Perceptual Computing:
Making Machines Sense, Perceive, and Interact
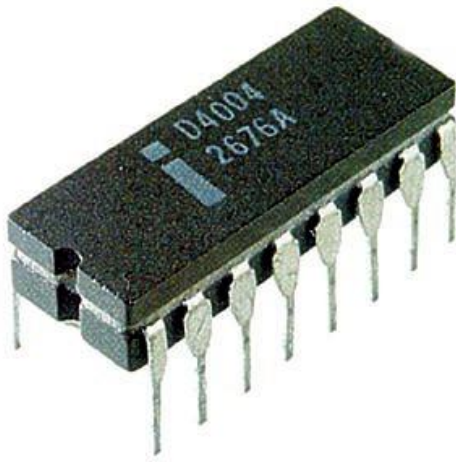
Achin Bhowmik, Ph.D.
VP & GM, Perceptual Computing Group
Intel Corporation
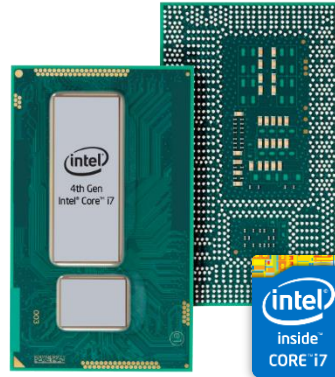
Perceptual Computing Mission

Add "Senses" to the "Brain": Eyes, Ears, Voice, Touch, Emotion and Context for a Natural, Intuitive and Immersive Life-like Experience

First, a quick look at the "Brain"

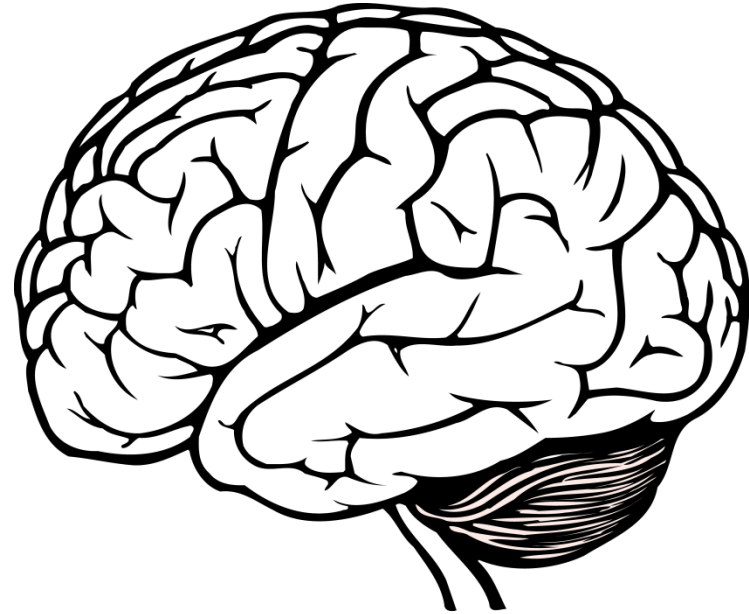Intel® 4004
1971

2.3K Transistors
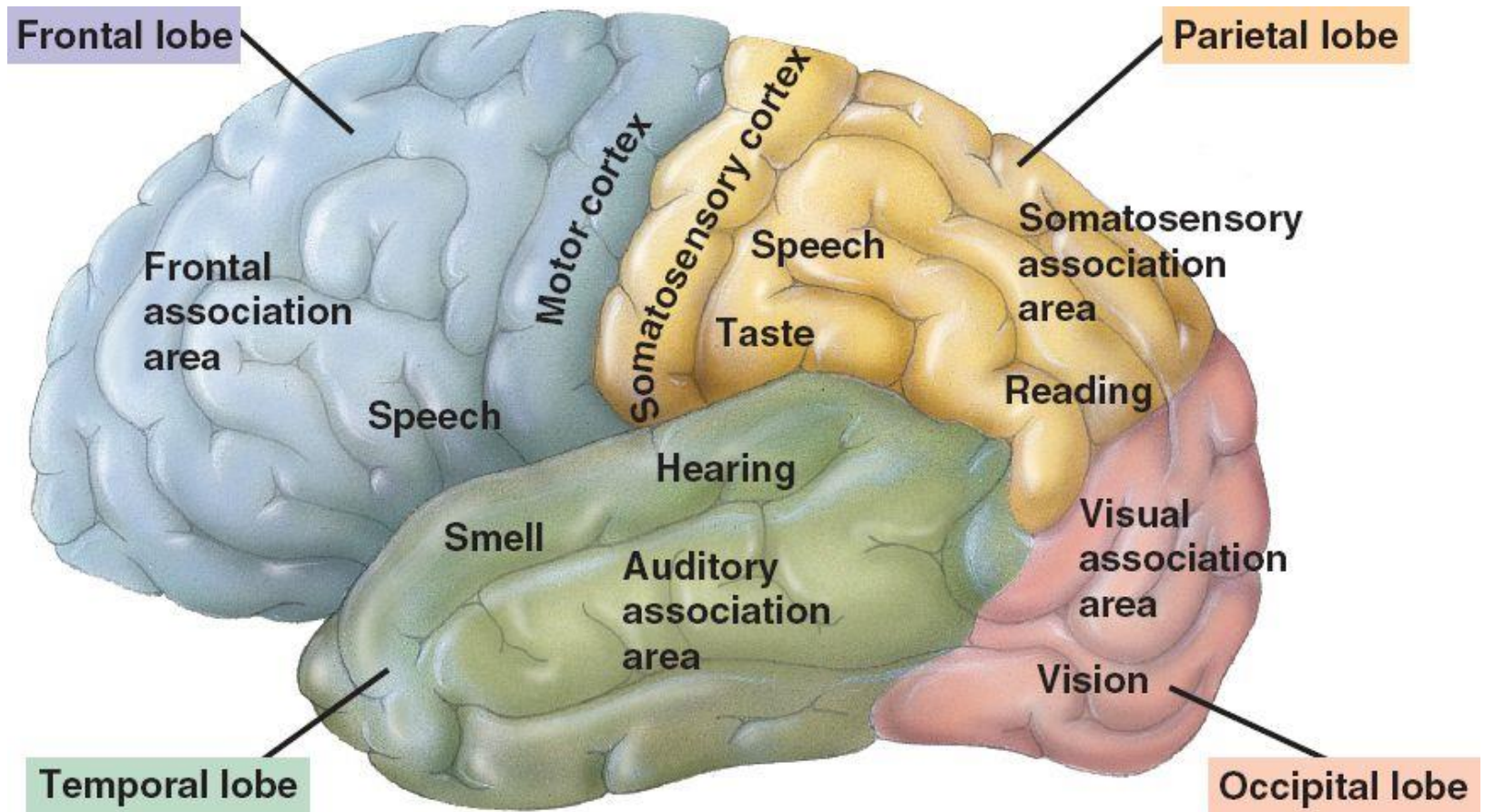740KHz

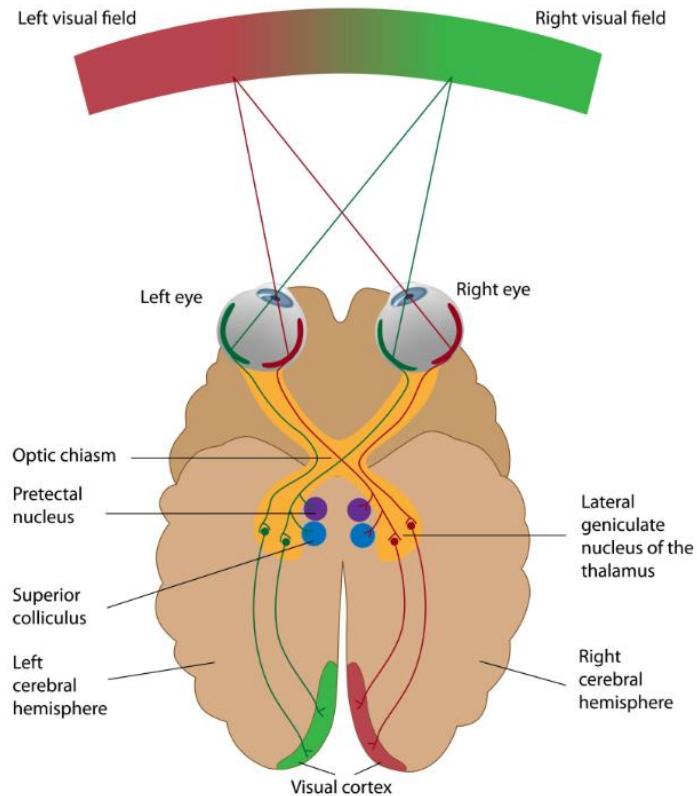Intel® Core™ Processor
2014

~2B Transistors
>3GHz

Homo Sapiens
Brain

~100B Neurons
<1KHz

Cerebral Cortex:
Sensory Processing Areas
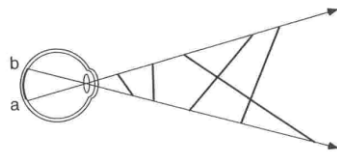
Visual Sensing

Human Visual System:
3D Vision

Traditional Computing Devices:
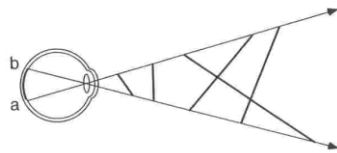2D Vision

2D Imaging Limitations

[R', G', B', x', y'] ⬅ [R, G, B, x, y, z]

Projections

$$\begin{pmatrix} 2D \\ point \end{pmatrix} = \begin{pmatrix} \text{Camera to pixel coord. trans. matrix} \end{pmatrix} \begin{pmatrix} \text{Perspective projection matrix} \end{pmatrix} \begin{pmatrix} \text{World to camera coord. trans. matrix} \end{pmatrix} \begin{pmatrix} 3D \\ point \end{pmatrix}$$

2D Imaging Limitations

[R', G', B', x', y'] ← [R, G, B, x, y, z]

Projections

???

$$\begin{pmatrix} \text{2D} \\ \text{point} \end{pmatrix} = \begin{pmatrix} \text{Camera to} \\ \text{pixel coord.} \\ \text{trans. matrix} \end{pmatrix} \begin{pmatrix} \text{Perspective} \\ \text{projection matrix} \end{pmatrix} \begin{pmatrix} \text{World to} \\ \text{camera coord.} \\ \text{trans. matrix} \end{pmatrix} \begin{pmatrix} \text{3D} \\ \text{point} \end{pmatrix}$$

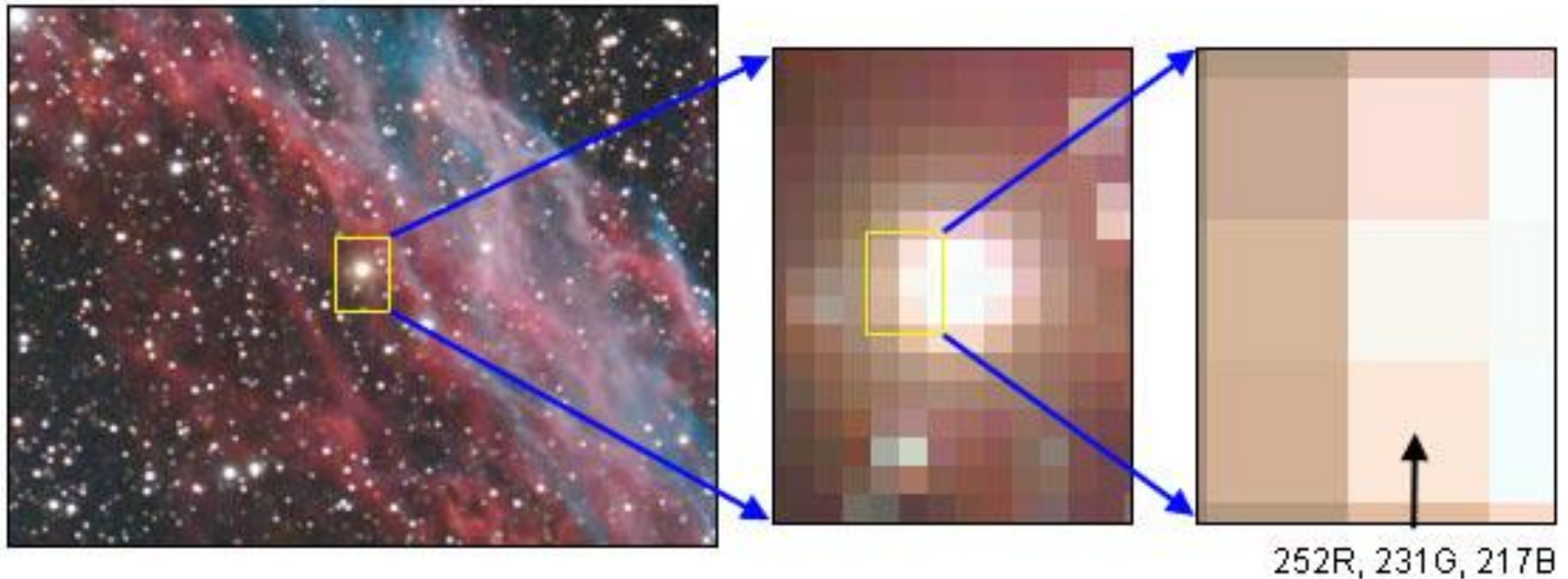# Life's Much Easier with Depth-Imaging
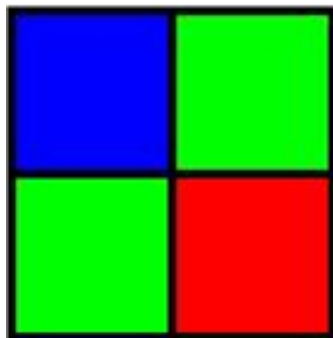

Color Image

Depth Image

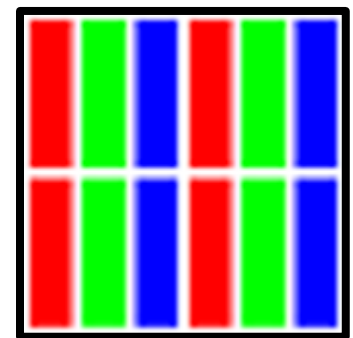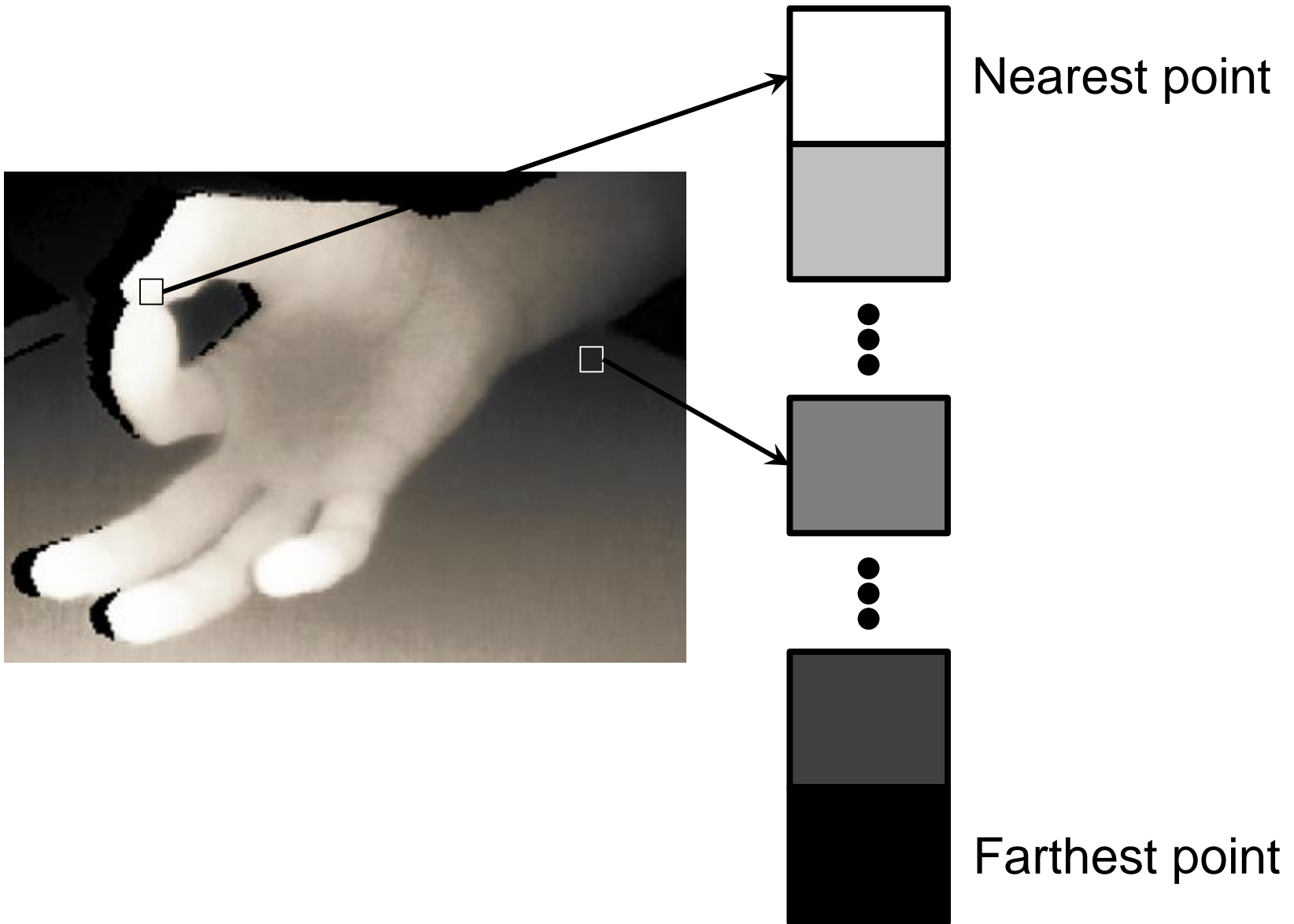*Captured with Intel® RealSense Camera*

Color Image vs. Depth Map

252R, 231G, 217B

Bayer
Pattern

RGB
Pattern

Color Image vs. <u>Depth Map</u>

Nearest point

Farthest point
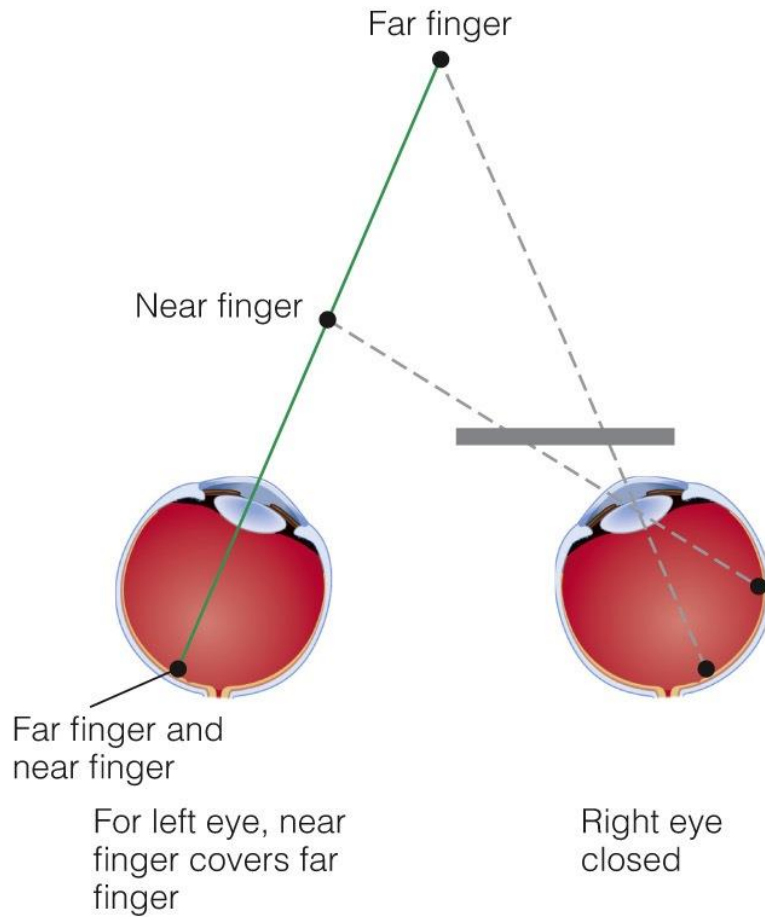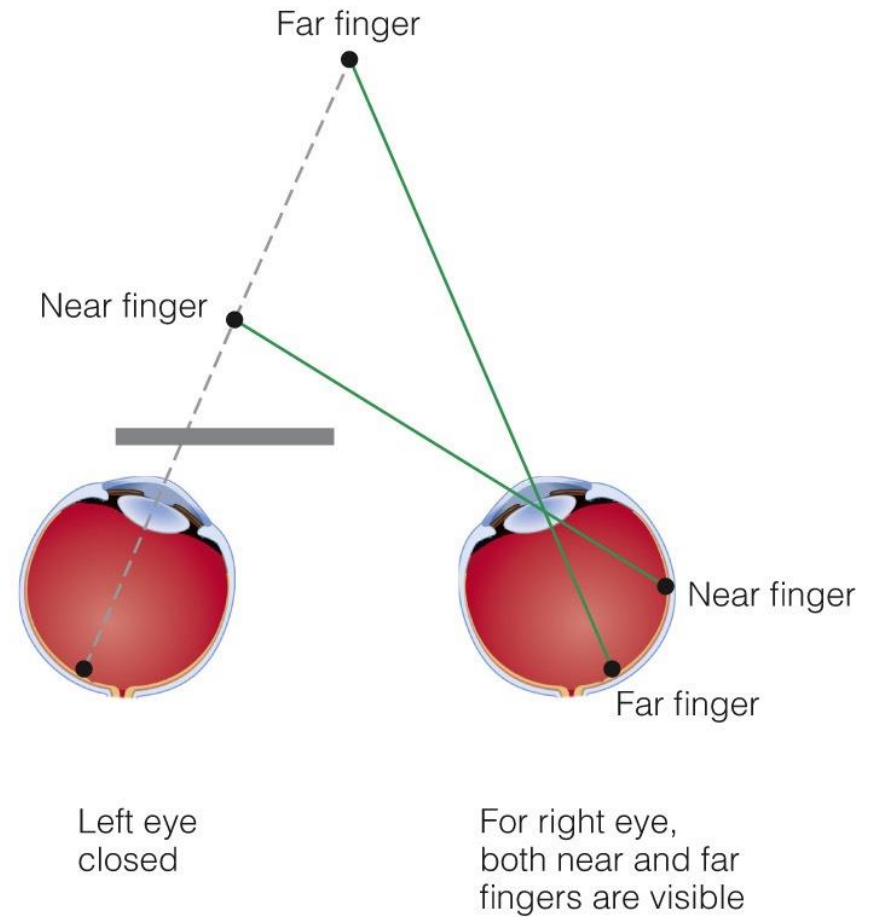
3D/Depth Imaging Techniques

- Stereo-3D Imaging

- Structured/coded light

- Time-of-flight

Binocular Depth Perception

For left eye, near finger covers far finger

Right eye closed
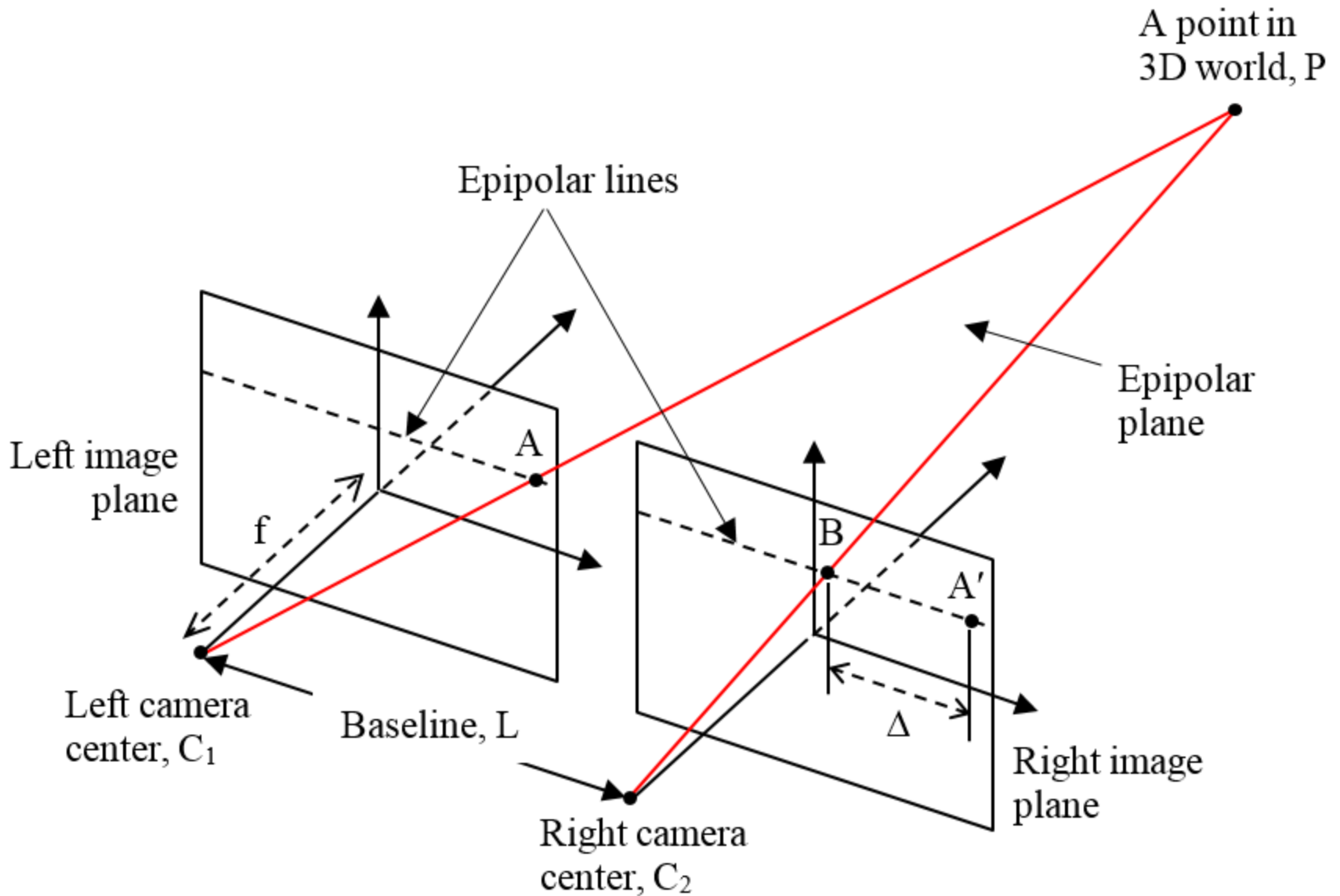
(a)

Left eye closed
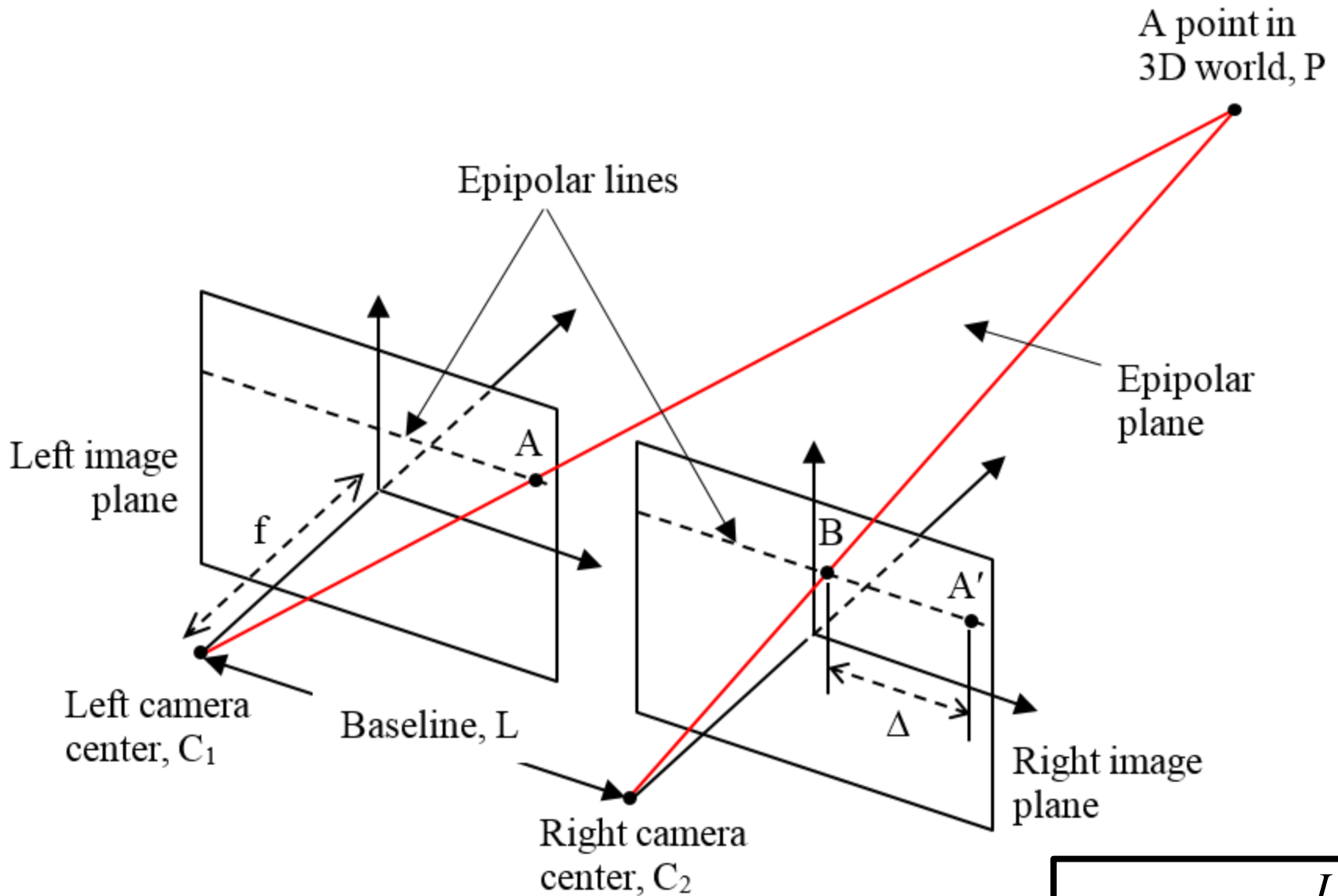
For right eye, both near and far fingers are visible

(b)

Stereo-3D Imaging

Δ = Binocular Disparity

Stereo-3D Imaging

A point in 3D world, P

Epipolar lines

Epipolar plane

Left image plane

f

Left camera center, C_1

Baseline, L

Right camera center, C_2

A

B

A'

Δ

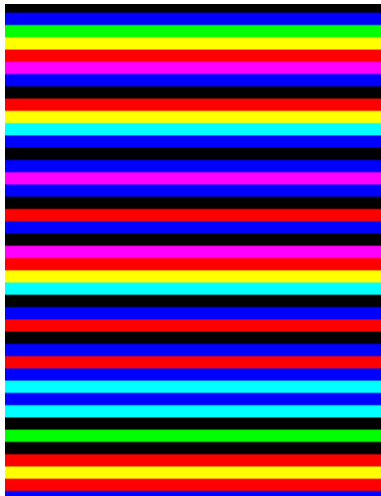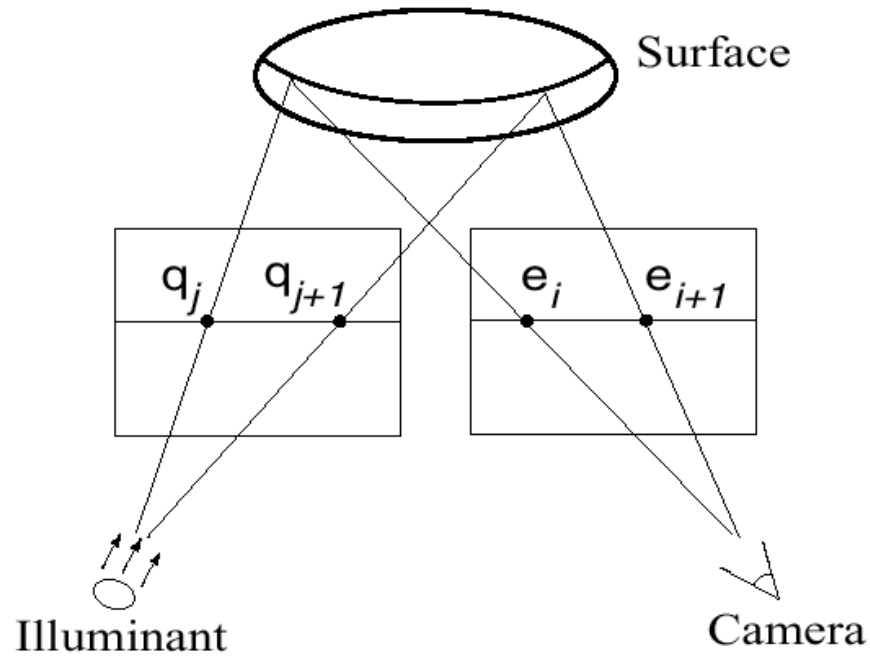Right image plane

Distance, $z = \dfrac{L \cdot f}{\Delta}$

3D imaging with structured light

L. Zhang, B. Curless, and S. M. Seitz. Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming. *3DPVT* 2002
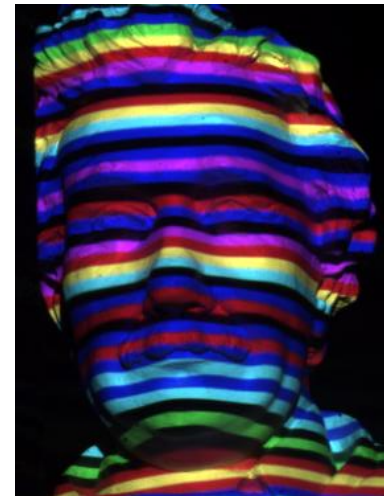
3D imaging with time-of-flight
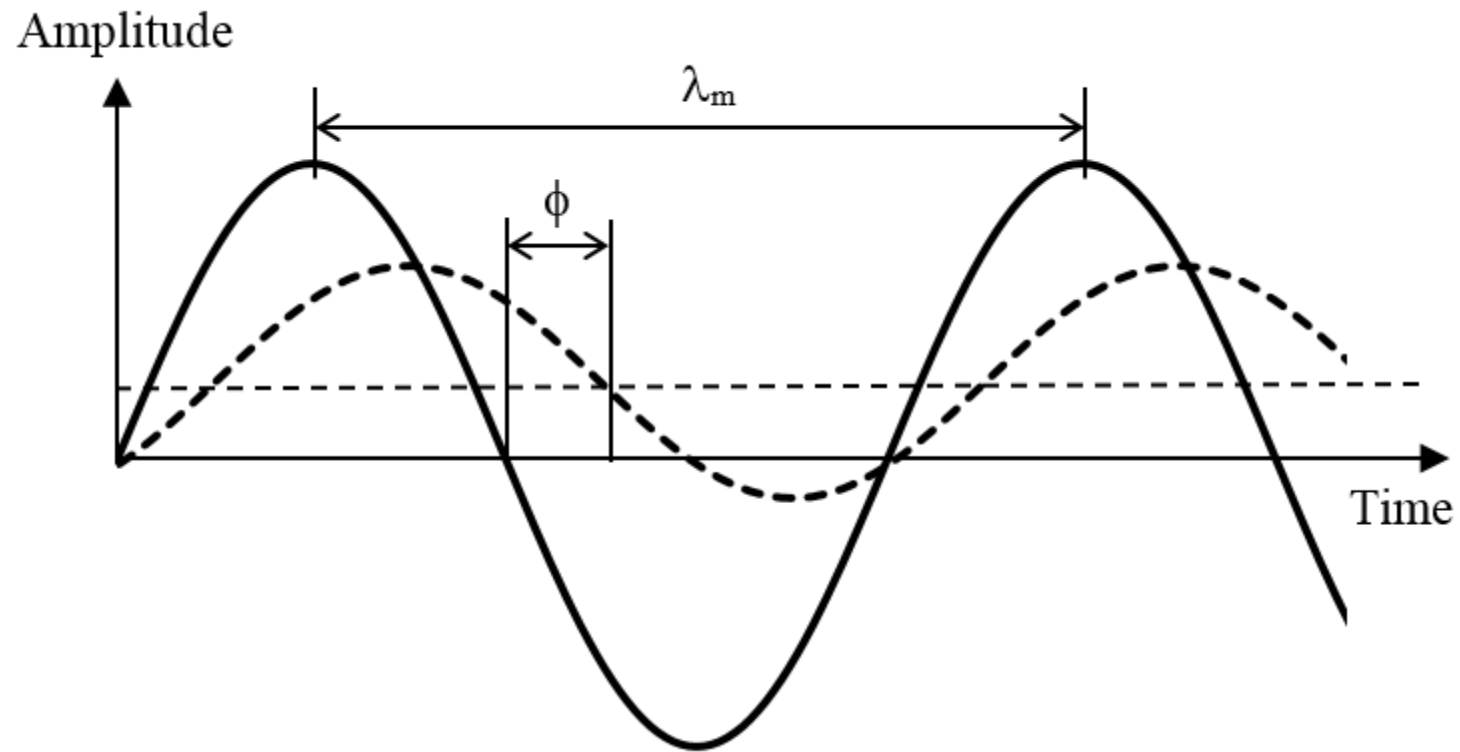
Towards Ubiquitous Proliferation: Key Requirements

- **Usages:** why should people care

- **Form-factor:** small enough for integration

- **Power:** easy on batteries

- **Cost:** affordable to consumers

Towards Ubiquitous Proliferation: Key Requirements

- **Usages:** why should people care

- **Form-factor:** small enough for integration

- **Power:** easy on batteries
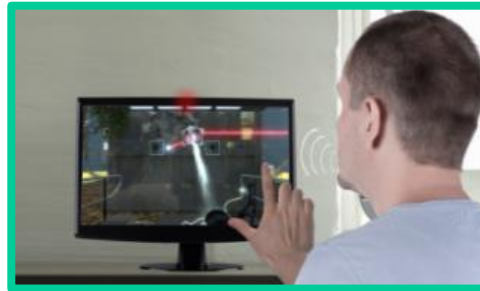
- **Cost:** affordable to consumers

PC Usages ("User-Facing" Config)
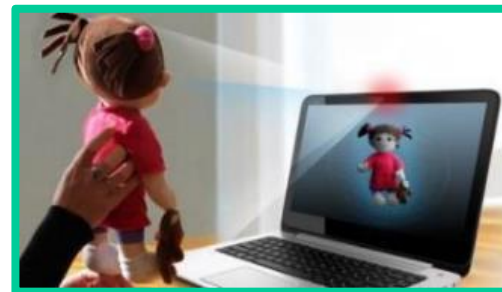
Immersive Collaboration

Gaming and Play

Interact Naturally

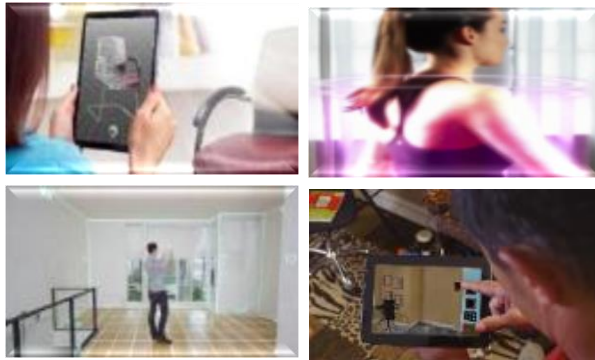**Learning and Edutainment**

Capture and Share

Mobile Usages ("World-Facing" Config)

Capture the World in 3D

Enhanced Photo & Video

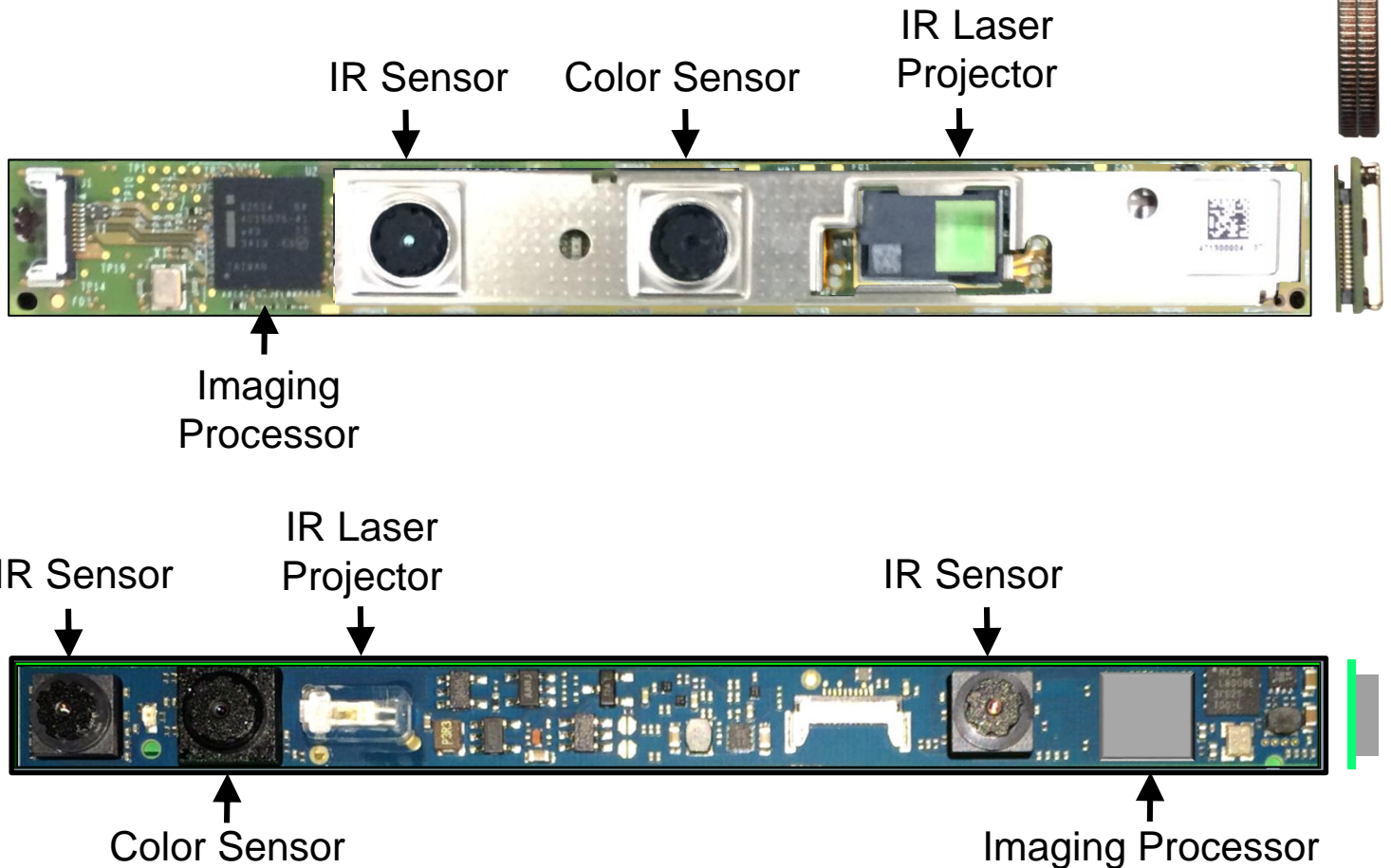### Immersive Gaming

Education & Training

Collaboration

Towards Ubiquitous Proliferation: Key Requirements

- **Usages:** why should people care

- **Form-factor:** small enough for integration

- **Power:** easy on batteries

- **Cost:** affordable to consumers
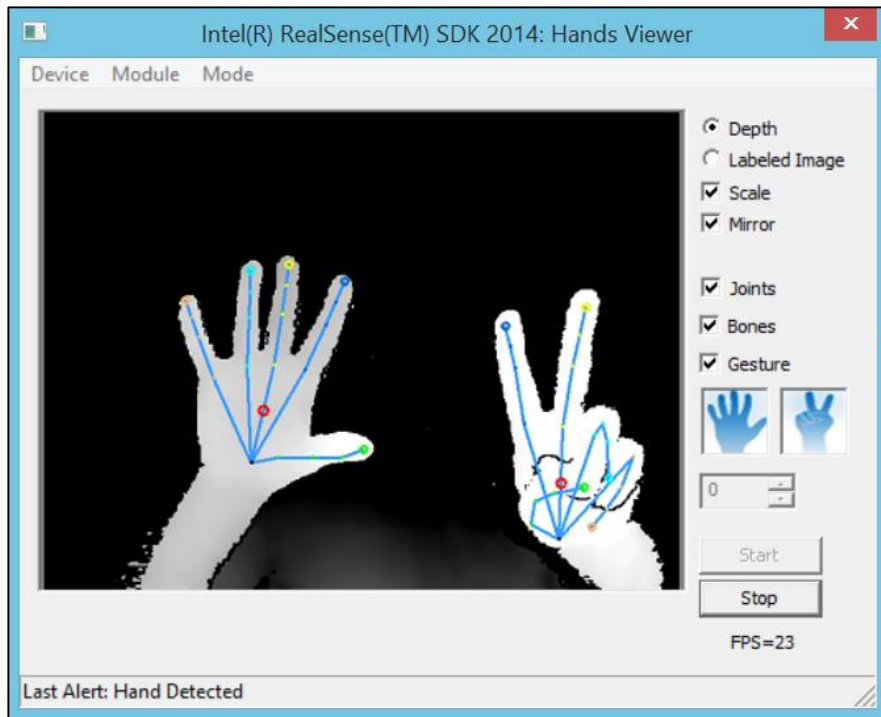
Intel® RealSense Cameras

IR Sensor   Color Sensor   IR Laser Projector

Imaging Processor

IR Sensor   IR Laser Projector   IR Sensor

Color Sensor   Imaging Processor
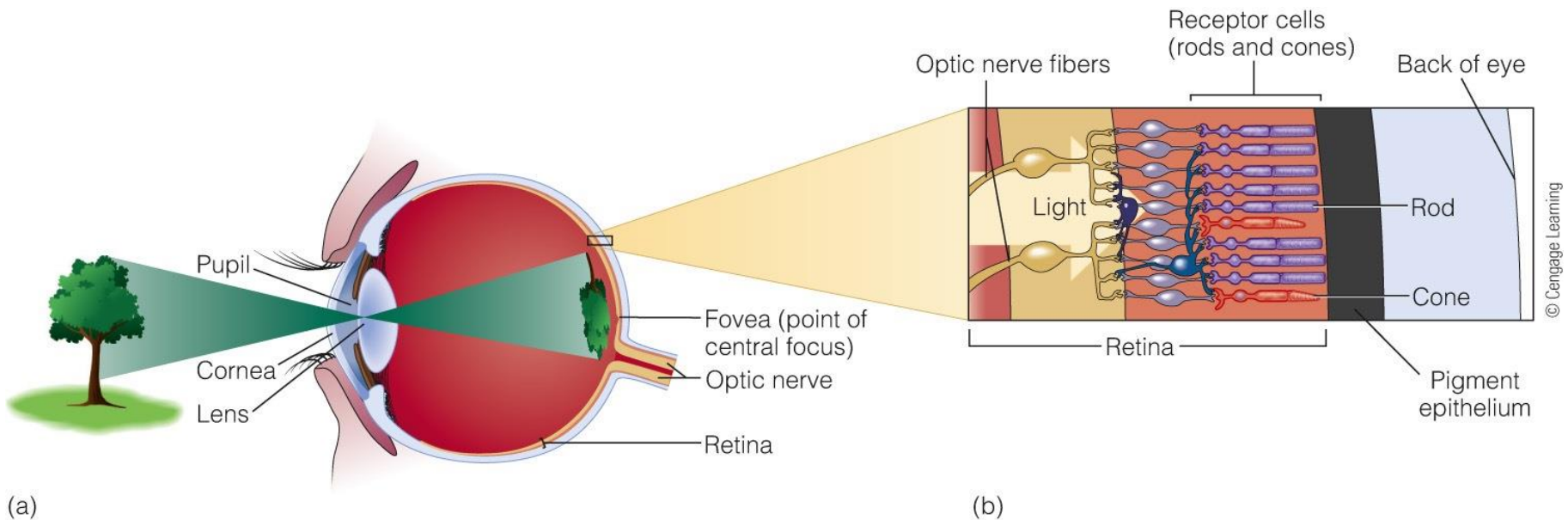
More info/SDK:
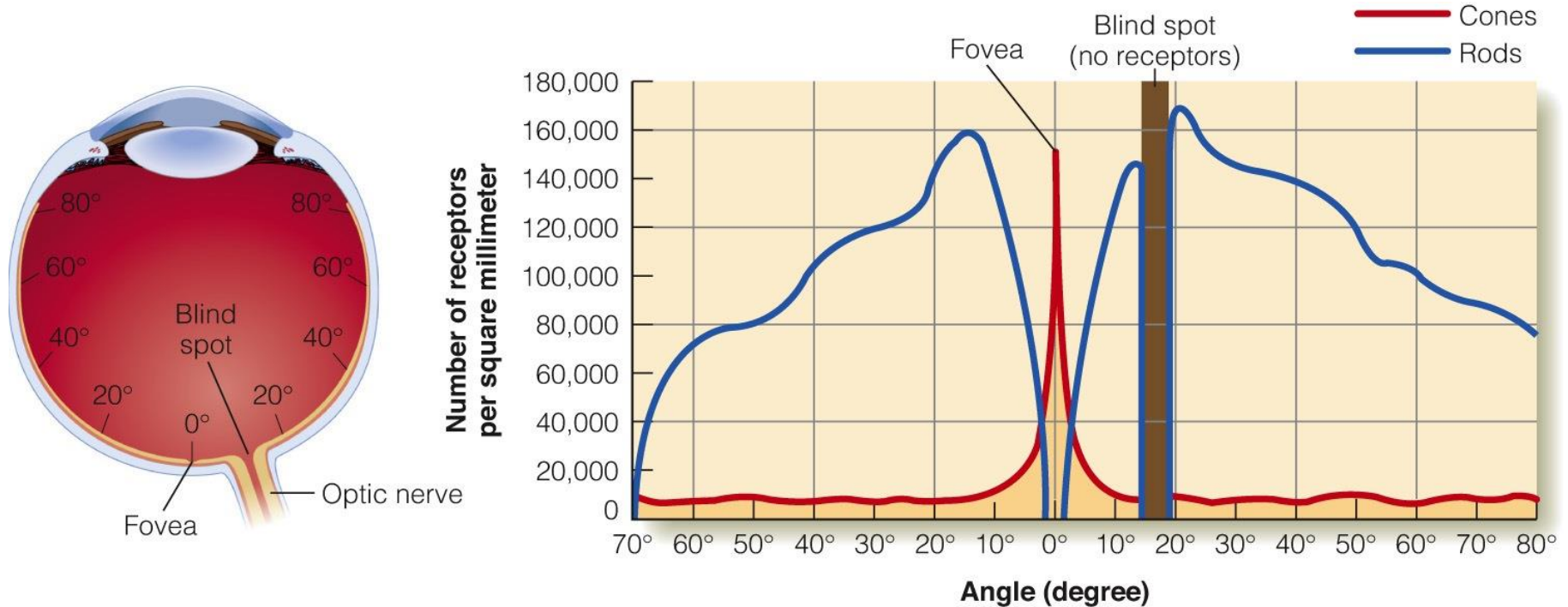
www.intel.com/RealSense

Q&A

Intel® RealSense SDK

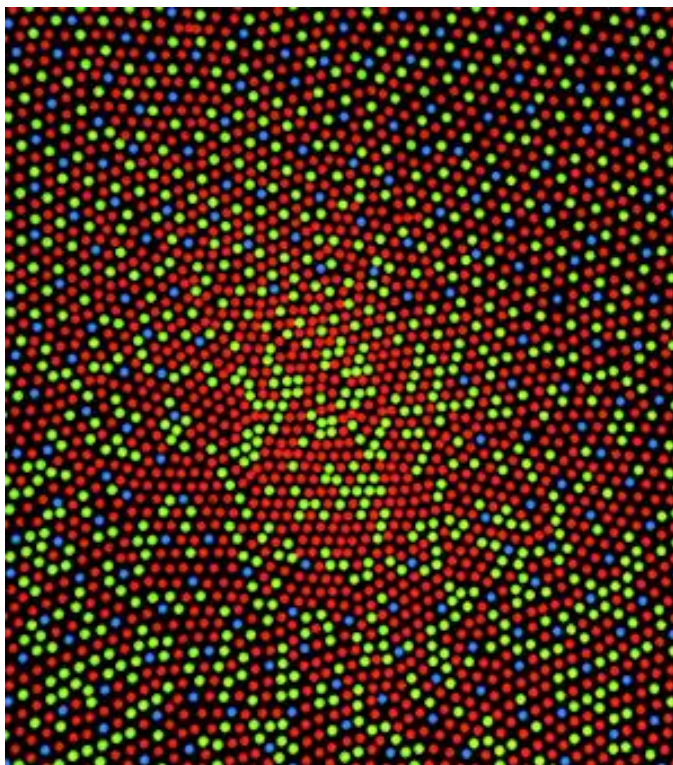# Transforming of Light Energy Into Electrical Energy
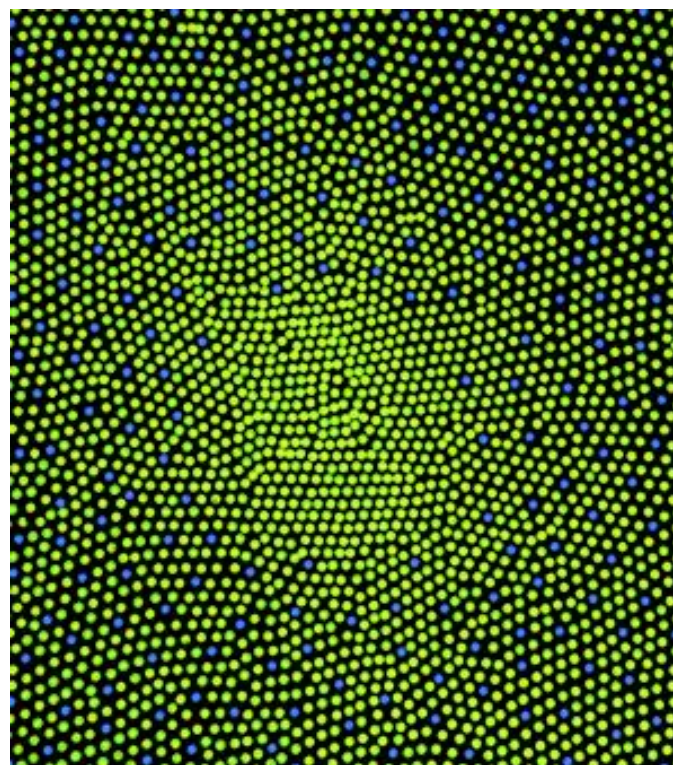
Distribution of Rods and Cones

Adapted from *Human Information Processing*, by P. Lindsay and D. A. Norman, 1977, 2nd ed., p. 126. Copyright © 1977 Academic Press, Inc. Adapted with permission.

- Fovea consists solely of cones.
- Peripheral retina has both rods and cones.
- More rods than cones in periphery.

Normal Vision   Color Blind

Ponzo Illusion

Courtesy of Mary Bravo

William Vann/www.edupic.net