Lecture 23:

Intro to Virtual Reality

Computer Graphics and
Imaging UC Berkeley
CS184/284A

Virtual Reality (VR) vs Augmented Reality (AR)

VR = virtual reality

- User is completely immersed in virtual world (sees only light emitted by display)

AR = augmented reality

- Display is an overlay that augments user's normal view of the real world (e.g., Terminator)

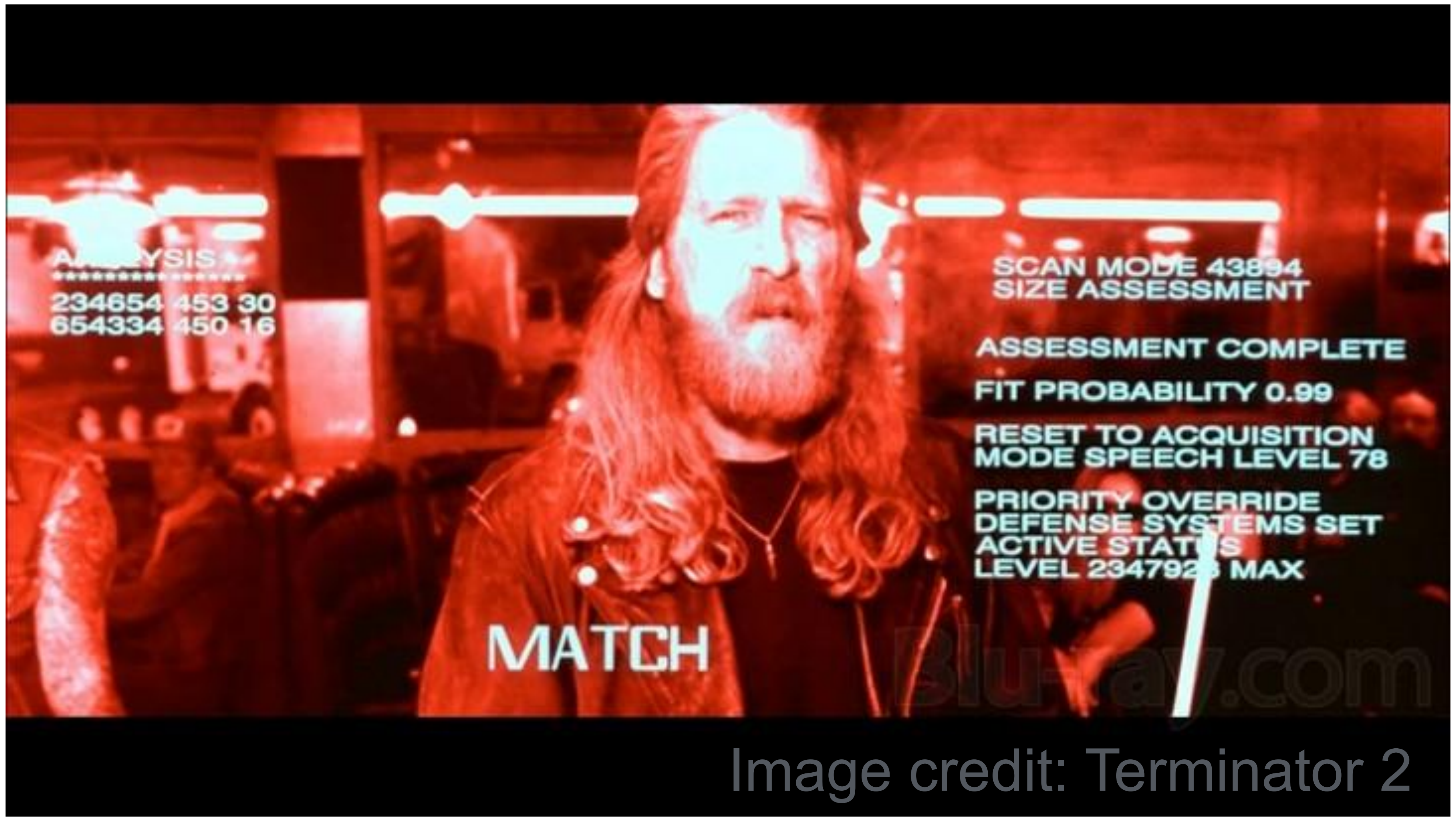


Image credit: Terminator 2

MR = mixed reality

- Blend of VR and AR

Ivan Sutherland's Virtual Reality Research in 1968

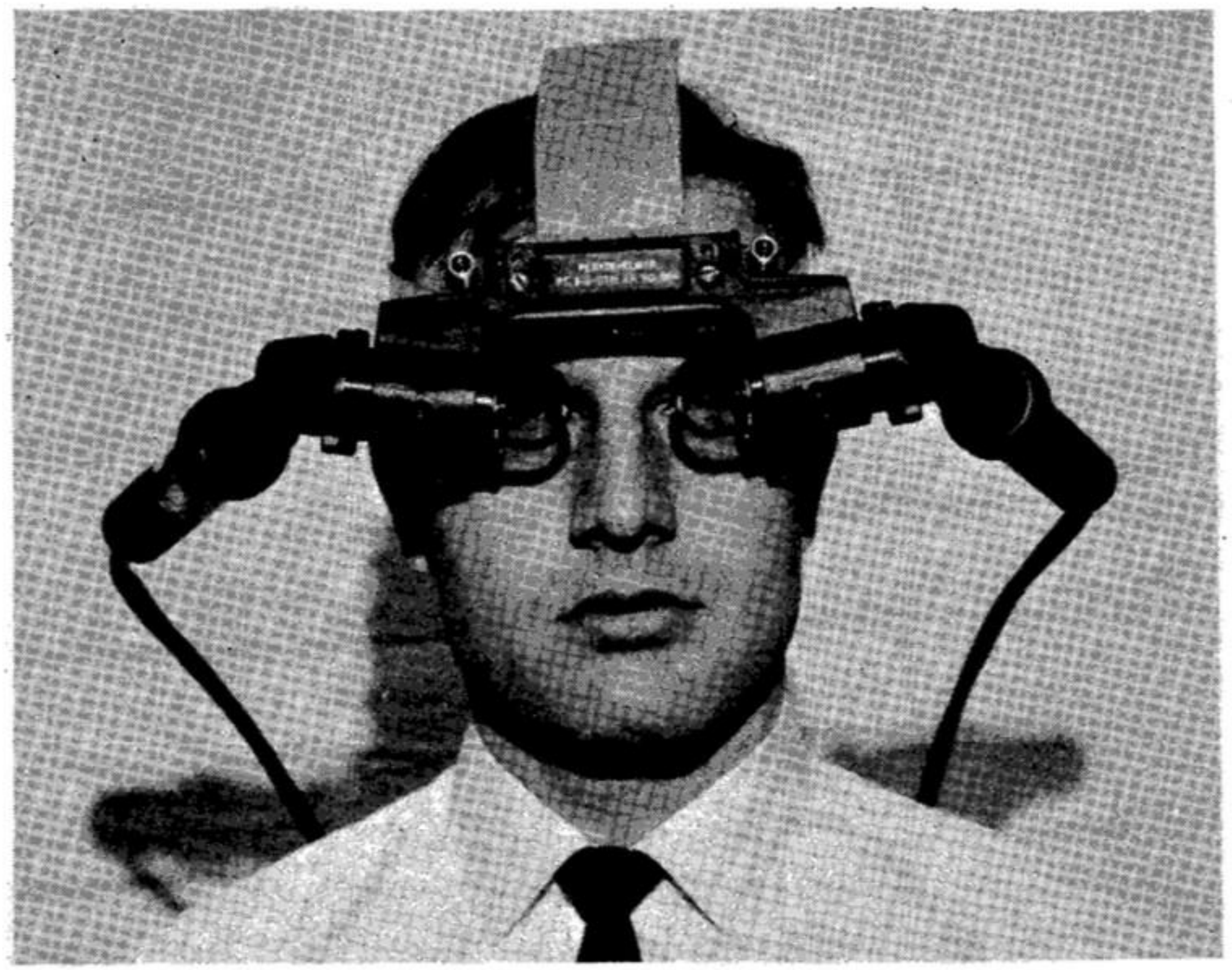


FIGURE 2—The head-mounted display optics with miniature CRT's

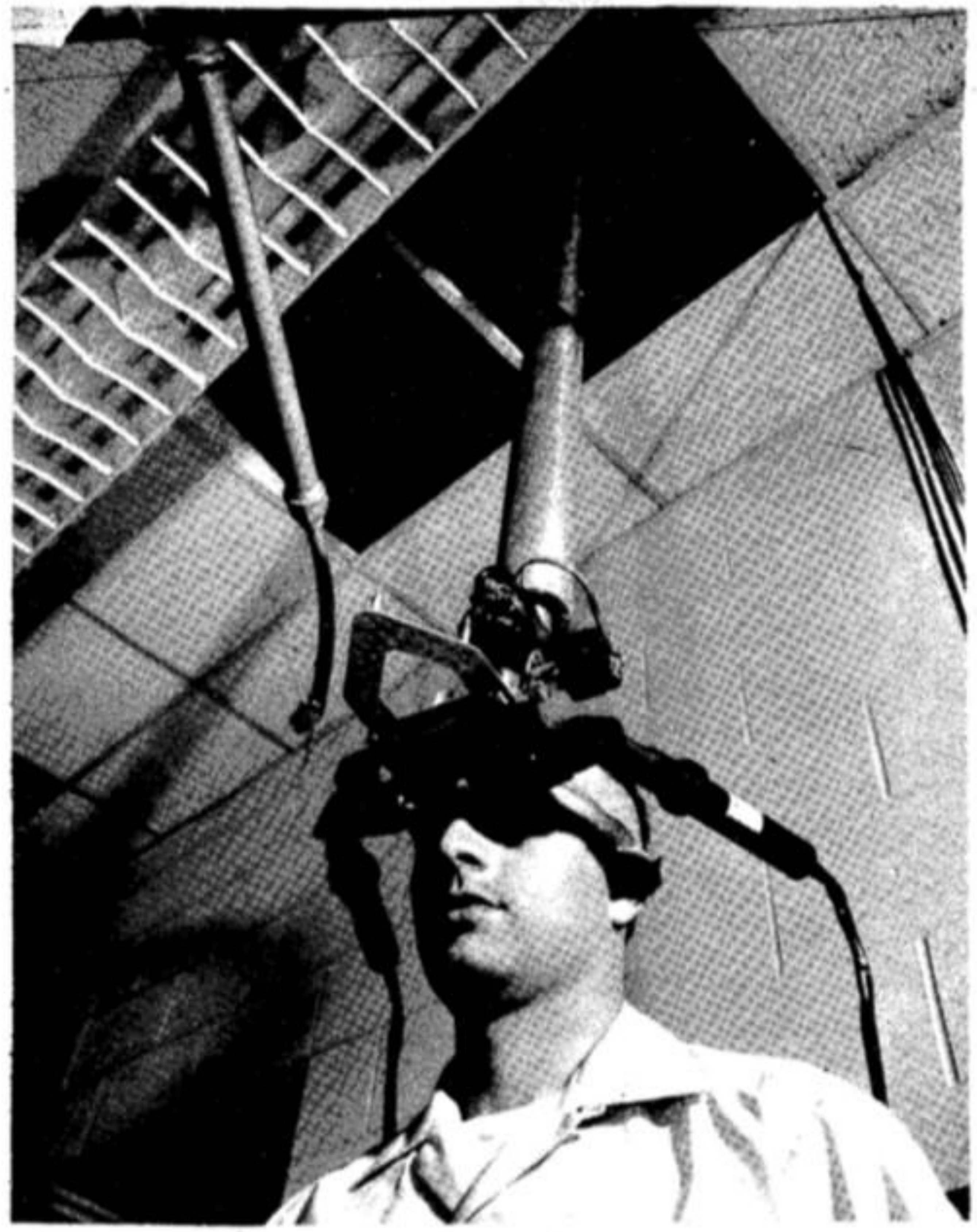


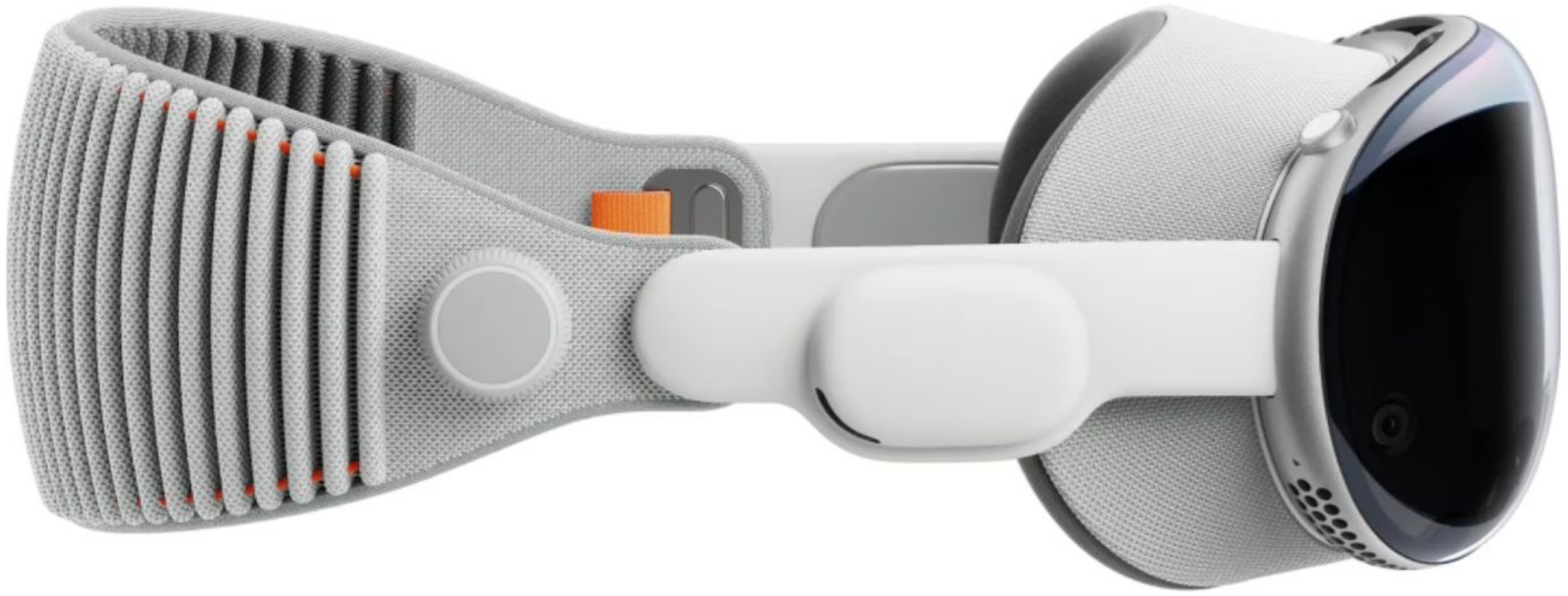
FIGURE 4—The ultrasonic head position sensor in use

VR Head-Mounted Displays (HMDs)

Meta Quest Sony

HTC
Vive

Apple Vision
Pro

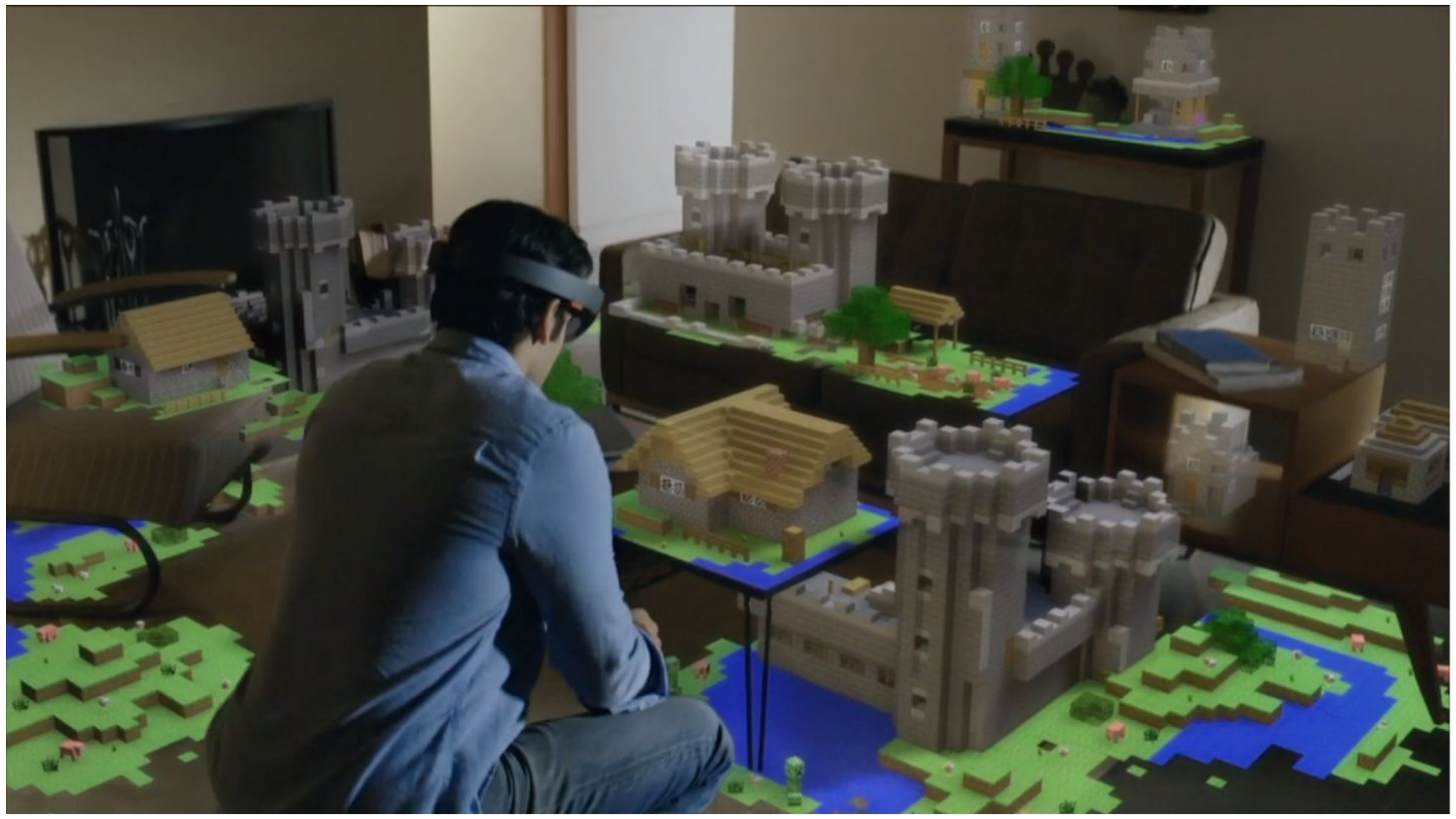


Also Valve Index, HP Reverb, etc…

AR Headsets

Microsoft Hololens

Magic Leap



“Mixed Reality” - VR with Passthrough Imaging

vive.com

Opaque VR headset, outward cameras, live imagery of real world + virtual 3D graphics

VR
Applications

VR Gaming

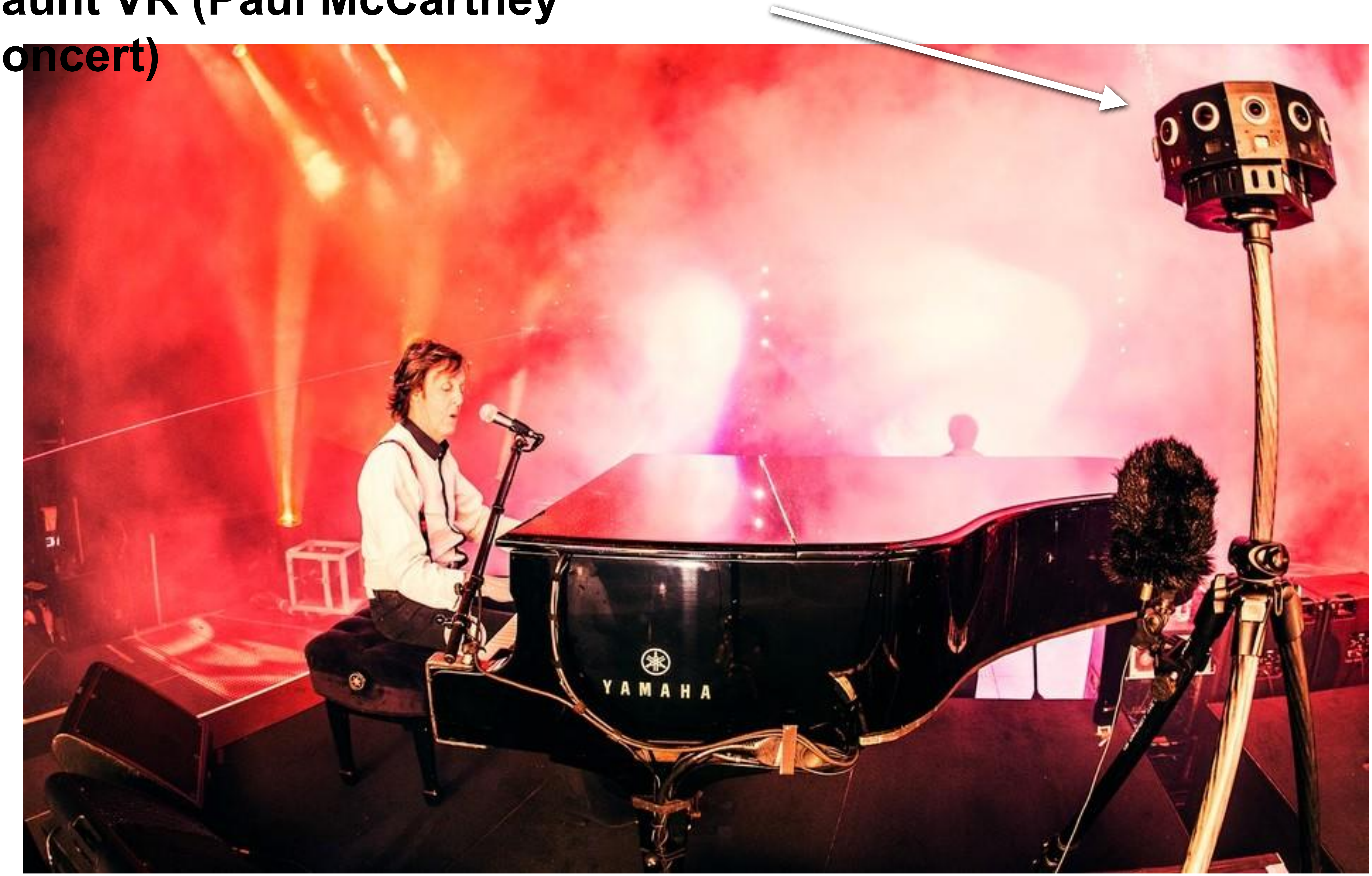
Star Wars Squadrons (EA)

VR Painting

Tilt Brush

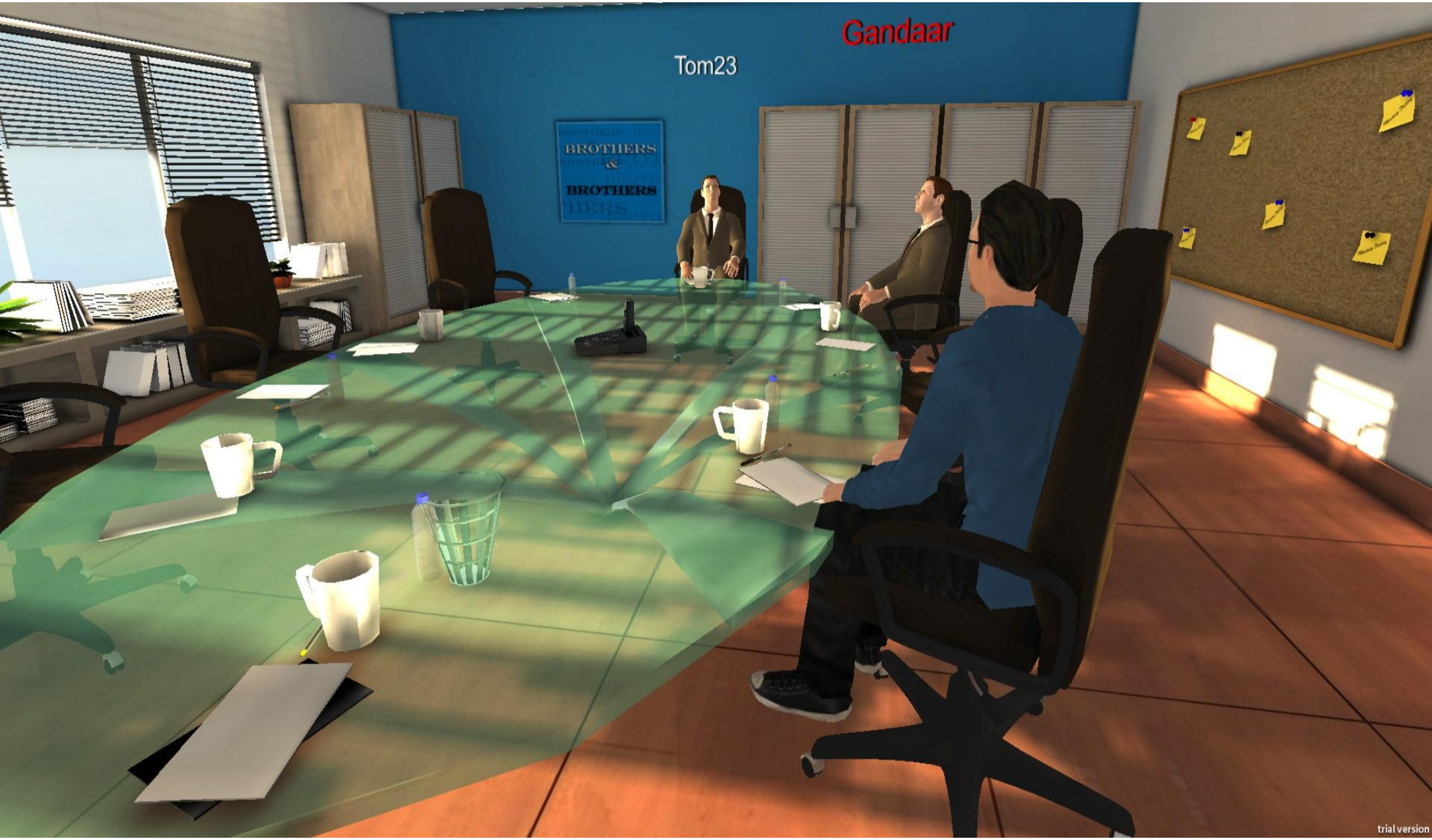
VR Video

Jaunt VR (Paul McCartney concert)



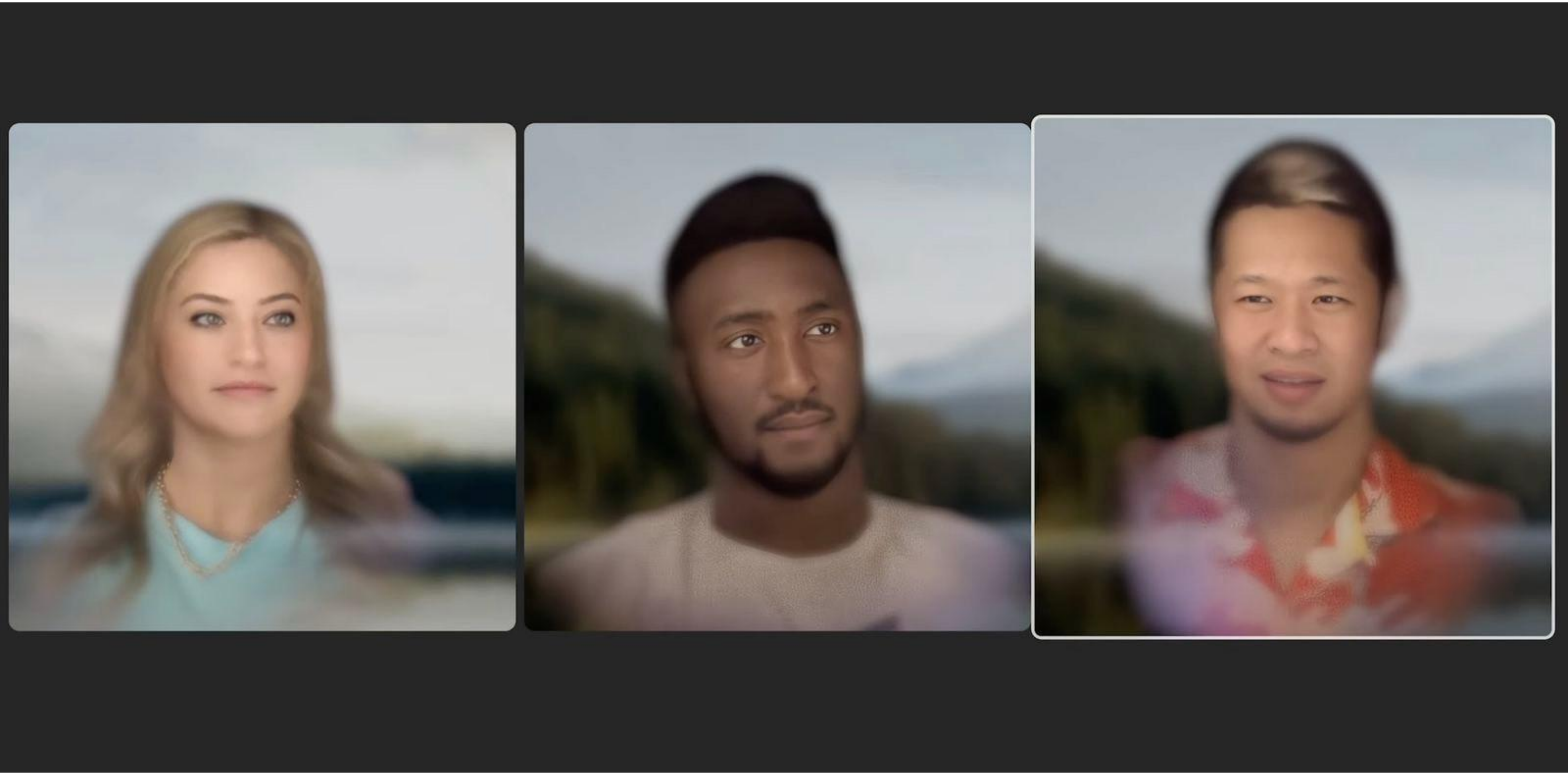
VR Video

VR Teleconference / Video Chat



http://vrchat.com/

VR Teleconference / Video Chat



Apple Vision Pro Personas

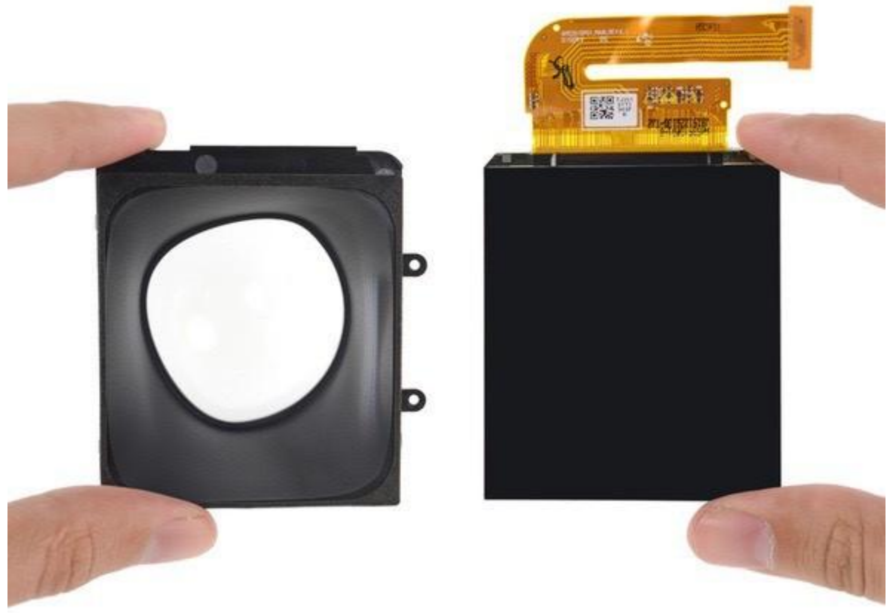
Image credit: Brian Tong

The Intimacy of VR Graphics

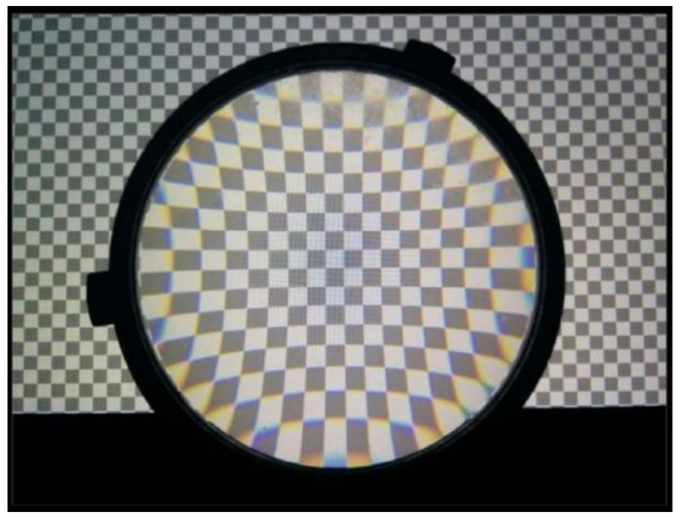
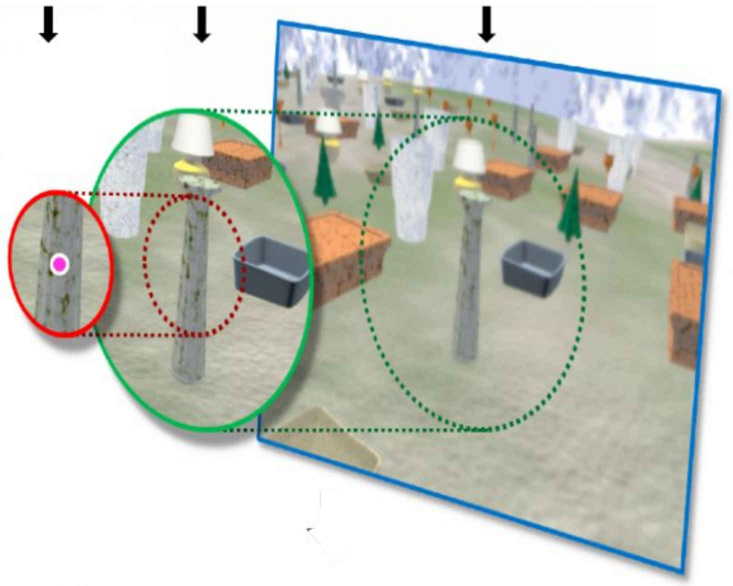
Google’s Tilt Brush on HTC Vive

Overview of VR Topics

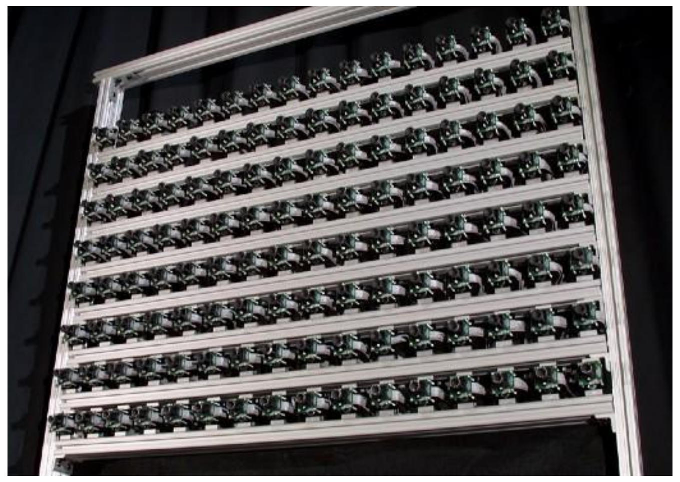
- VR Displays



- VR Rendering

- VR Imaging



VR
Displays

Field of View

Regular 2D panel displays have windowed FOV

- User orients themselves to the physical window of the display

VR/AR displays provide 360 degree FOV

- Displays attached to head
- Head orientation is tracked physically
- Rendered view synchronized to head orientation in realtime (much more on this later)

3D Visual Cues

Panel displays give 3D cues from monocular rendering

- Occlusion, perspective, shading, focus blur, …
    - Uses z-buffer, 4x4 matrices, lighting calculation, lens calculations…

VR/AR displays add further 3D cues

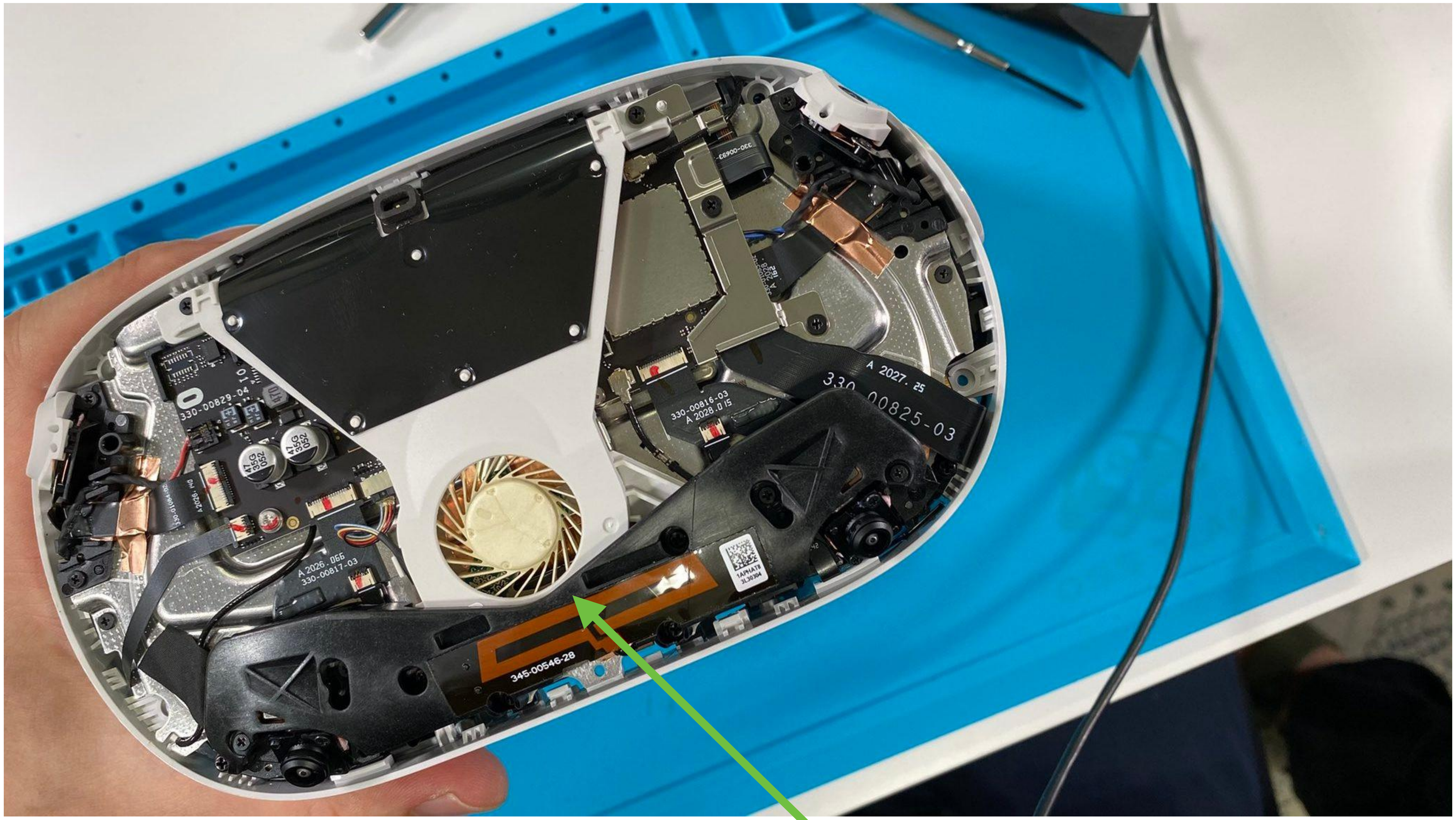
- Stereo: different perspective view in left/right eyes
    - Physically send different images into each eye
- Parallax (user-motion): different views as user moves
    - Uses head-tracking technology coupled to perspective rendering

VR Headset Components

Oculus Quest 2 Headset (2020)

Oculus Quest 2 Headset

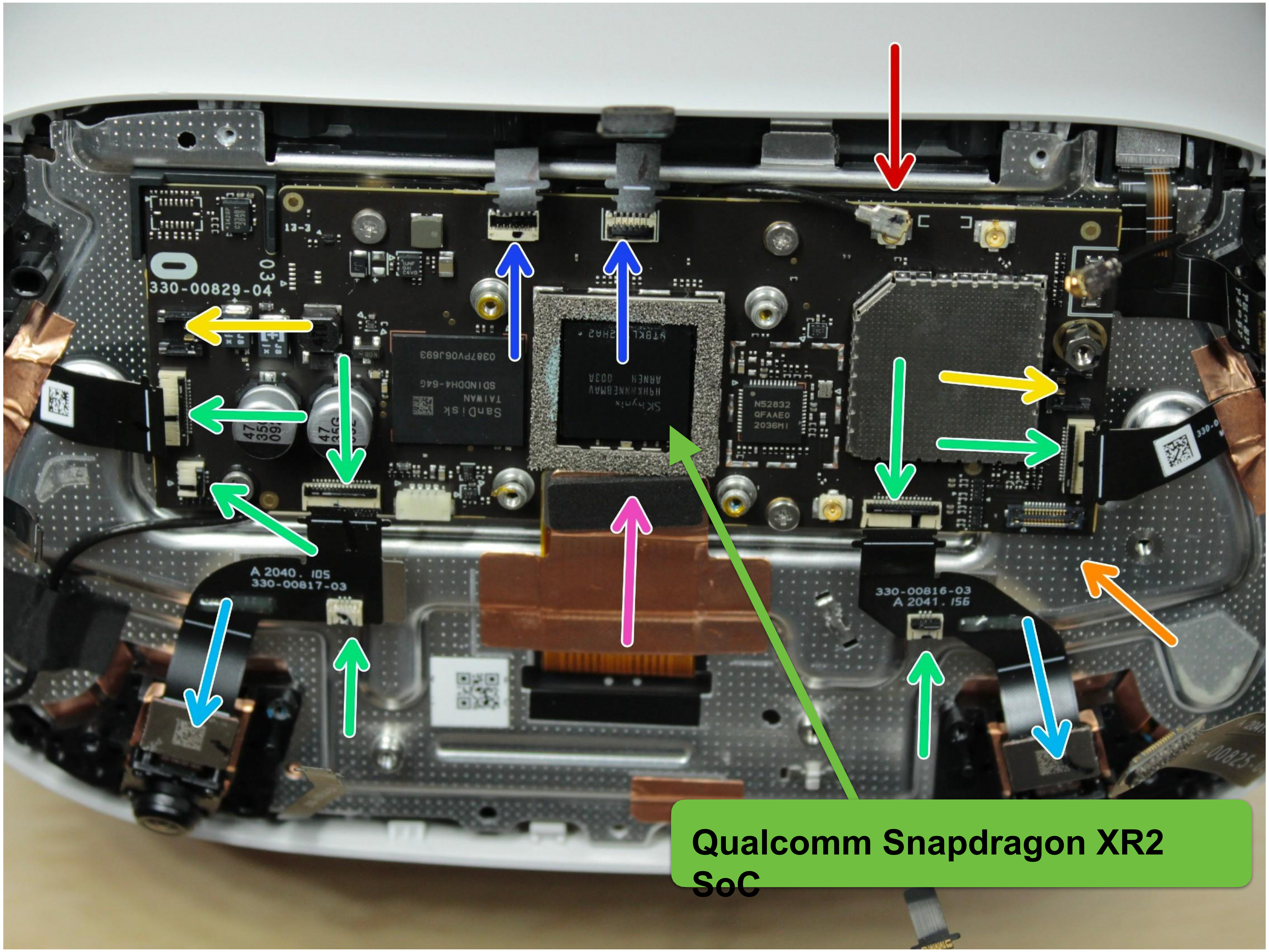
Oculus Quest 2 Headset



Fan

Image credit: ifixit.com

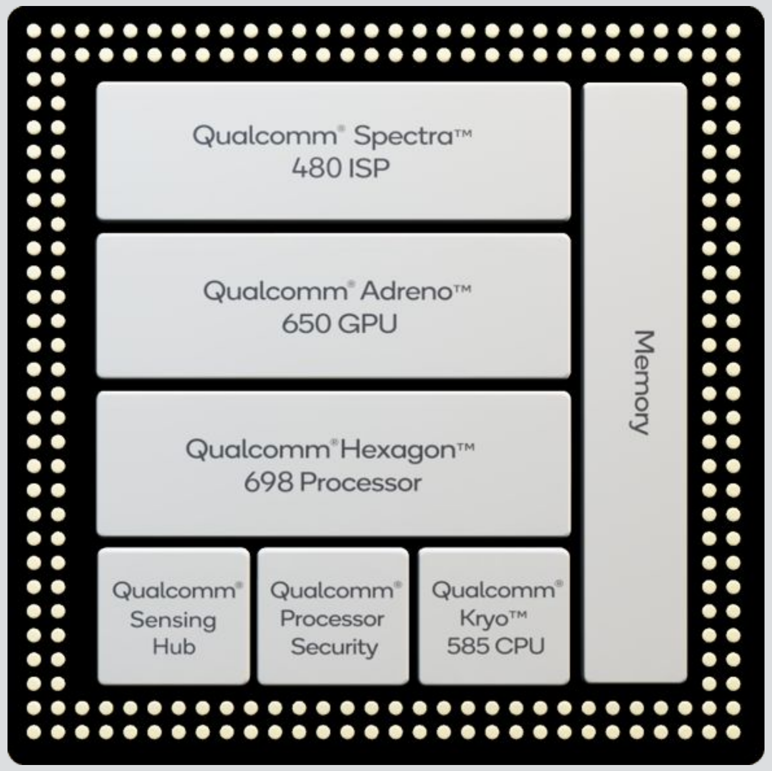
Oculus Quest 2 Headset



Qualcomm Snapdragon XR2 SoC

Image credit: ifixit.com

Oculus Quest 2 Headset (Snapdragon SoC)



Qualcomm® Spectra™ 480 ISP

Qualcomm® Adreno™ 650 GPU

Qualcomm® Hexagon™ 698 Processor

Qualcomm® Sensing Hub

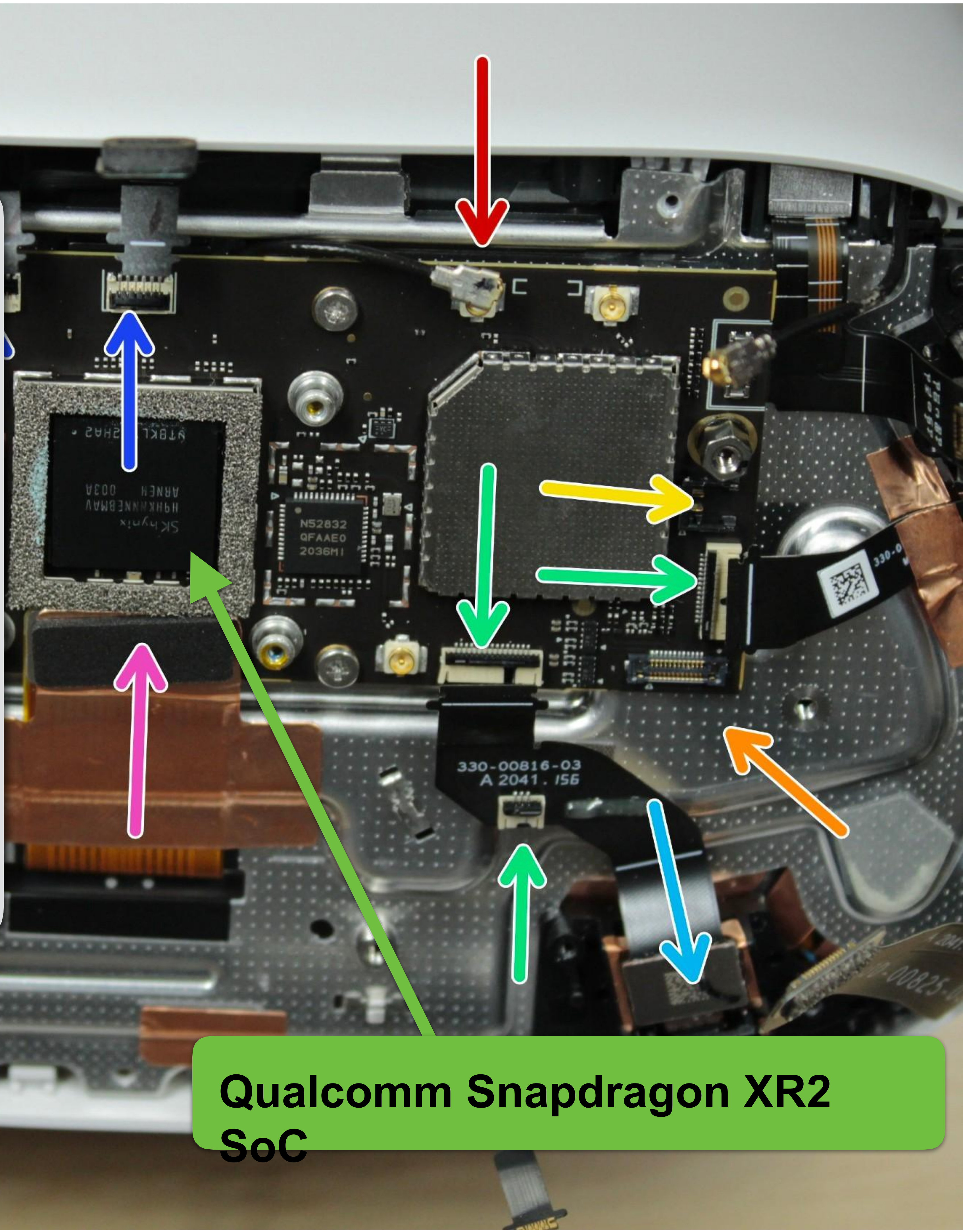
Qualcomm® Processor Security

Qualcomm® Kryo™ 585 CPU

Memory

This diagram is from Snapdragon 865

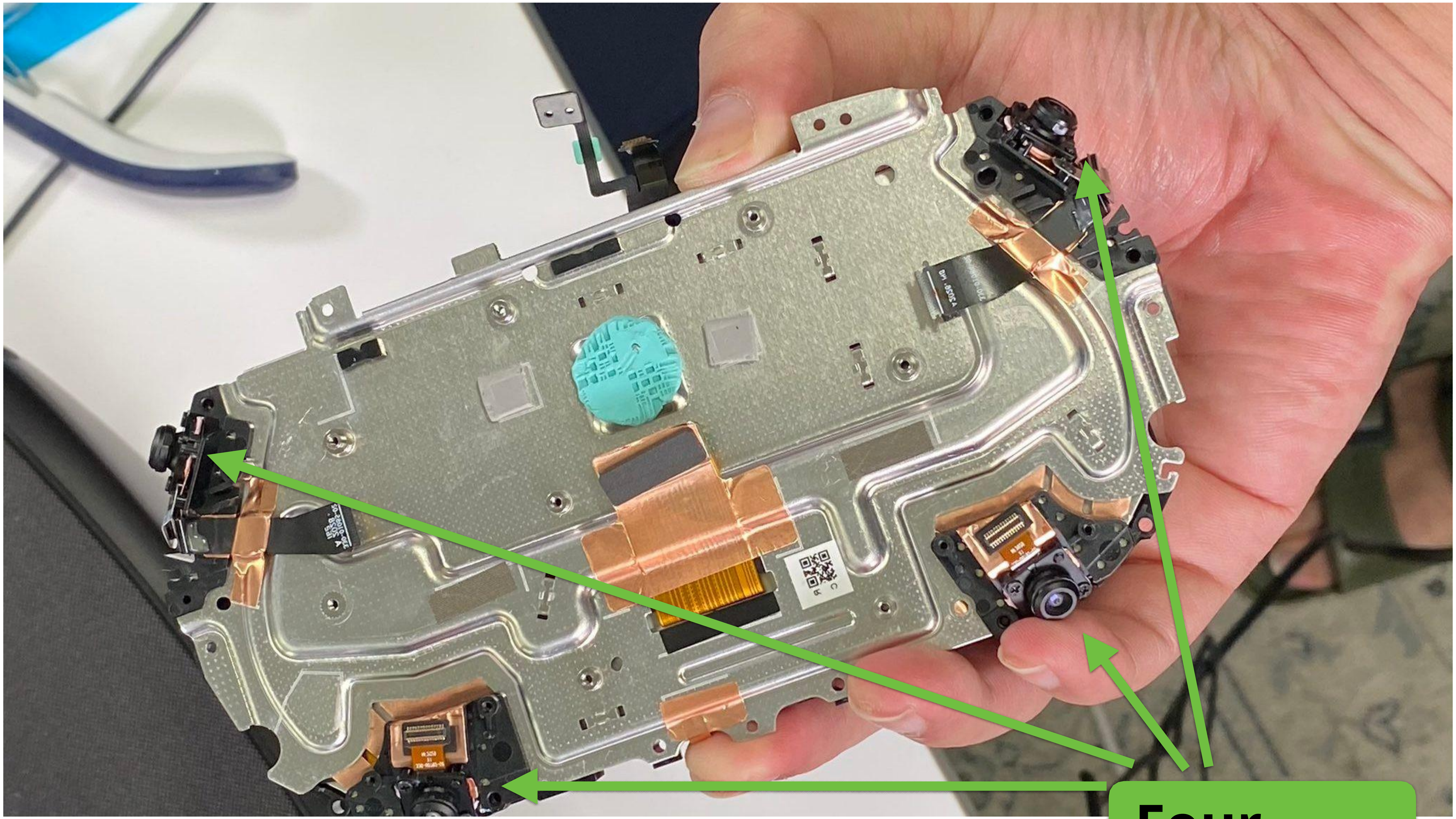
4 high-performance cores
4 low-performance (low energy) cores
Image processor + DSP
Multi-core graphics processor (GPU) — up to 3000 x 3000 display @ 90
Hz Additional processor for sensors (IMU etc)
Can process inputs from up to seven simultaneous video camera
streams



Qualcomm Snapdragon XR2 SoC

Image credit: ifixit.com

Oculus Quest 2 Headset



Four cameras

Image credit: ifixit.com

Oculus Quest 2 Headset (Lens Assembly)

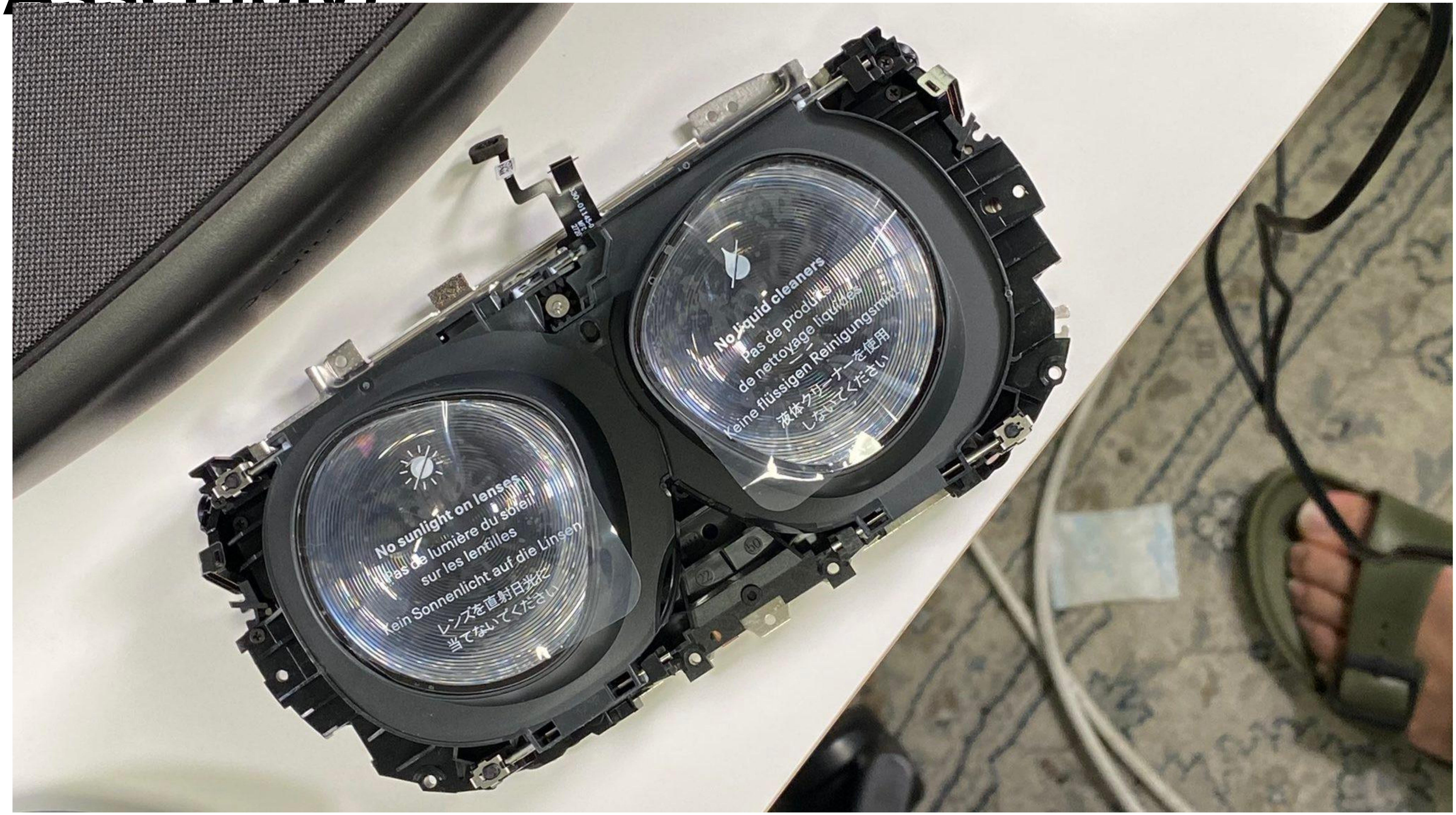
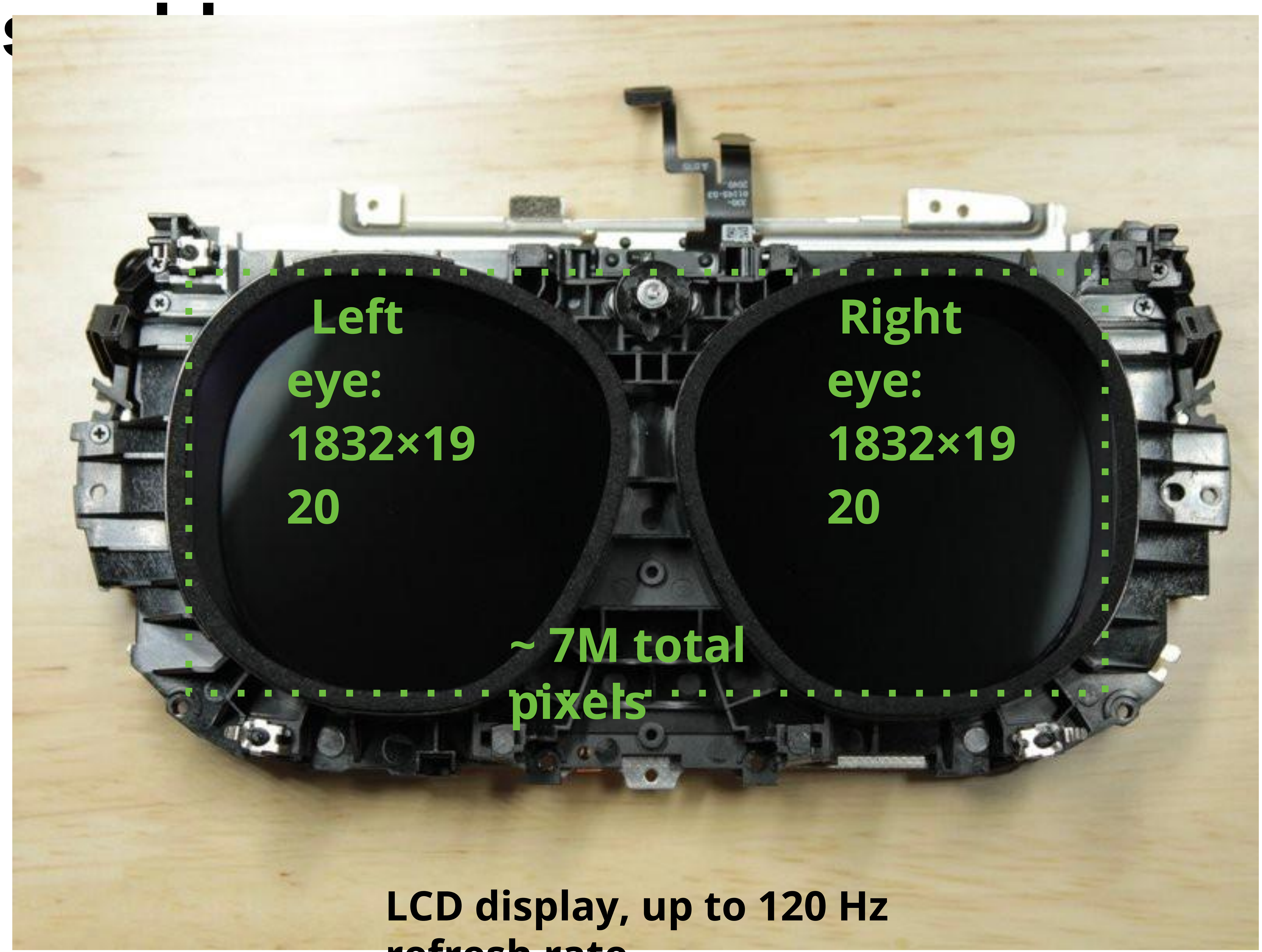


Image credit: ifixit.com

Oculus Quest 2 Display + Lens Ass



Left eye: 1832×1920

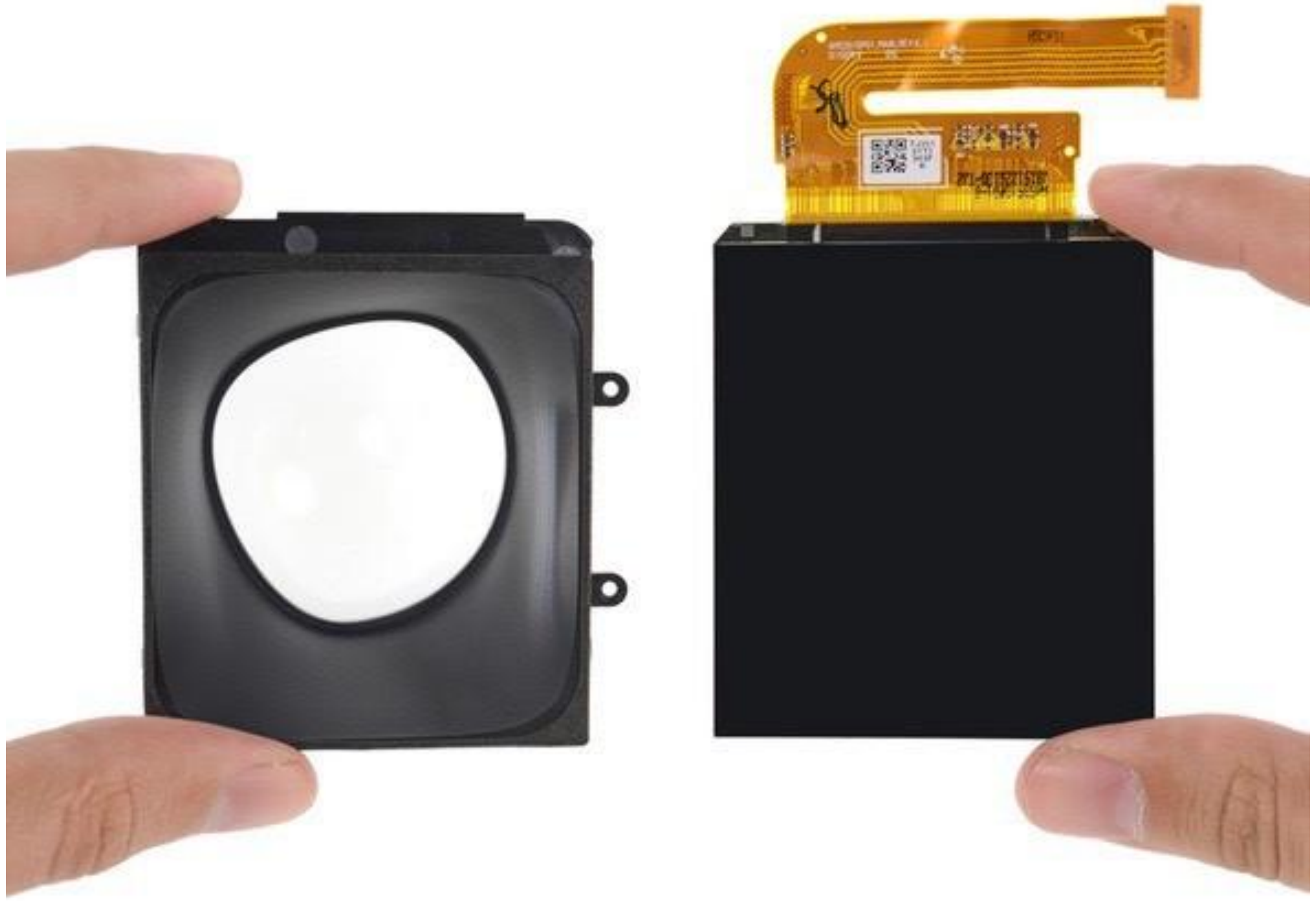
Right eye: 1832×1920

~ 7M total pixels

LCD display, up to 120 Hz refresh rate.

Image credit: ifixit.com

Oculus Rift (2016)



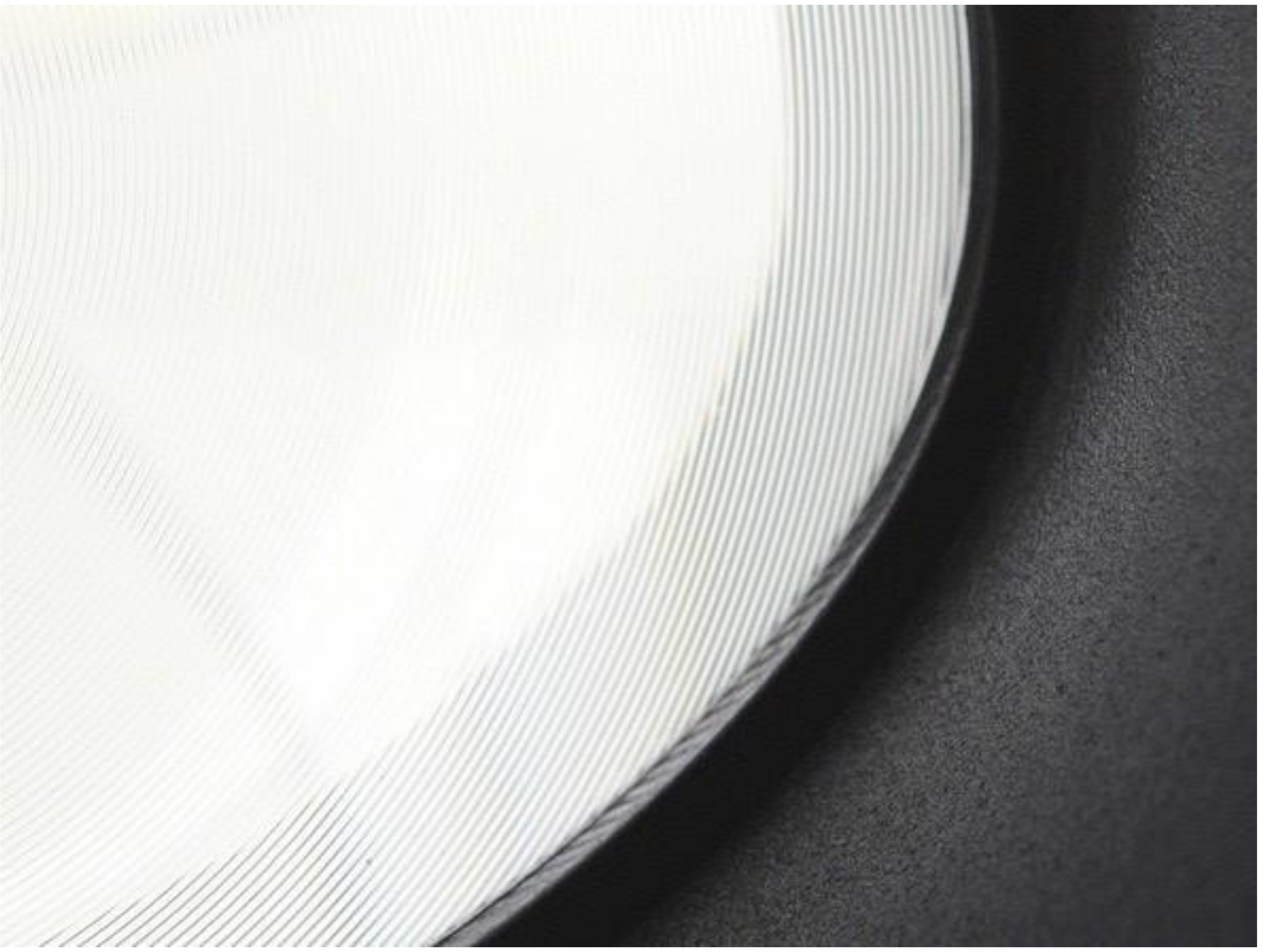
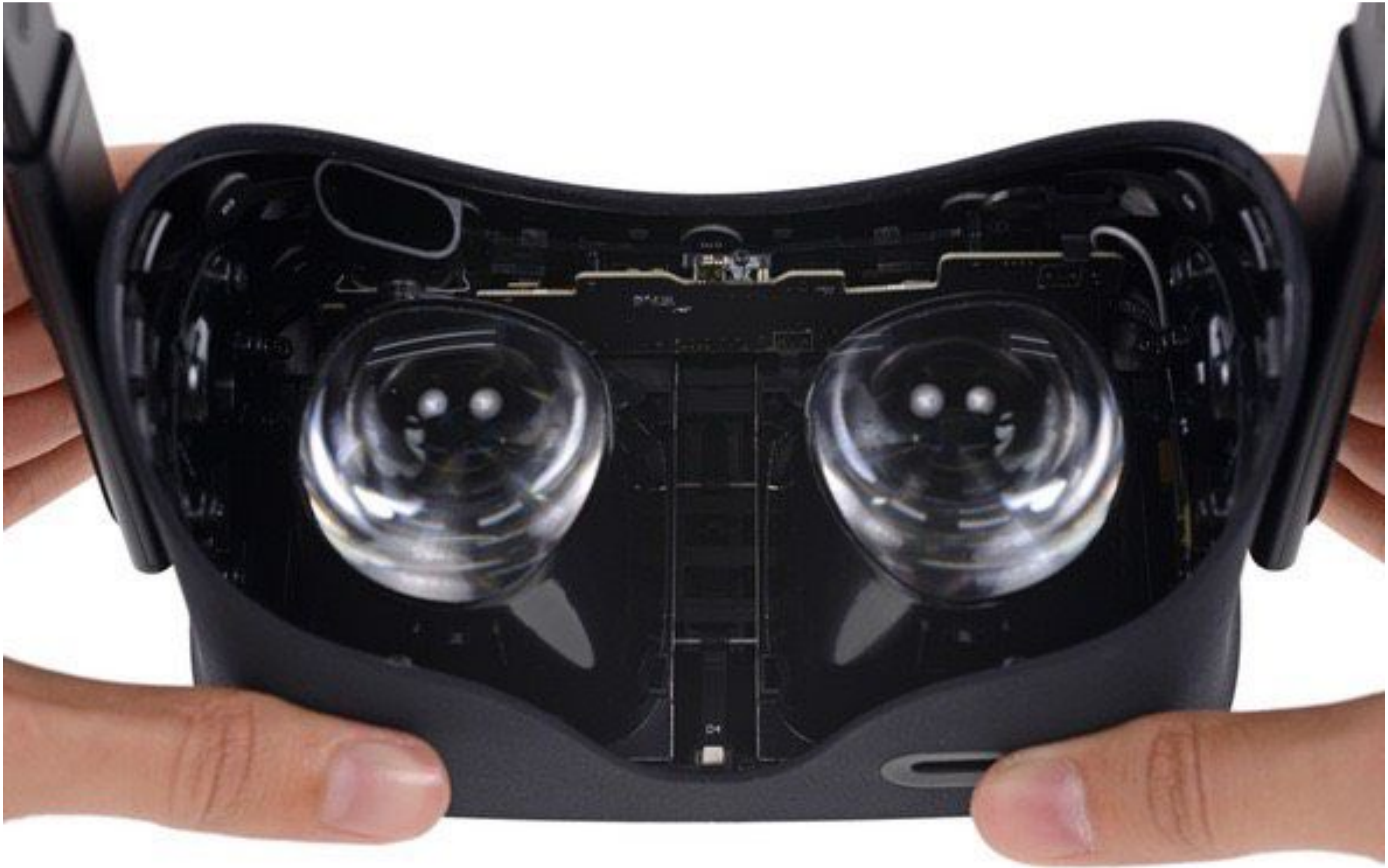


Image credit: ifixit.com https://www.ifixit.com/Teardown/Oculus+Rift+CV1+Teardown/60612

Oculus Rift (2016)



Intra-ocular distance adjustment

Image credit: ifixit.com

Oculus Rift (2016)

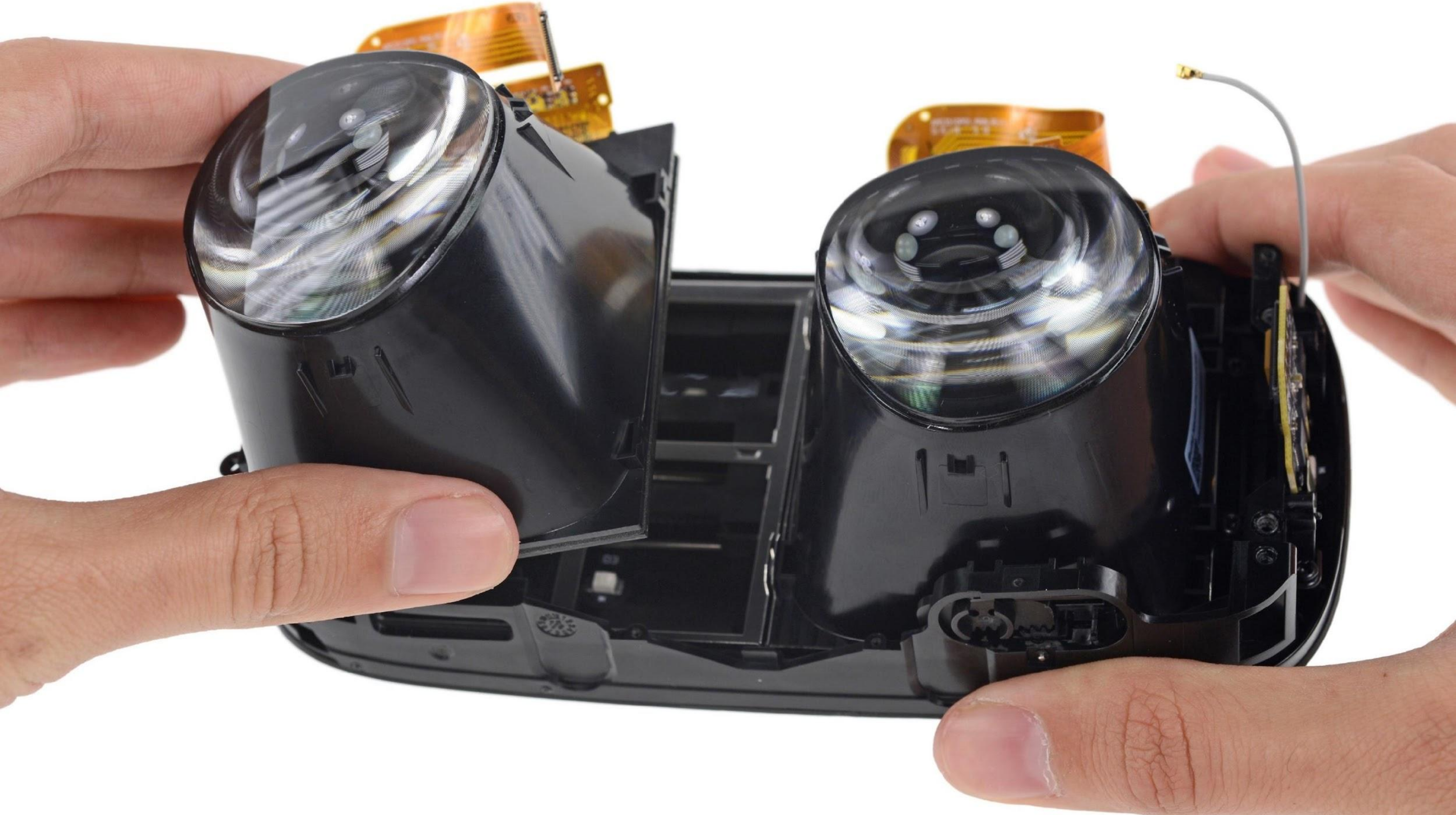
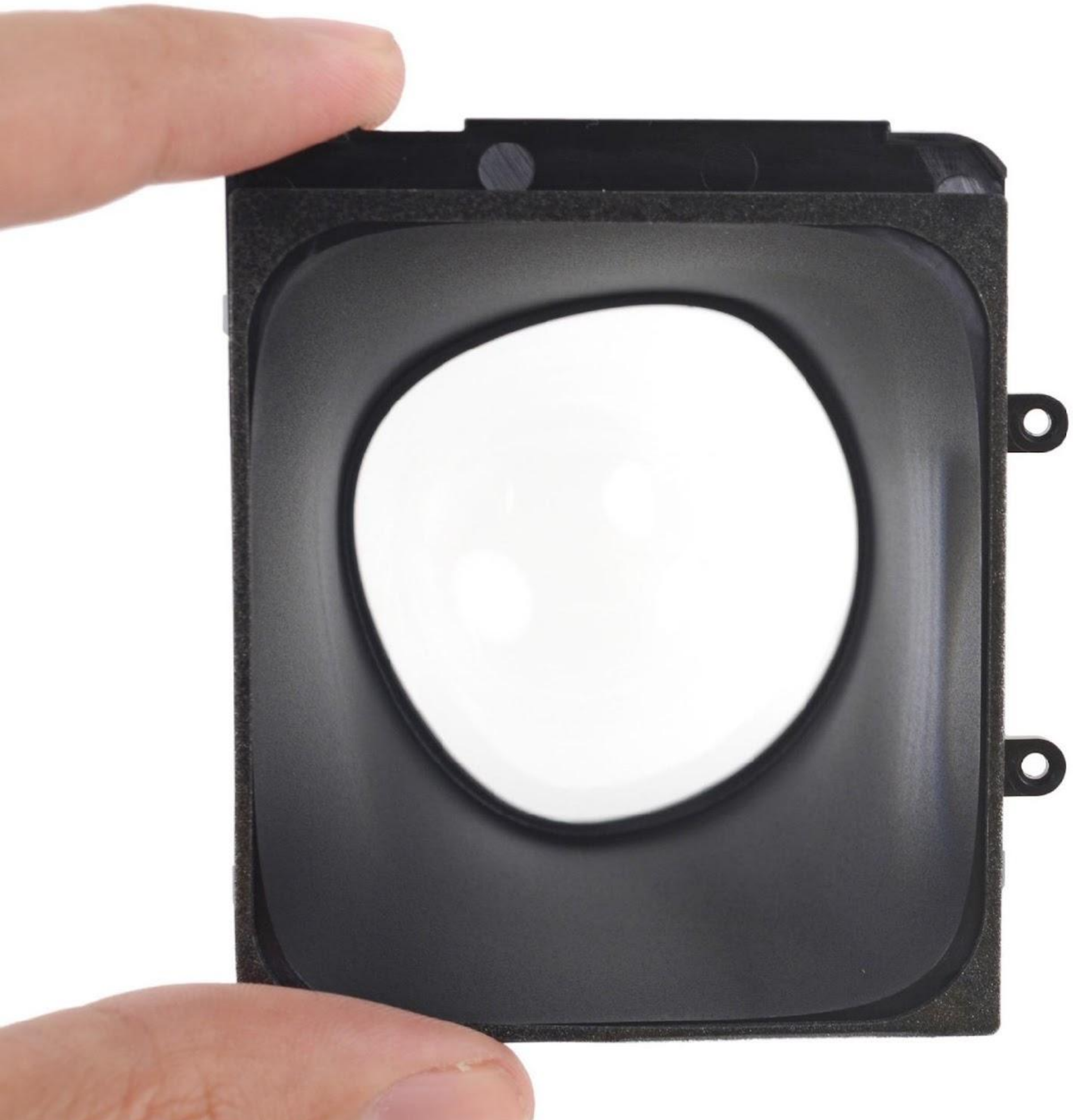


Image credit: ifixit.com

Oculus Rift (2016)

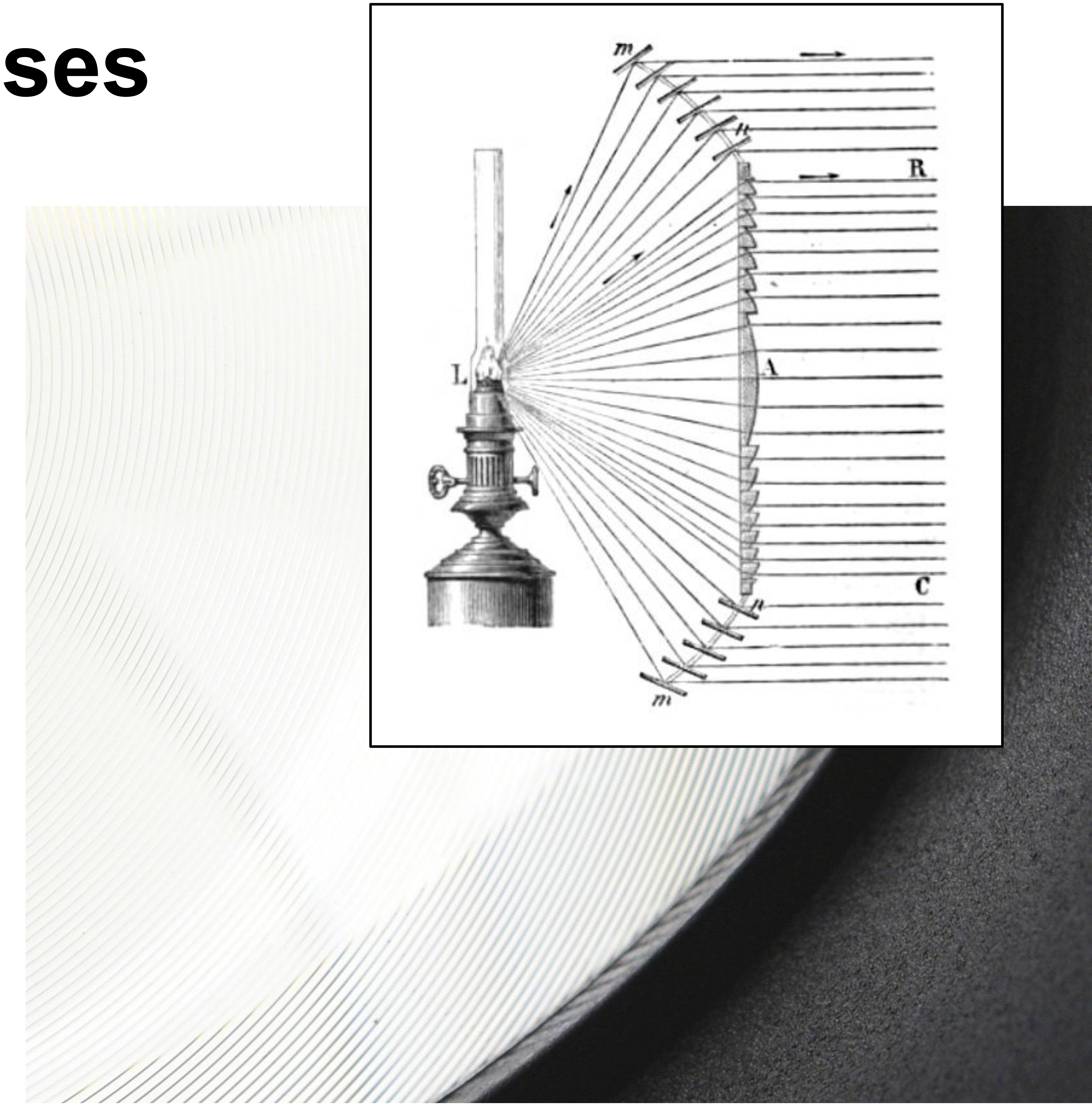


Fresnel eyepiece lens

1080x1200 display, 90 Hz

Image credit: ifixit.com

Oculus Rift Lenses



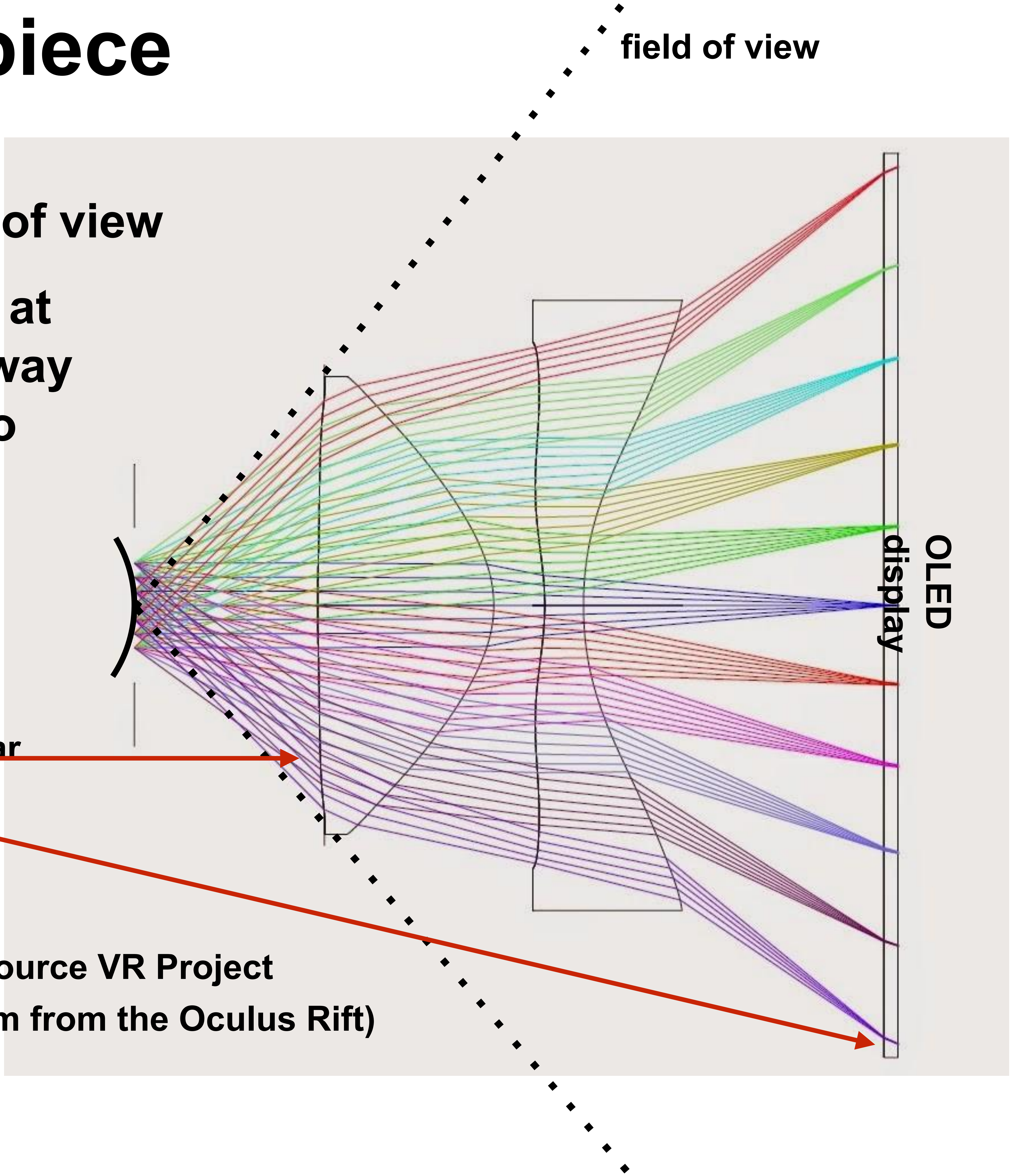
Fresnel eyepiece lens

Image credit: ifixit.com

Role of Eyepiece Lenses

1. Create wide field of view
2. Place focal plane at several meters away from eye (close to infinity)

Note: parallel lines reaching eye converge to a single
point on display (eye eye accommodates to plane near infinity)

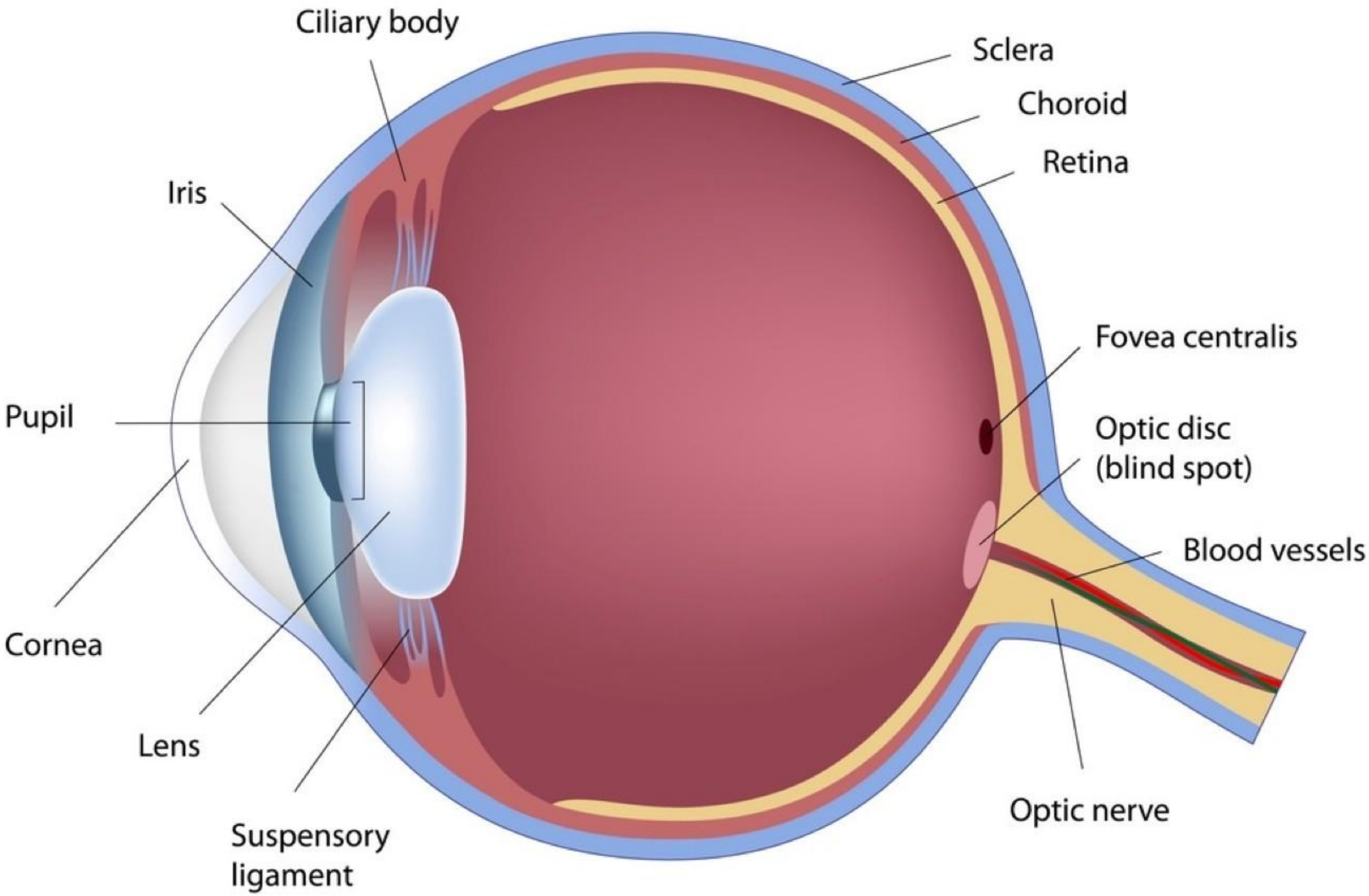


Lens diagram from Open Source VR Project (OSVR) (Not the lens system from the Oculus Rift)

http://www.osvr.org/

Display Requirements Derive From Human Perception

Anatomy of The Human Eye



Display Requirements

Derive From Human Perception

Example 1: Color

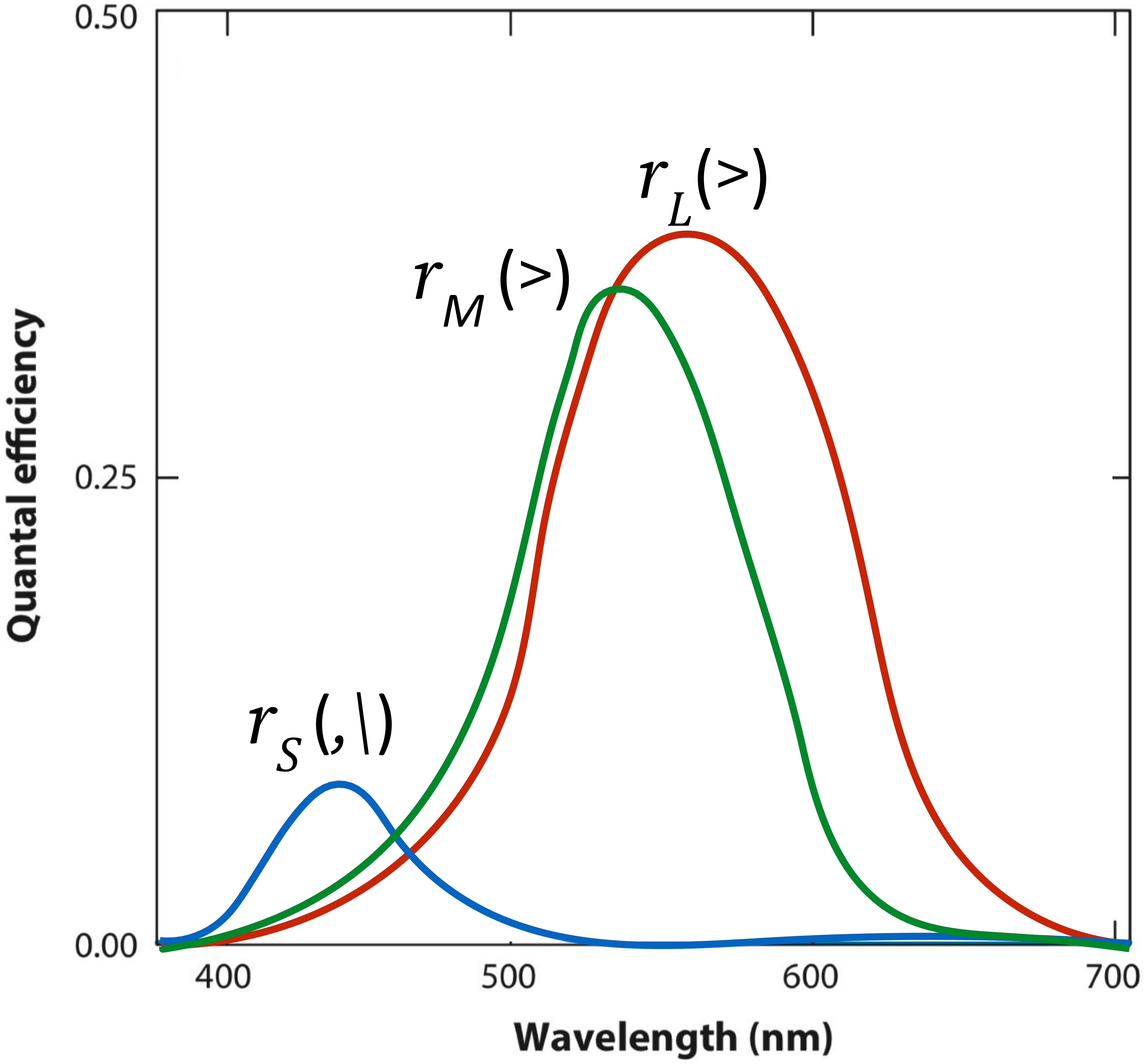
Recall: Spectral Response of Human Cone Cells

Instead of one detector as before, now we have three detectors (S, M, L cone cells), each with a different spectral response curve

$$S = \int r_S(\lambda) s(\lambda)\, d\lambda$$

$$M = \int r_M(\lambda) s(\lambda)\, d\lambda$$

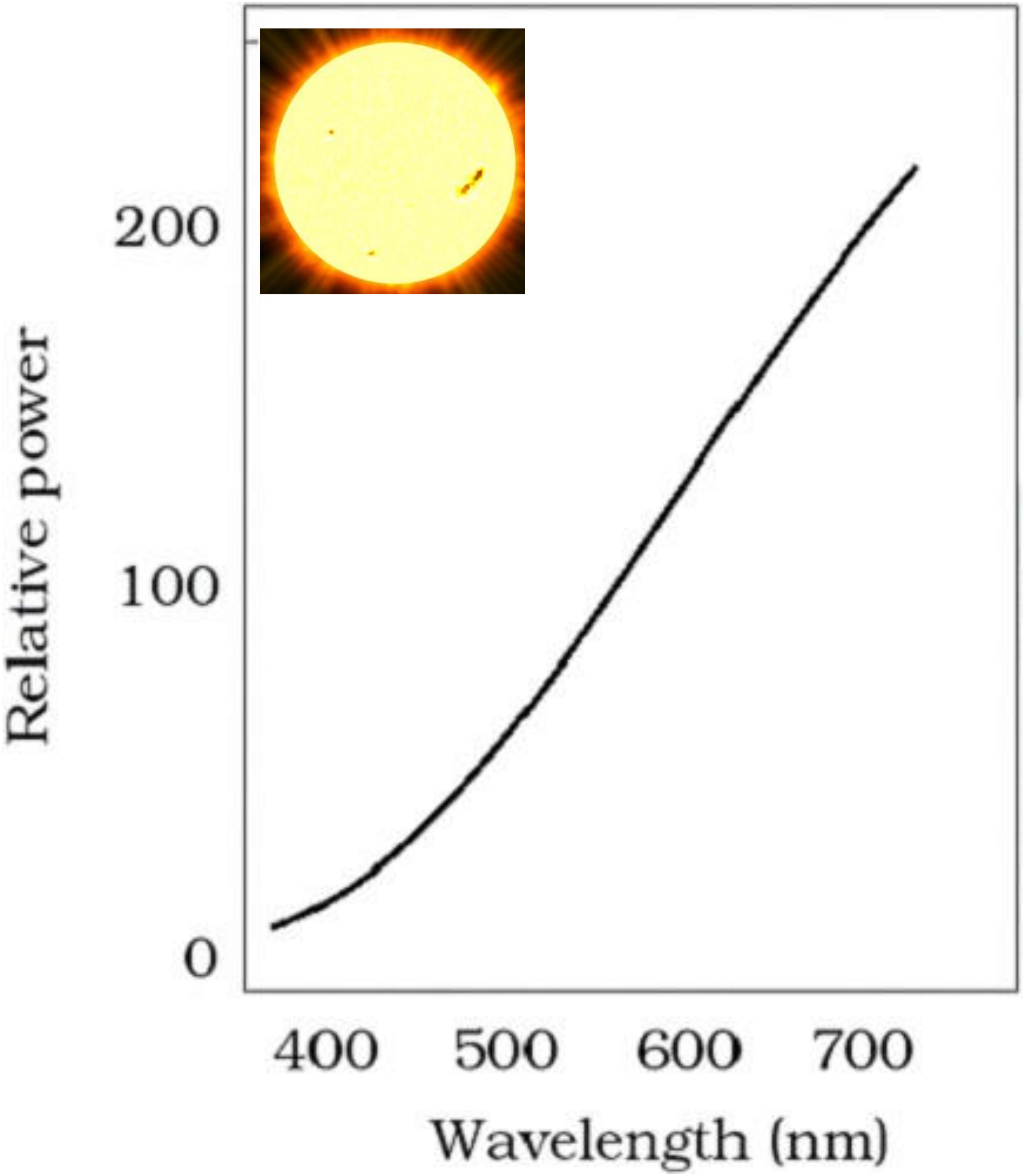
$$L = \int r_L(\lambda) s(\lambda)\, d\lambda$$

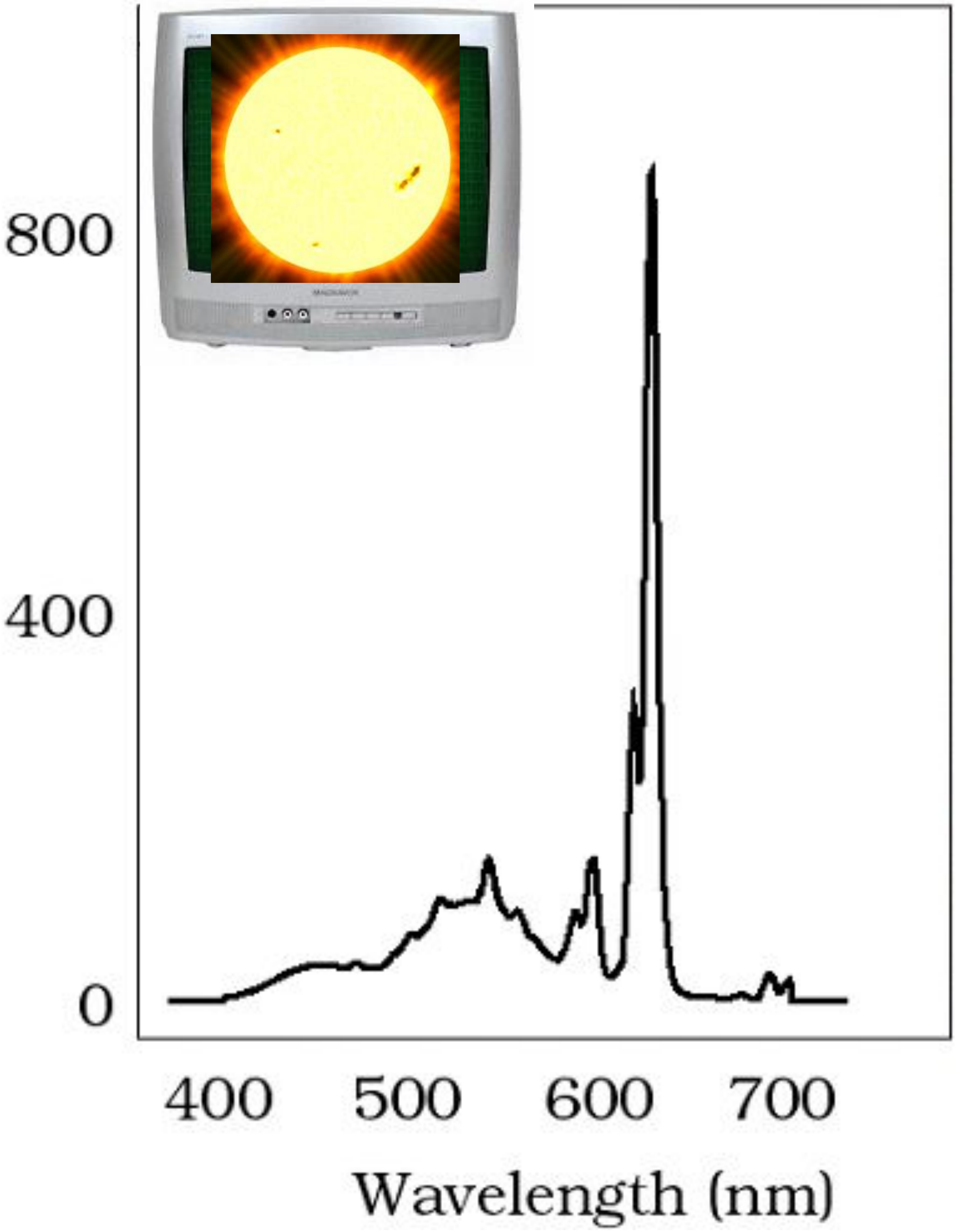


Brainard, Color and the Cone Mosaic, 2015.

Recall: Metamerism

Color matching is an important illusion that is understood quantitatively

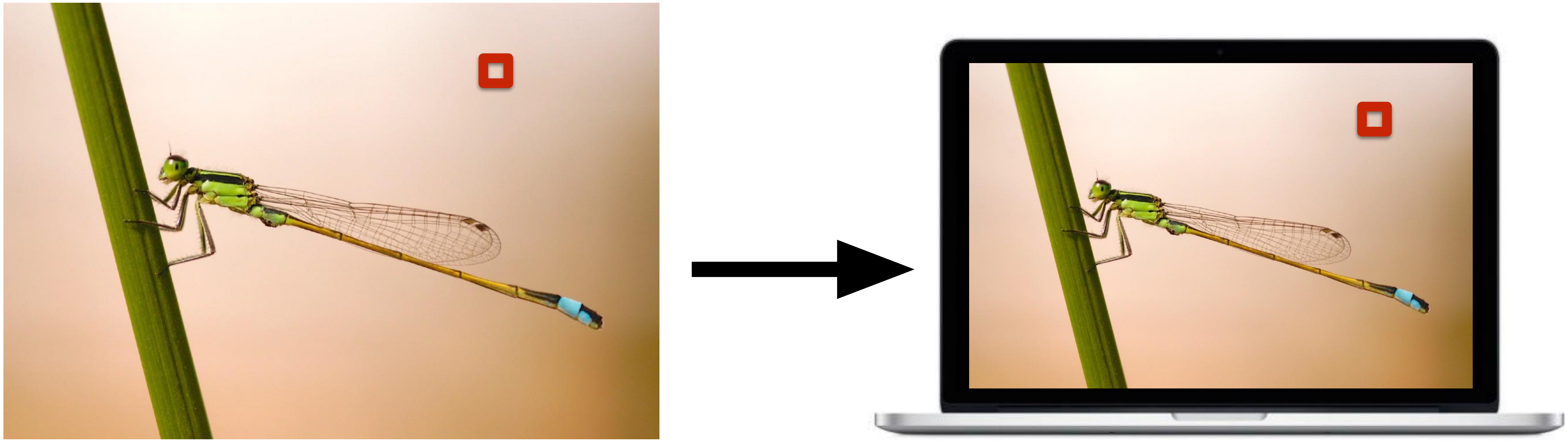




Brian Wandell

g

Recall: Color Reproduction



Target real spectrum $s(,\backslash)$

Display outputs spectrum

$R\,s_R(>) + Gs_G(>) + B\,s_B(>)$

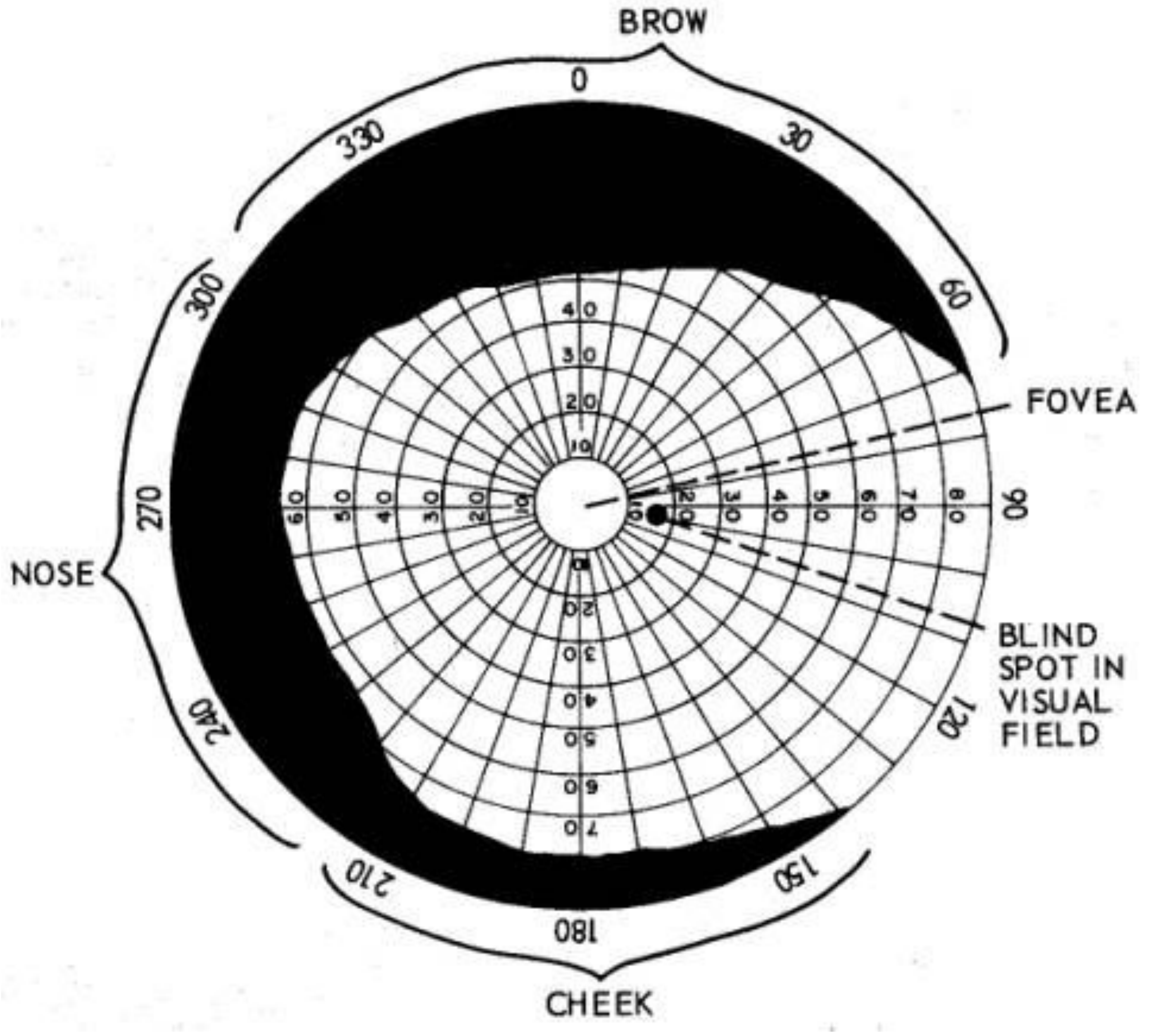
Goal: at each pixel, choose R, G, B values for display so that the output color matches the appearance of the target color in the real world.

Display Requirements Derive From Human Perception

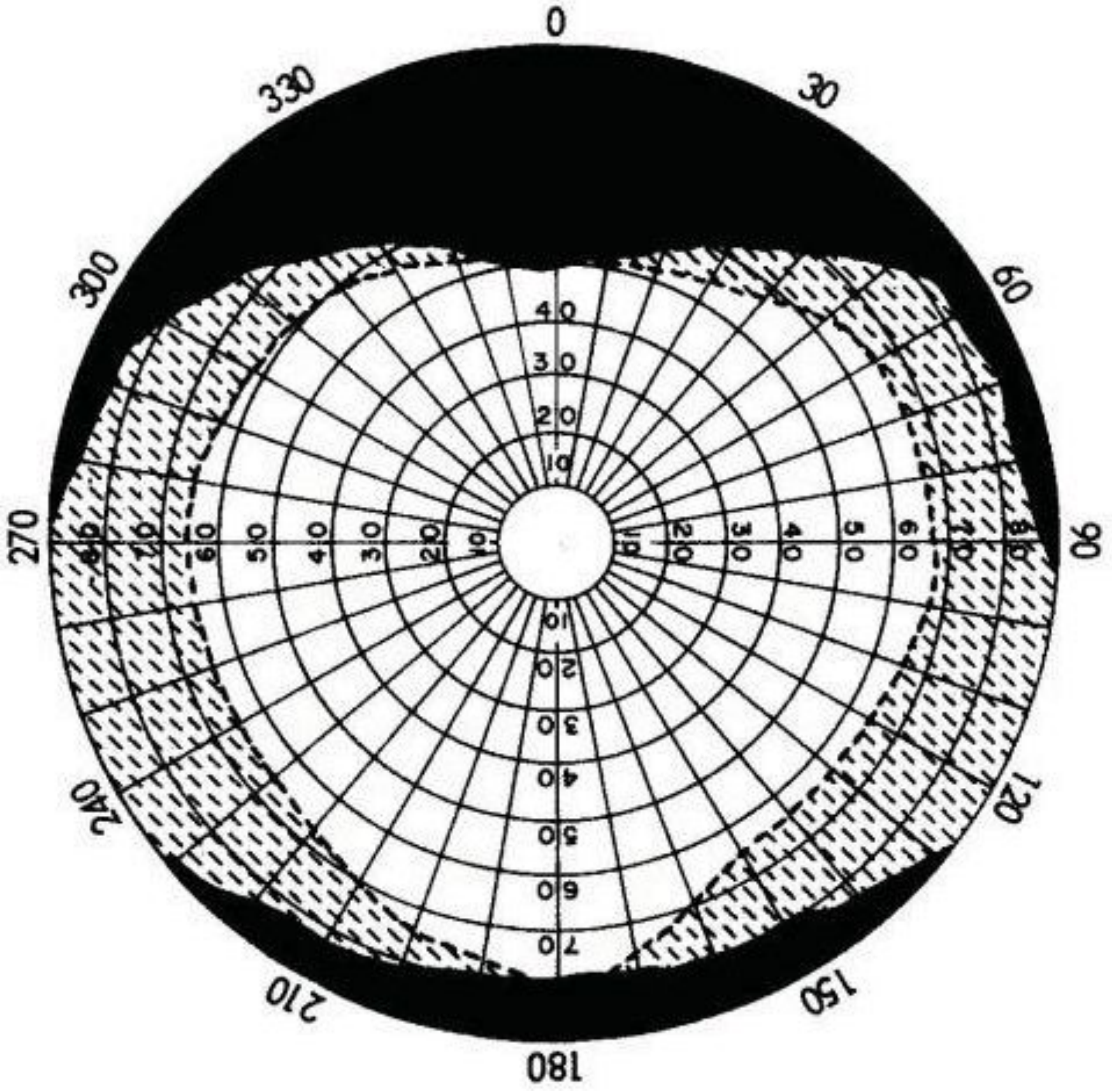
Example 2: Field of View & Resolution

Human Visual Field of View

Ruch & Fulton, 1960



monocular visual field

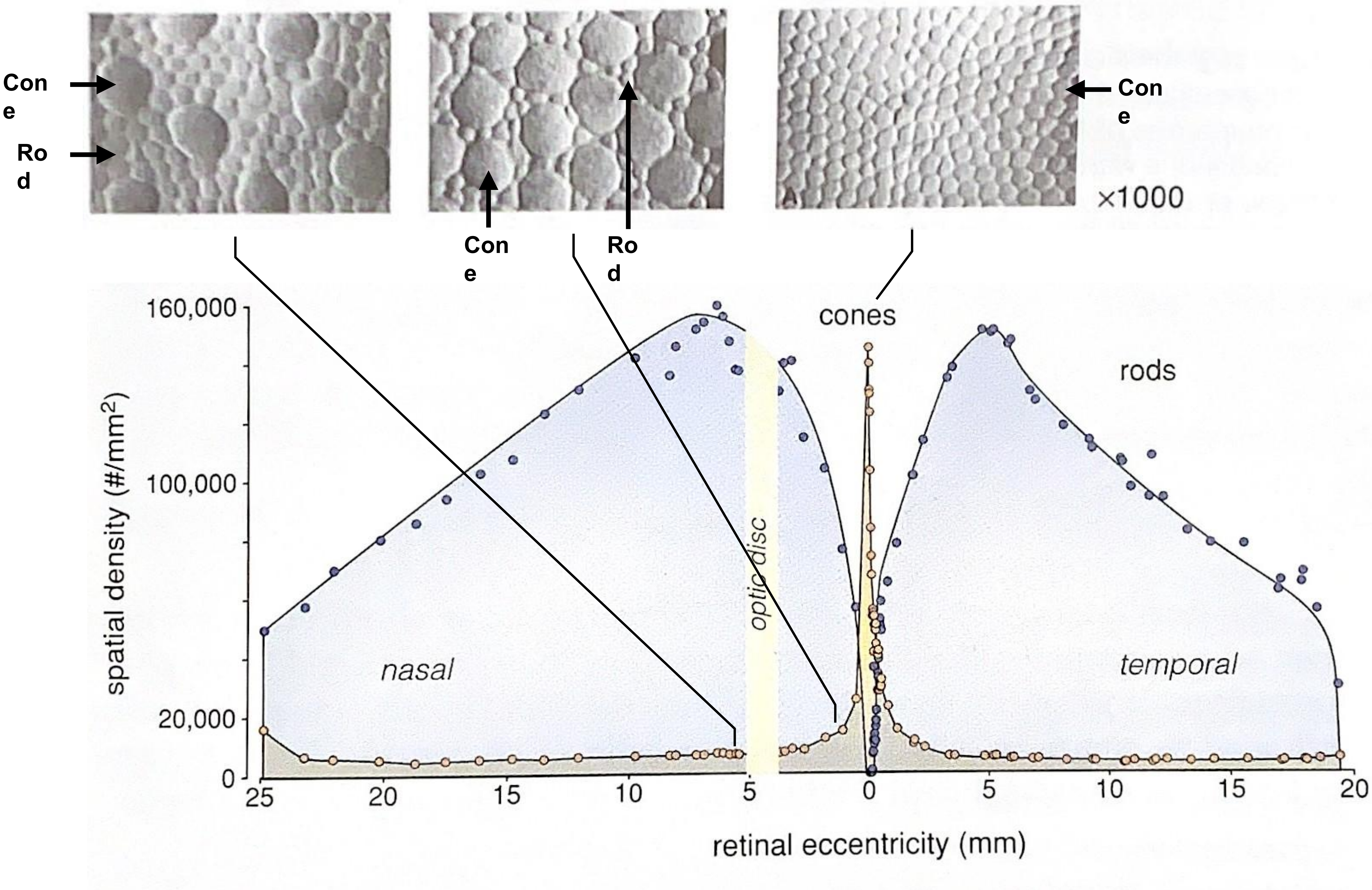


binocular visual field

Human: ~160° view of field per eye (~200° overall) (Note: does not account for eye’s ability to rotate in socket)

Slide credit: Gordon Wetzstein

Recall: Photoreceptor Size and Distribution Across Retina



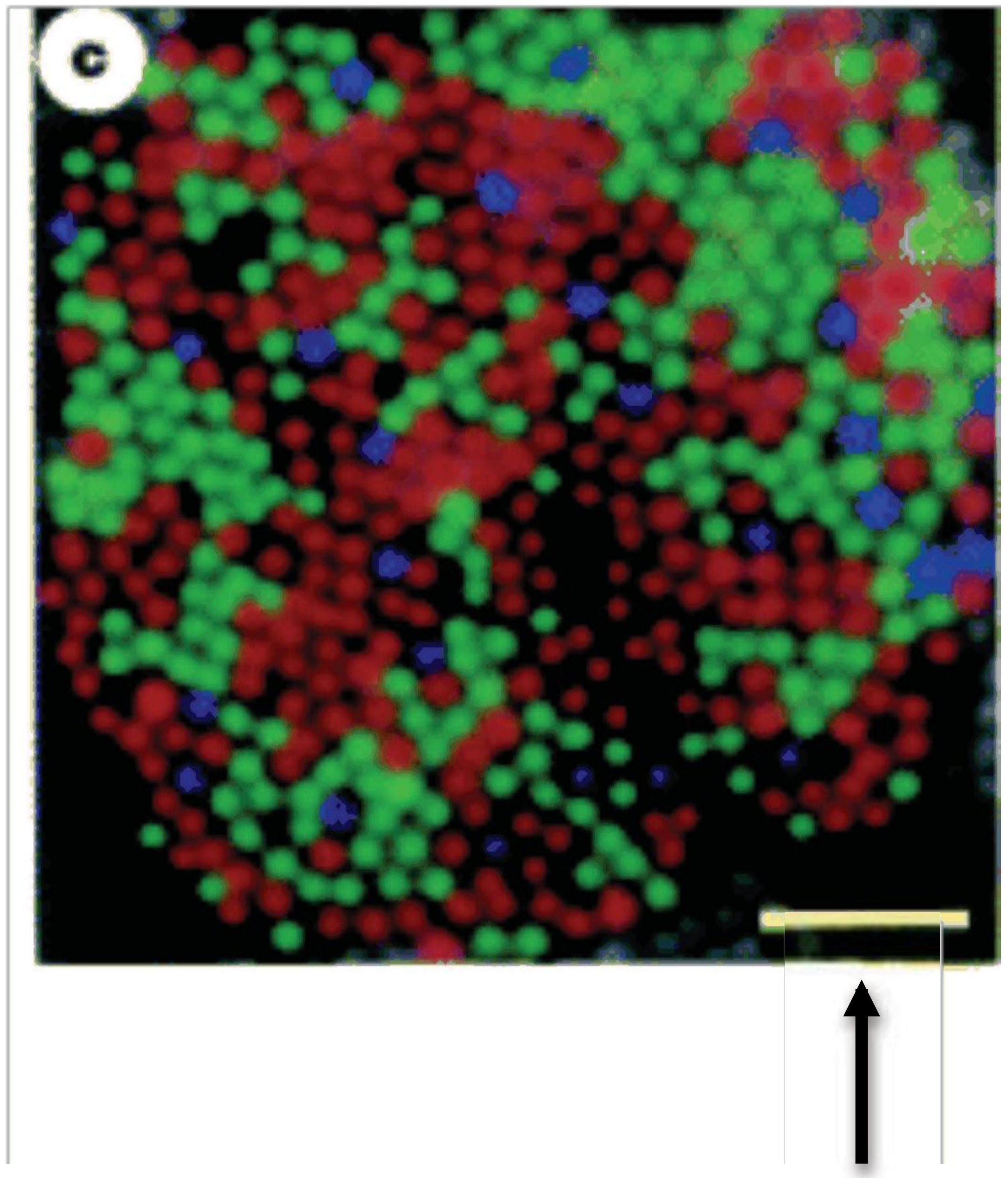
after Østerberg, 1935; as modified by Rodieck, 1988

Rodieck, p. 42

Visual Acuity

each photorecepter

~ 1 arc min (1/60 of a degree)

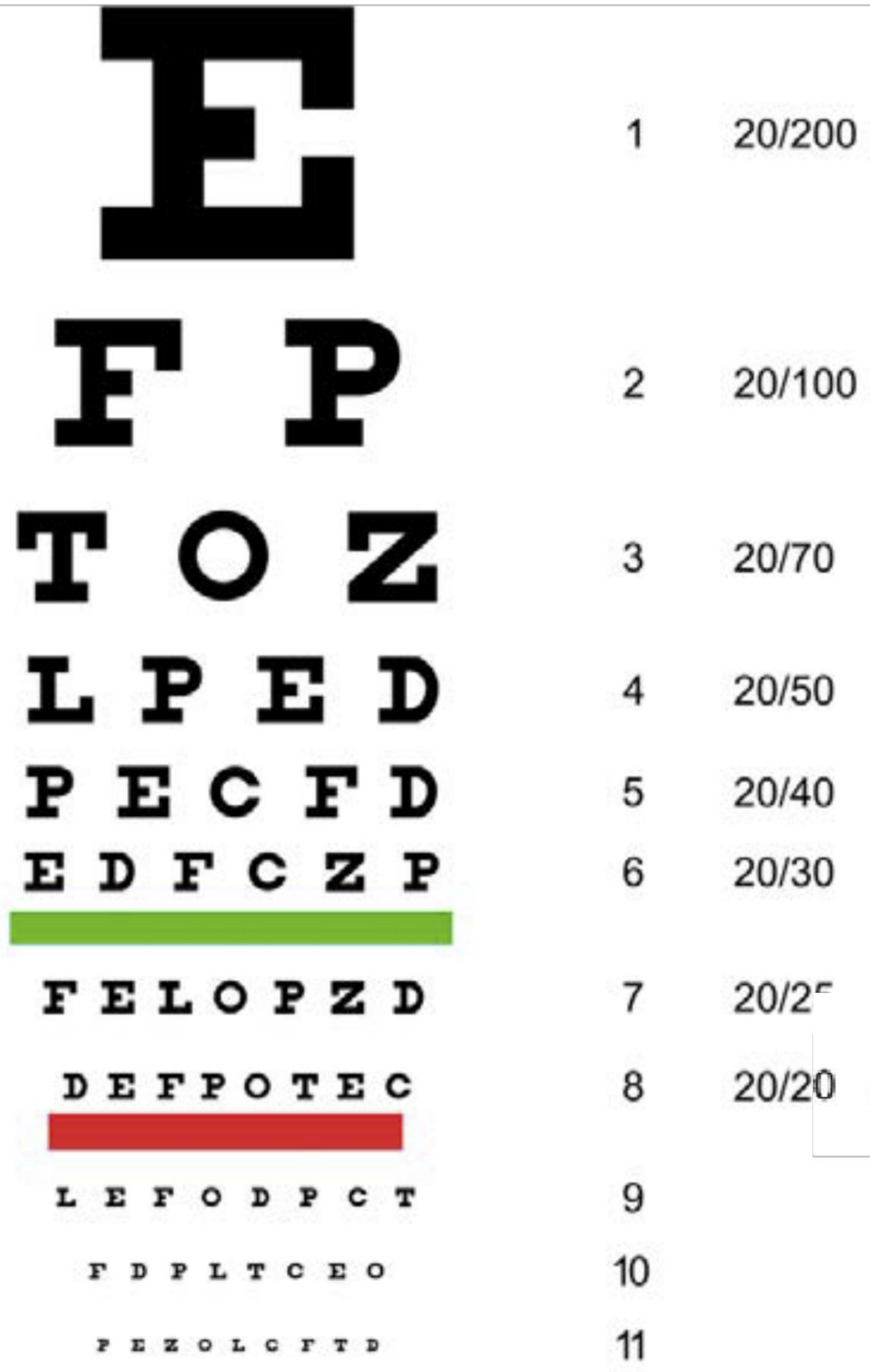


5 arcmin visual angle

Roorda & Williams, 1999, Nature

Slide credit: Gordon Wetzstein

Visual Acuity



Snellen chart

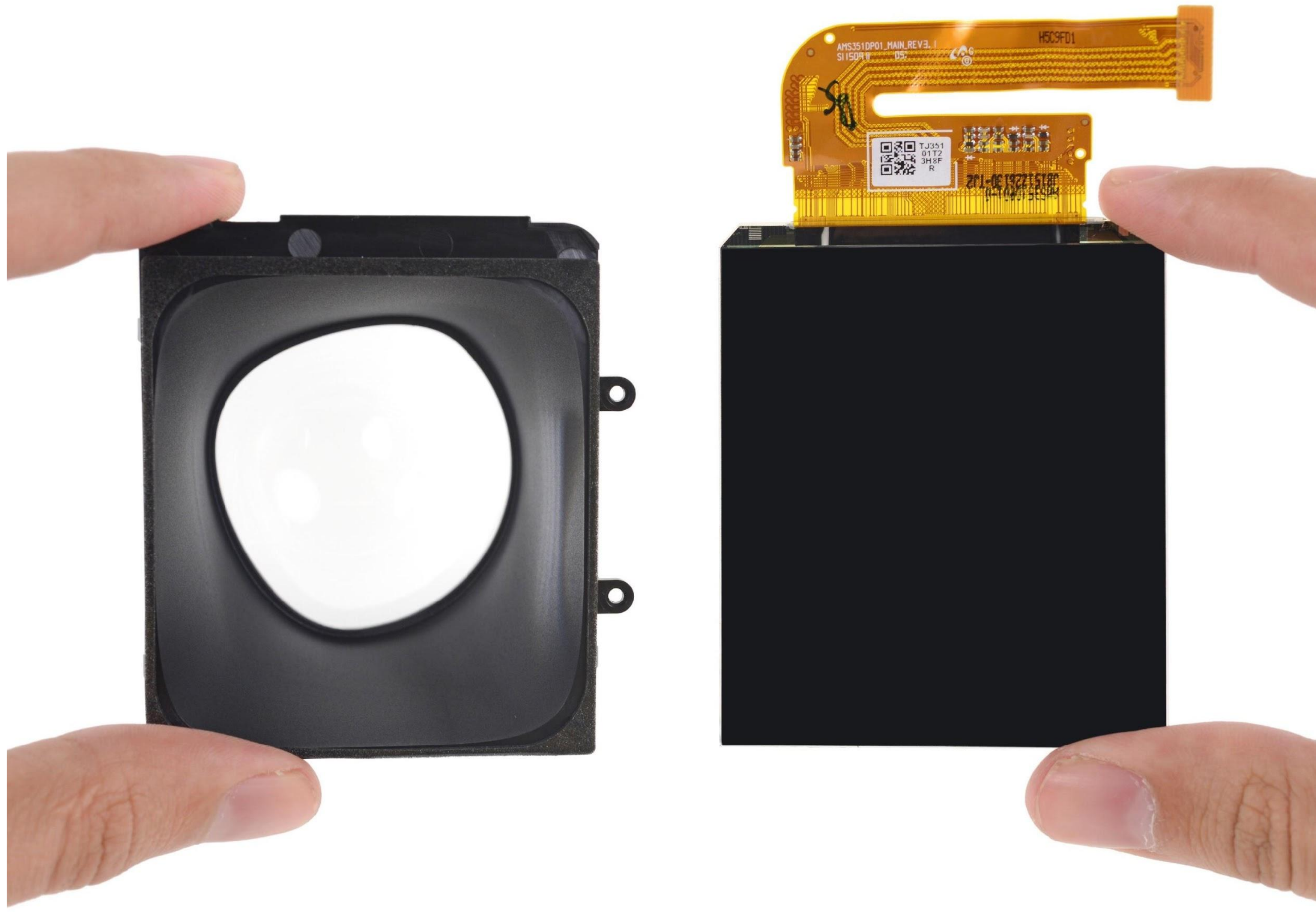
characters are 5 arc min, need to resolve 1 arc min to read

Slide credit: Gordon Wetzstein

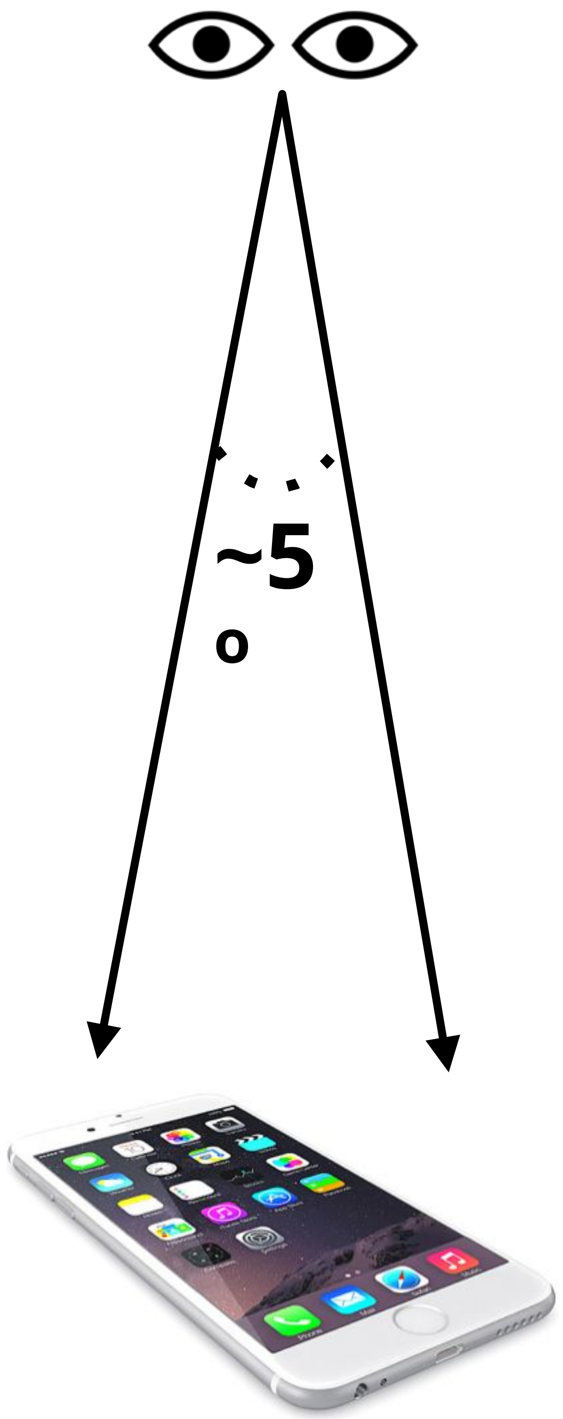
Current VR Headset Field of View and Resolution

Example: HTC Vive Pro 2

- Field of view: approximately 100° per eye
  - Resolution: 2448 x 2448 (6MP) pixel display
- About 24 pixels per degree (as opposed to ~60 samples for 20/20 vision)
- [Note: VR headsets exist up to 2880x2720 (7.8MP) now]



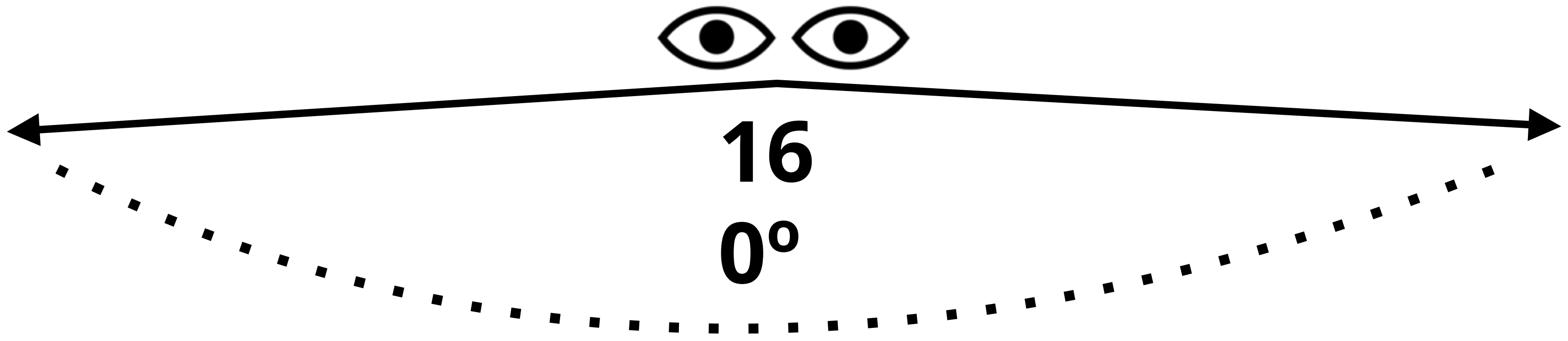
A VR Display at Human Visual Acuity



iPhone 6: 4.7 in “retina” display:

1.3 MPixel

326 ppi → ~60 ppd



Human: ~160° view of field per eye (~200° overall) (Note: does not account for eye’s ability to rotate in socket)

Future “retina” VR display:

~ 8K x 8K display per eye (50 ppd)

= 128 MPixel

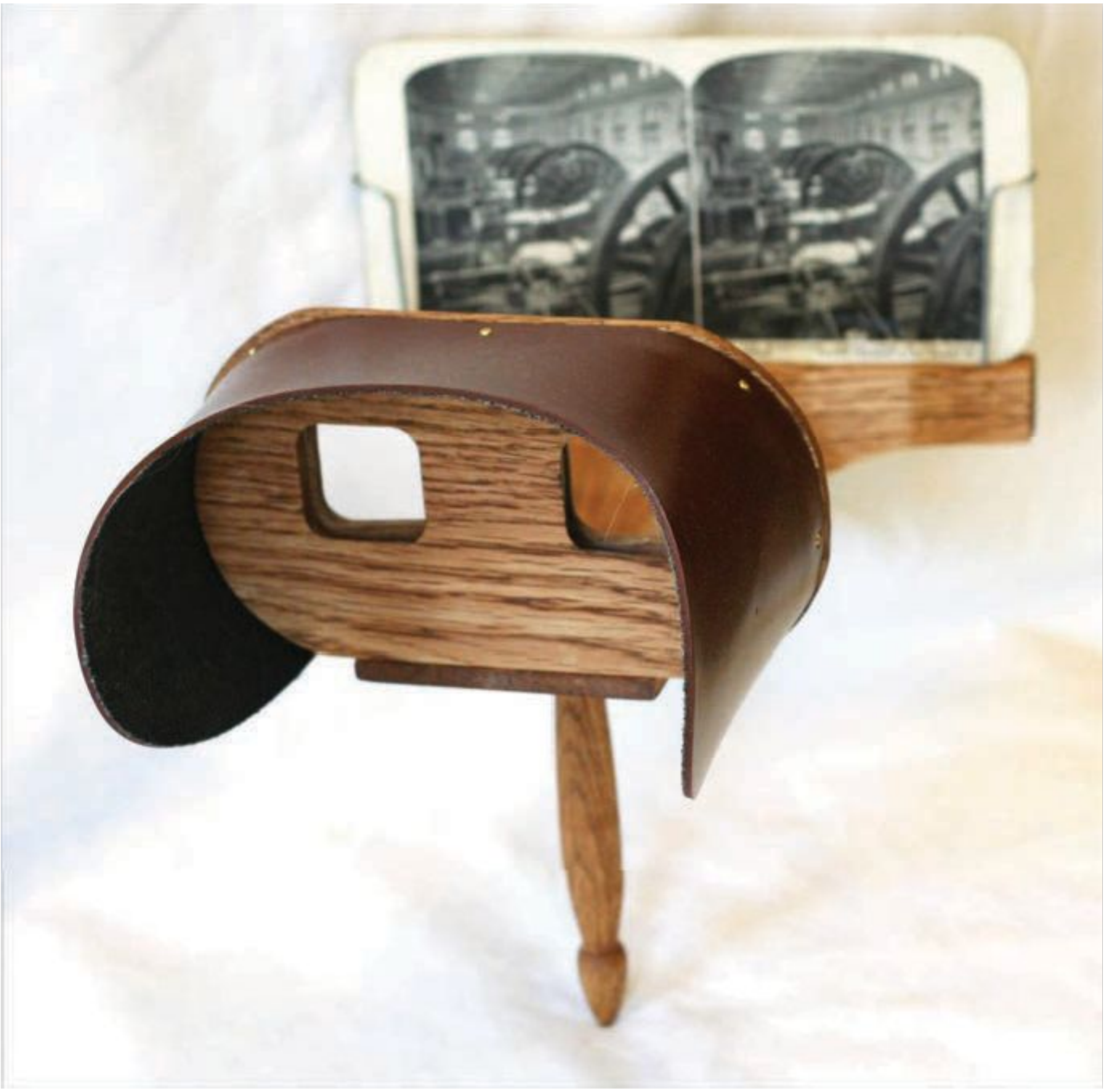
Strongly suggests need for eye tracking and foveated rendering (eye can only perceive detail in 5° region about gaze point)

Eye icons designed by SuperAtic LABS from the thenounproject.com

Display Requirements Derive From Human Perception

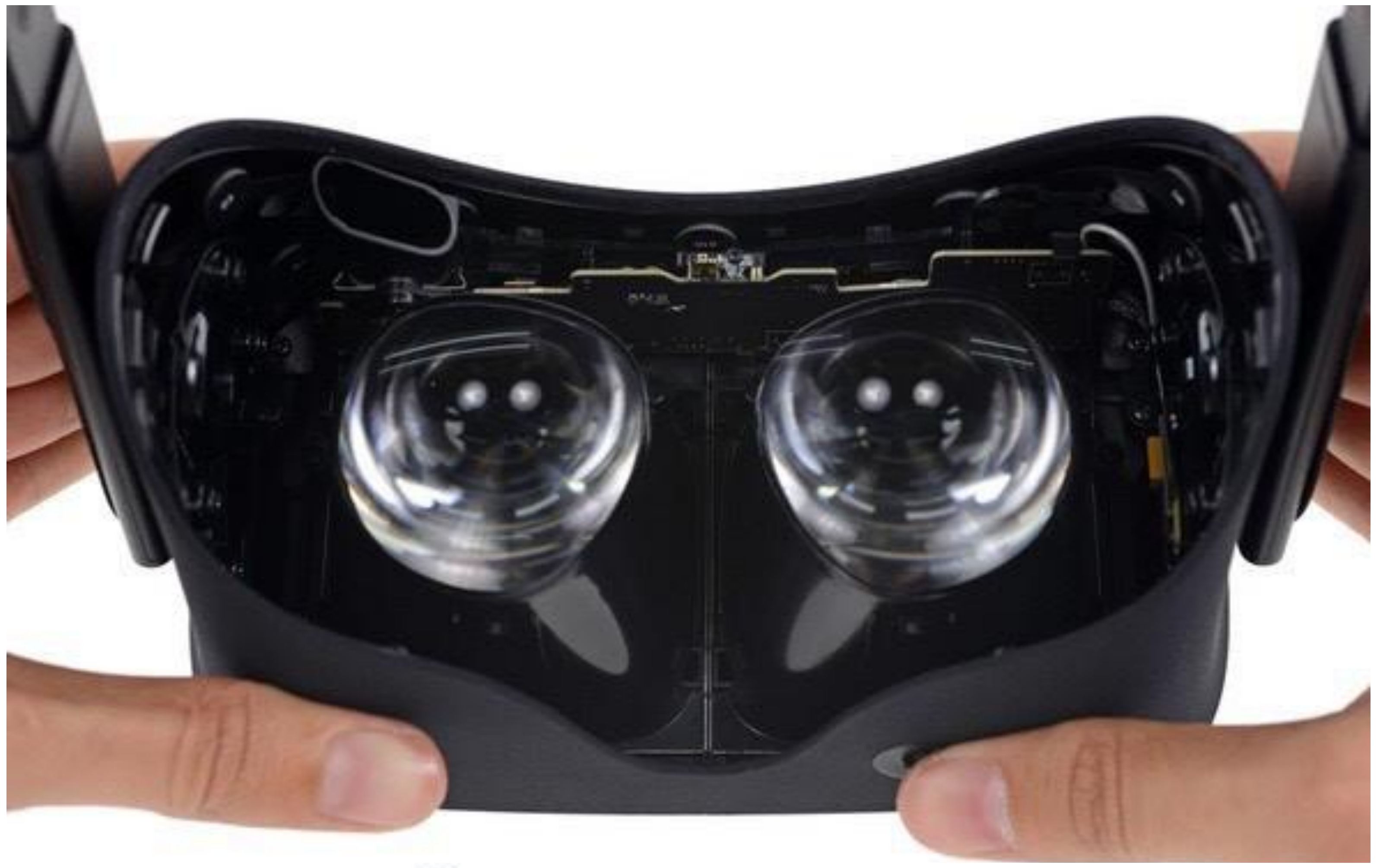
Example 3: Binocular Stereo and Eye Focus ("Accommodation")

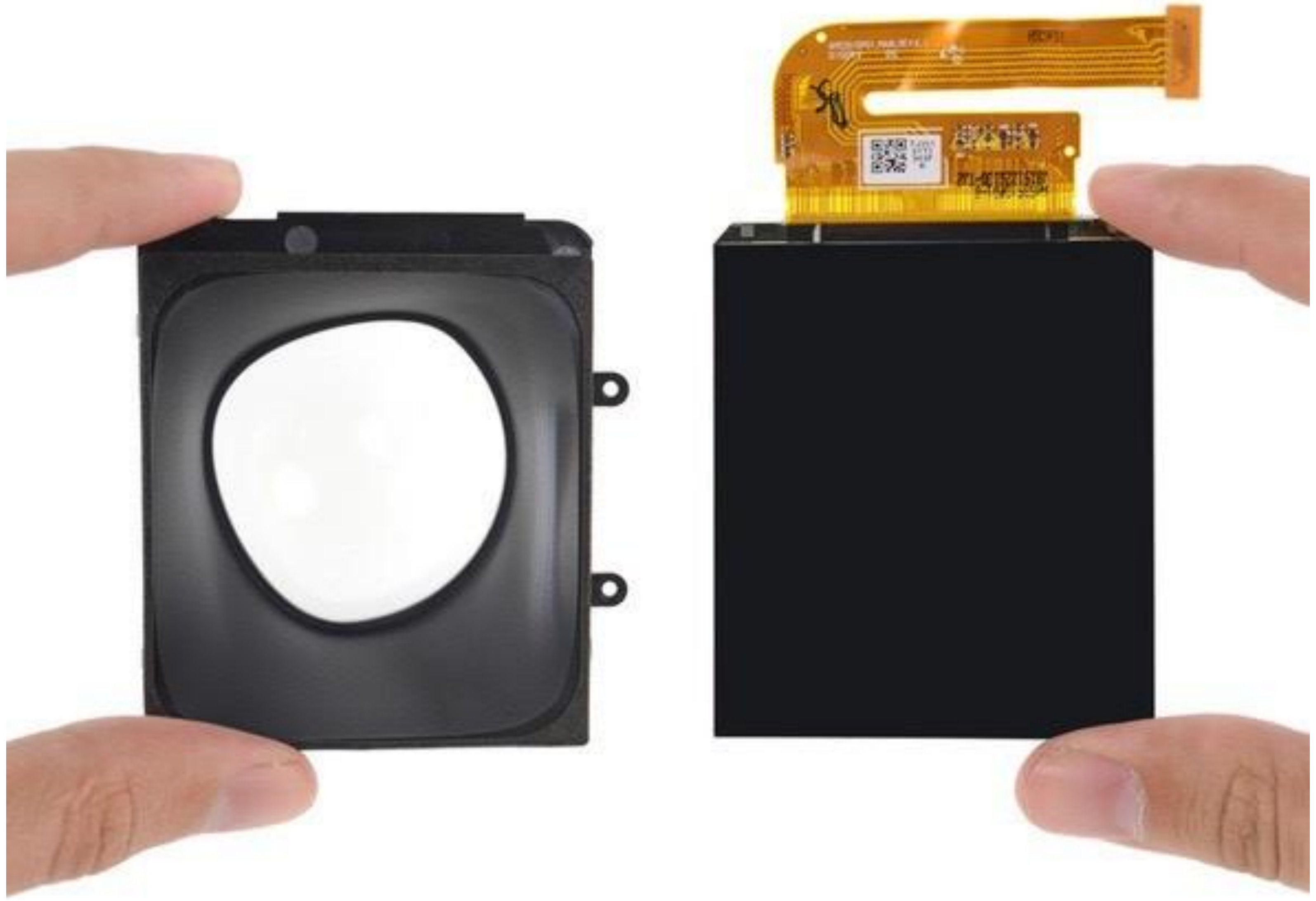
Two Eyes: Two Views

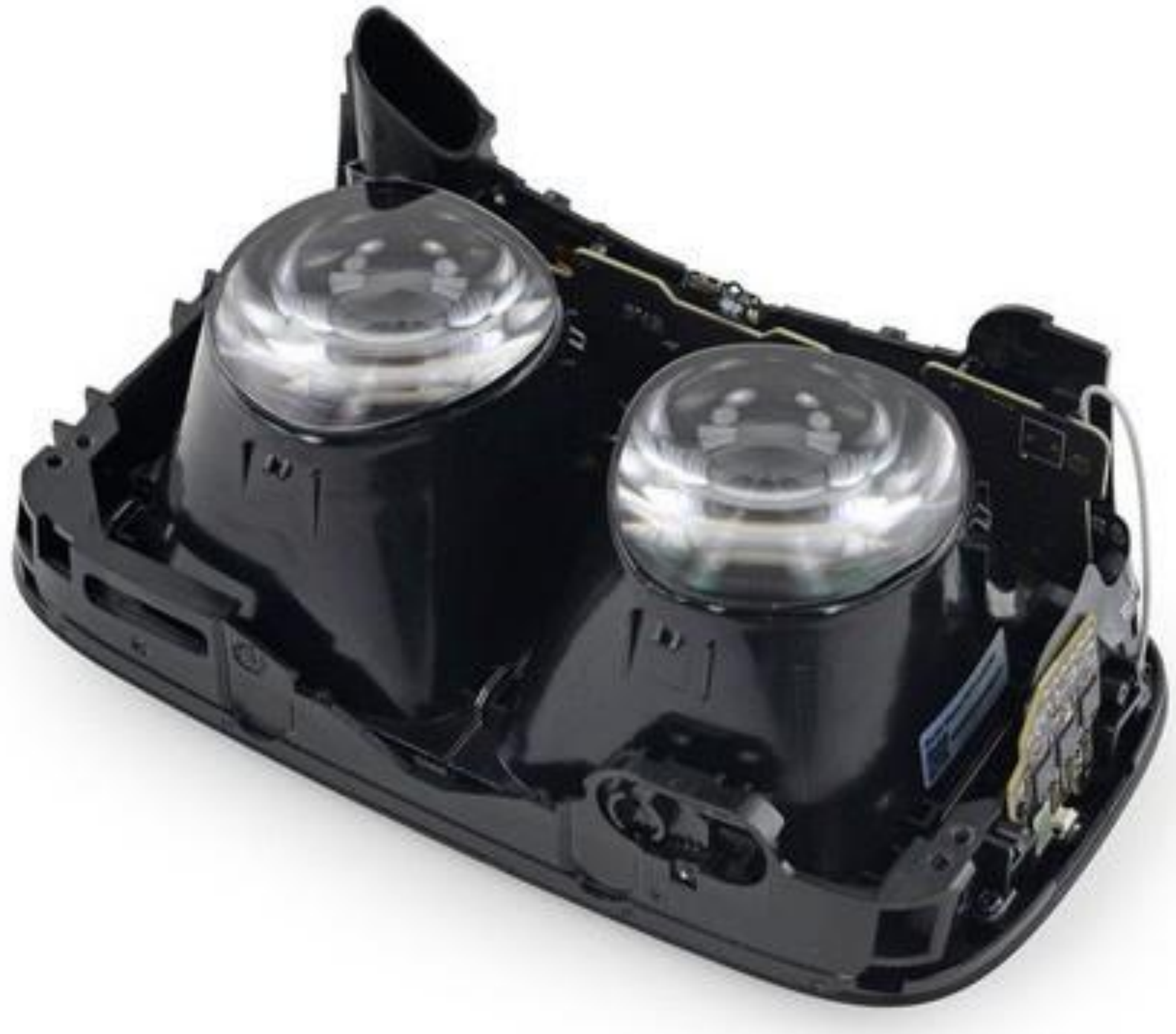


Charles Wheatstone stereoscope, 1838

Recall: Current VR HMD Optical Design







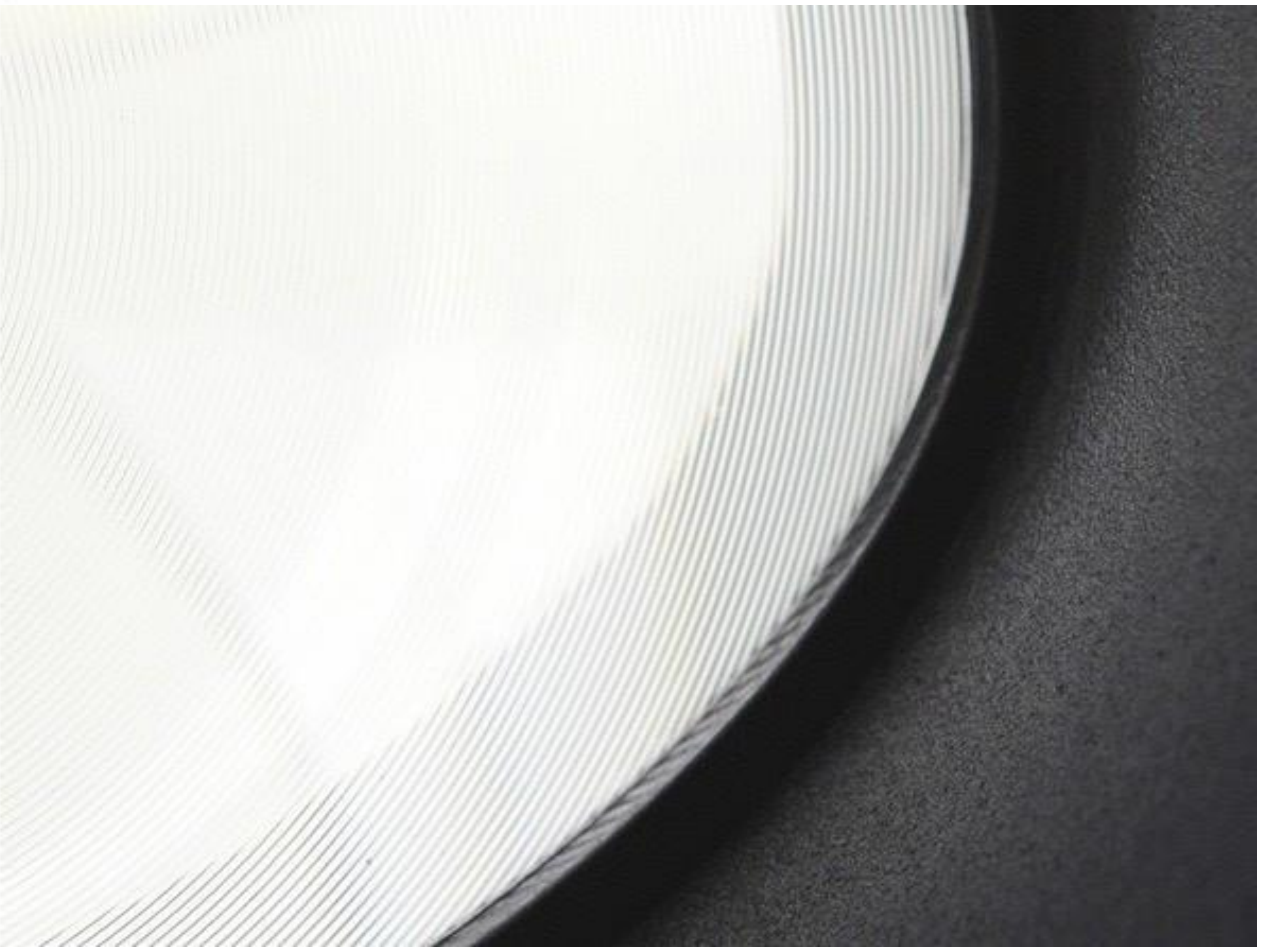
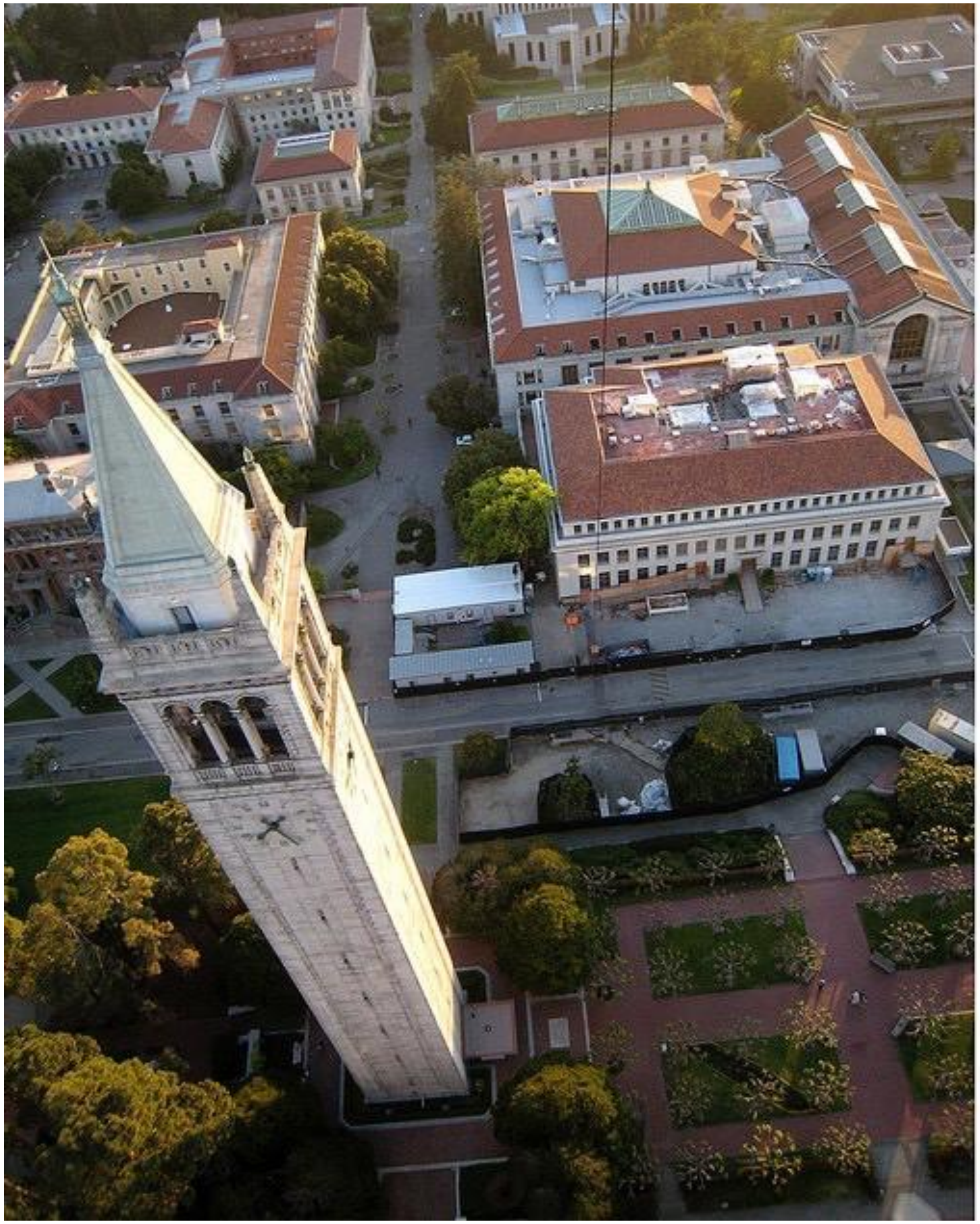
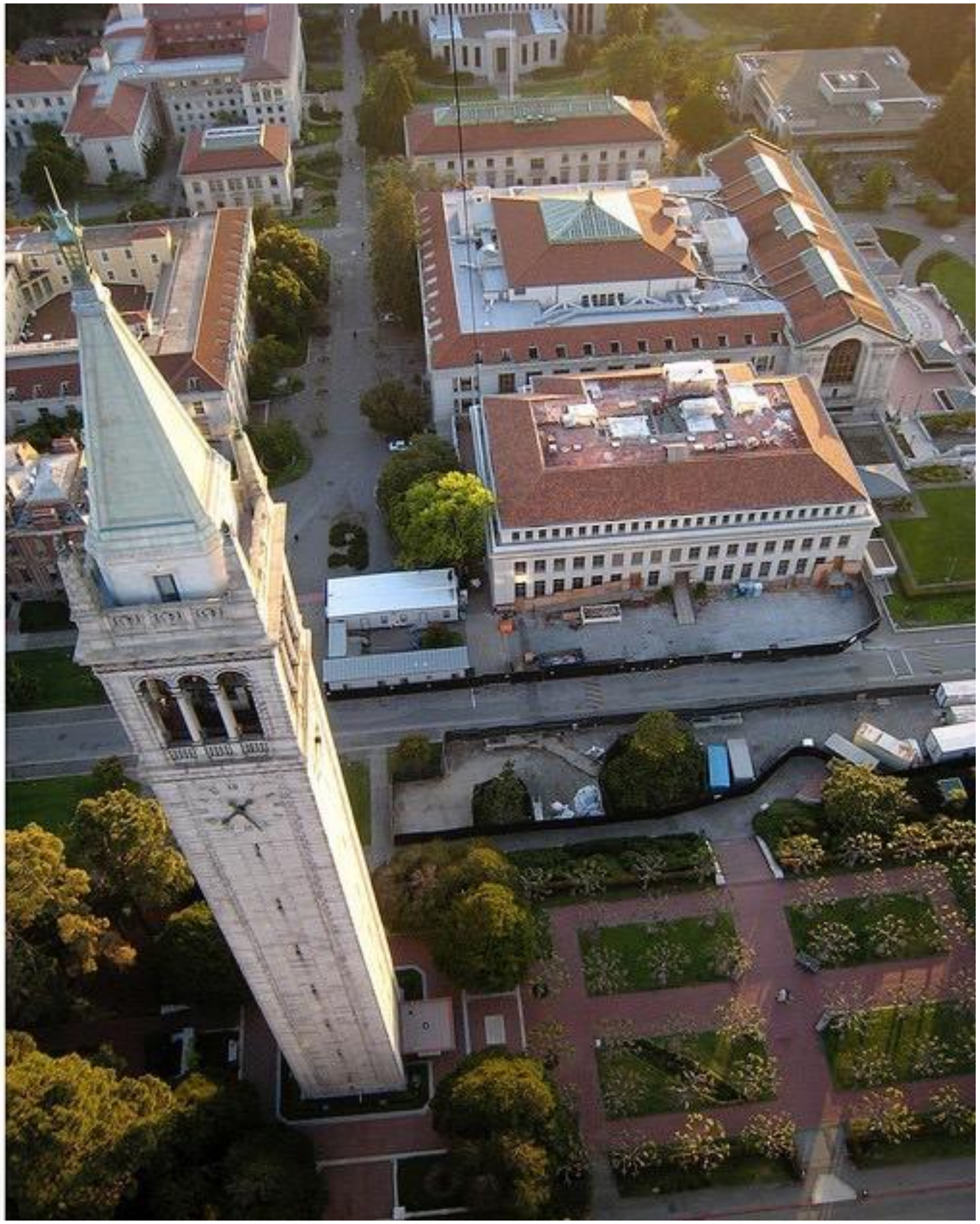


Image credit: ifixit.com https://www.ifixit.com/Teardown/Oculus+Rift+CV1+Teardown/60612

Stereo Vergence

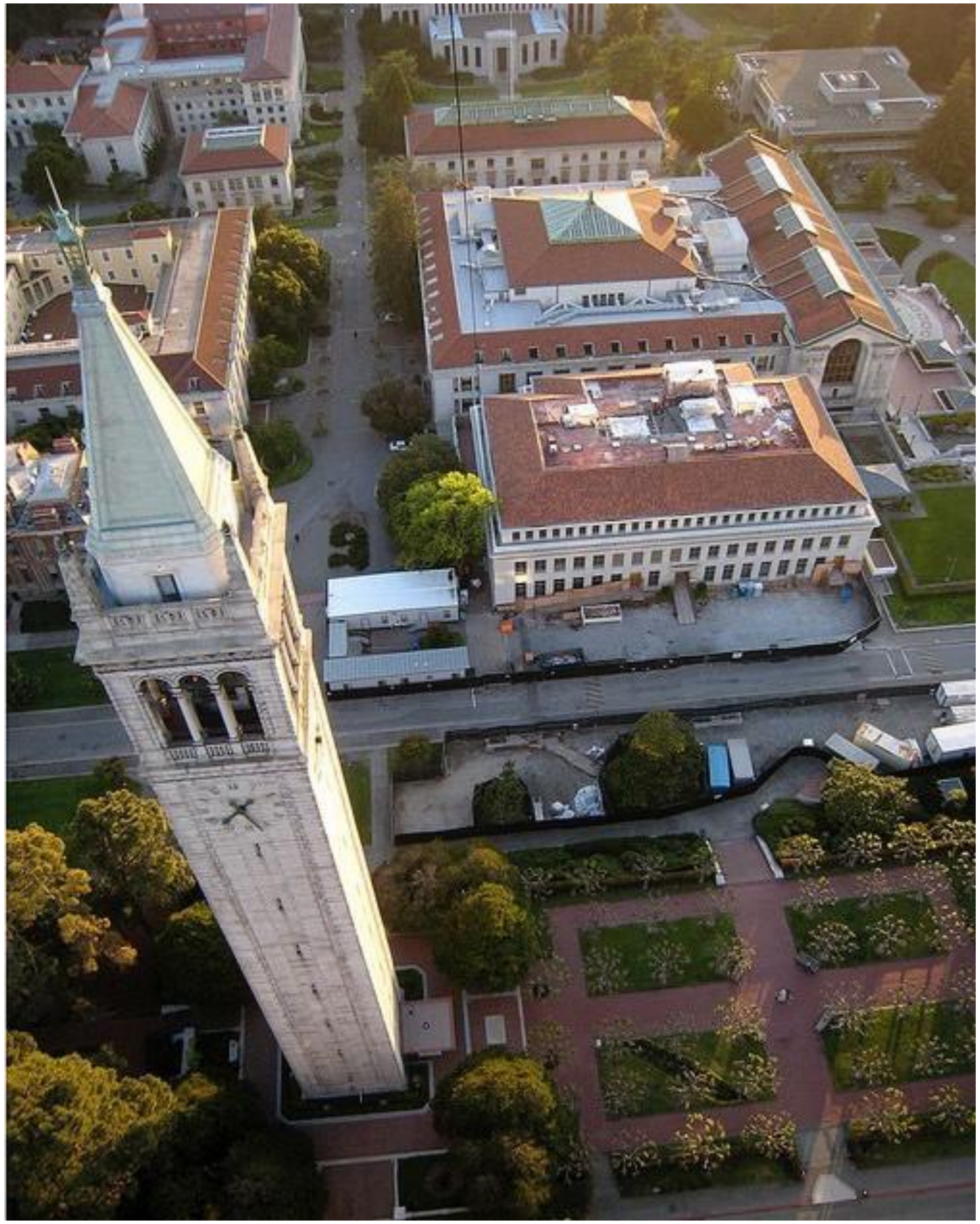


Left-eye perspective

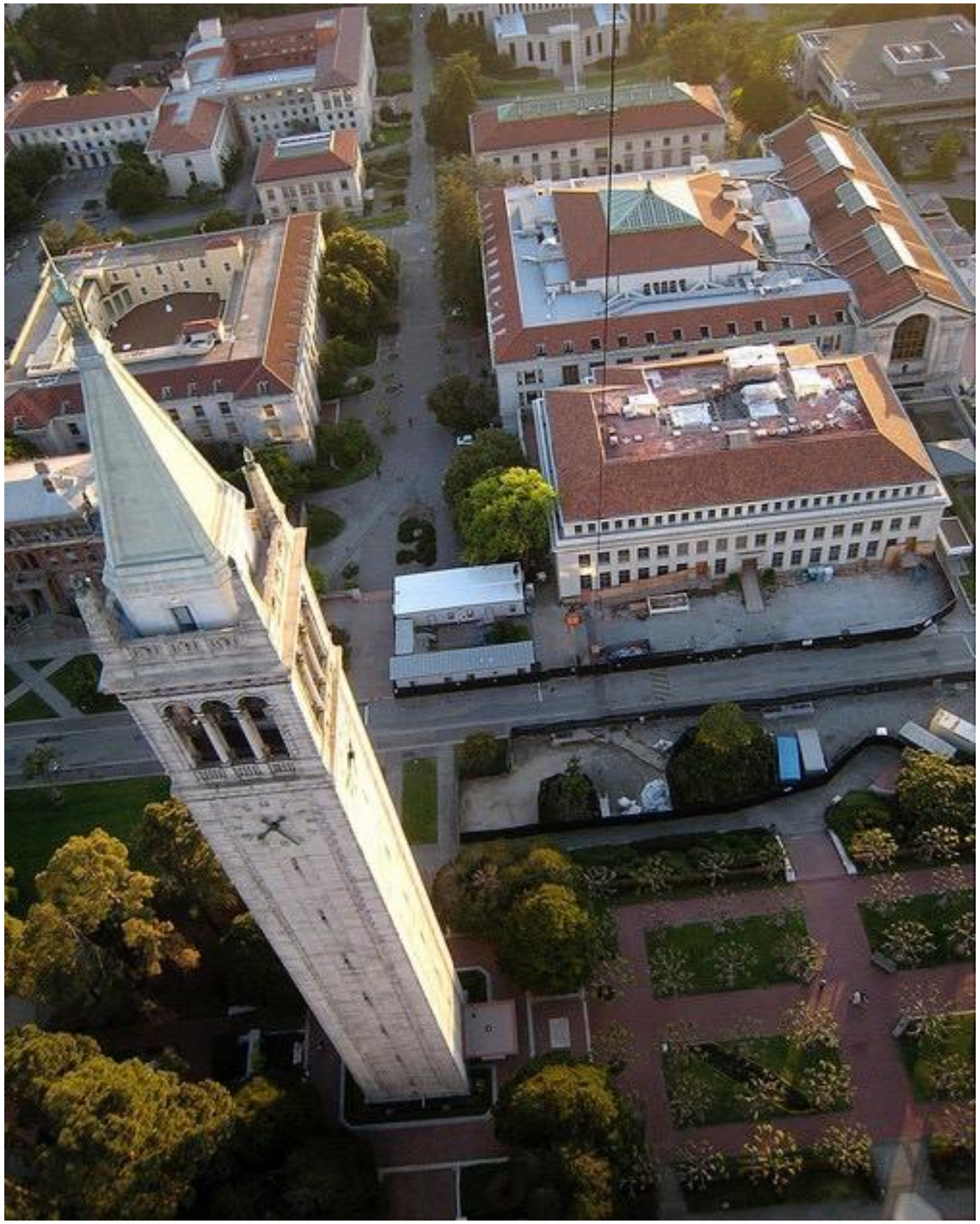


Right-eye perspective

Stereo Vergence

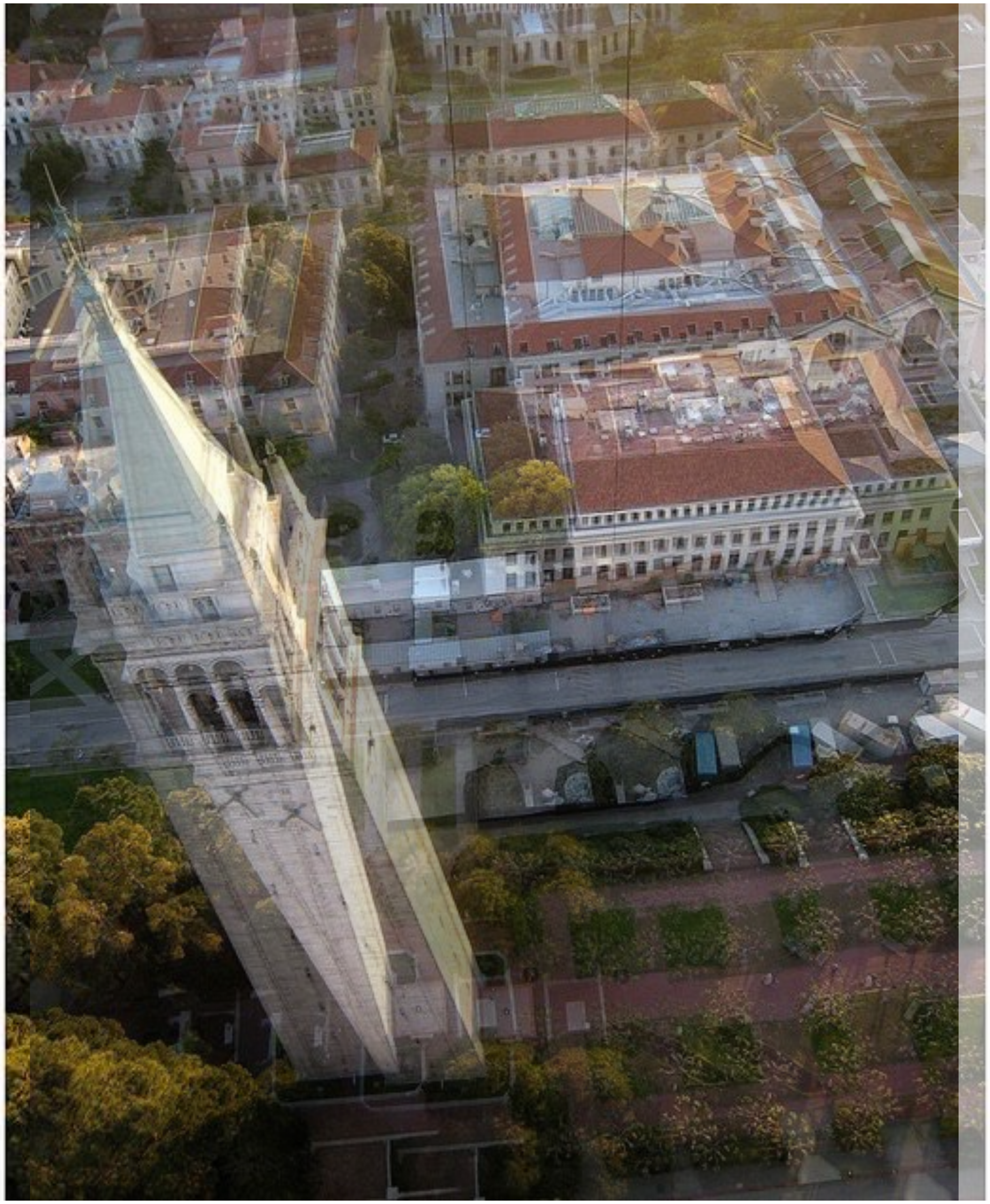


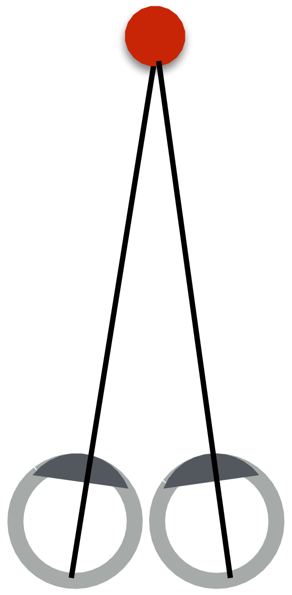
Left-eye perspective



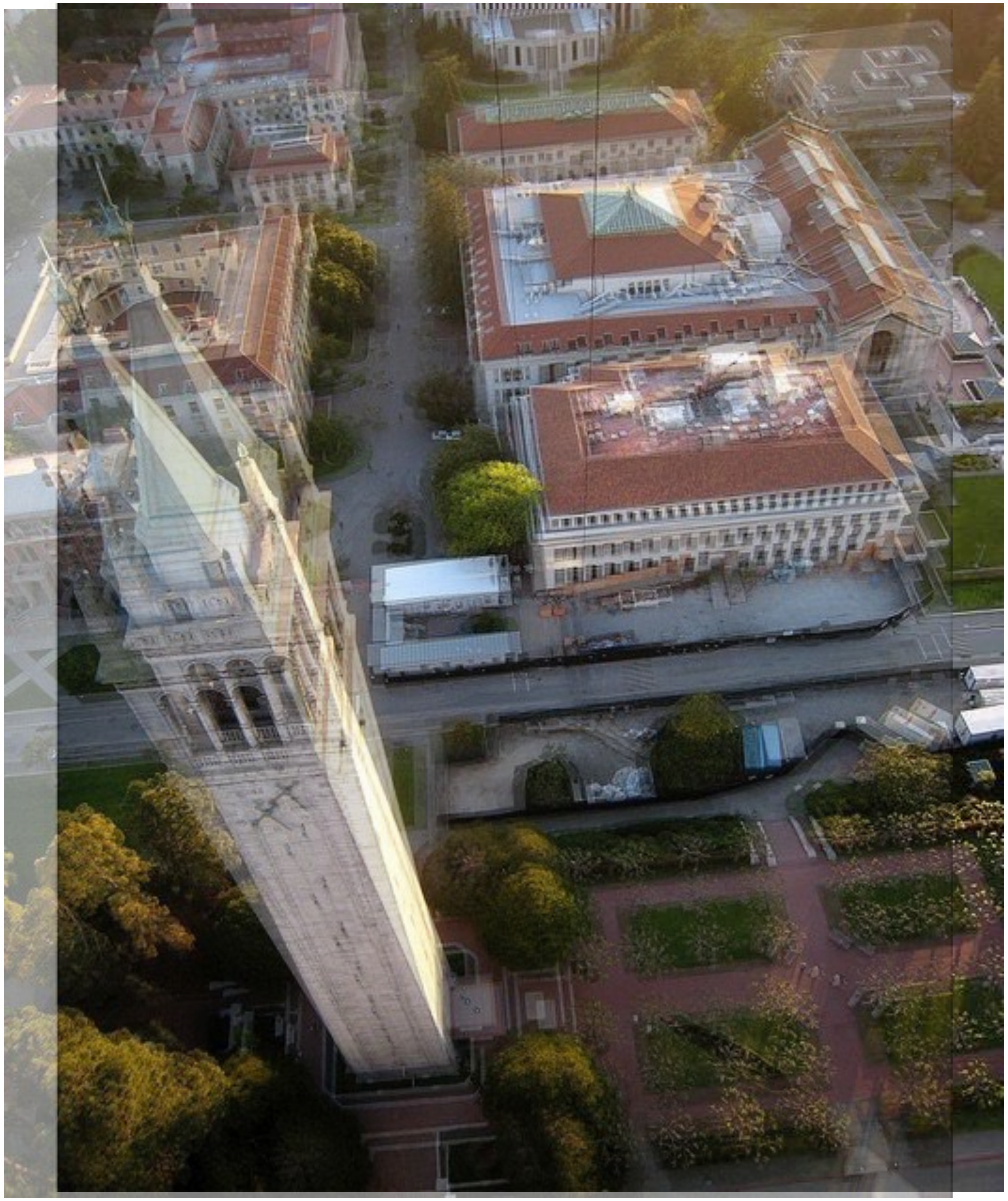
Right-eye perspective

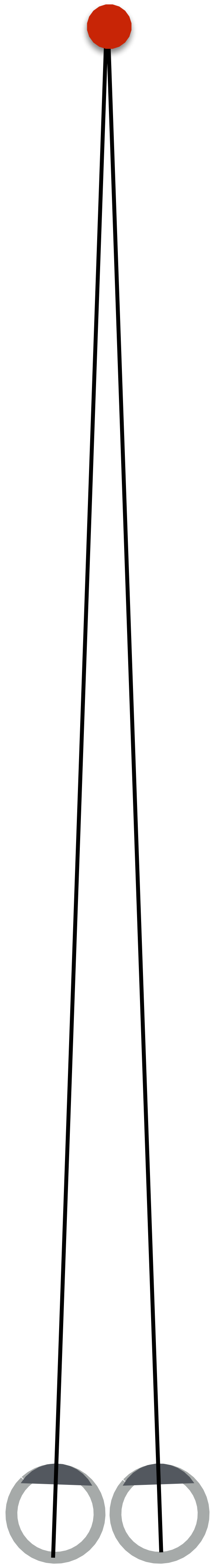
Stereo Vergence





Stereo Vergence

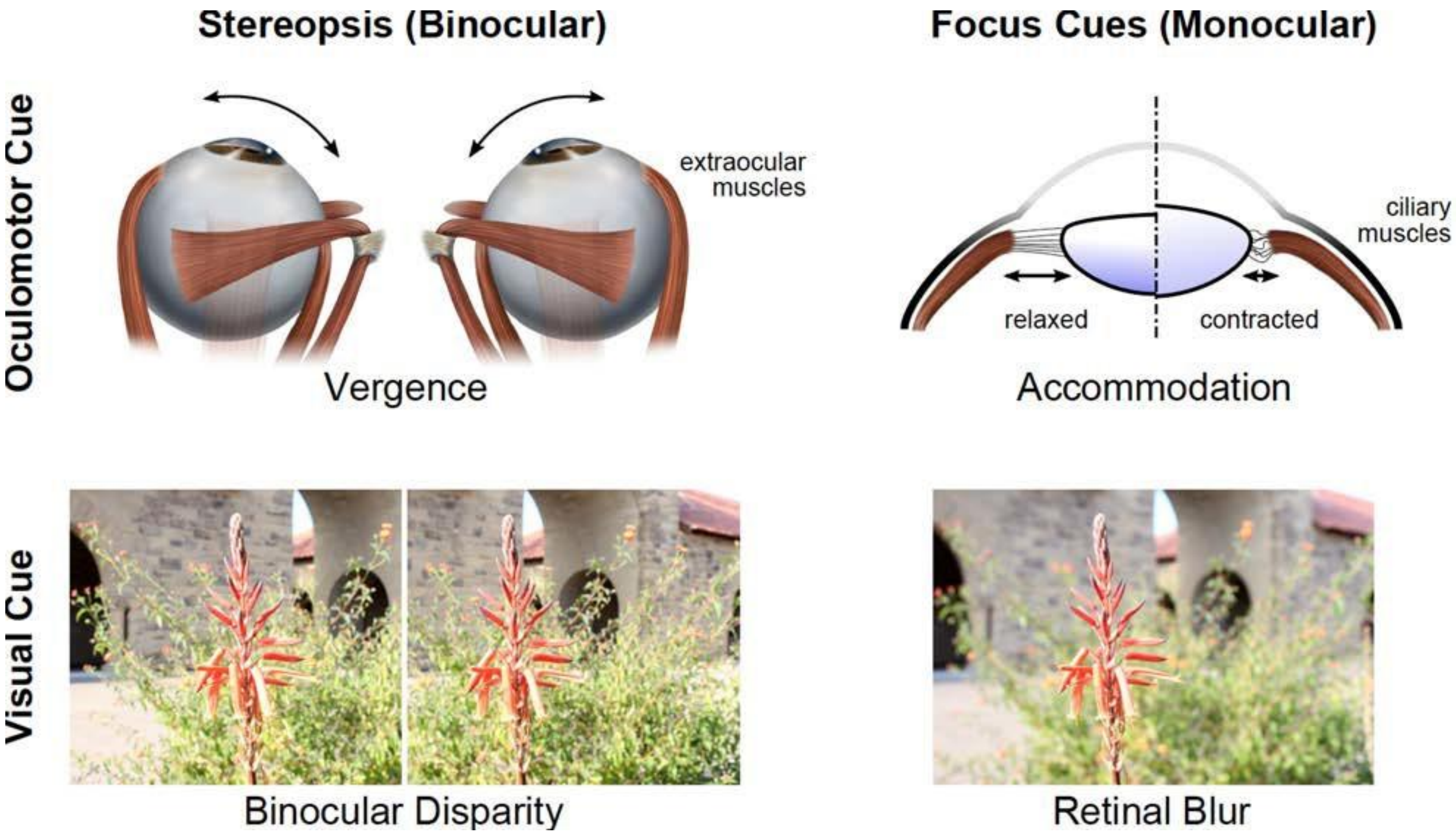




Stereo

- Passive (no tracking of eyes)
- Present each eye with perspective view corresponding to that eye's location relative to the other eye
- Eyes will con(verge) by rotating physically in sockets in order to bring closer and further objects into physical alignment on retina

Human Eye Muscles and Optical Controls

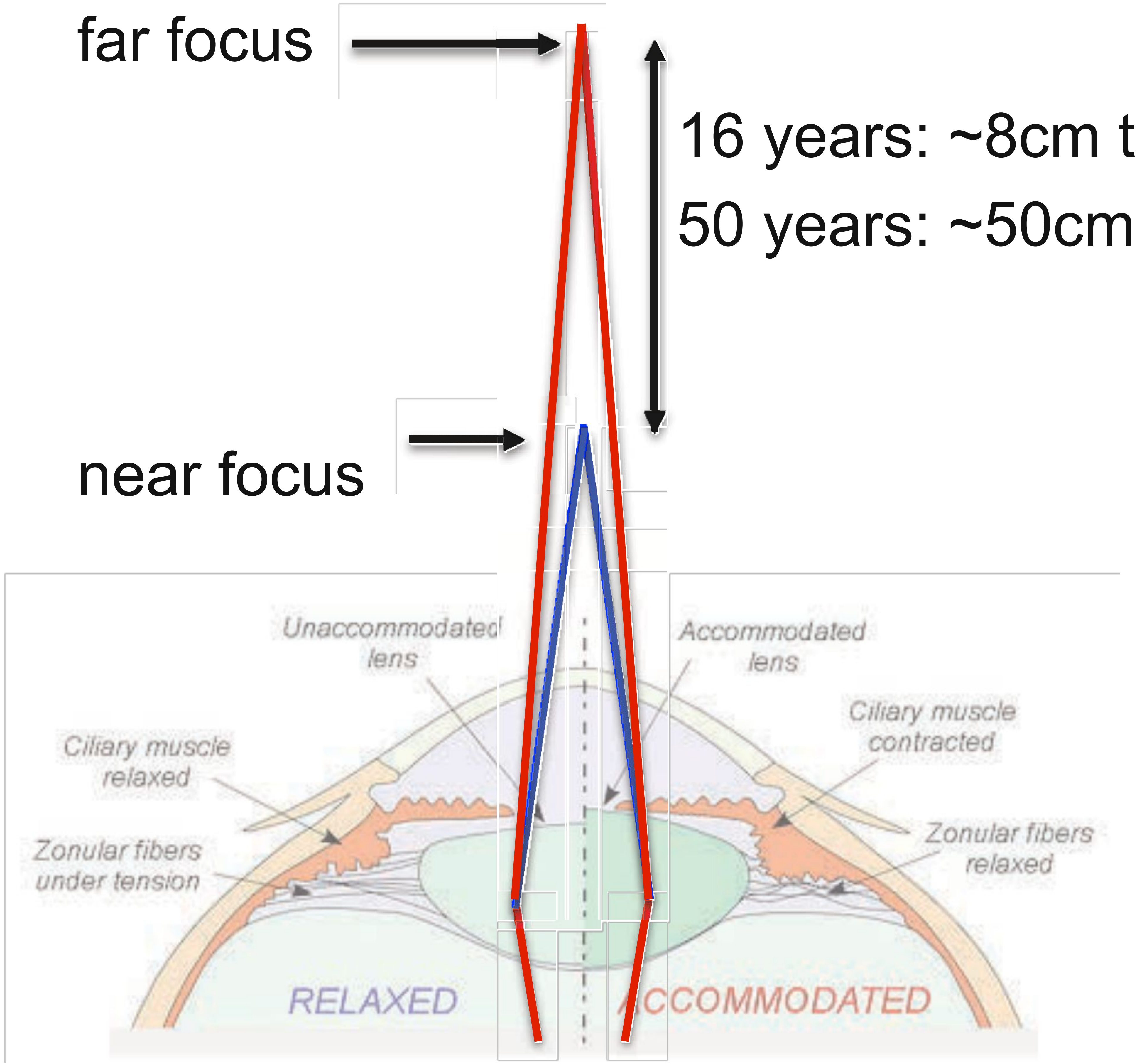


Slide credit: Gordon Wetzstein

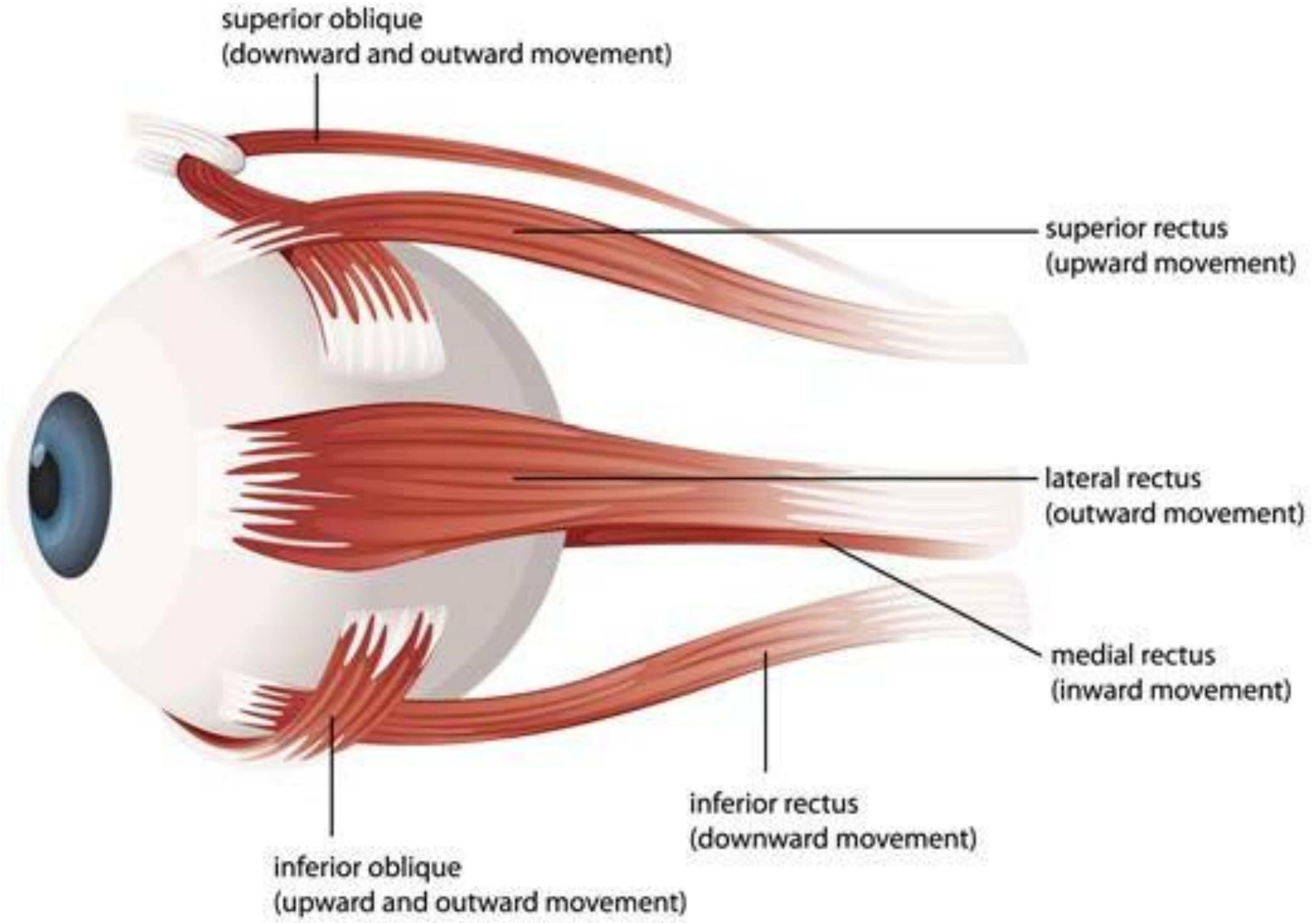
Human Eye Muscles and Optical Controls

16 years: ~8cm to "

50 years: ~50cm to " (mostly irrelevant)



adithyakiran.wordpress.com

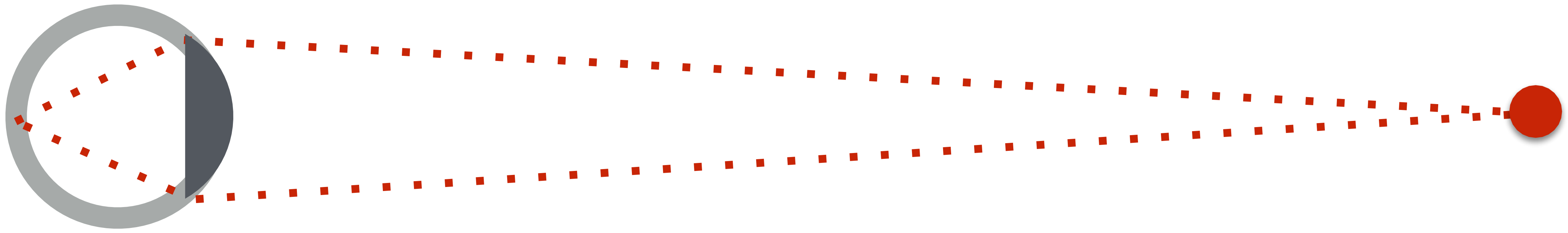

superior oblique (downward and outward movement)
superior rectus (upward movement)
lateral rectus (outward movement)
medial rectus (inward movement)
inferior rectus (downward movement)
inferior oblique (upward and outward movement)

Slide credit: Gordon Wetzstein

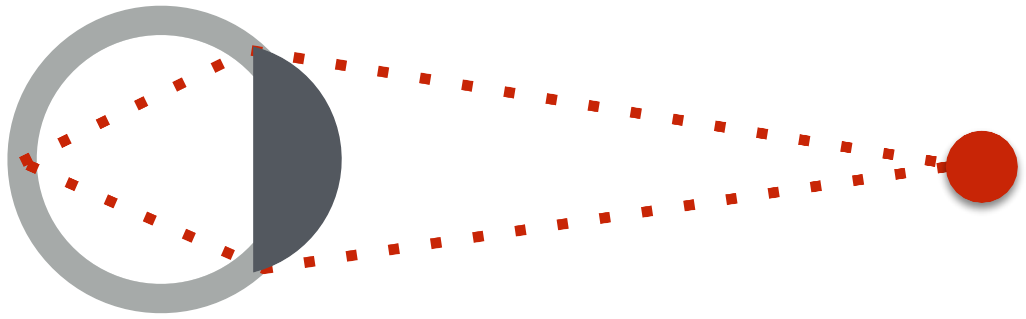
Accommodation and Vergence

Accommodation: changing the optical power of the eye (lens) to focus at different distances

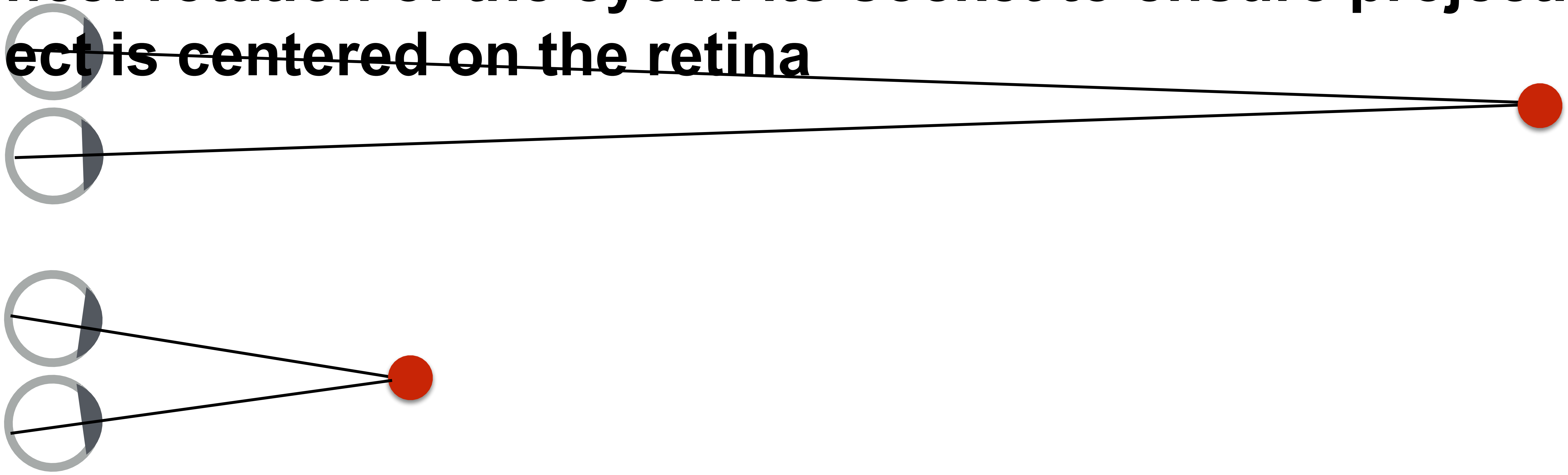
Eye accommodated to focus on a distant object



Eye accommodated to focus on a nearby object



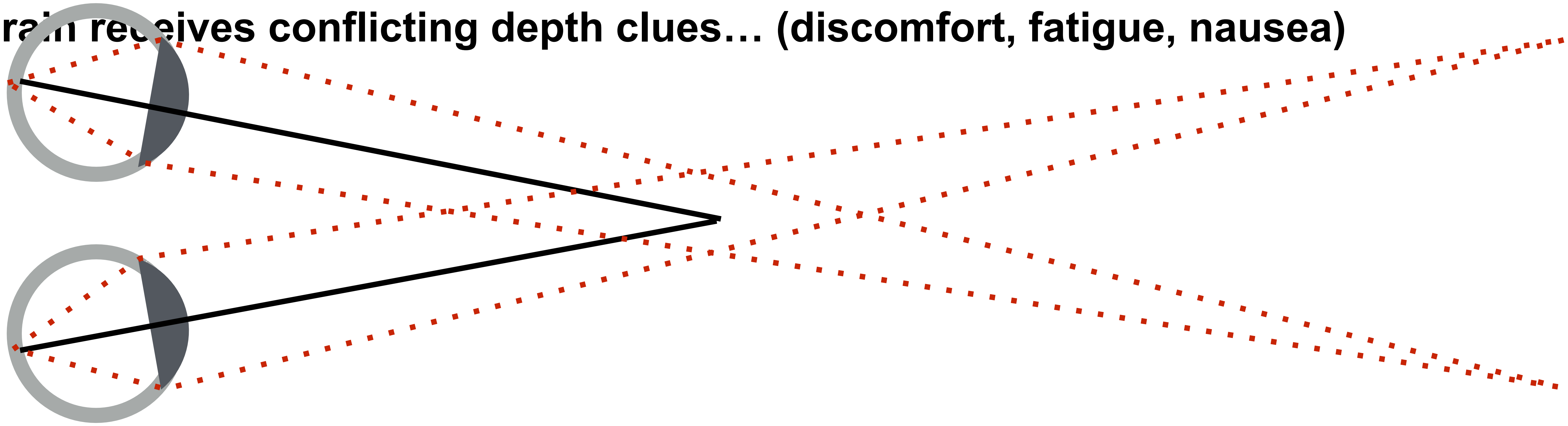
Vergence: rotation of the eye in its socket to ensure projection of object is centered on the retina



Accommodation – Vergence Conflict

Given design of current VR displays, consider what happens when objects are up-close to eye in virtual scene

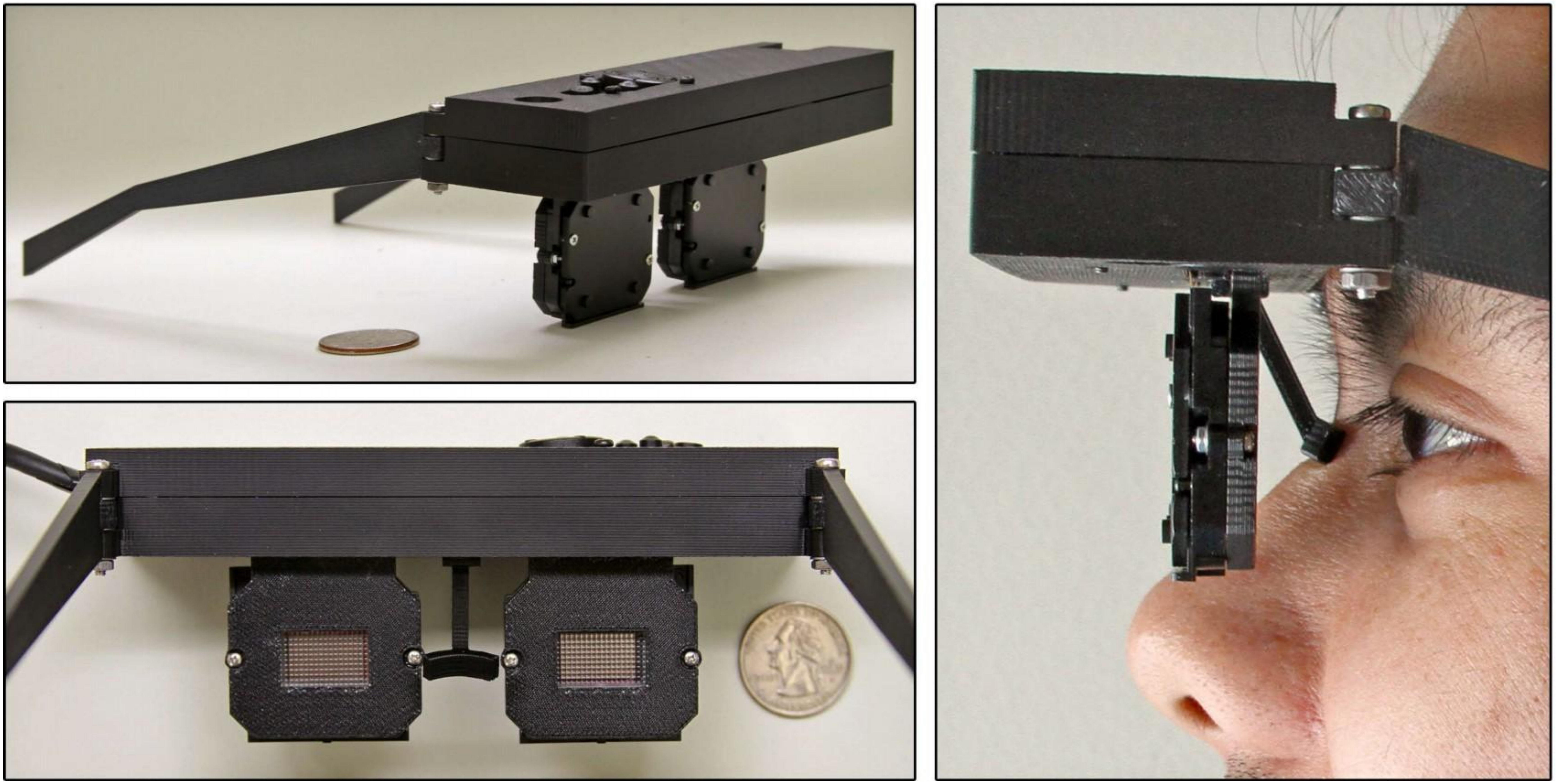
- Eyes must remain accommodated to far distance (otherwise image on screen won't be in focus)
- But eyes must converge in attempt to fuse stereoscopic images of object up close
- Brain receives conflicting depth clues… (discomfort, fatigue, nausea)



This problem stems from nature of display design. If you could just make a display that emits the light field that would be produced by a virtual scene, then you could avoid the accommodation - vergence conflict…

Aside: Research on Near-Eye Light Field Displays

Goal: recreate light field in front of eye



Lanman and Luebke, SIGGRAPH Asia 2013.

Display Requirements Derive From Human Perception

Example #4: Motion Parallax from Head Motion

Google Cardboard: Tracking Using Headset Camera

Tracking uses gyro / rear- facing camera to estimate user’s viewpoint

- 2D rotation tracking generally works well
- 3D positional tracking was a challenge in general environments

Environment-Supported Vision-Based Tracking?

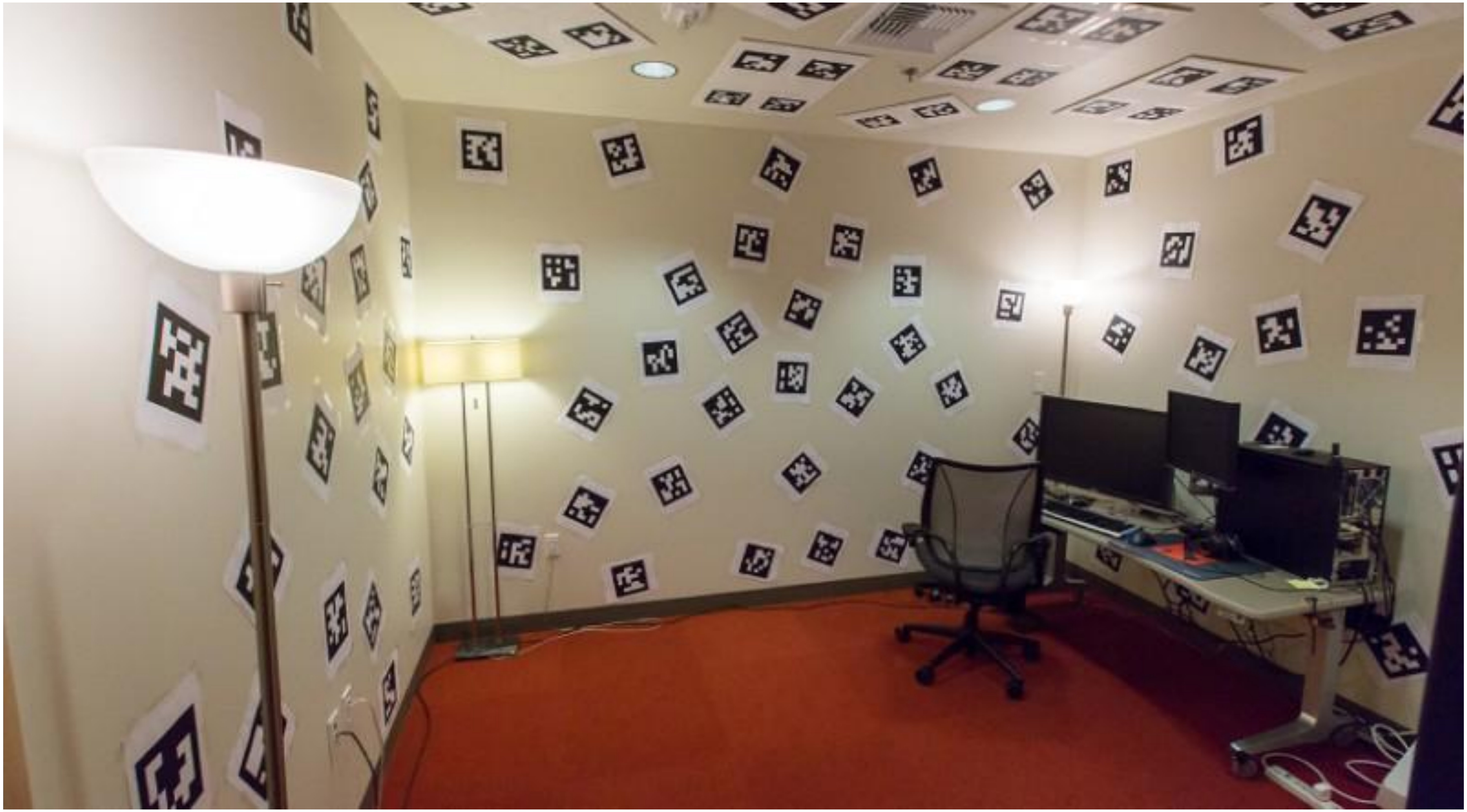


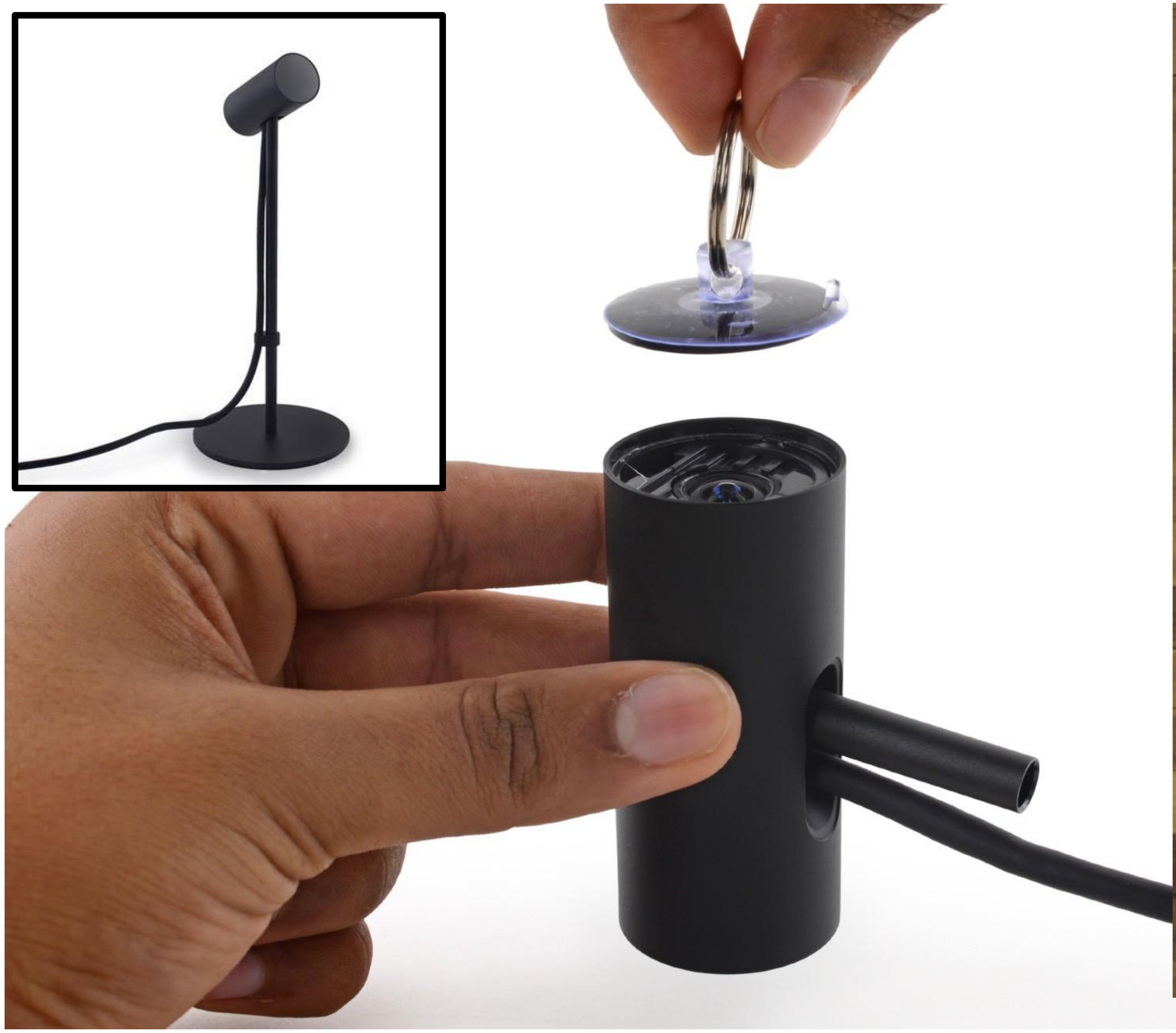
Image credit: gizmodo.com

Early VR test room at Valve, with markers positioned throughout environment

Oculus Rift IR LED Tracking System

Oculus Rift + IR LED sensor

Oculus Rift IR LED Tracking Hardware



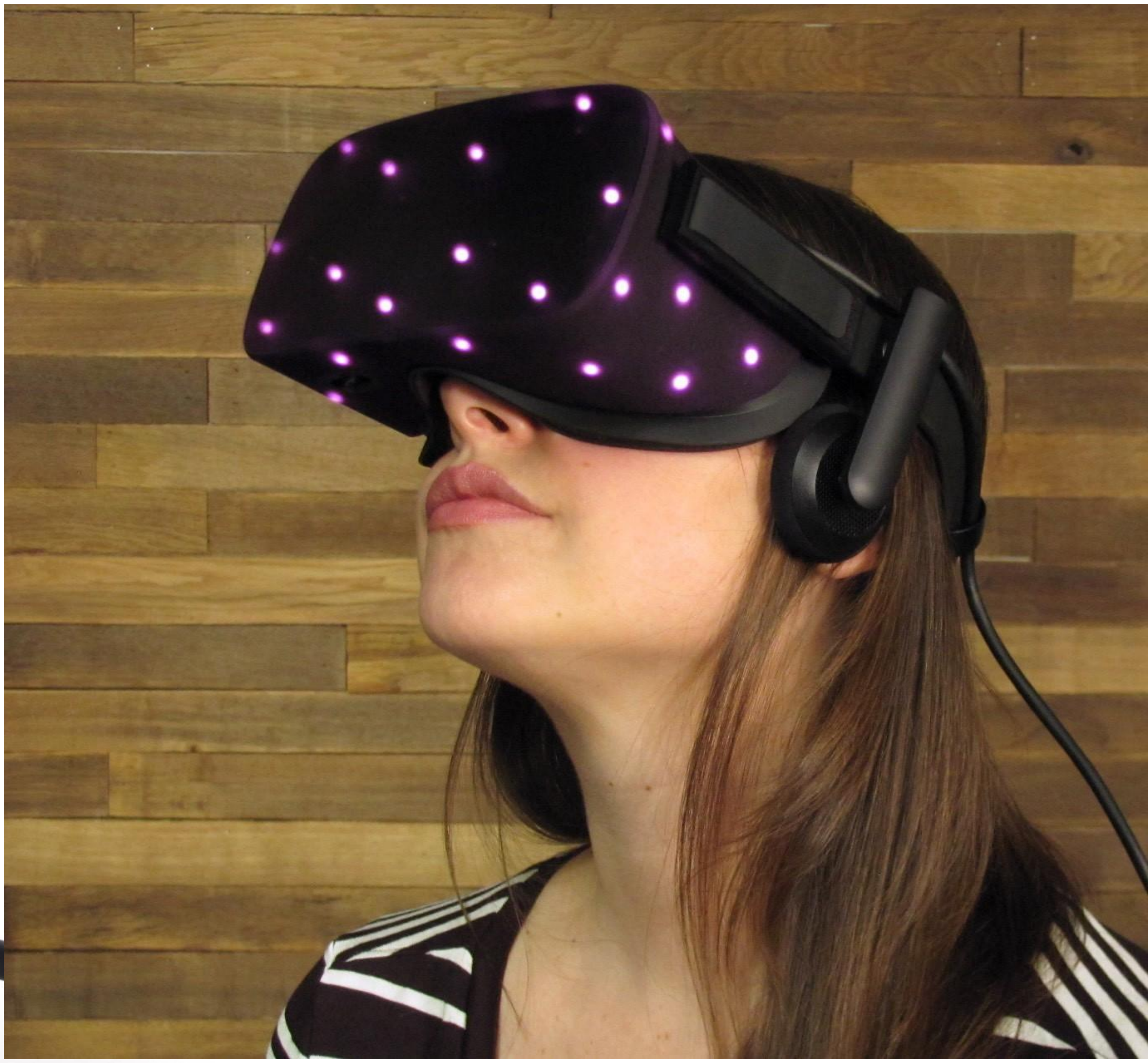
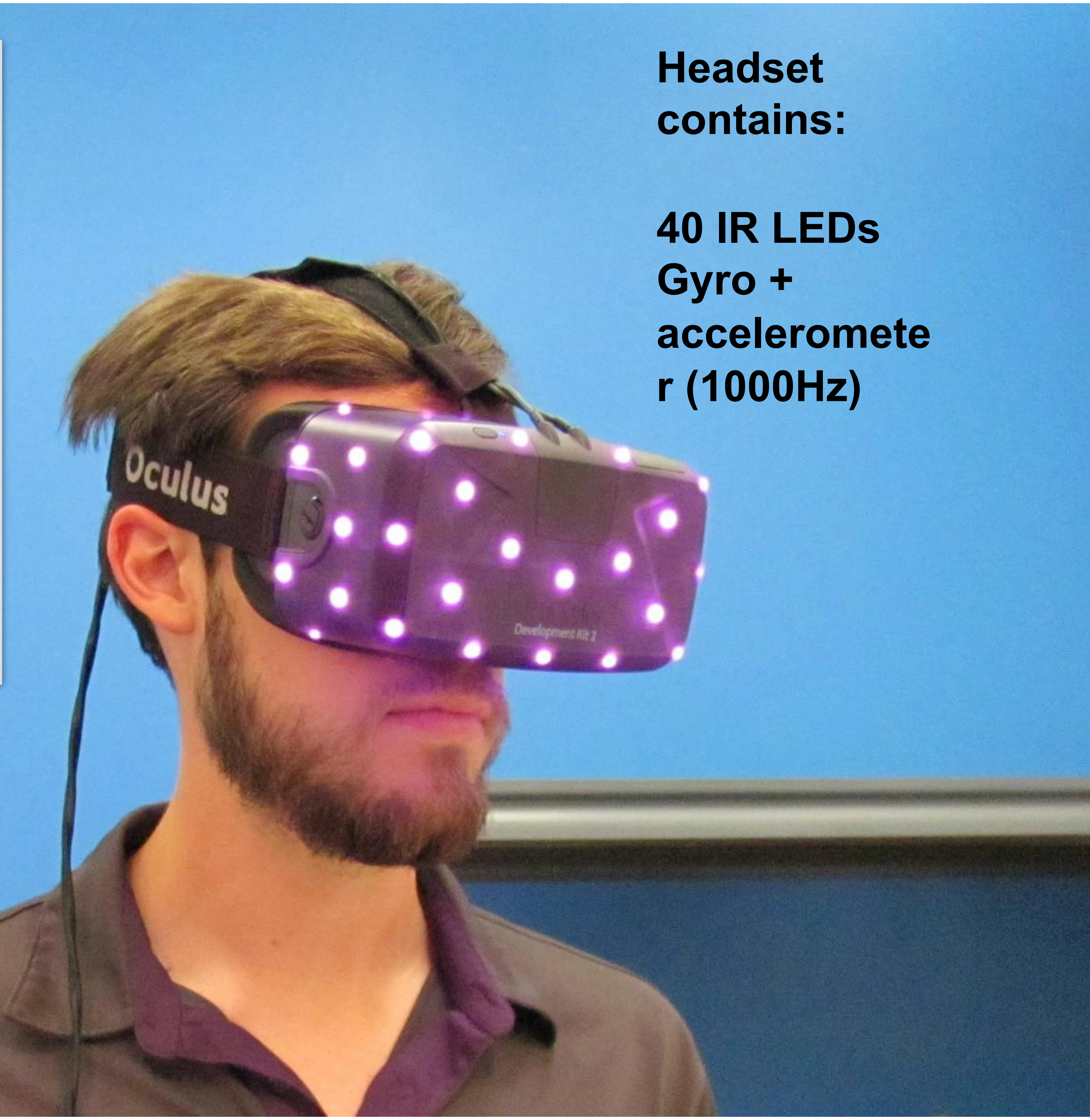


Photo taken with IR-sensitive camera

https://www.ifixit.com/Teardown/Oculus+Rift+Constellation+Teardown/61128

Oculus Rift LED Tracking System (DK2)

External 60Hz IR Camera



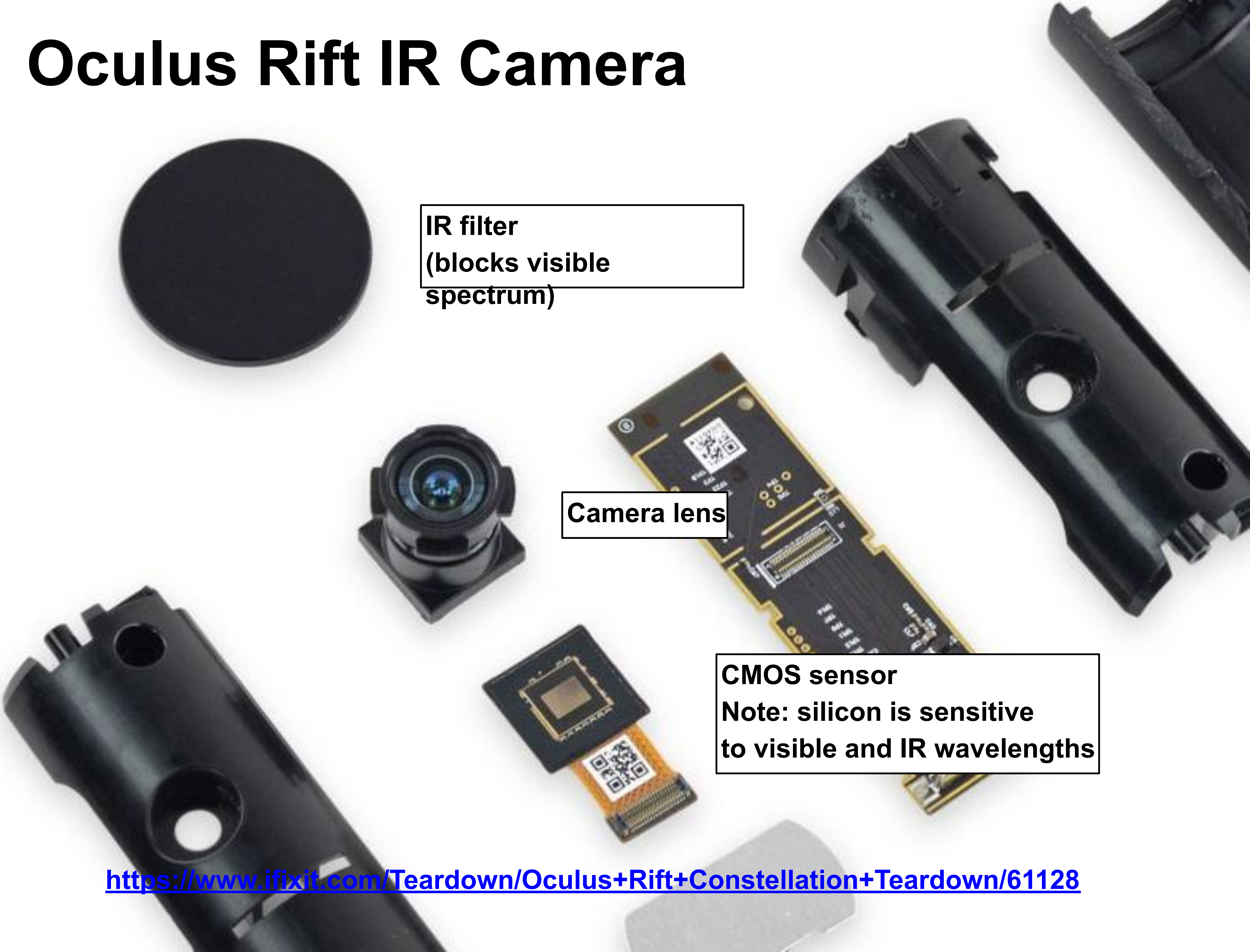
Headset contains:

40 IR LEDs
Gyro + acceleromete
r (1000Hz)

Photo taken with IR-sensitive camera (IR LEDs not visible in real life)

Image credit: ifixit.com

Oculus Rift IR Camera



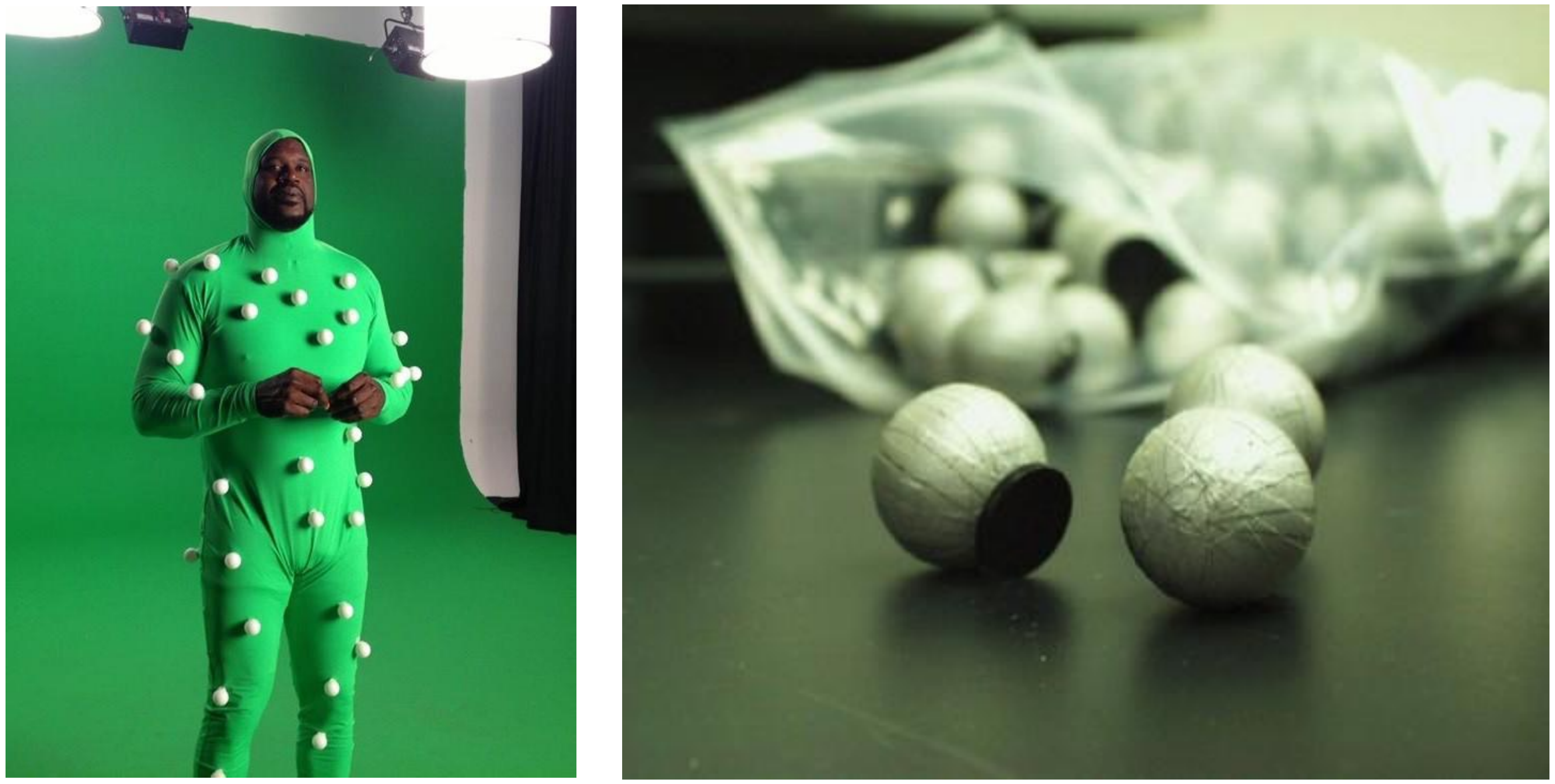
IR filter
(blocks visible
spectrum)

Camera lens

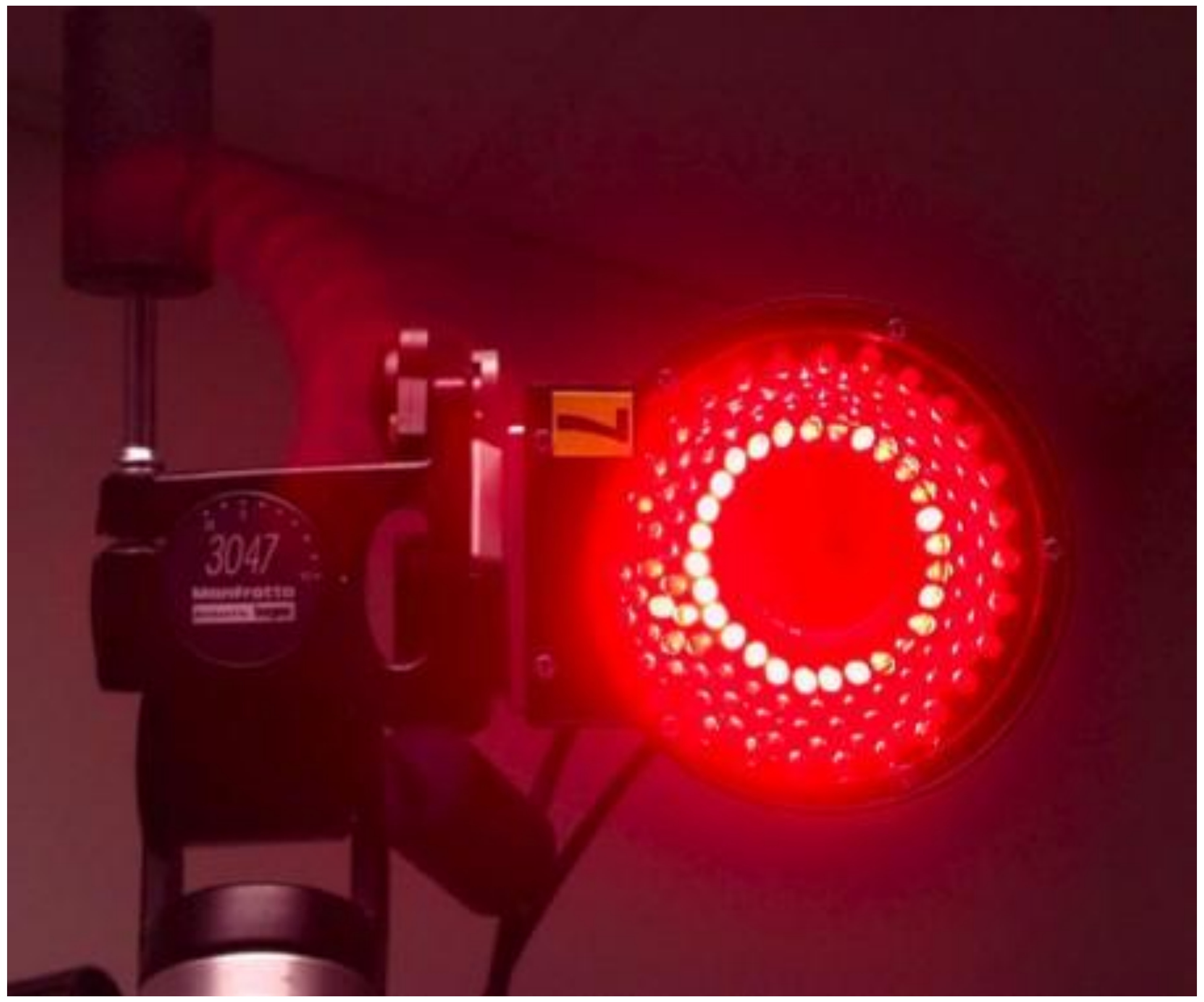
CMOS sensor
Note: silicon is sensitive
to visible and IR wavelengths

https://www.ifixit.com/Teardown/Oculus+Rift+Constellation+Teardown/61128

Recall: Passive Optical Motion Capture



Retroflective markers attached to subject



IR illumination and cameras

- Markers on subject
- Positions by triangulation from multiple cameras
- 8+ cameras, 240 Hz, occlusions are difficult

Slide credit: Steve Marschner

Active Optical Motion Capture

- **Each LED marker emits unique blinking pattern (ID)**

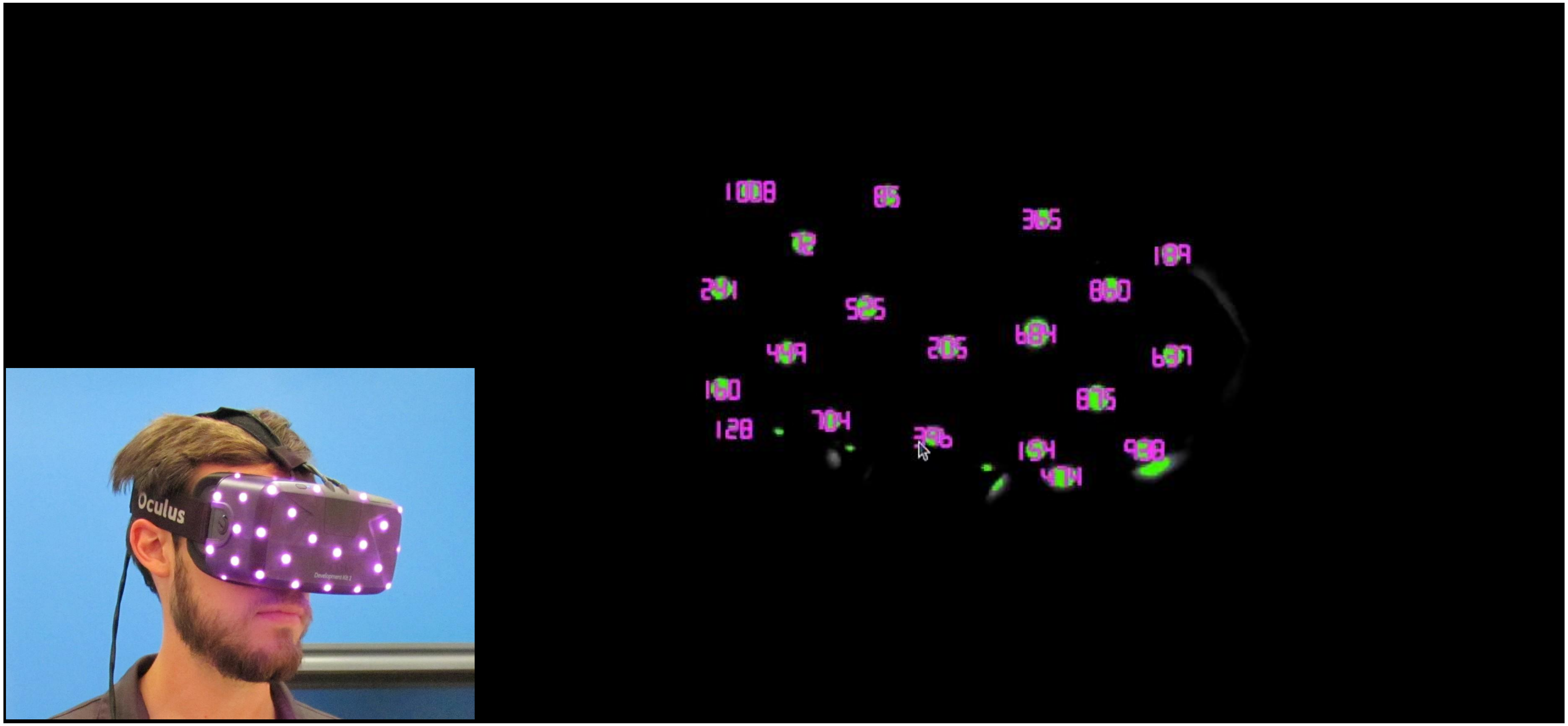
- **Reduce marker ambiguities / unintended swapping**

- **Ha re Ds**

Phoenix Technology

Phase Space

Oculus Rift Uses Active Marker Motion Capture



Credit: Oliver Kreylos, https://www.youtube.com/watch?v=O7Dt9Im34OI

- **Motion capture: unknown shape, multiple cameras**
- **VR head tracking: known shape, single camera**

6 DOF Head Pose Estimation

Head pose: 6 degrees of freedom (unknowns)

- 3D position and 3D rotation of headset (e.g. can represent as 4x4 matrix)

Inputs:

- Fixed: relative 3D position of markers on headset (e.g. can represent each marker offset as 4x4 matrix)
- Fixed: camera viewpoint (ignoring distortion, also a 4x4 projective mapping of 3D scene to 2D image)
- Each frame: 2D position of each headset marker in

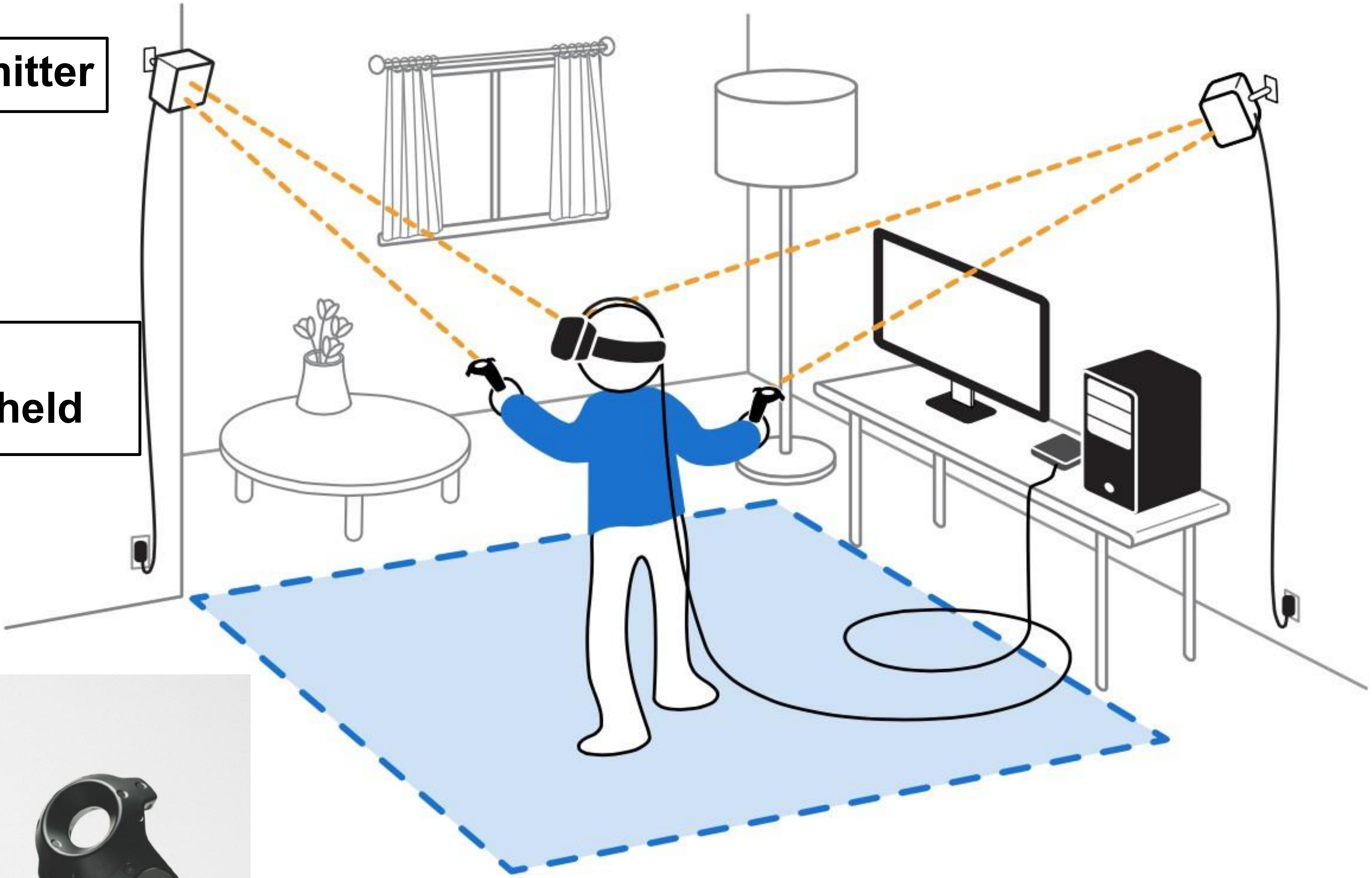
image Pose calculation:

- Write down equations mapping each marker to image pixel location as a function of 6 degrees of freedom
- Solve for 6 degrees of freedom (e.g. least squares)

HTC Vive Tracking System (“Lighthouse”)

Structured light transmitter

Photodiode arrays on headset and hand-held controllers



Vive Headset & Controllers Have Array of IR Photodiodes

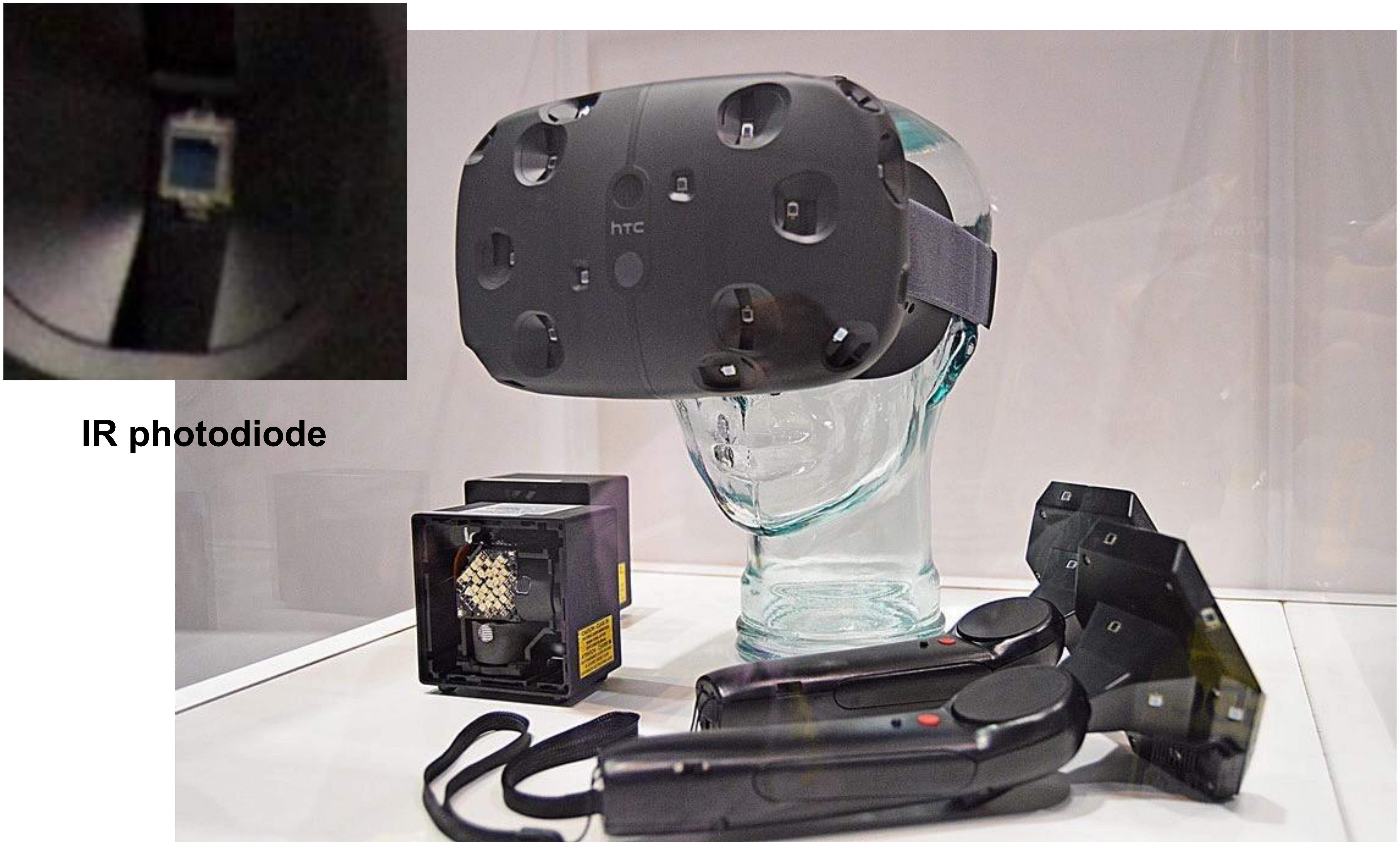
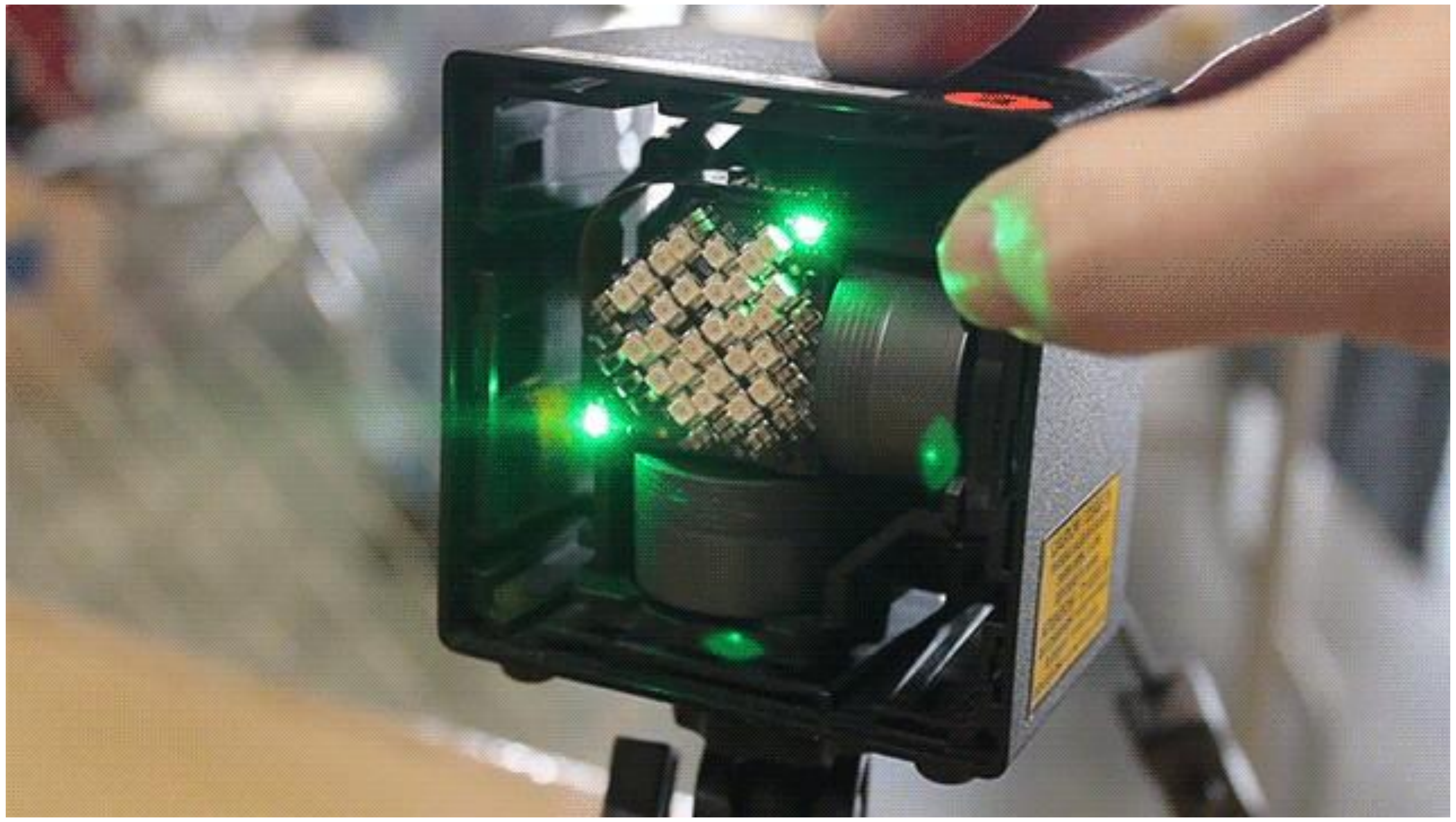
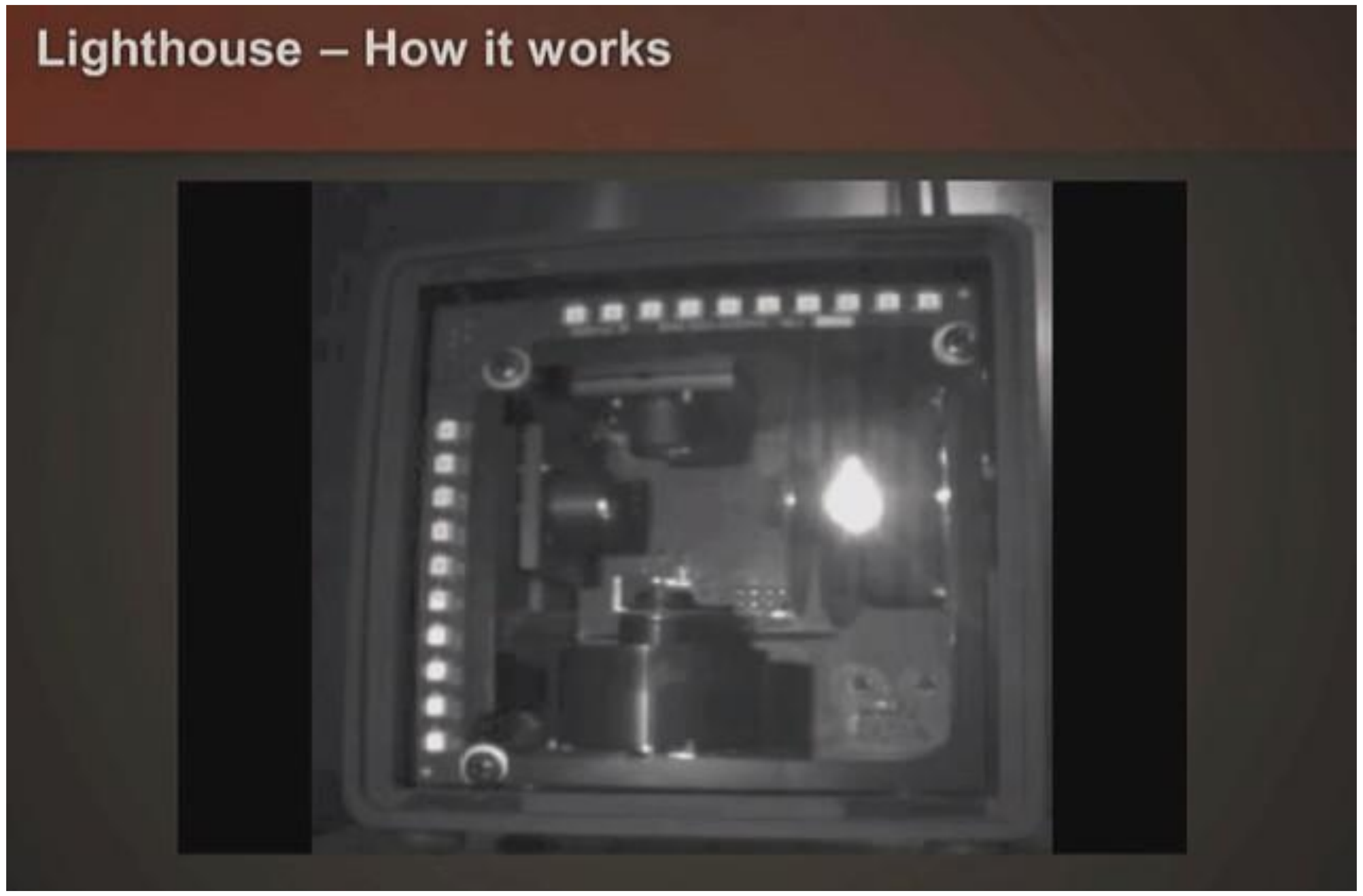


Image credit: uploadvr.com

(Prototype) Headset and controller are covered with IR photodiodes

HTC Vive Structured Light Emitter (“Lighthouse”)





Light emitter contains array of LEDs (white) and two spinning wheels with lasers

Sequence of LED flash and laser sweeps provide structured lighting throughout room

Credit: Gizmodo:
http://gizmodo.com/this-is-how-valve-s-amazing-lighthouse-tracking-technol-1705356768

HTC Vive Tracking System

For each frame, lighthouse does the following:

- LED pulse, followed by horizontal laser sweep
- LED pulse, followed by vertical laser sweep

Each photodiode on headset measures time offset between pulse and laser arrival

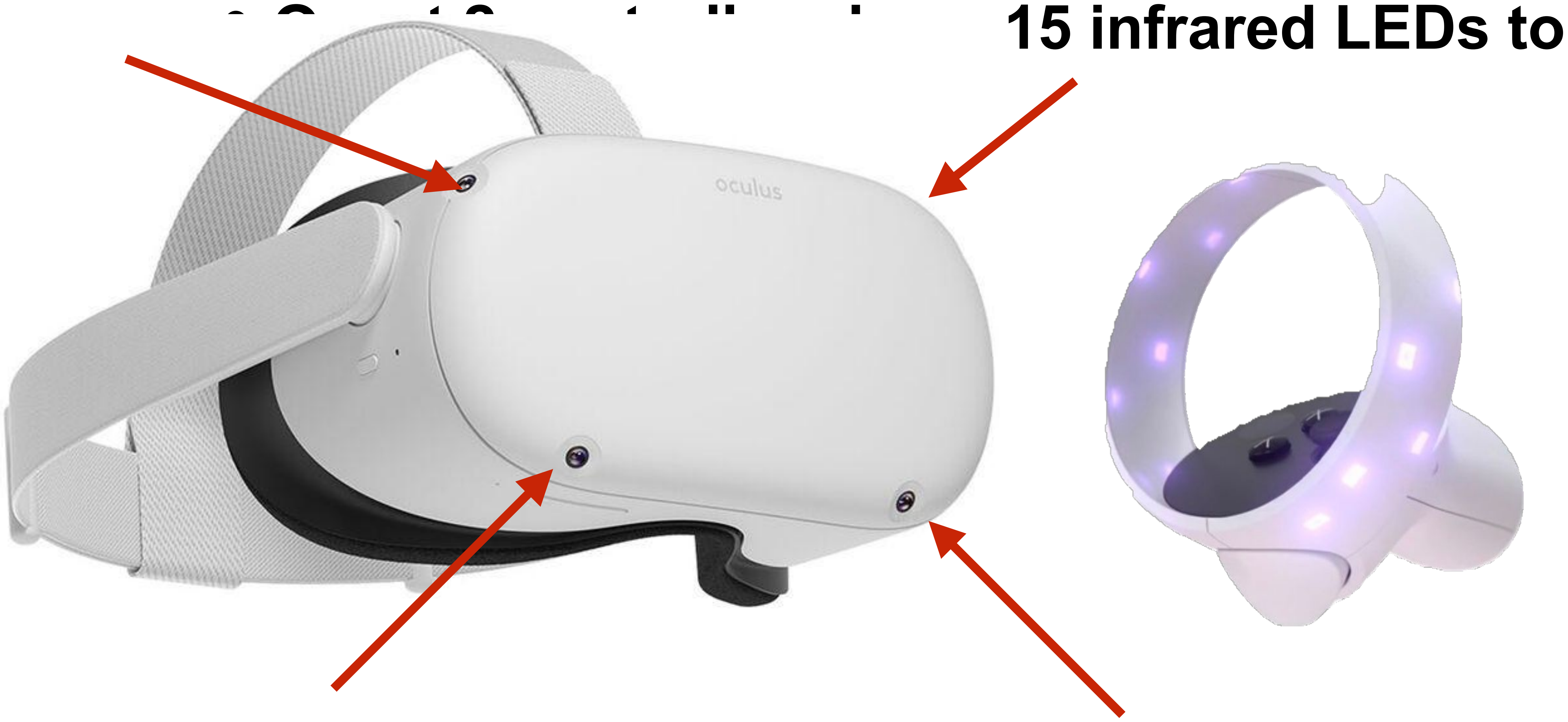
- Determines the x and y offset in the lighthouse's field of view
- In effect, obtain an image containing the 2D location of each photodiode in the world
  - (Can think of the lighthouse as a virtual "camera")

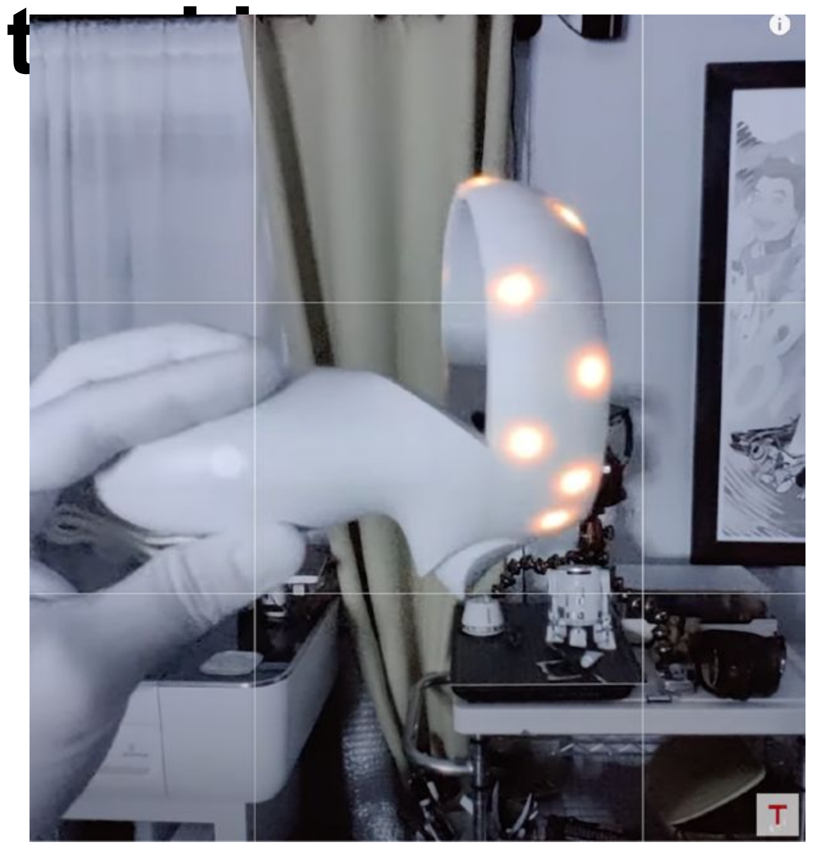
HTC Vive Tracking System (“Lighthouse”)

Credit: rvdm88 / youtube. https://www.youtube.com/watch?v=J54dotTt7k0

Many Modern Systems Use “Inside Out” Tracking

- Wide-angle cameras look outward from headset
- Use computer vision (SLAM) to estimate 3D structure of world and position/orientation of camera in the world
- These cameras also track the position/orientation of the controllers
  - Quest 2 controllers have 15 infrared LEDs to aid tracking





View of controller through infrared camera

(credit Adam Savage’s Testbed)

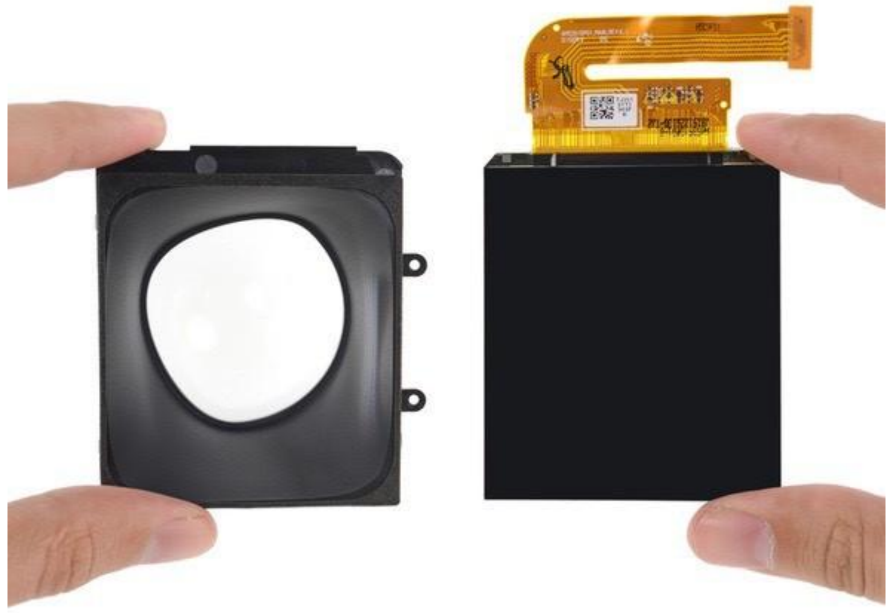
Tracking Summary

Looked at a few tracking methods

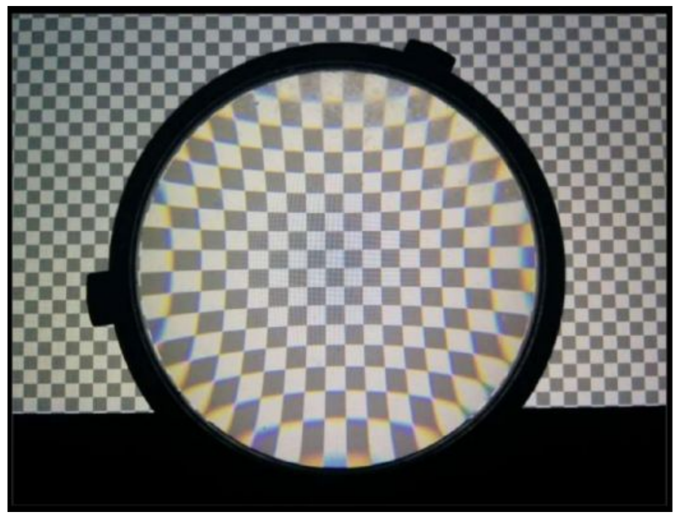
- Camera on headset + computer vision + gyro
- External camera + marker array on headset
- External structured light + sensor array on headset
- "Inside out" tracking

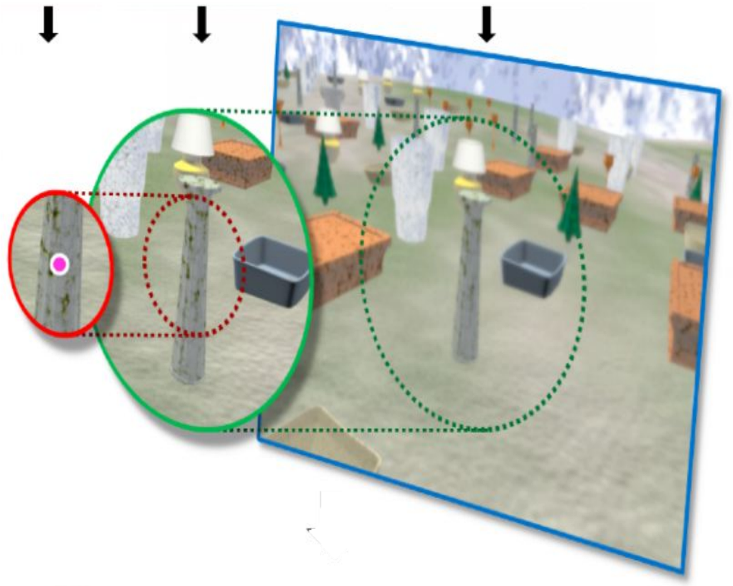
Overview of VR Topics

- VR Displays

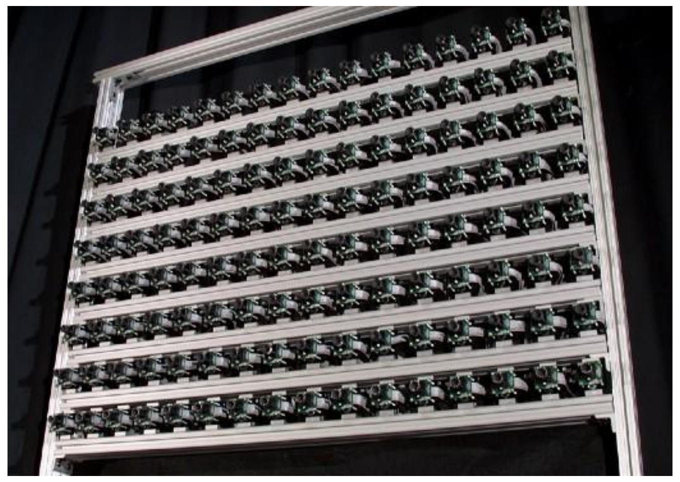


- VR Rendering



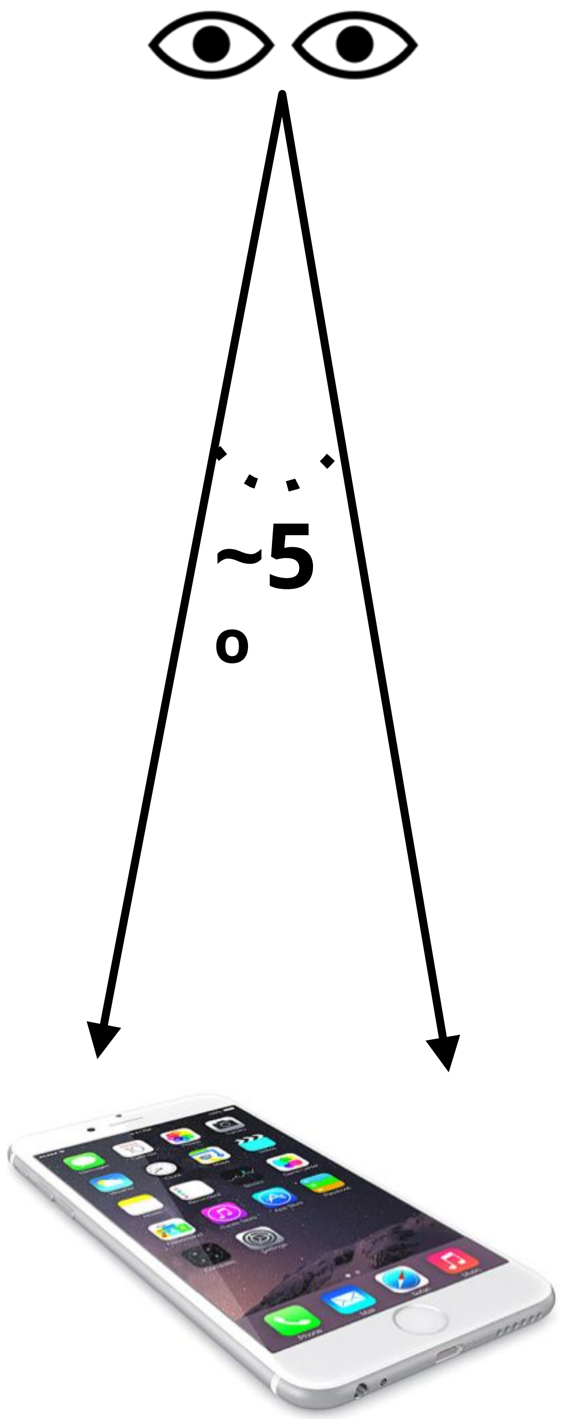


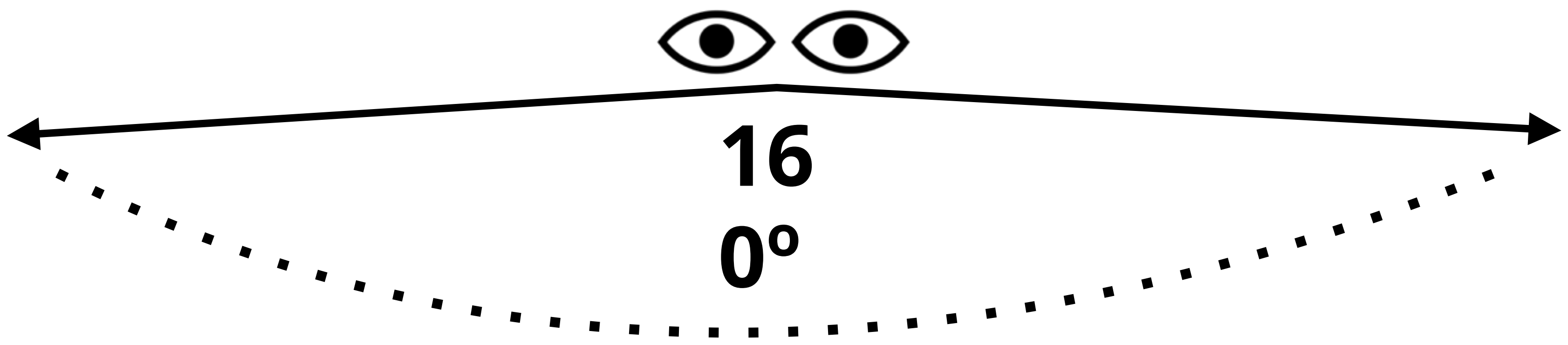
- VR Imaging



Rendering Latency in VR

A VR Display at Human Visual Acuity





Human: ~160° view of field per eye (~200° overall) (Note: does not account for eye's ability to rotate in socket)

Future "retina" VR display:
~ 8K x 8K display per eye (50 ppd)
= 128 MPixel

iPhone 6: 4.7 in "retina" display:
1.3 MPixel
326 ppi → ~60 ppd

Strongly suggests need for eye tracking and foveated rendering (eye can only perceive detail in 5° region about gaze point)

Eye icons designed by SuperAtic LABS from the thenounproject.com

Latency Requirements in VR Are Challenging

The goal of a VR graphics system is to achieve “presence”, tricking the brain into thinking what it is seeing is real

Achieving presence requires an exceptionally low-latency system

- What you see must change when you move your head!
- End-to-end latency: time from moving your head to the time new photons hit your eyes
  - Measure user’s head movement
  - Update scene/camera position
  - Render new image
  - Transfer image to headset, then transfer to display in headset
  - Actually emit light from display (photons hit user’s eyes)
- Latency goal of VR: 10-25 ms
  - Requires exceptionally low-latency head tracking
  - Requires exceptionally low-latency rendering and display

Thought Experiment: Effect of Latency

Consider 2,000 x 2,000 display spanning 100° field of view

- 20 pixels per degree

Assume:

- You move your head 90° in 1 second (only modest speed)
- End-to-end latency of system is a slow 33 ms (1/30 sec)

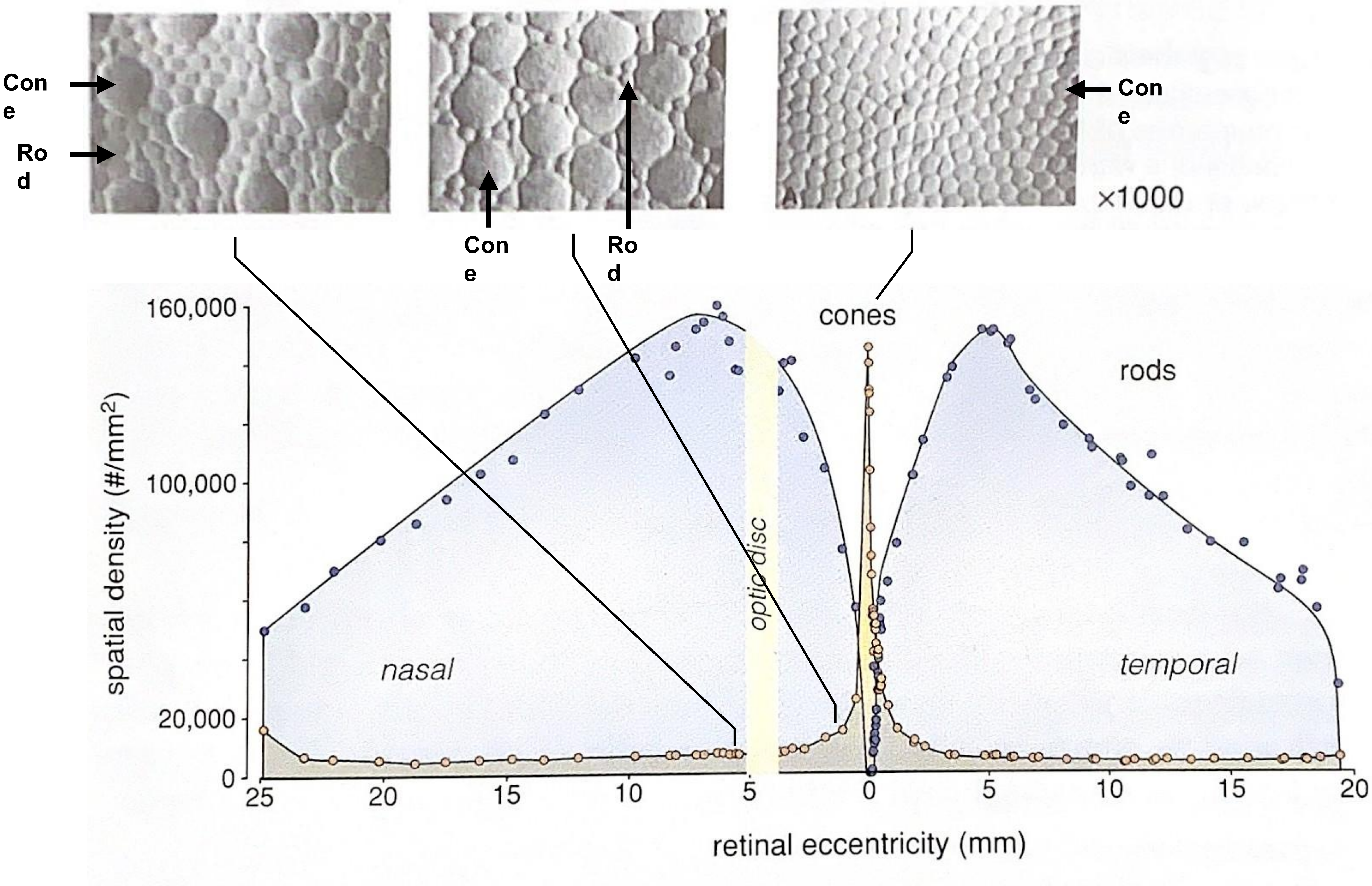
Result:

- Displayed pixels are off by 3.0° ~ 60 pixels from where they would be in an ideal system with 0 latency

Example credit: Michael

Rendering Challenge: Low Latency and High Resolution Require High Rendering Speed

Recall: Photoreceptor Size and Distribution Across Retina



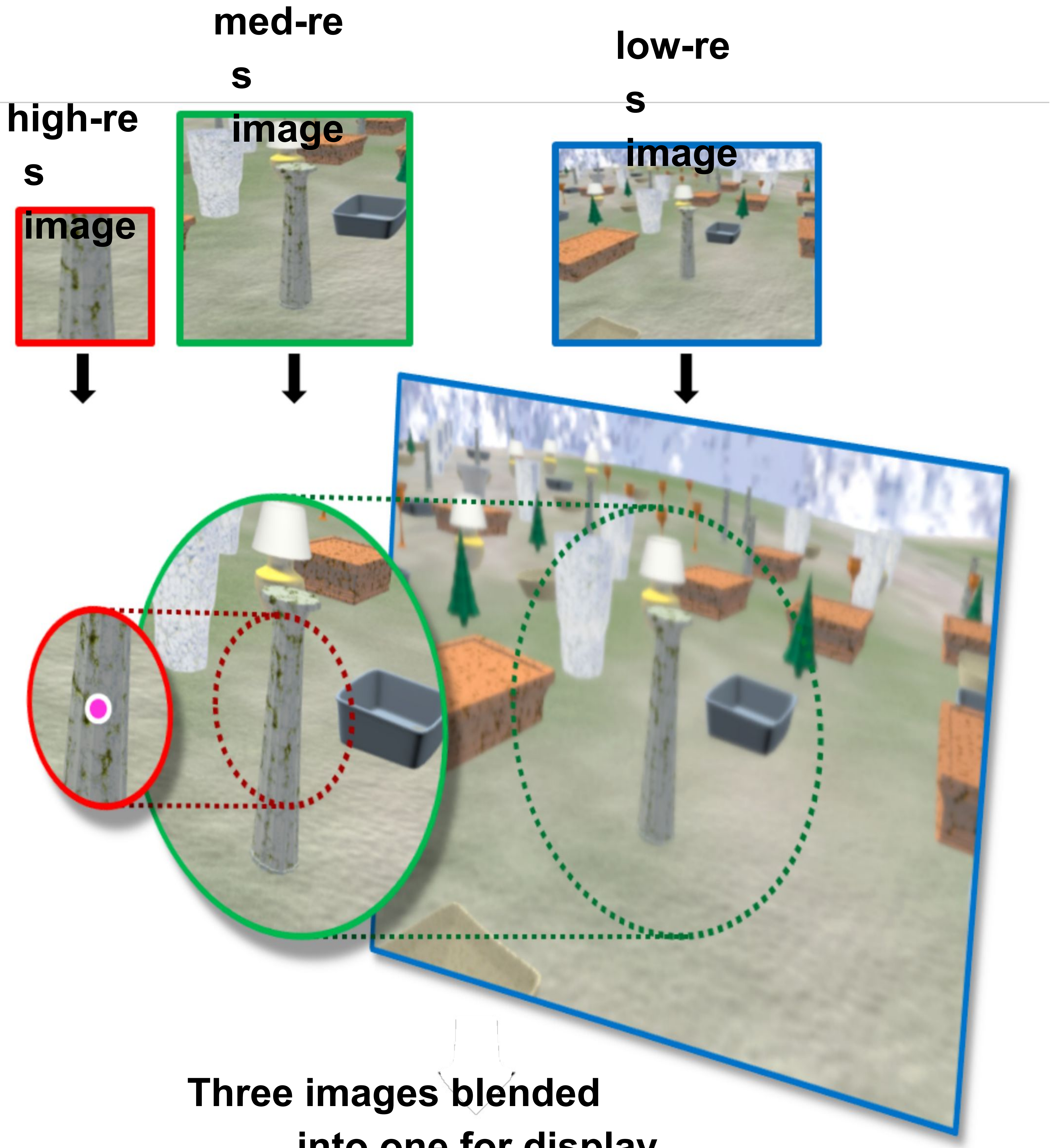
after Østerberg, 1935; as modified by Rodieck, 1988

Rodieck, p. 42

Foveated Rendering

Idea: track user’s gaze, render with increasingly lower resolution farther away from gaze point

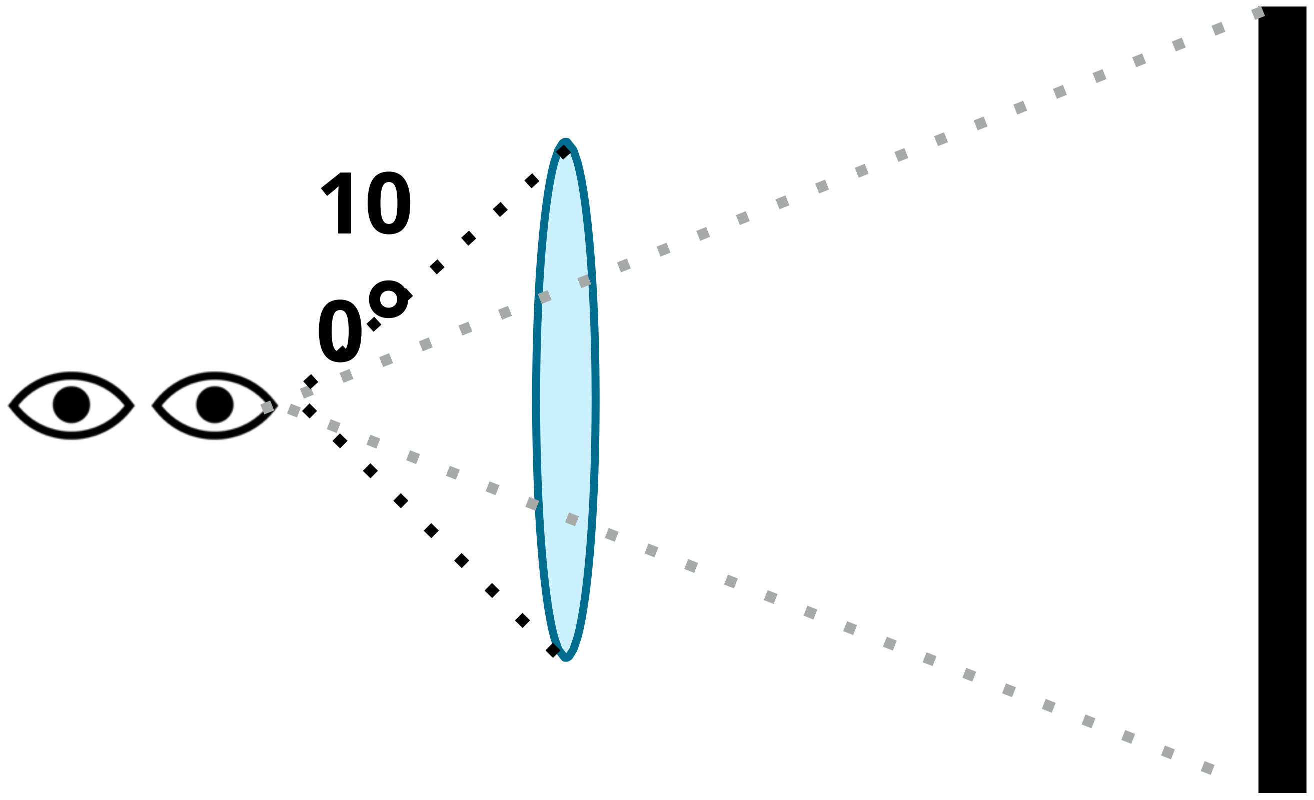
VR headset with eye tracker: HTC Vive Pro Eye



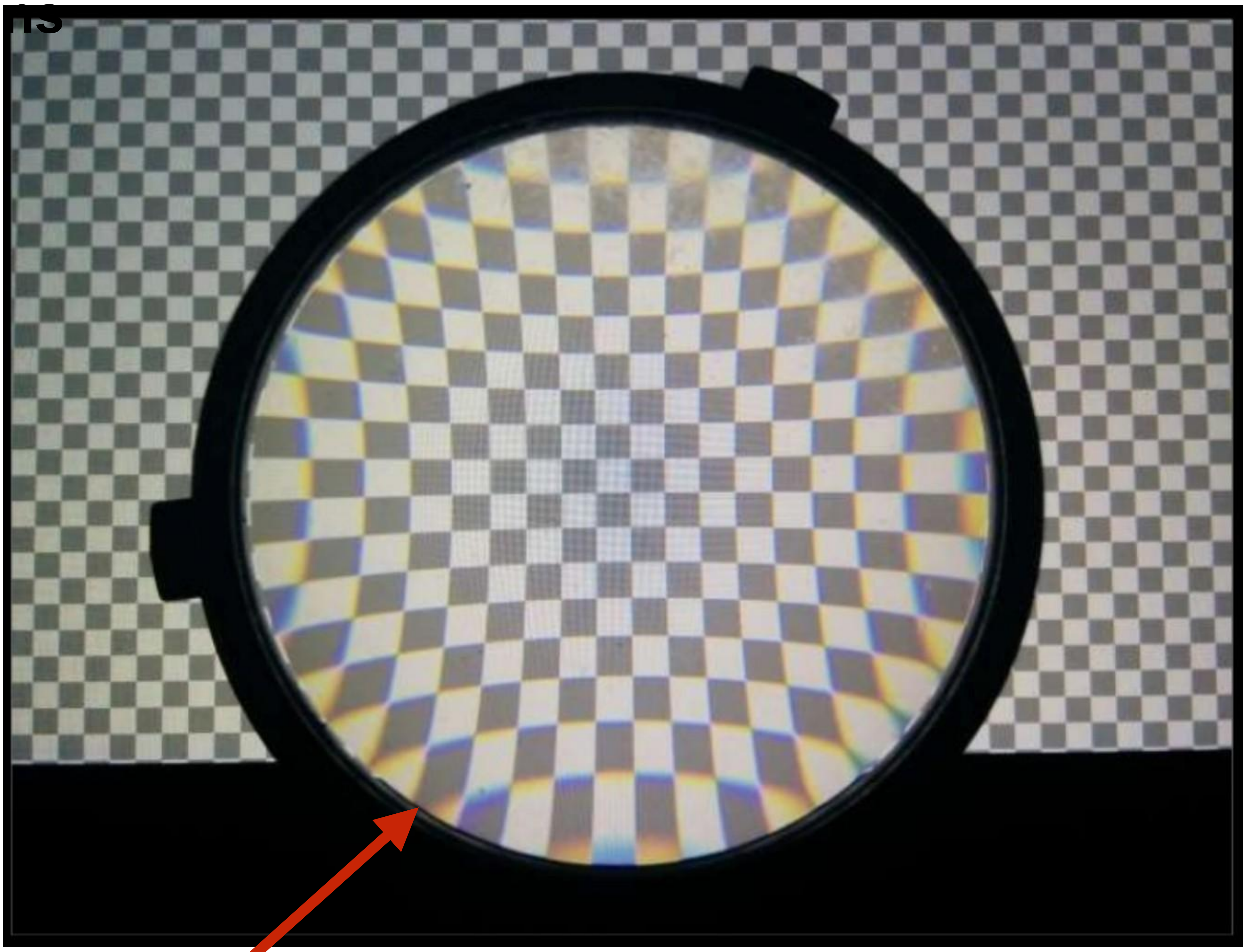
Three images blended into one for display

Rendering Challenge: Optical Distortion in VR Headset Viewing

Requirement: Wide Field of View



View of checkerboard through Oculus Rift (DK2) lens

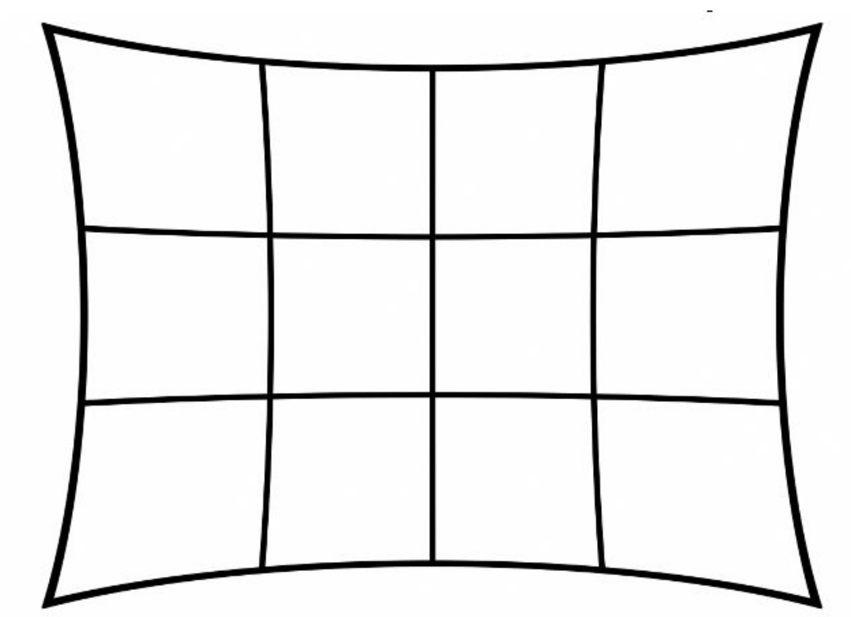


Lens introduces distortion

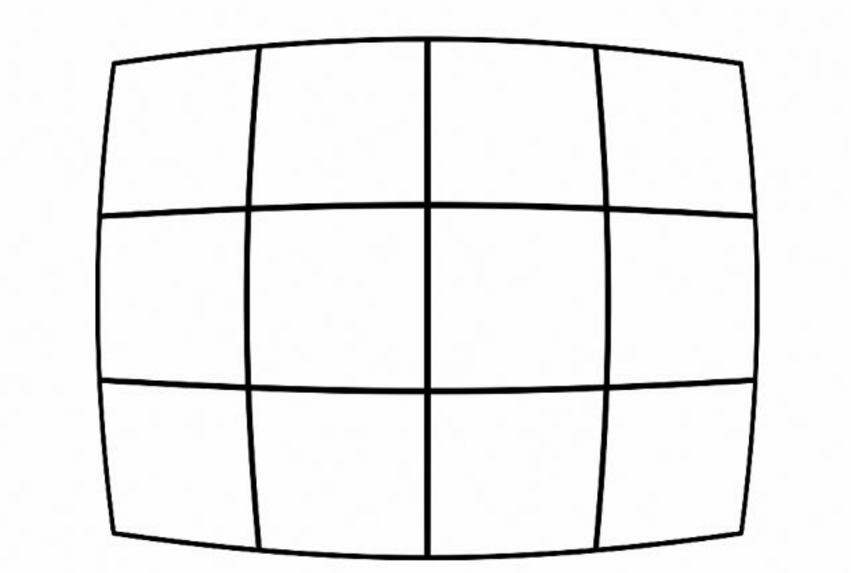
- Pincushion distortion
- Chromatic aberration (different wavelengths of light refract by different amount)

Icon credit: Eyes designed by SuperAtic LABS from the thenounproject.com Image credit: Cass Everitt

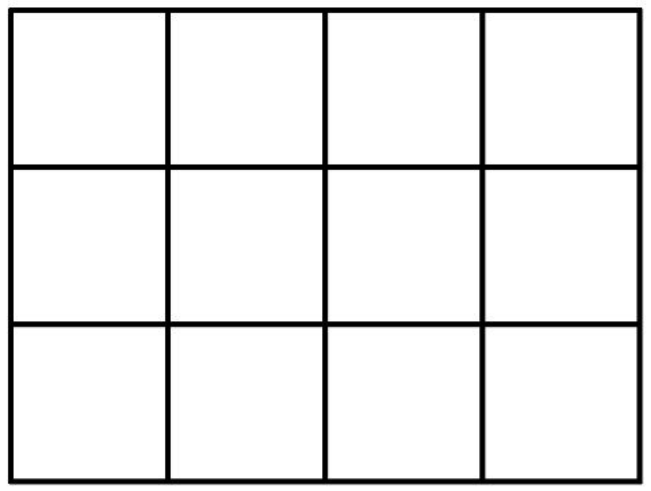
Software Correction of Lens Distortion in Photography



Pincushion distortion

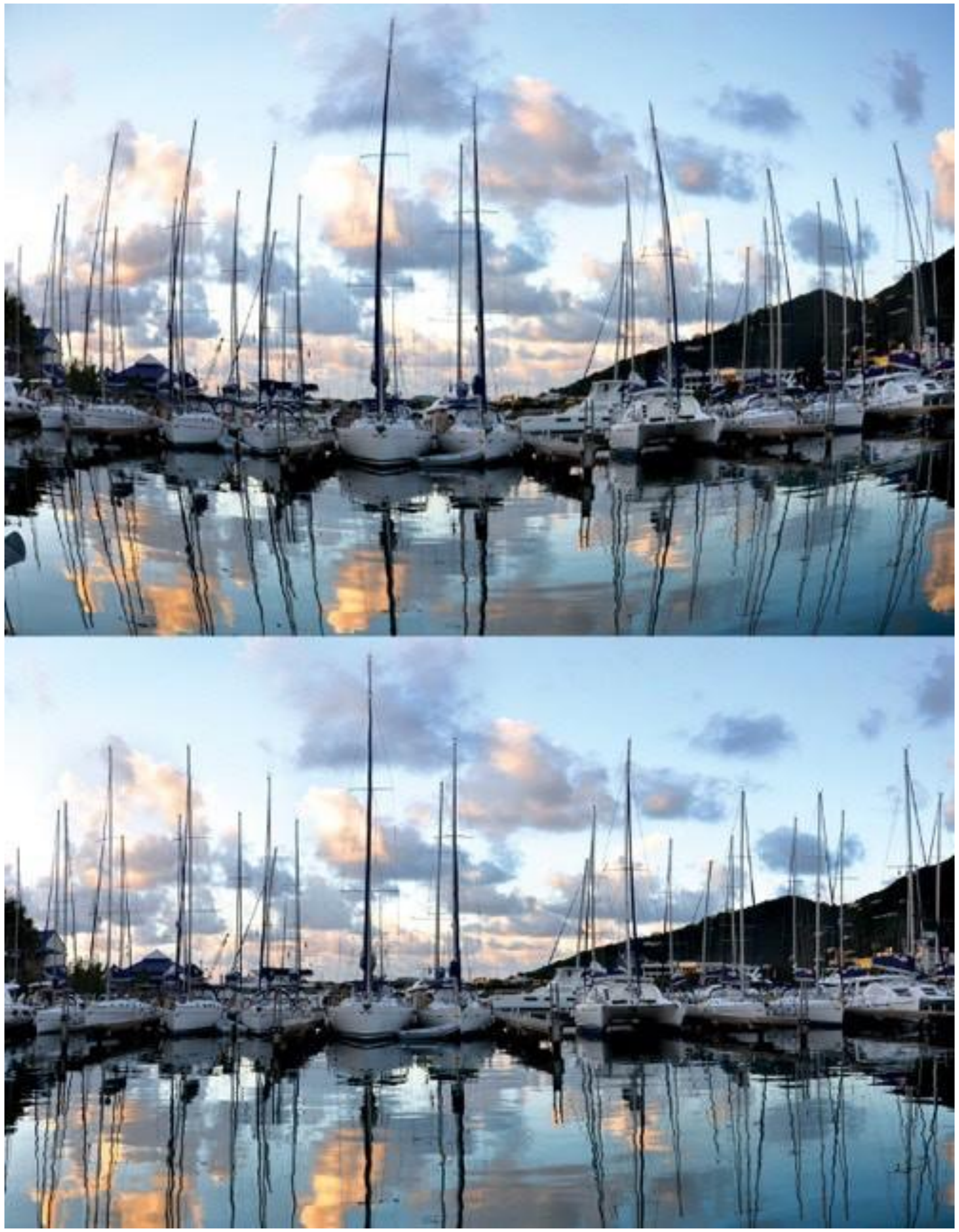


Barrel distortion



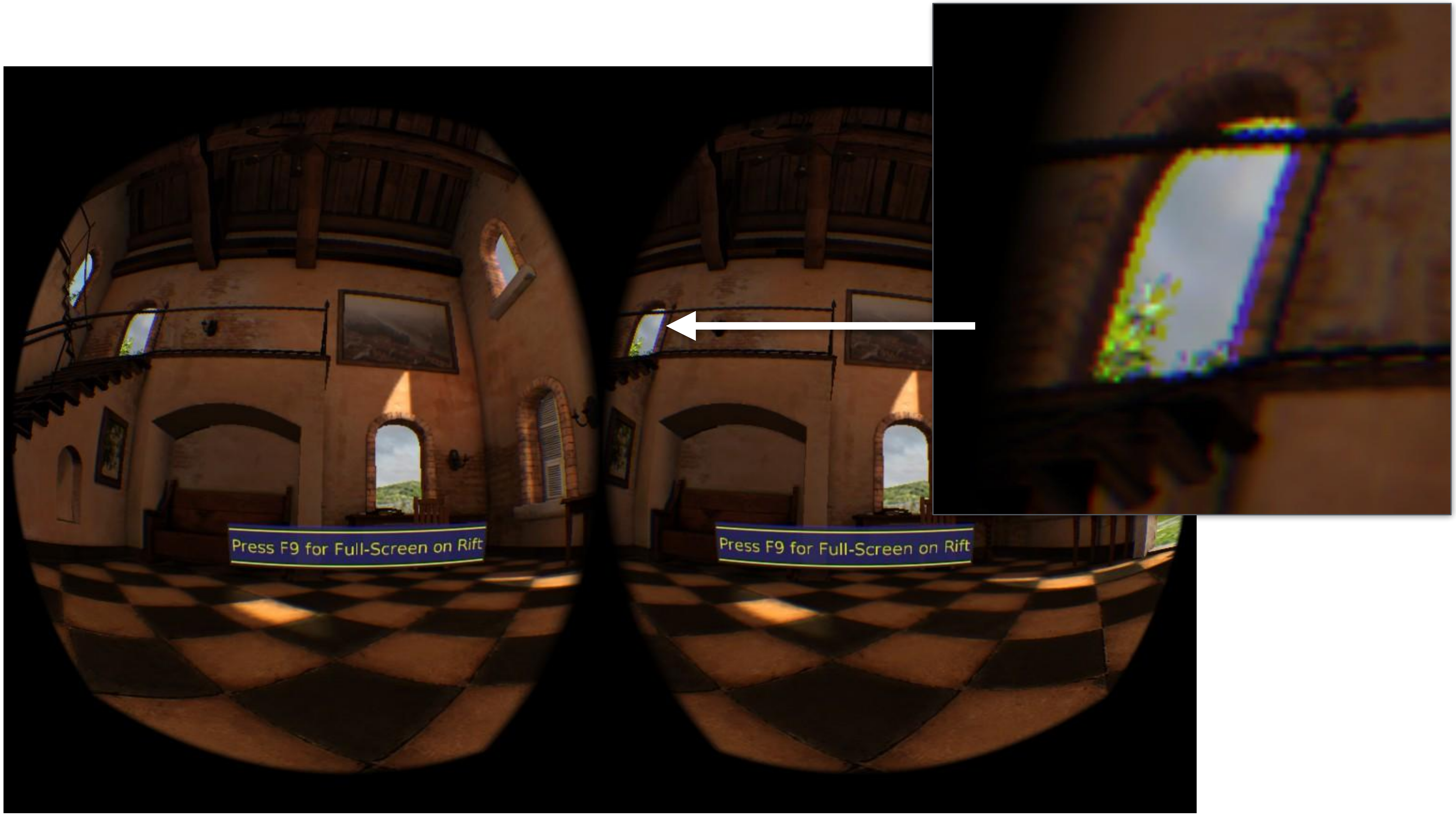
Rectilinear

m43photo.blogspot.com



Credit: The Photoshop Creative Team
http://blog.photoshopcreative.co.uk

Software Compensation of Lens Distortion in VR Rendering


Press F9 for Full-Screen on Rift
Press F9 for Full-Screen on Rift

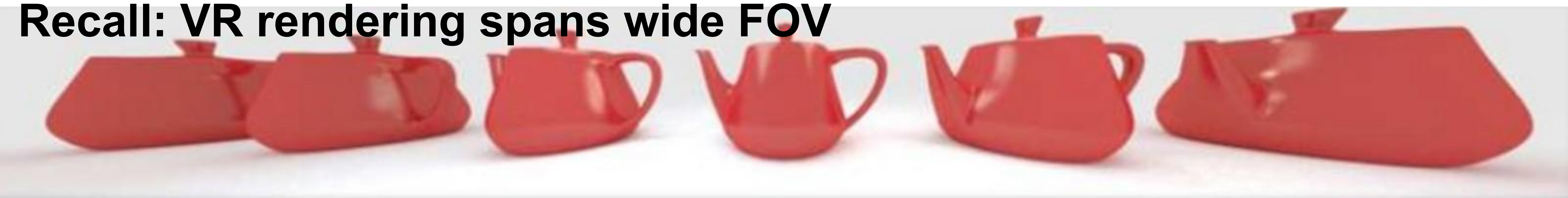
Step 1: Render scene using traditional graphics pipeline at full resolution for each eye
Step 2: Warp images in manner that scene appears correct after physical lens distortion

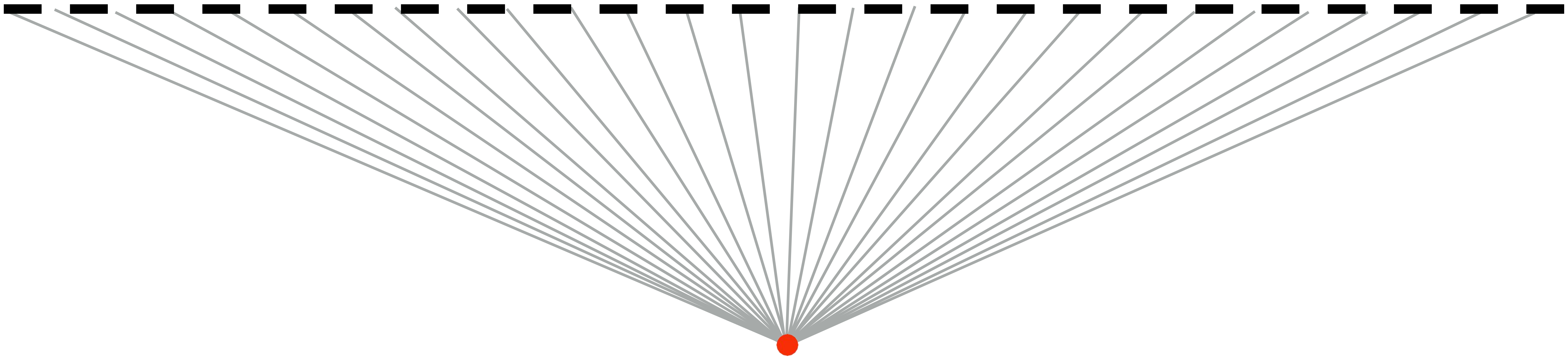
(Can use separate distortions to R, G, B to approximately correct chromatic aberration)

Related Challenge: Rendering via Planar Projection

Recall: rasterization-based graphics is based on perspective projection to plane

- **Distorts image under high FOV, as needed in VR rendering**
- **Recall: VR rendering spans wide FOV**





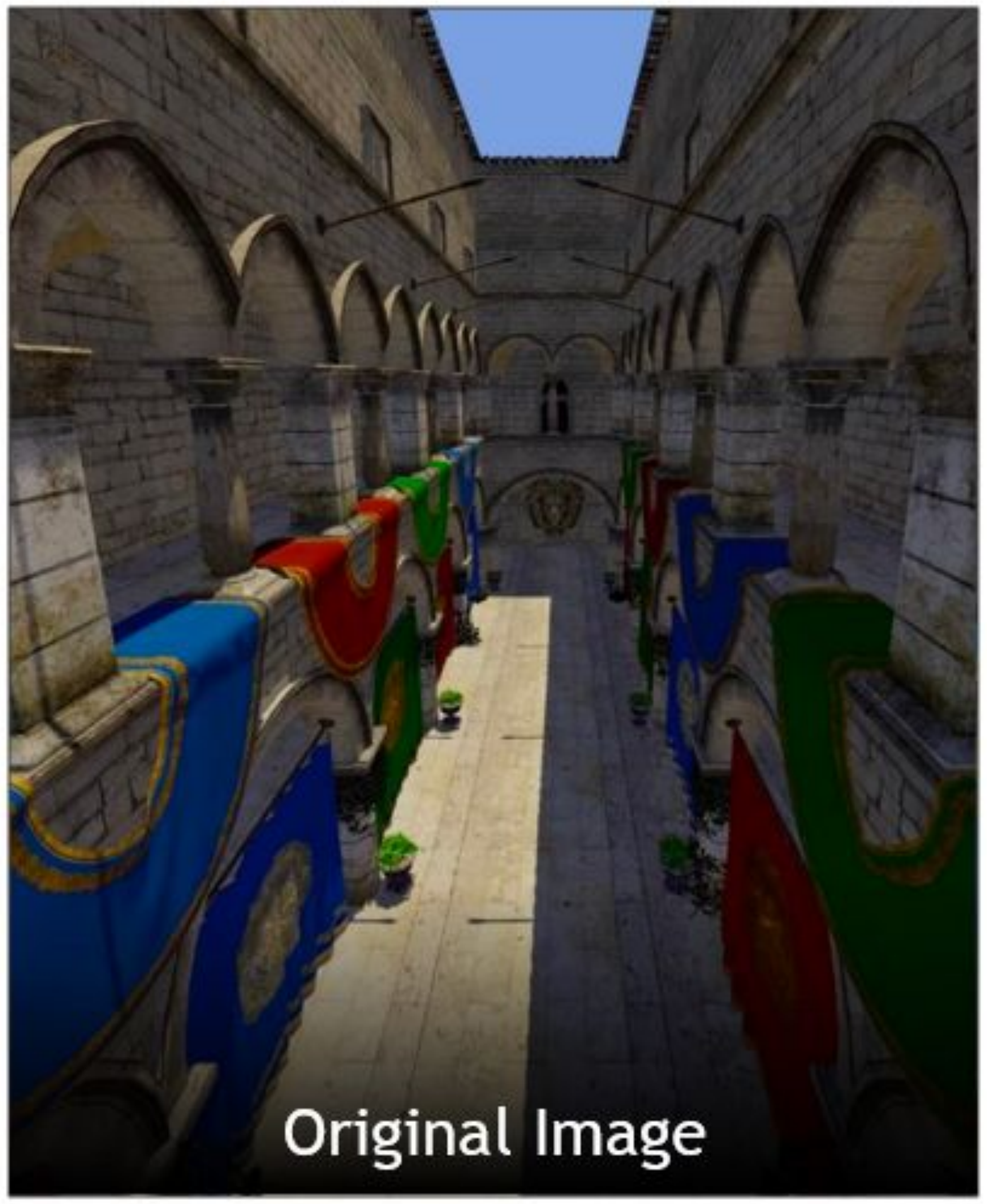
Pixels span larger angle in center of image (lowest angular resolution in center)

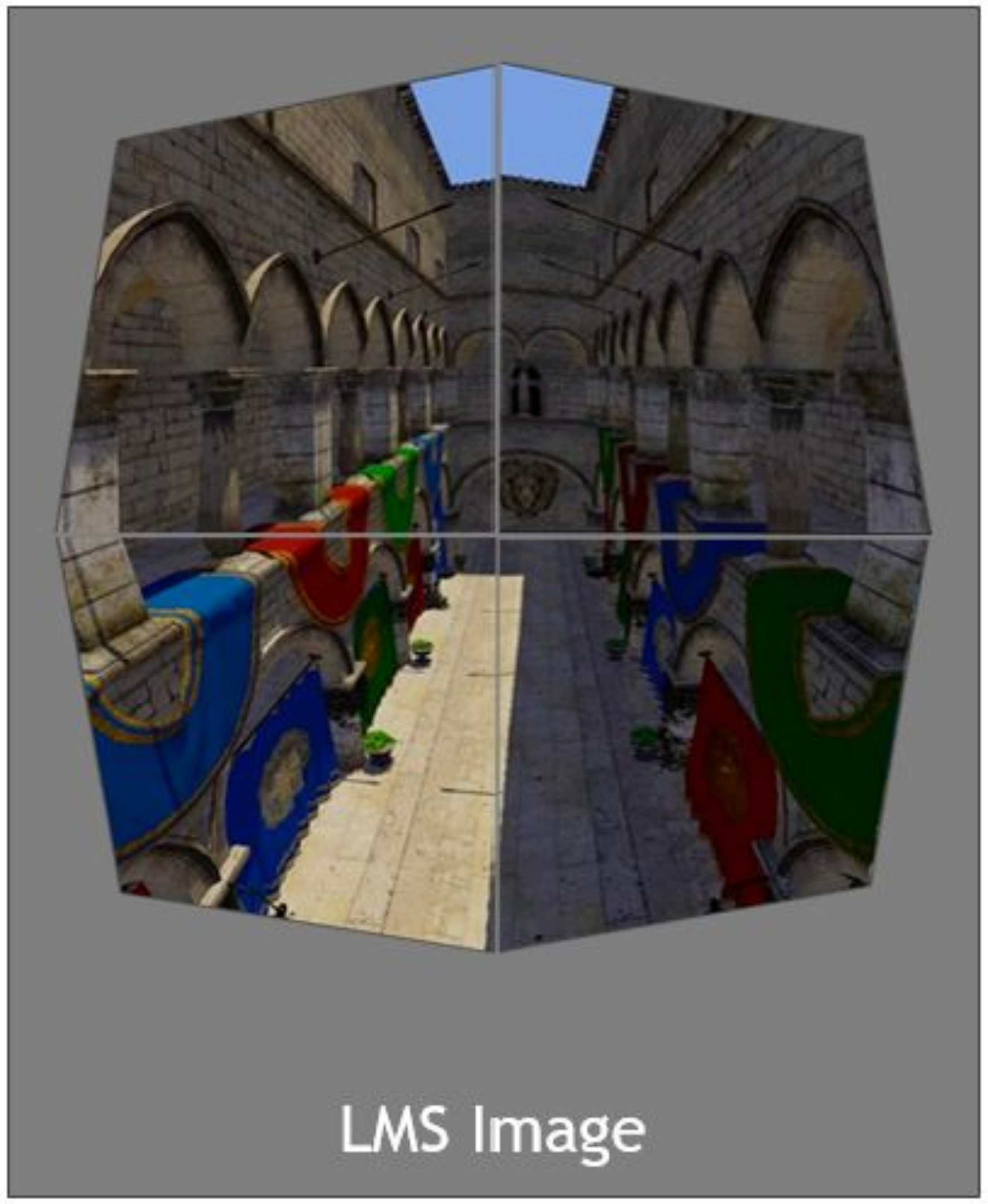
Potential solution space: curved displays, ray casting to achieve uniform angular resolution, rendering with piecewise linear projection plane (different plane per tile of screen)

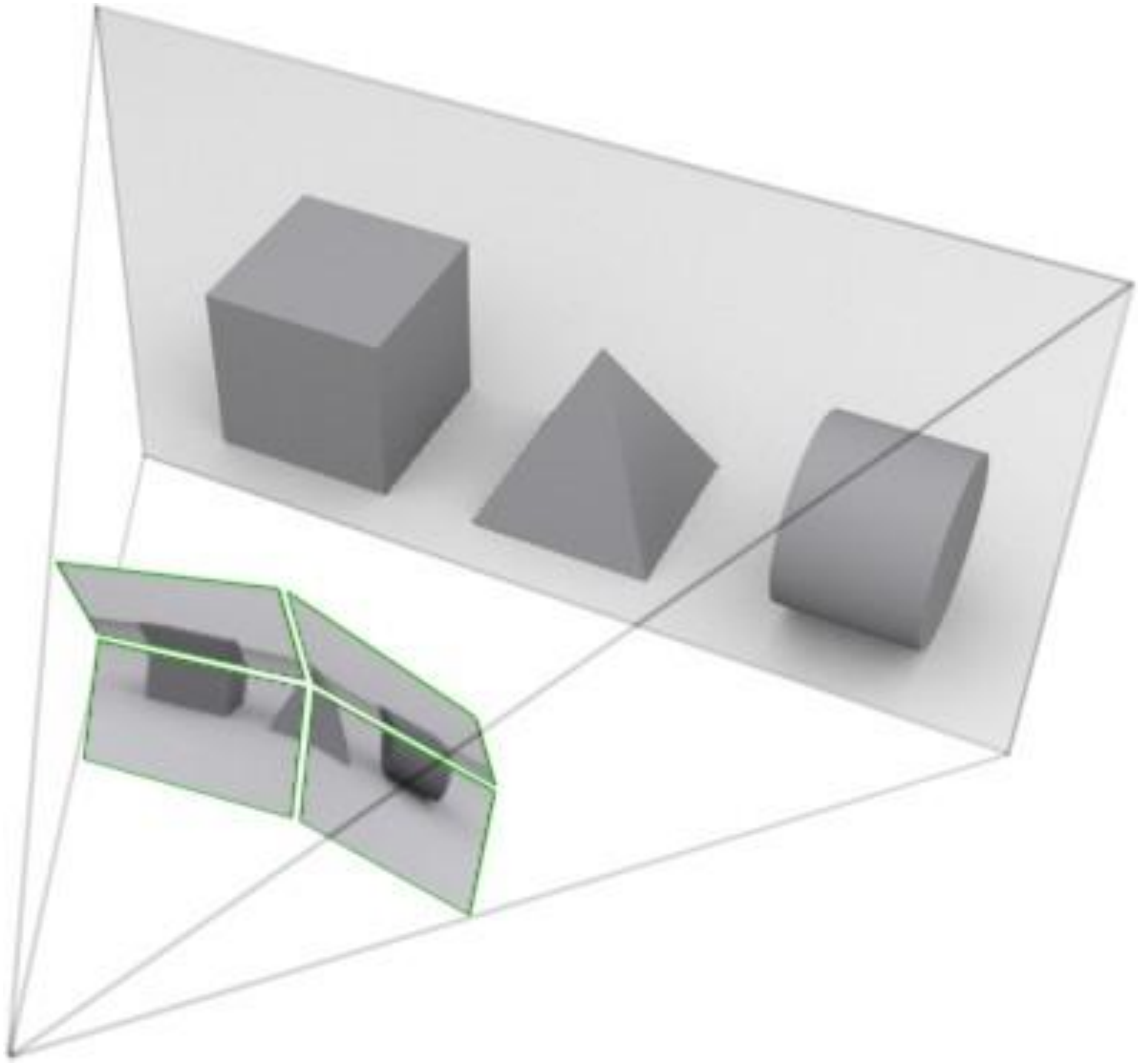
A Recent Implementation: Lens Matched Shading

Render scene with four viewports, each has different projection matrix

"Compresses" scene in the periphery (fewer samples), while not affecting scene near center of field of view





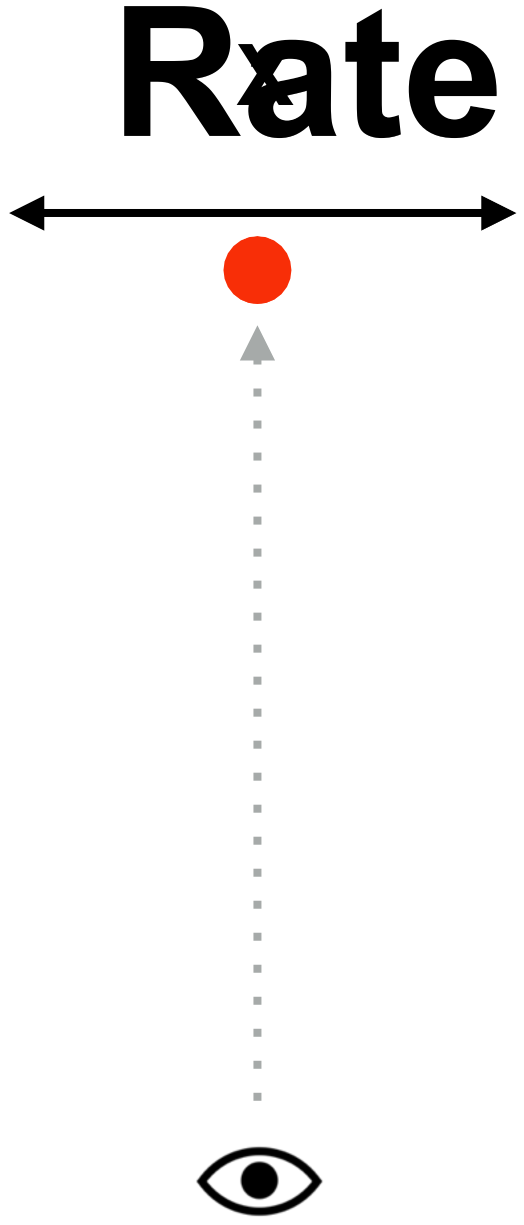


[Image credit: NVIDIA]

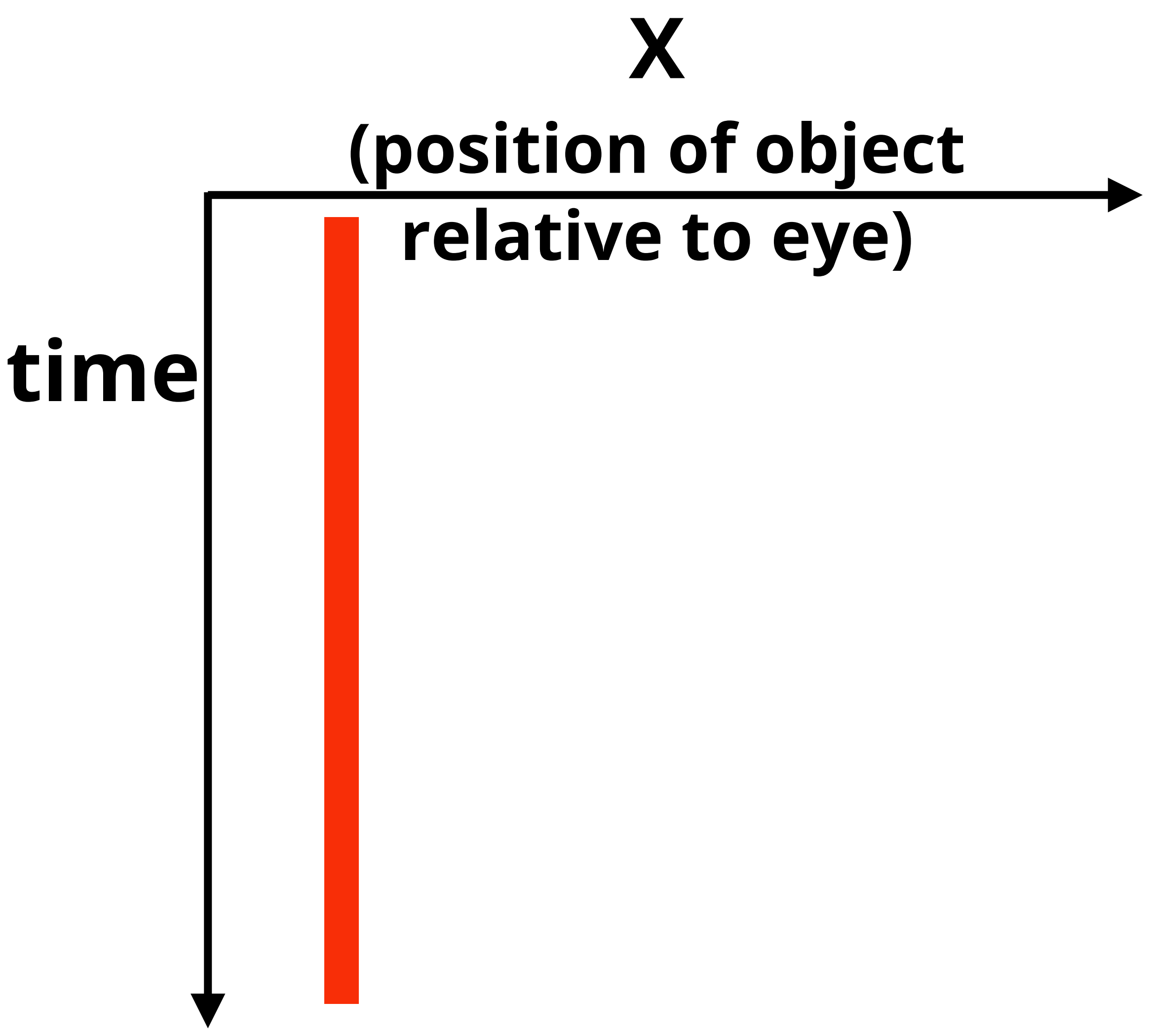
Slide credit: Kayvon Fatahalian

Rendering Challenge: Eye Motion And Finite Rendering Rate

Consider Finite VR Display Refresh Rate

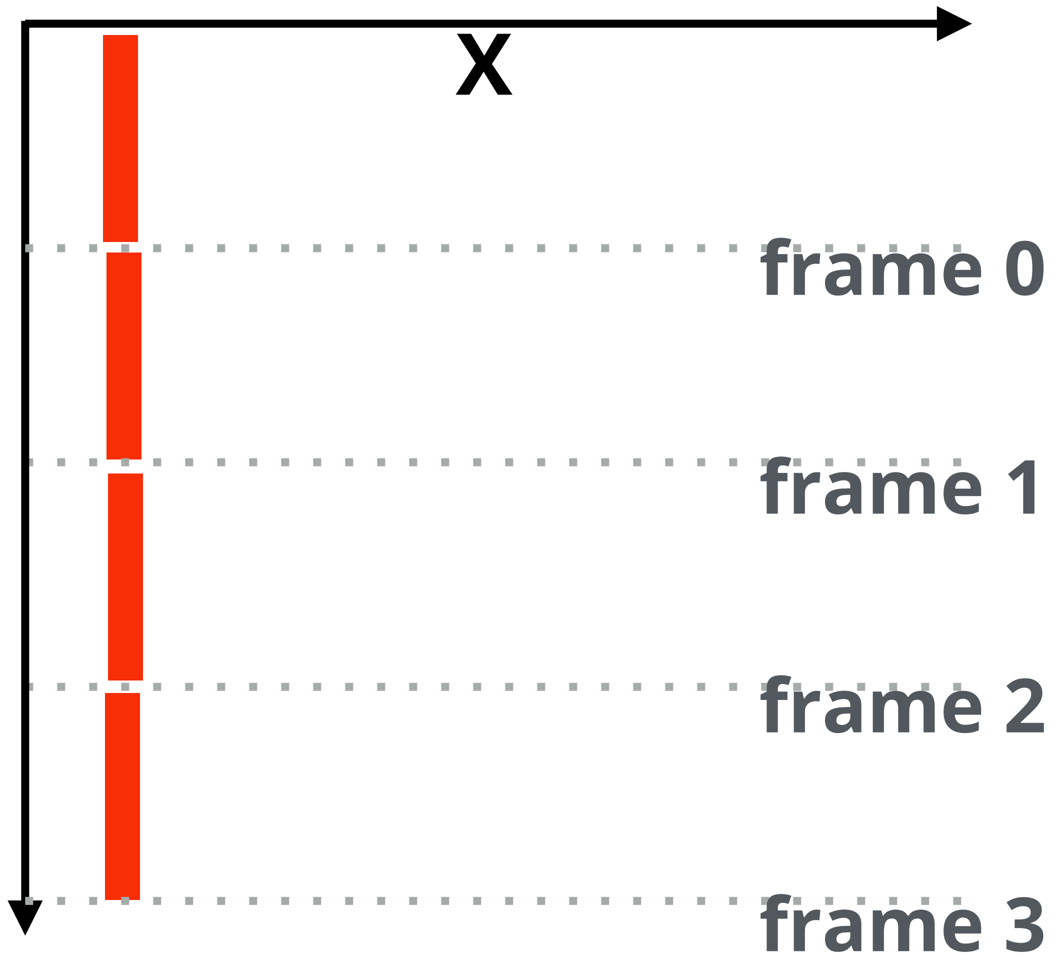


Reality (continuous)



- Red object fixed;
- Eye gaze fixed

VR (discrete display refresh)

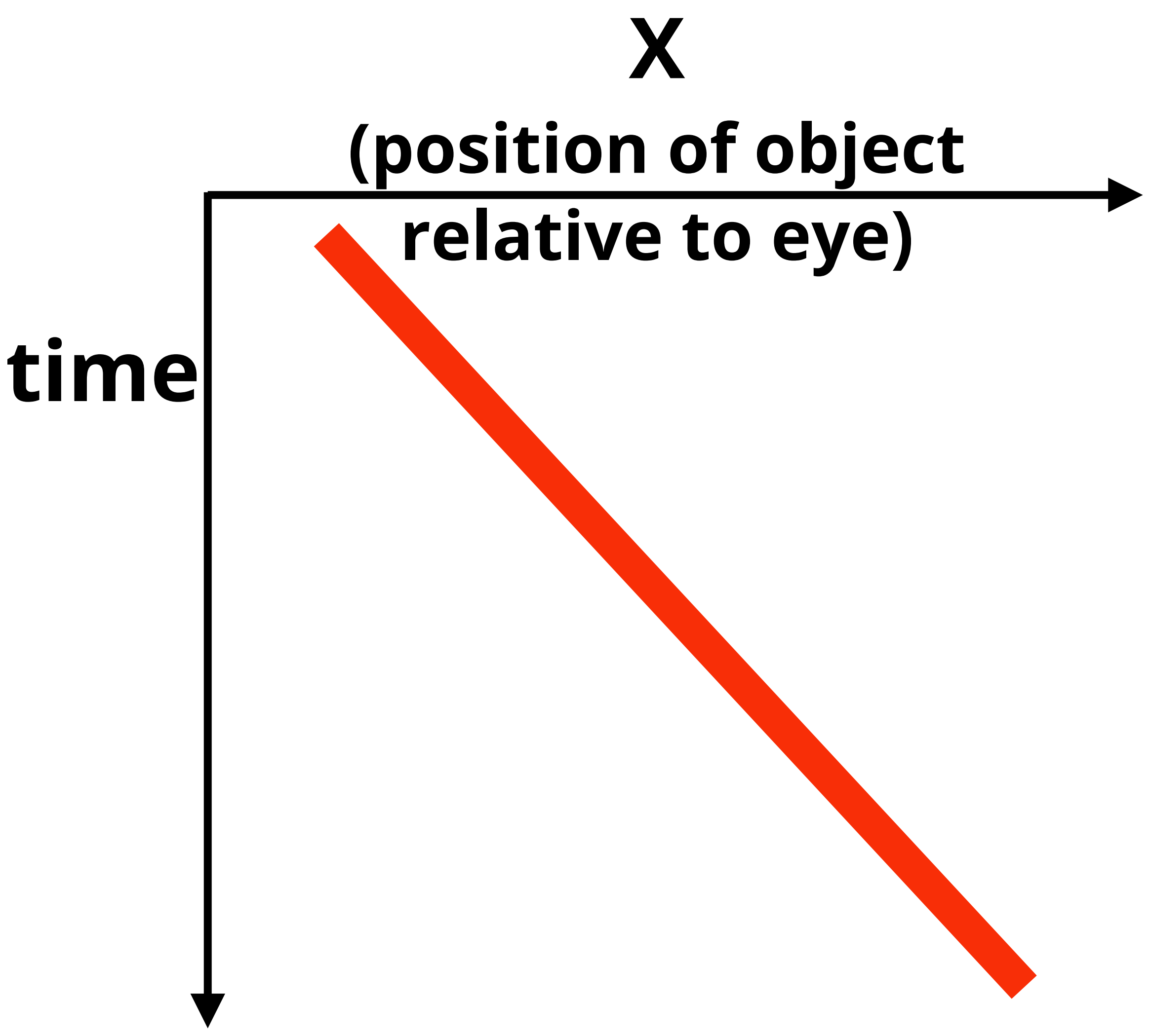


- Light from display (light updates every frame)

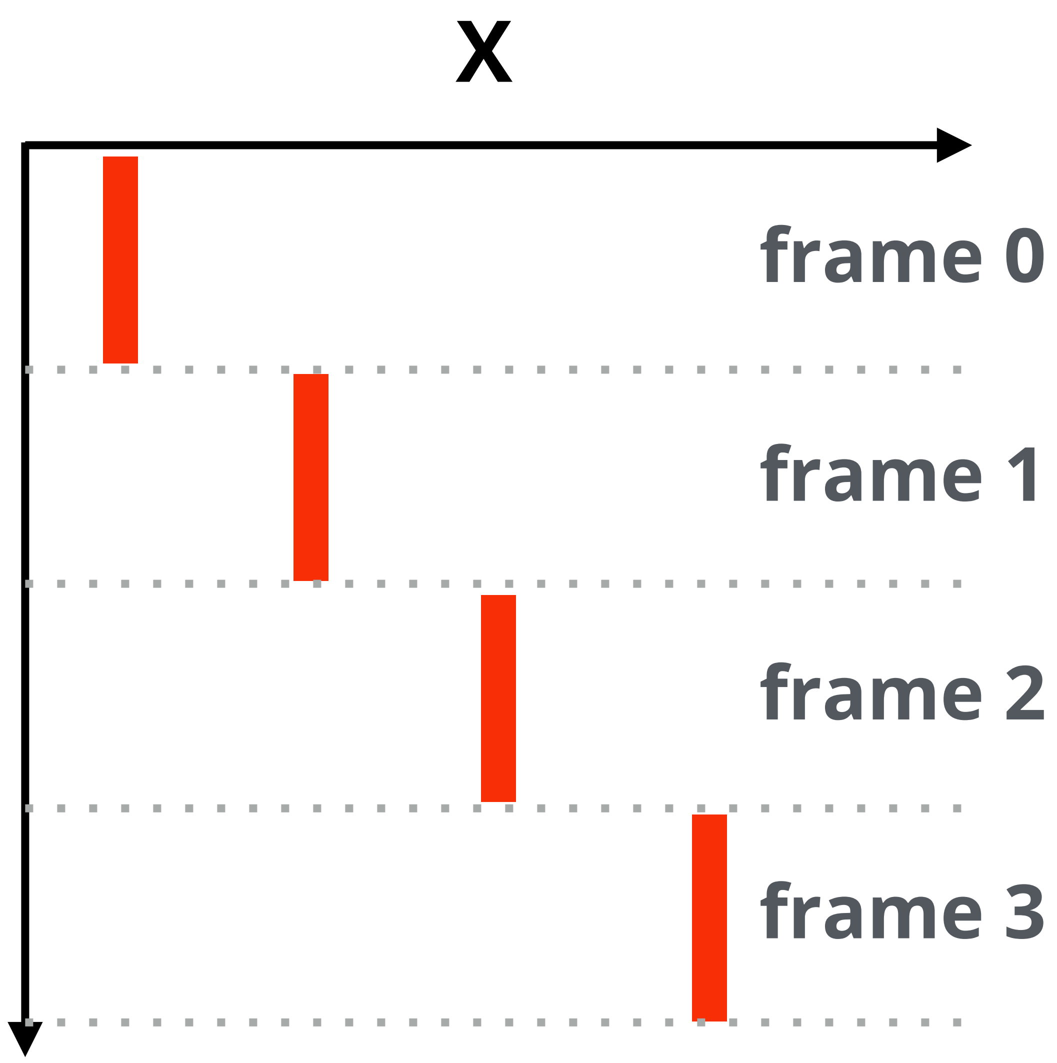
Spacetime diagrams adopted from presentations by Michael Abrash

Case 2: Object Moving Relative to Eye

Reality (continuous)



VR (discrete display refresh)

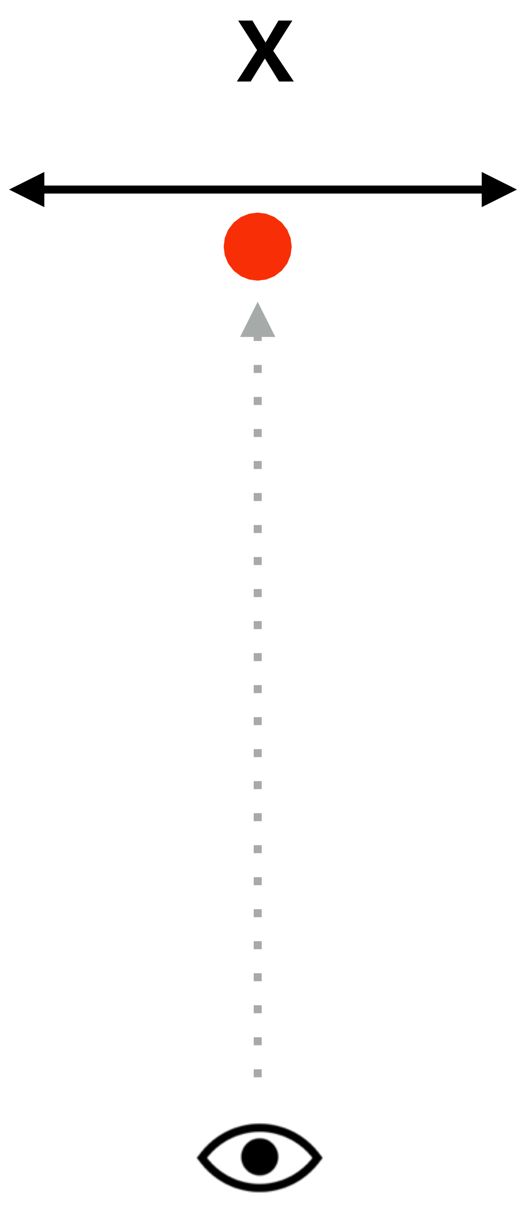


- Red object moving left to right;
- Eye gaze fixed

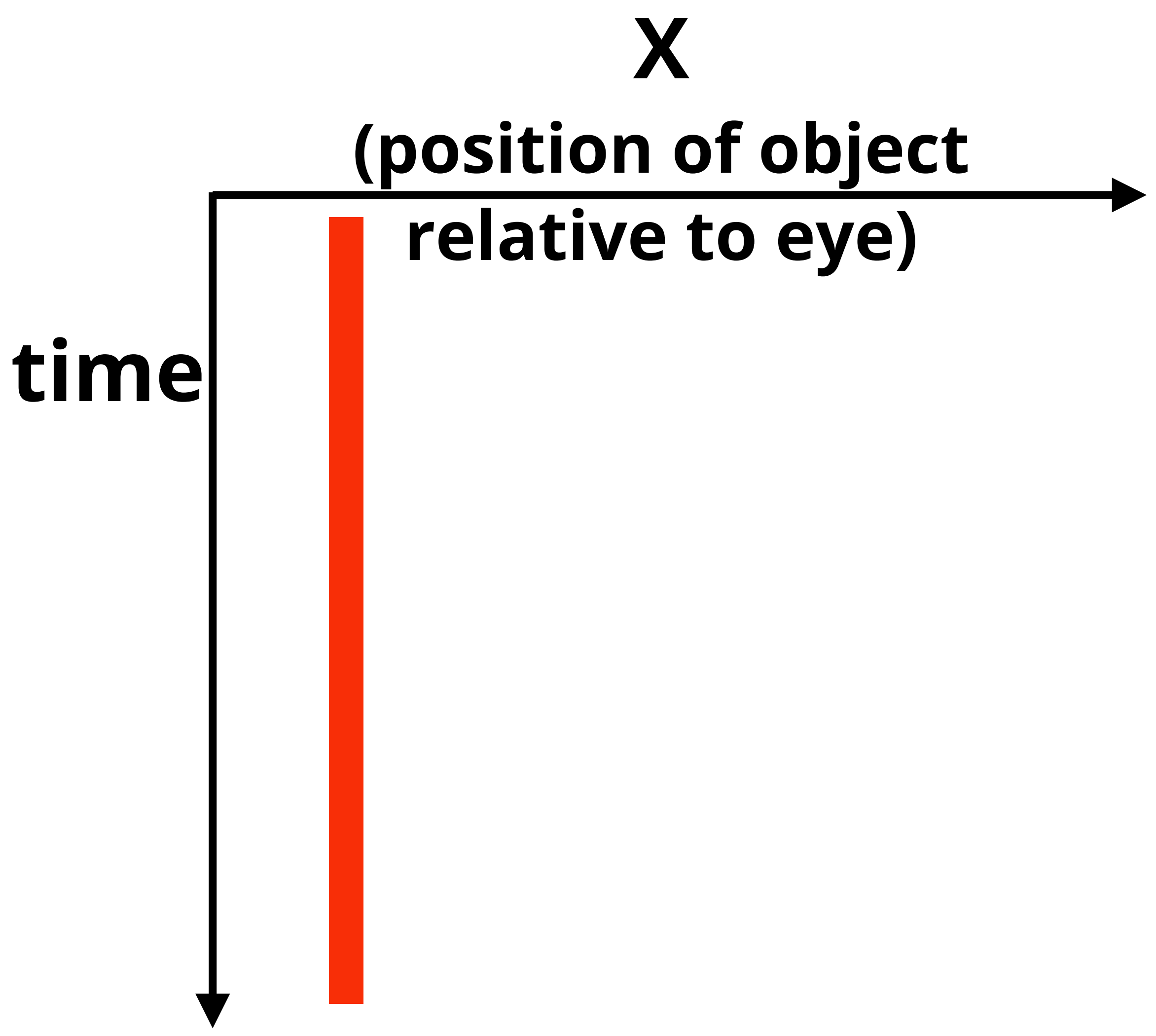
- Effect: time discretization
- OK: same perceptual effect as on regular 2D displays

Spacetime diagrams adopted from presentations by Michael Abrash

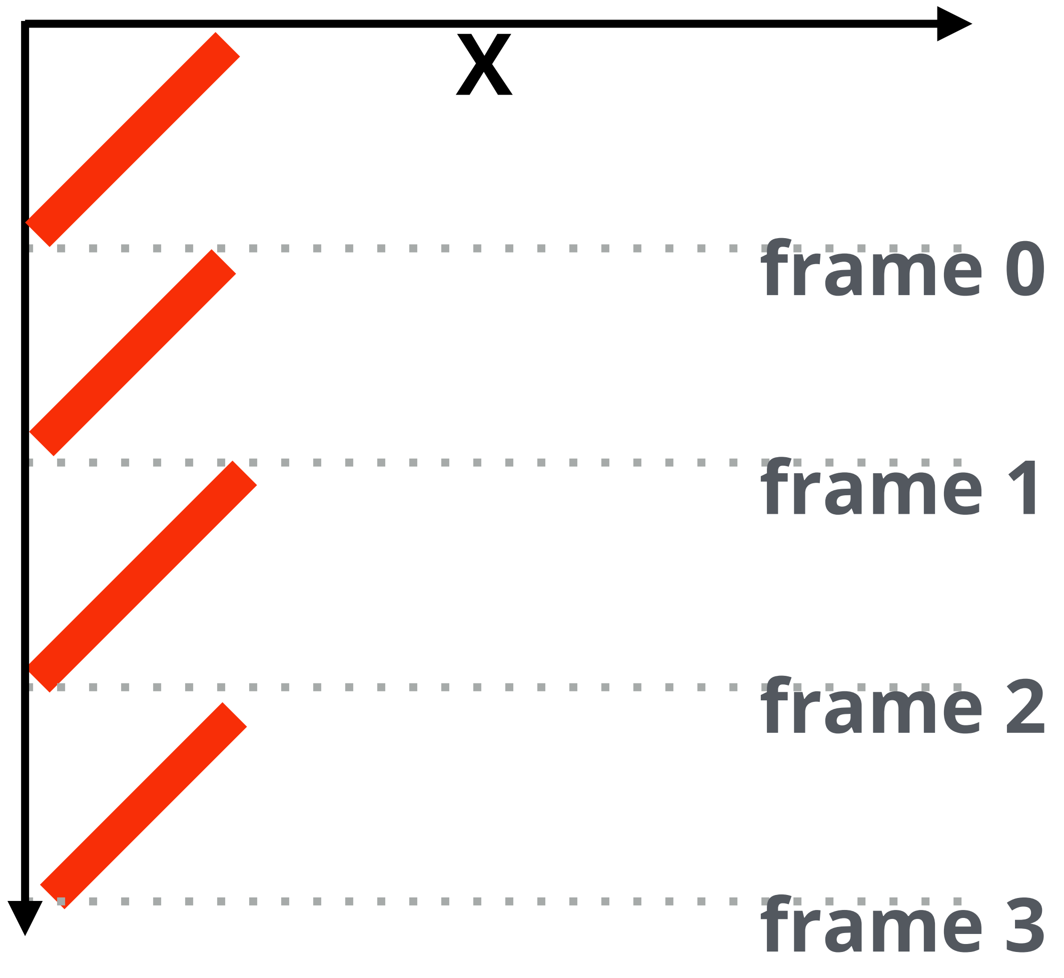
Case 3: Eye Moving to Track Moving Object



Reality (continuous)



VR (discrete display refresh)

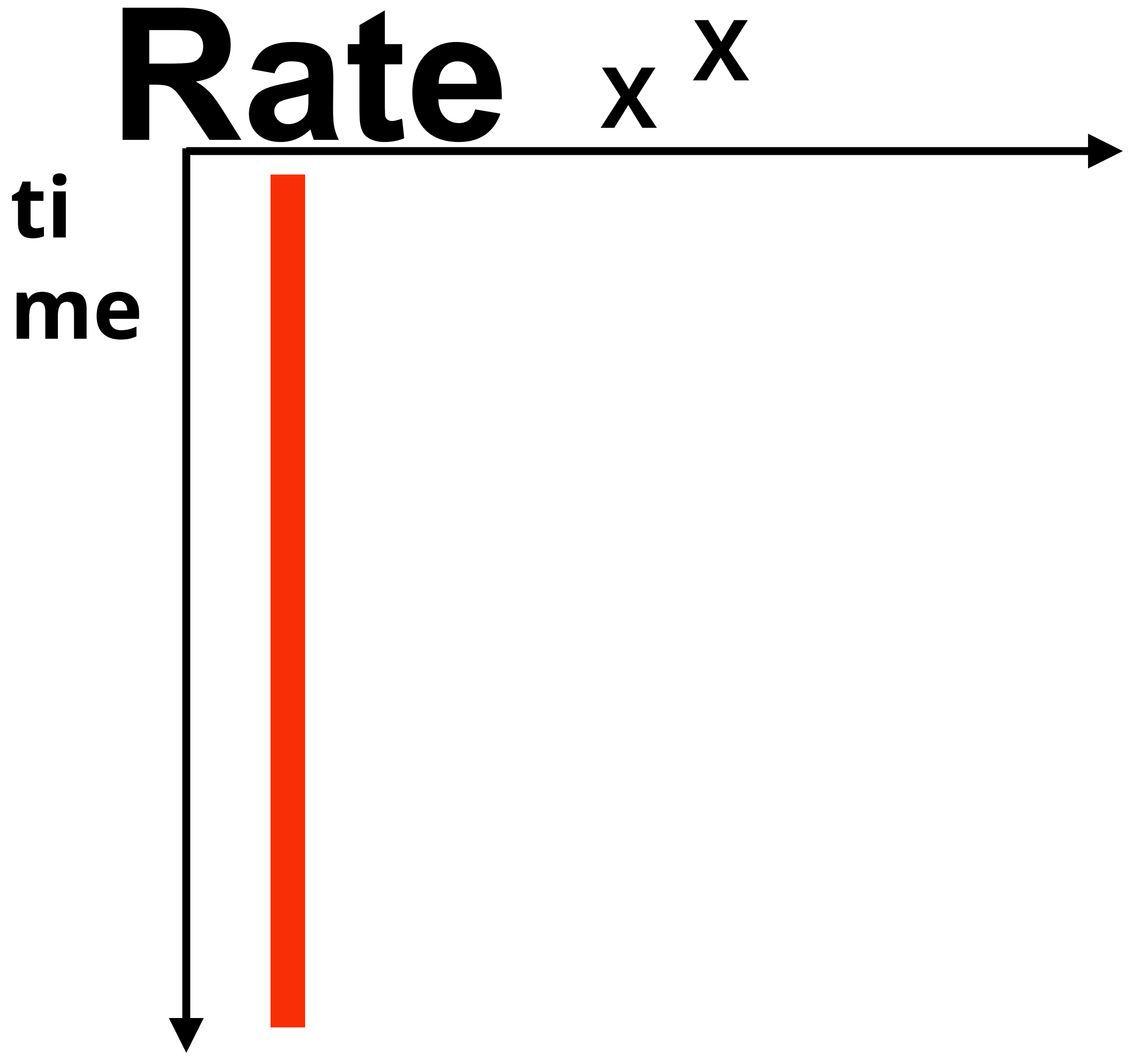


- Red object moving left to right;
- Eye gaze moving left to right to track object

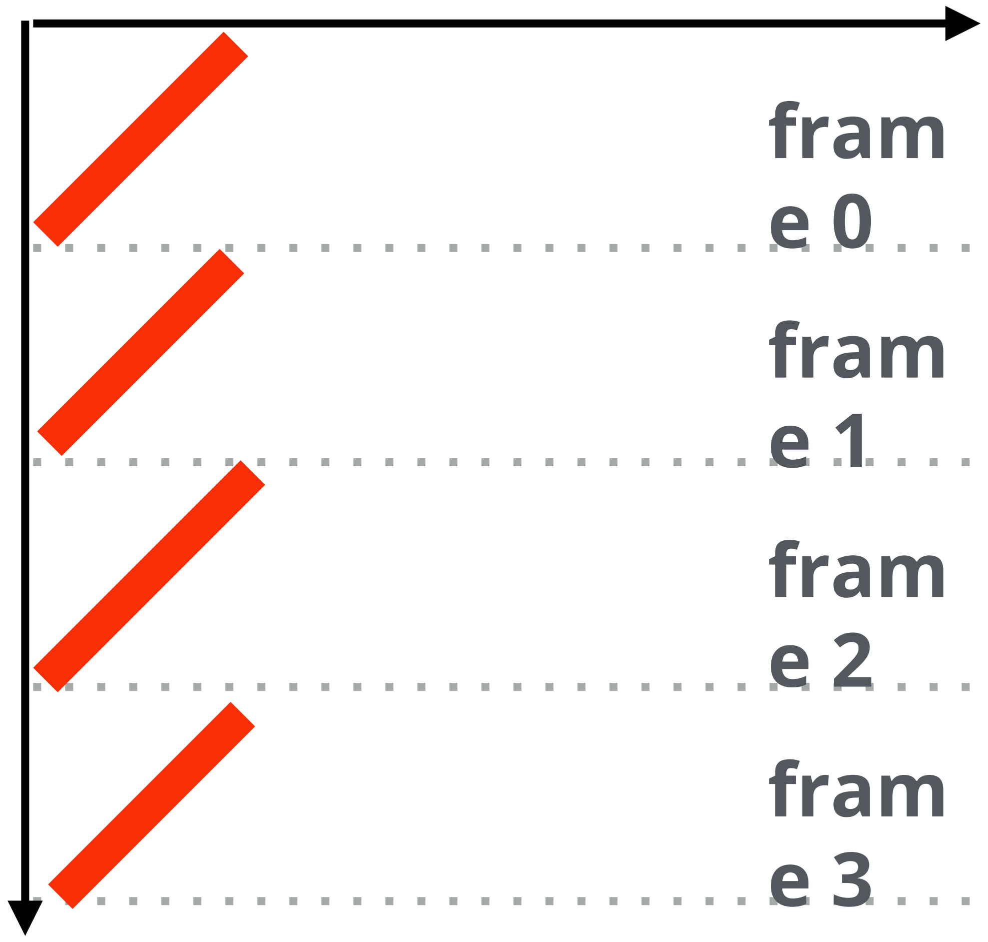
- Eye is moving continuously <u>relative to display</u>
- During each frame, image of object lags eye motion

Spacetime diagrams adopted from presentations by Michael Abrash

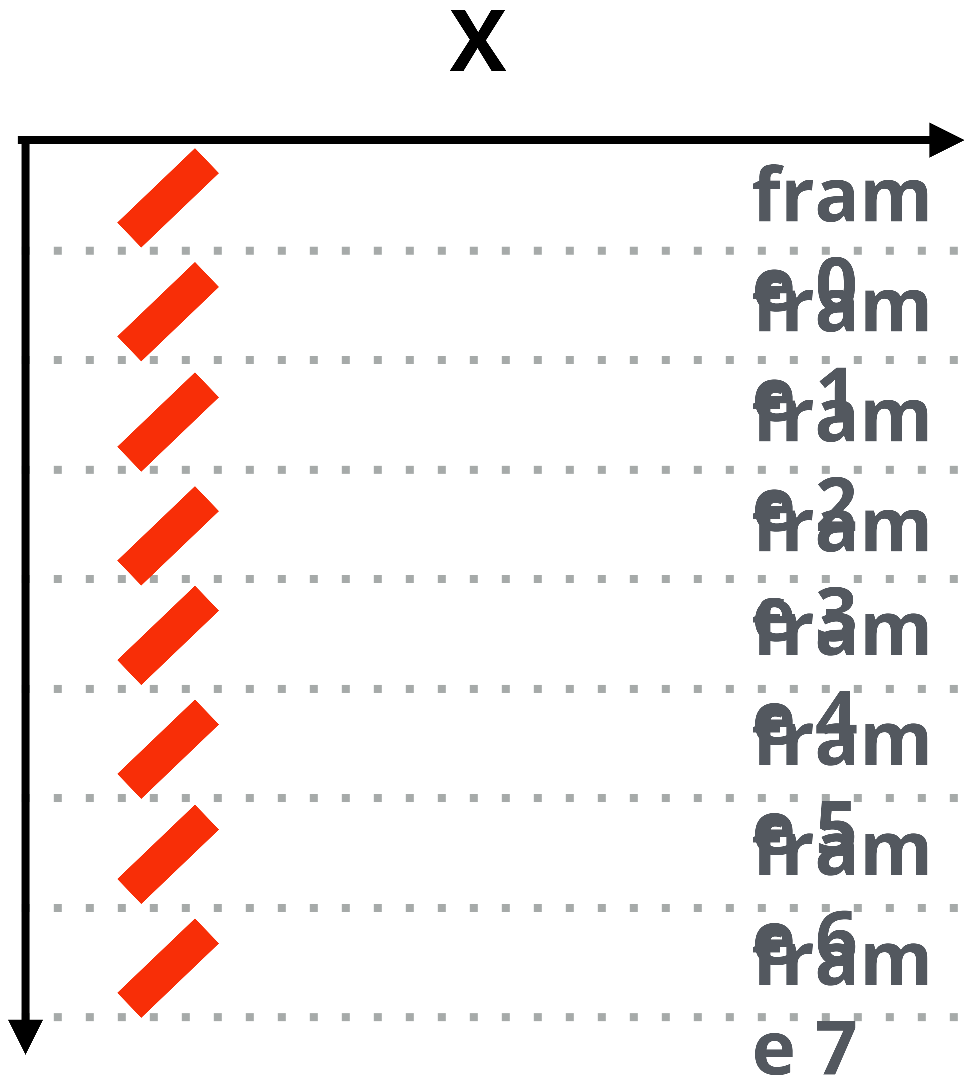
Reducing Judder: Increase Frame Rate



Continuous ground truth



Origina l judder



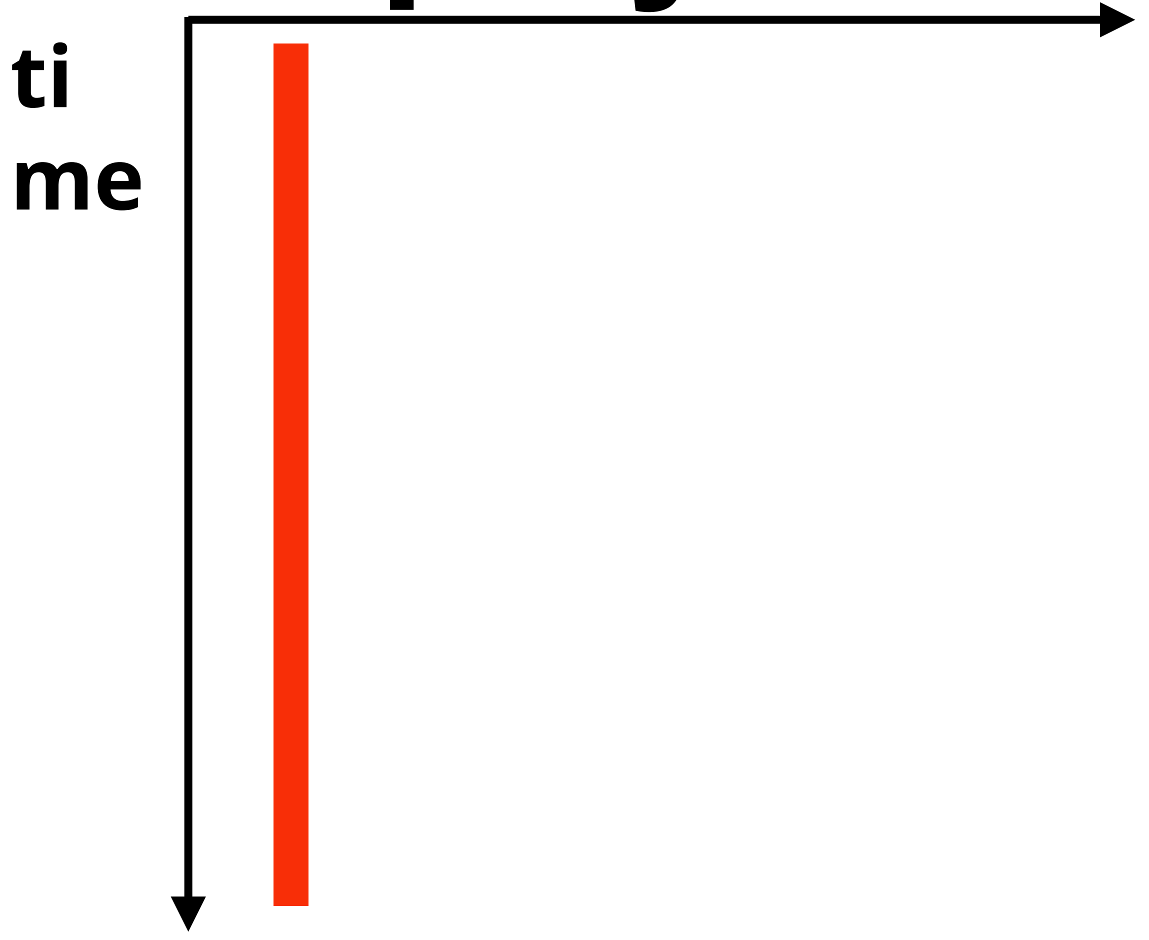
High frame-rate judder

Higher frame rate (right-most diagram)

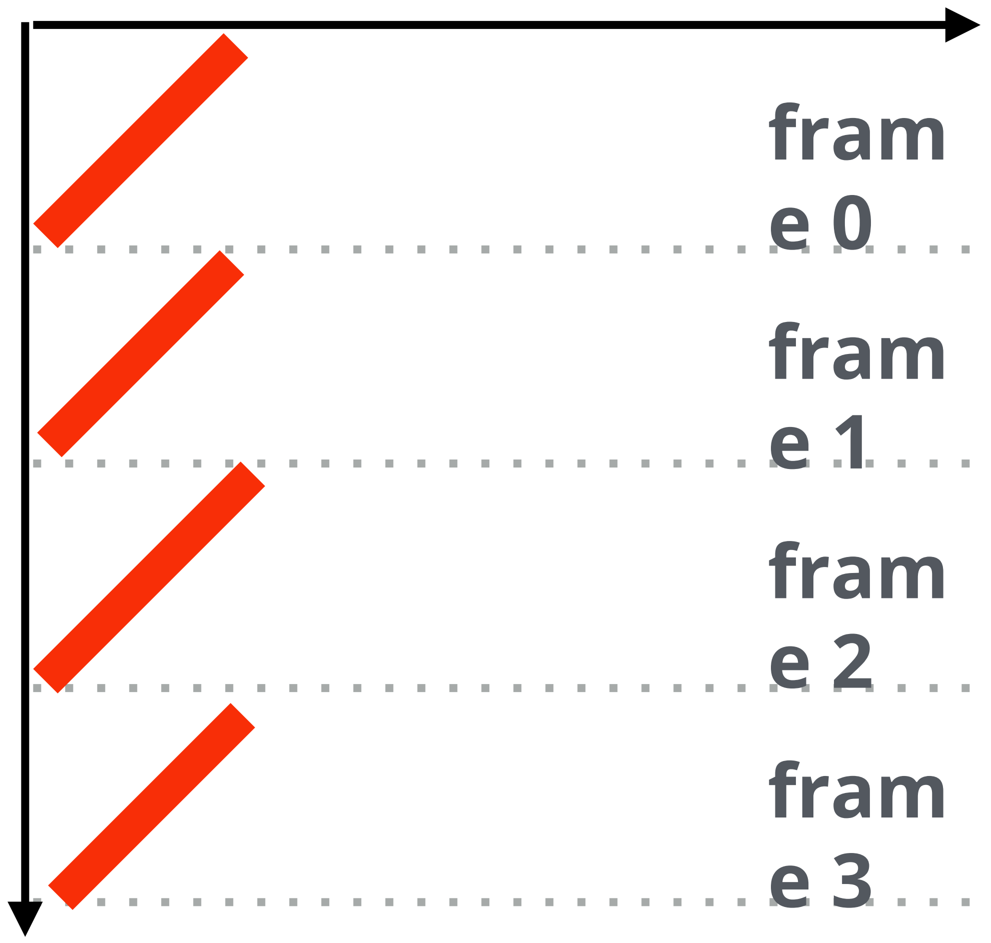
- Closer approximation of ground truth

Spacetime diagrams adopted from presentations by Michael Abrash

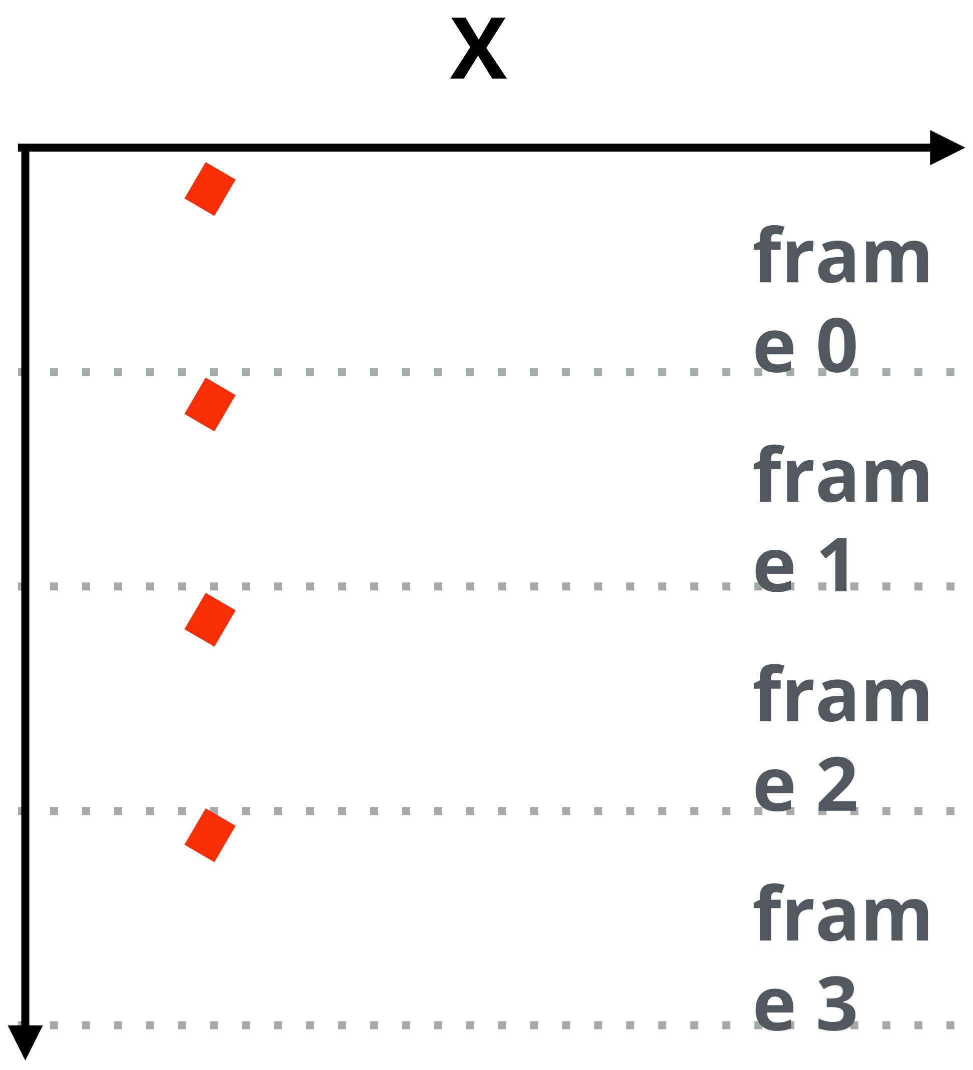
Reducing Judder: Low Persistence Display



Continuous ground truth



Original judder



Low-persistence display

Low-persistence display: pixels emit light for small fraction of frame

- Oculus DK2 OLED low-persistence display:
  - 75 Hz frame rate = ~13 ms per frame
  - Pixel persistence = 2-3 ms

Spacetime diagrams adopted from presentations by Michael Abrash

Rendering Challenge:
High-Quality vs
Low-Latency

Problem: High-Quality Rendering Can Be Slow

Constraints:

- Battery-powered device
- High-resolution outputs for both

  eyes Implication:

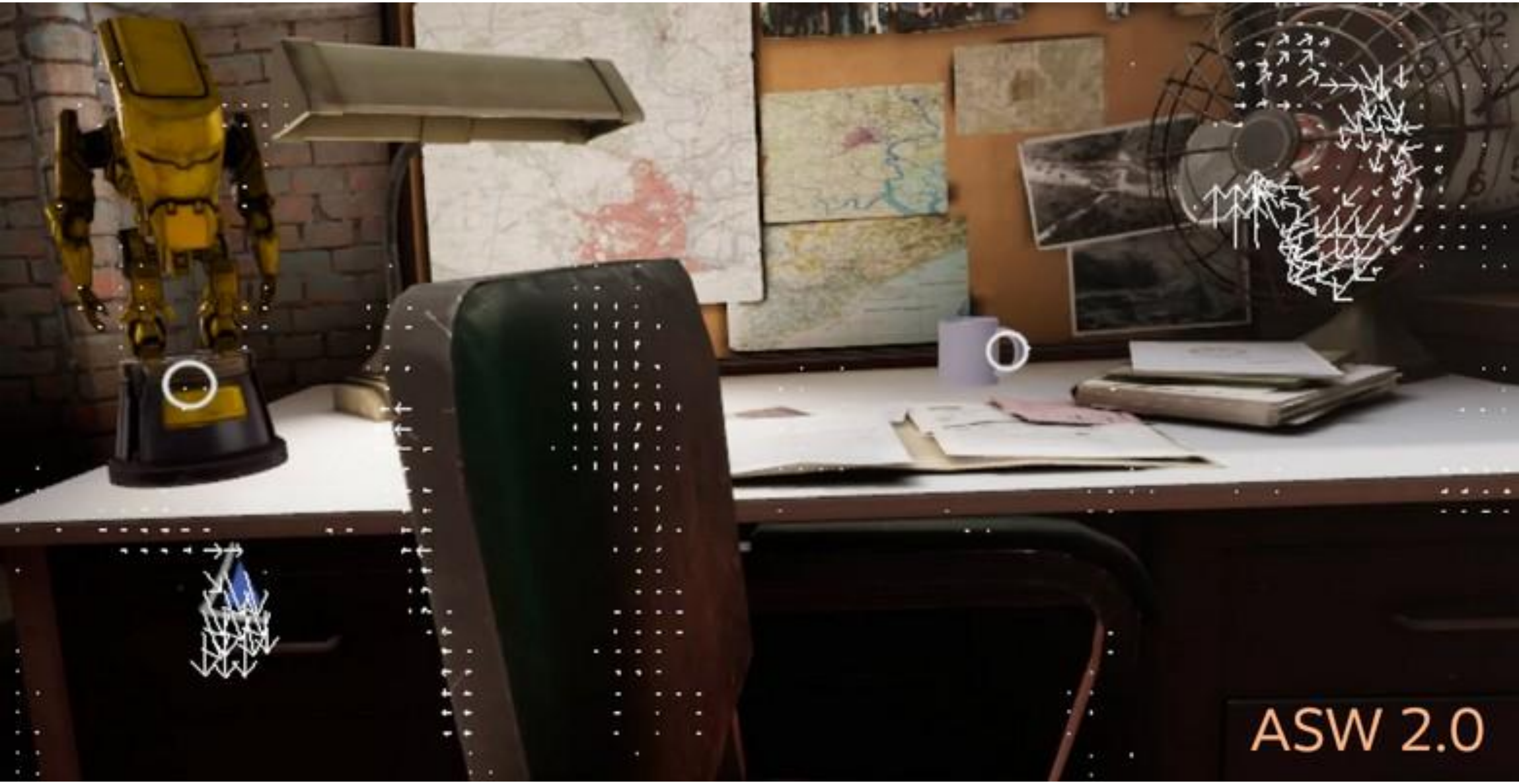
- Can take significant time to render a frame
- This increases latency, can cause motion sickness
- This can reduce frame refresh rate

Modern VR Engineering Solution: Reprojection

Key Ideas:

- Decouple slow, high-quality rendering of frames from fast “reprojection” immediately before display
- The high-quality frame uses then-current head- tracking (which may be stale by end of render)
- Reprojection occurs extremely close in time to physical display, and warps the most recent high-quality frame to the very latest head-tracking data
- Accurate reprojection warp requires both rendered image, its depth map, and potentially motion derivatives (e.g. optical flow)

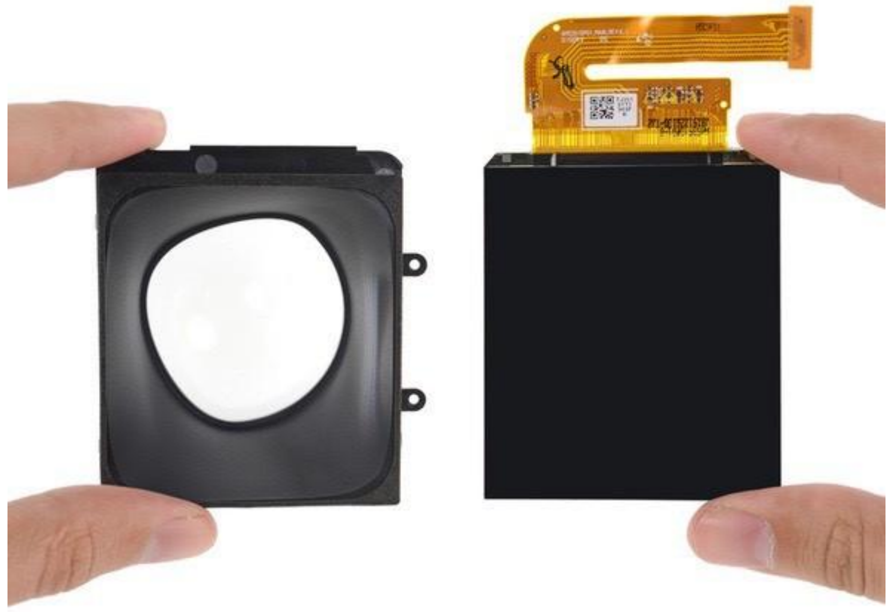
Modern VR Engineering Solution: Reprojection



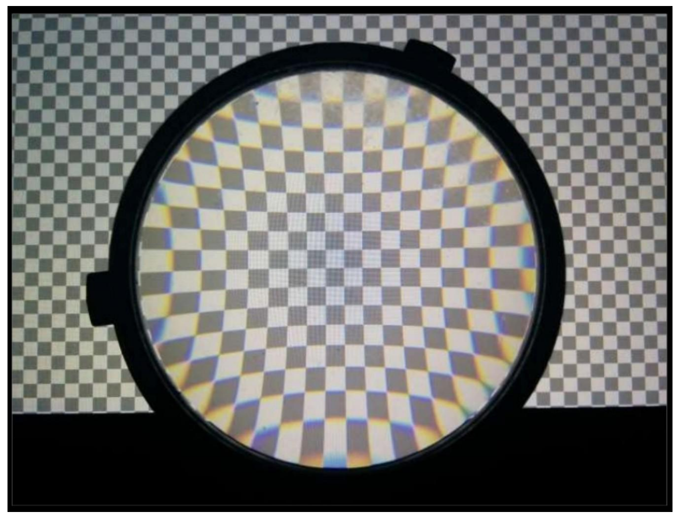
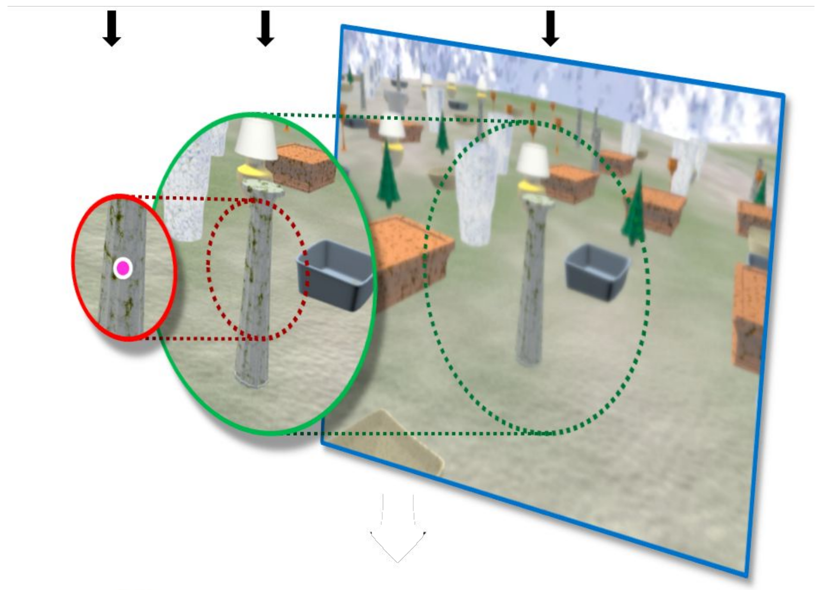
Overview of VR Topics

Areas we will discuss over next few lectures

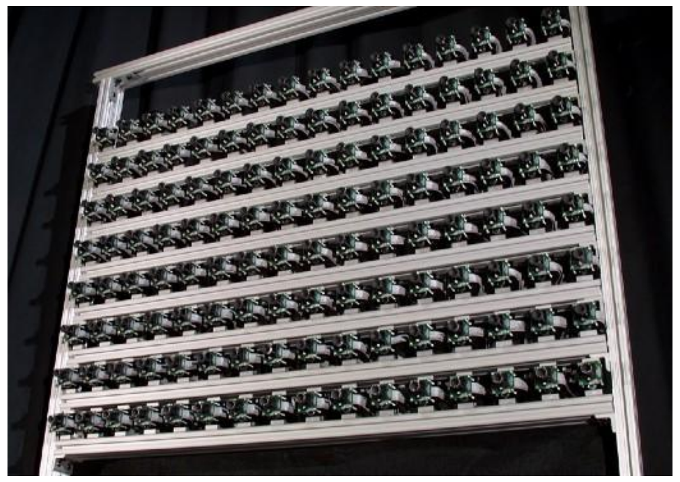
- VR Displays



- VR Rendering

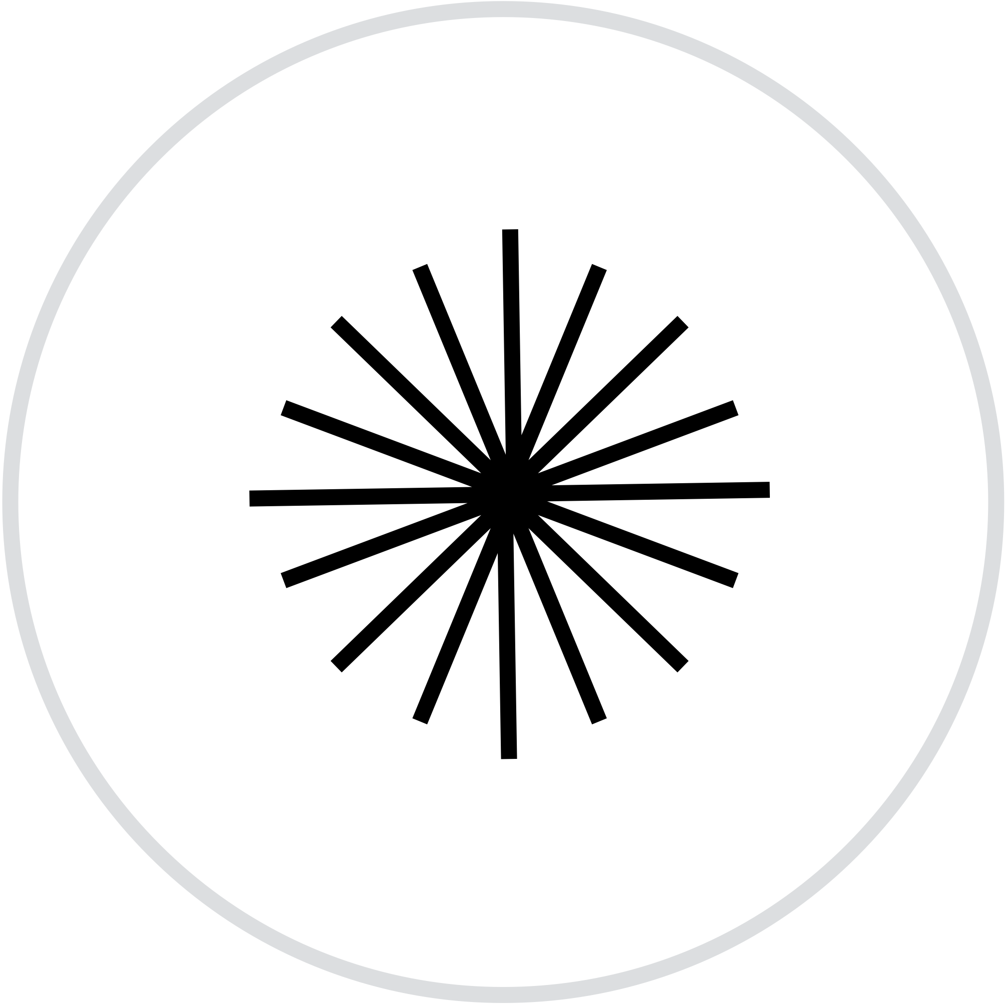
- VR Imaging

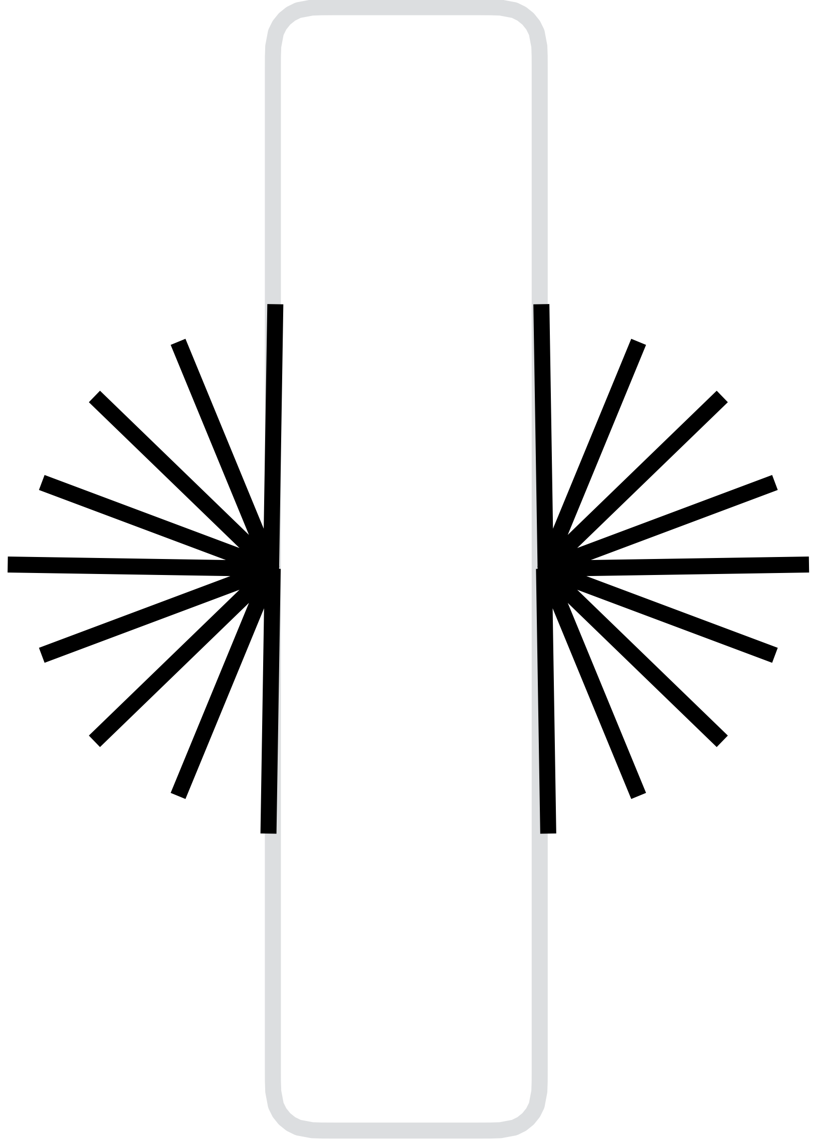


Spherical Imaging (Monocular 360)

Dual Fisheye

Stitching Challenges





Want this

ray

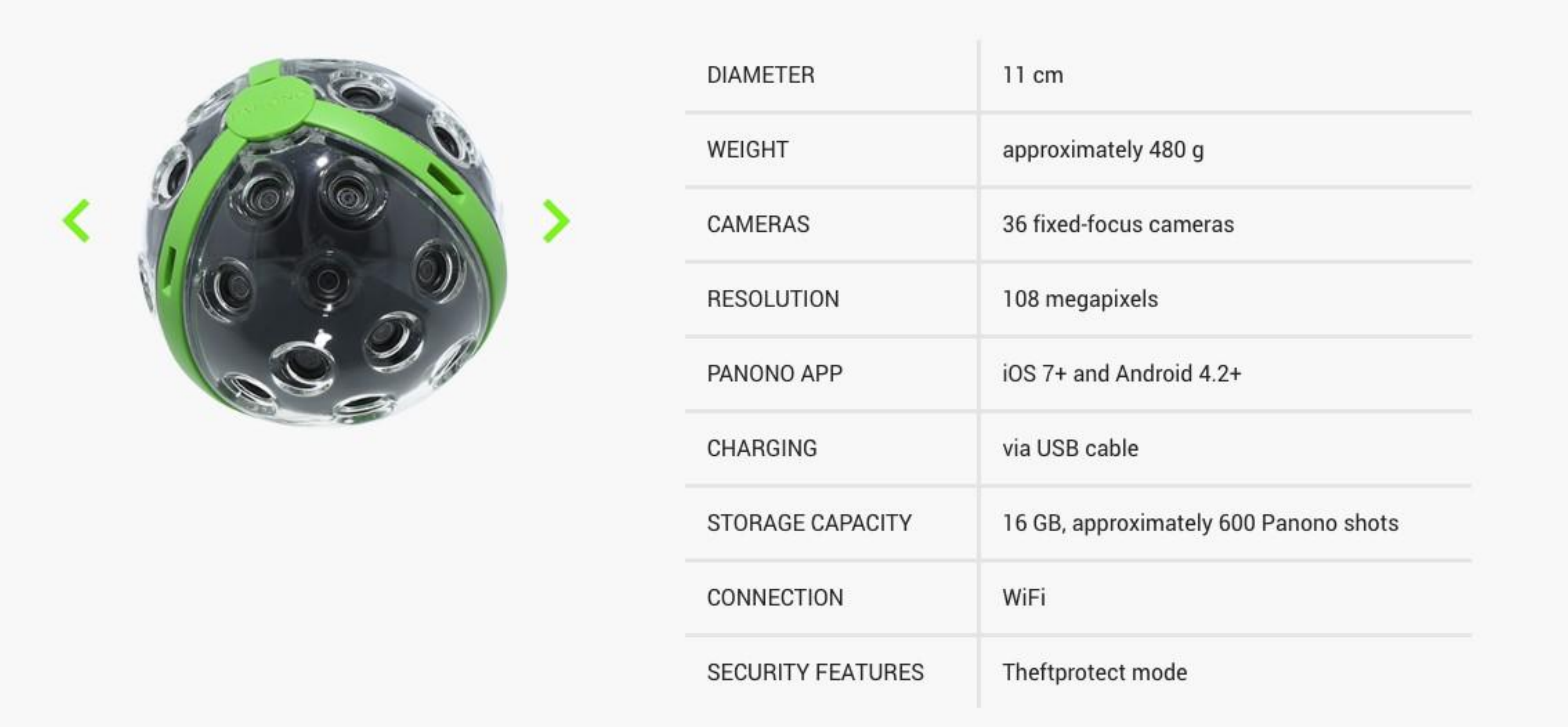
sampling

Get this

ray

sampling

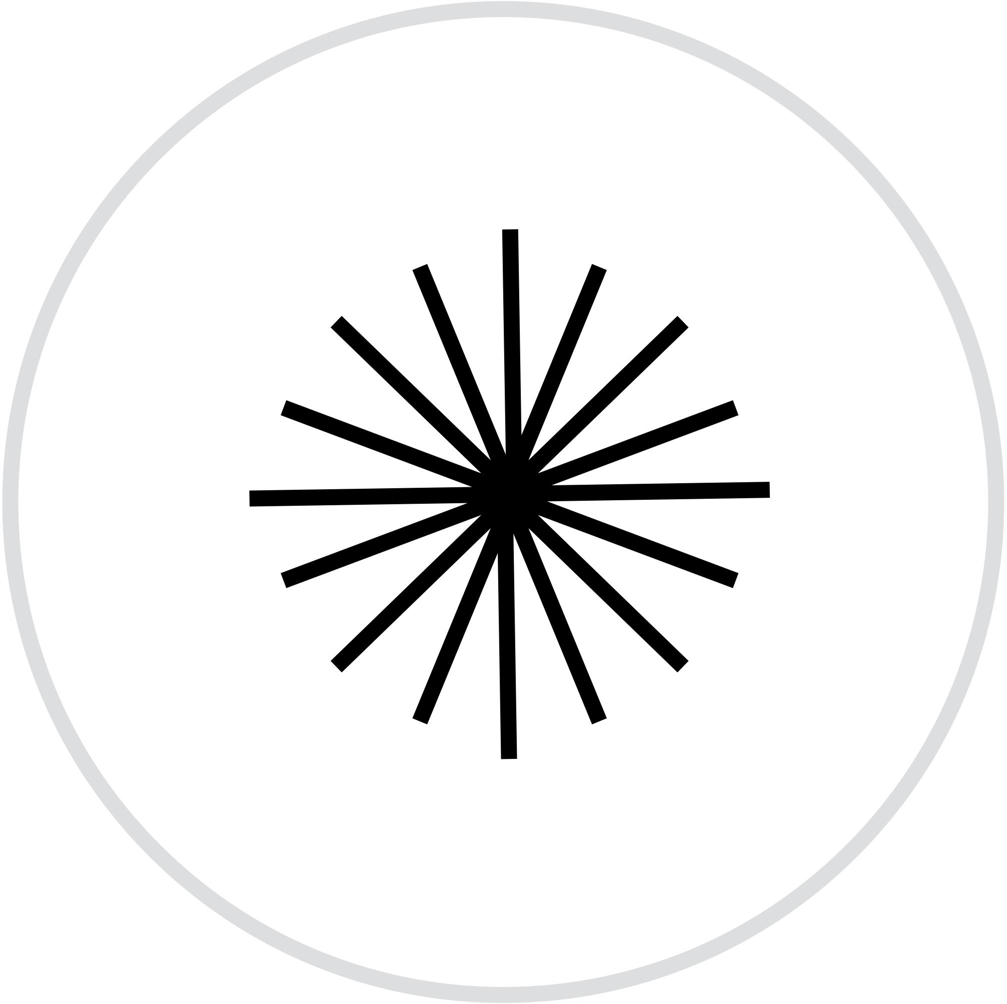
Spherical Array of Cameras



DIAMETER	11 cm
WEIGHT	approximately 480 g
CAMERAS	36 fixed-focus cameras
RESOLUTION	108 megapixels
PANONO APP	iOS 7+ and Android 4.2+
CHARGING	via USB cable
STORAGE CAPACITY	16 GB, approximately 600 Panono shots
CONNECTION	WiFi
SECURITY FEATURES	Theftprotect mode

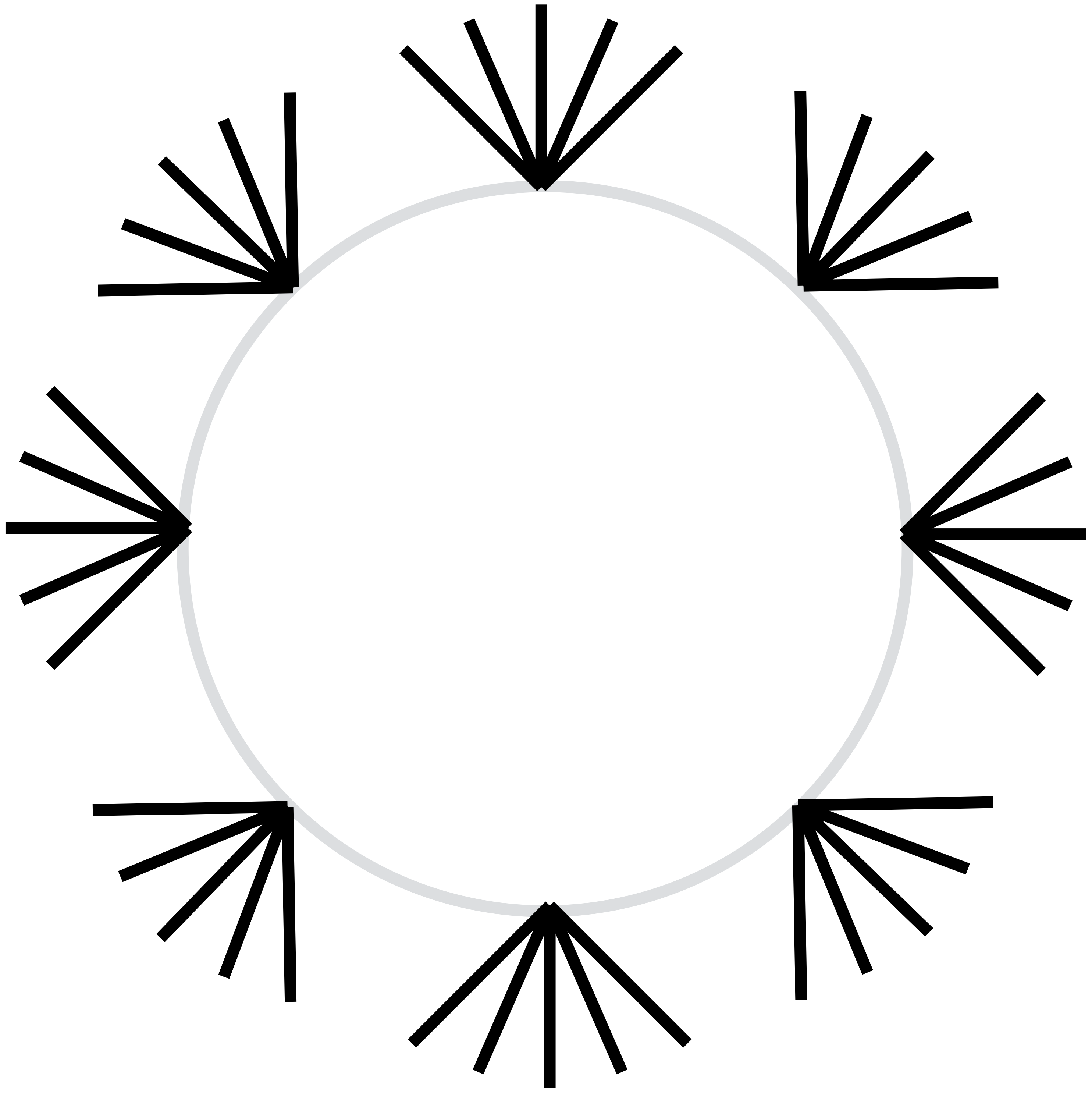
Panono 360 degree Camera

Stitching Challenges



Want this

ray
sampling



Get this

ray
sampling

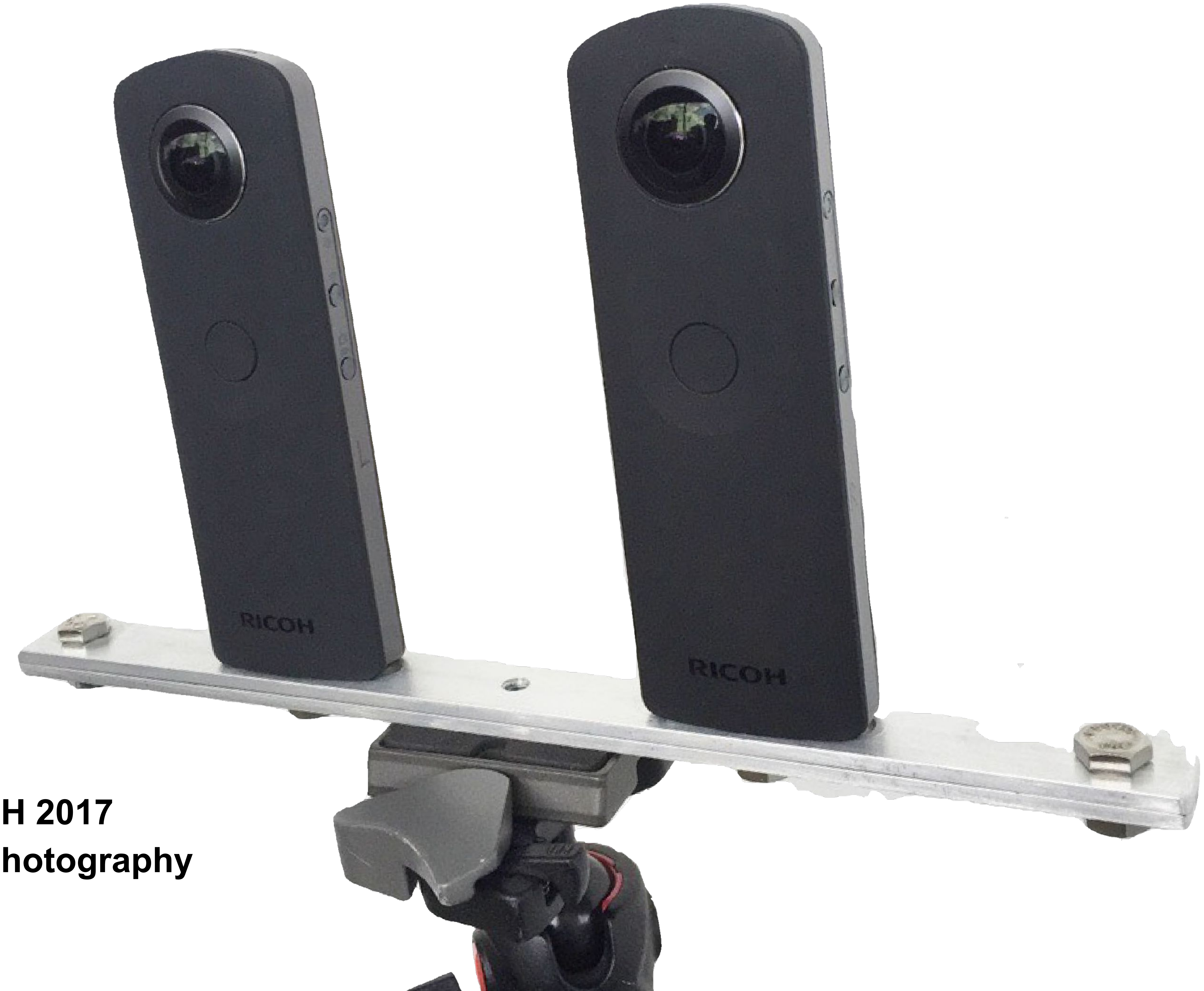
High Quality Stitching Solution Uses Computer Vision

Use computer vision techniques:

- Detect image features (like SIFT features)
- Correlate features across frames (transform)
- Warp to align frames and blend

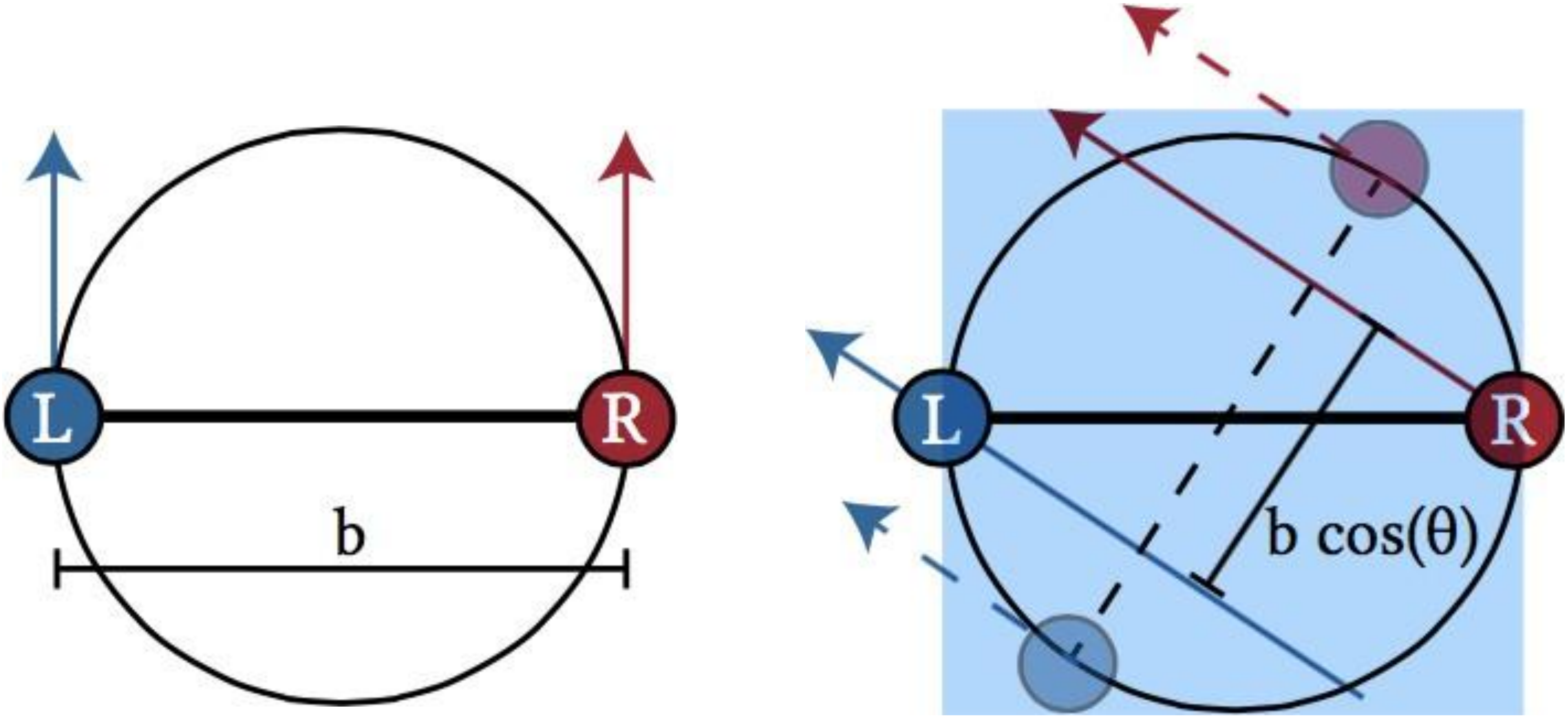
Spherical Stereo Imaging

Two Eyes — Two Spherical Cameras?



Matzen et al. SIGGRAPH 2017
Low-Cost 360 Stereo Photography
and Video Capture

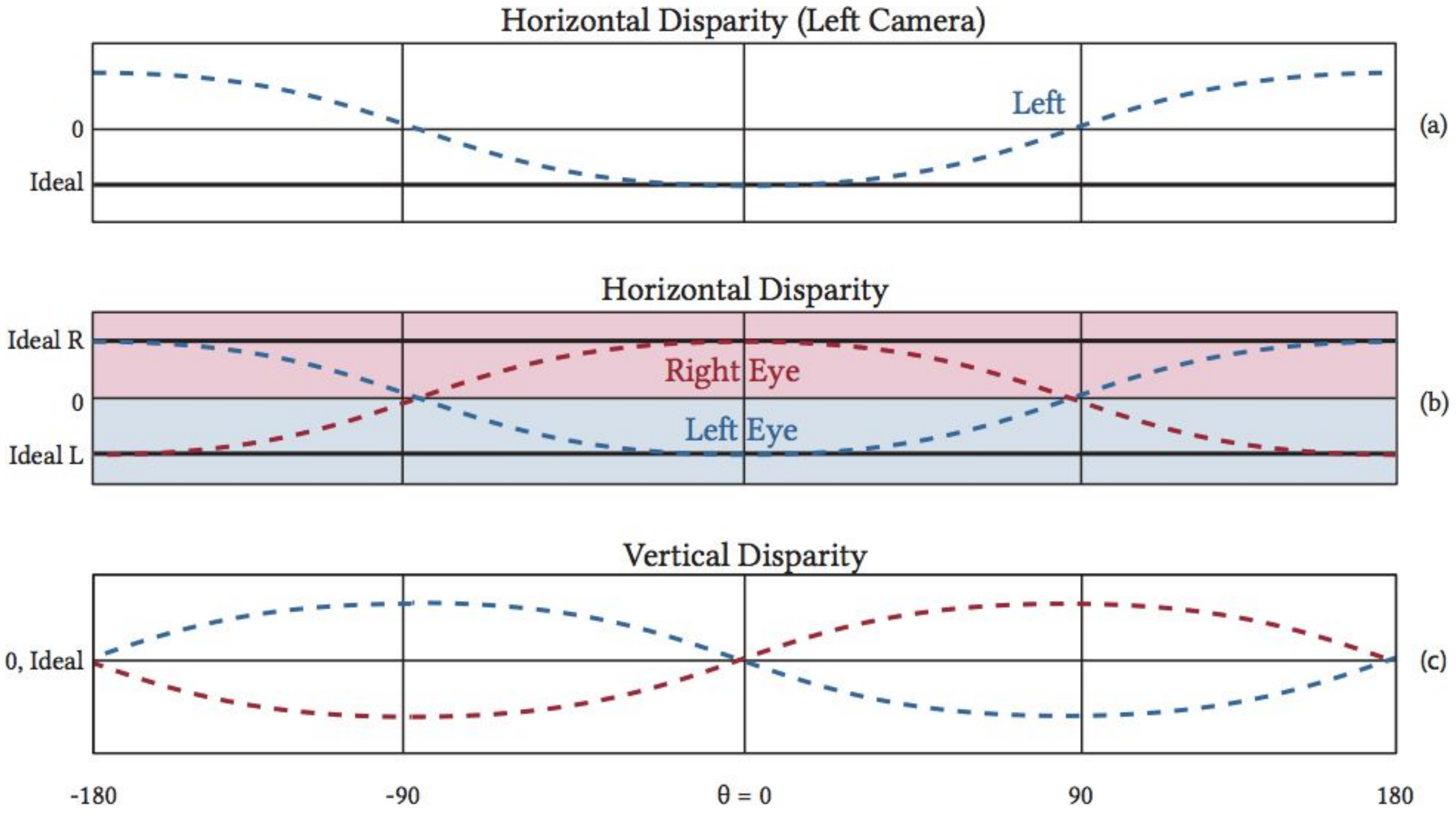
Problem: Stereo Baseline Fluctuates With View Angle



Apparent stereo baseline decreases by cos(θ) if rays are mapped directly

Matzen et al. SIGGRAPH 2017

Problem: Both Horizontal and Vertical Disparities Fluctuate



Matzen et al. SIGGRAPH 2017

Problems

- Disparity: incorrect baseline as view angle changes
- Occlusion: each camera blocks the other’s view!

Matzen et al. SIGGRAPH 2017

Partial Solution: Computational Photography

3D reconstruction

- **Computer vision on stereo**

  views Disparity correction

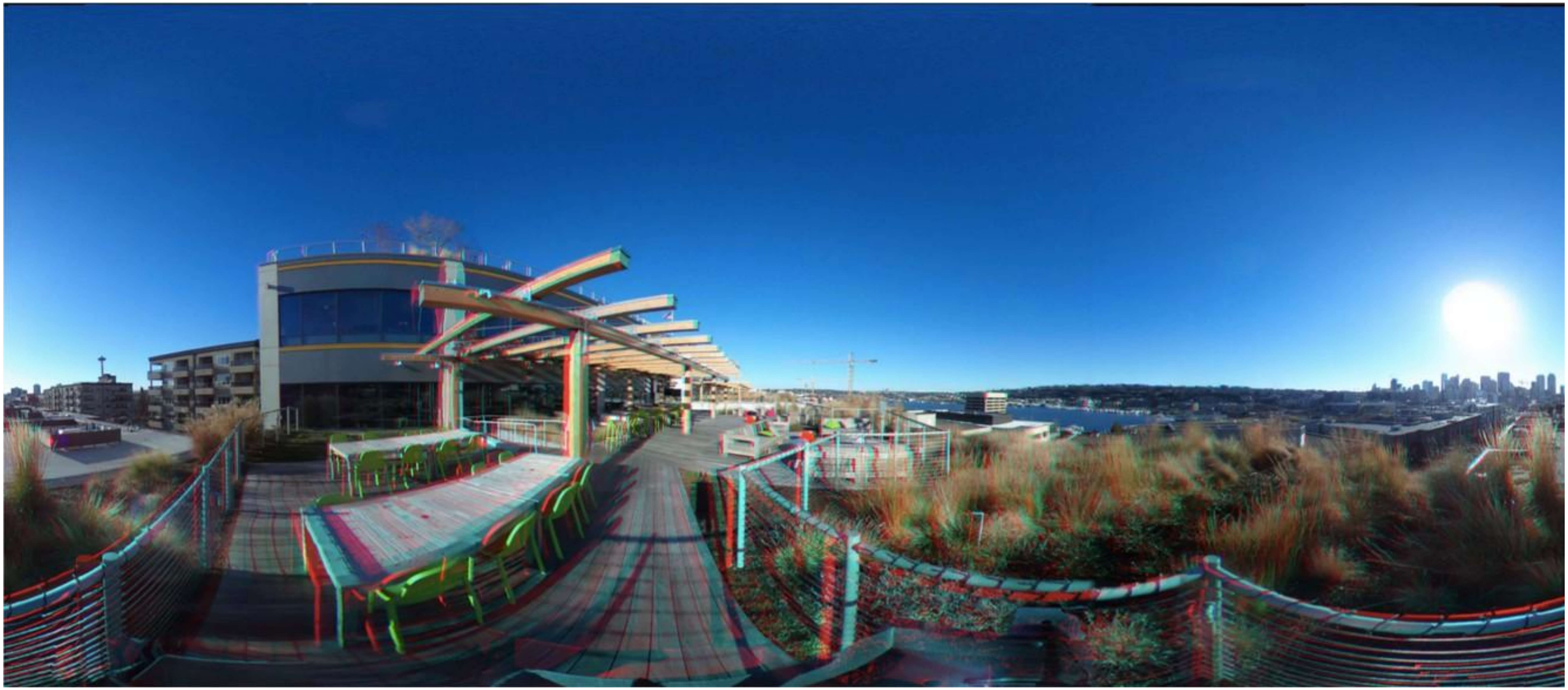
  - **Use 3D model to correct stereo disparities**
    - **e.g. amplify horizontal disparities by 1/cos(θ)**
- **Flip views when facing**

  backwards Hole filling

Matzen et al. SIGGRAPH 2017

  - **Cut out view of other camera, and fill hole with**

Spherical Stereo Result



Matzen et al. SIGGRAPH 2017

What Pairs of Viewpoint Positions Do We Want To Sample?

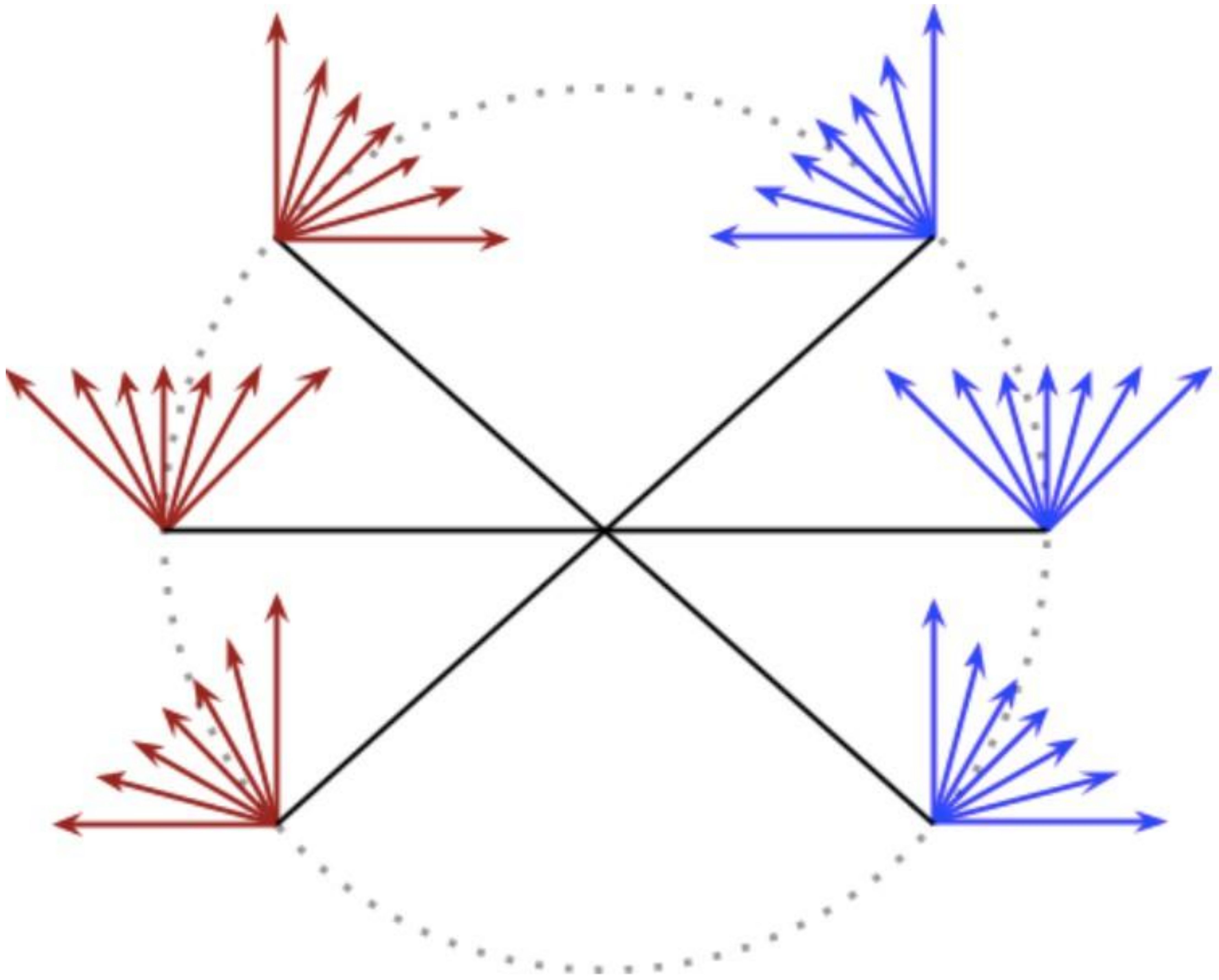
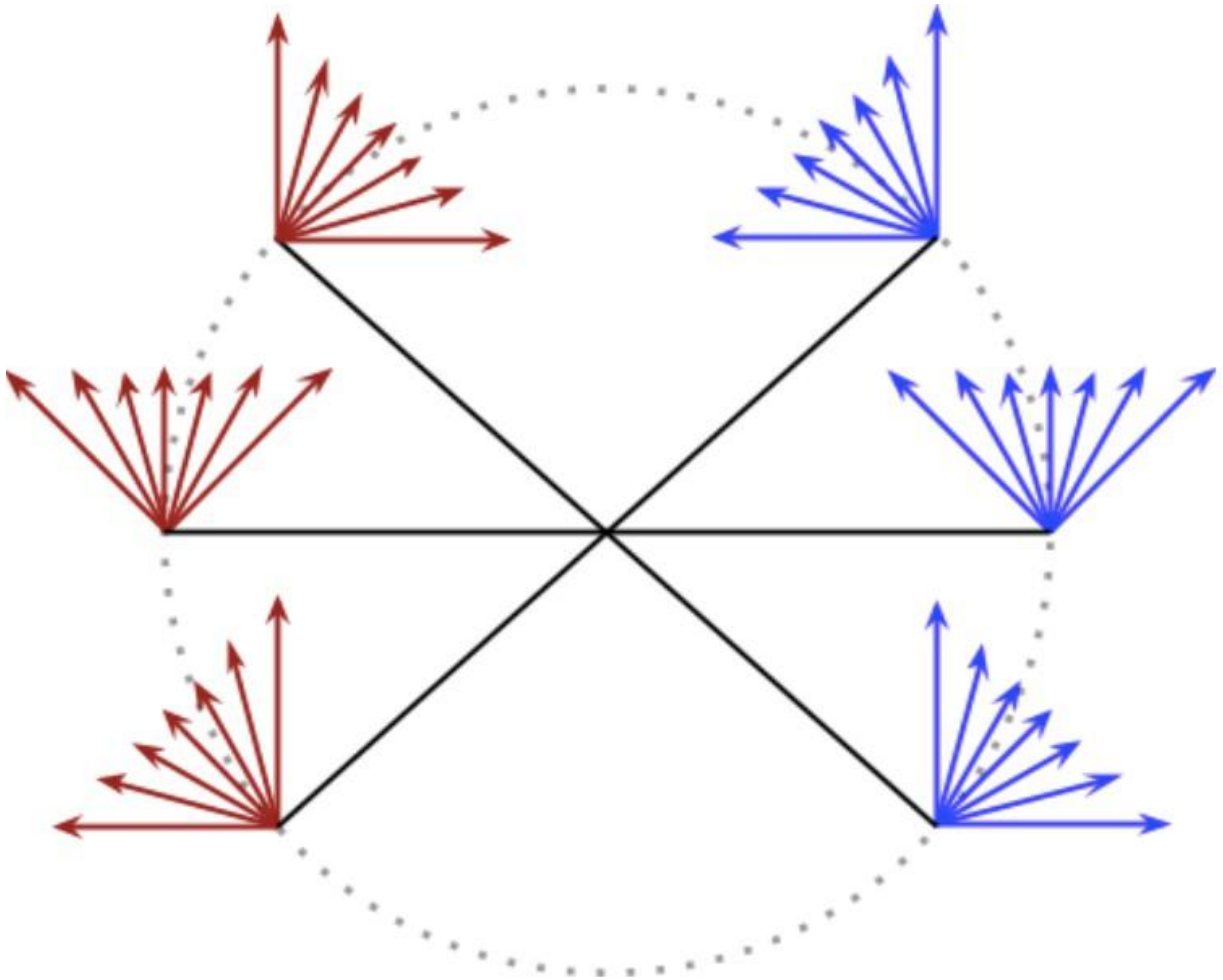


Image Credit: Google Inc.

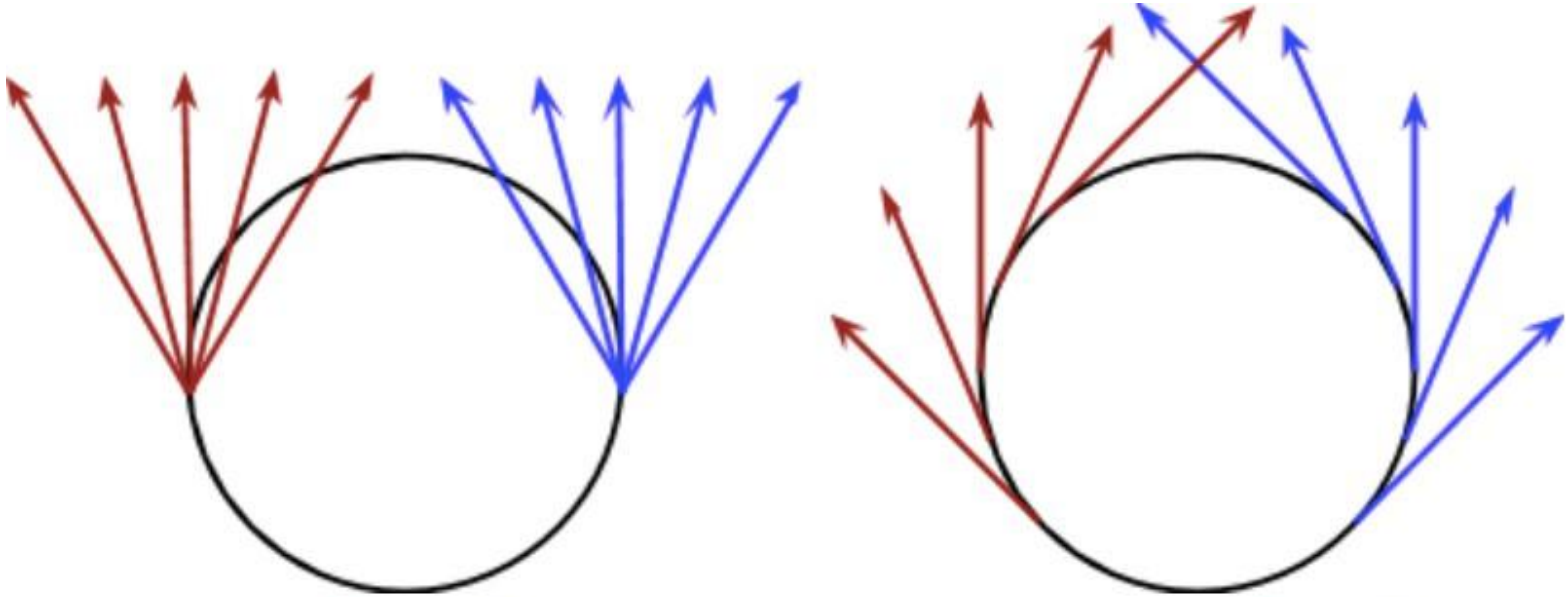
Idea: Spin a Pair of Cameras About Midpoint



Store a set of movie pairs (one per angle) But that’s a lot of data

Image Credit: Google Inc.

Omni-Directional Stereo Approximation

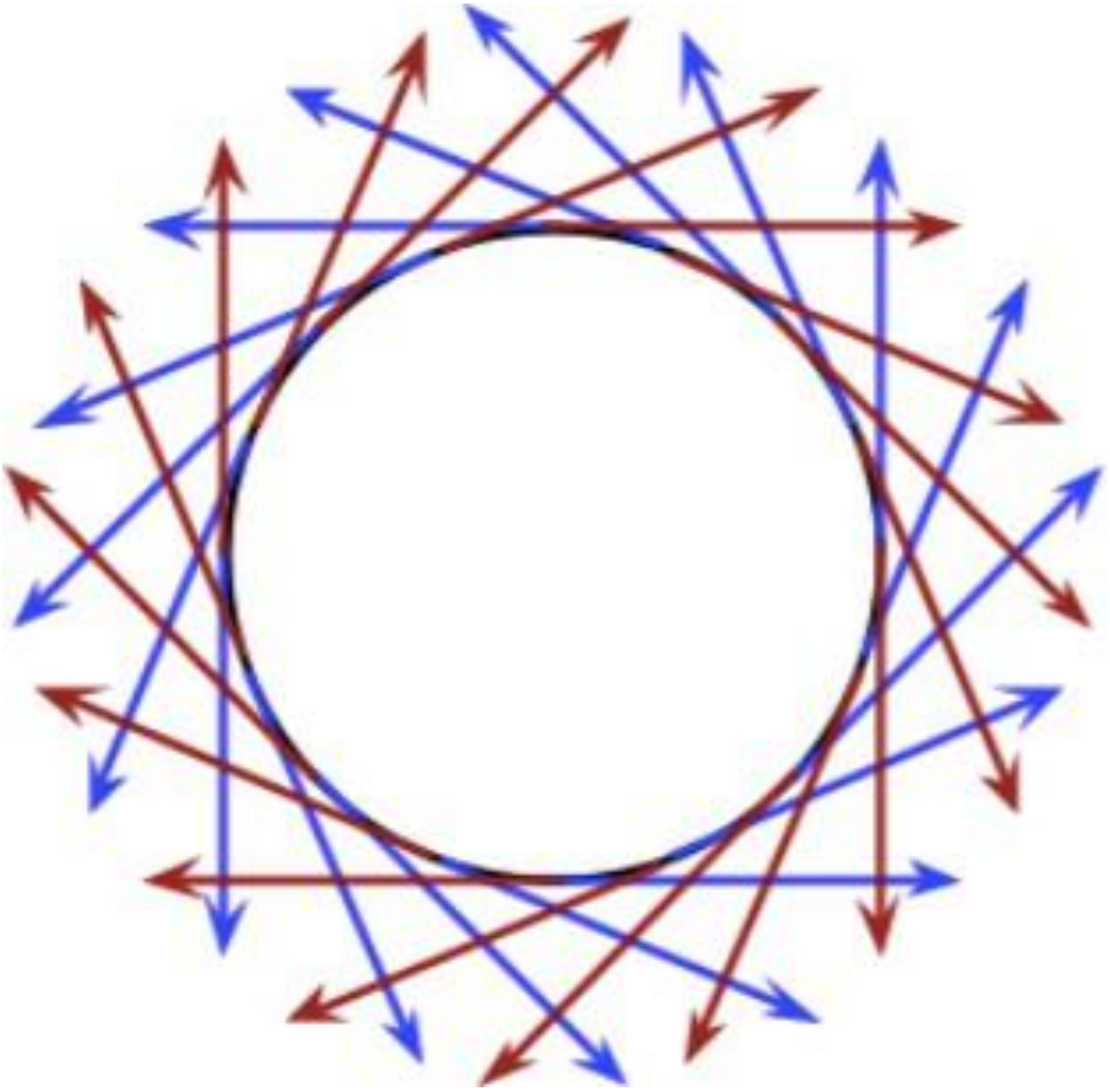


full-frame **left** and **right** eyes

ODS-approximated **left** and **right** eyes

Image Credit: Google Inc.

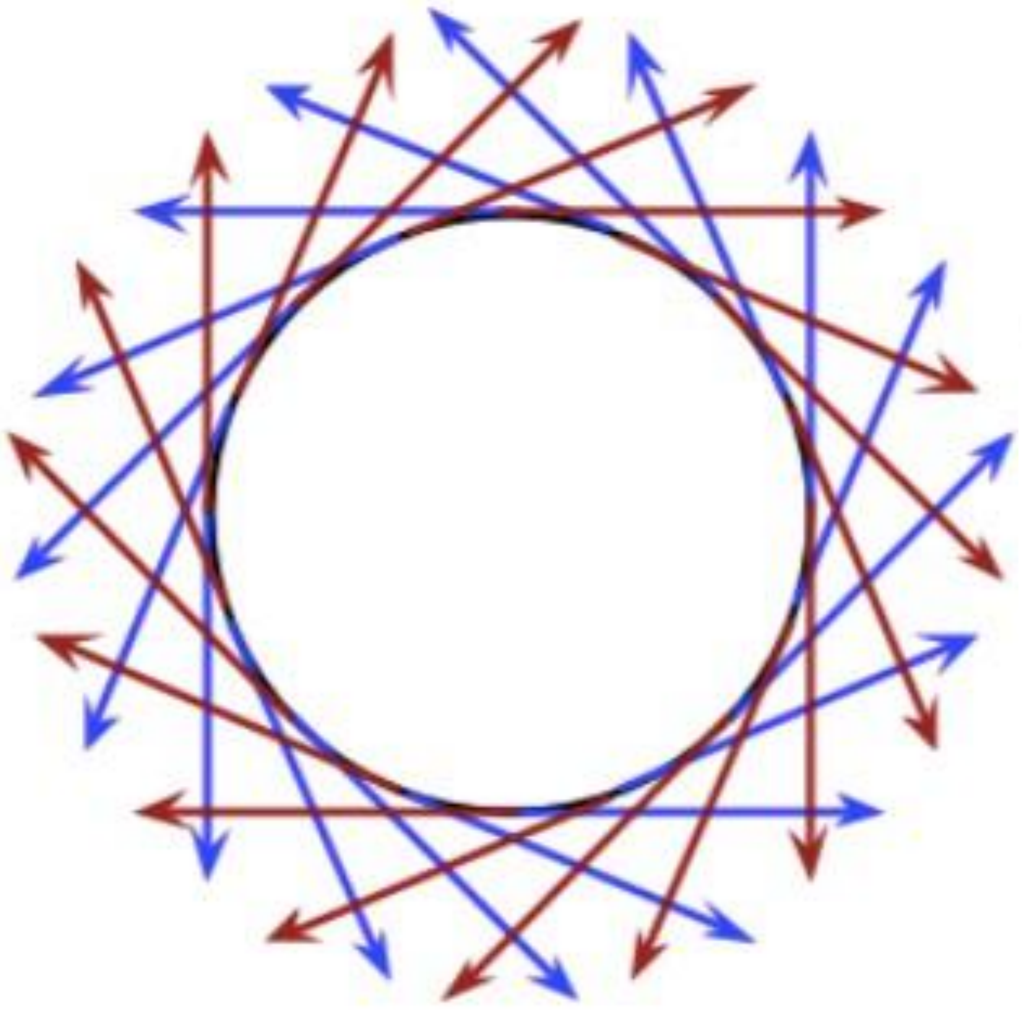
Omni-Directional Stereo Approximation

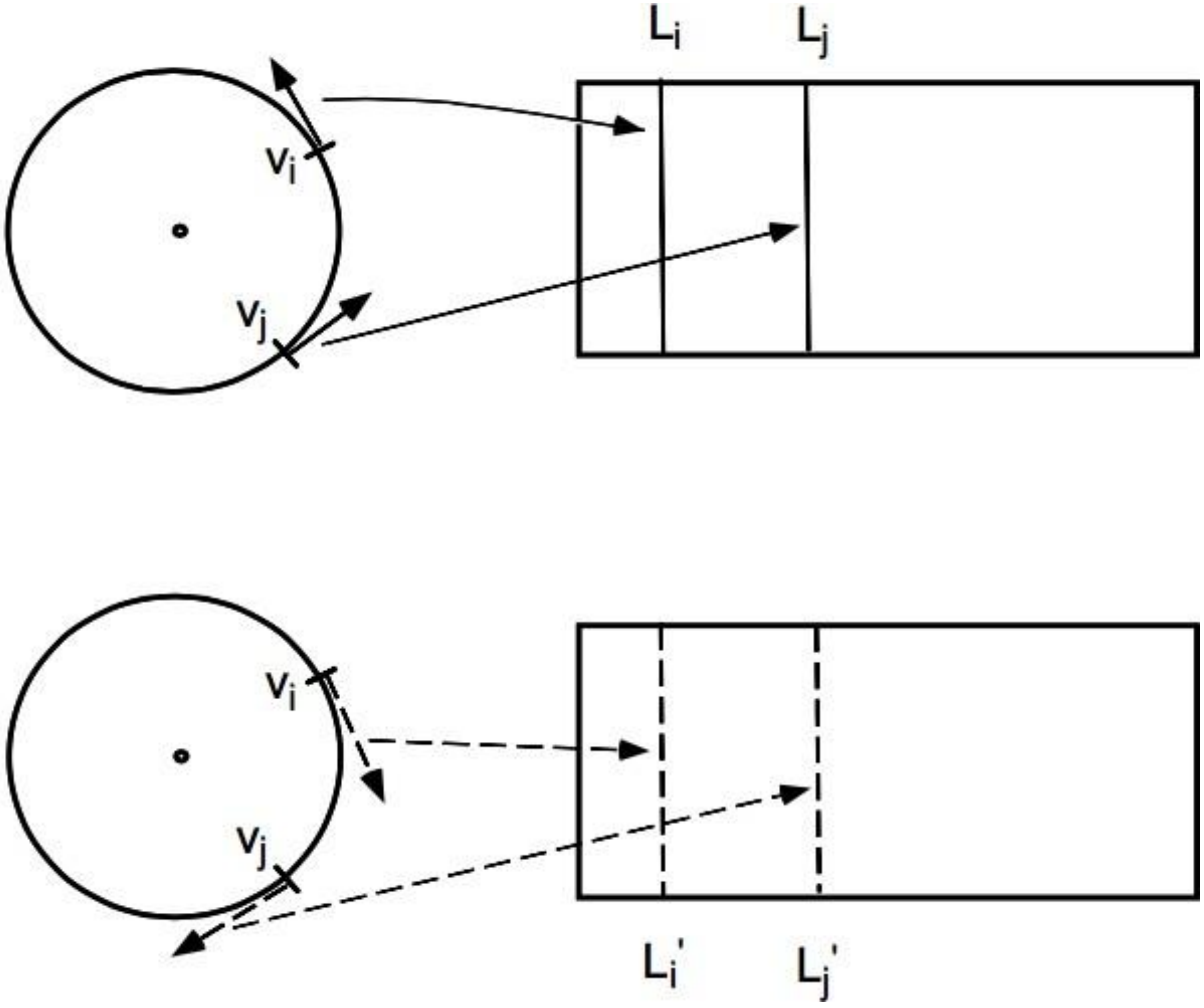


Extended to be omnidirectional

Image Credit: Google Inc.

Spinning Camera





Concentric Mosaics
Shum and He, SIGGRAPH 1999

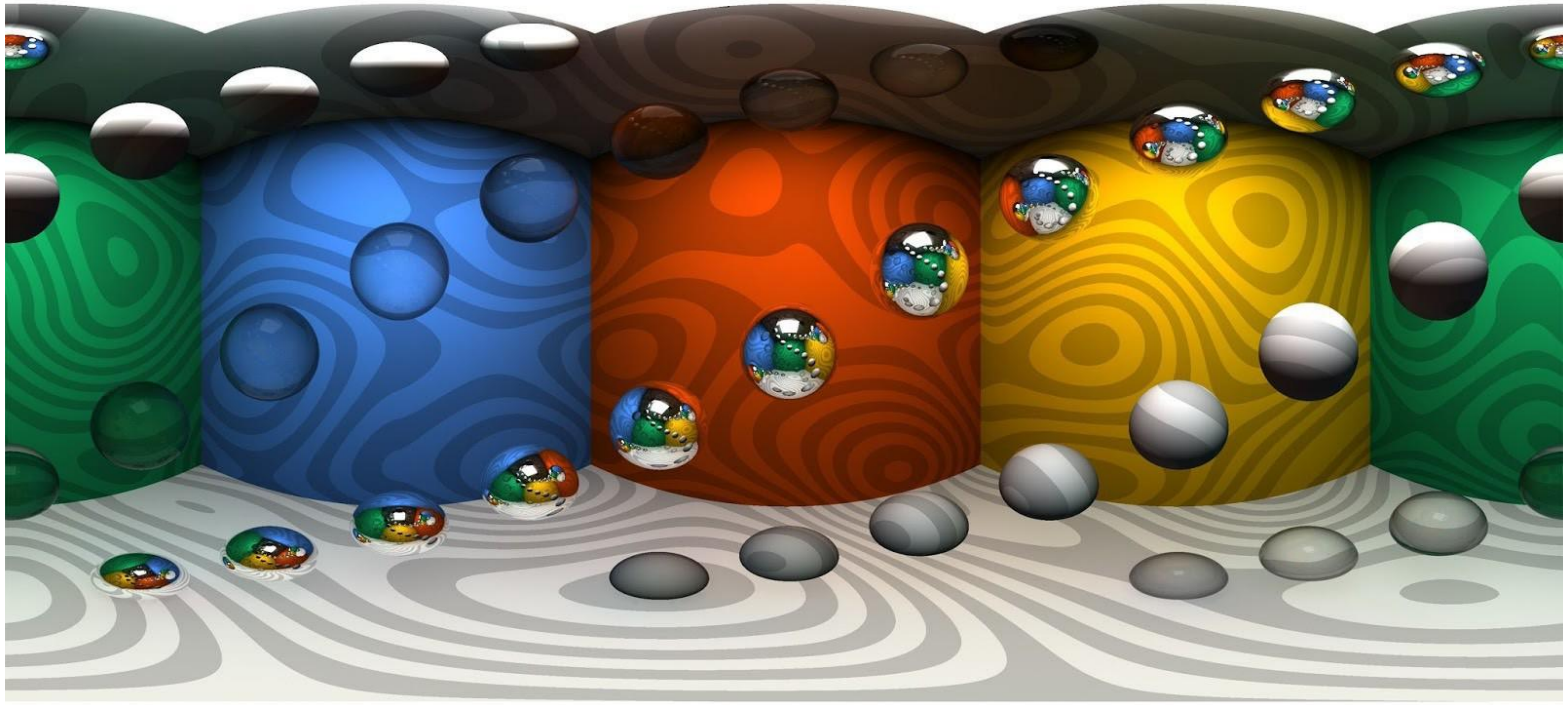
Omni-Directional Stereo Representation

Encode left/right views as just two spherical images

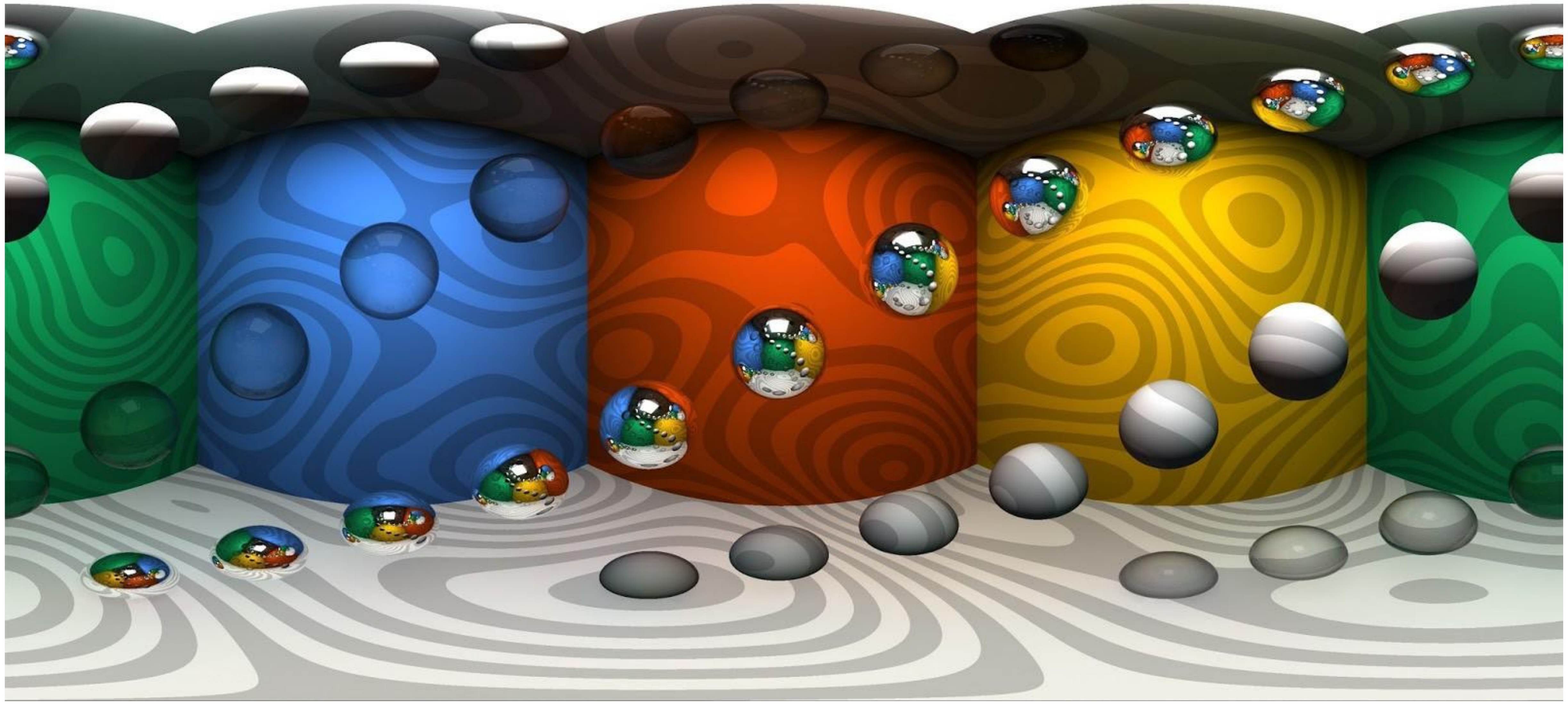
- **Render left and right views for each angular view independently, with regular viewing software**
- **Efficient and compact, but this is an approximation**
  - **Straight lines may appear slightly curved**
  - **Vertical disparity for close objects incorrect**

Example (Rendered)

Left Eye

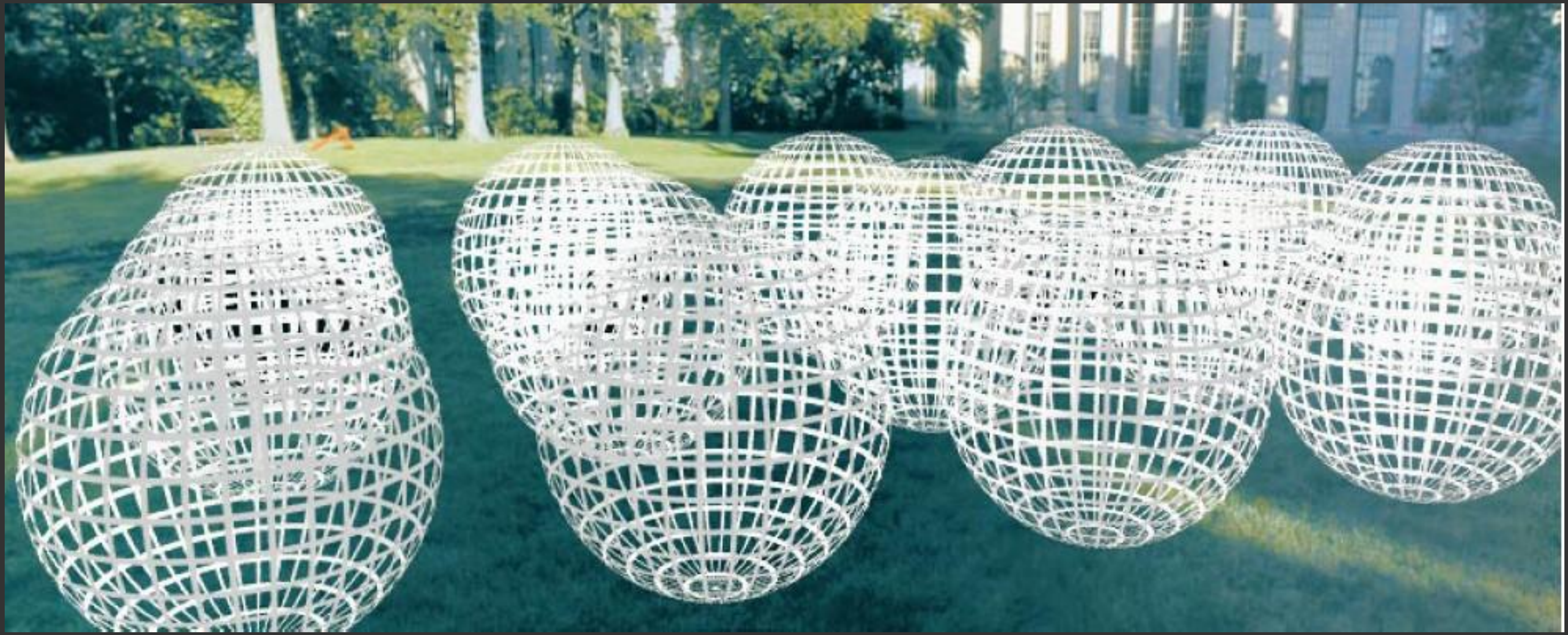


Right Eye



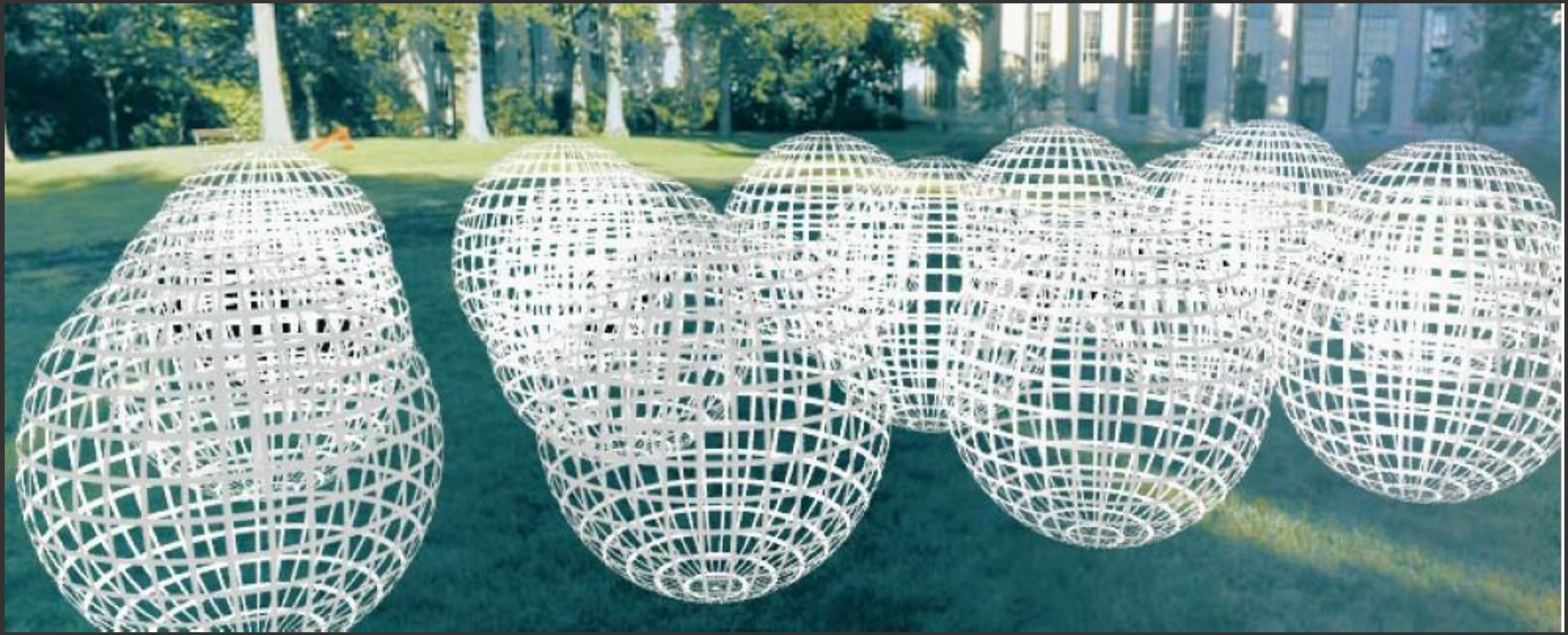
Moving-Viewpoint Imaging
(Full Plenoptic Function?)

The 5D Plenoptic Function



$$P\,(\checkmark, \text{¢}, V_x, V_y, V_z)$$

4D Light Field



$$P\,(\checkmark,\ rp,\ V_x,\ V_y\,) = P\,(u,\ v,\ s,\ t)$$

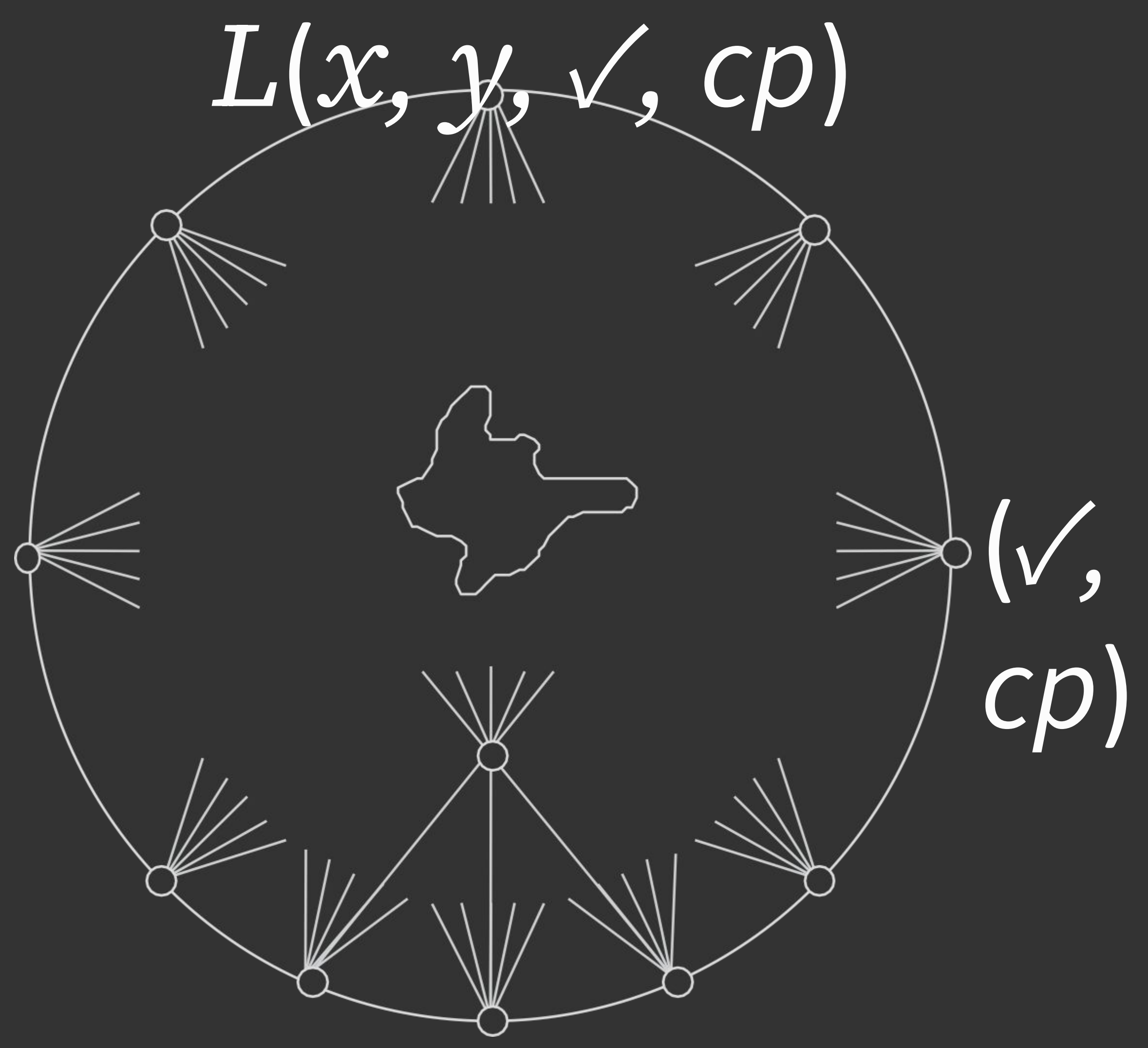
- **In a region of free-space, 5D plenoptic function simplifies to 4D because light is constant along a ray**

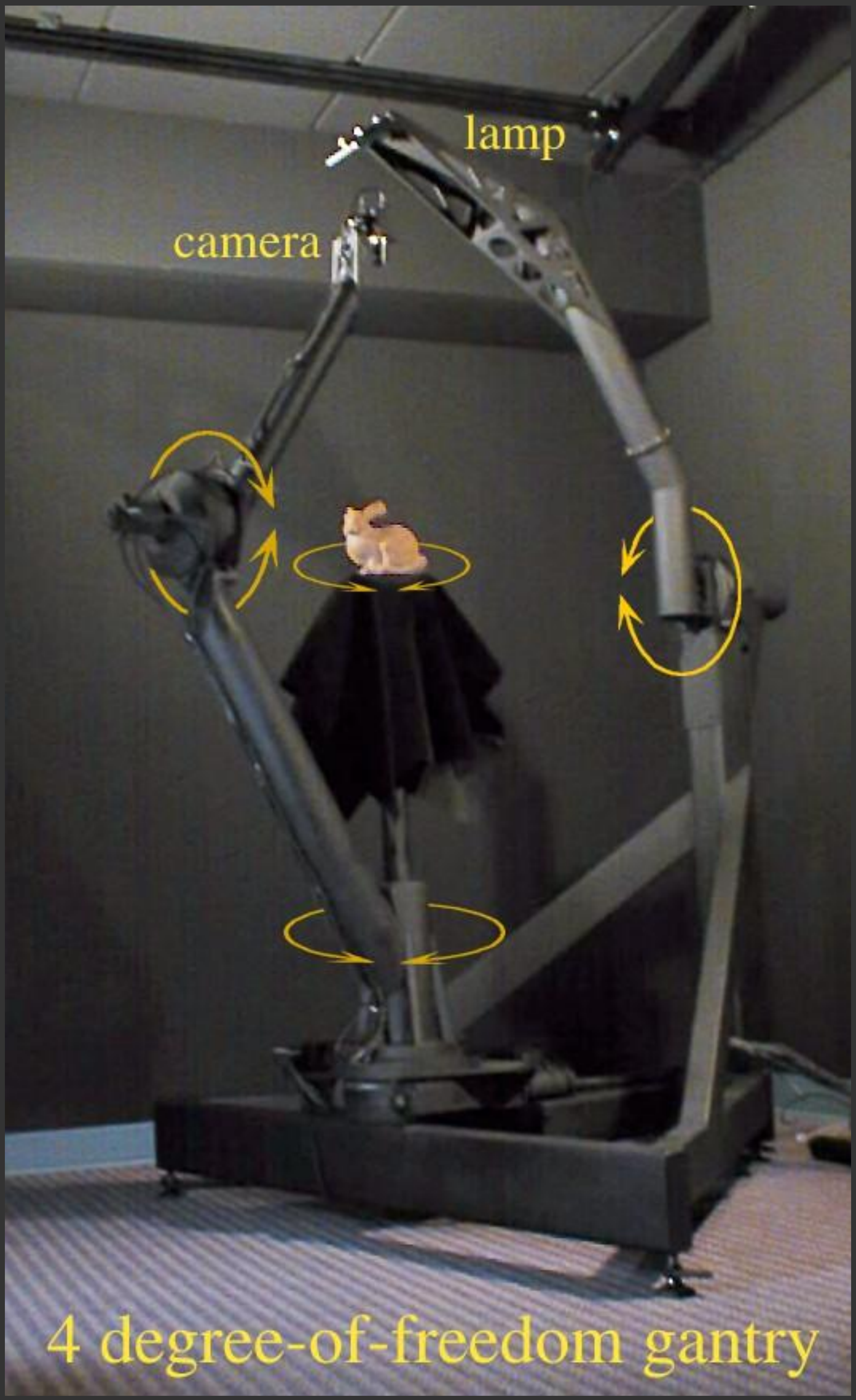
Light Field Capture Robot

Original light field rendering paper

Take photographs of an object from all points on an enclosing sphere

Captures all light leaving an object – like a hologram

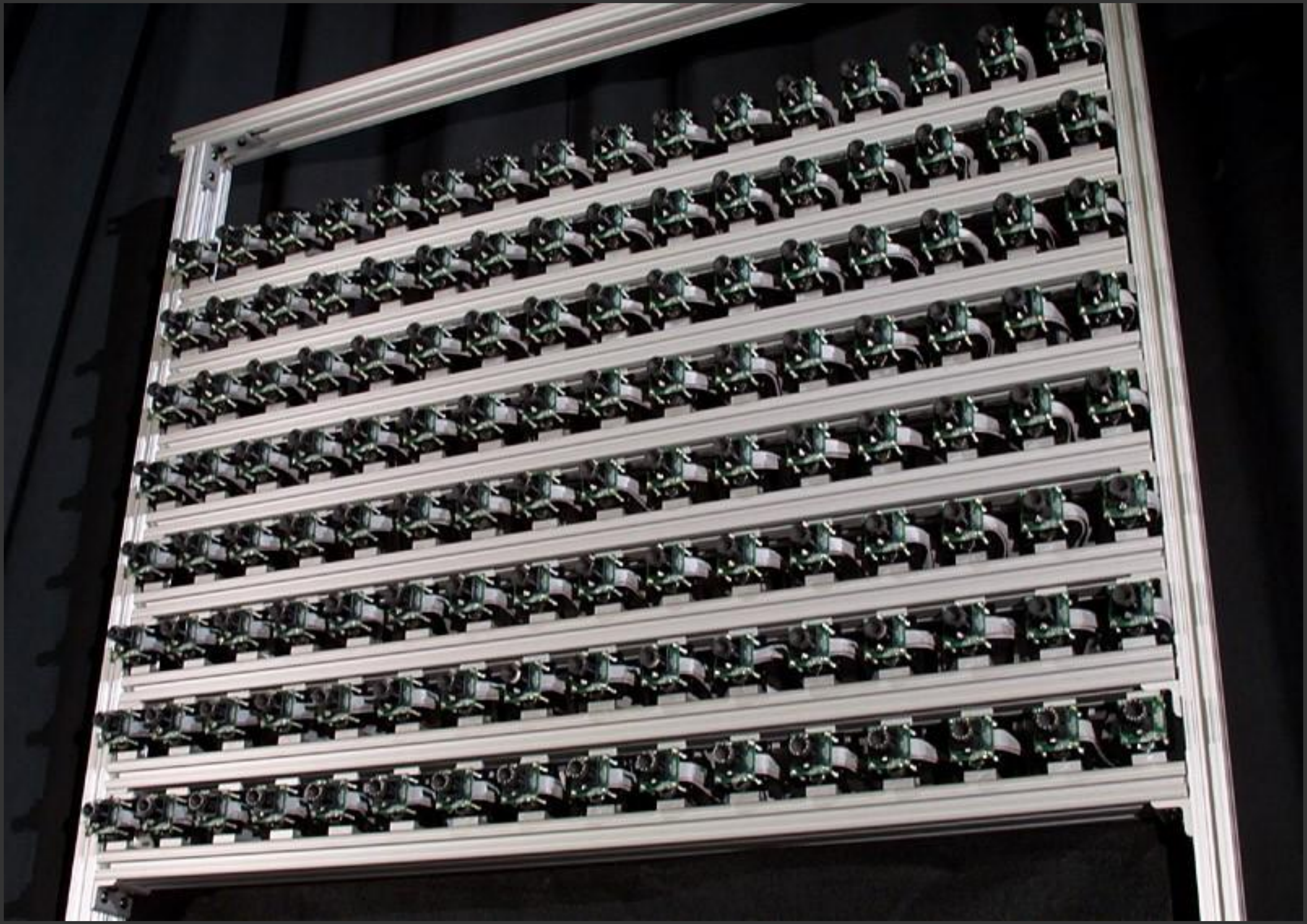




[Levoy & Hanrahan 1996] [Gortler et al. 1996]

Slide credit: Pat Hanrahan

Multi-Camera Array ⇒ 4D Light Field



[Wilburn et al. SIGGRAPH 2005]

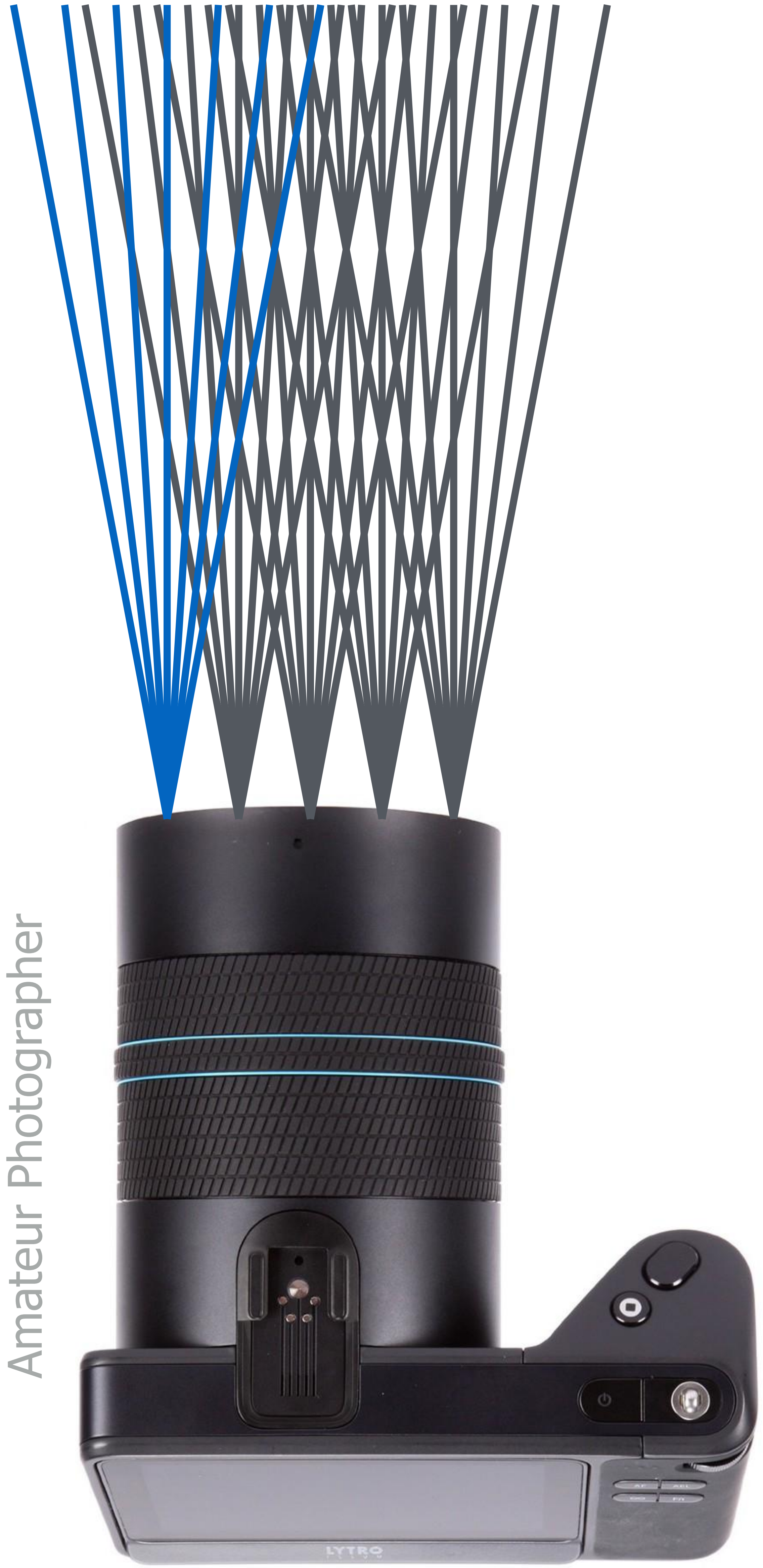
Slide credit: Pat Hanrahan

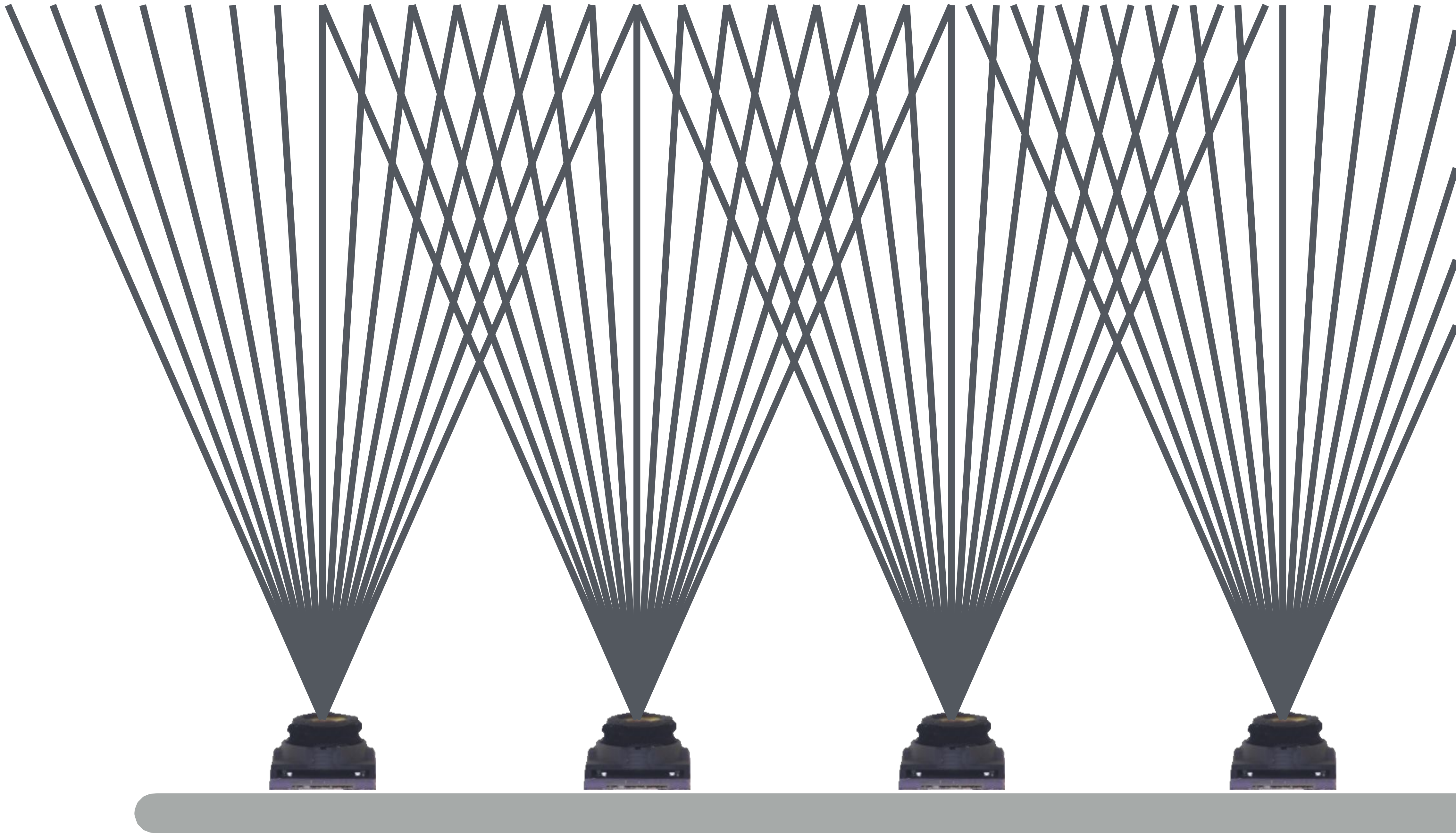
[Wilburn et al. SIGGRAPH 2005]

Handheld 4D Light Field Camera (Plenoptic Camera)

Lytro Gen-2 Light Field Camera

Handheld Light Field Camera vs Camera Array





Camera array: e.g. 10x10 views distributed across large planar support

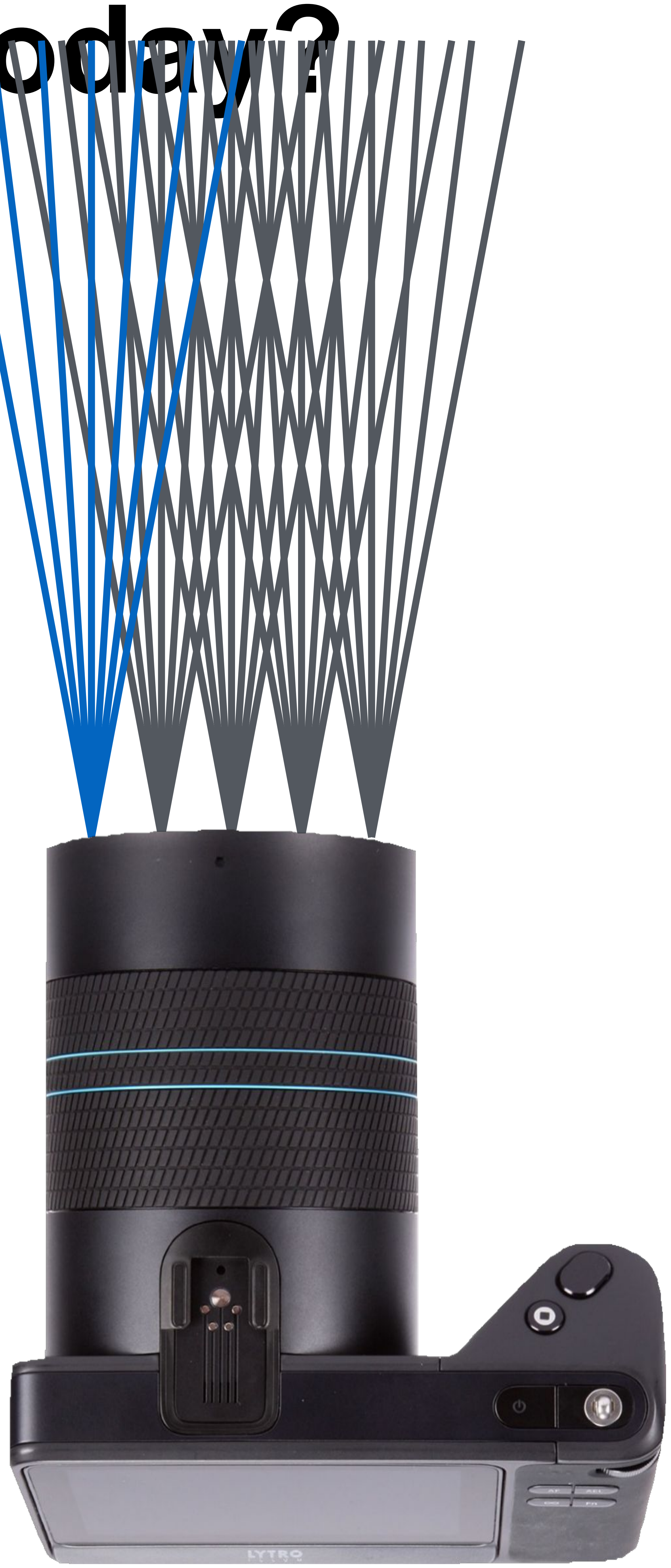
Plenoptic camera: e.g. 14x14 views distributed across small lens pupil Note: antialiased across views, unlike camera array

The Intimacy of VR Graphics

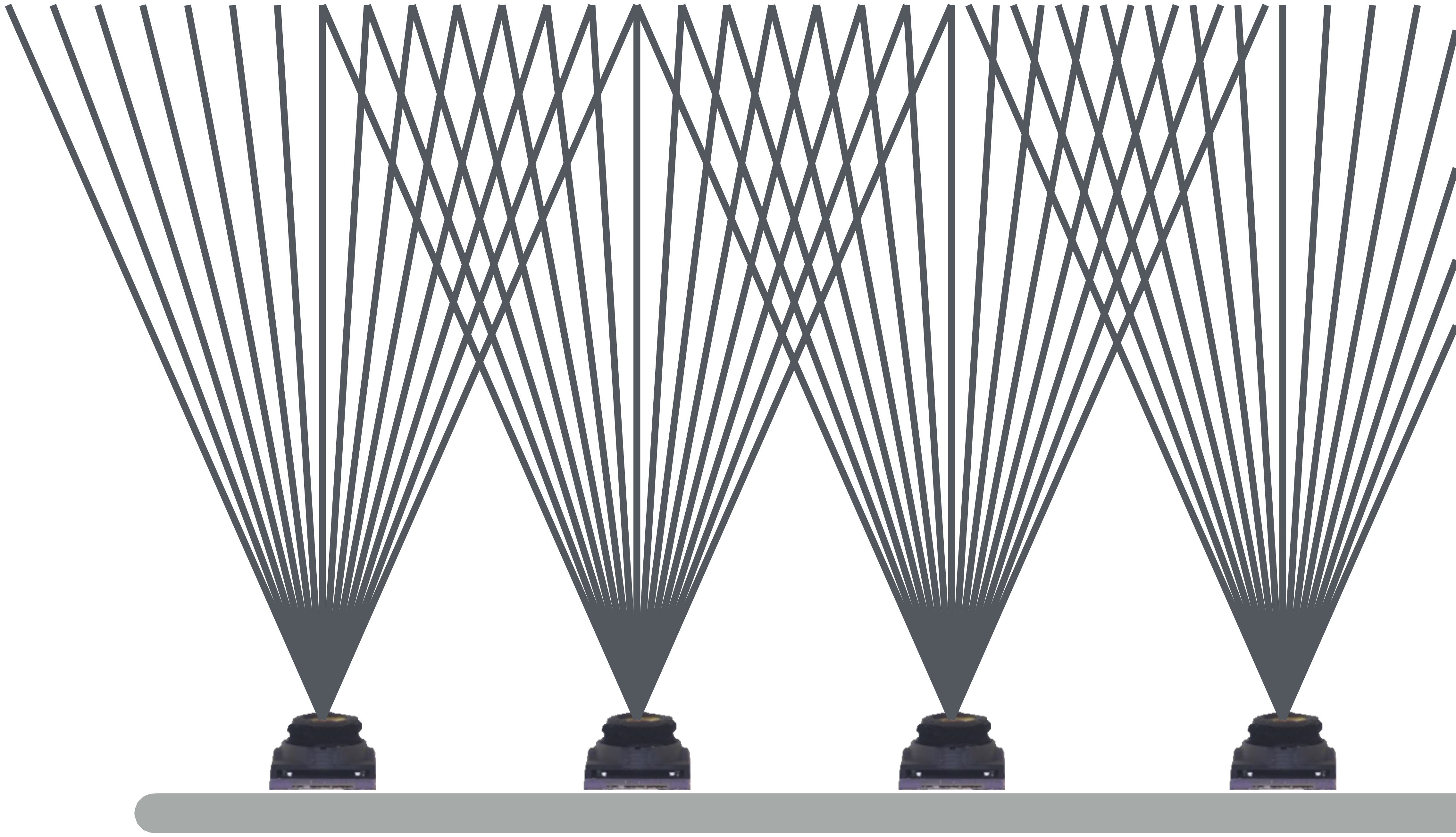
Google's Tilt Brush on HTC Vive

A Challenge: Intimate Proximity in VR Imaging

How Dense Are Camera Views Today?



Amateur Photographer

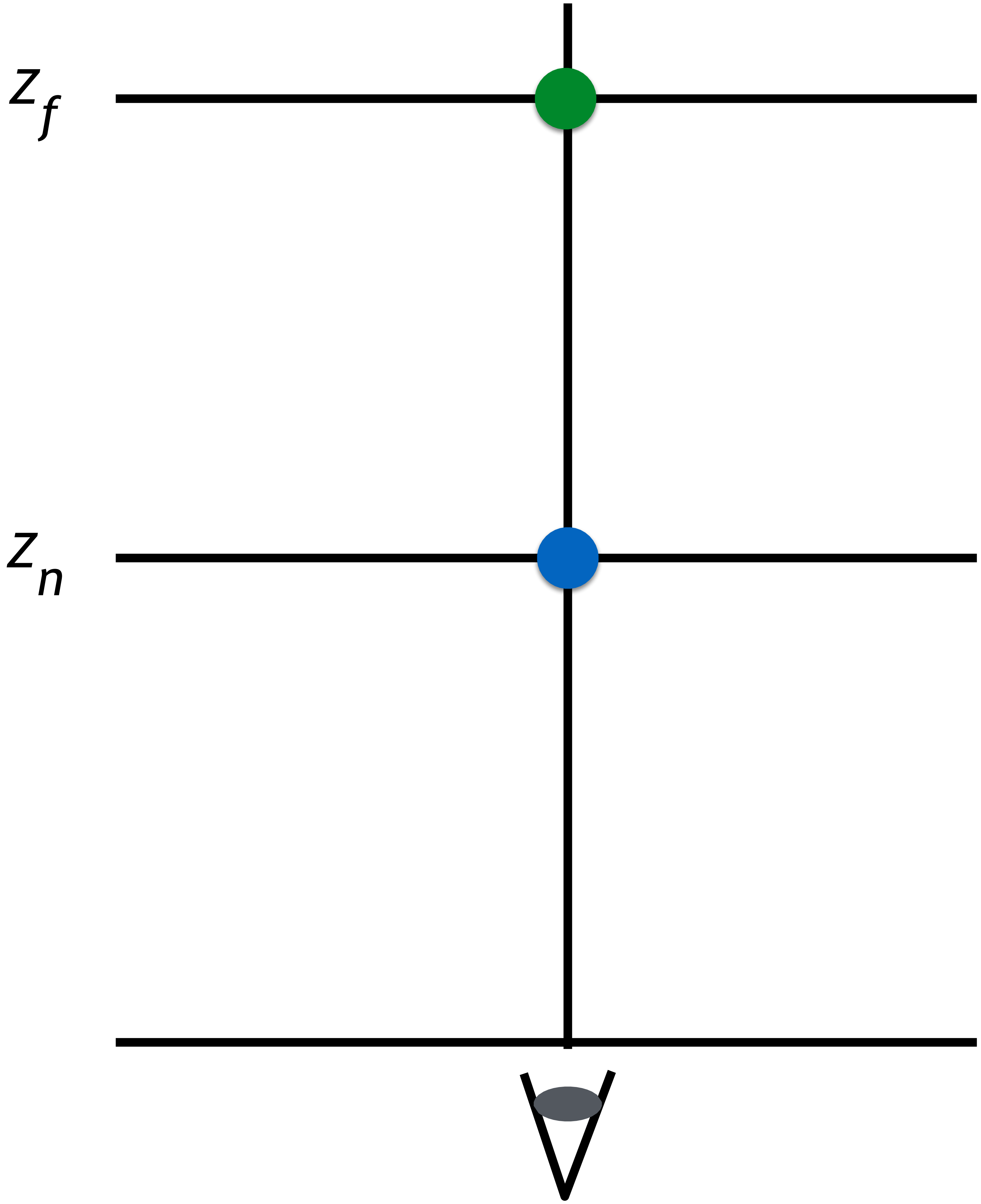


Multi-camera
arrays: Plenoptic
cameras:

50 - 100 views
100 - 200 views

How Dense Must Cameras Views Be?

How Dense Must Camera Views Be?

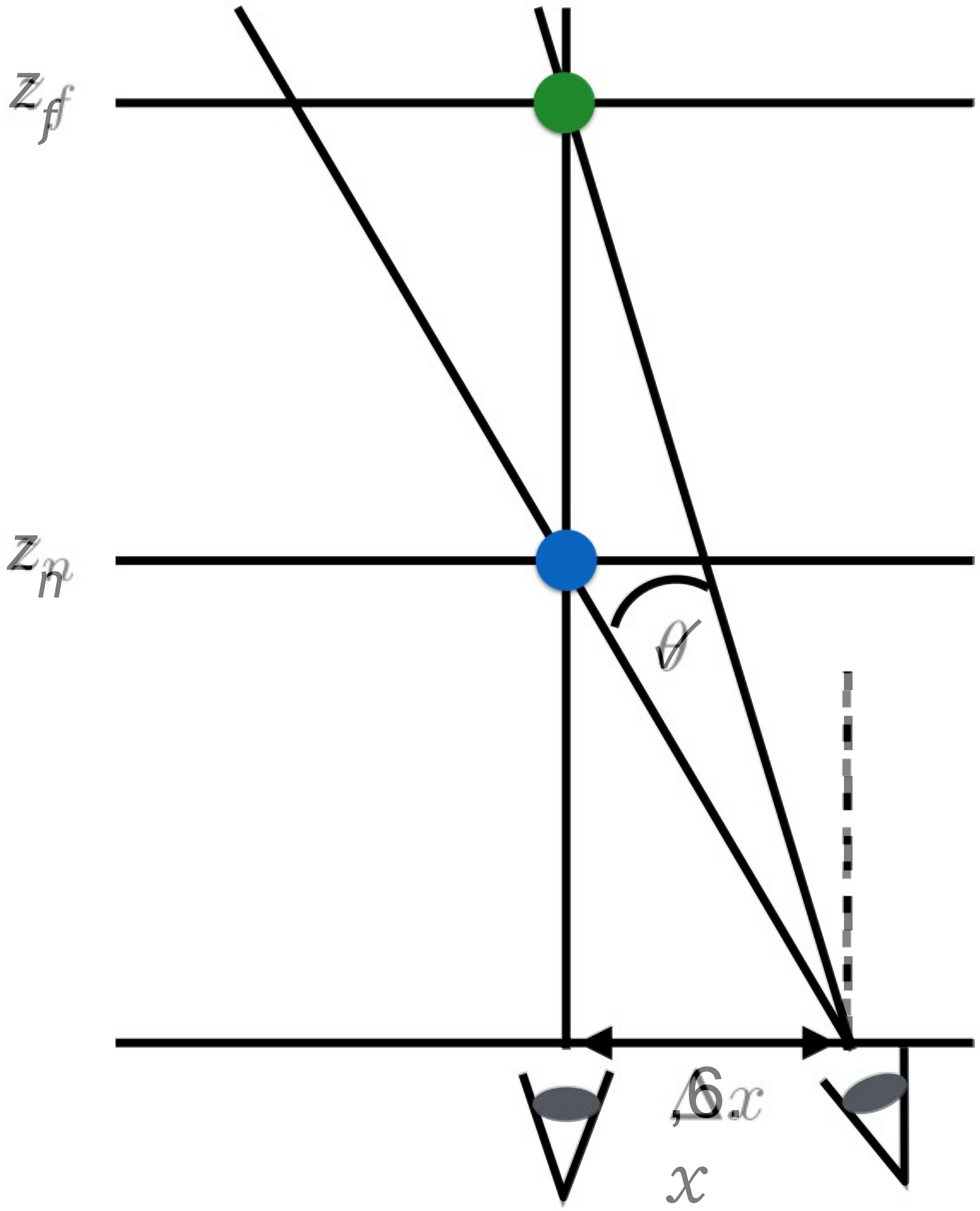


Child in lap, front to back of head

$z_n = 0.3\text{m}$

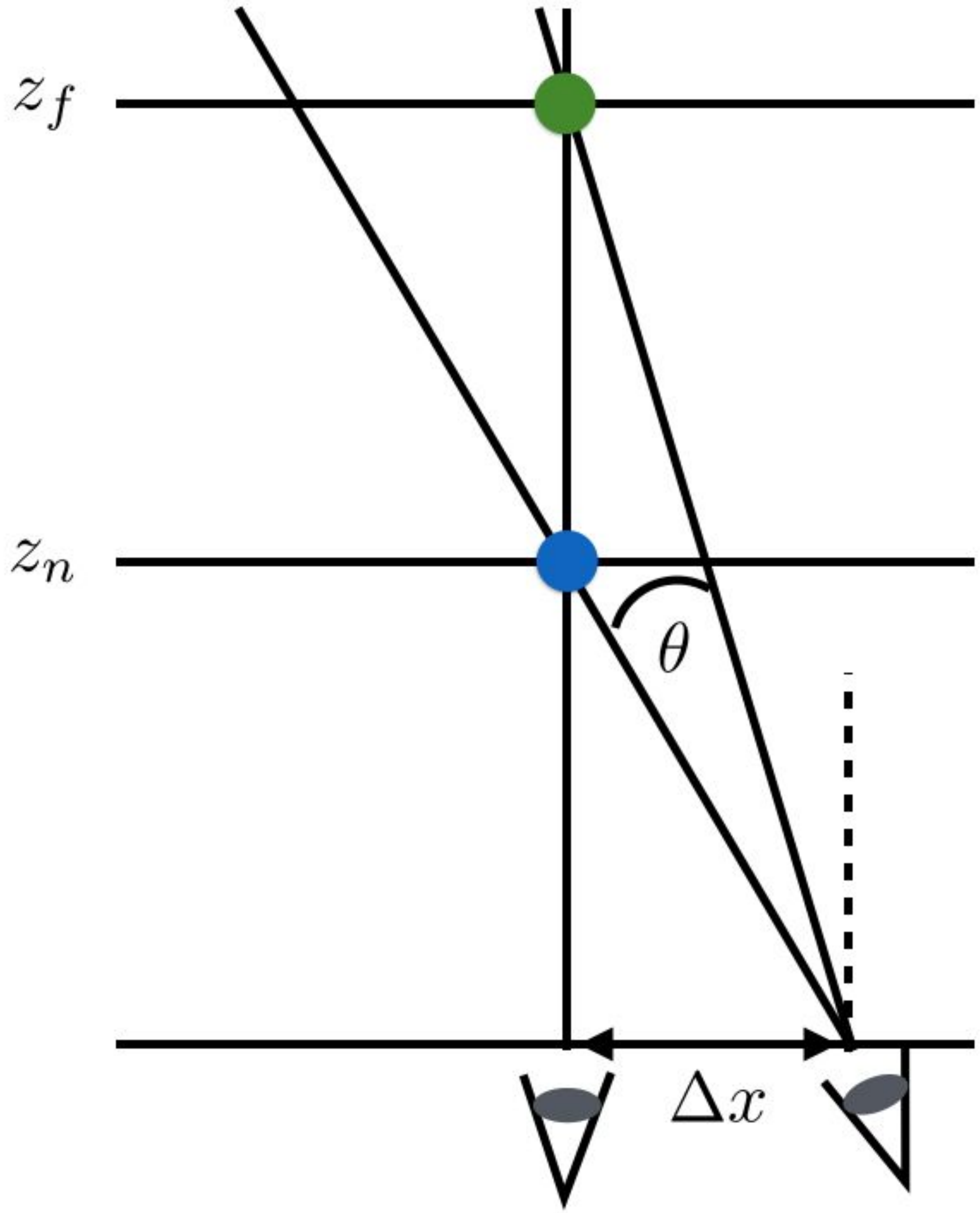
$z_f = 0.6\text{m}$

How Dense Must Camera Views Be?



What is the minimum lateral eye movement Δx so that we can visually distinguish the close and far features?

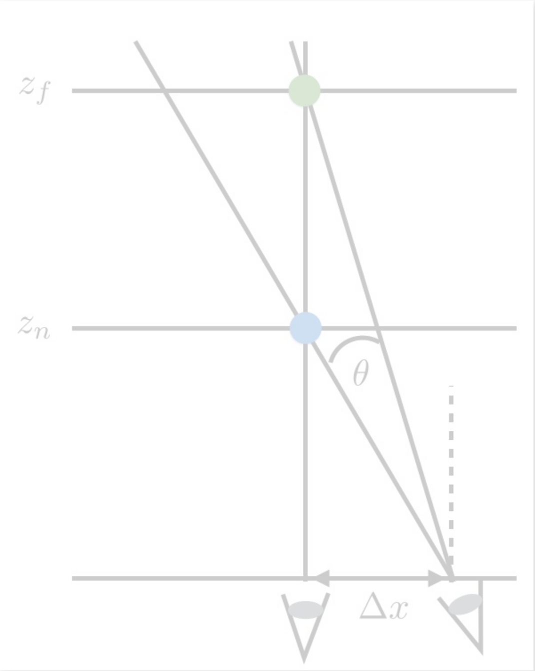
How Dense Must Camera Views Be?



20/20 vision: ✓ ⇡ $(1/60)^{0}$

Curent HMDs: ✓ ⇡ $(1/10)^{o}$

How Dense Must Camera Views Be?



Solving for minimum lateral motion:

$$x = \frac{(z_f - z_n) - {}^{p}(z_f - z_n)^2 - 4 \tan^2 \checkmark z_n z_f}{2 \tan \checkmark}$$

Child in lap, front to back of head

$z_n = 0.3\mathrm{m}$

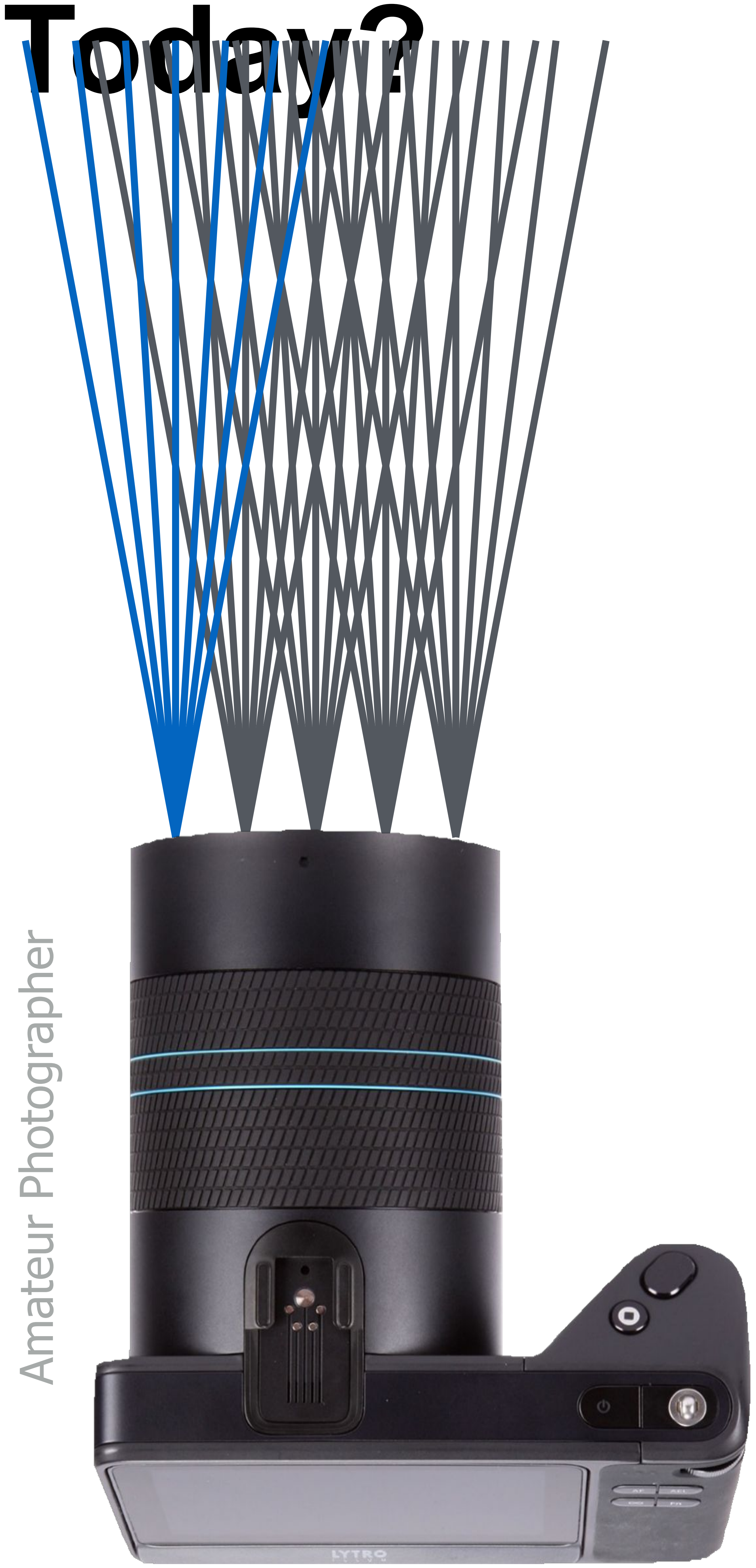
$z_f = 0.6\mathrm{m}$

20/20 vision: $\theta \approx (1/60)^\circ \implies \Delta x \approx (1/1719)\mathrm{ft}$

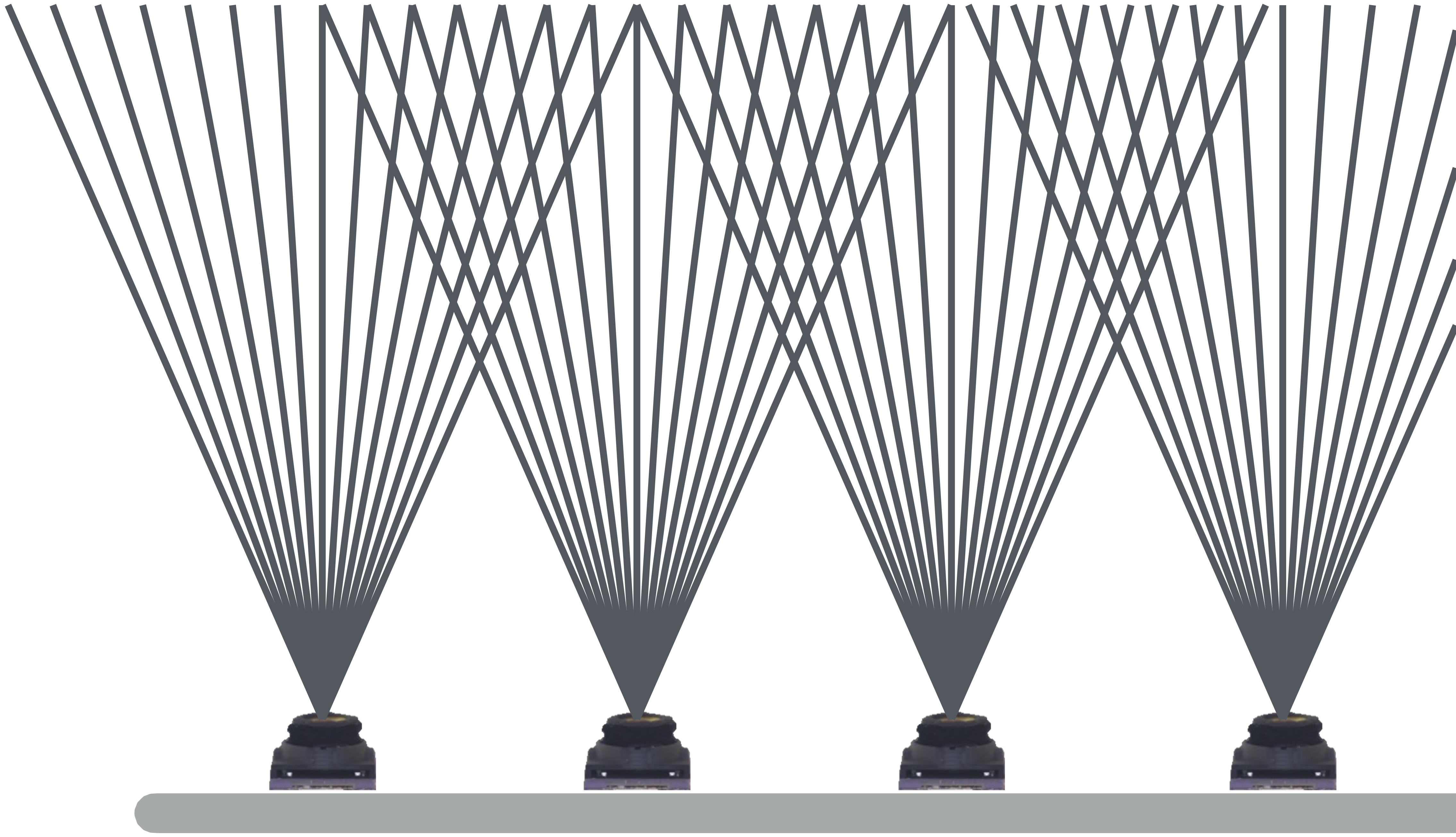
Curent HMDs: $\theta \approx (1/10)^\circ \implies \Delta x \approx (1/286)/\mathrm{ft}$

20/20 vision:	**millions of views per square foot**
Current HMDs:	**a hundred thousand views per square foot**

How Dense Are Camera Views Today?

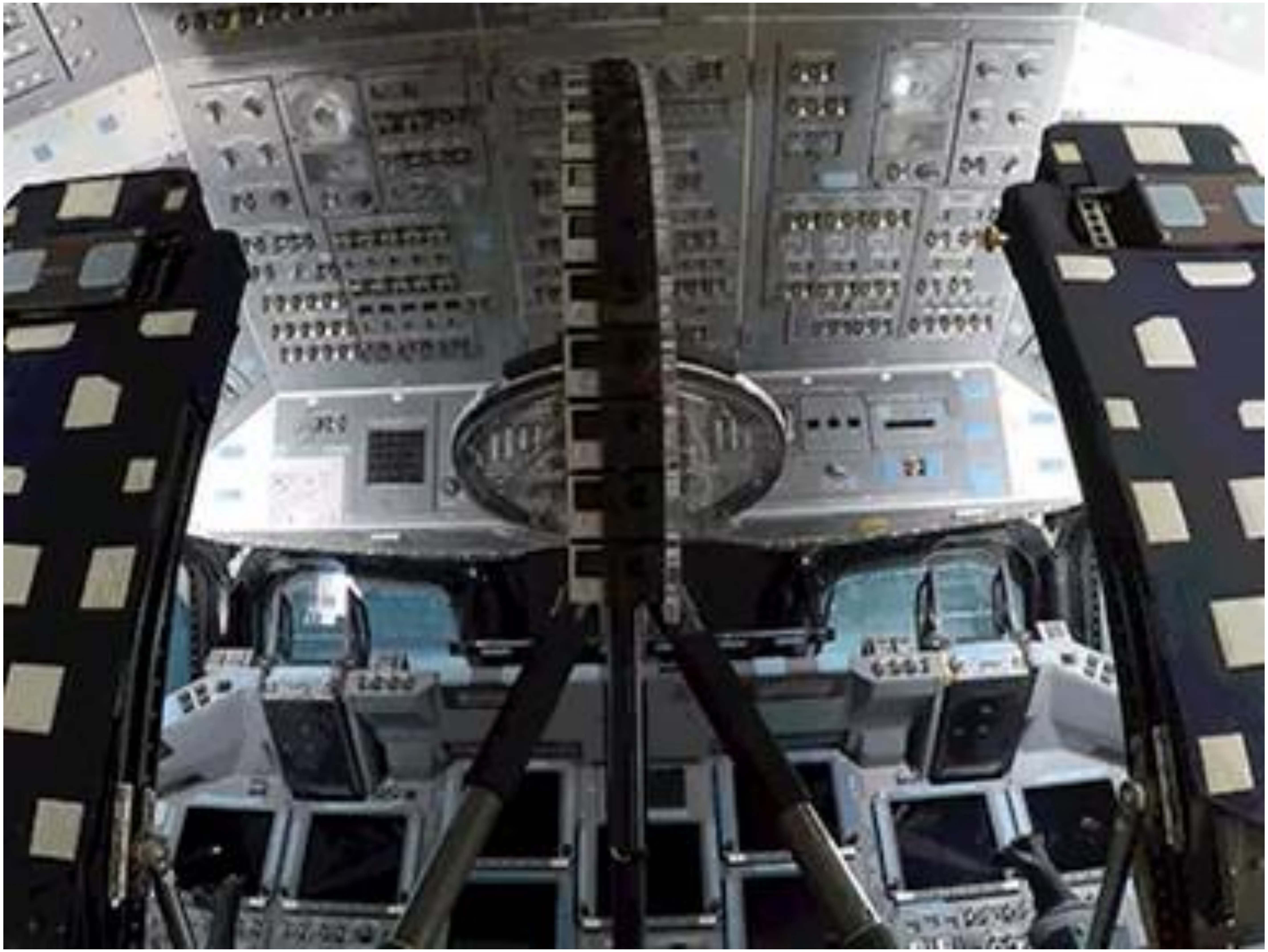


Amateur Photographer



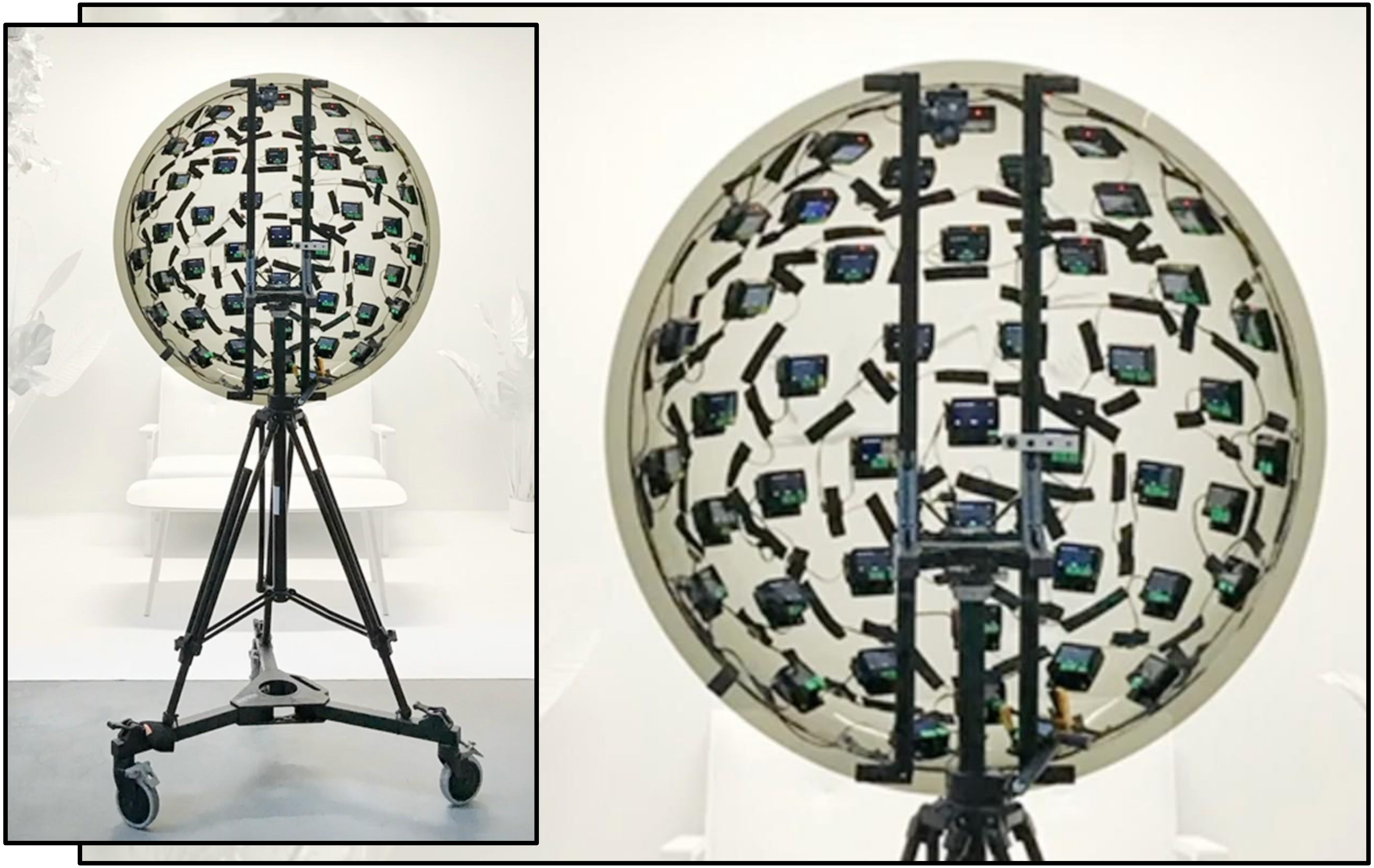
Multi-camera arrays:	50 - 100 views
Plenoptic cameras:	100 - 200 views

Google VR Camera Rig



Paul Debevec, Google

Imaging for Virtual Reality



Google 6 DOF Light Field Camera. Broxton et al. 2019.

Imaging for Virtual Reality

Google 6 DOF Light Field Camera. Broxton et al. 2019.

Recent Breakthrough - Novel View Interpolation

Inputs: sparsely sampled images of scene

Outputs: *new* views of same scene

NeRF, 2020, Mildenhall, Srinivasan, Tancik, et al.

VR Headset Imaging

VR Headset Imaging

Outward facing cameras

- **Live imaging of real world to enable rendering of “passthrough” imagery into VR**

- **“Inside-out” head tracking**
- **Hand / body tracking for user**

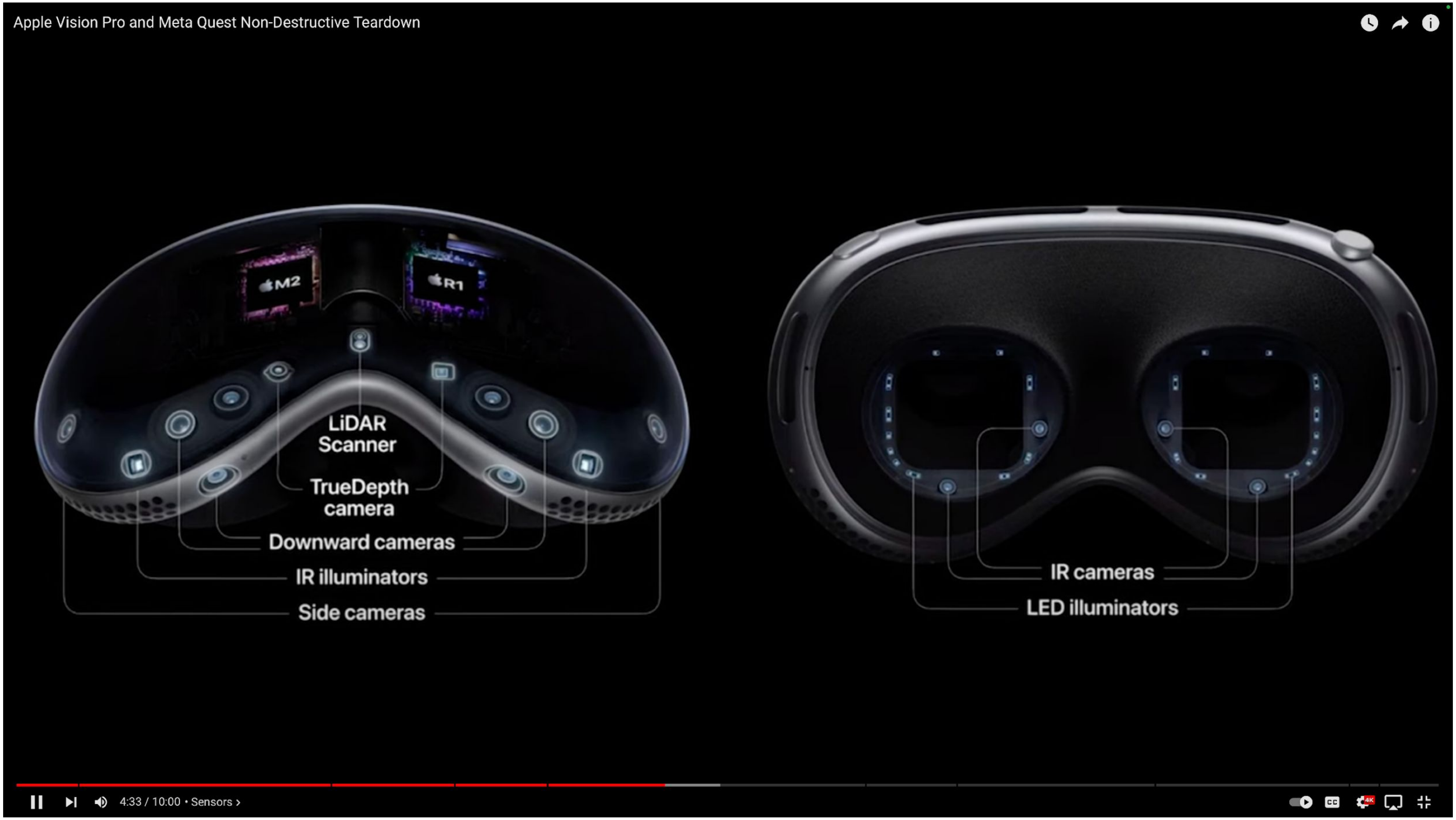
  interfaces Inward facing cameras

- **Image the eyes, face**

- **Detect gaze point for user interfaces**

- **Detect facial expressions for virtualizing user into VR to render for others (e.g. telepresence)**

Apple Vision Pro Sensors - Outward and Inward

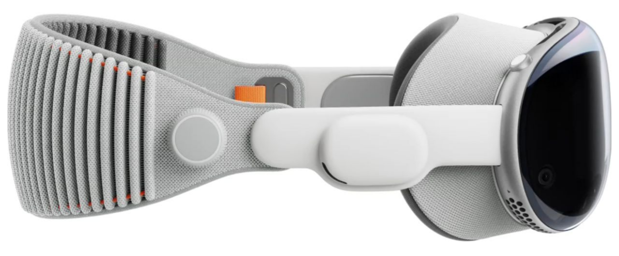


Apple Vision Pro and Meta Quest Non-Destructive Teardown

Apple Vision Pro, Lumafield teardown

https://youtu.be/V5_gja9O-GE?si=AfSOAyOTcN_Qv-n
n

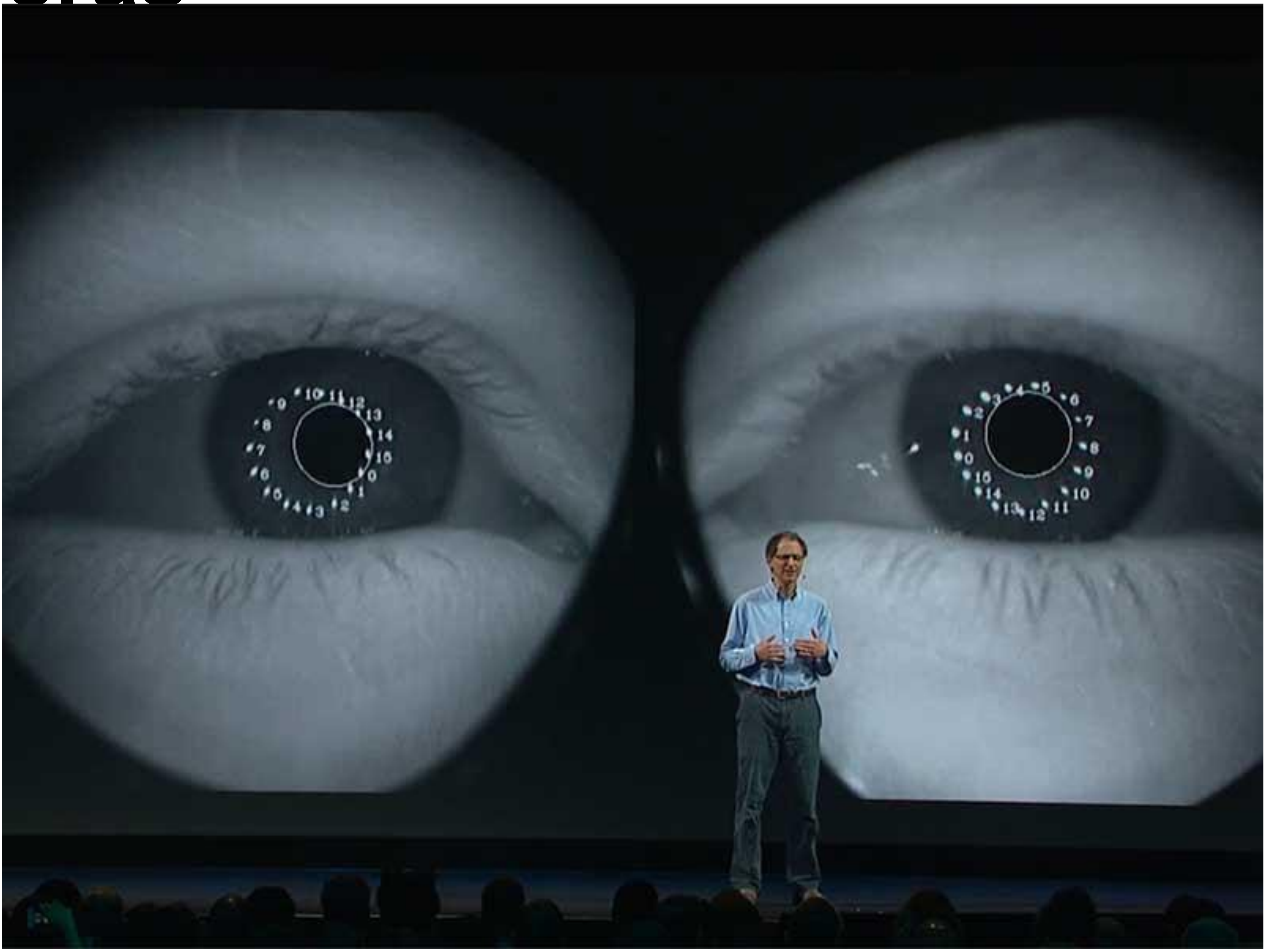
“Mixed Reality” - VR with Passthrough Imaging



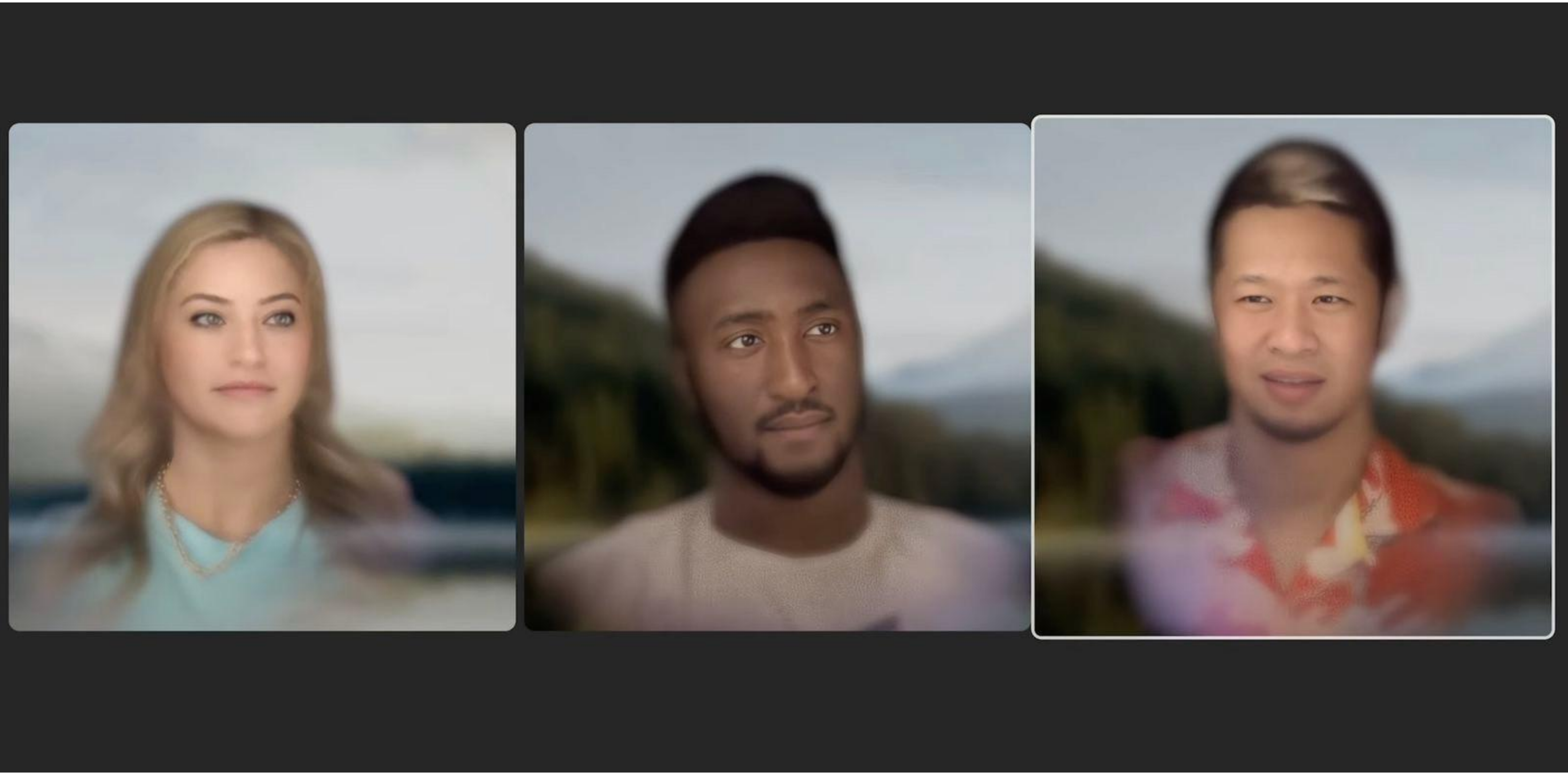
vive.com

Opaque VR headset, outward cameras, live imagery of real world + virtual 3D graphics

VR Headset Inward-Facing Cameras



VR Teleconference / Video Chat



Apple Vision Pro Personas

Image credit: Brian Tong

Things to Remember

VR presents many new graphics

challenges! Displays

- Head-pose tracking with high accuracy and low latency

Rendering

- Low-latency, high resolution & frame-rate, wide field of view, …

Imaging

- 360 spherical, stereo, light field

Acknowledgments

This slide set contain contributions from:

- Kayvon Fatahalian
- David Forsyth
- Pat Hanrahan
- Angjoo Kanazawa
- Steve Marschner
- Ren Ng
- James F. O’Brien
- Mark Pauly