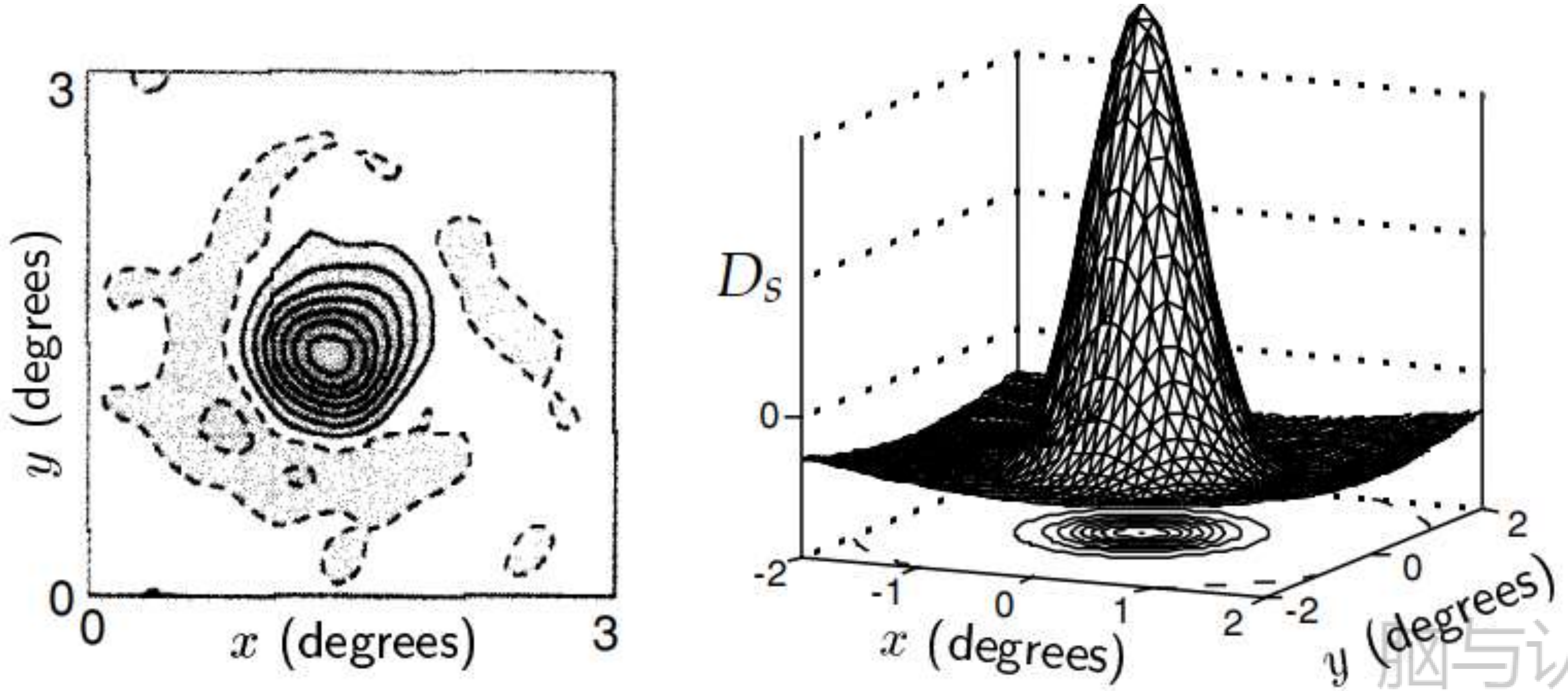




Computational model of ganglion/LGN receptive fields

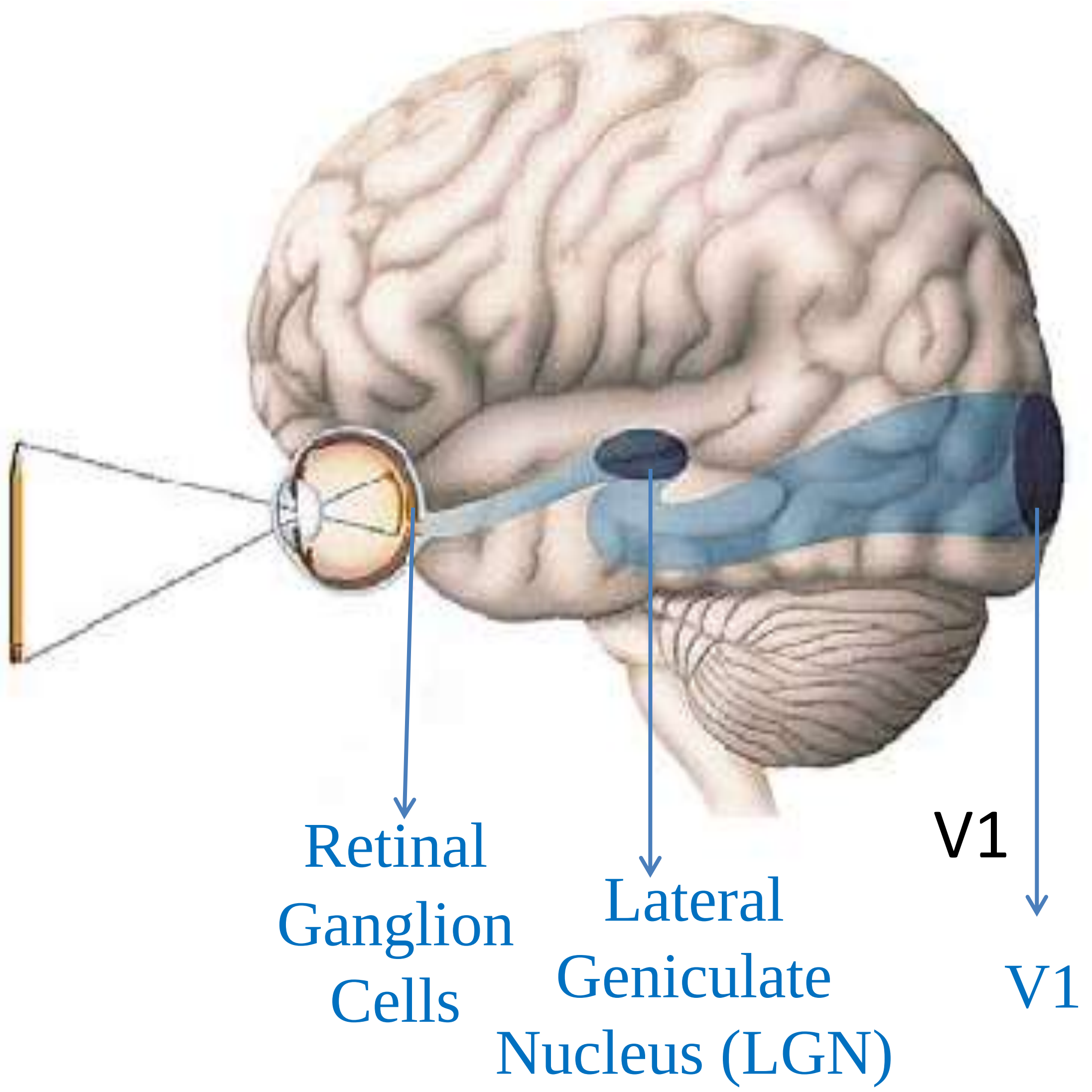
Difference-of-Gaussians

$$D_s(x,y) = \pm\left(\frac{1}{2\pi\sigma_{\mathrm{cen}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\mathrm{cen}}^2}\right) - \frac{B}{2\pi\sigma_{\mathrm{sur}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\mathrm{sur}}^2}\right)\right)$$

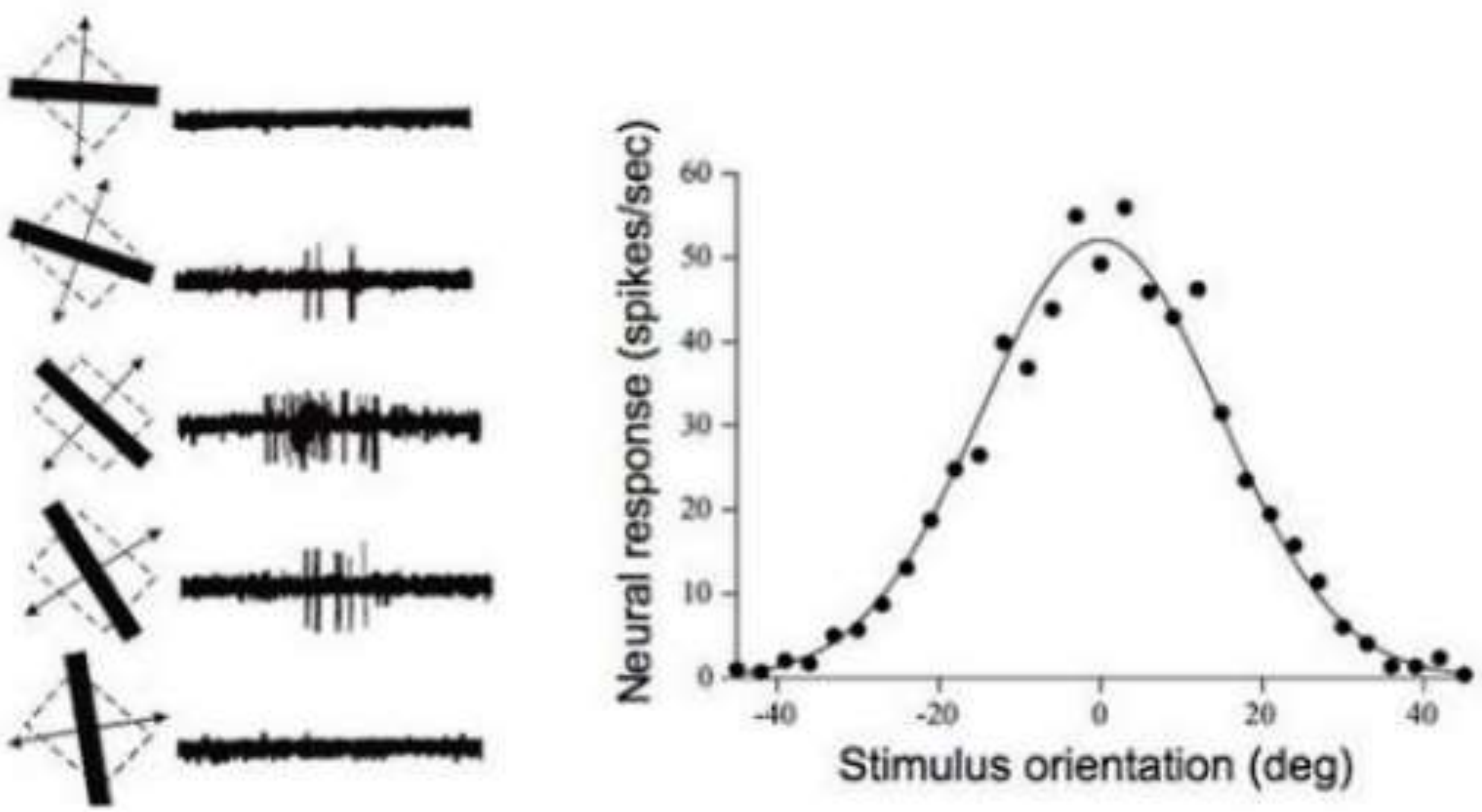


Orientation RF of V1 cells

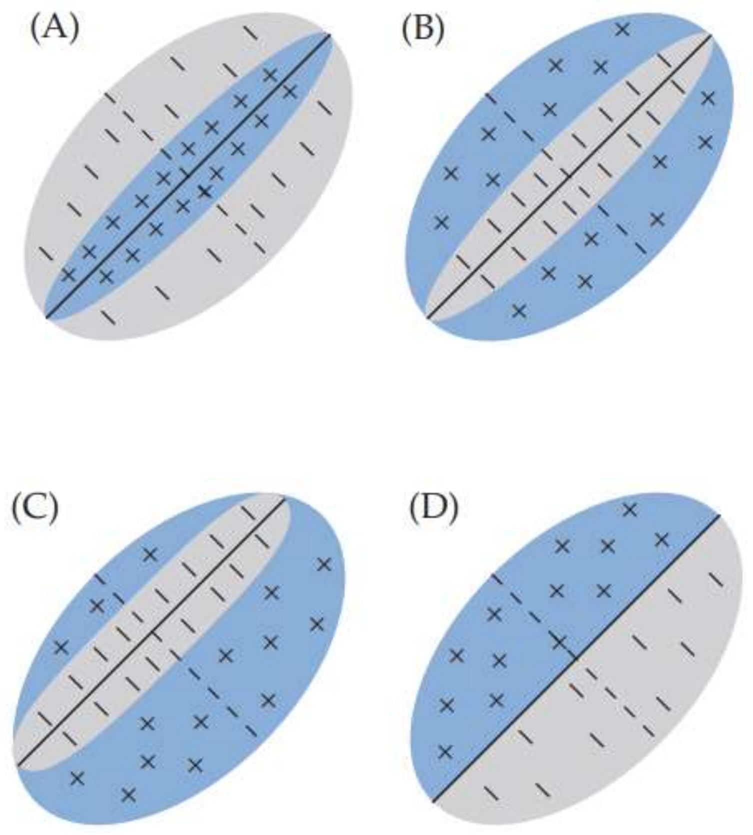
Visual pathway



Tuning curve of V1 cell



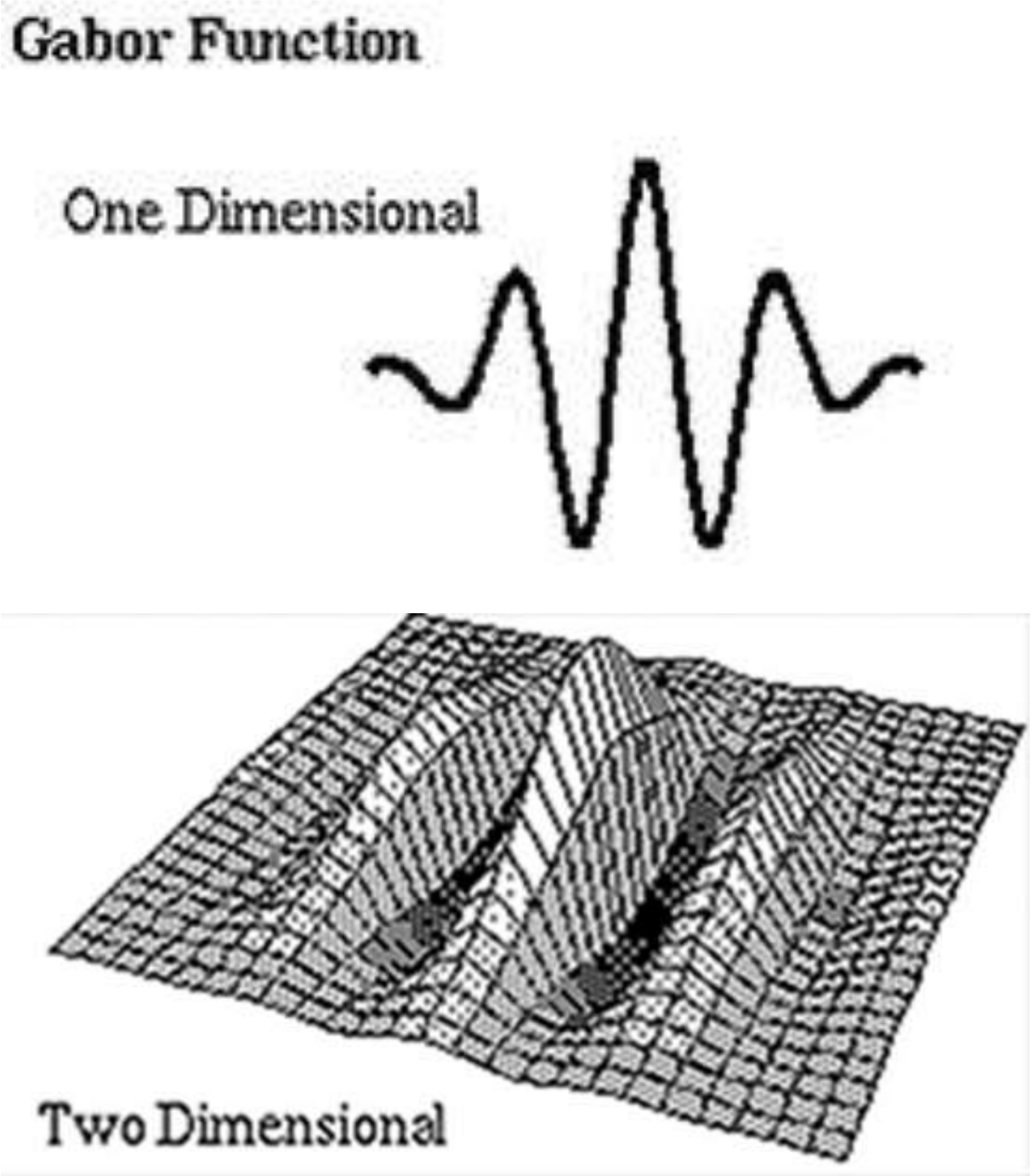
RF of V1 cell

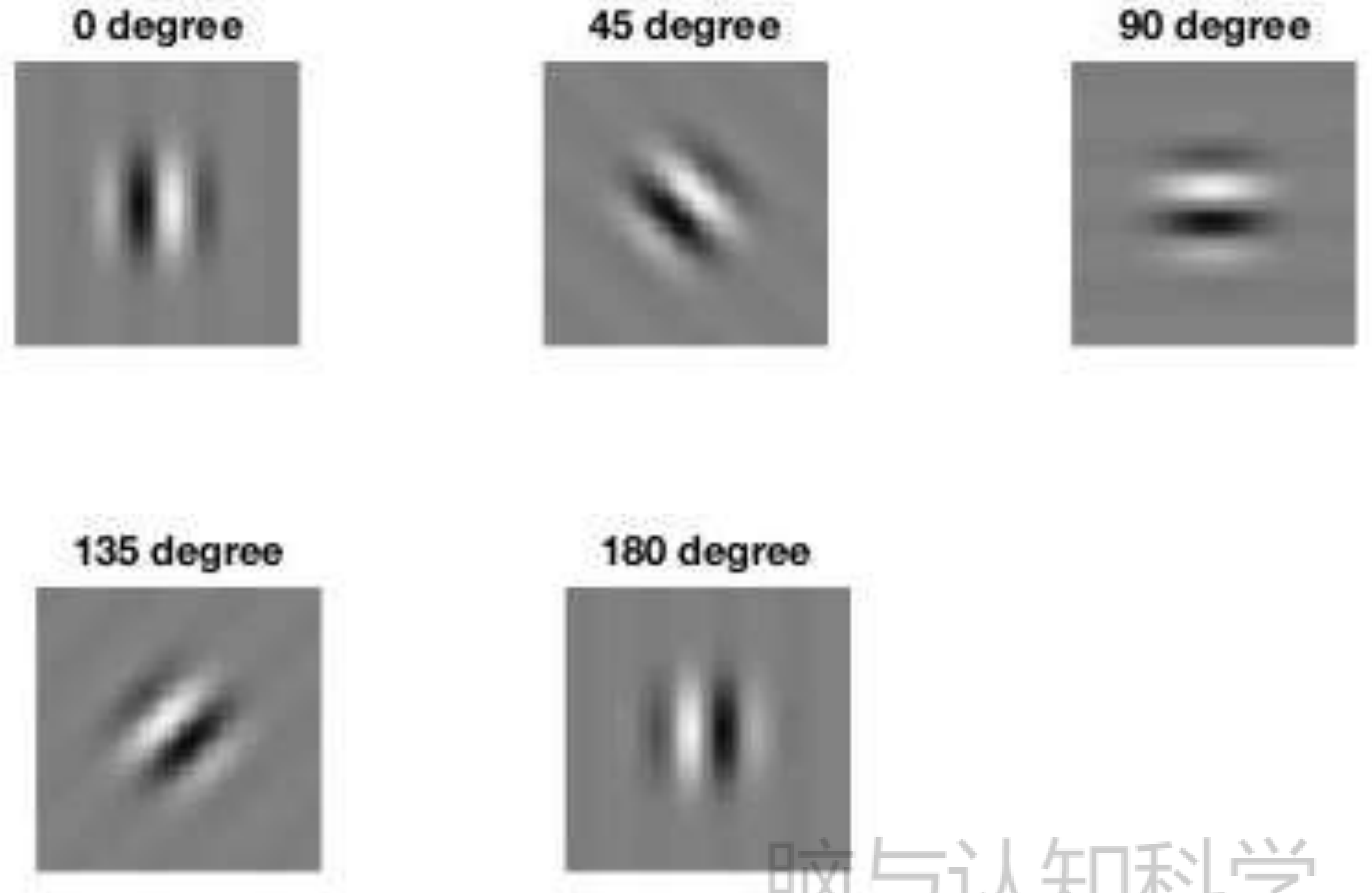


Computational model of V1 oriented RFs

$$G(x,y) = exp(-\frac{|\mathbf{r}|^2}{2\sigma_s^2})\ cos[(\mathbf{r}\cdot\omega_{\mathbf{s}})+\theta] \qquad exp(-\frac{x^2+y^2}{2\sigma_s^2}) = \ exp(-\frac{|\mathbf{r}|^2}{2\sigma_s^2})$$

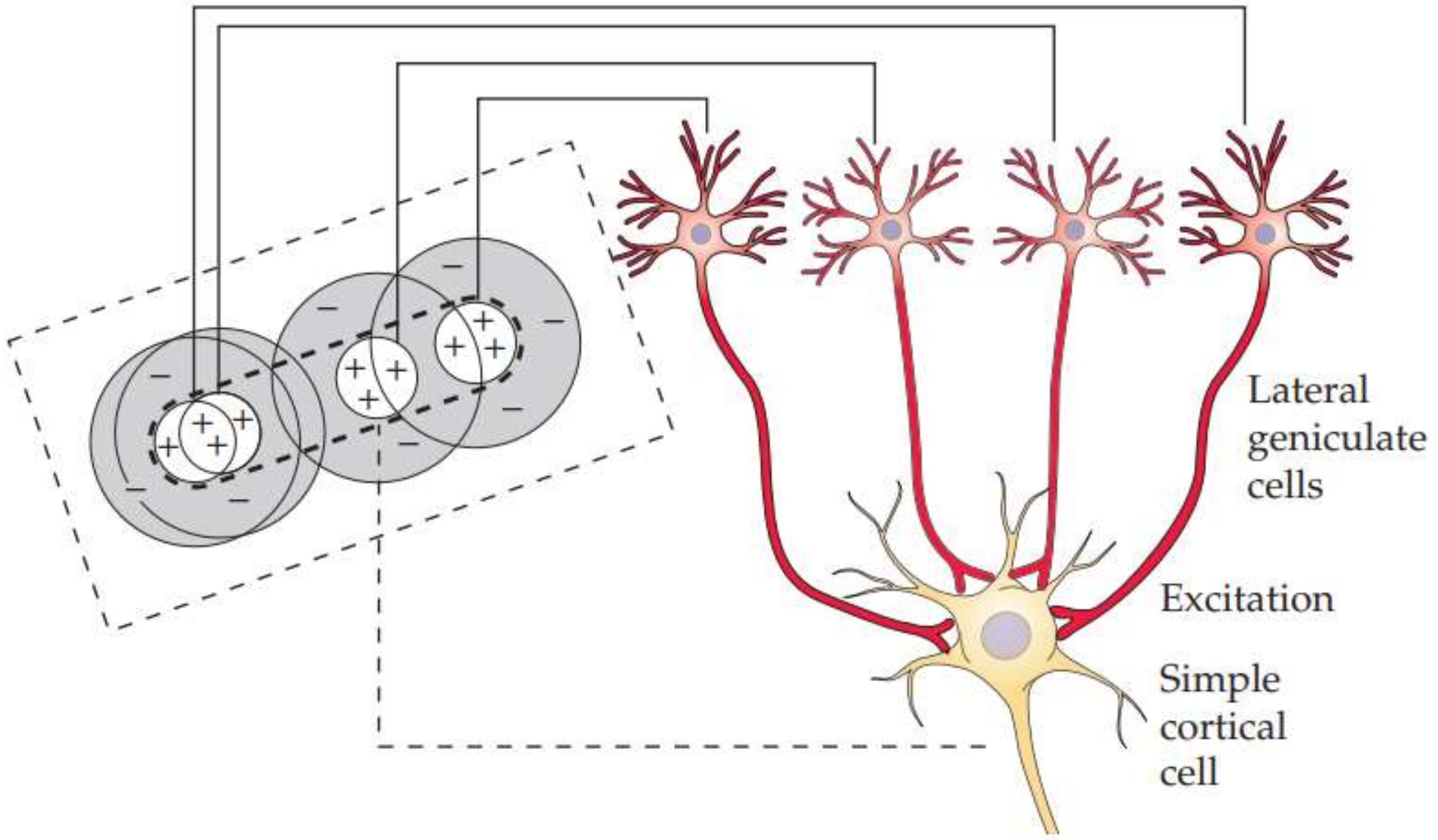
$$\begin{cases} \omega_s = \sqrt{\omega_x^2+\omega_y^2} \\ \mathbf{n} = [cos\ \phi,\ sin\ \phi] \end{cases}$$

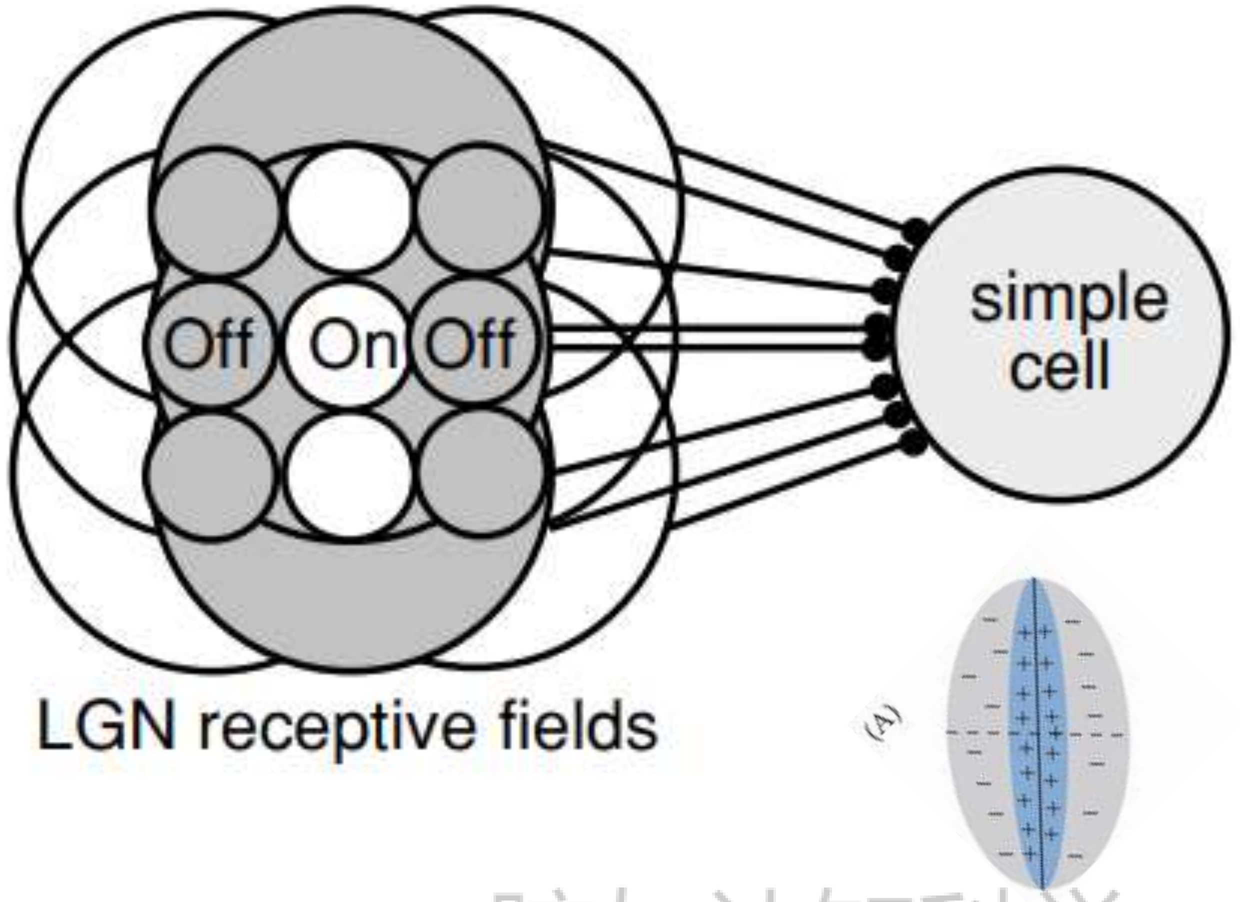




From LGN to V1 receptive fields

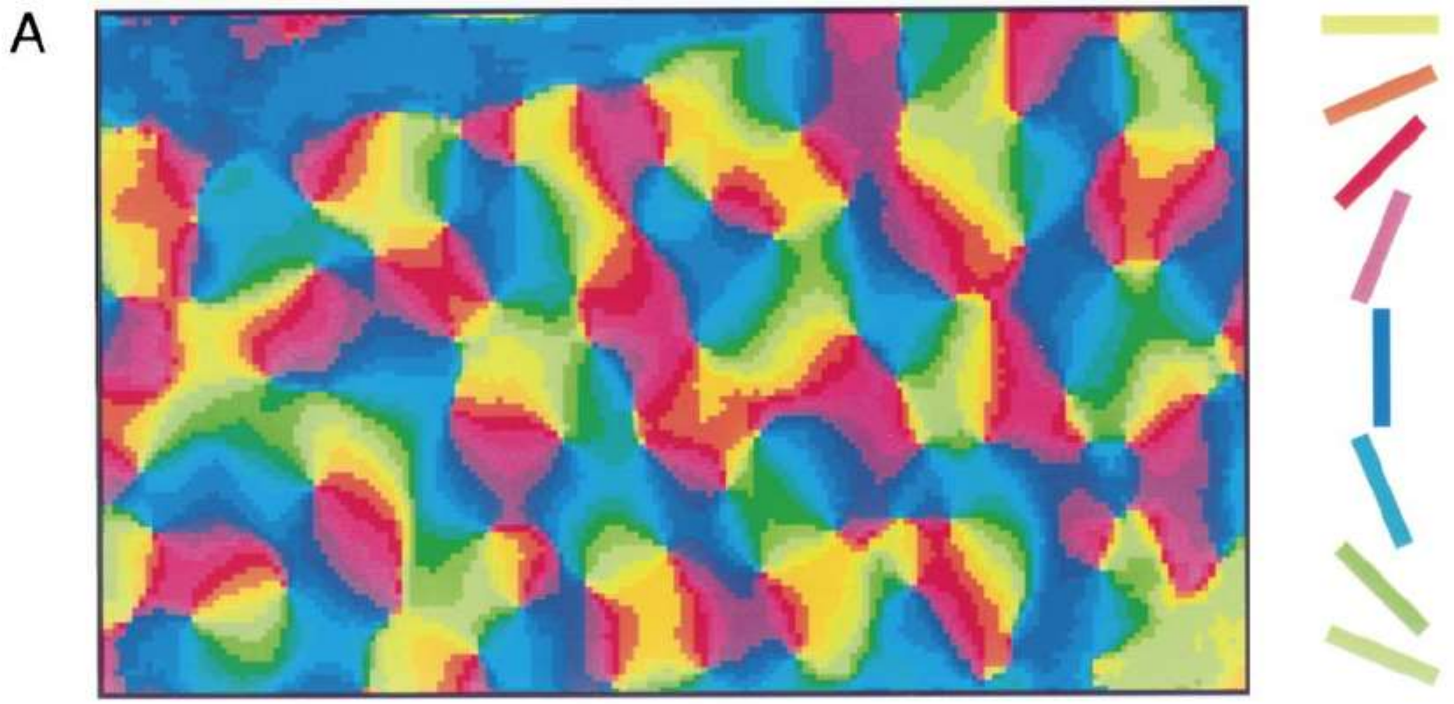
The oriented receptive fields of V1 simple neurons could be generated by summing the input from appropriately selected LGN neurons.



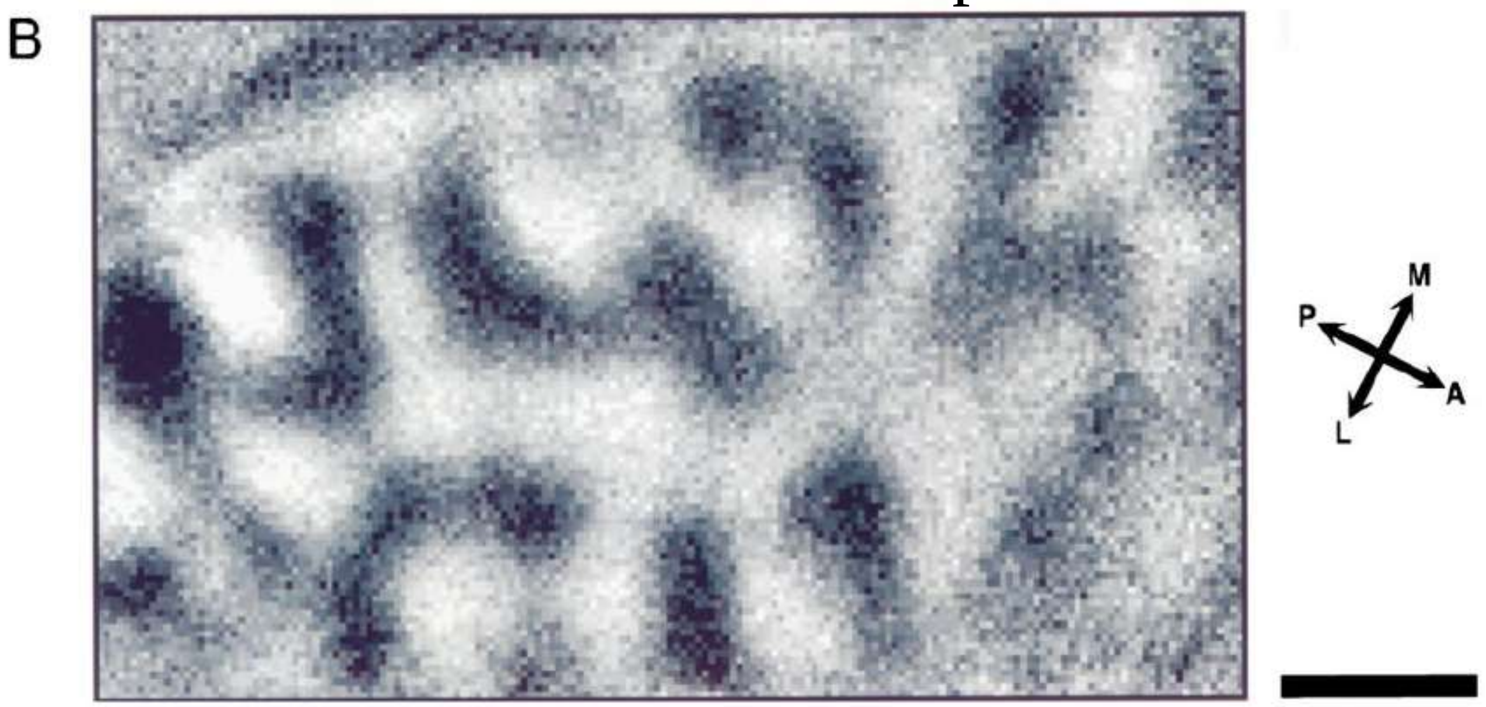


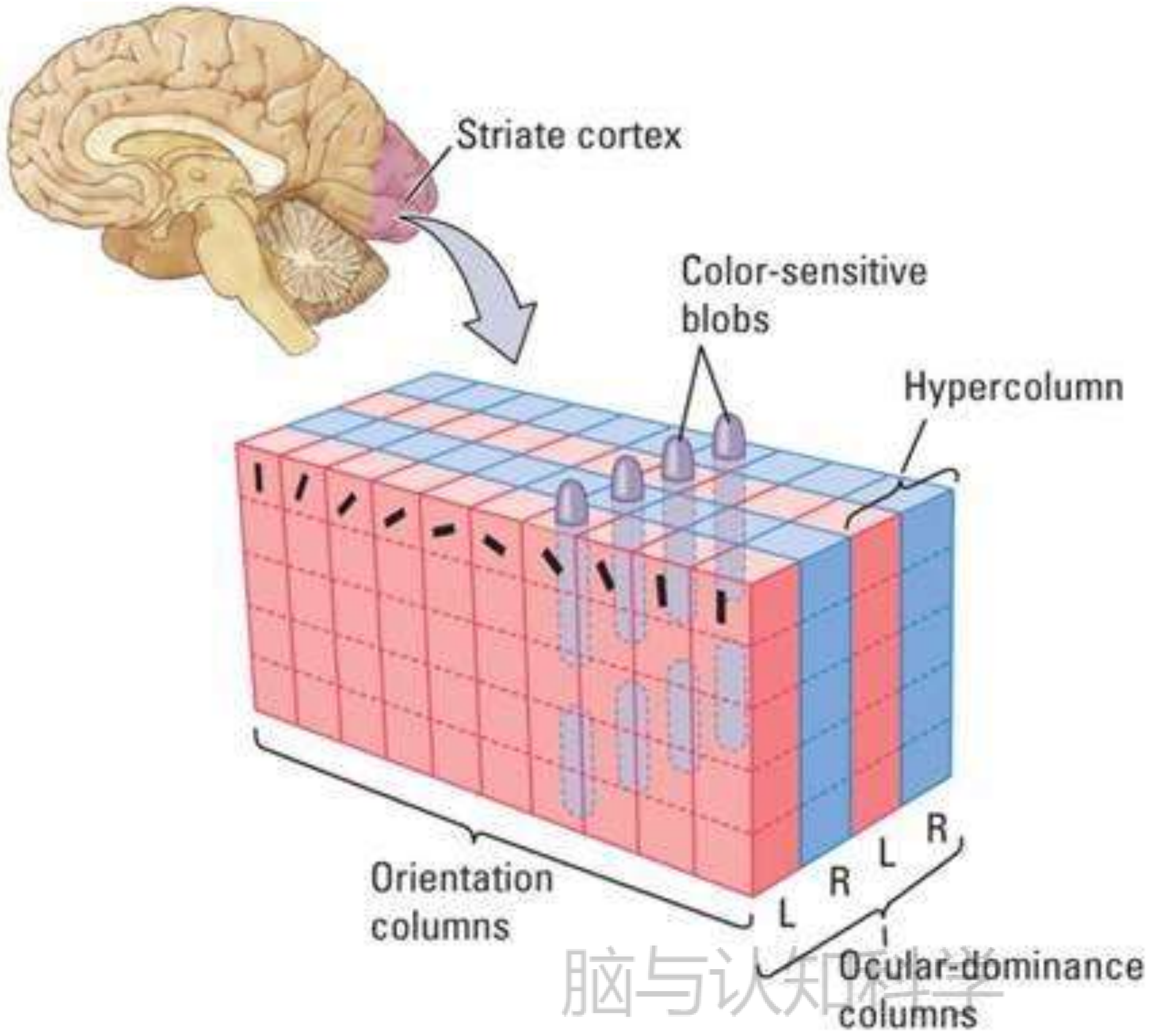
Functional column and hypercolumns in V1

Orientation preference map

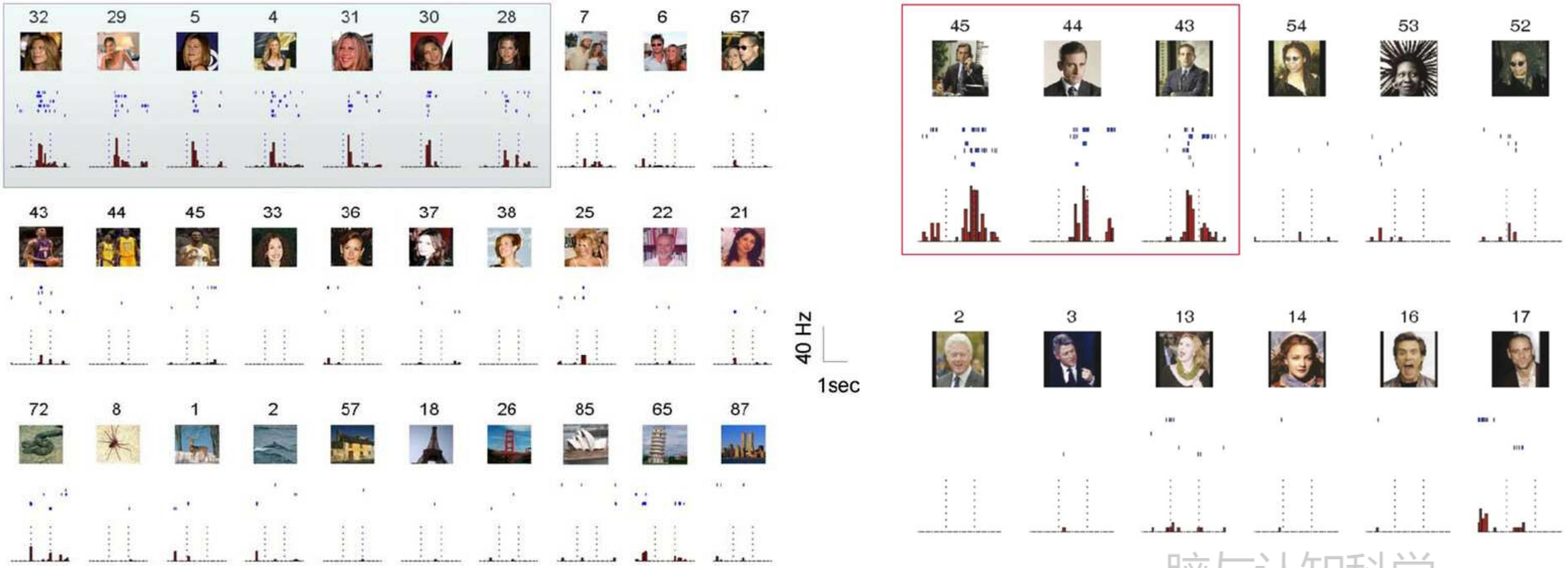


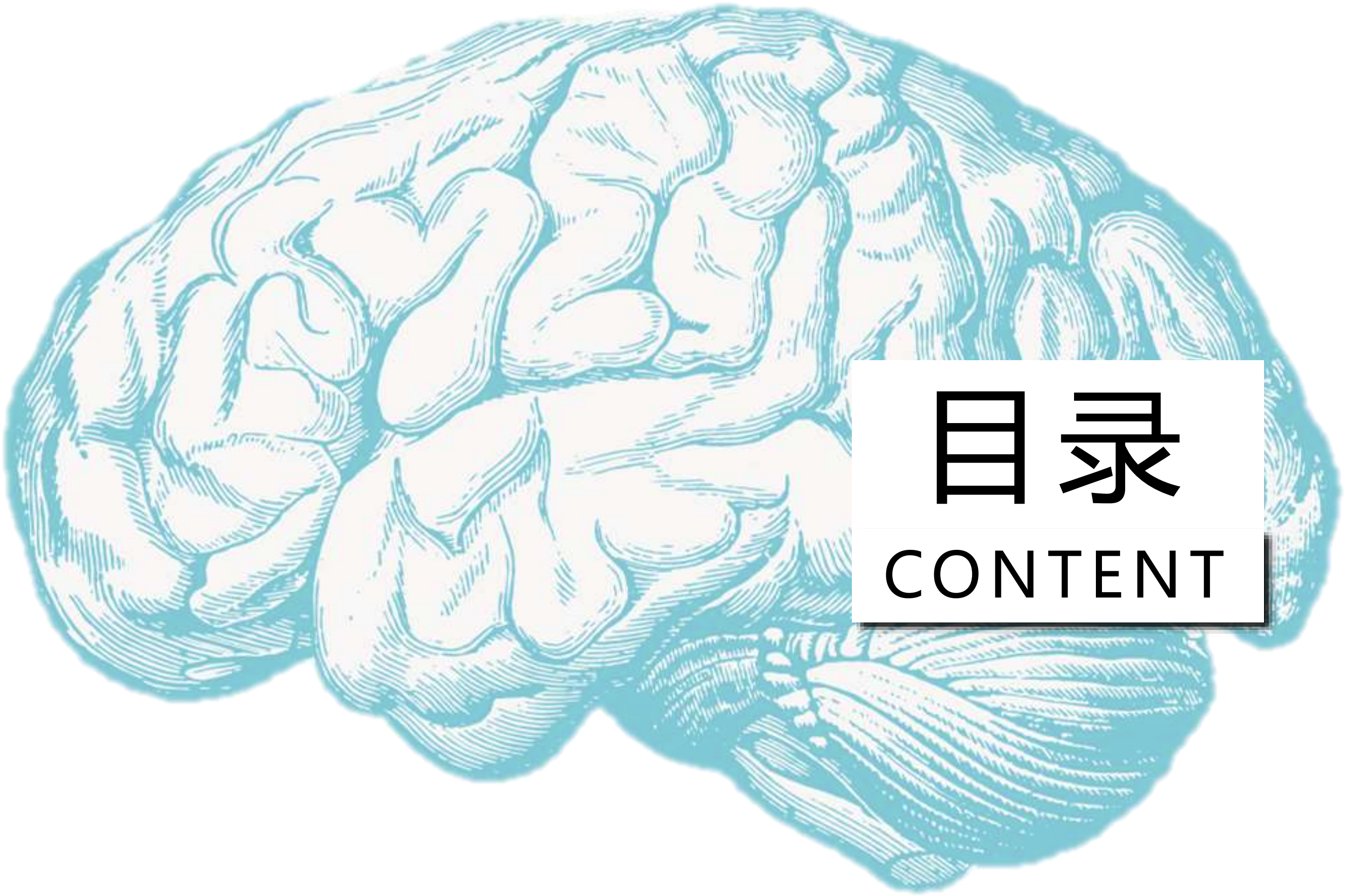
Ocular dominance map





Face neuron in the right anterior hippocampus





目录

CONTENT

A 视神经元的功能

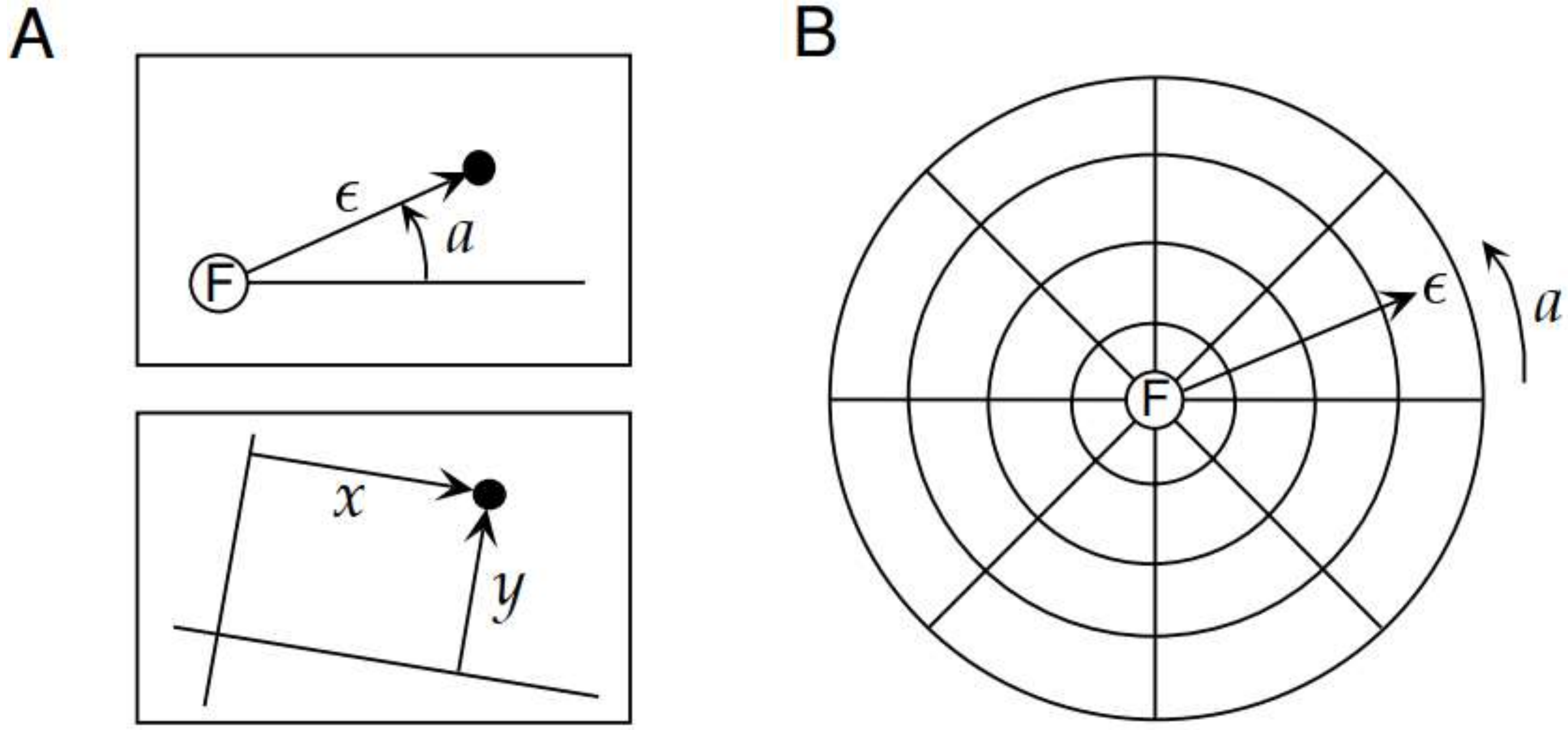
B 视觉脑功能分区

C 视觉计算模型

Coordinate systems to the visual field

- Locations on visual sphere can be represented using the longitude and latitude angles. The “north pole” of the sphere is located at the fixation point, the image point that focuses onto the fovea or center of the retina.

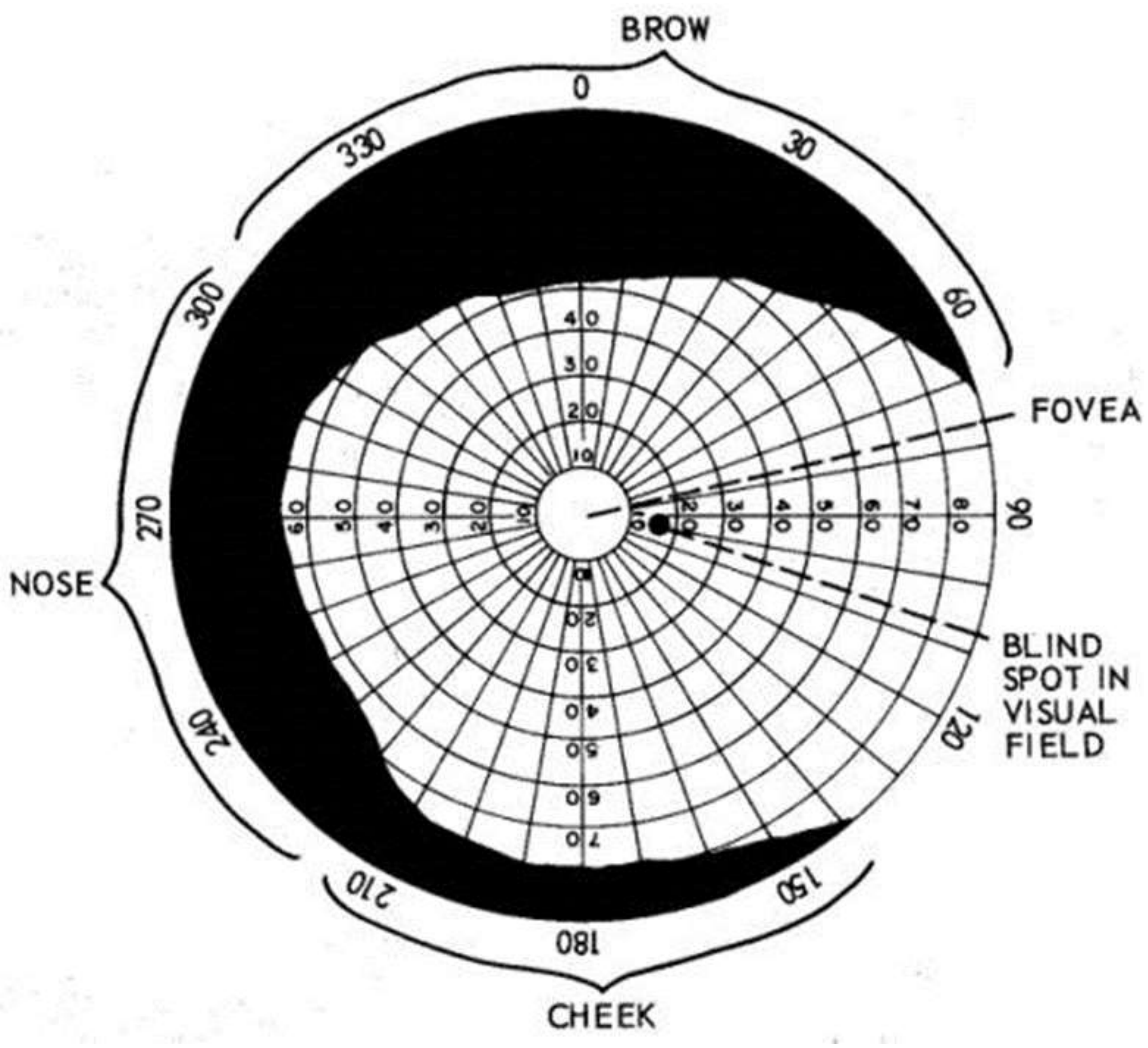
- If the screen is not too large in experiment, the eccentricity and azimuth angles can be approximately coincide with polar coordinates on the screen.



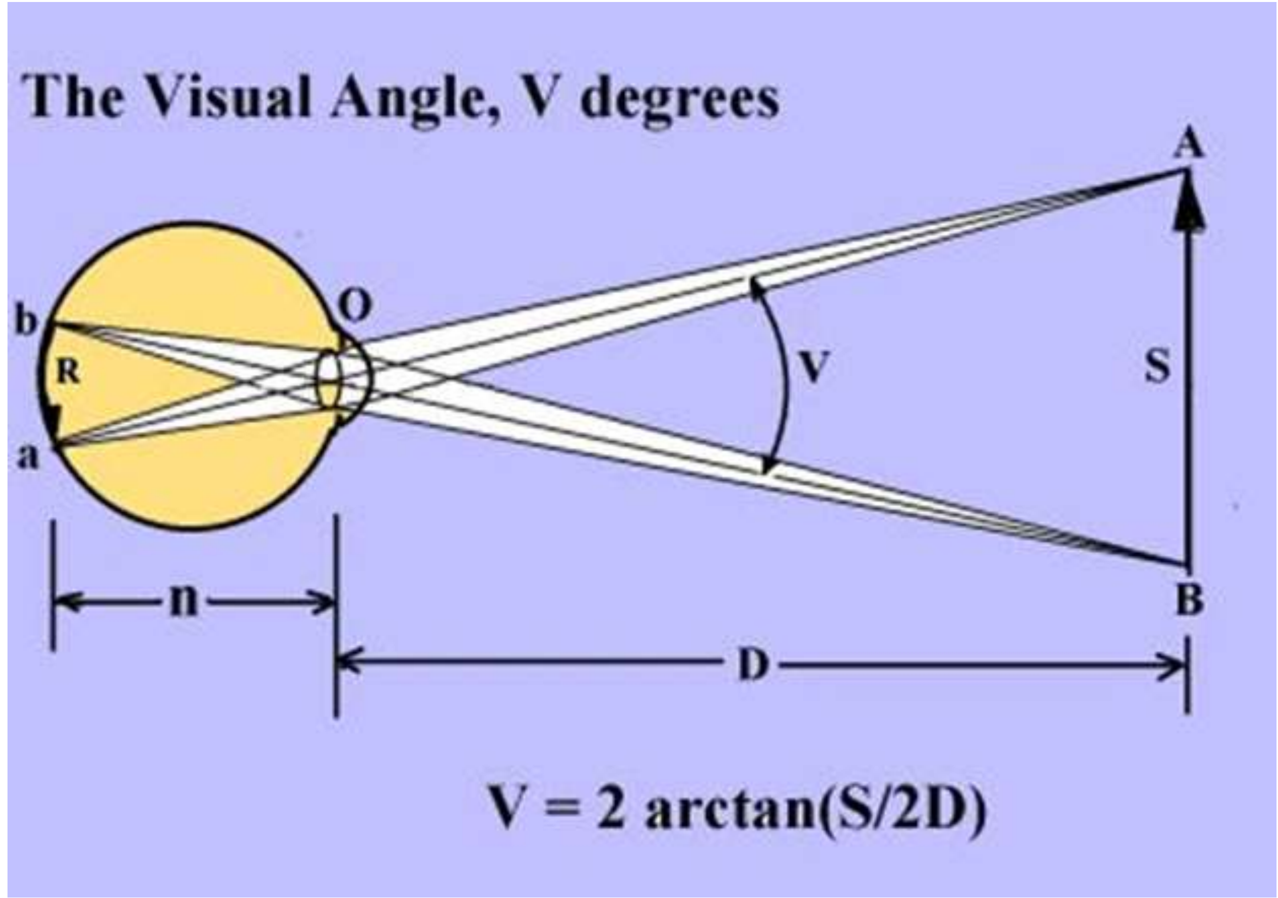
$-90^\circ \le a \le +90^\circ$ for ϵ from 0° to about 70°

The visual field

Left VF **Right VF**

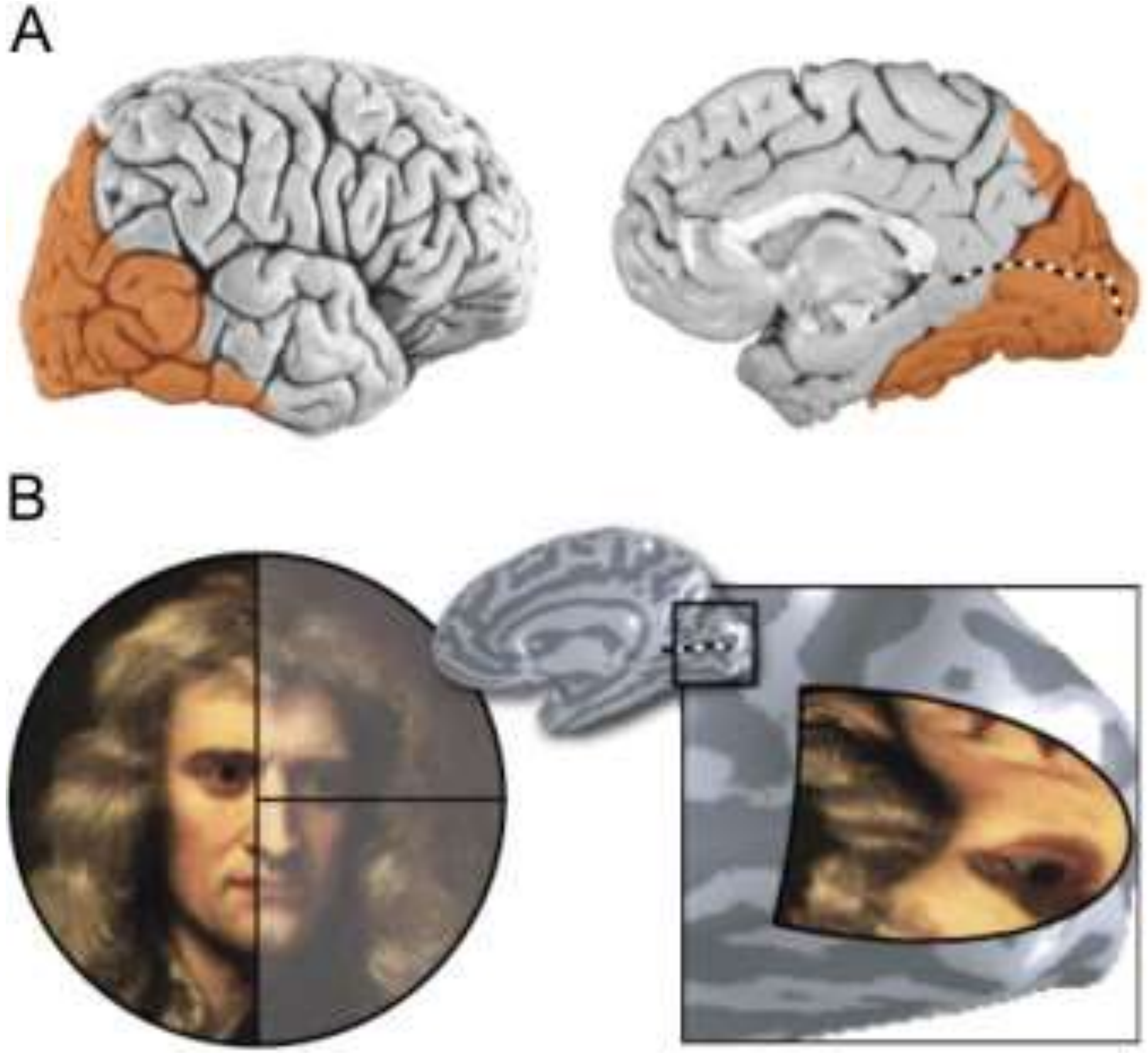


Right eye visual field



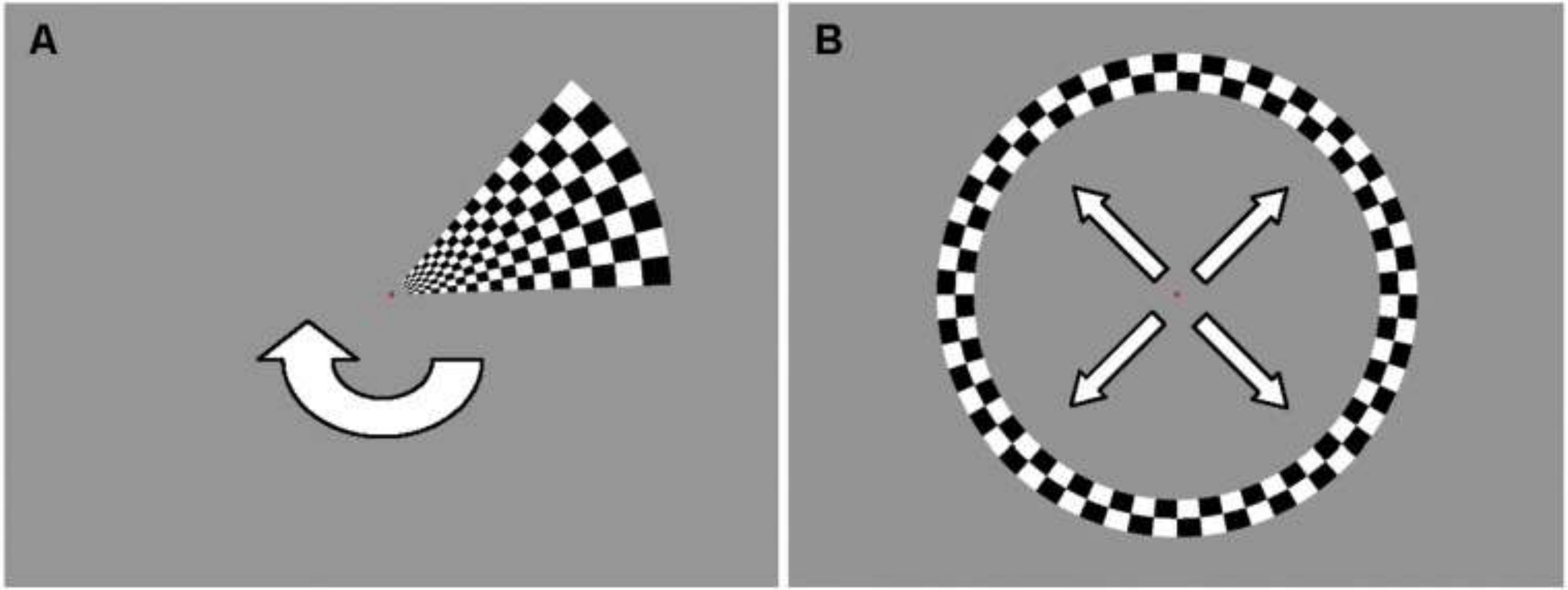
Visual angle

Visual field map(retinotopic mapping)



The visual world is mapped onto the cortical surface in a topographic manner: **neighboring points in a visual image evoke activity in neighboring regions of visual cortex.**

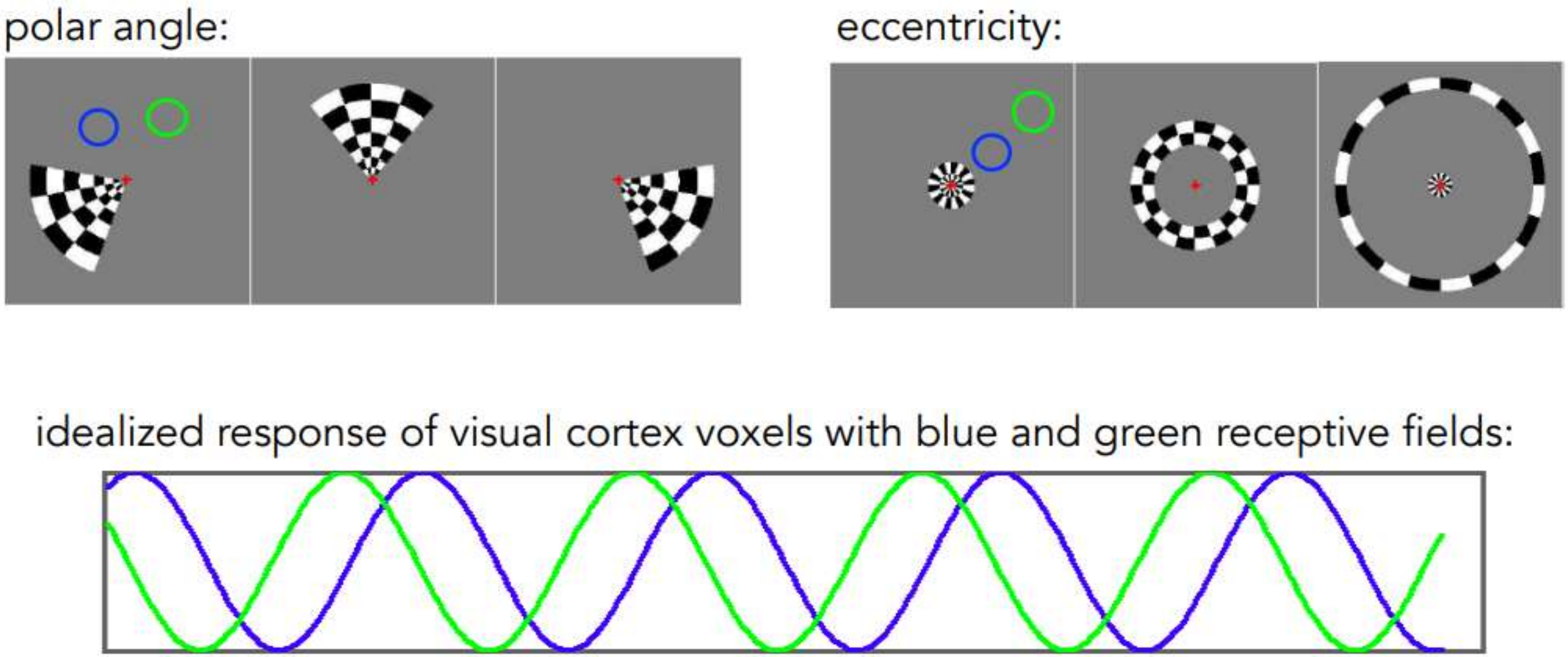
Phase encoding for retinotopic mapping



Polar angle mapping.

Eccentricity mapping

Phase-encoded mapping

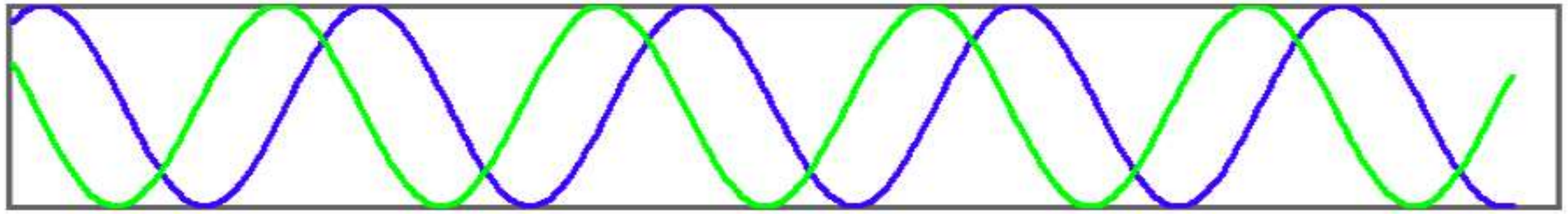


- Periodic stimuli map out visual space in polar angle in eccentricity
- Voxel responses are periodic, having the same frequency but different phases
- So, parameter of interest is response *phase* because it tags visual field location

Phase-encoded mapping

polar angle: eccentricity:

Stimulus is cyclical, so response is cyclical



how to measure the phase of a cyclical function?

1) Fourier analysis (phase of a sinusoid)

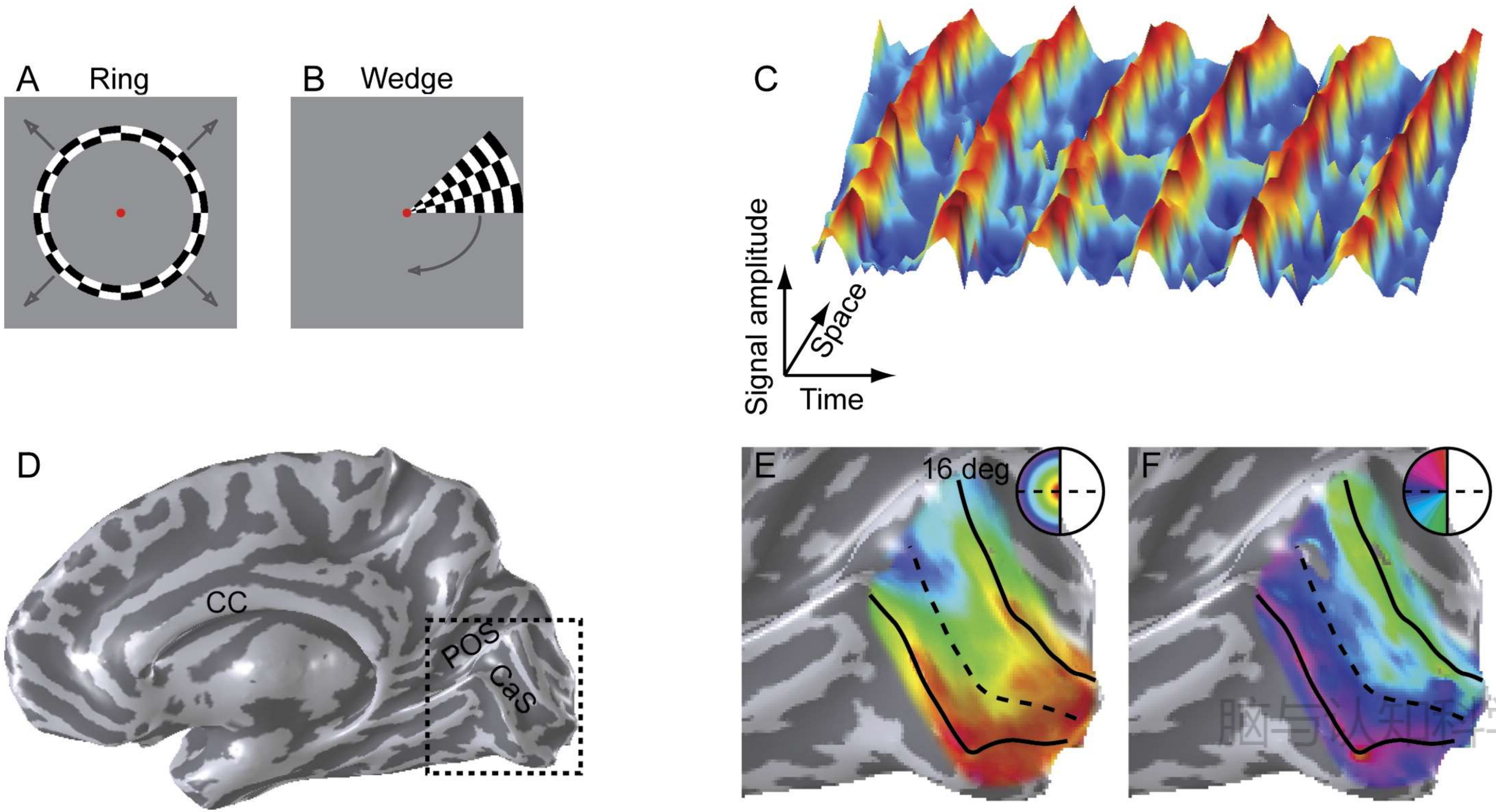
2) Cross-correlation (phase of a specified model function)

difference from standard GLM analysis?

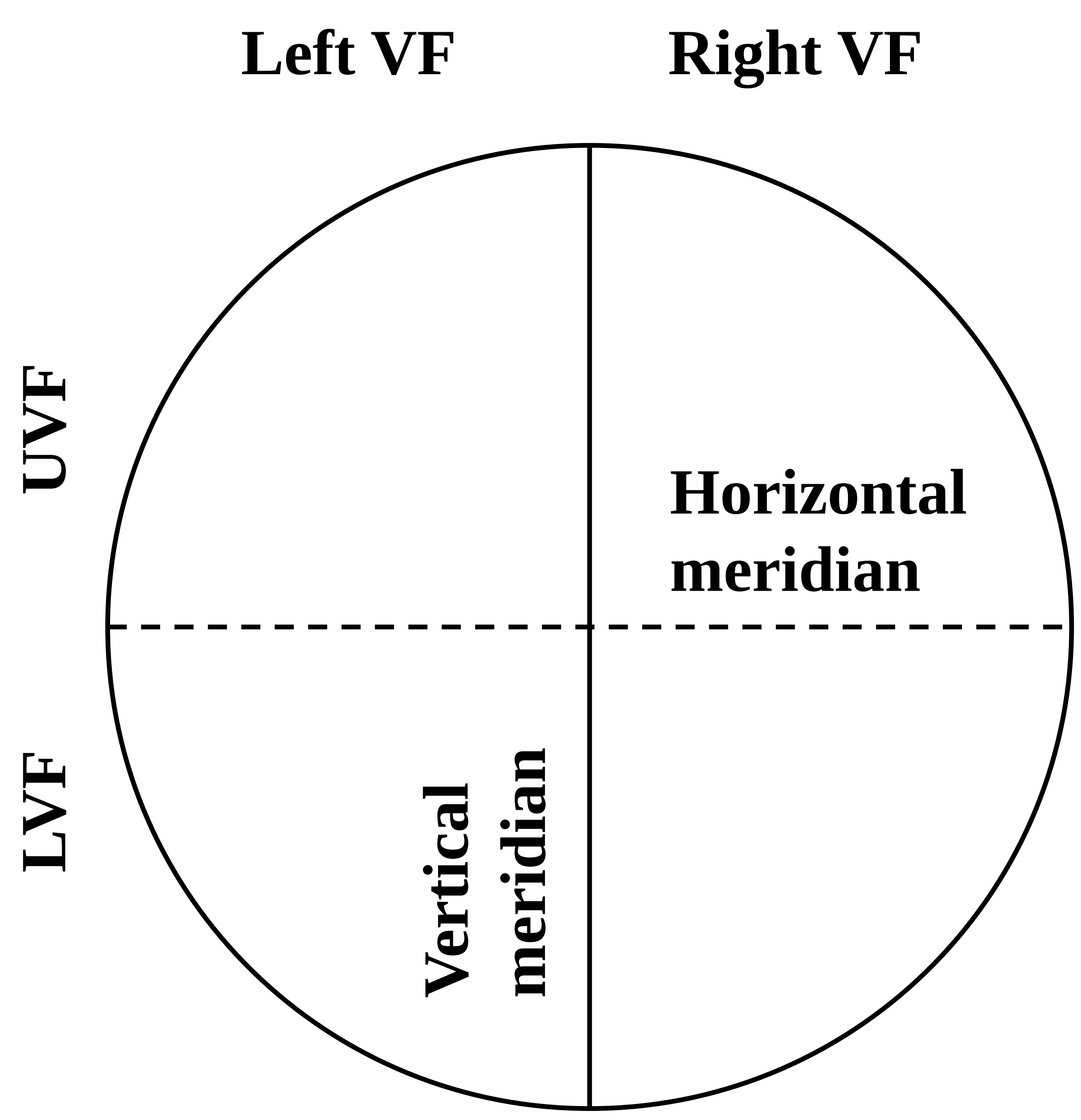
1) key parameter is response phase, not amplitude

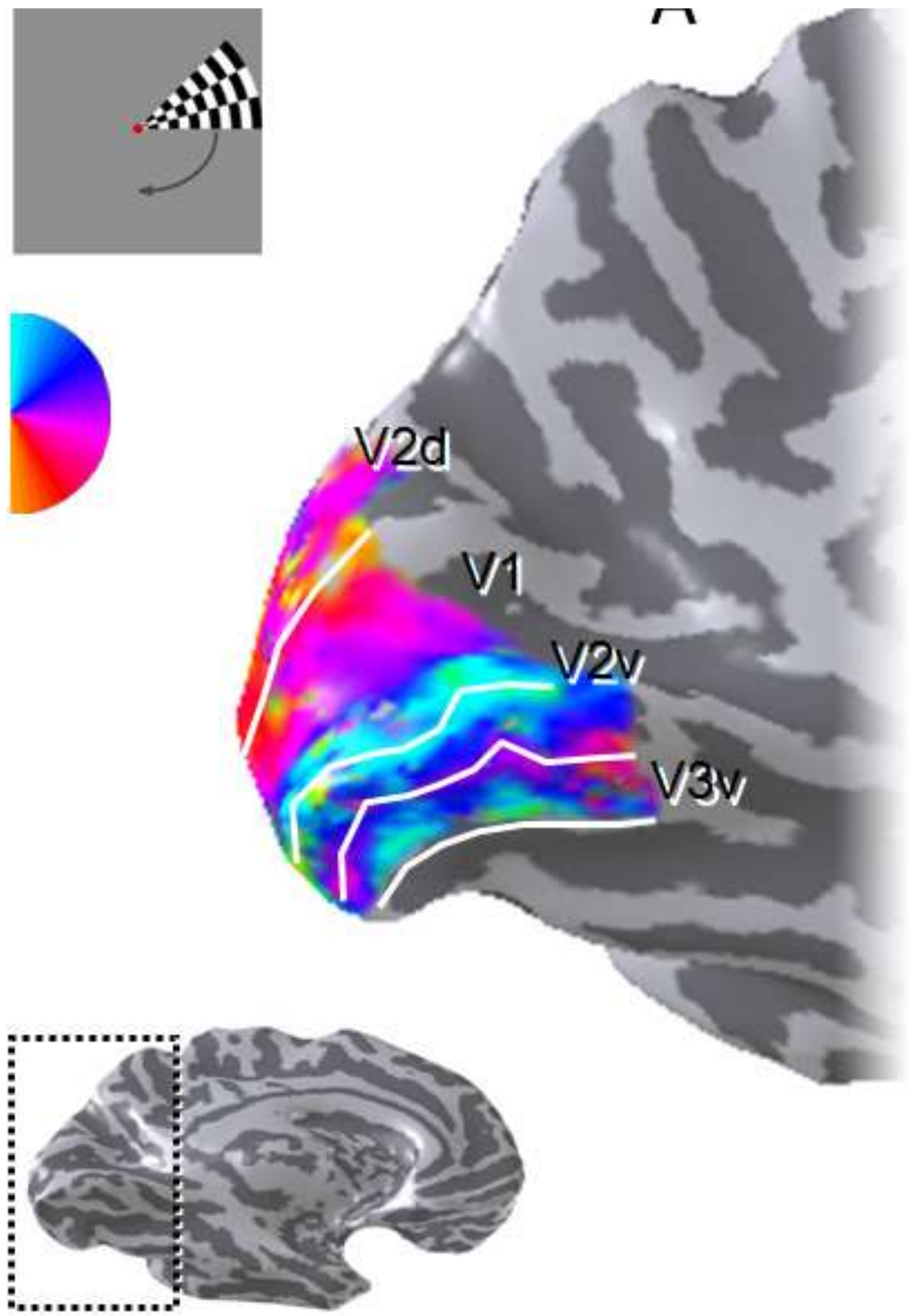
2) no contrast, overlapping responses contributes to efficiency of design

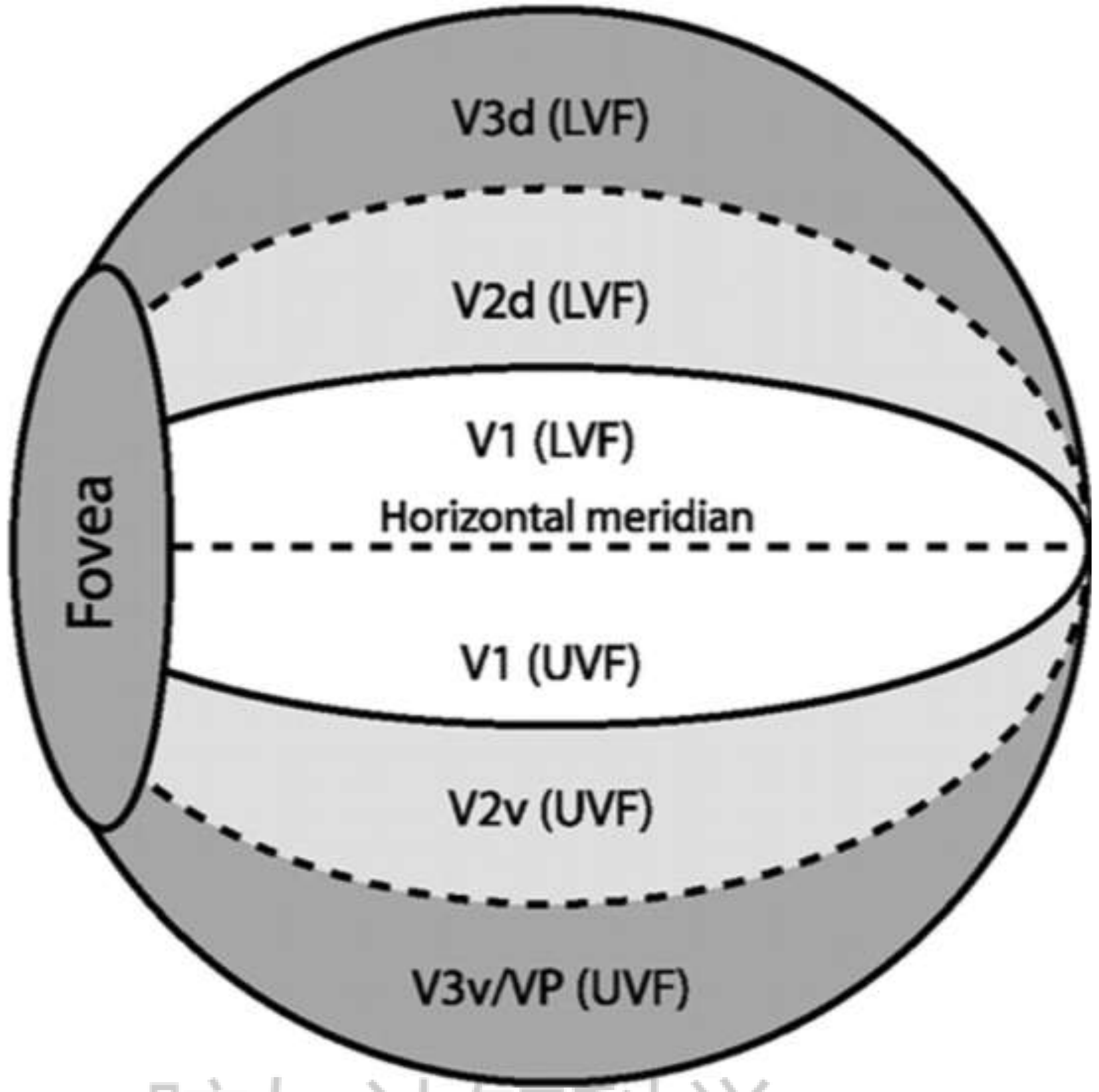
Phase-encoded mapping



Primary maps: V1, V2, and V3

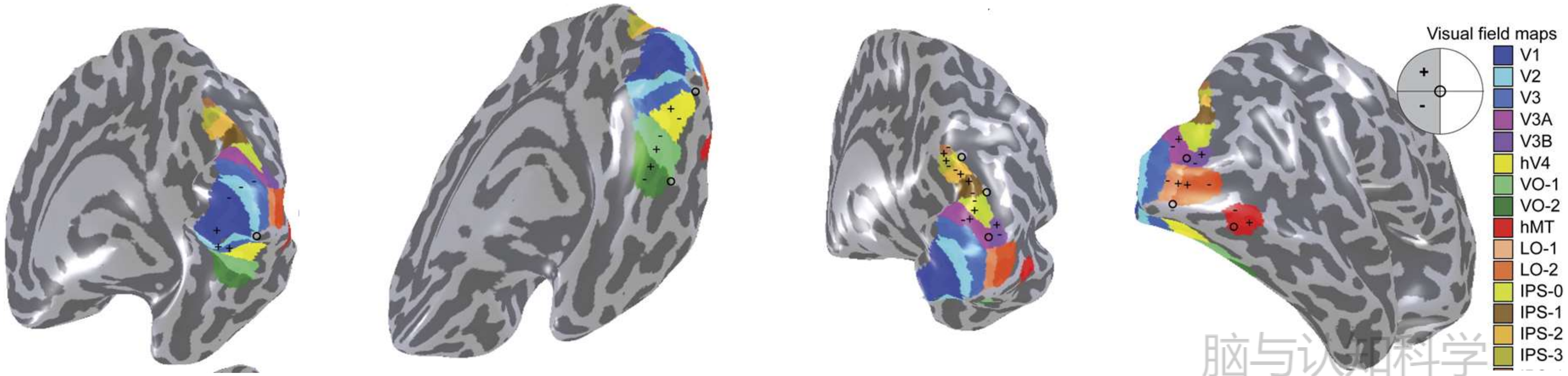




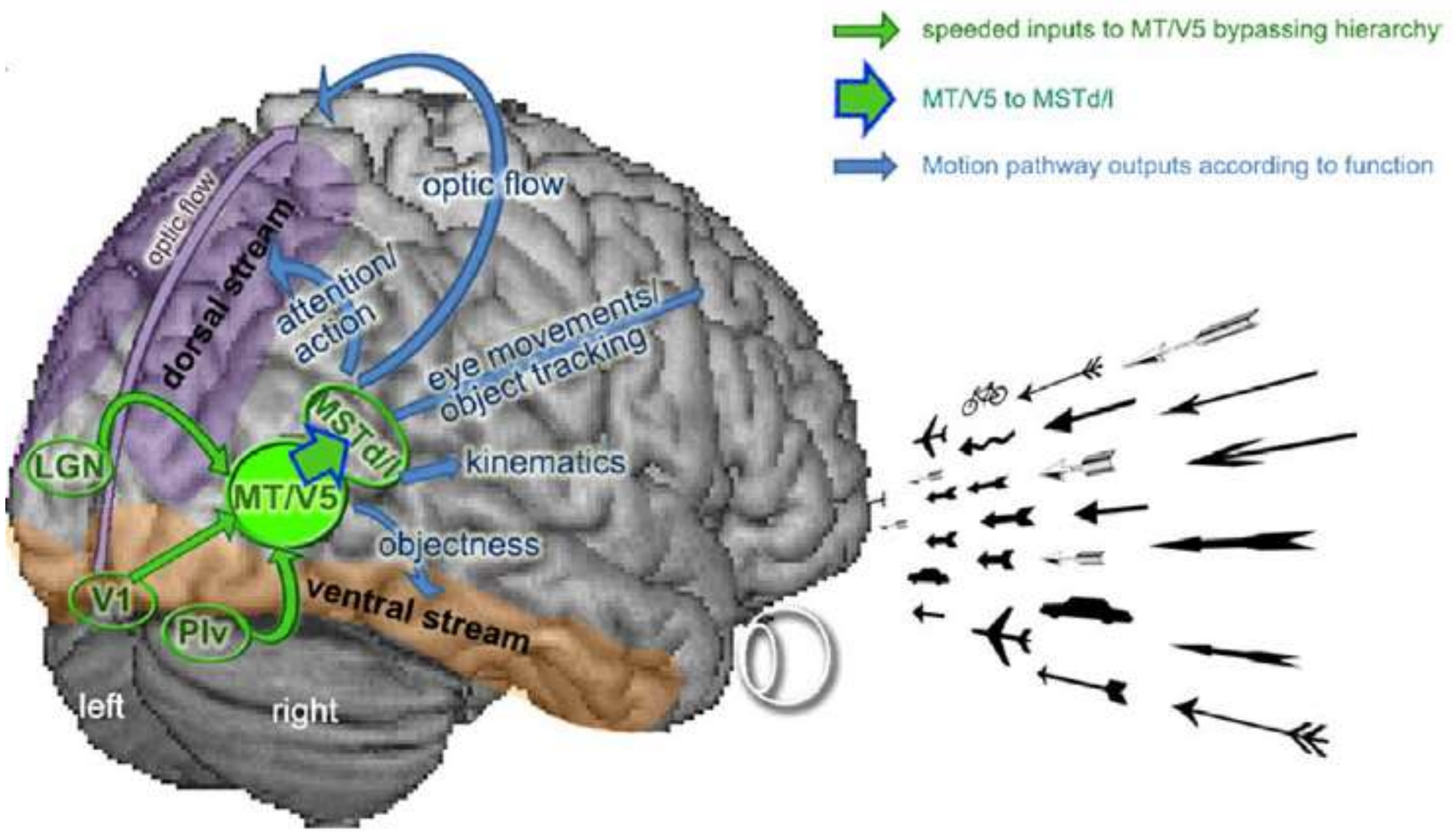


Multiple visual field maps

- A visual field map contains no more than a **single** representation for each point in the visual field.
- A visual field map represents a **substantial part** of the visual field.
- A visual field map represents a substantial part of the visual field in an **orderly** fashion;
- A visual field map should be **consistent** across individual subjects

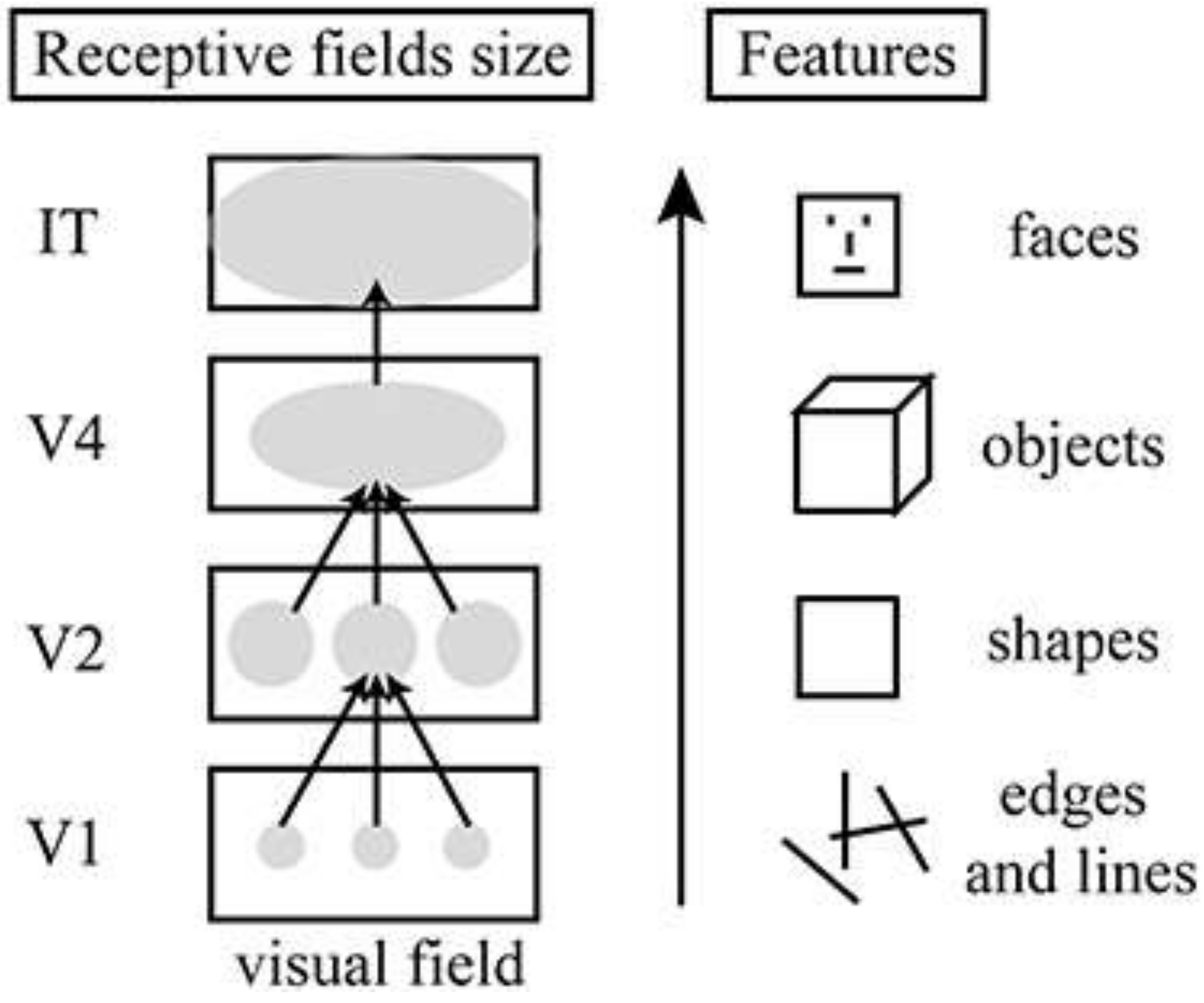


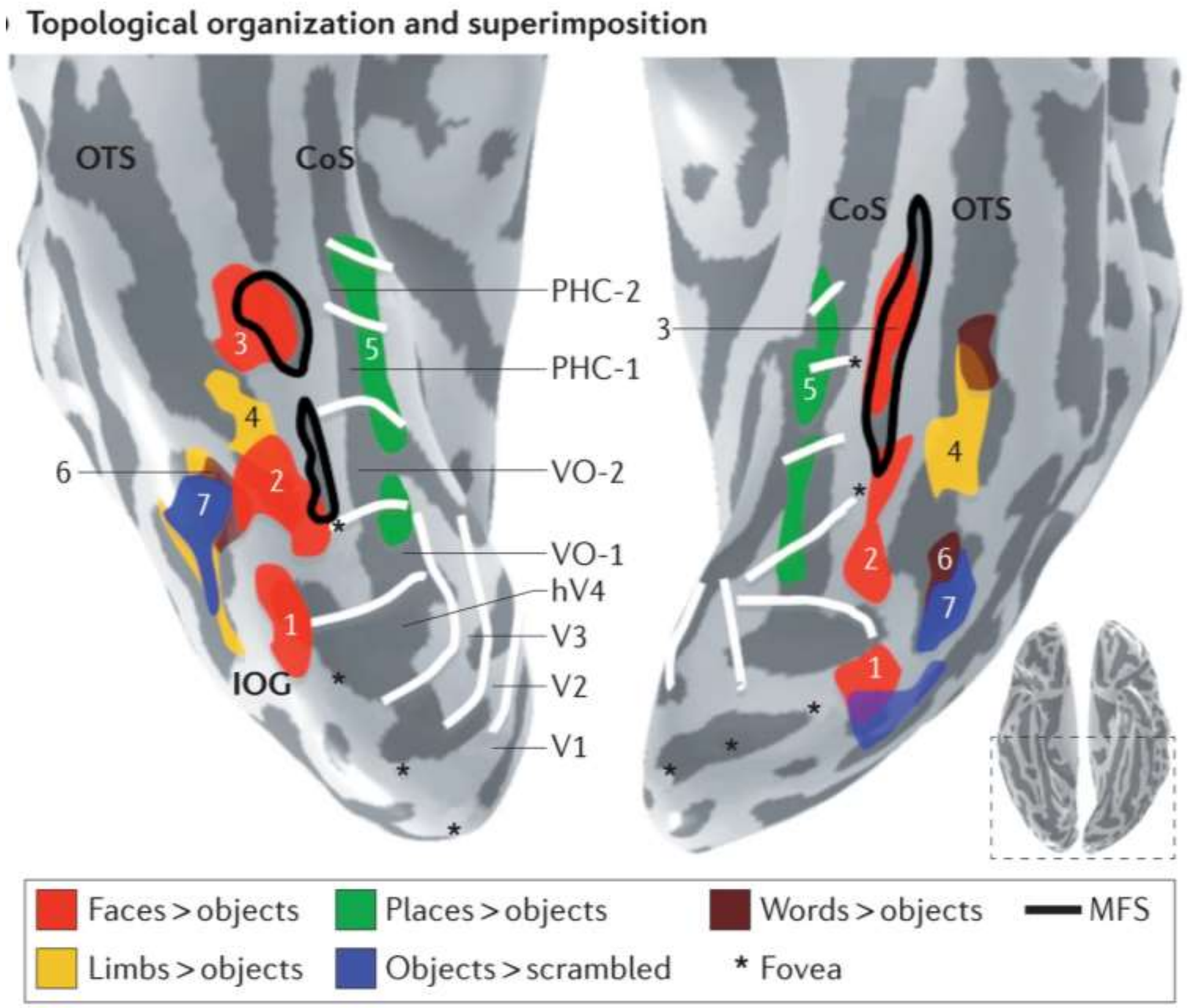
V5/MT: motion perception



- All of neurons in V5/MT are sensitive to motion.
- Bilateral damage of V5/MT results in visual akinetopisa（运动失认症）.
- Biological motion, and movement in other sensory modality do not rely on V5/MT

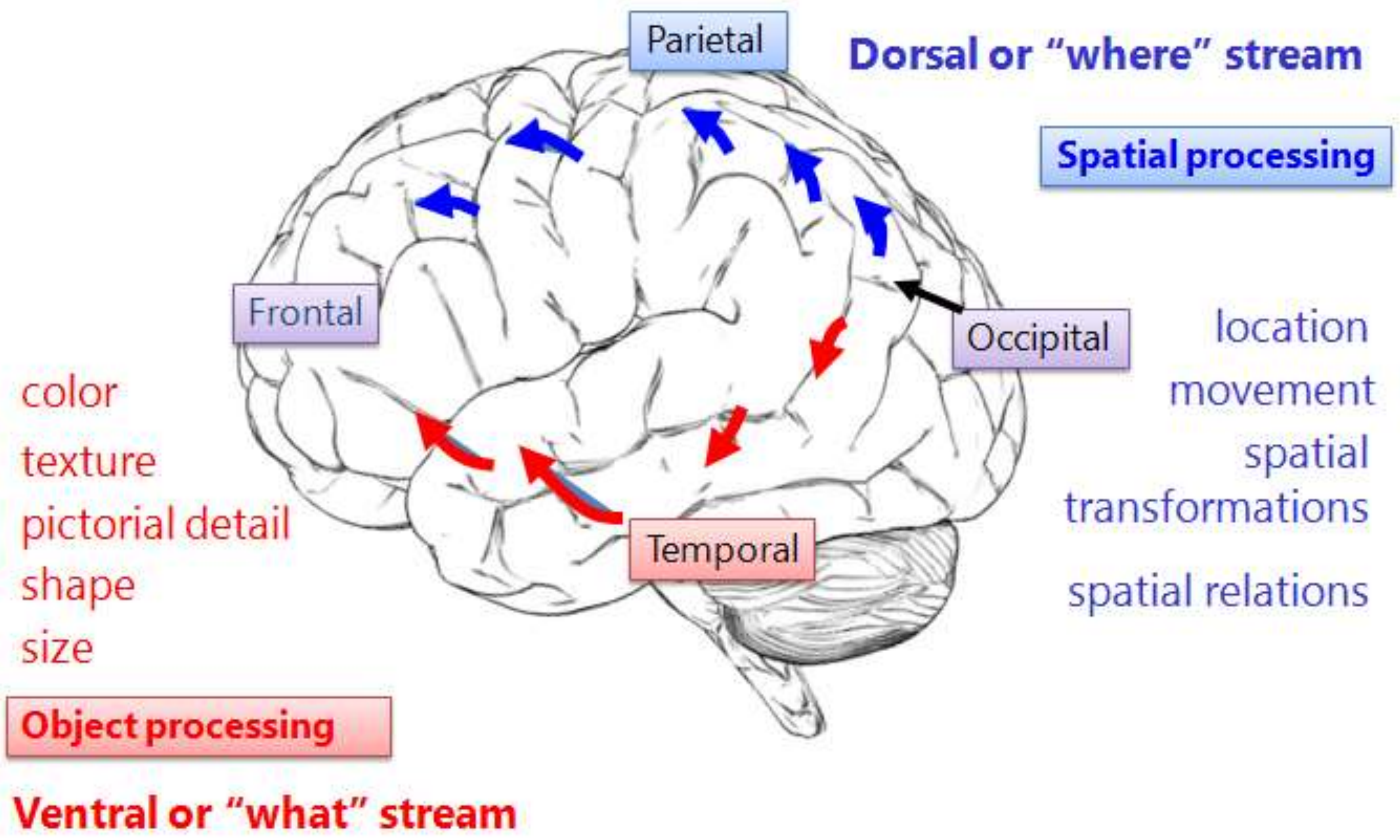
Category specificity visual areas

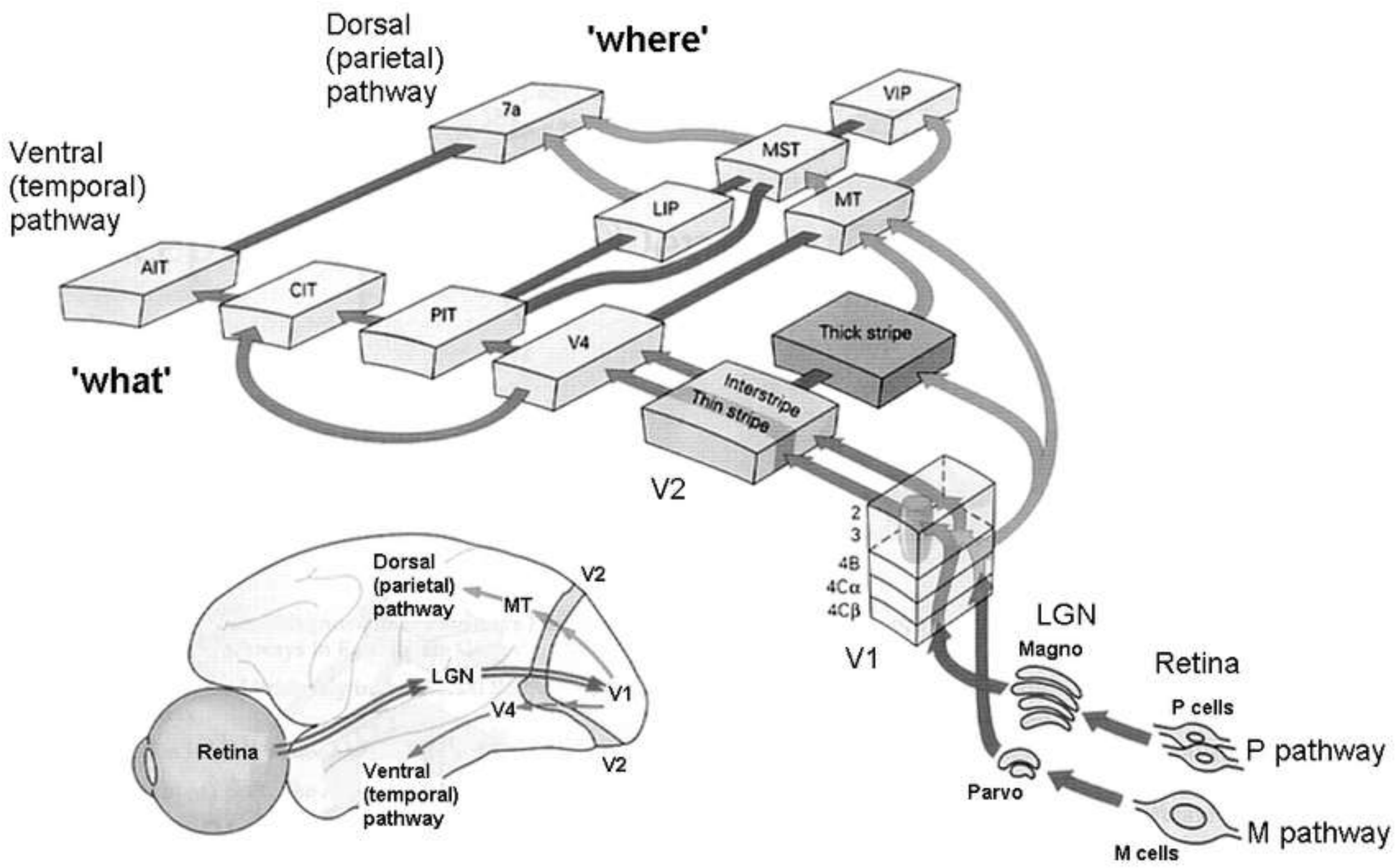




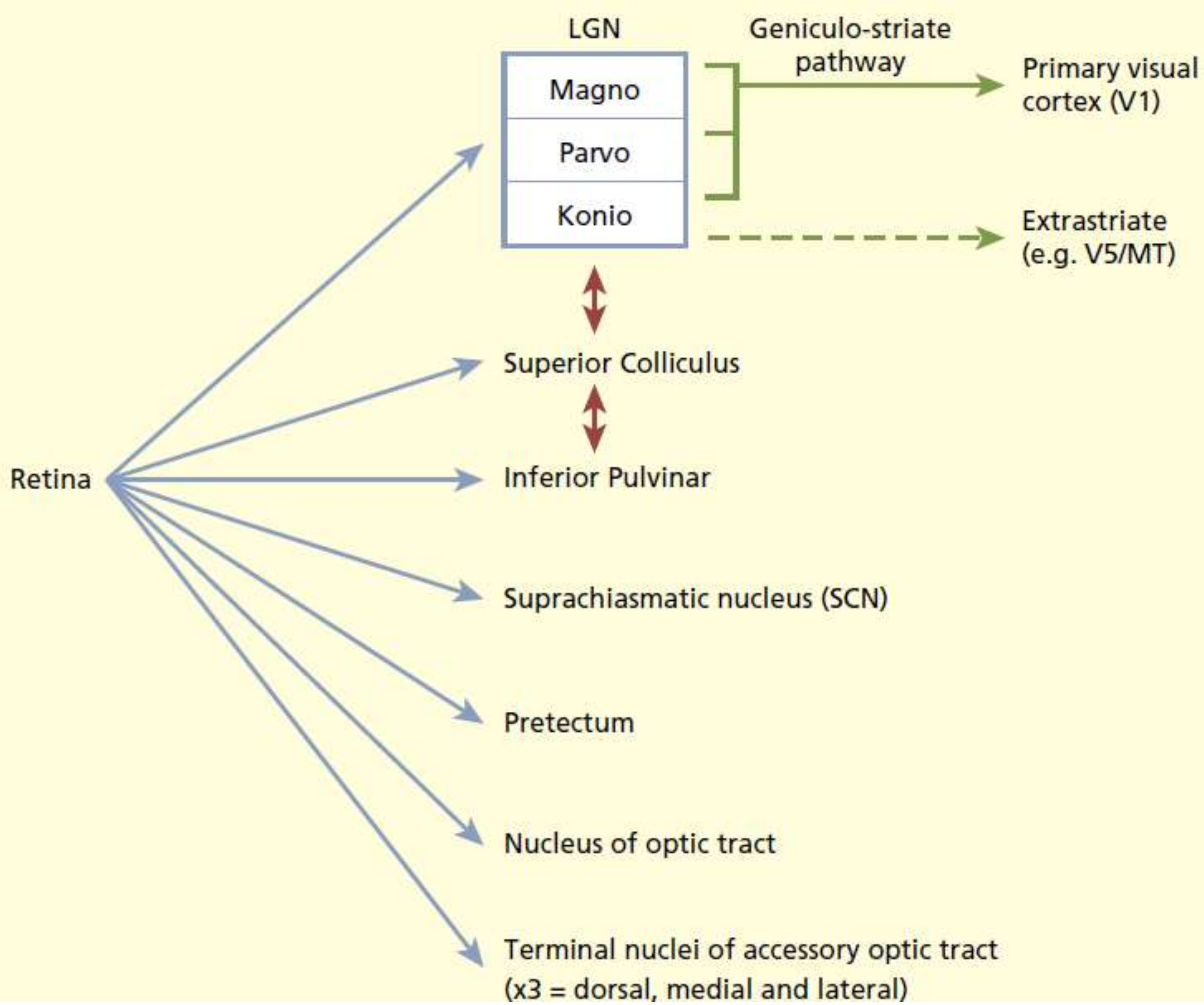
Dorsal and ventral visual stream

腹侧通路(Ventral Stream):沿着大脑皮层的枕颞叶分布，包括纹状体皮层、前纹状体皮层和下颞叶，主要功能是物体的识别
背侧通路(Dorsal Stream):沿着枕顶叶分布，包括纹状体皮层、前纹状体皮层和下顶叶，主要功能是空间位置和运动的识别



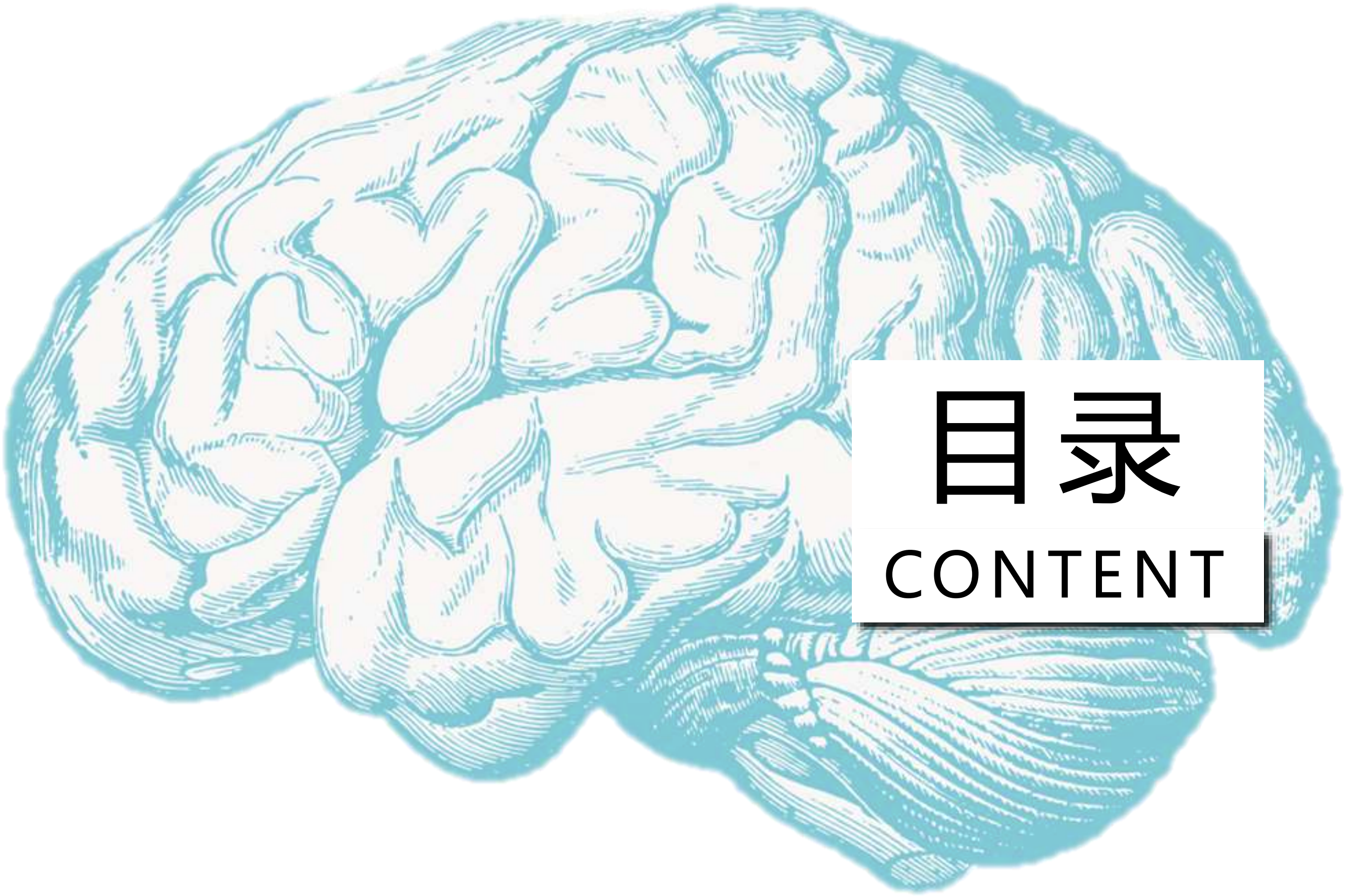


Multiple routes from the retina to the brain



1. The pathway via the lateral geniculate nucleus to V1 is the most well understood and appears to make the largest contribution to human visual perception.

2. The other routes are evolutionarily more ancient. Evolution appears not to have replaced these routes with "better" ones, but has retained them and added new routes that enable finer levels of processing.

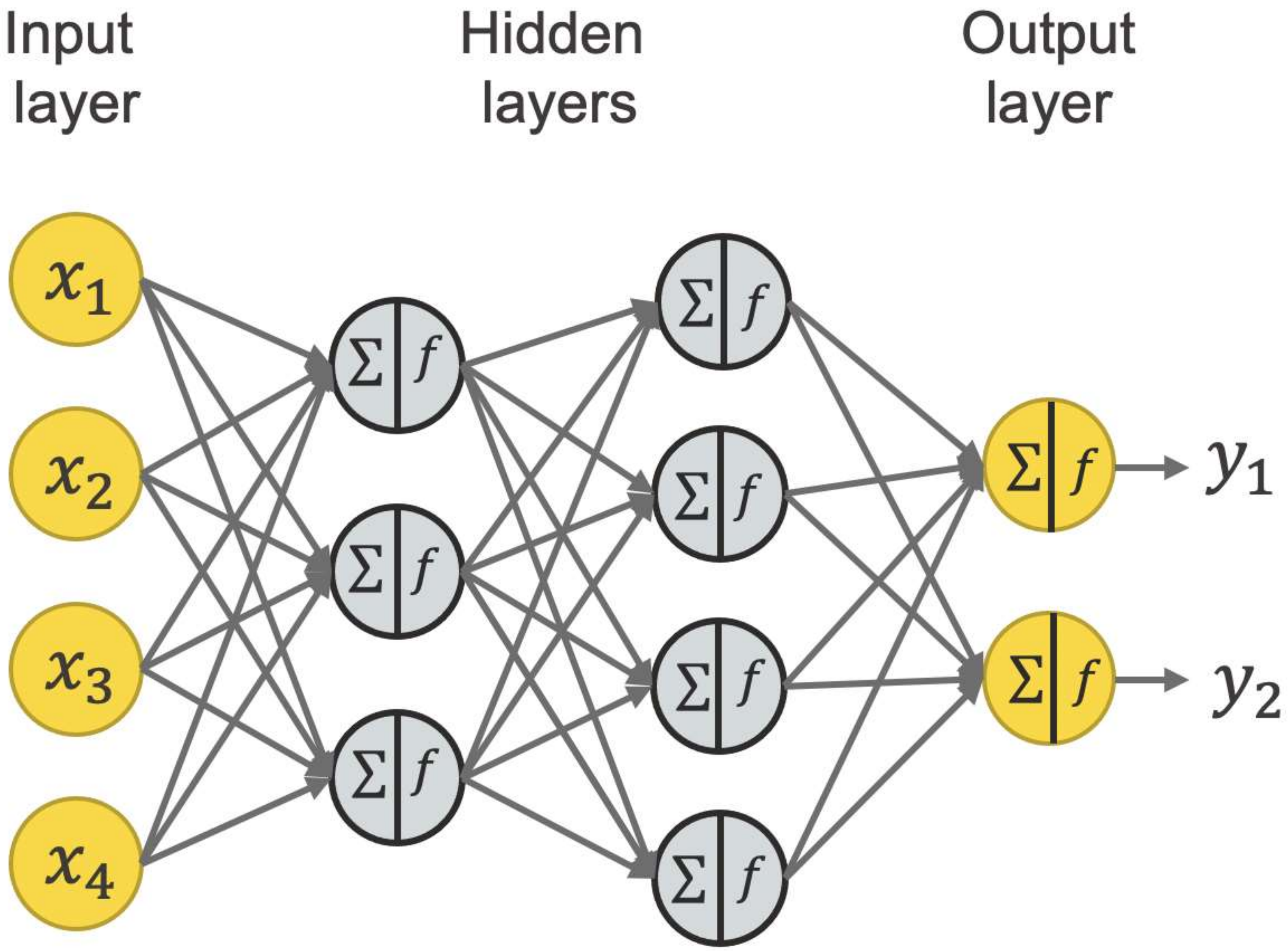


目录
CONTENT

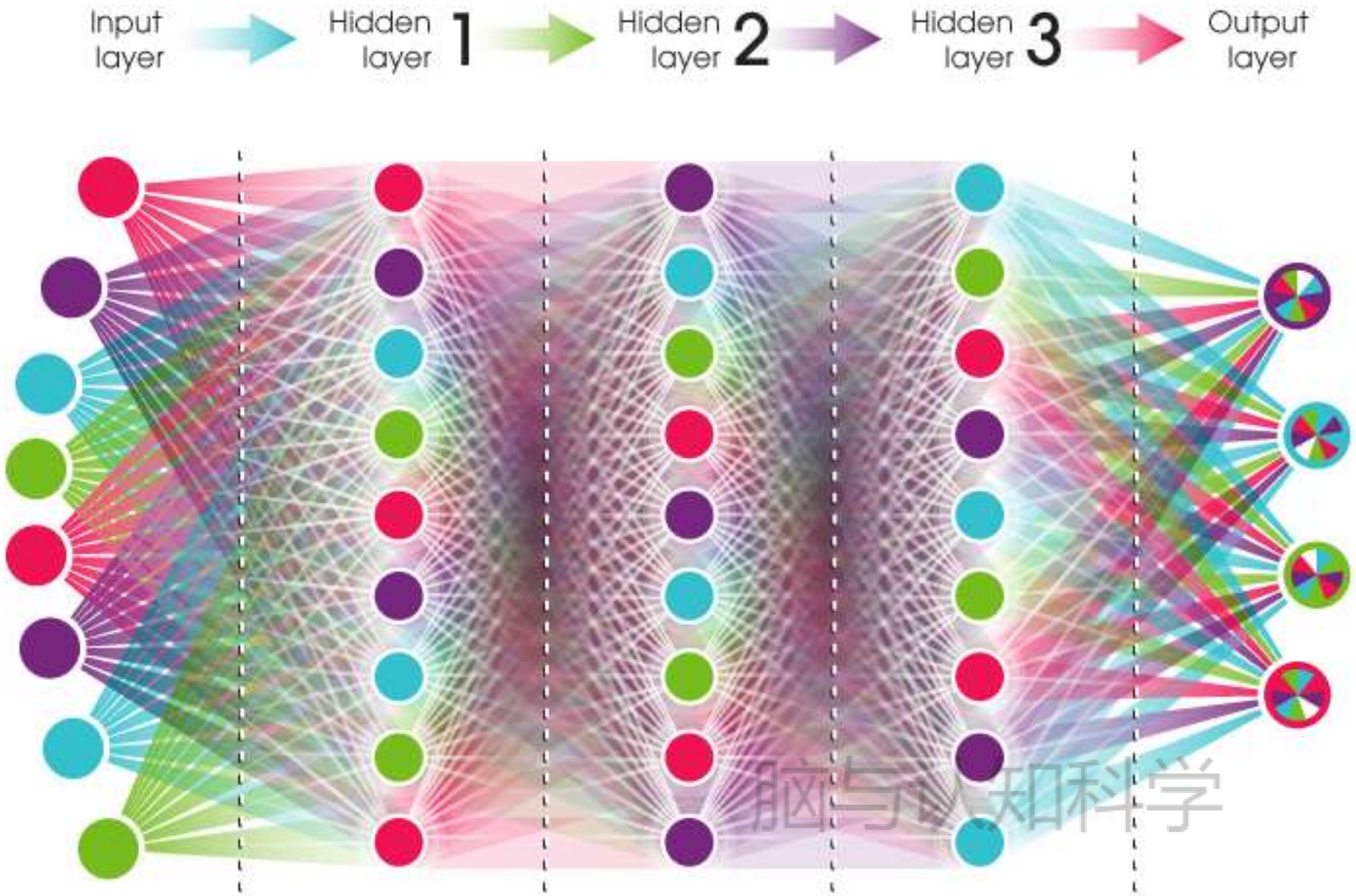
Deep neural network

A DNN is an ANN with multiple hidden layers between the input and output layers. DNNs can model more complex non-linear relationships than shallow ANNs.

Shallow neural network

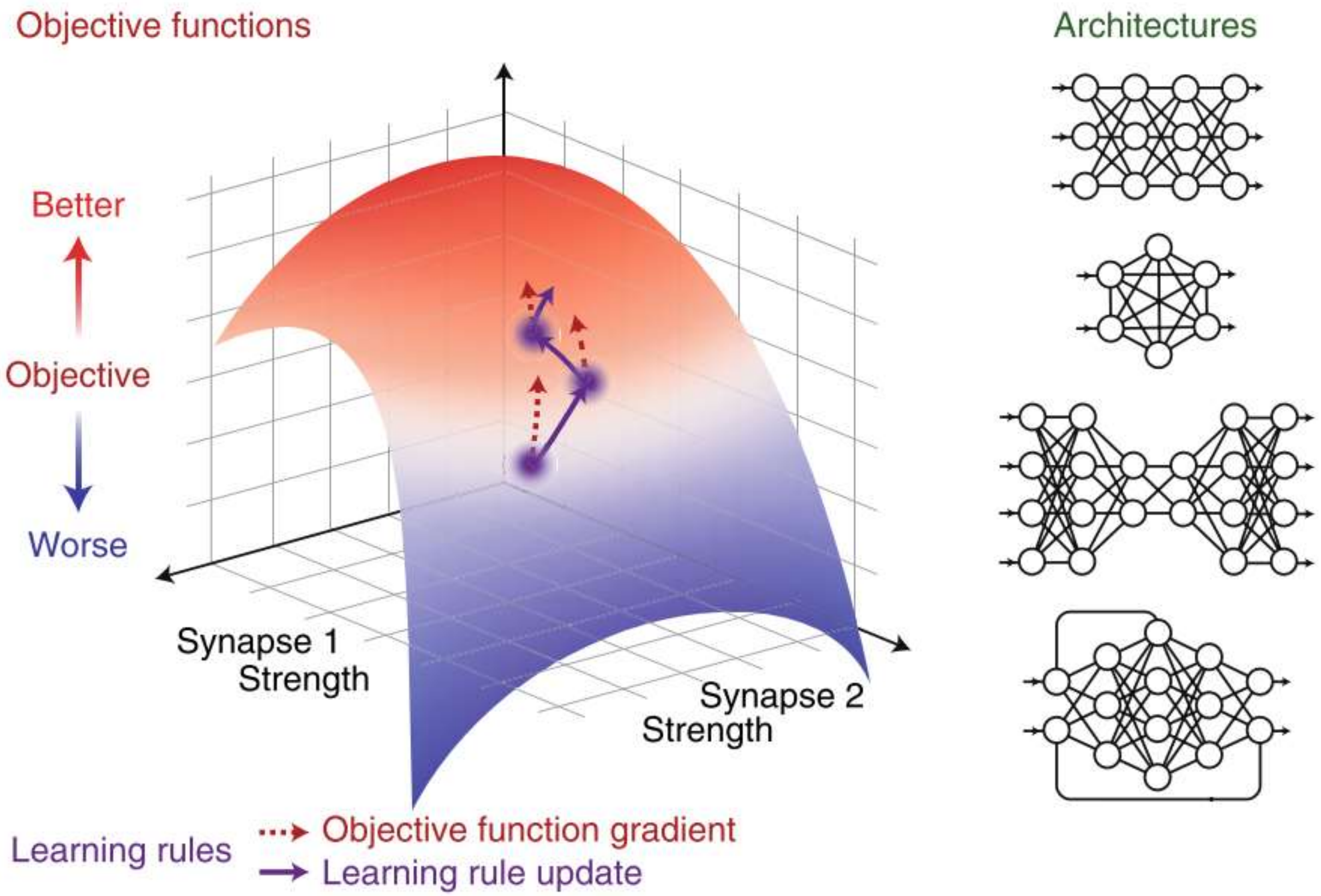


Deep neural network



Three core components of ANN

When designing ANNs, we do not craft the specific computations performed by the network. Instead we specify three components.



Architectures specify the arrangement of units in the network and determine the flow of information.

Objective functions quantify the performance of the network on a task.

Learning rules involve finding synaptic weights that maximize or minimize the objective function.

Objective functions of ANN

The network is trained to optimize the value of an objective function or, commonly, minimize the value of a loss function.

$$L = \frac{1}{N}\sum_i L\left(\boldsymbol{y}^{(i)}, \boldsymbol{y}_{\text{target}}^{(i)}\right),$$

where $L\left(\boldsymbol{y}^{(i)}, \boldsymbol{y}_{\text{target}}^{(i)}\right)$ quantifies the difference between the target output and the actual output.

Loss for regression $$L = MSE = \frac{1}{n}\sum_i^n (\hat{y}_i - y_i)^2$$

Loss for classification

Two classes $$L = \frac{1}{N}\sum_i L_i = \frac{1}{N}\sum_i -[y_i \cdot log(p_i) + (1 - y_i) \cdot log(1 - p_i)]$$

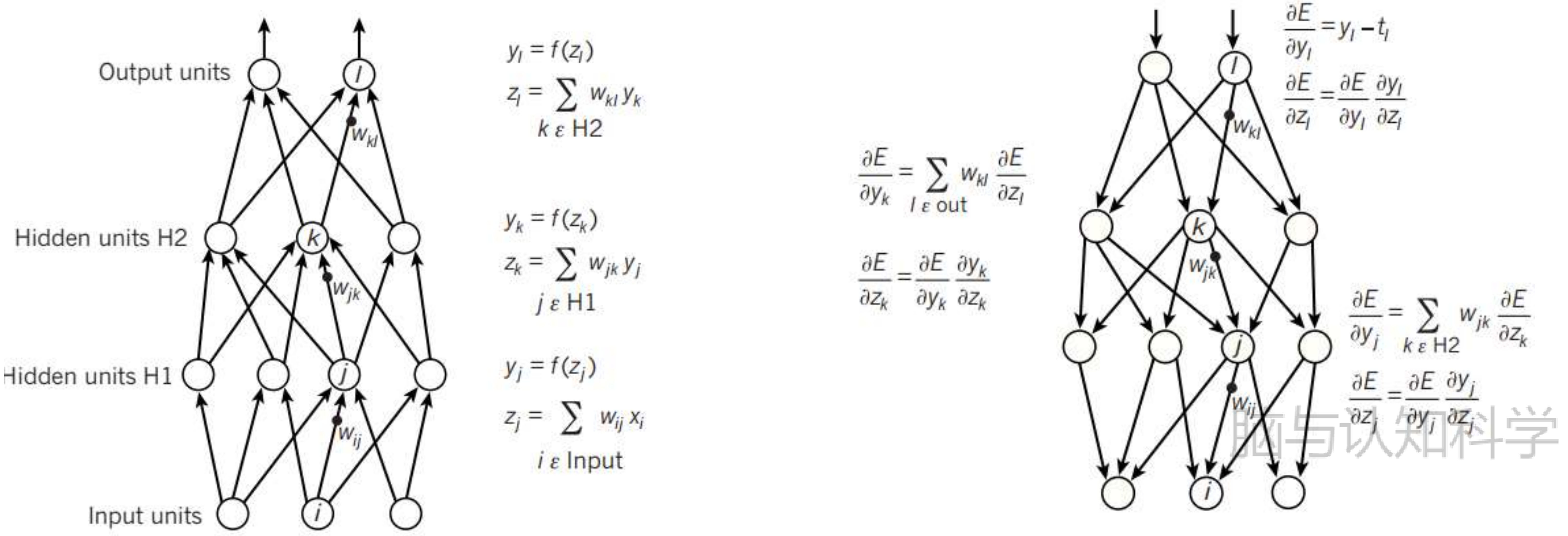
Multiple classes $$L = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i \sum_{c=1}^{M} y_{ic} \log(p_{ic})$$

Learning rules of ANN

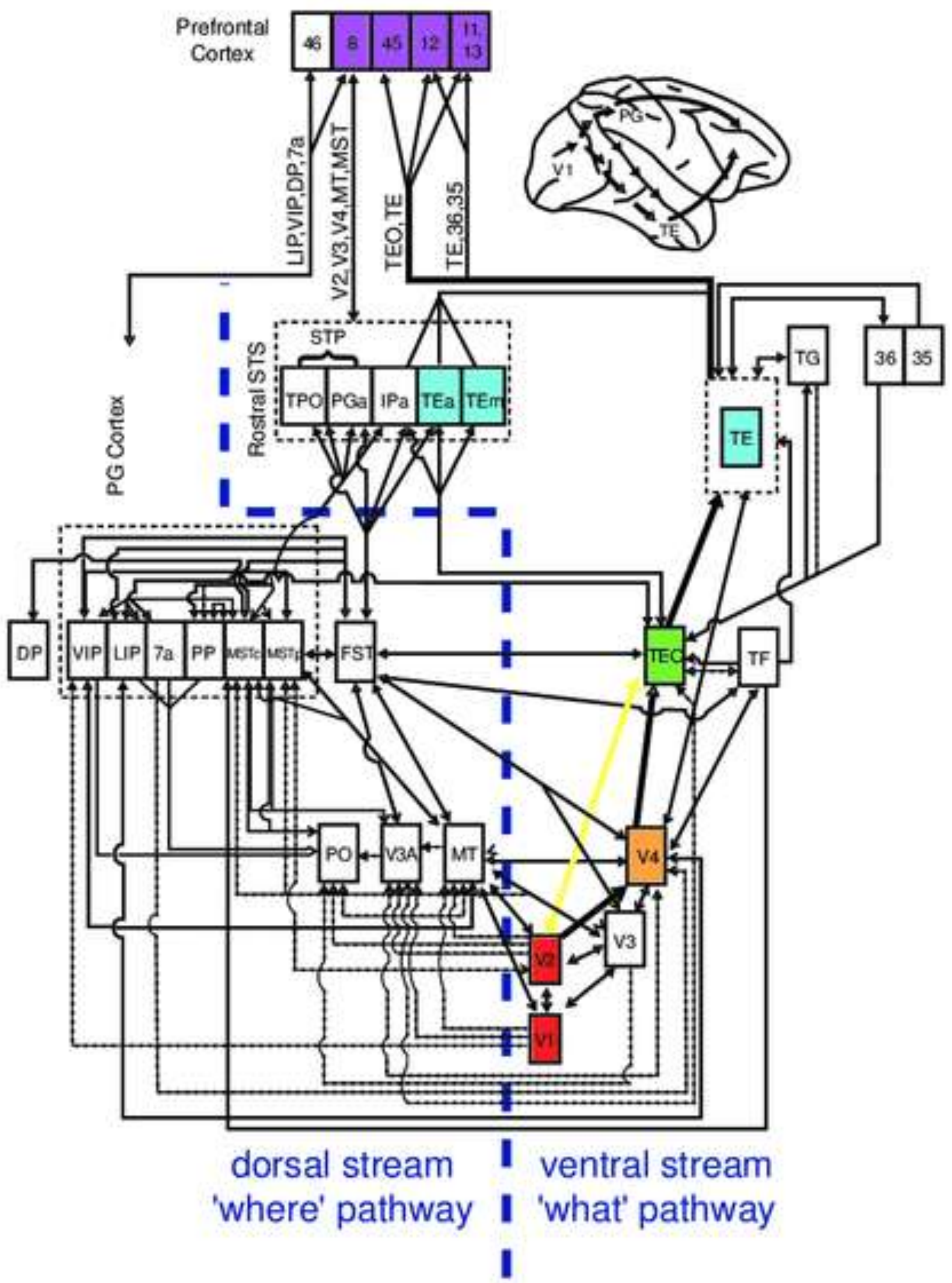
The signature method of training in deep learning is **stochastic gradient descent (SGD)**

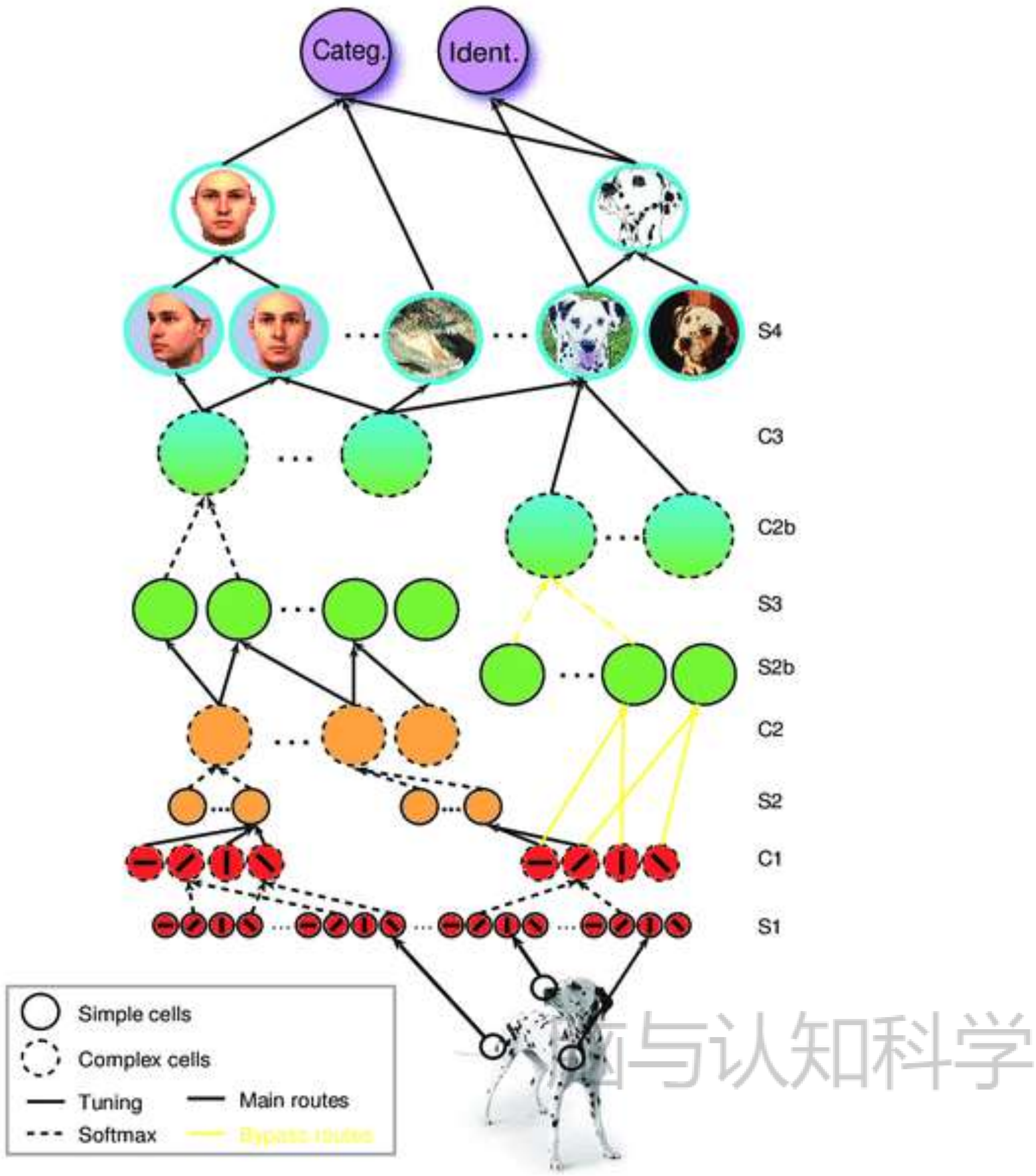
$$\Delta\boldsymbol{\theta} = -\eta\frac{\partial L}{\partial\boldsymbol{\theta}}.$$

Backpropagation algorithm uses the chain rule to recursively calculate gradients backwards from the output for deep ANNs.

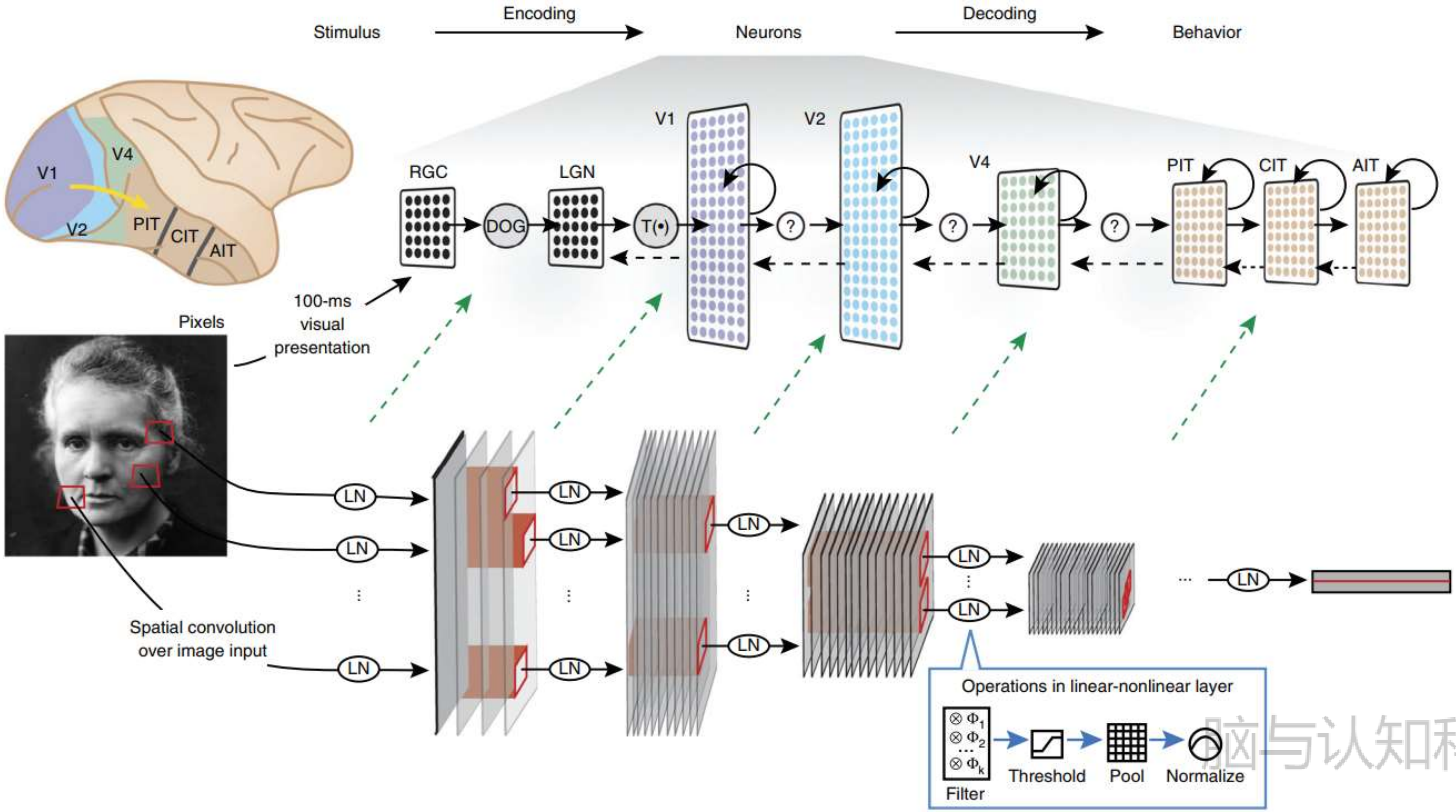


Distributed hierarchical processing in the primate cerebral cortex



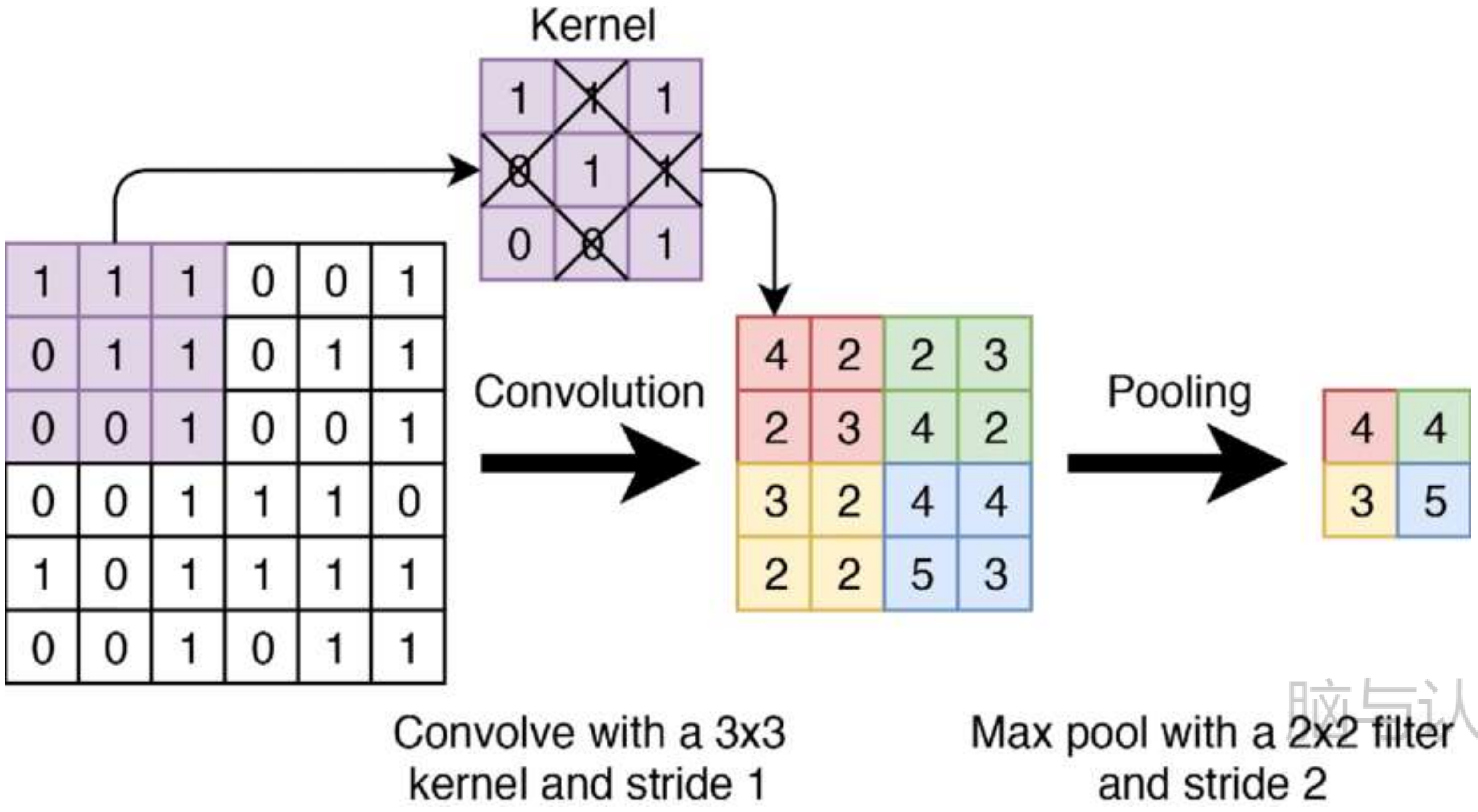


Deep convolutional neural network



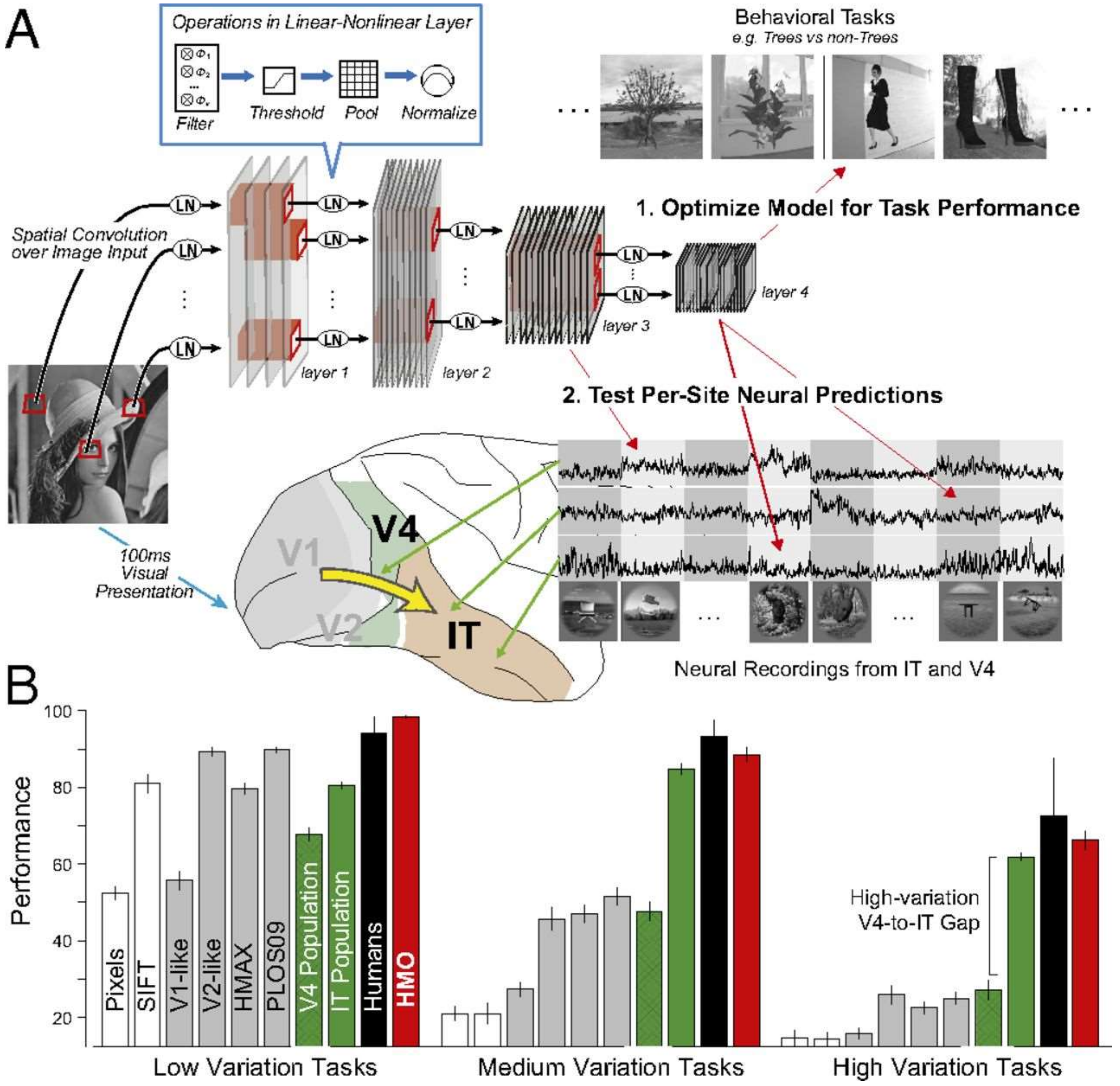
LN operations in CNN

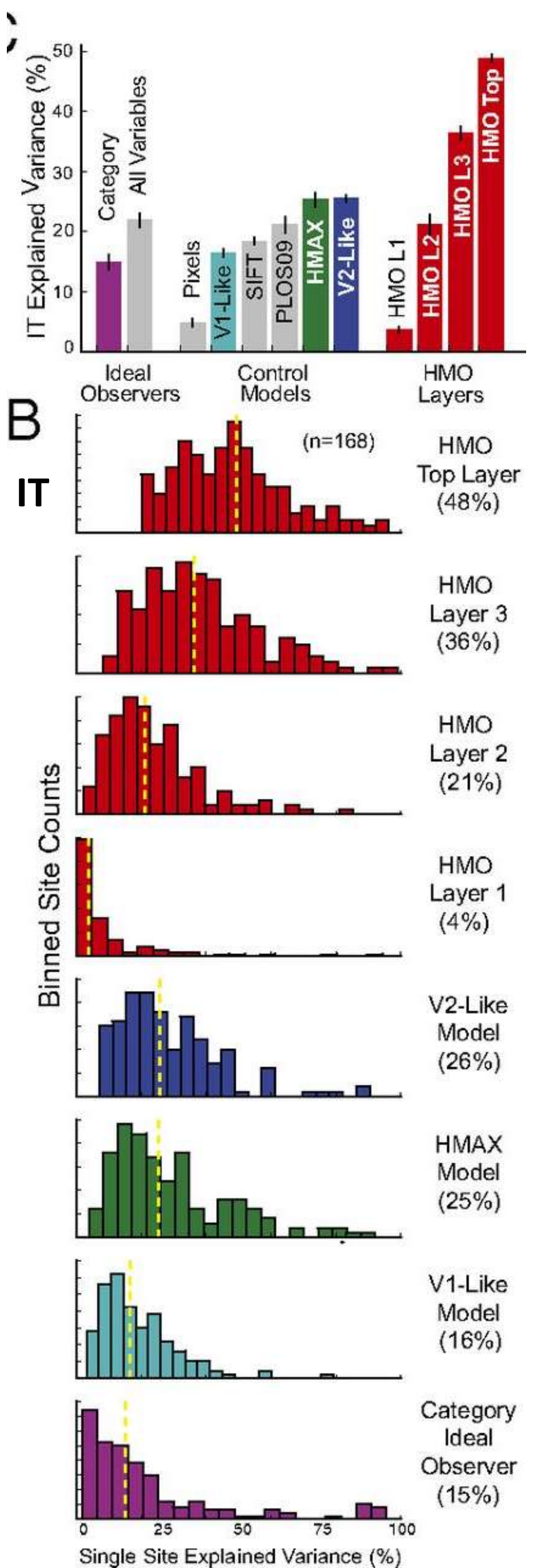
Each convolutional layer is composed of simple and neuronally plausible basic operations, including linear filtering, thresholding, pooling, and normalization

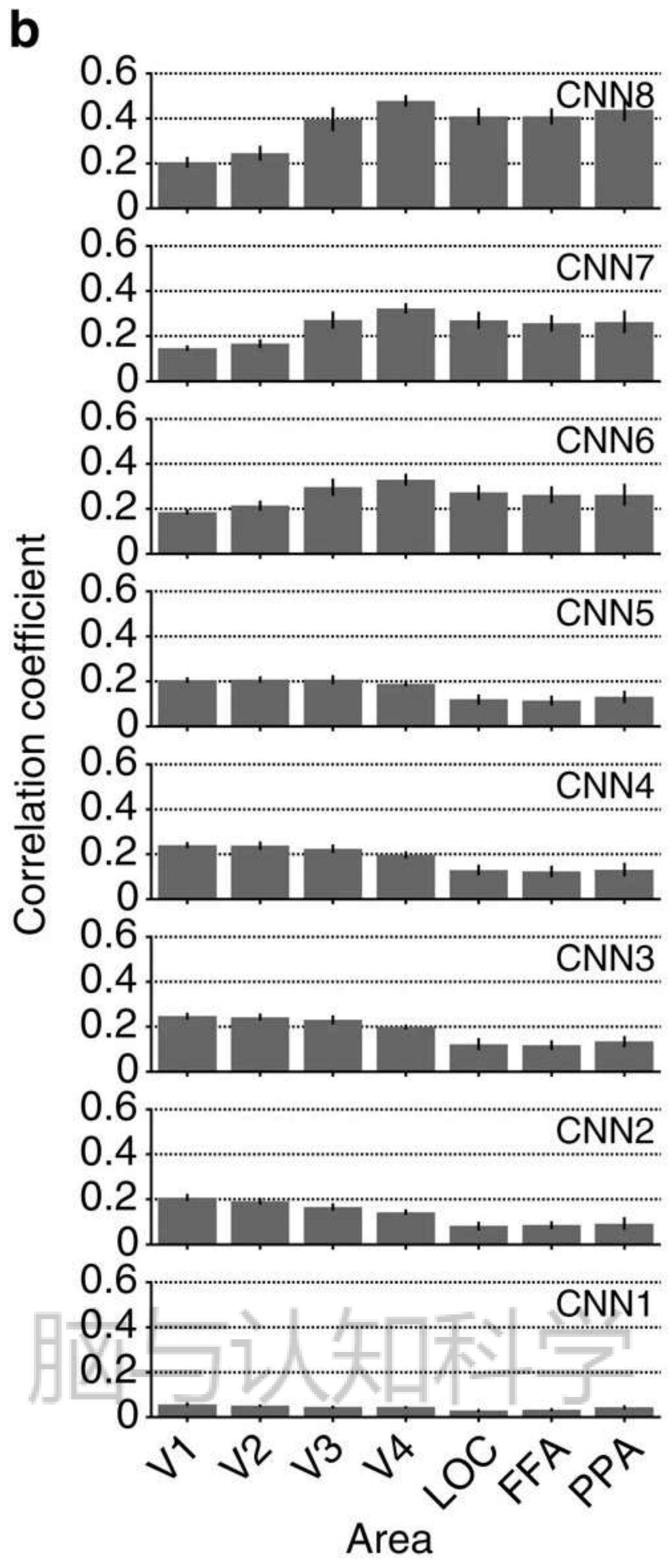


Representation correspondence between DNN and visual cortex

Yamins DL, et al. PNAS,2014.

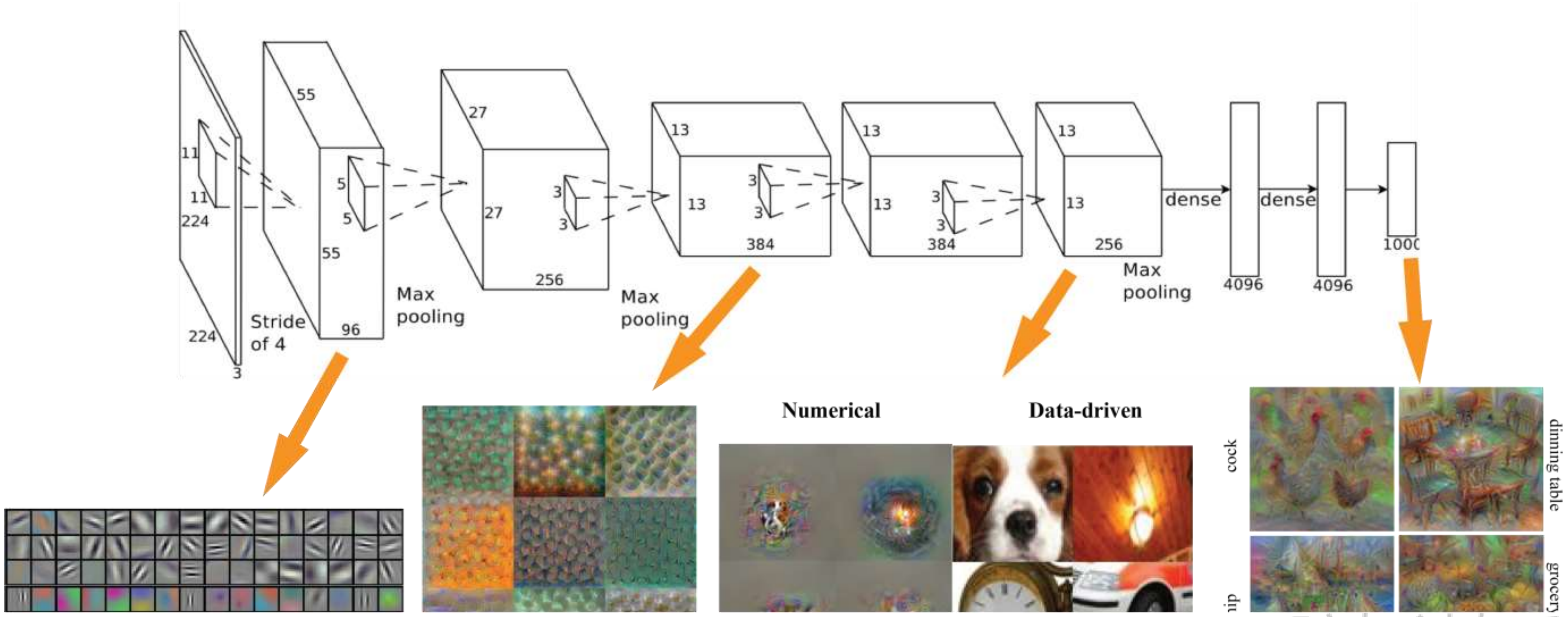






Open the black box

Internal representations of CNN provide the best current models of representations of visual images in inferior temporal cortex in humans and monkeys.



本章关键知识点

1. 不同视觉神经元的功能差异体现在其感受野的不同。
2. 视觉系统由多个不同功能区构成，不同功能区相互连接形成了视觉脑网络。
3. 大脑视觉加工是通过对不同属性分而治之进而再重新整合来完成的，但其机制尚不完全清楚。
4. 视觉加工中存在what 和where 两条视觉通路，每条通路遵循分布式层级加工的原则。
5. 对生物视觉系统的理解，有助于构建类脑人工视觉系统；反之，人工视觉系统的进步，也有助于理解生物视觉系统。

思考题

1、你认为视觉系统的主要功能是什么？

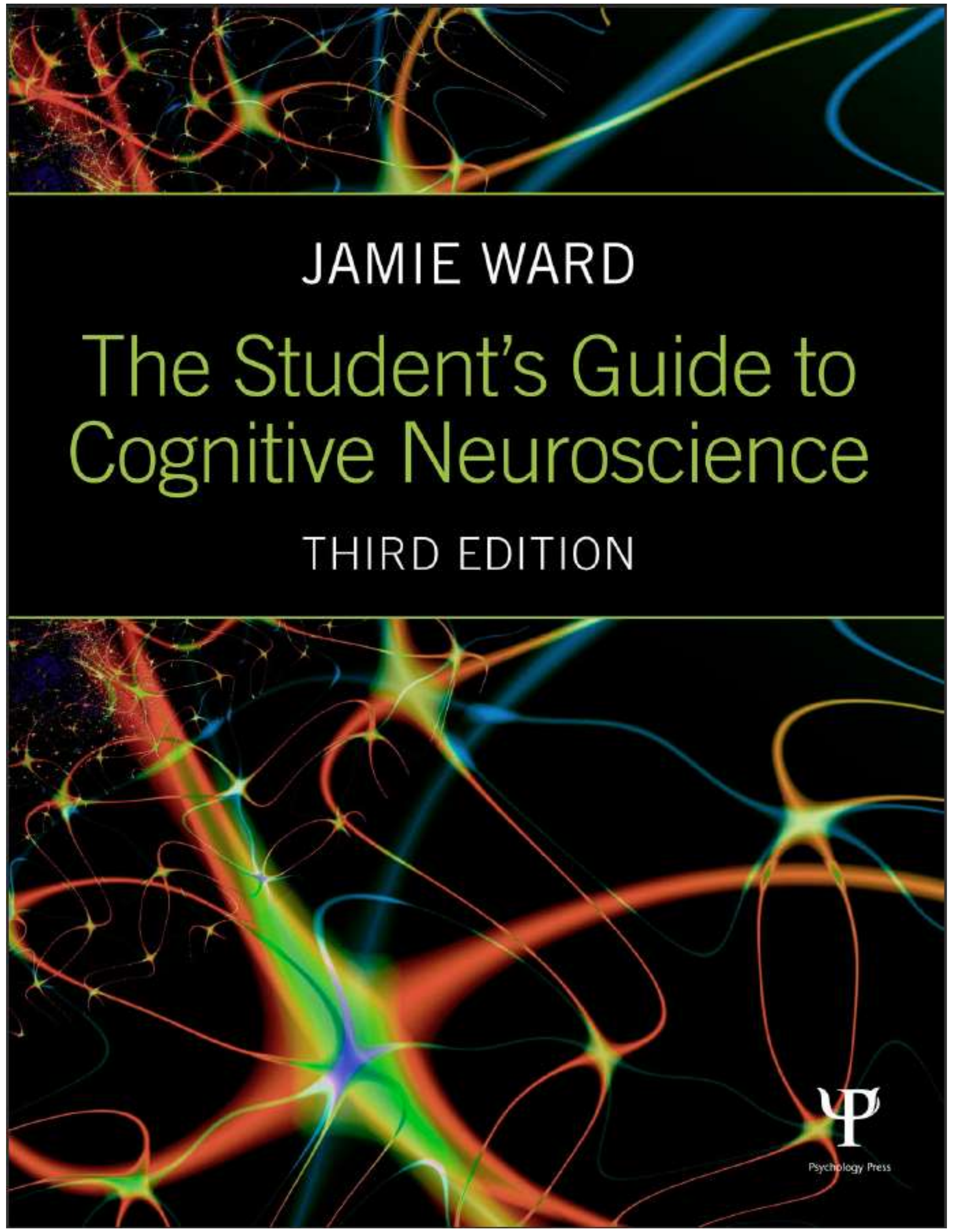
2、你认为目前视觉认知神经研究中，存在哪些不足？假设没有任何技术限制，你觉得应该怎么研究视觉系统？

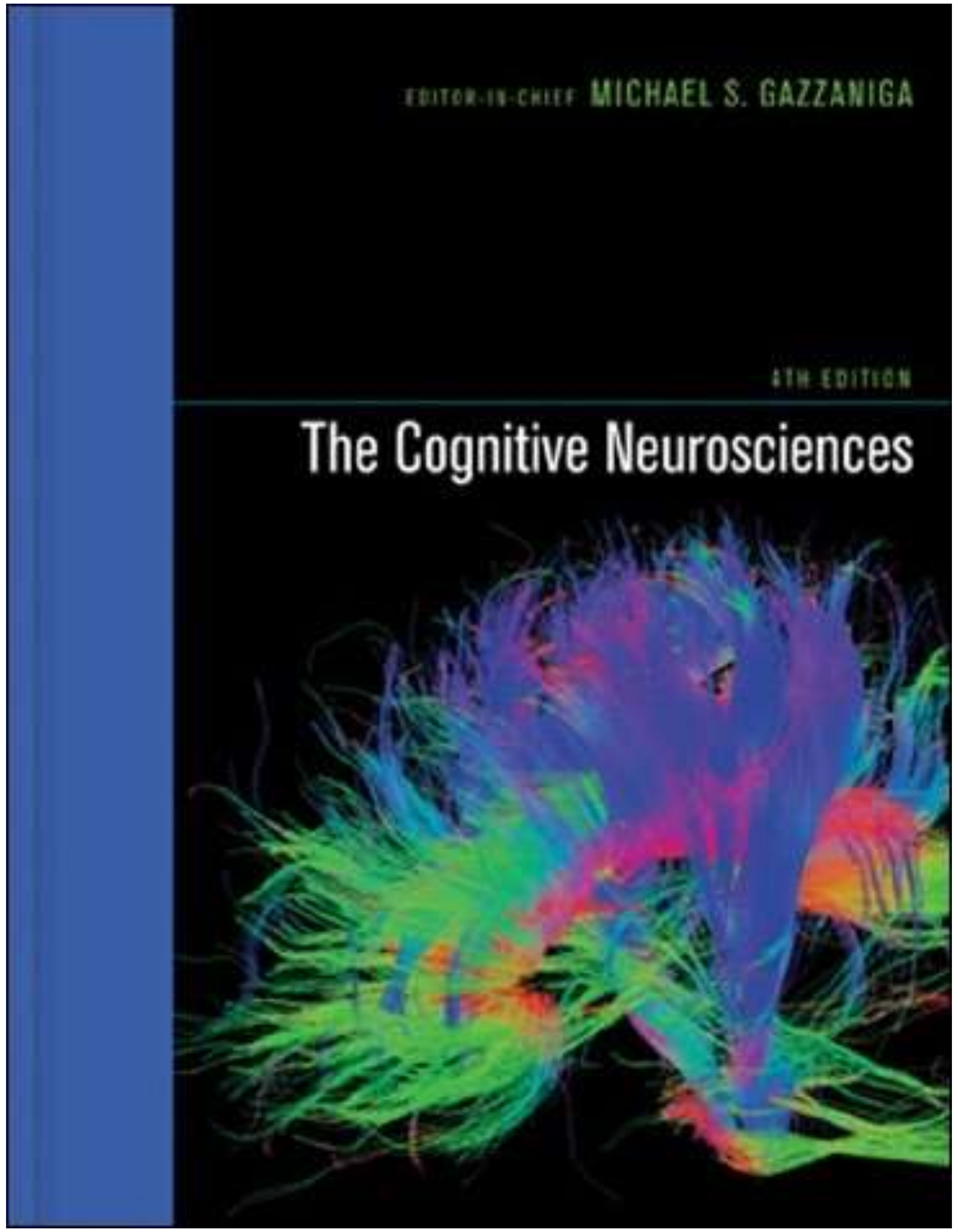
3、视网膜拓扑映射是低级视觉区的典型组织原则，高级视觉脑区是否仍服从这个原则么？

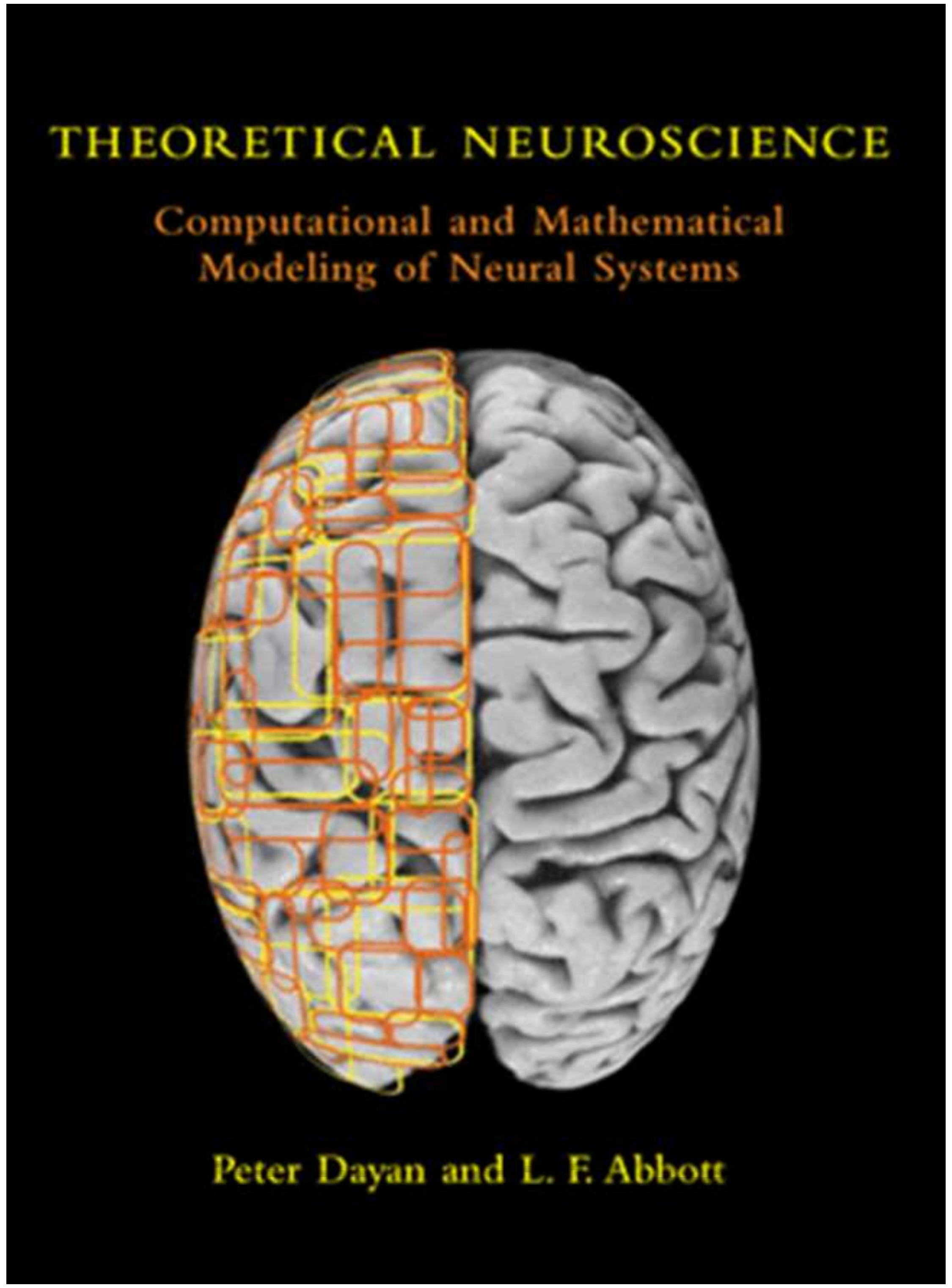
4、人脑视觉认知神经的研究，未来对人工智能可能会产生哪些帮助？

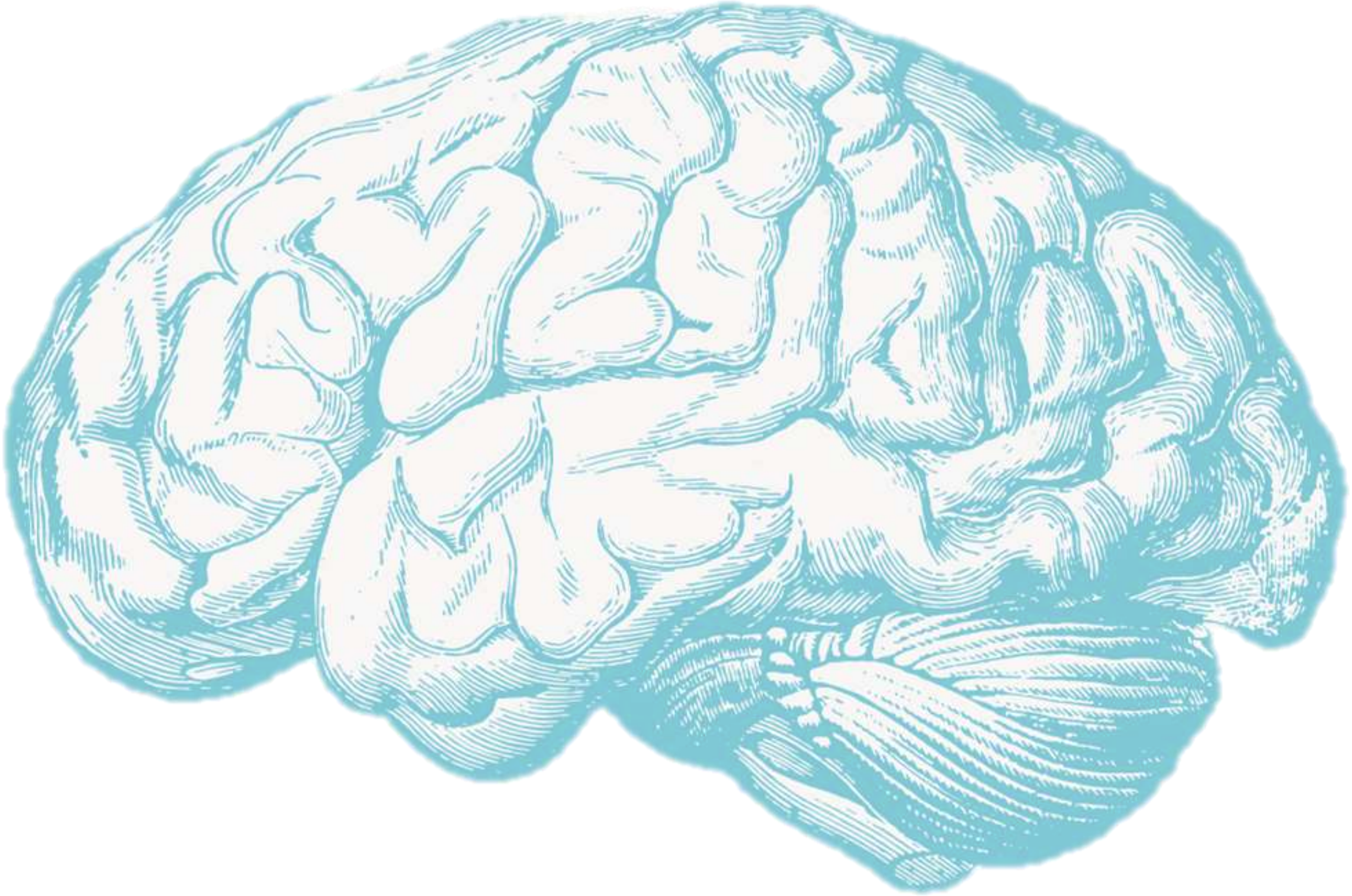
5、本章孤立地介绍了视觉系统，然而视觉输入是认知加工的起点，它和那些认知工功能相关？

推荐书籍


JAMIE WARD
The Student's Guide to Cognitive Neuroscience
THIRD EDITION
Psychology Press


EDITOR-IN-CHIEF MICHAEL S. GAZZANIGA
4TH EDITION
The Cognitive Neurosciences


THEORETICAL NEUROSCIENCE
Computational and Mathematical Modeling of Neural Systems
Peter Dayan and L. F. Abbott



【下一讲 】
脑与认知加工：
感知运动


Q&A