计算机视觉与模式识别

Computer Vision and Pattern Recognition

-- Physiology and Classic Theories of Vision

(Overview)

袁泽剑

人工智能与机器人研究所

Institute of Artificial Intelligence and Robotics

Email:yuan.ze.jian@xjtu.edu.cn

科学馆102室

What is (computer) vision?

- What does it mean, to see? The plain man's answer (and Aristotle's, too). would be, to know **what** is **where** by looking. -- David Marr, Vision (1982)

- **Visual Perception:** The process of acquiring knowledge about environmental objects and events by extracting information from the light they emit or reflect. -- S.E. Palmer, Vision Science (1999).

- **Computer Vision**: The study of how computers can be programmed to extract useful information about the environment from optical images --S.E. Palmer, Vision Science (1999)

计算机能跟人类视觉感知相媲美吗？

- ➢在特定领域（视觉精细计算）
  - ✓图像处理（恢复/编解码）
  - ✓基于几何的重建（拓扑、射影、仿射、欧氏）
  - ✓精细测量\视觉伺服(结构化环境)
  - ✓计算机图形\可视化
- ➢在深度认知与高层视觉信息利用方面
  - ✓目标识别\场景解析与语义描述
  - ✓上下文关系\多层异构信息集成与推理
  - ✓视觉信息的层次抽象、表示与转换

关键因素

- ✓ **Reality and Illusion / Measurement vs. Perception**
- ✓ **Local vs. global information**
- ✓ **Perception could be ambiguous**
- ✓ **Vision is ill-posed**
- ✓ **Visual Representations & Semantic Concept**

- **What do humans care about?**
  - **Verification**
  - **Detection**
  - **Identification**
  - **Object categorization**
  - **Scene and context categorization**
  - **Rough 3D layout, depth ordering**

- The world behind the image

Challenges for View-based Representation

- View point variation/ illumination/ occlusion/ scale
- Deformation/ object intra-class variation
- Background clutter
- Complexity

提纲

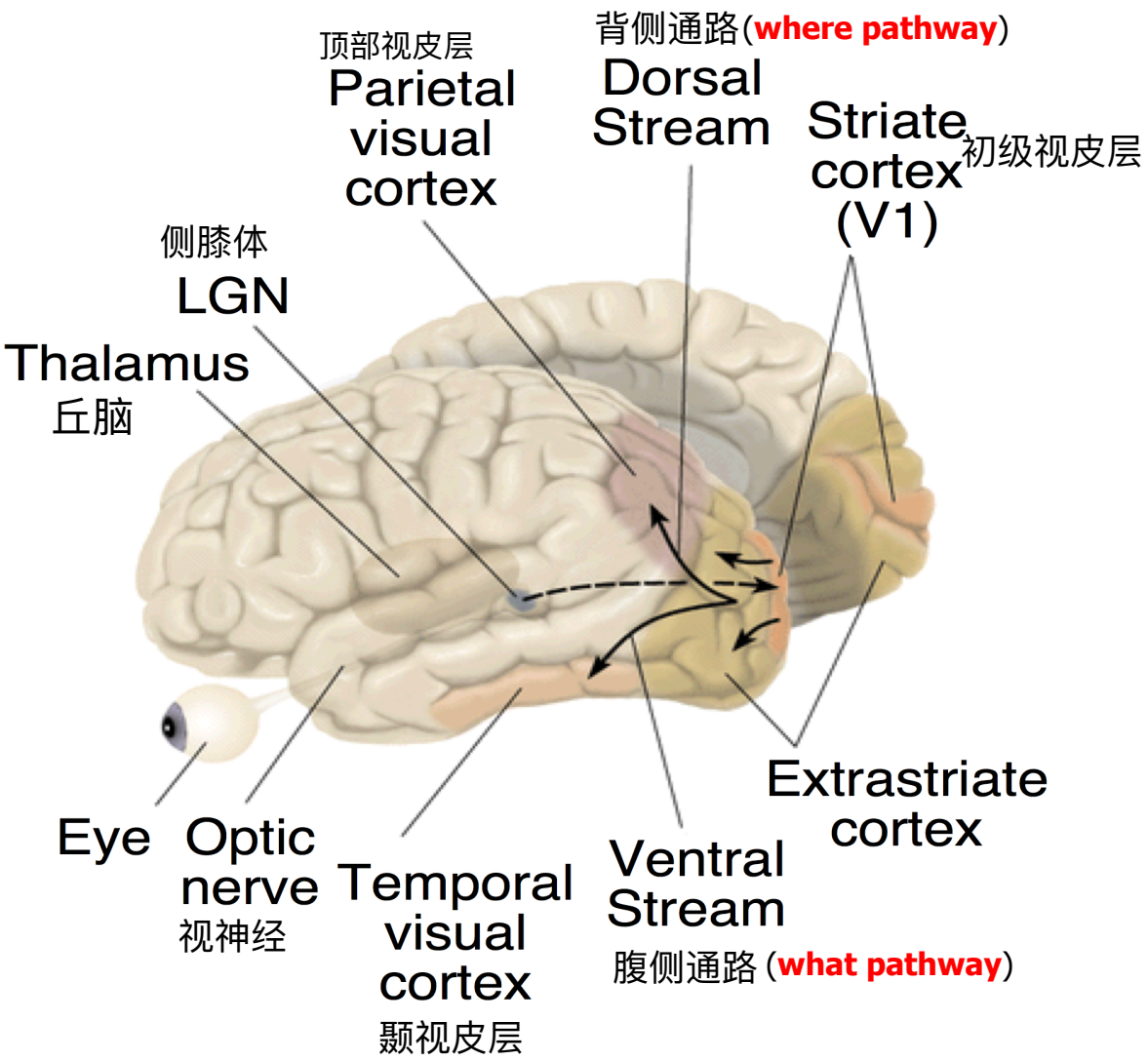
- 视觉生理学简介
- 视觉的经典理论
- 主要的参考文献

[1] D. Marr, **Vision**----A computational investigation into the human representation and processing of visual information, W.H. Freeman and Company San Francisco, 1982.

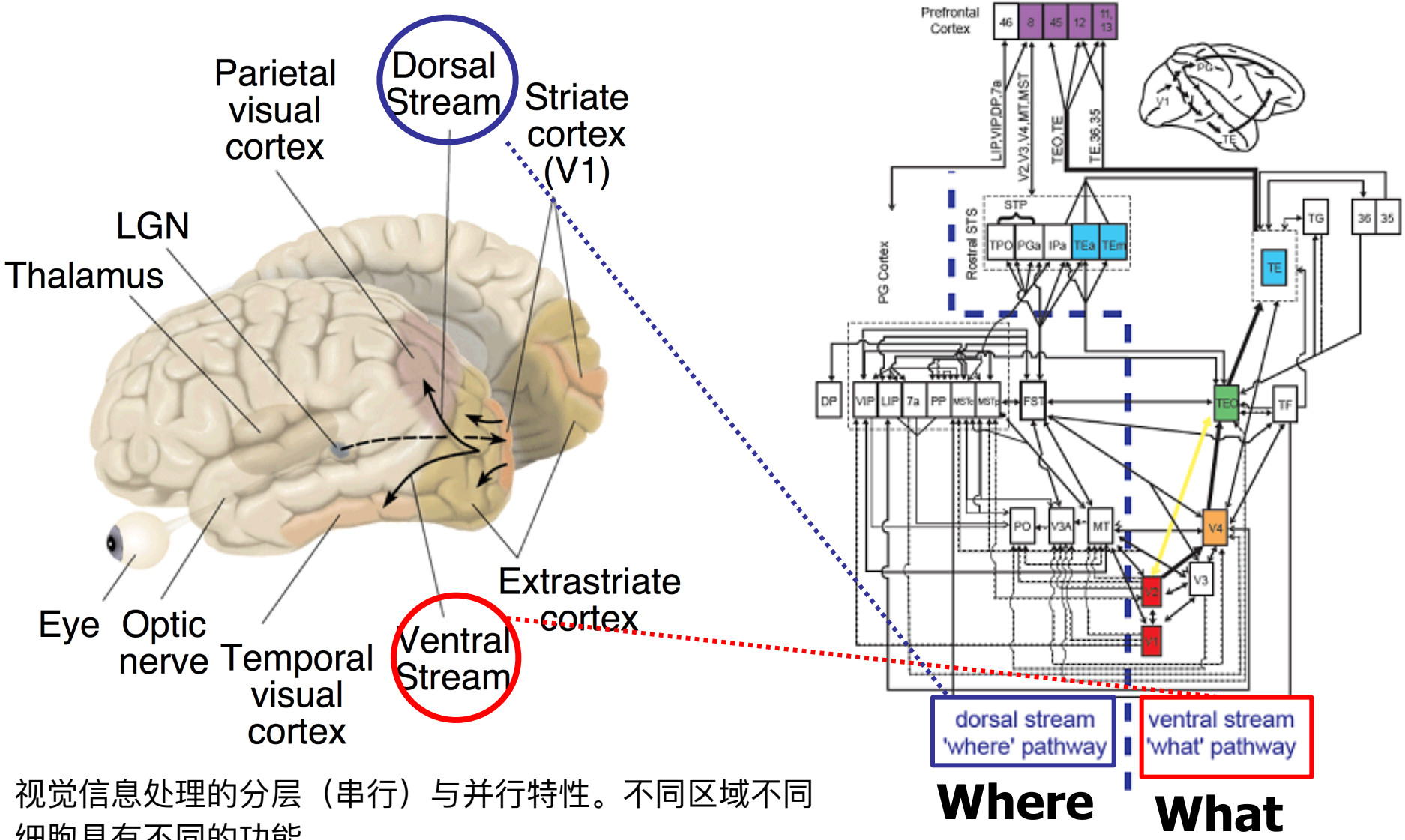
[2] S.E. Palmer, **Vision Science** – Photons to Phenomenology, the MIT Press, 1999.

Visual Physiology

- Human Visual System （ HVS ）



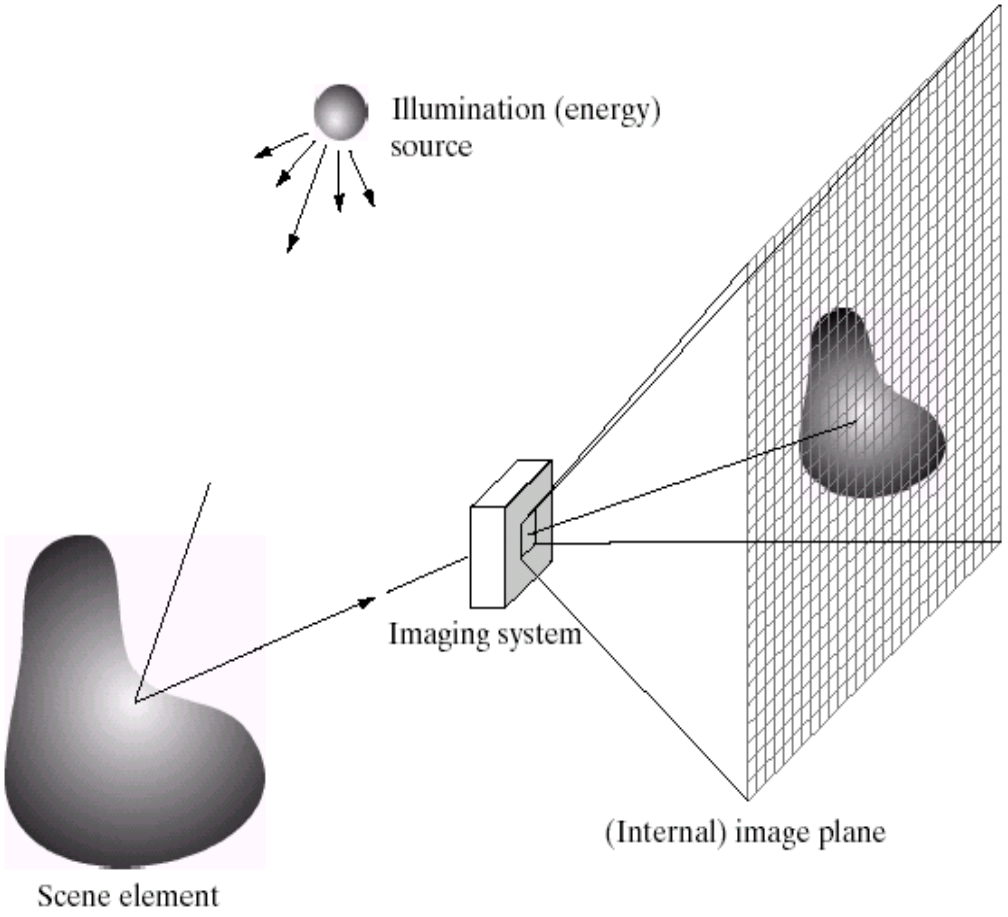
Human Visual System

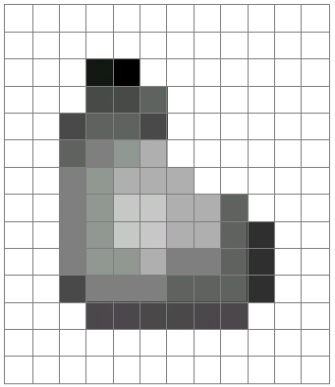


视觉信息处理的分层（串行）与并行特性。不同区域不同细胞具有不同的功能。

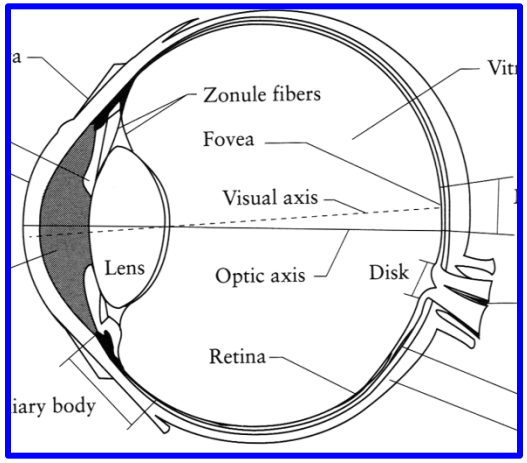
Visual Physiology

- Imaging





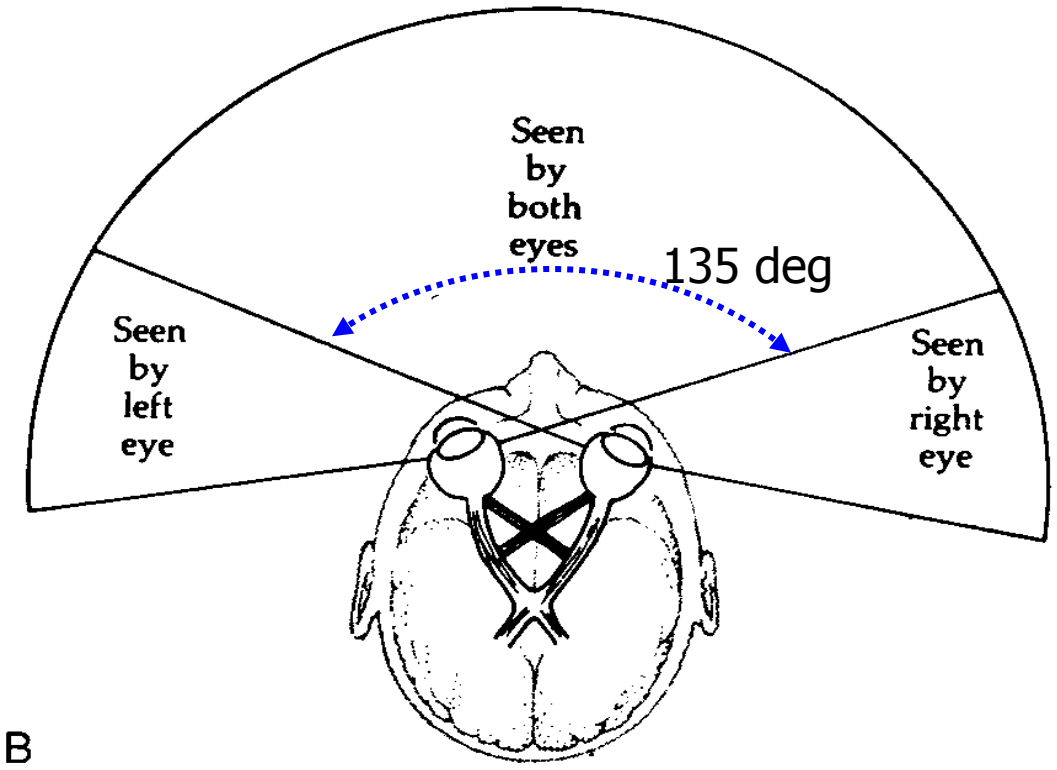
Digital Camera



The Eye

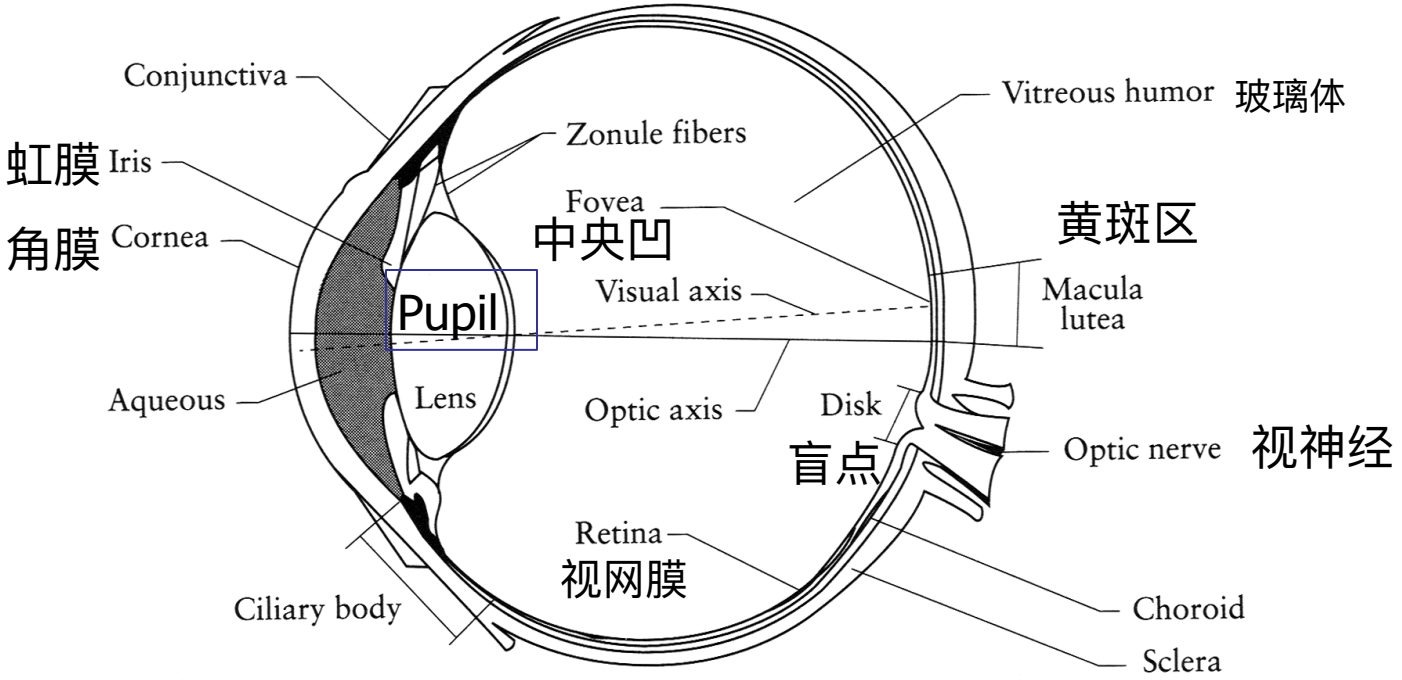
Visual Physiology

- **Visual Field**
  - Monocular Visual Field: 160 deg (H) X 135 deg (V)
  - Binocular Visual Field: 200 deg (H) X 135 deg (V)



Visual Physiology

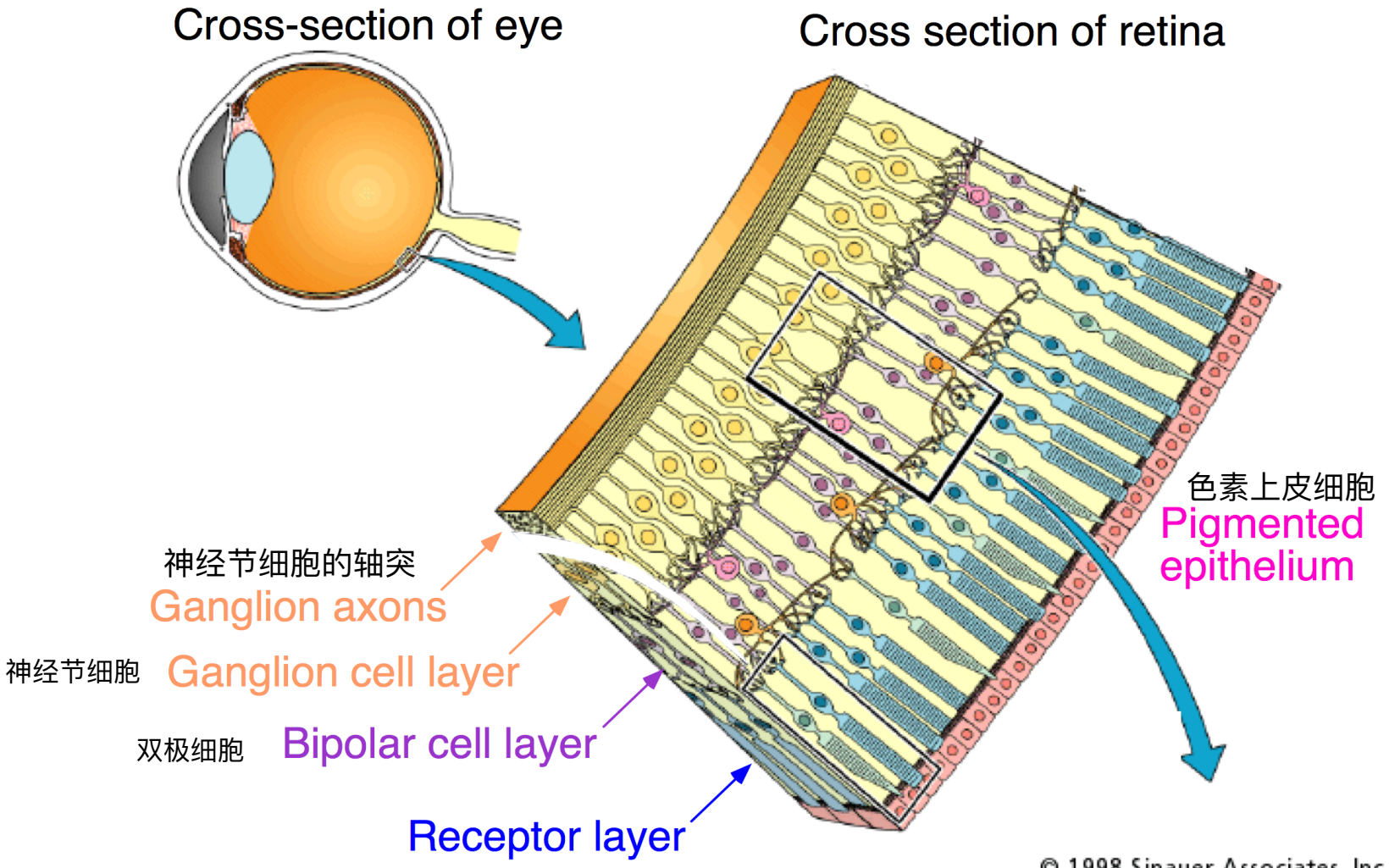
- The human eye (Camera)



- **Iris** - colored annulus with radial muscles
- **Pupil** - the hole (aperture) whose size is controlled by the iris
- What's the "film"?
  - **Photoreceptor** cells (rods and cones) in the **retina**

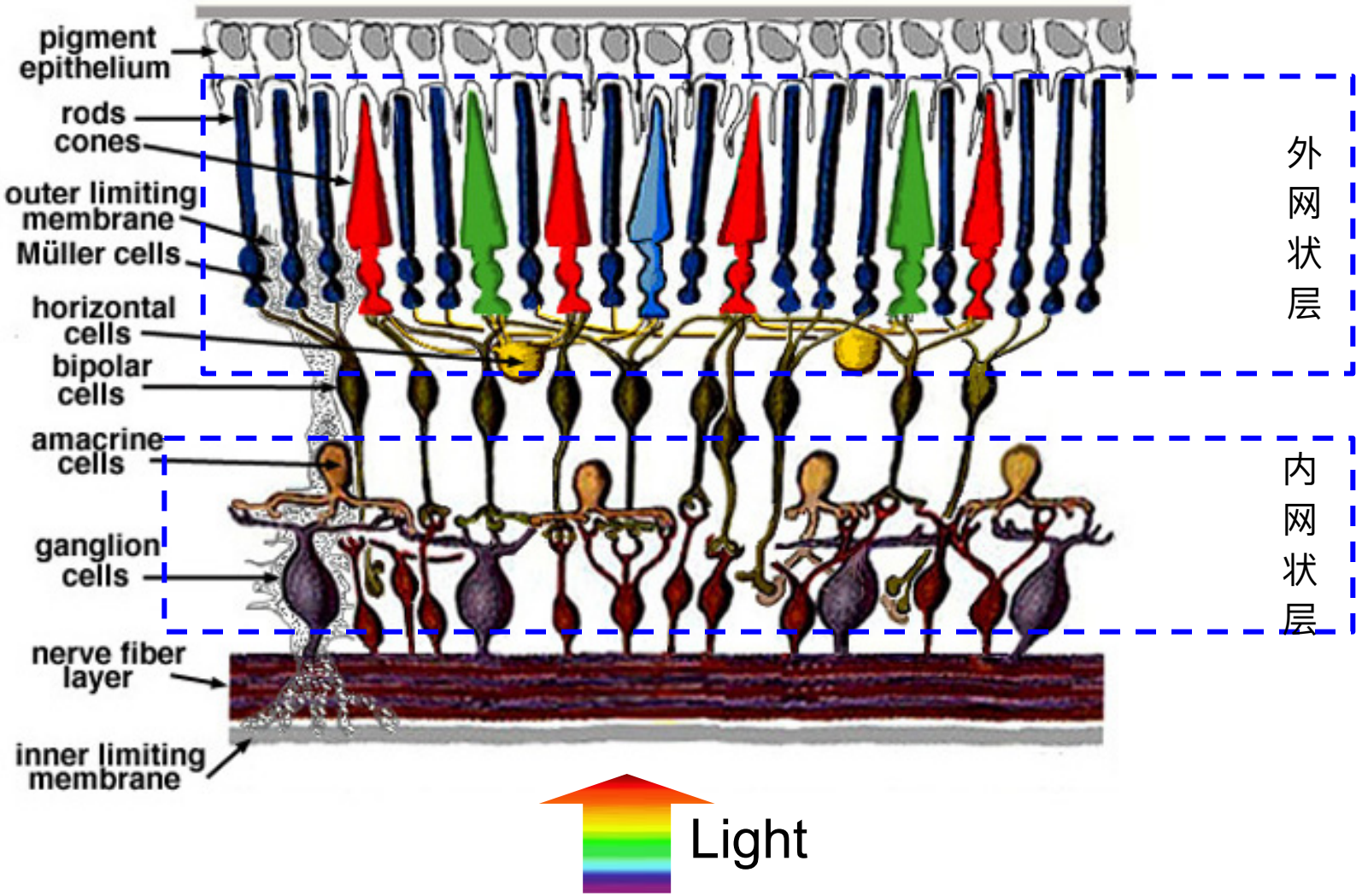
Visual Physiology

- **The Retina**



Visual Physiology

- **Retina**



Visual Physiology

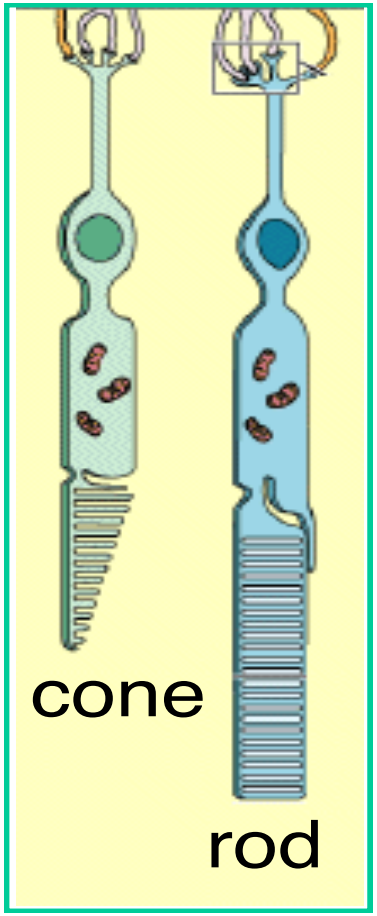
- **Two types of light-sensitive receptors**

Cones

- ✓ cone-shaped
- ✓ less sensitive
- ✓ operate in high light
- ✓ color vision

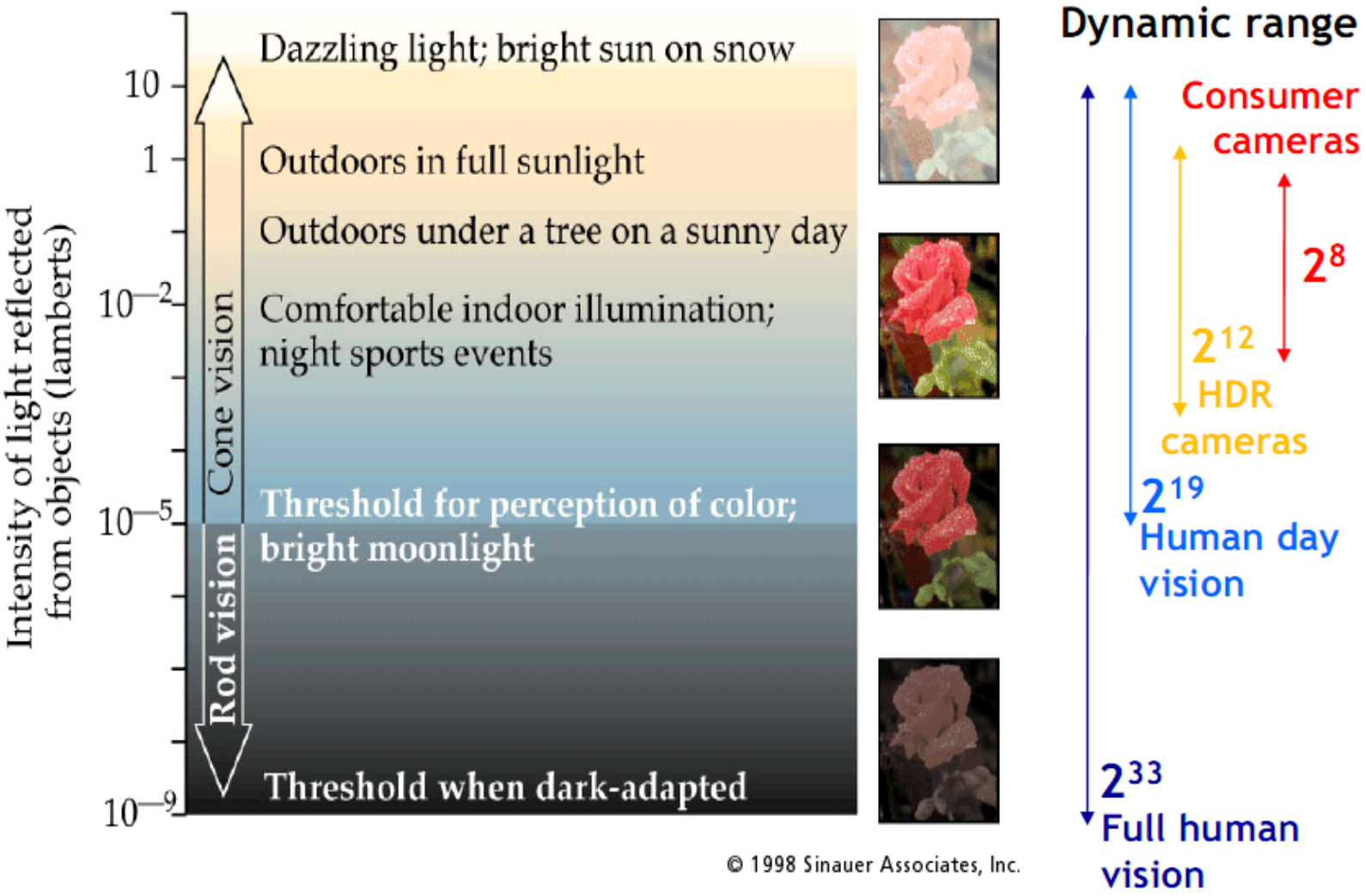
Rods

- ✓ rod-shaped
- ✓ highly sensitive
- ✓ operate at night
- ✓ gray-scale vision



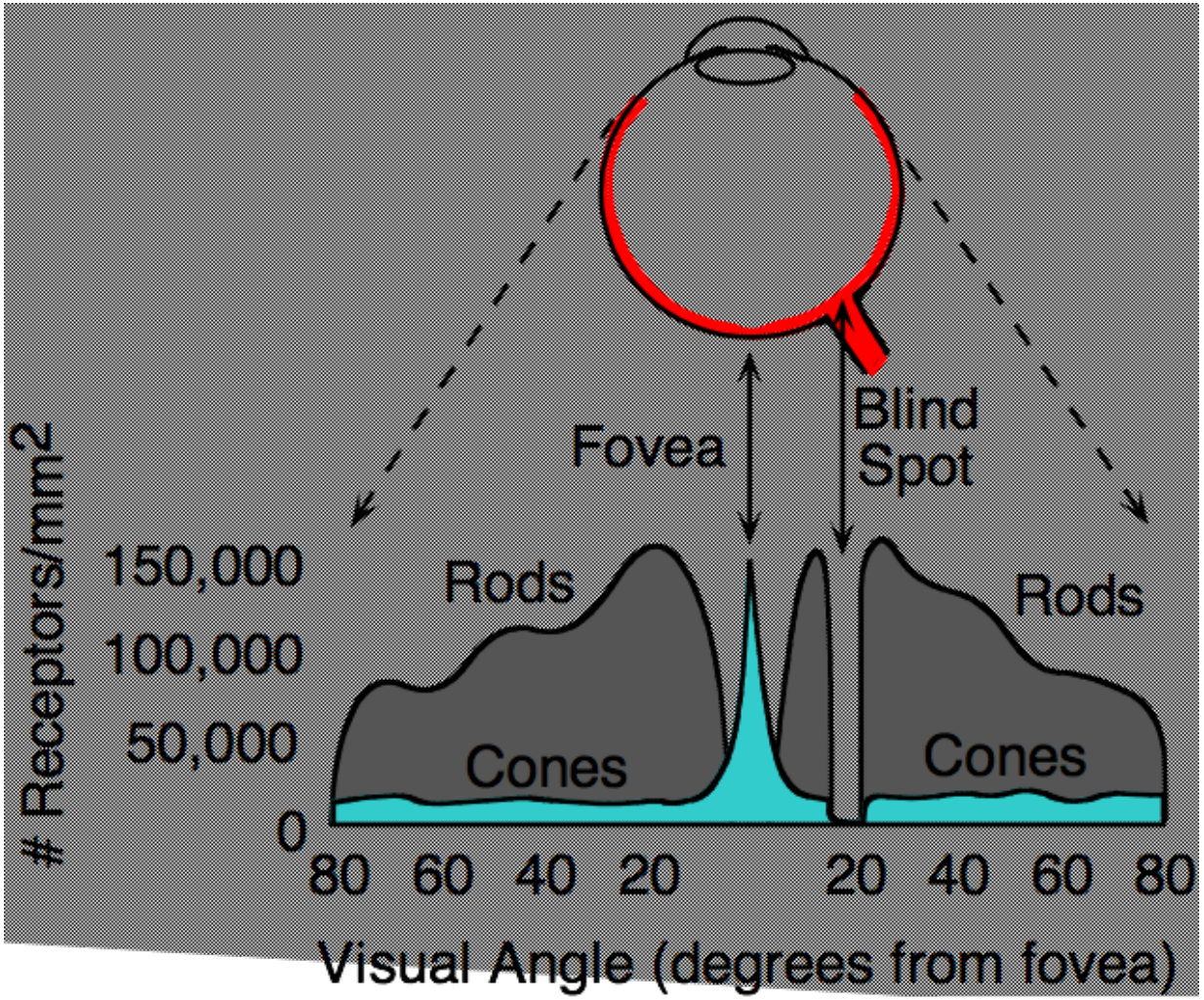
Visual Physiology

- **Rod / Cone sensitivity**



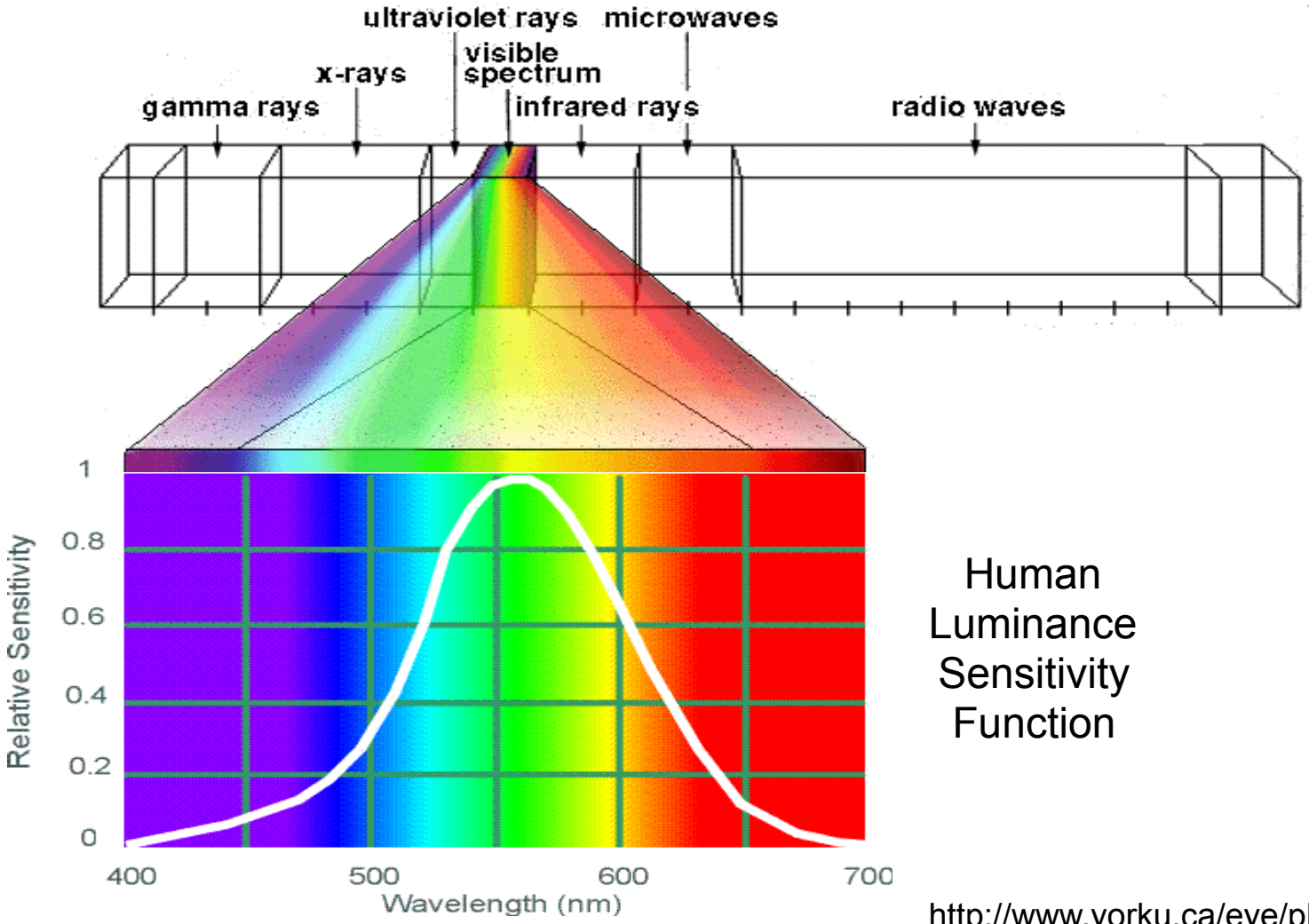
Visual Physiology

- **Distribution of Rods and Cones**



Visual Physiology

- **Electromagnetic Spectrum**



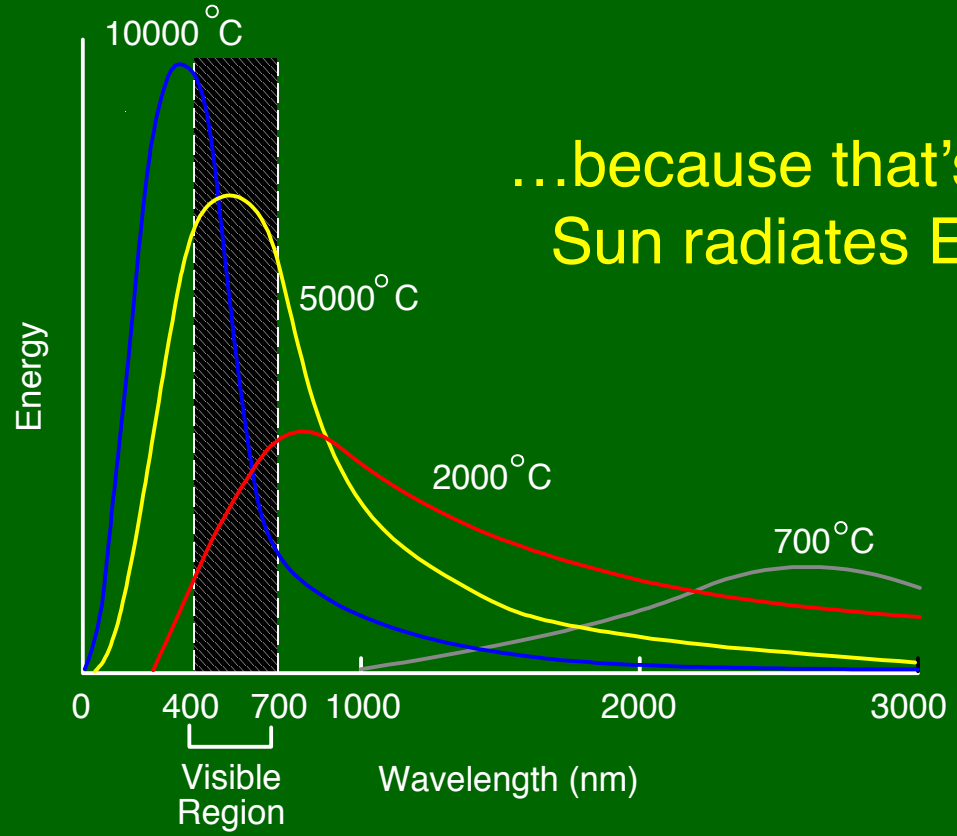
Human Luminance Sensitivity Function

http://www.yorku.ca/eye/photopik.htm

Visual Physiology

- **Visible Light**

Why do we see light of these wavelengths?

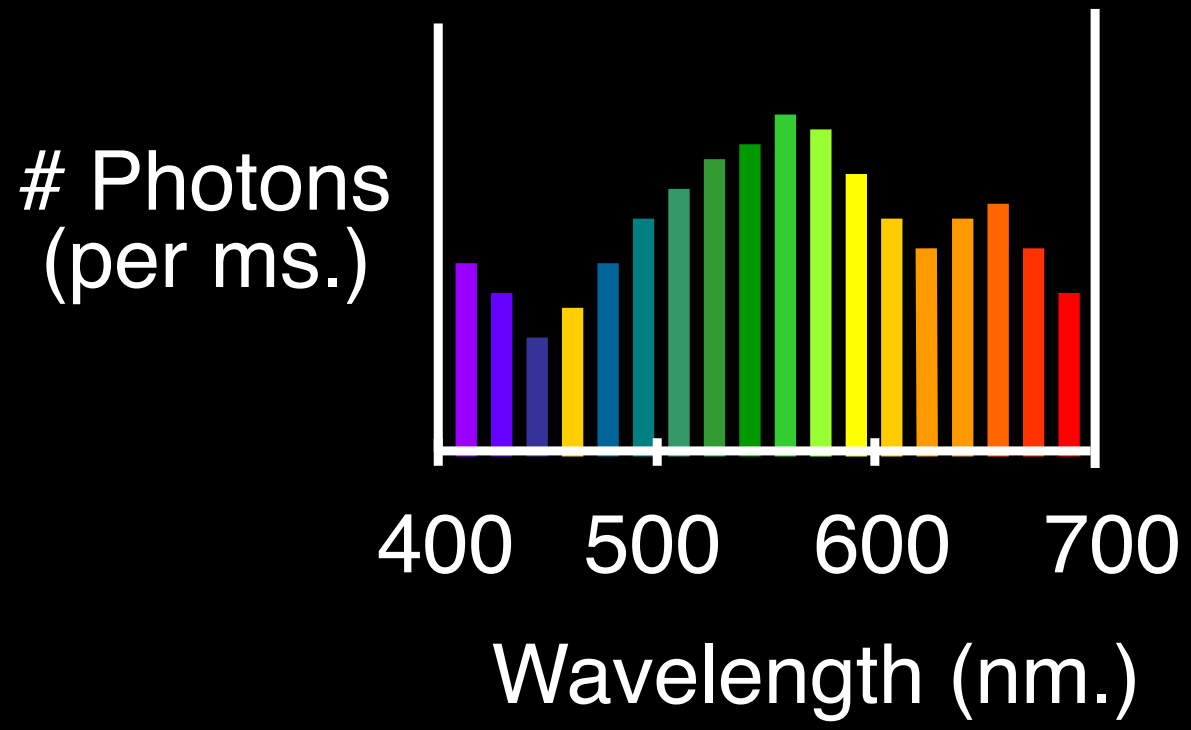


…because that's where the Sun radiates EM energy

Visual Physiology

- **The Physics of Light**

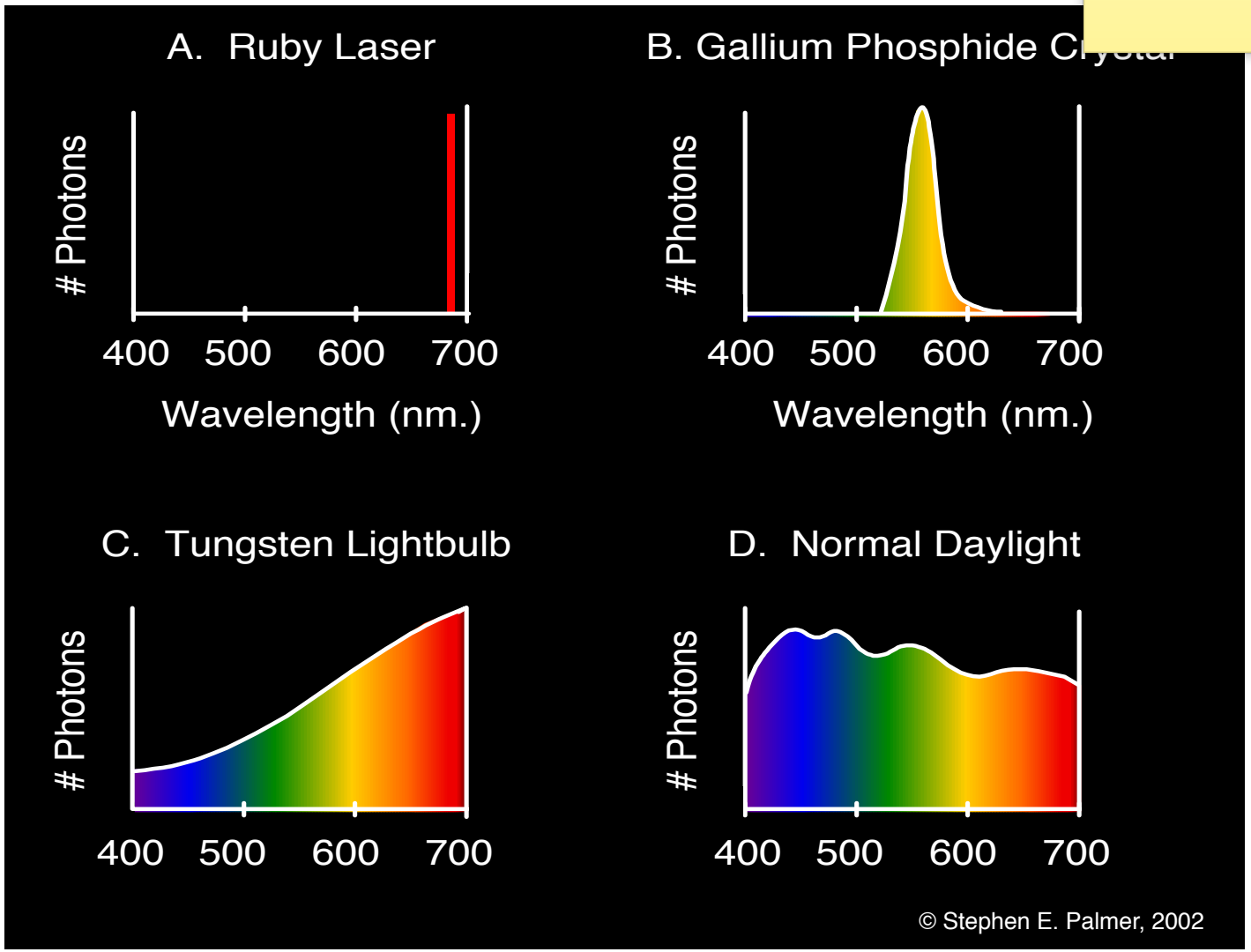
Any patch of light can be completely described physically by its spectrum: the number of photons (per time unit) at each wavelength 400 - 700 nm.



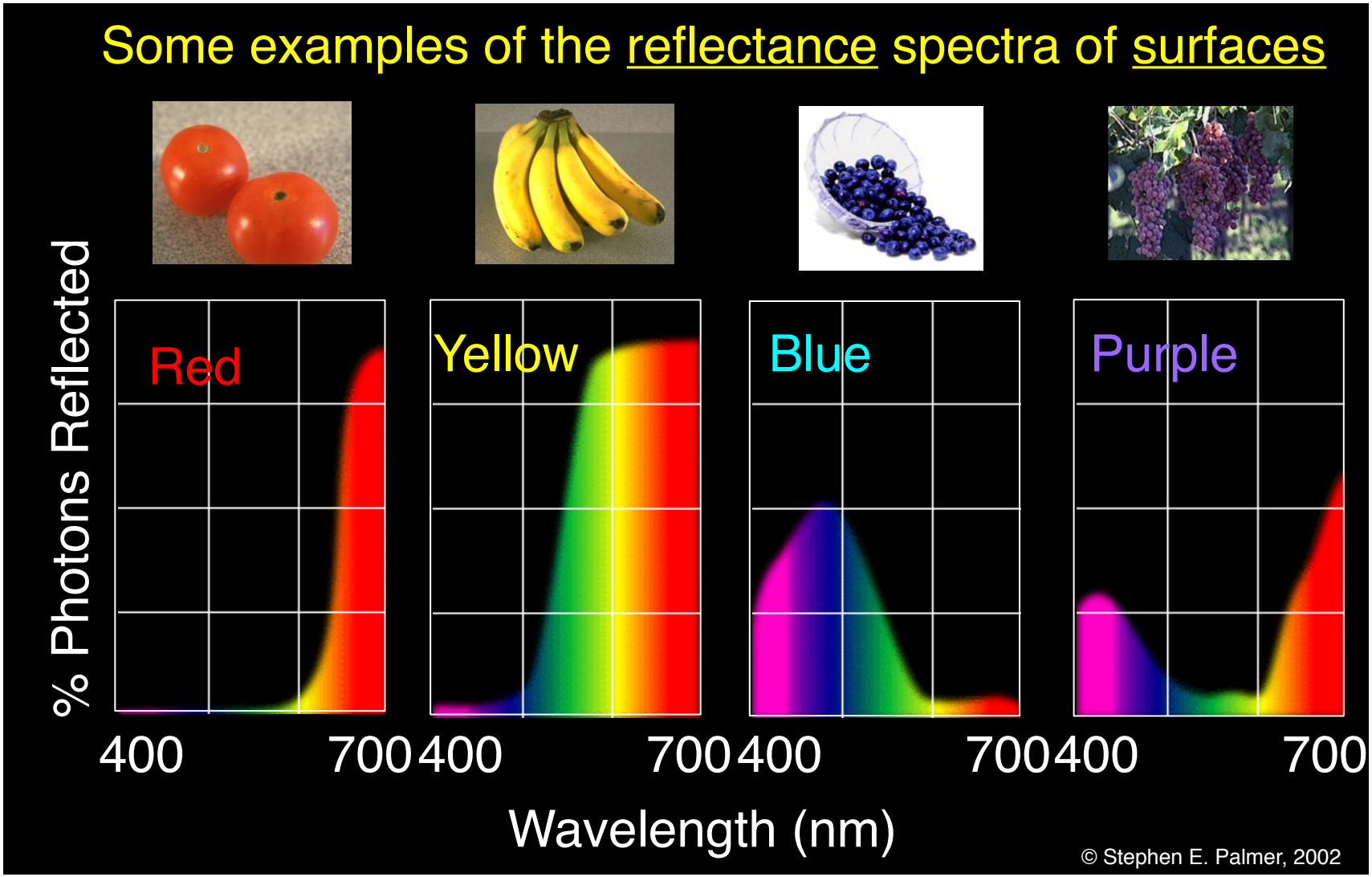
Visual Physiology

红宝石
磷化镓晶体
钨丝灯泡

- **The spectra of light sources**

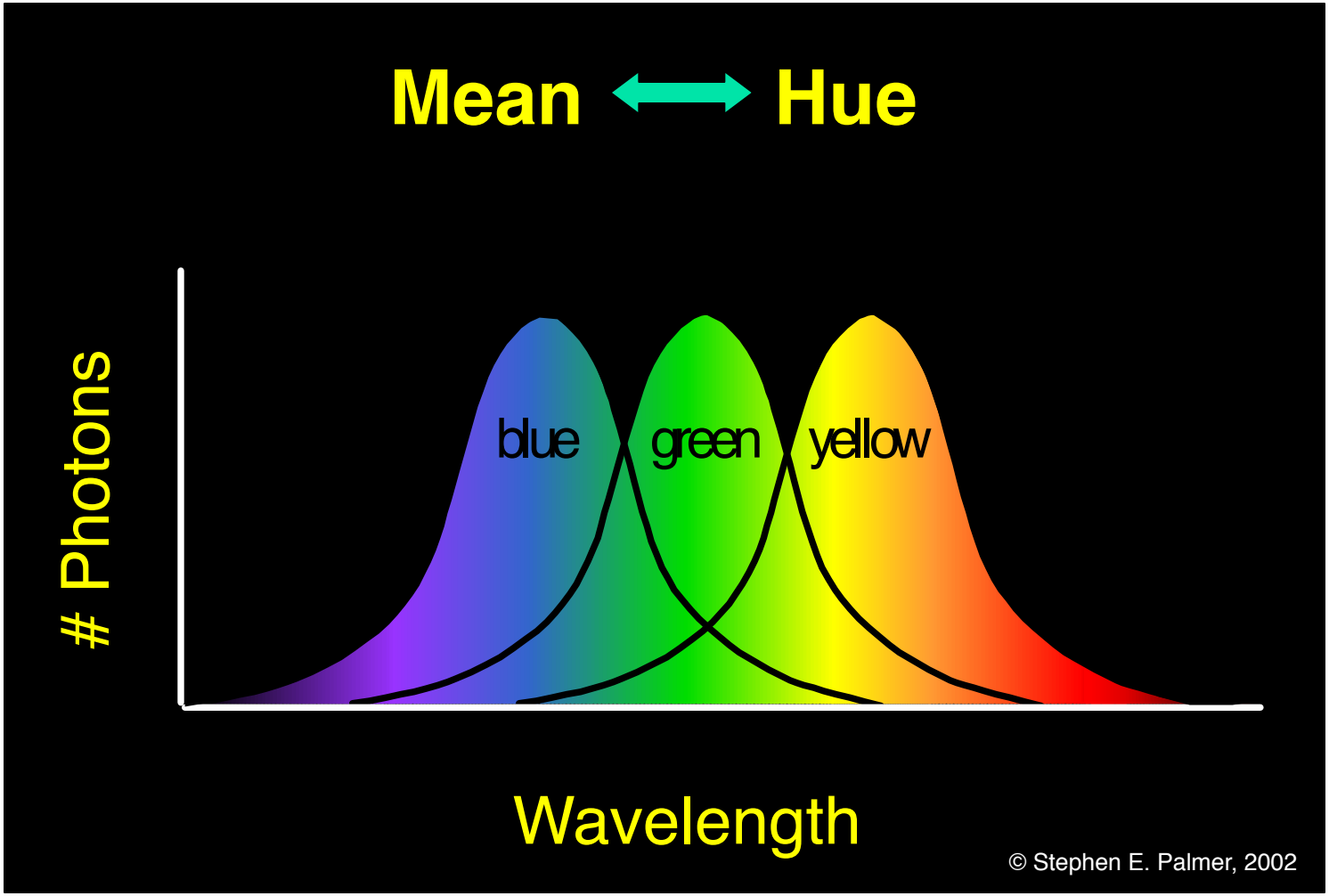


Visual Physiology



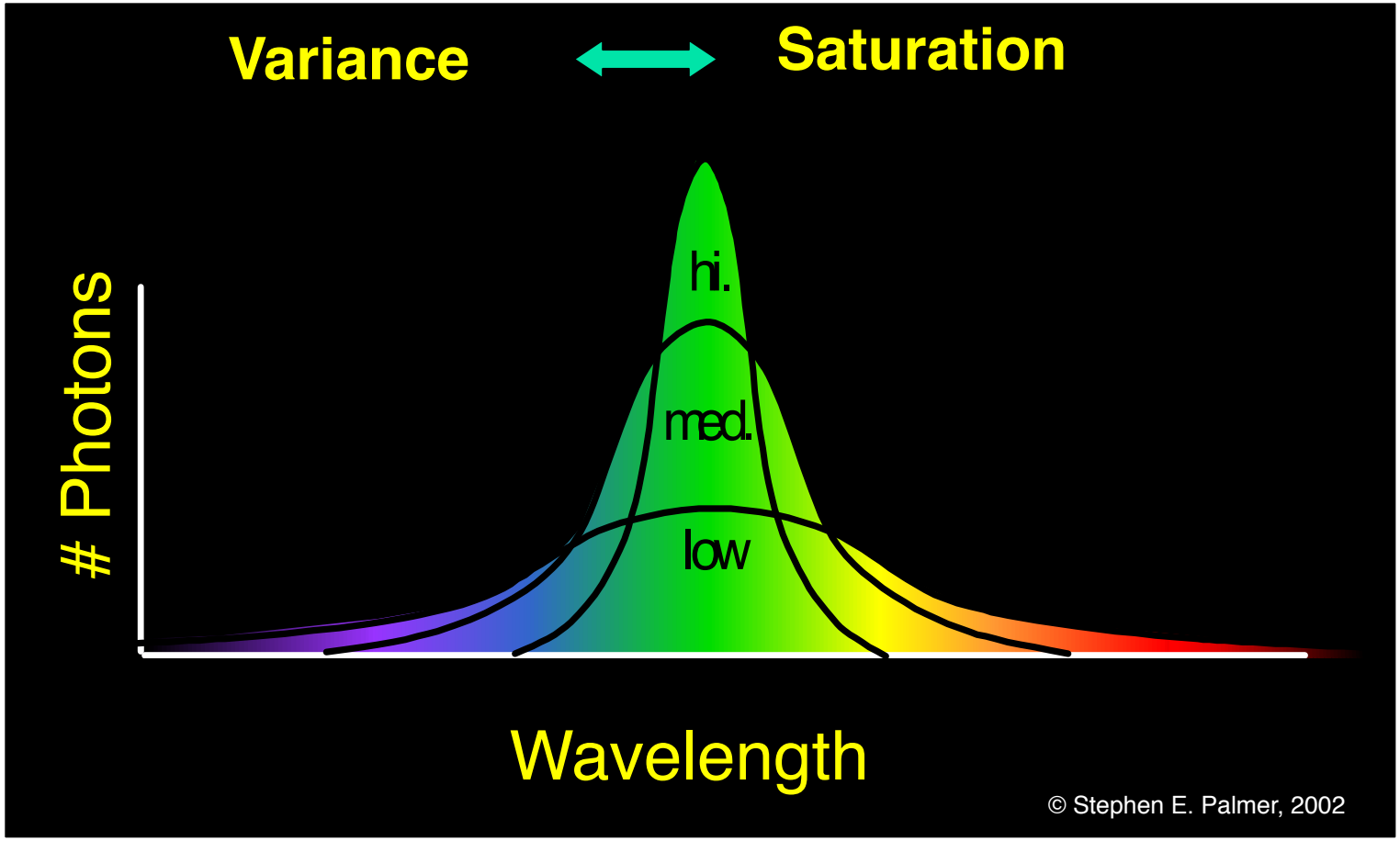
Visual Physiology

- **The Psychophysical Correspondence**



Visual Physiology

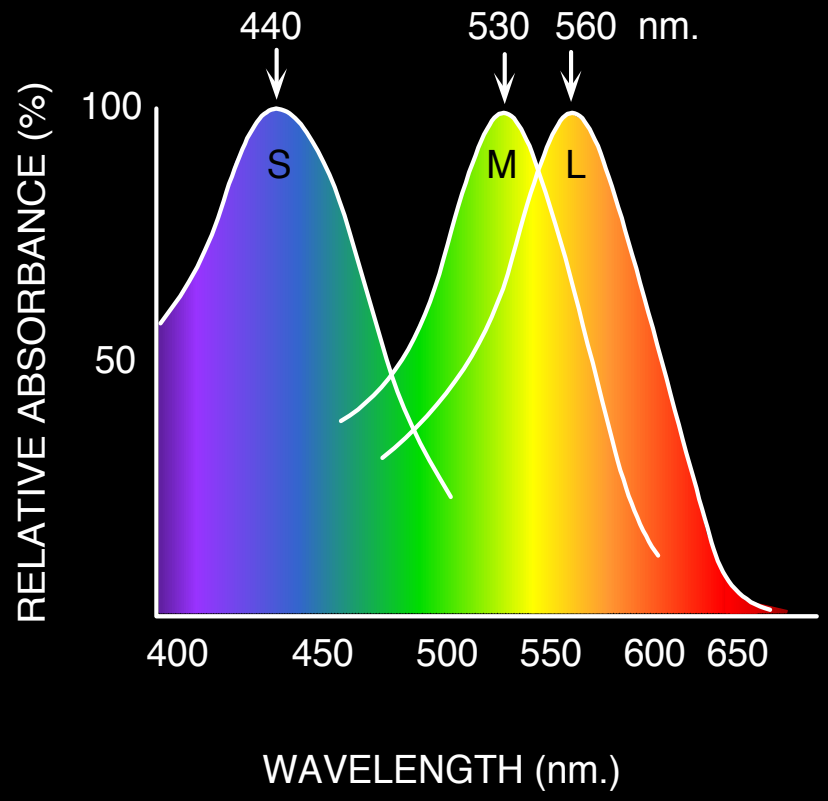
- **The Psychophysical Correspondence**

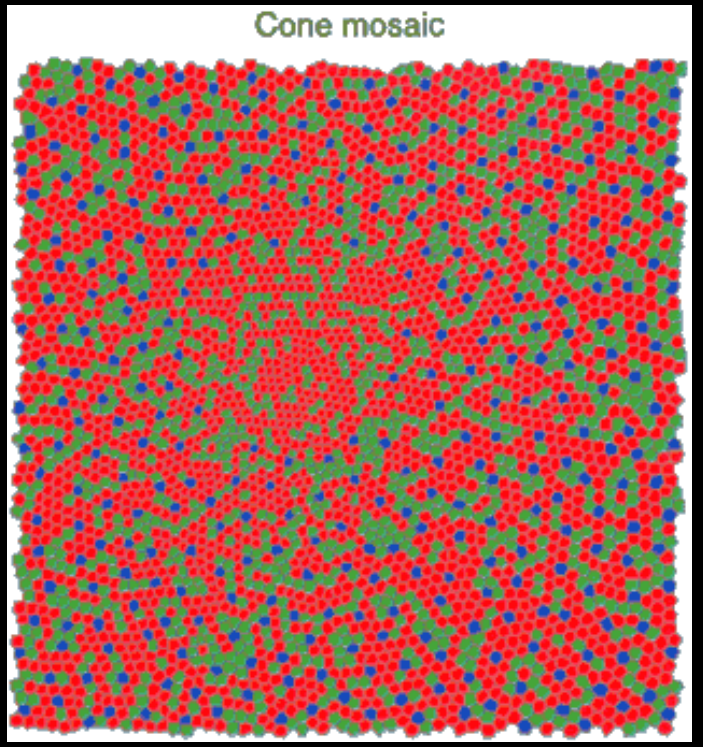


Visual Physiology

Physiology of Color Vision

Three kinds of cones:

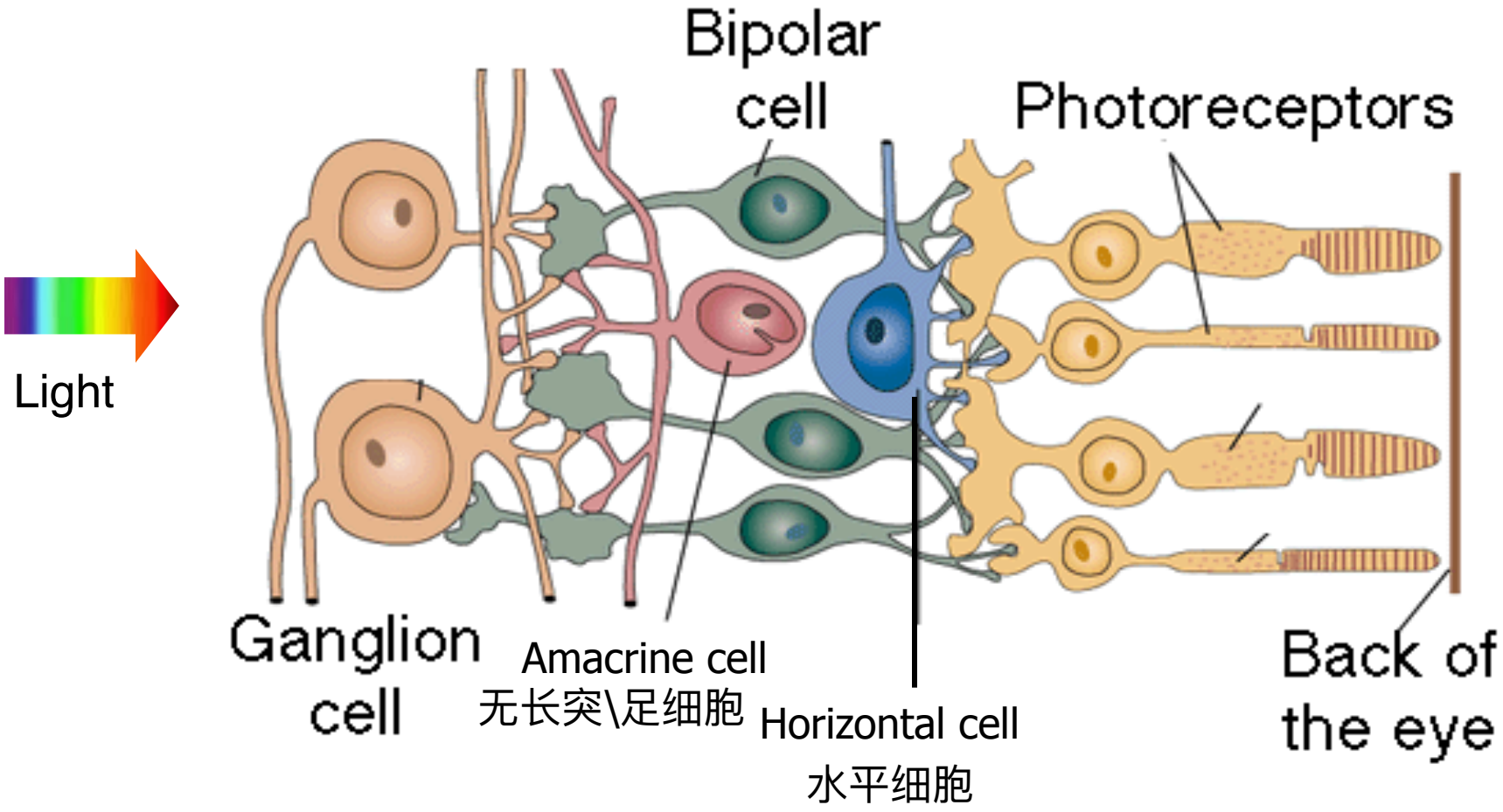




- Ratio of L to M to S cones: approx. 10:5:1
- Almost no S cones in the center of the fovea

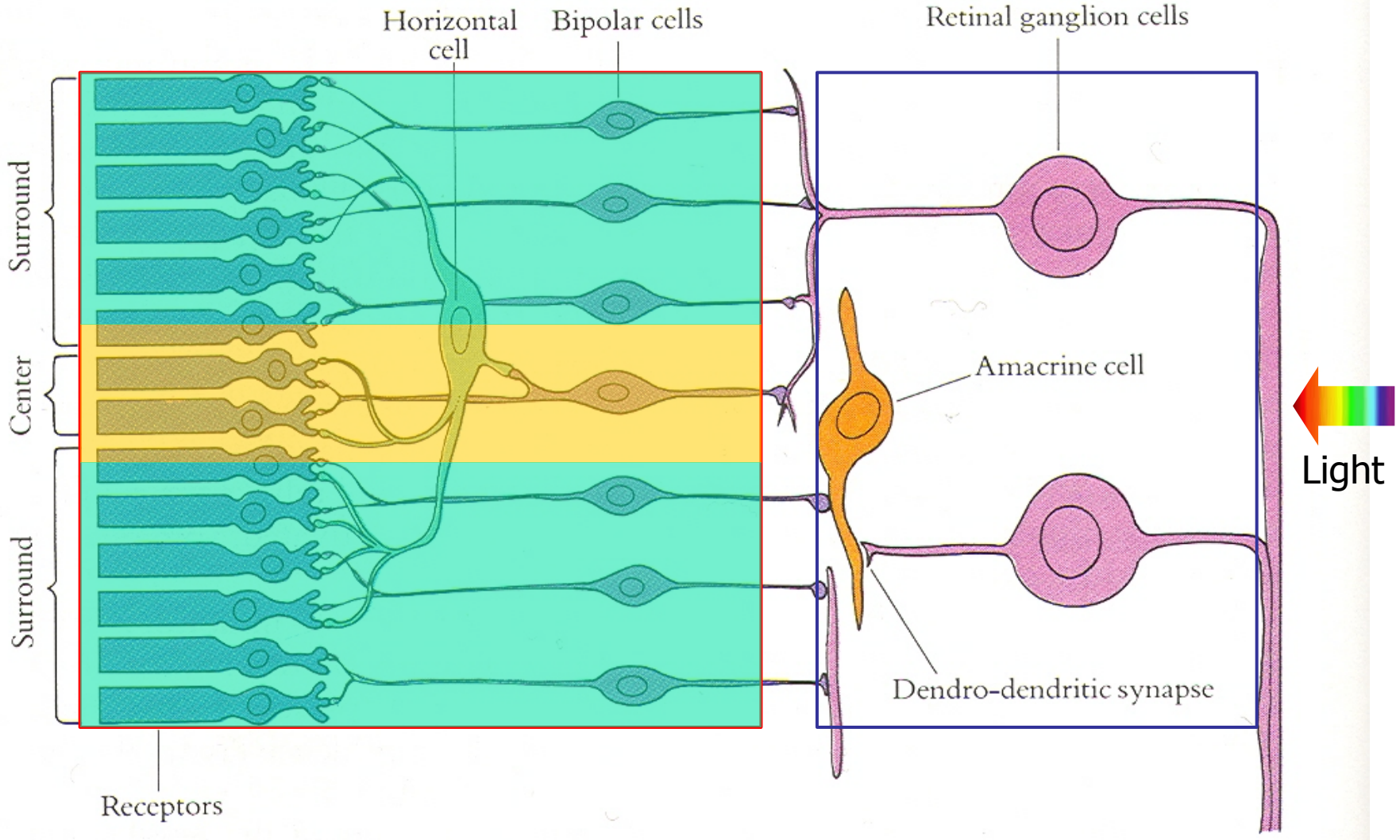
Visual Physiology

- **Retinal Processing**



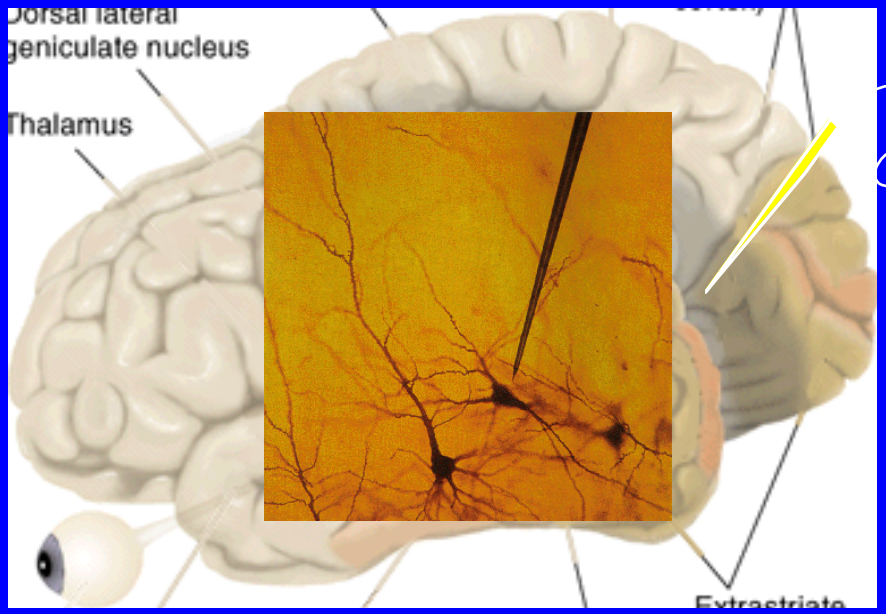
Visual Physiology

- **Retinal Receptive Fields**



Visual Physiology

- **Single Cell Recording**



Microelectrode

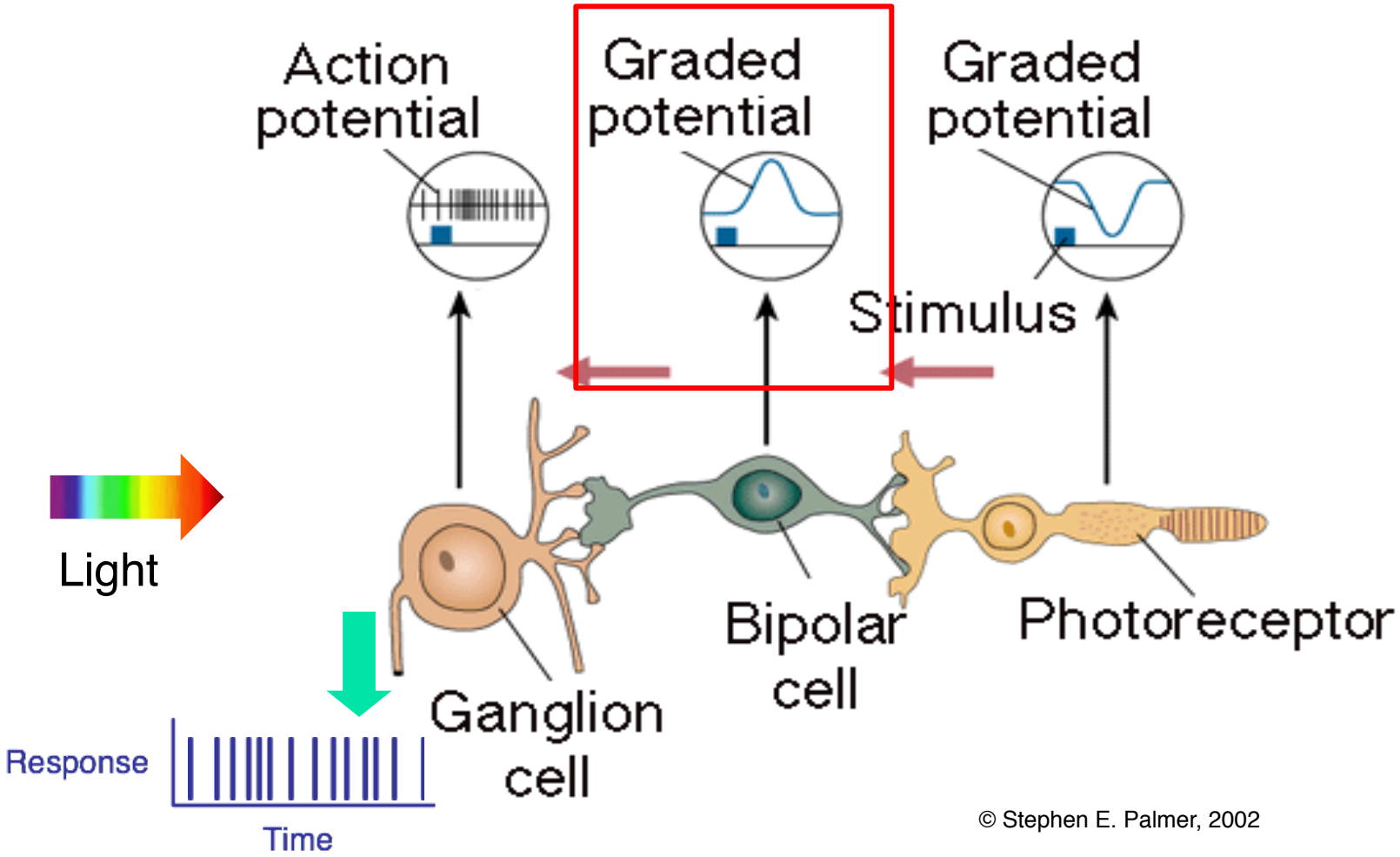
Amplifier

Electrical response
(action potentials)

mV

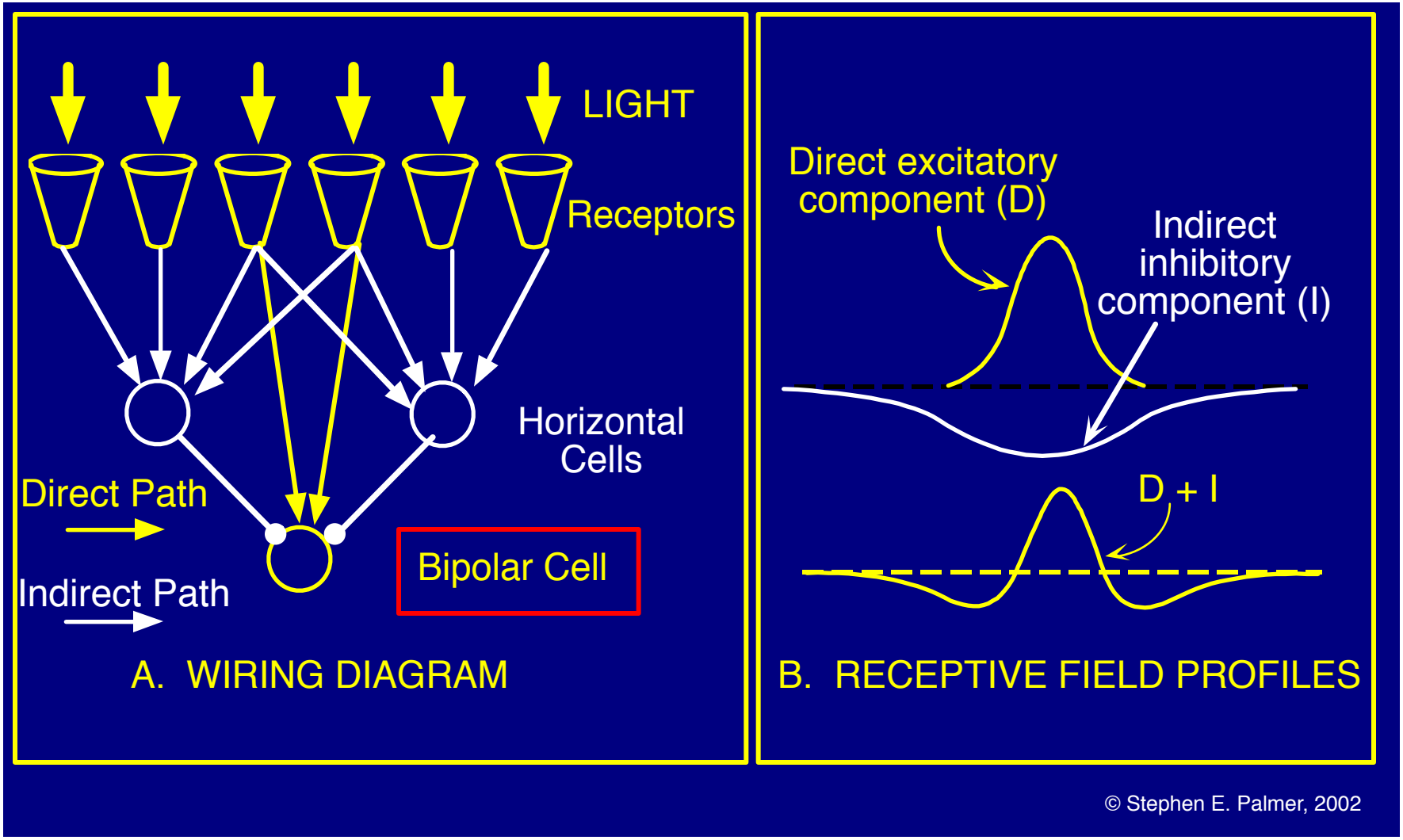
Visual Physiology

- **Receptive field structure in Bipolar Cells**



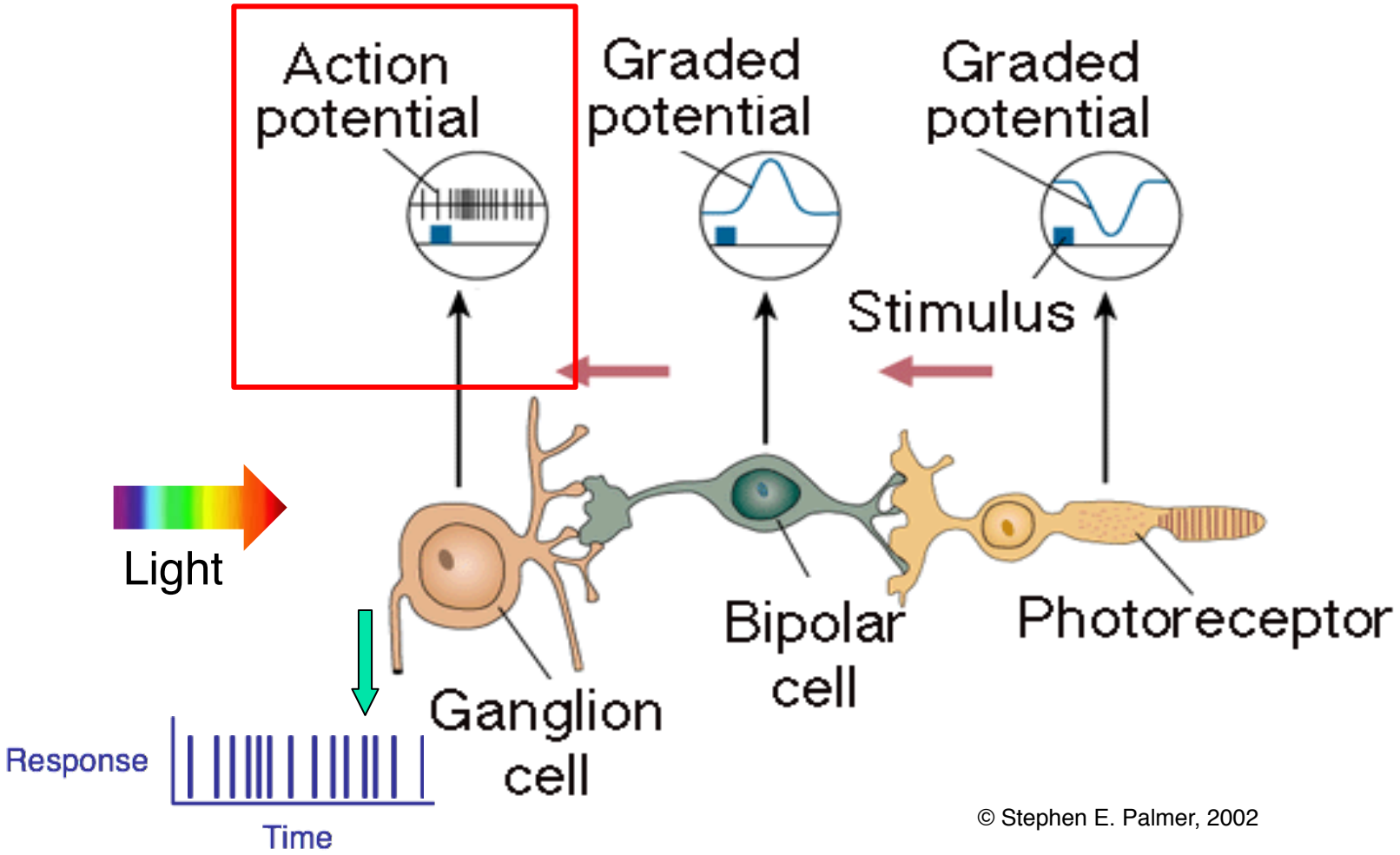
Visual Physiology

- Receptive field structure in **Bipolar Cells**



Visual Physiology

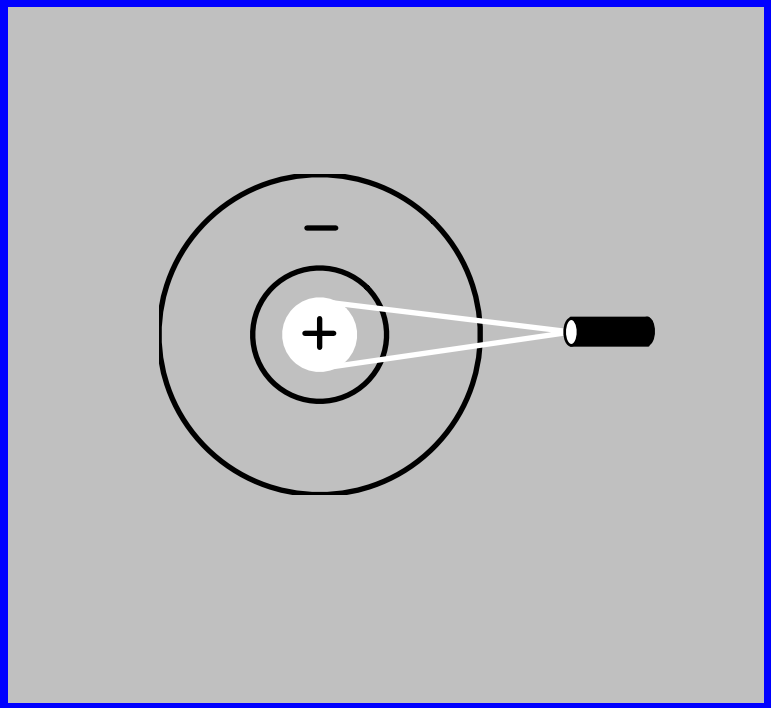
· Receptive field structure in Ganglion Cells

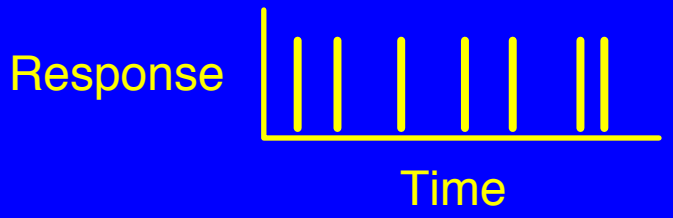


Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround





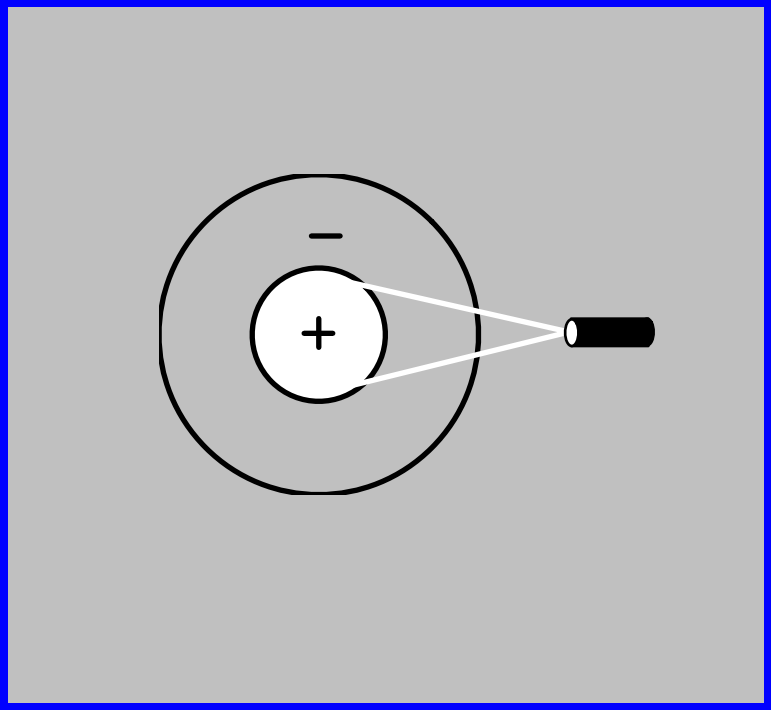
Stimulus condition

Electrical response

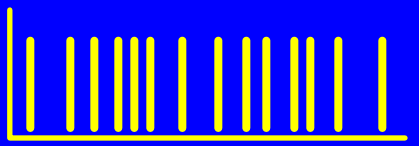
Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround



Response



Time

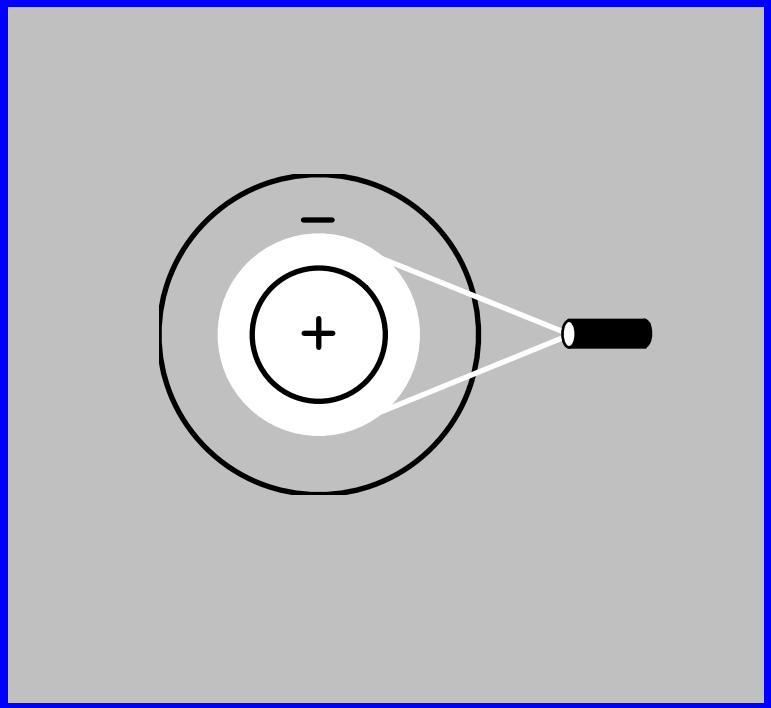
Stimulus condition

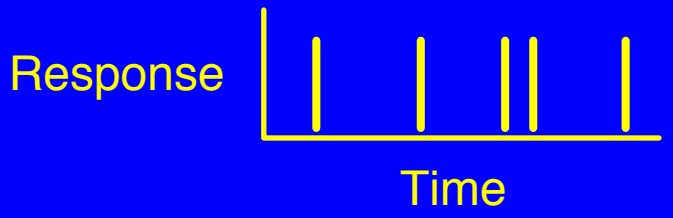
Electrical response

Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround





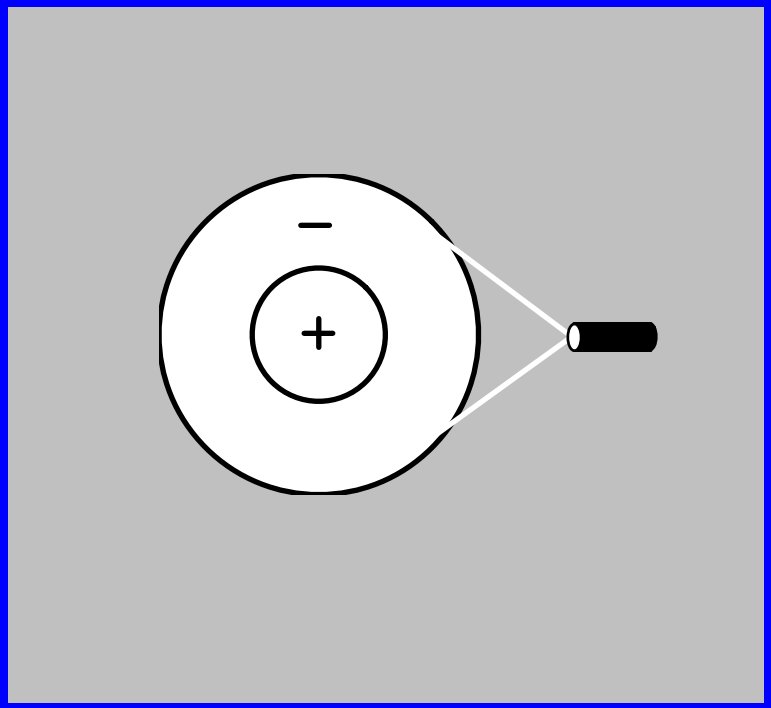
Stimulus condition

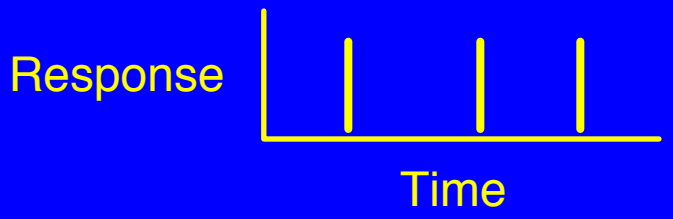
Electrical response

Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround





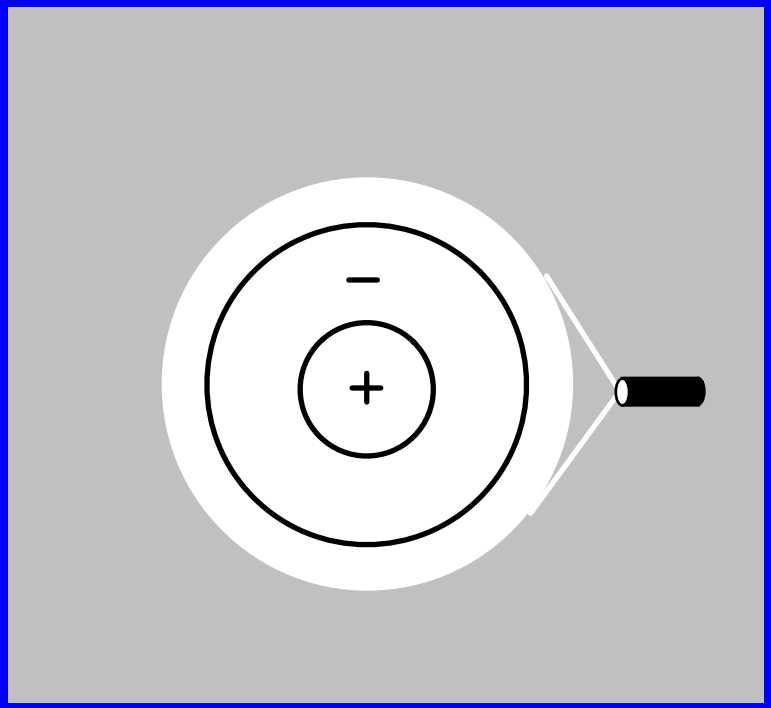
Stimulus condition

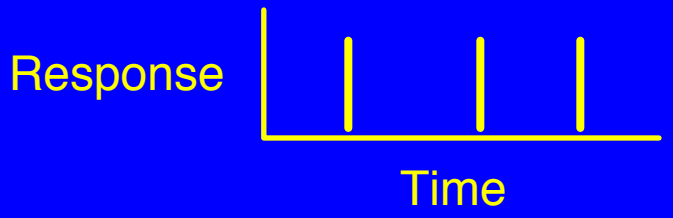
Electrical response

Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround





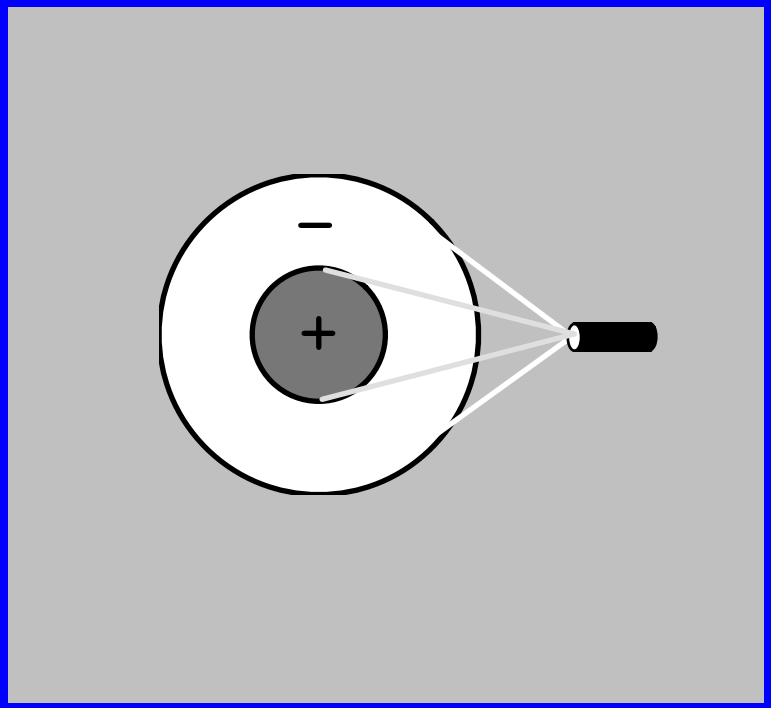
Stimulus condition

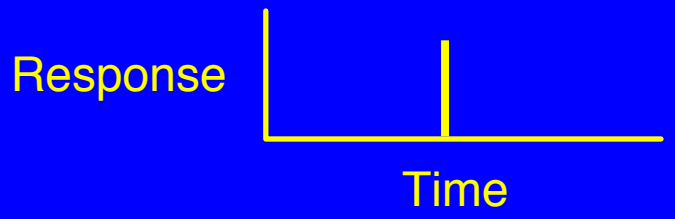
Electrical response

Visual Physiology

- **Receptive field structure in Ganglion Cells**

On-center Off-surround



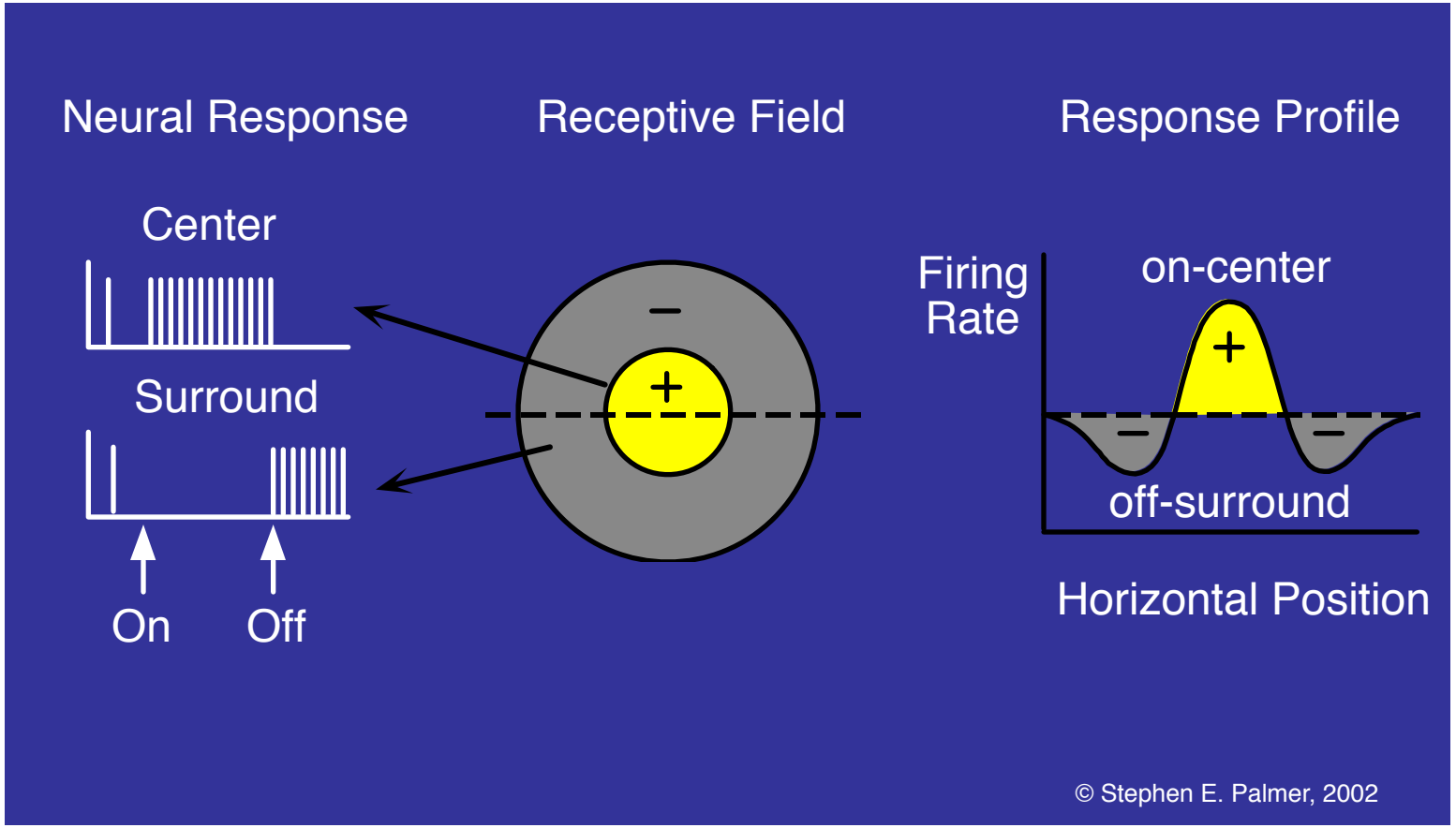


Stimulus condition

Electrical response

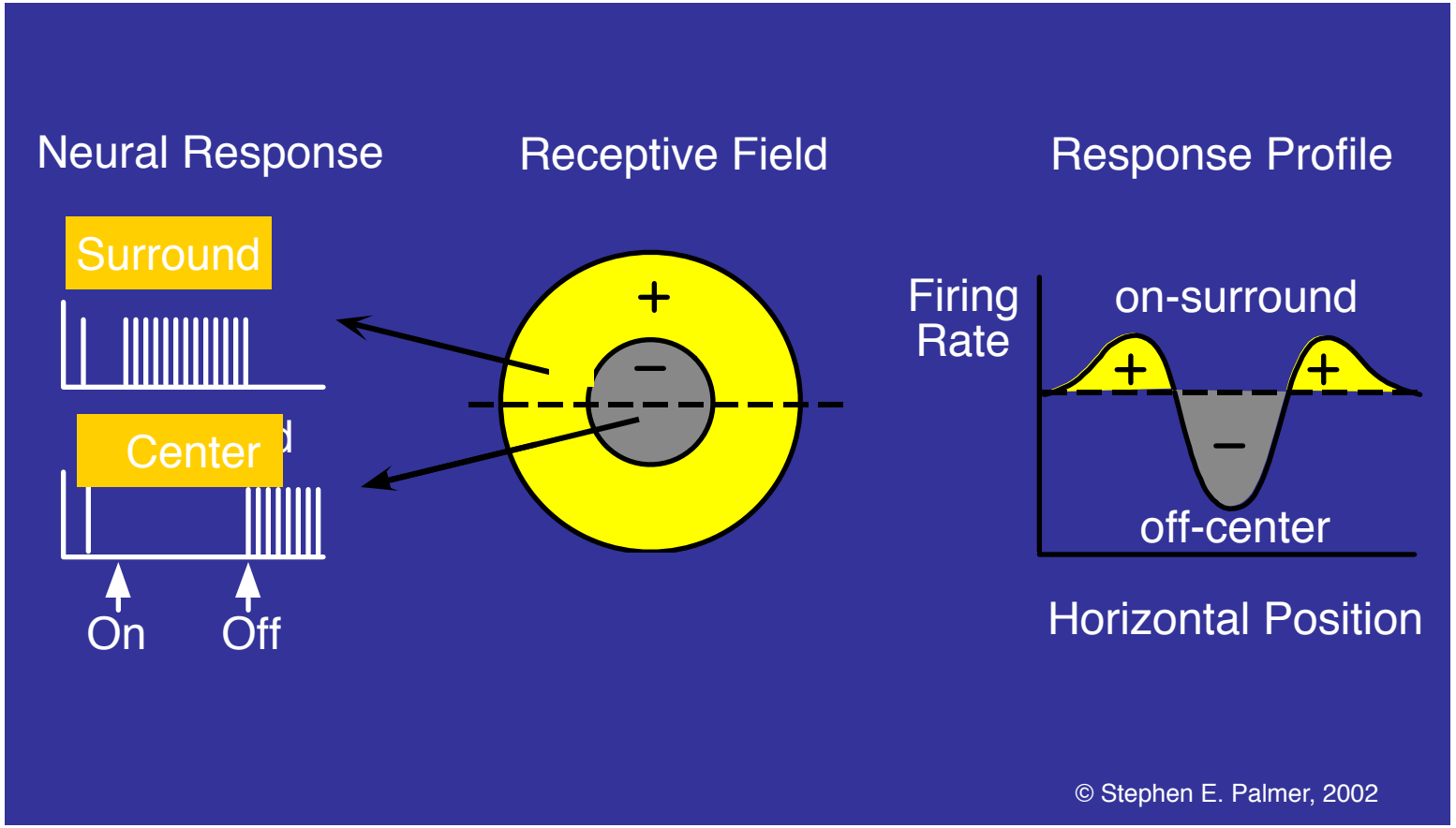
Visual Physiology

- **RF of On-center Off-surround cells**



Visual Physiology

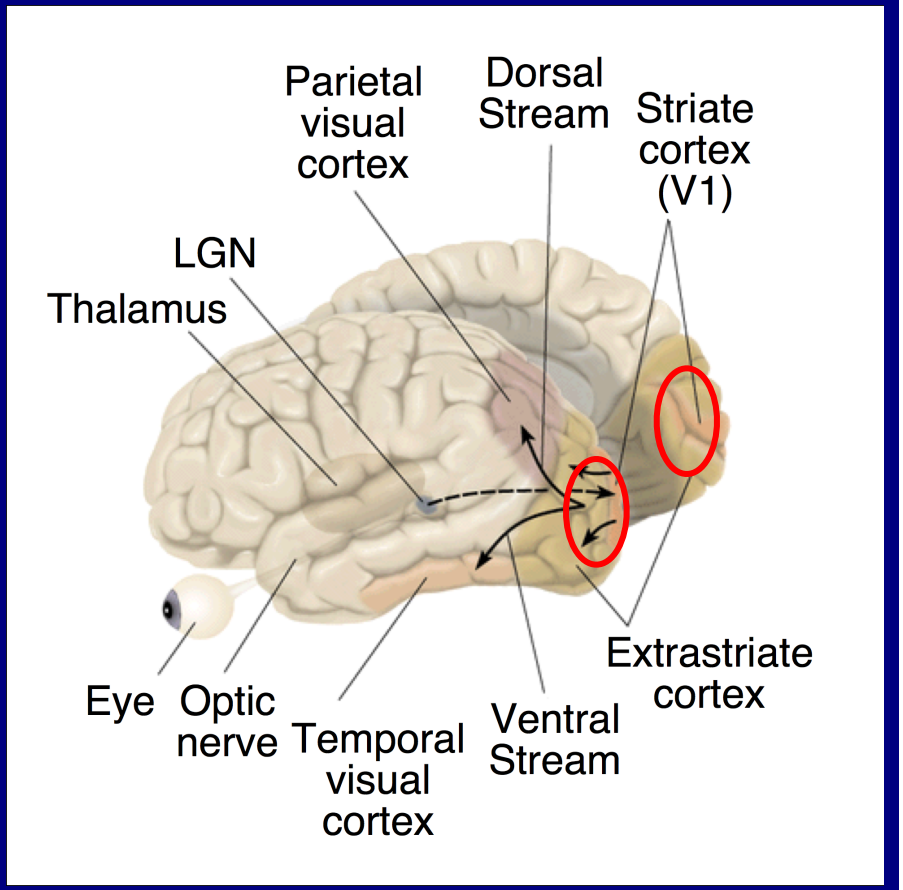
- **RF of Off-center On-surround cells**



Visual Physiology

Cortical Area V1

1、Primary visual cortex
2、Striate cortex
3、Brodman's area 17



Visual Physiology

Single-cell recording from visual cortex

David Hubel & Thorston Wiesel

Visual Physiology

▪ Cortical Receptive Fields

Three classes of cells in V1

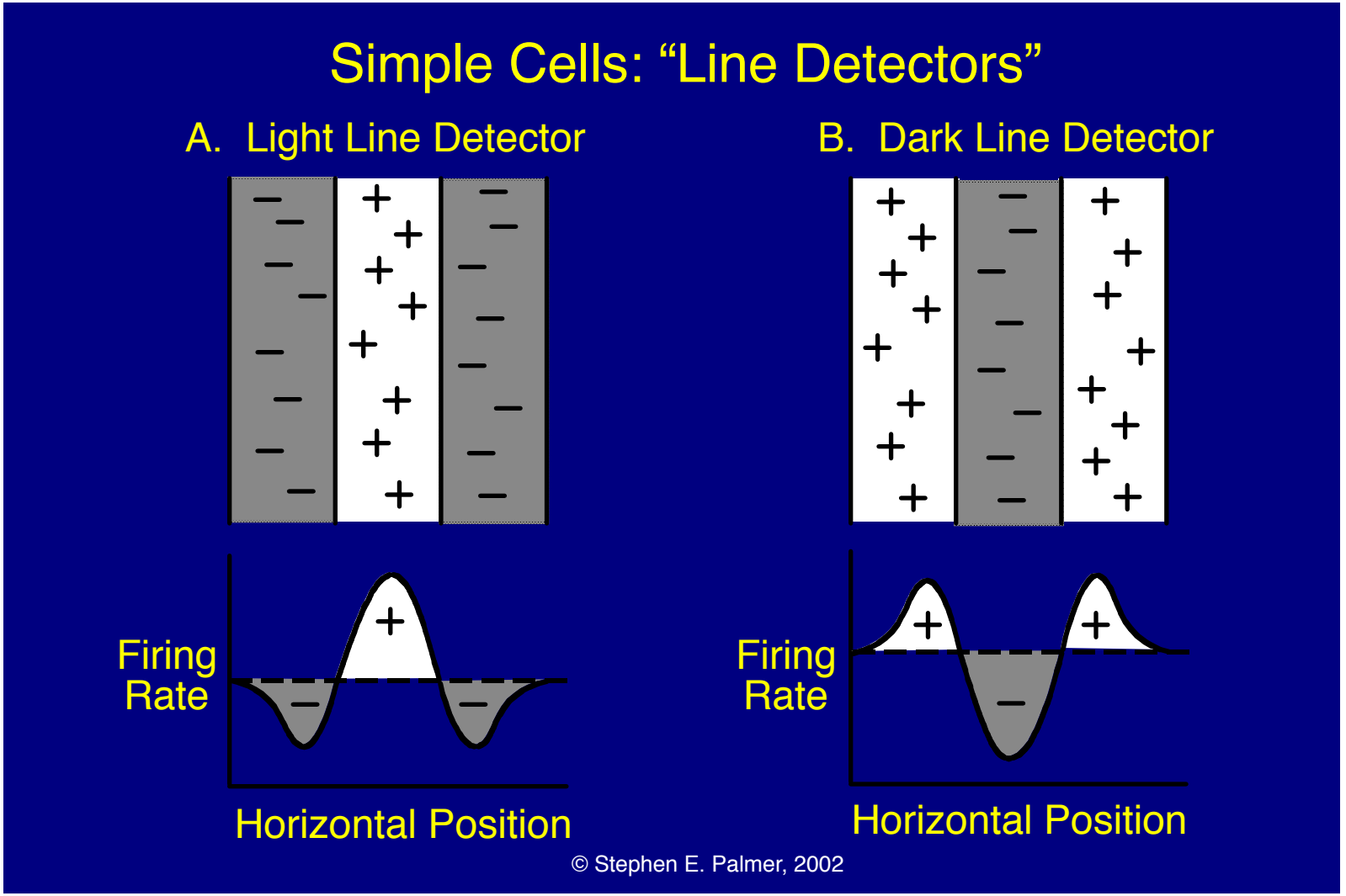
Simple cells

Complex cells

Hypercomplex cells

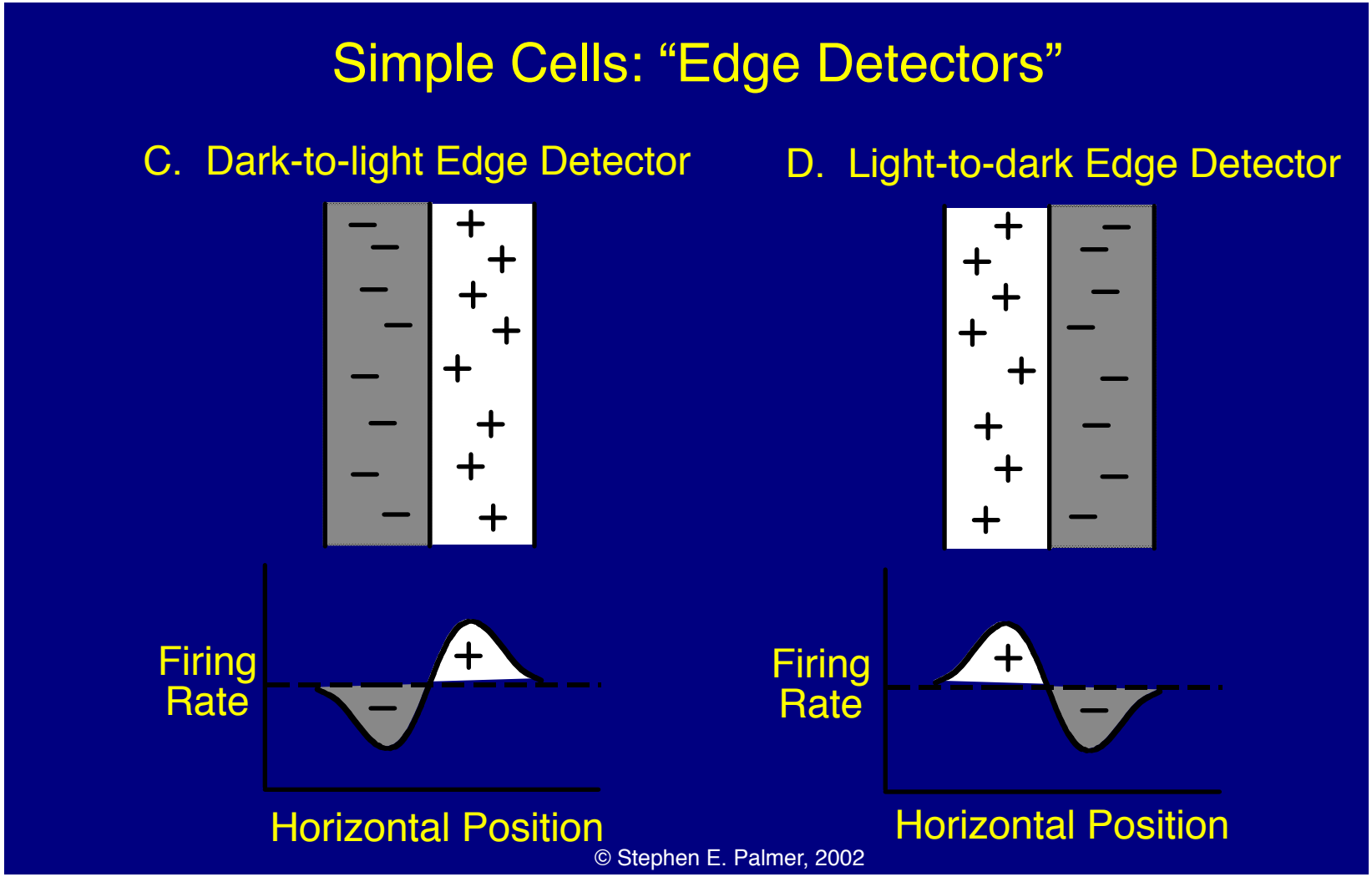
Visual Physiology

Cortical Receptive Fields



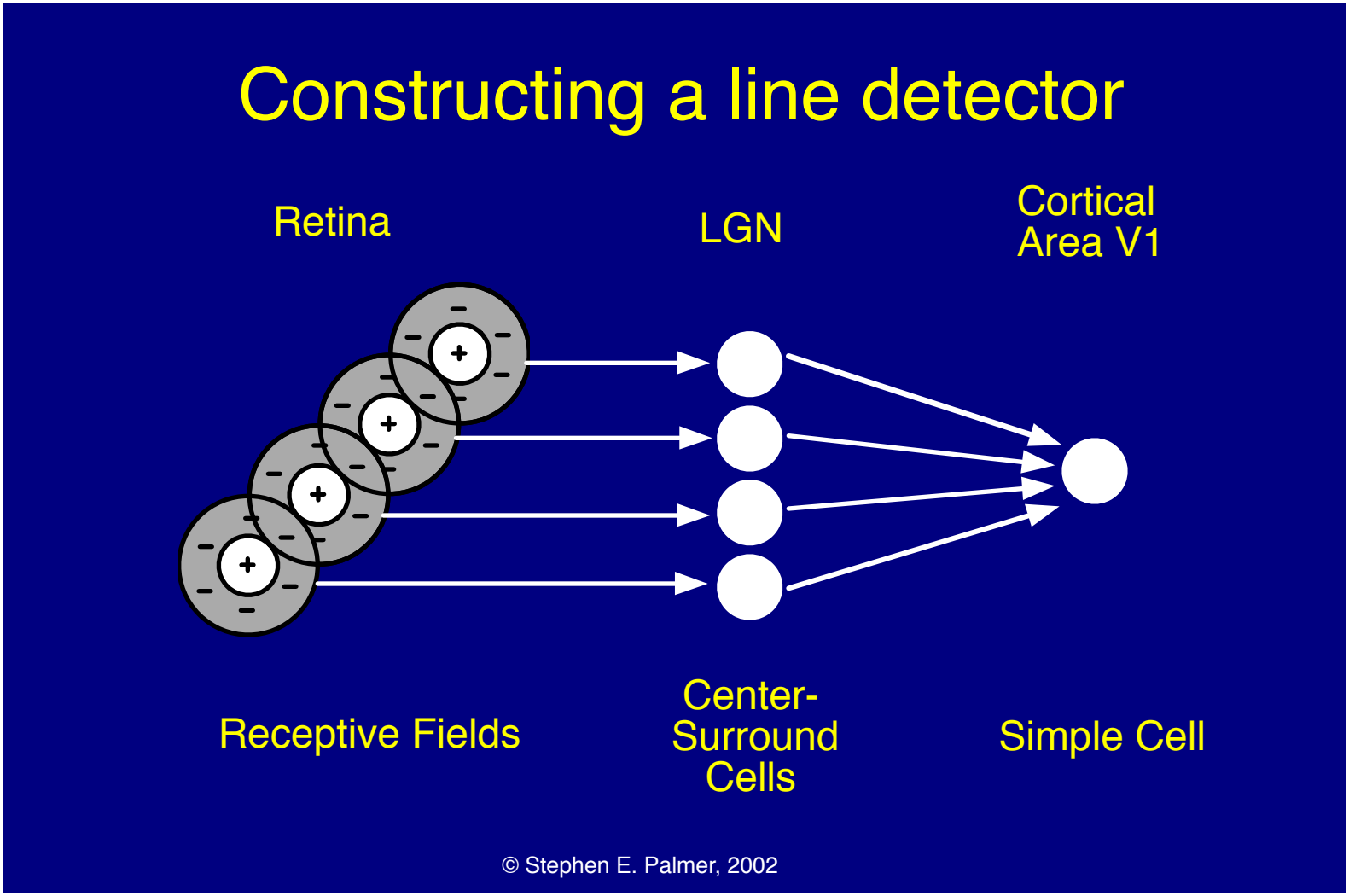
Visual Physiology

Cortical Receptive Fields



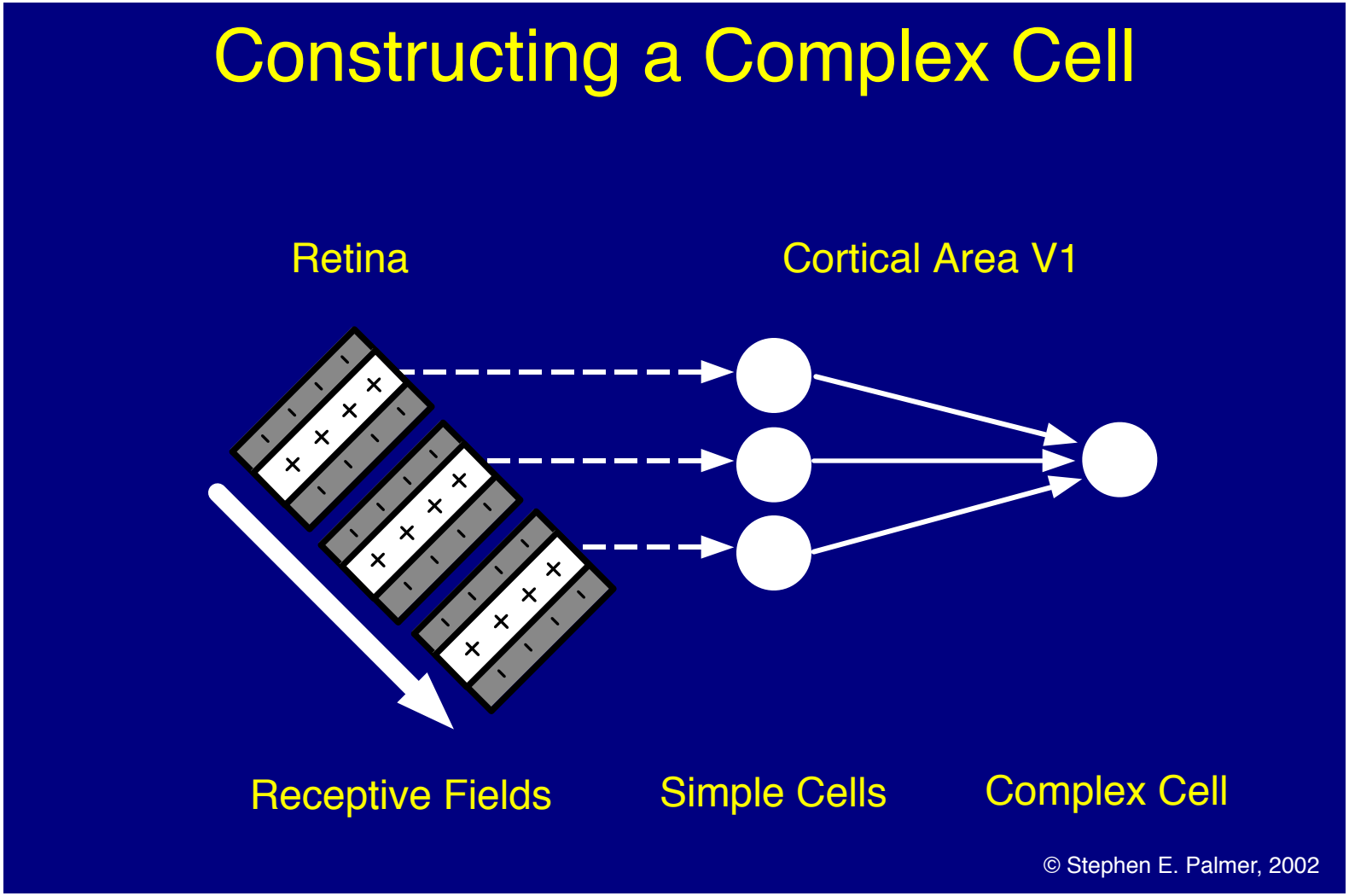
Visual Physiology

Cortical Receptive Fields



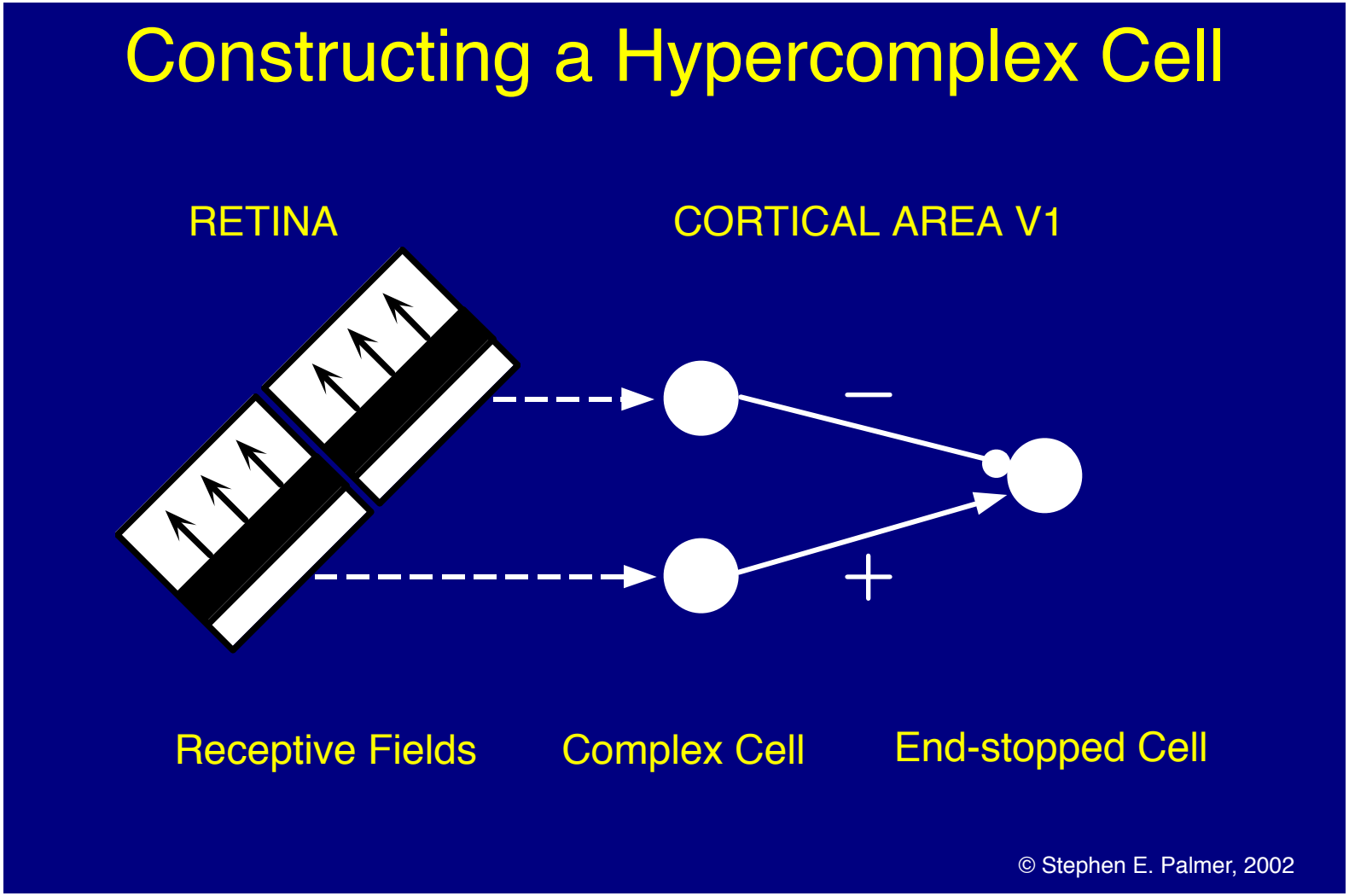
Visual Physiology

- **Cortical Receptive Fields**



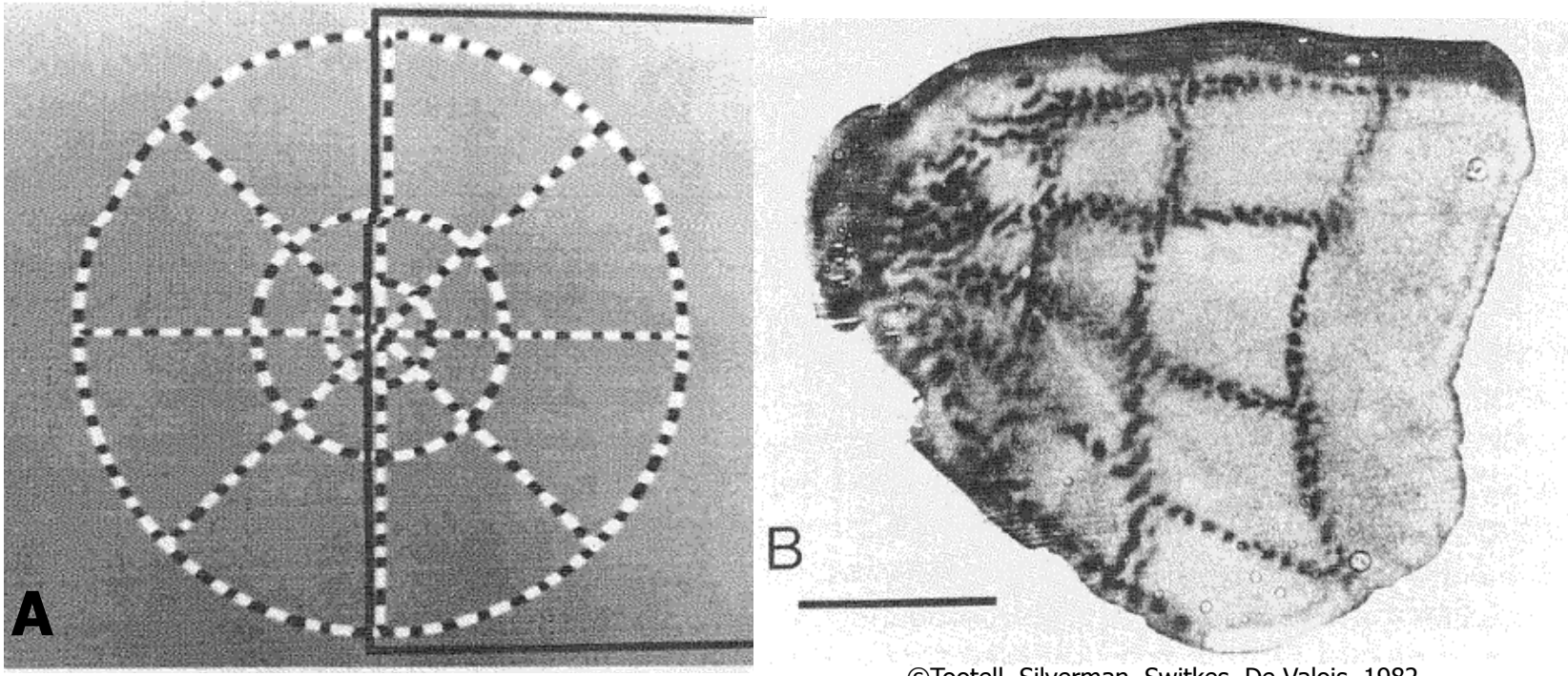
Visual Physiology

Cortical Receptive Fields



Visual Physiology

- **Mapping from Retina to V1**

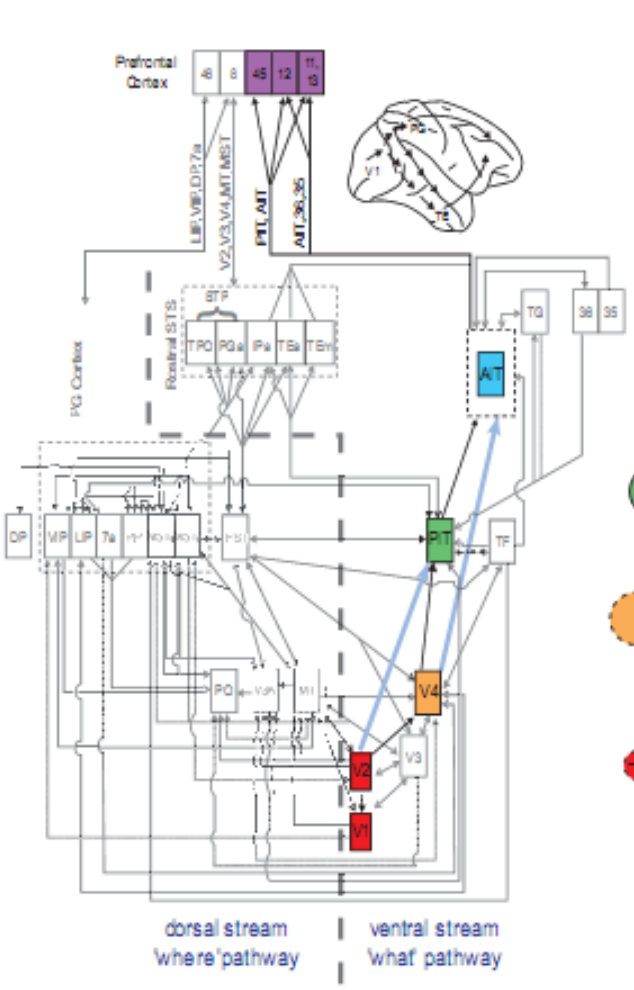


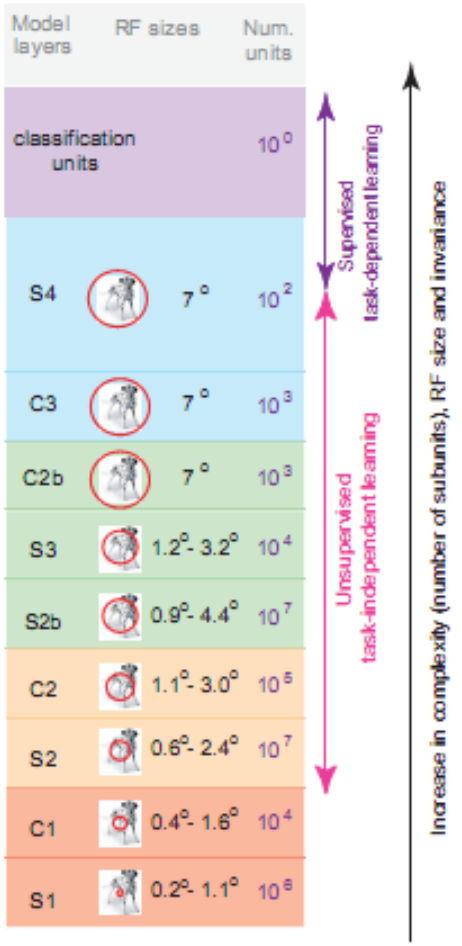
©Tootell, Silverman, Switkes, De Valois, 1982.

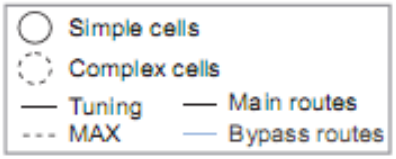
The retinotopic map in area V1 of a macaque monkey

基于分层前馈的信息处理结构

A feedforward architecture accounts for rapid categorization







(Poggio, PNSA, 2007)

Theories of Vision

- **Structuralism** （结构主义）
- **Gestaltism** （完形心理学）
- **Ecological Optics** （生态光学）
- **Constructivism** （构造主义）

Theories of Vision

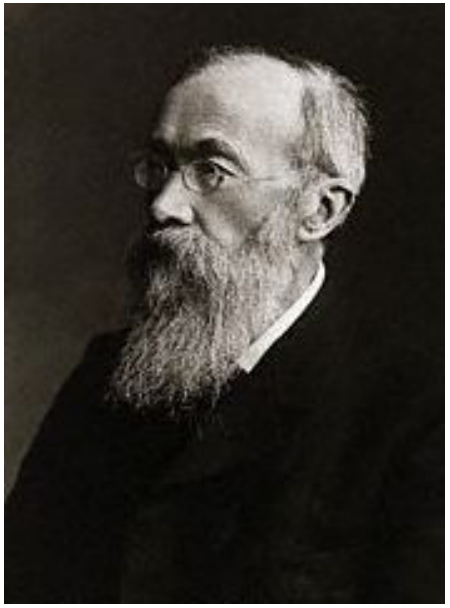
Structuralism（结构主义）

Structuralism(结构主义)

Perception results from the association of basic sensory atoms in memory via repeated, prior joint occurrences.

----Proposed by Wilhelm Wundt.

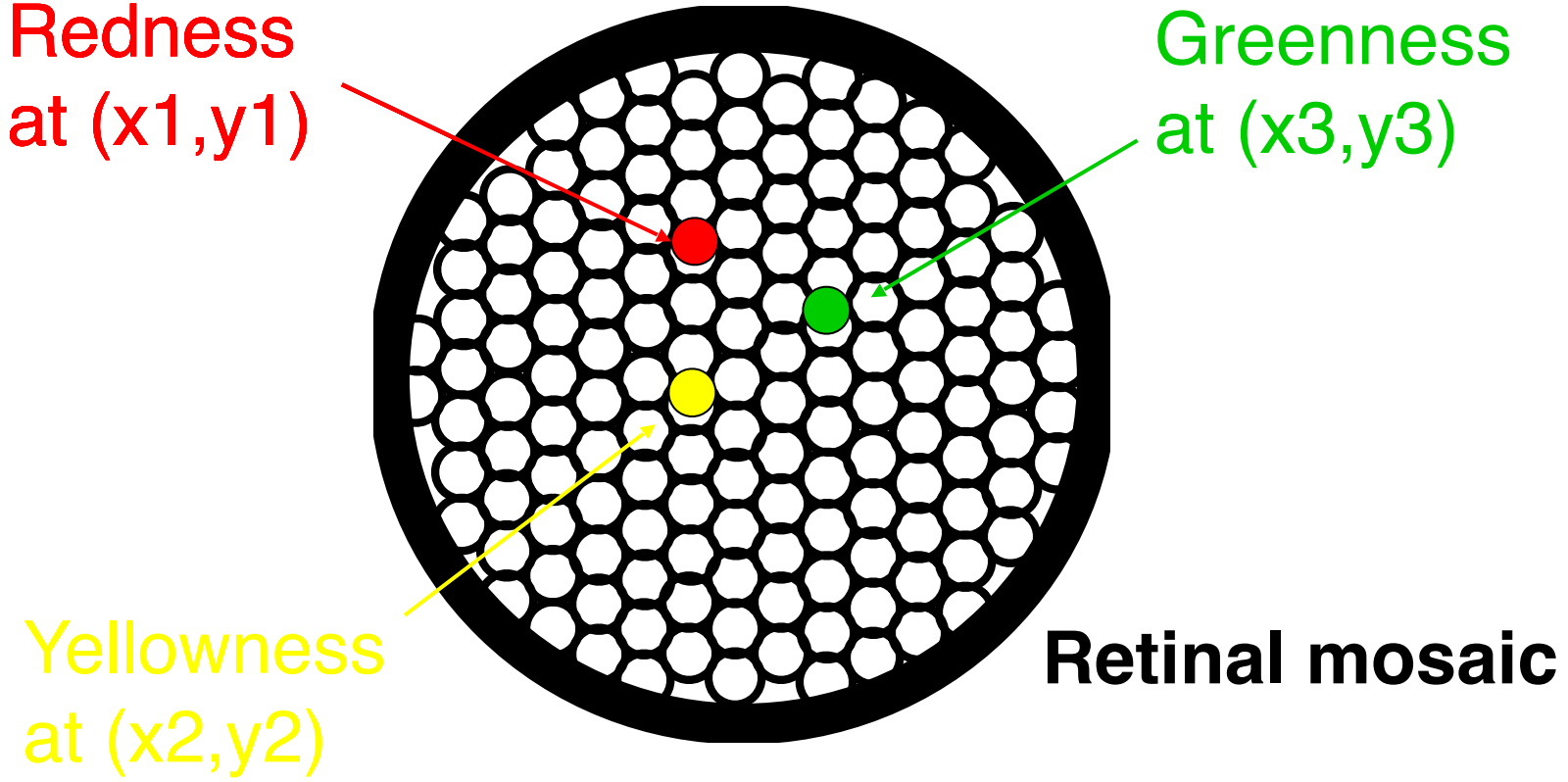
感知结果来自记忆体中的基本传感原子通过重复与预先联合共生的形式相互关联。



冯特（Wilhelm Wundt，1832～1920）为科学心理学的创始者，实验心理学之父，创建了一支国际心理学专业队伍，奠定构造主义心理学派的基础，感情三度说的倡议者，也是使心理学脱离哲学范畴成为一门独立科学的巨擘。

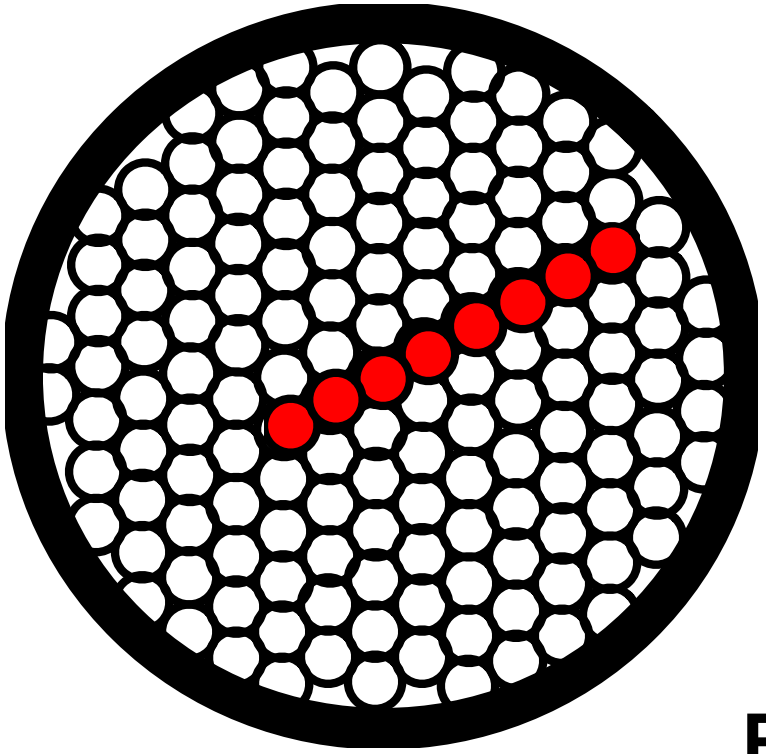
Theories of Vision

Sensory Atoms



Theories of Vision

Perceptual Complexes

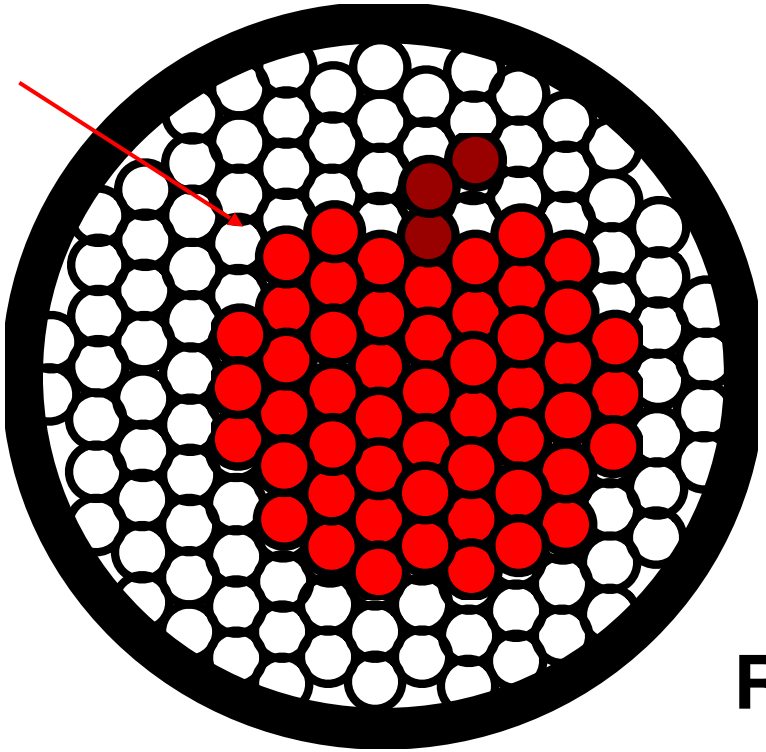


Retinal mosaic

Theories of Vision

Perceptual Complexes

Red apple
at (x0,y0)

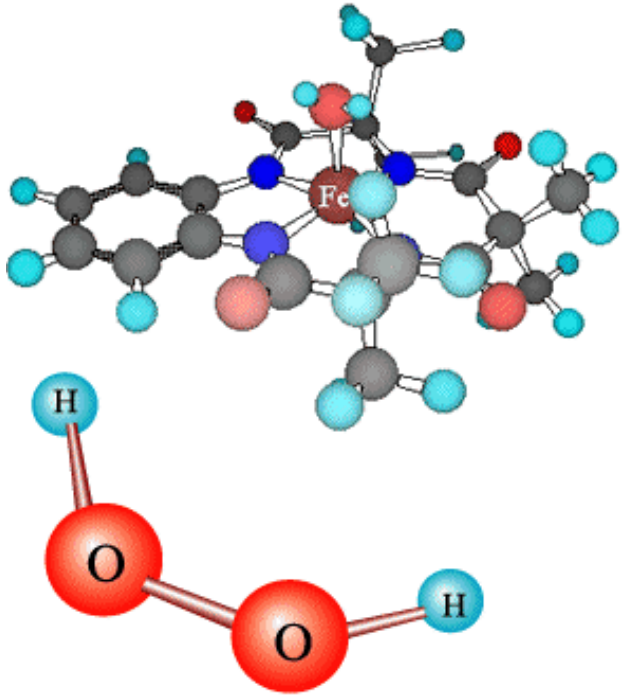


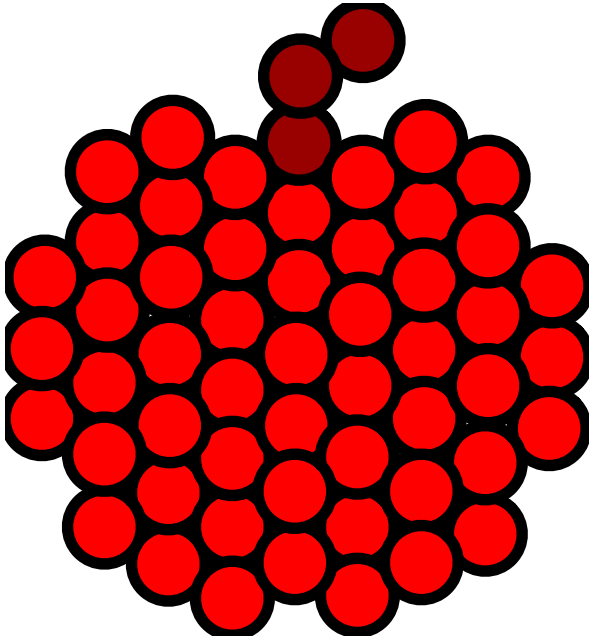
Retinal mosaic

Theories of Vision

Chemical Analogy

Perceptions are made of basic sensory experiences just as molecules are made of basic atoms.





Theories of Vision

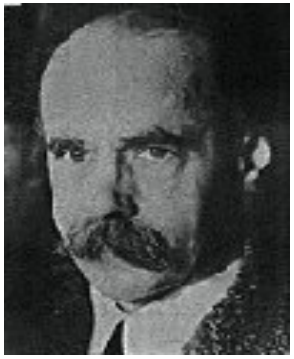
Gestaltism（完形心理学）

Perception results from the interaction between the intrinsic structure of the stimulus and the intrinsic structure of the brain.

主要观点：

感知结果来源于激励的本质结构与大脑的本质结构之间的交互作用。

对表象的体验要保持表象的本来面目，不能将它分解为感觉元素，并认为对表象的体验是整体的或完形的。



Max Wertheimer (韦特海默)

Wolfgang Köhler (苛勒)

Kurt Koffka (考夫卡)

Theories of Vision

Principles of Gestalt Theory

Holism（整体论）:

The whole is different from the sum of its parts.

Prägnanz（简单原则）:

The percept will be as "good" as the prevailing conditions allow, I.e. simplest explanation

Nativism（先天论）:

Not a total rejection of learning, but rejection of Its primacy.

不是完全拒绝后天学习，但拒绝其首要地位。

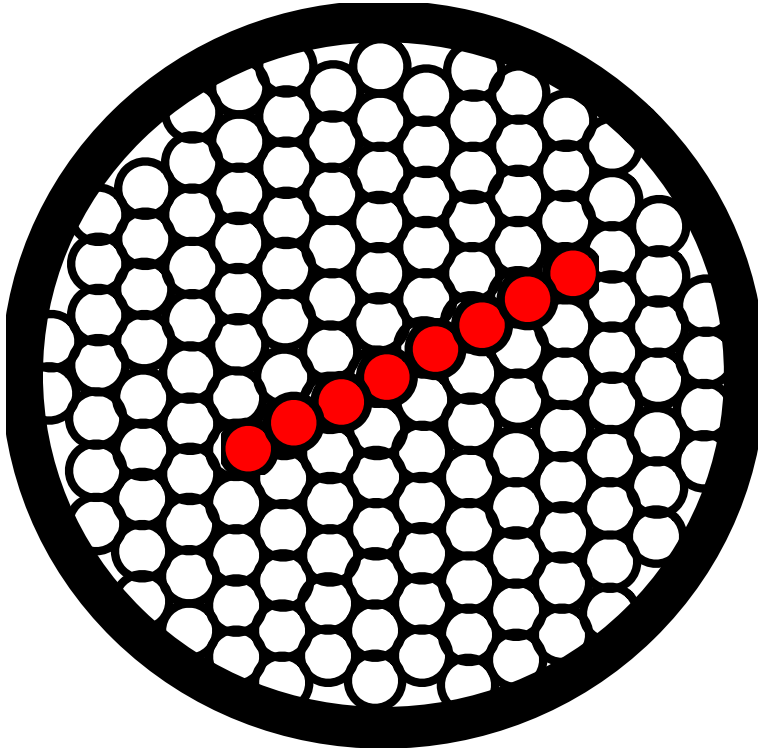
Theories of Vision

Holism: The whole is different from the sum of its parts.

Emergent properties:

Features of a configuration that are not features of its components, e.g.:

- length
- orientation
- curvature
- closure
- connectedness



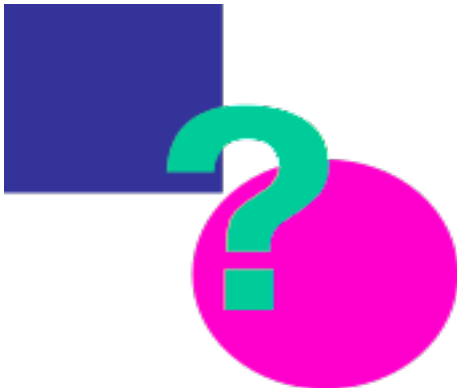
整体论观点：整体不同于部分之和，具有涌现特性。

Theories of Vision

Theories of Vision

Prägnanz: the percept will be as "good" as the prevailing conditions allow

What is this?



square & circle?

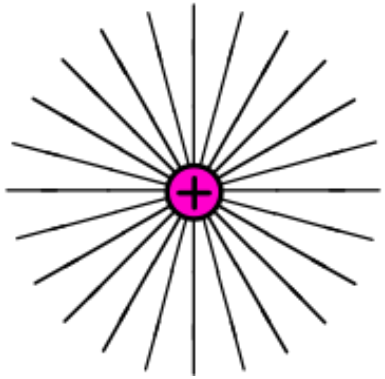
square & pacman?

squigit & pacman?

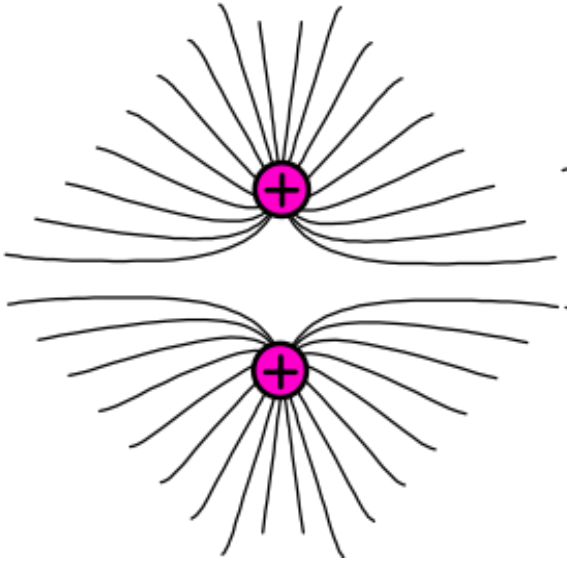
etc….

Theories of Vision

Field Theoretic Analogy

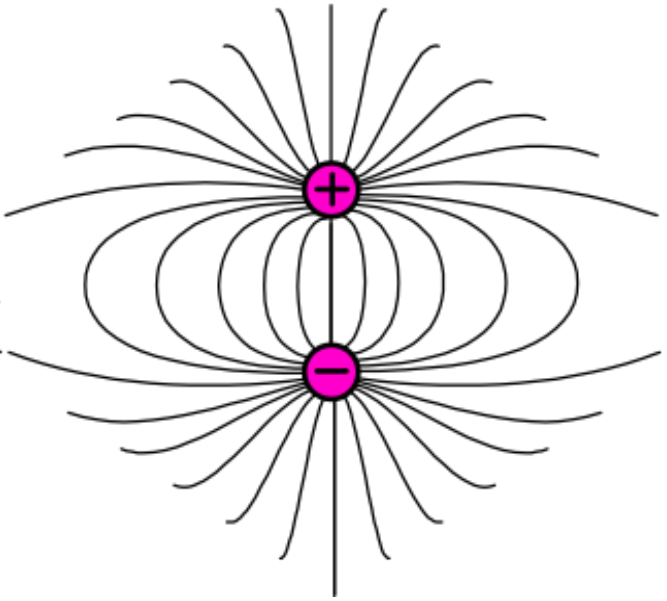


Field of a single positive charge



Field of a pair of positive charges

Repulsion



Field of a positive & a negative charge

Attraction

Theories of Vision

Ecological Optics（生态光学）

Perception is the direct apprehension of the visible environment by extracting invariants in the dynamic ambient optic array and the affordances of objects.

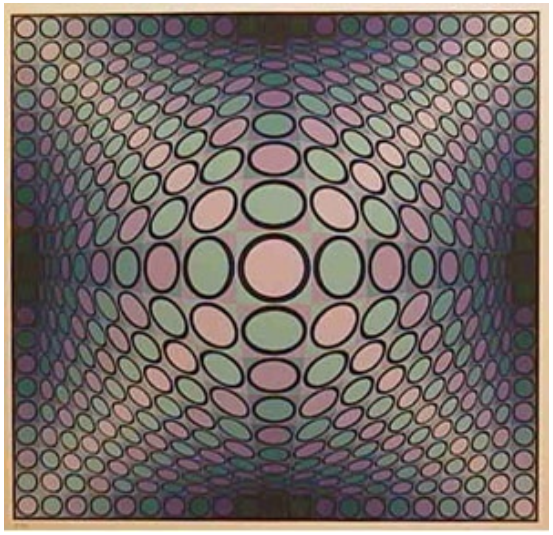
美国心理学家J.J.吉布森于20世纪50年代提出的一种视知觉理论。它与传统的知觉理论不同，不承认知觉中经验和推理的作用，而认为知觉是人从环境得到刺激信息的直接感悟。即,直接从可见环境中提取不变性的结果。

Theories of Vision

Texture Gradients are gradual changes in the size & shape of texture elements in an image when a plane recedes in depth.

Tile floor

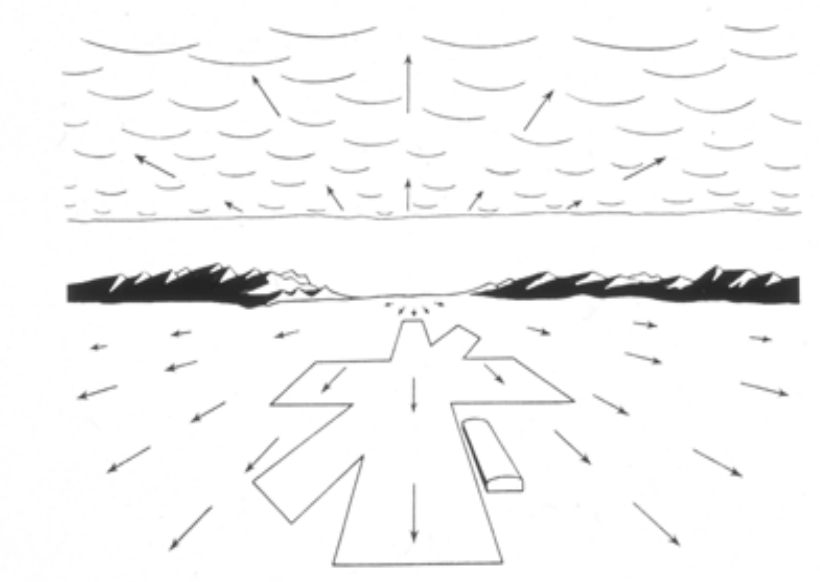
Wheat field



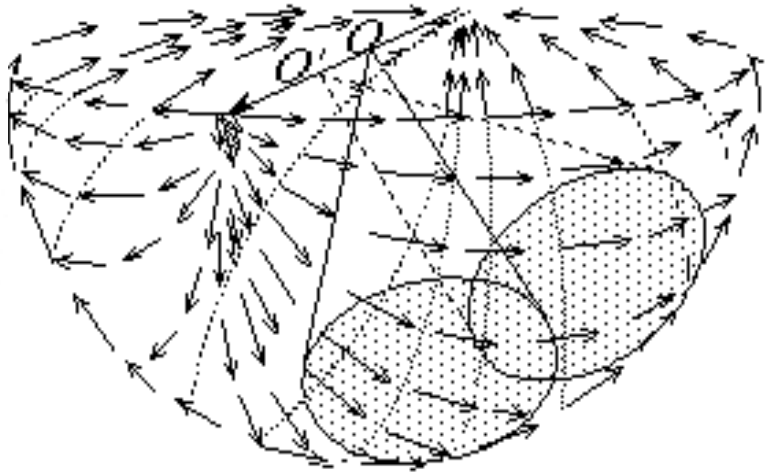
Geometric surface

Theories of Vision

Motion Gradients are gradual changes in the speed and direction of elements in optic flow within the dynamic ambient optic array.



Plane landing



Bird flying

Theories of Vision

Constructivism（构造主义）

Perception is the result of unconscious inferences about the scene most likely to have caused the retinal image or event.

Hermann von Helmholtz originated the idea of **unconscious inference** and the likelihood principle.

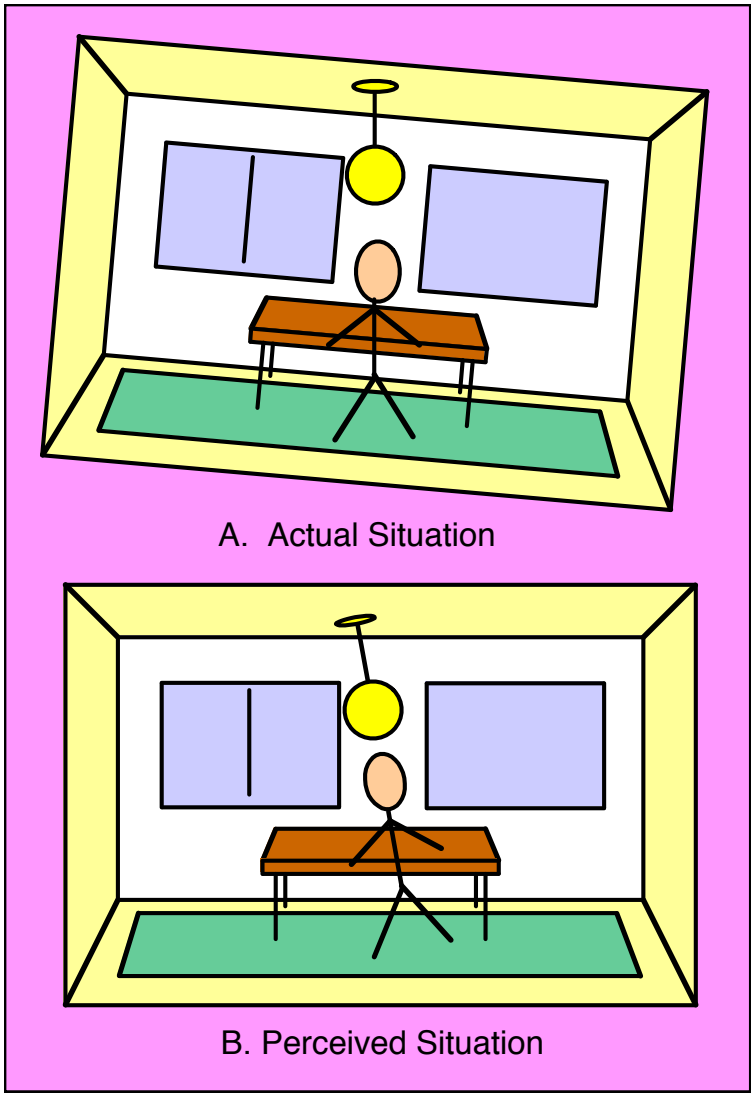
十九世纪最伟大的科学家之一，德国的物理学家、生理学家兼心理学家。发展颇富盛名的扬－赫尔姆霍茨三色说；其生理学的突出成就为实验心理学的建立打下了根基。

Theories of Vision

Unconscious Inference: the process of recovering environmental information by logically combining **retinal information with heuristic assumptions**.

Tilted room illusion:
If you assume that the walls and floor of the room are vertical and horizontal, then you must be tilted —and you feel that way!

Theories of Vision



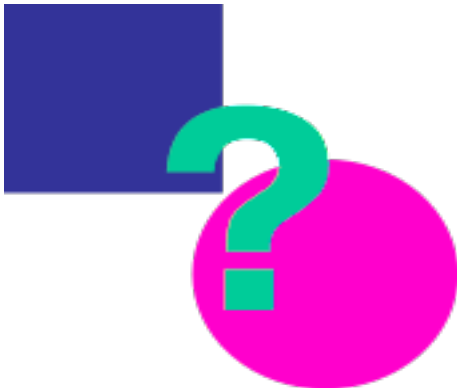
简单原则与似然。

Theories of Vision

Prägnanz vs. Likelihood

What governs what we see: goodness or probability?

What is this?



square & circle?

square & pacman?

squigit & pacman?

etc….

Theories of Vision

THEORY	NATIVISM vs. EMPIRICISM	ATOMISM vs. HOLISM	ORGANISM vs. ENVIRONMENT	PRINCIPAL ANALOGY	METHOD
Structuralism	Empiricism	Atomism	Organism	Chemistry	Trained Introspection
Gestaltism	Nativism	Holism	Organism	EM Fields	Naïve Introspection
Ecological Optics	Nativism	Holism	Environment	Mechanical Resonance	Ecological Analysis
Constructivism	(both)	(both)	(both)	Logical Inference	Likelihood Analysis

Theories of Vision

Which theory is correct?

Probably none of them!

Or maybe all of them, to some degree!

It's good to keep them in mind, when designing your algorithms!

Theories of Vision

A brief history of visual information processing

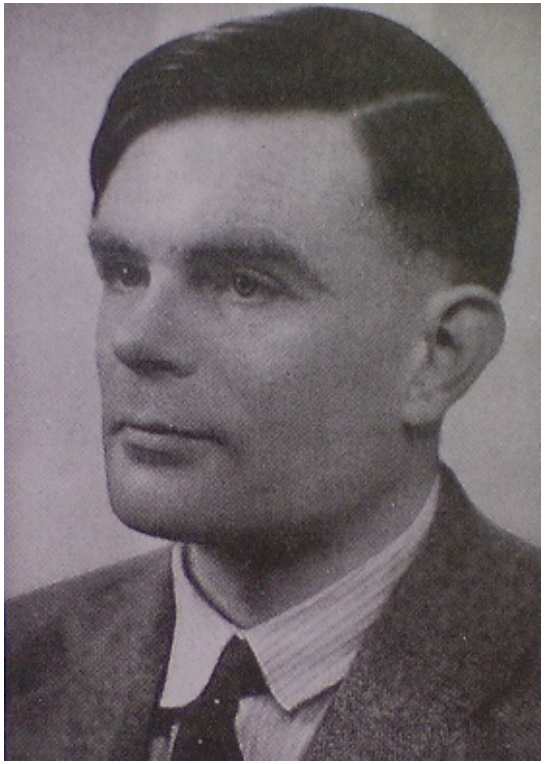
Computer Vision

- ✓Turing Machines(1930-1937); ENIAC(1946)
- ✓**AI** (simulating intelligent thought,Turing,1950)
- ✓Difficult intellectual tasks(playing chess, proving mathematical theorems — Newell & Simon,1958,1963)
- ✓Perceive the environment visually (1960s---)

Two important developments

- ✓Tested on real images of real objects. (Theories)
- ✓Explicit theories

Changed the theoretical branch of vision science.



(Alan Turing)
A brilliant British mathematician

The first insight: "vision is extremely difficult."

Theories of Vision

Three Levels of Information Processing [1]

➢ The Computational Level

(Computational Theory)

✓The goal of the computation

✓The logic of strategy

➢The Algorithmic Level

(Representation & Algorithm)

✓How to implement the computational theory

✓The representation for input and output

✓The algorithm for the transformation

➢ The Implementational Level

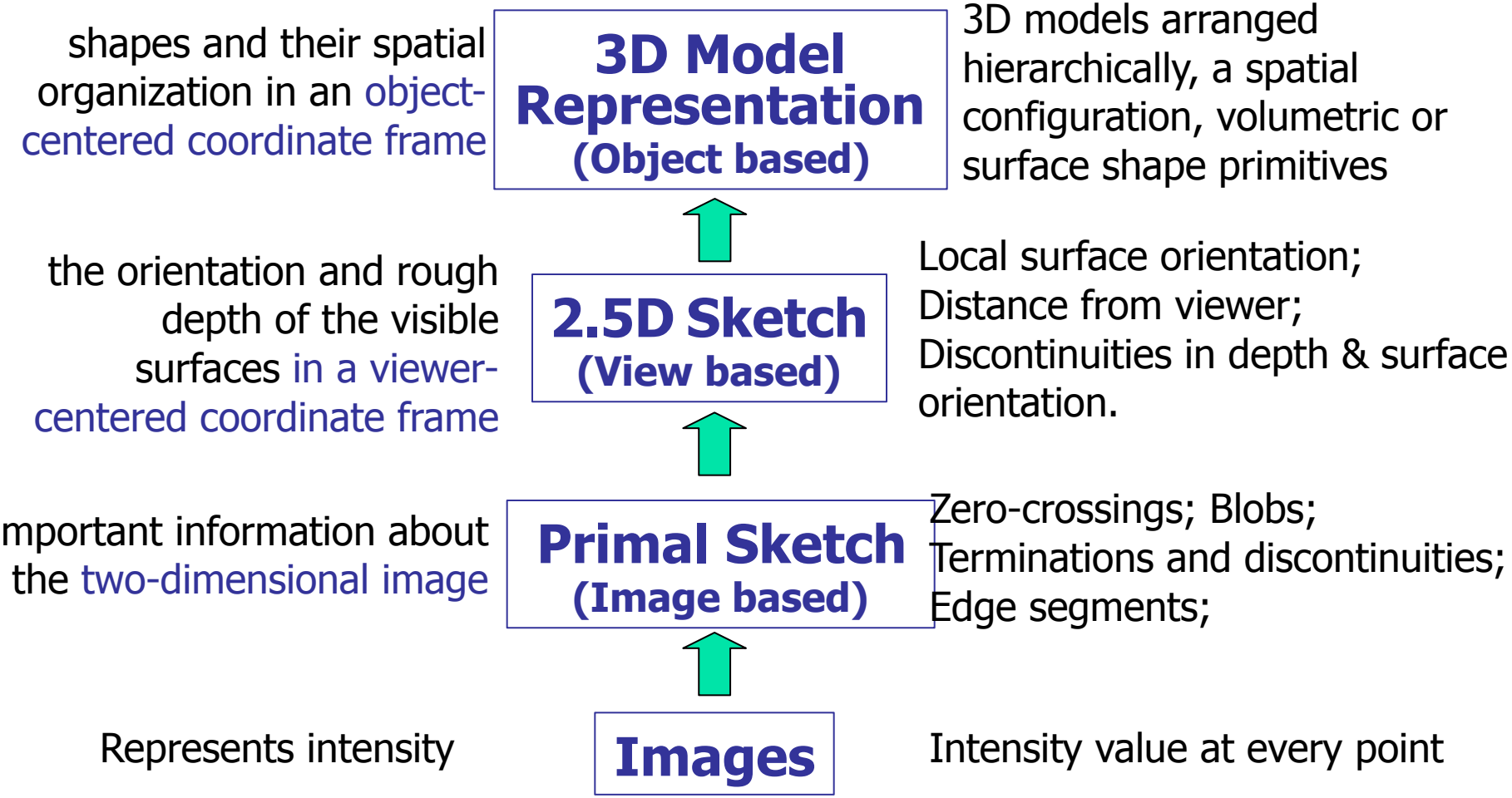
(Hardware Implementation)

✓Realized physically

David Marr

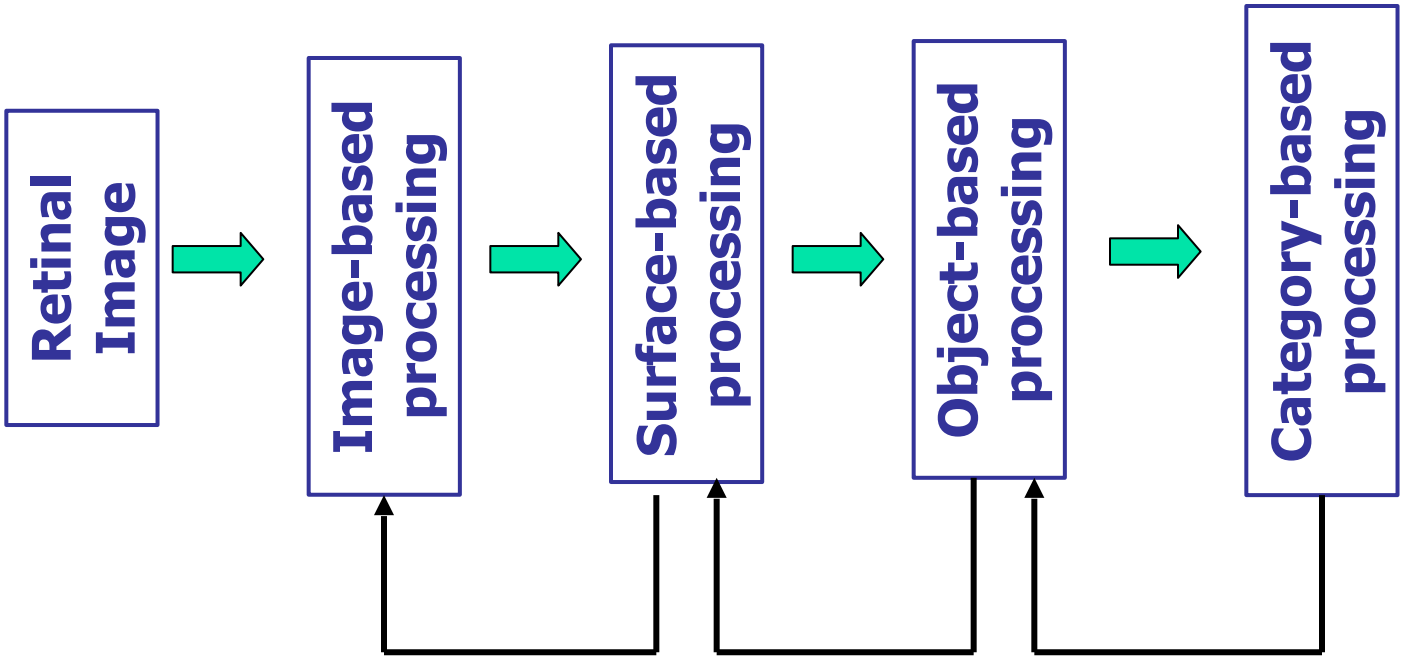
Theories of Vision

- **A Representational Framework for Vision**[1]



Theories of Vision

- **Four major stages at the algorithmic Level** [2]

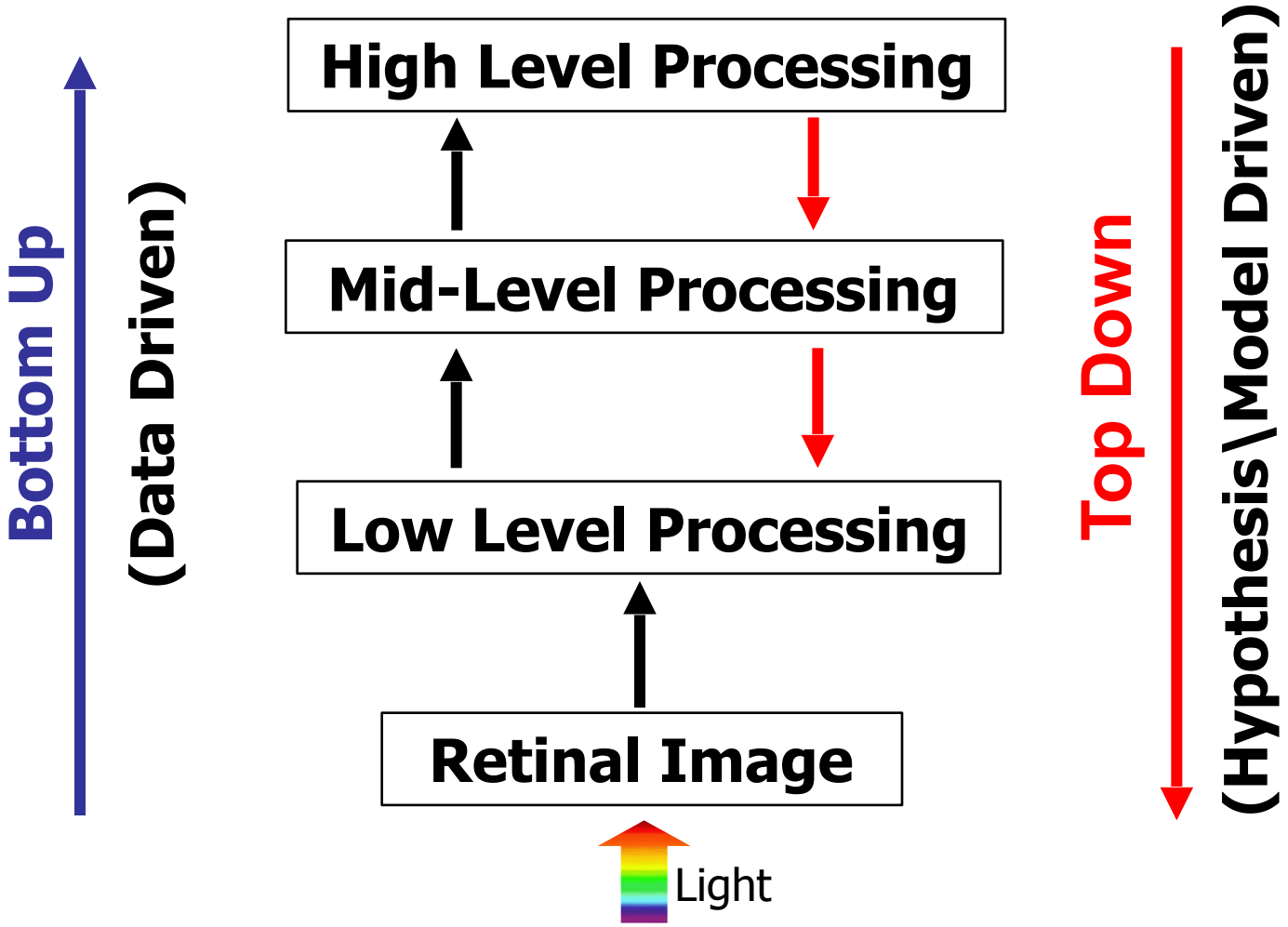


Visual Perception

The four stages provide a fairly general and robust framework for understanding vision as a computational process.—S.E. Palmer, Vision Science.

Theories of Vision

- **Top-Down vs Bottom Up**



参考文献

[1] D. Marr, **Vision**----A computational investigation into the human representation and processing of visual information, W.H. Freeman and Company San Francisco, 1982.

[2] S.E. Palmer, **Vision Science** – Photons to Phenomenology, the MIT Press, 1999.