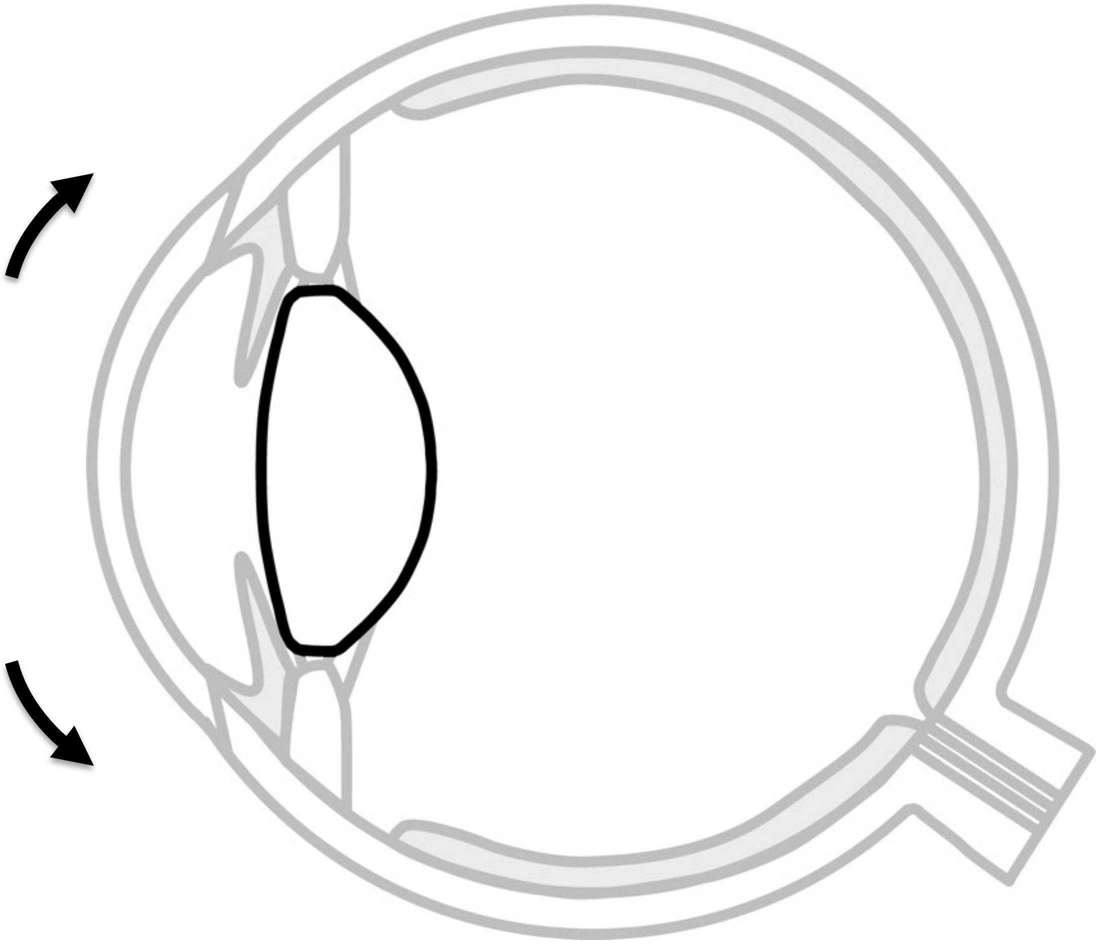
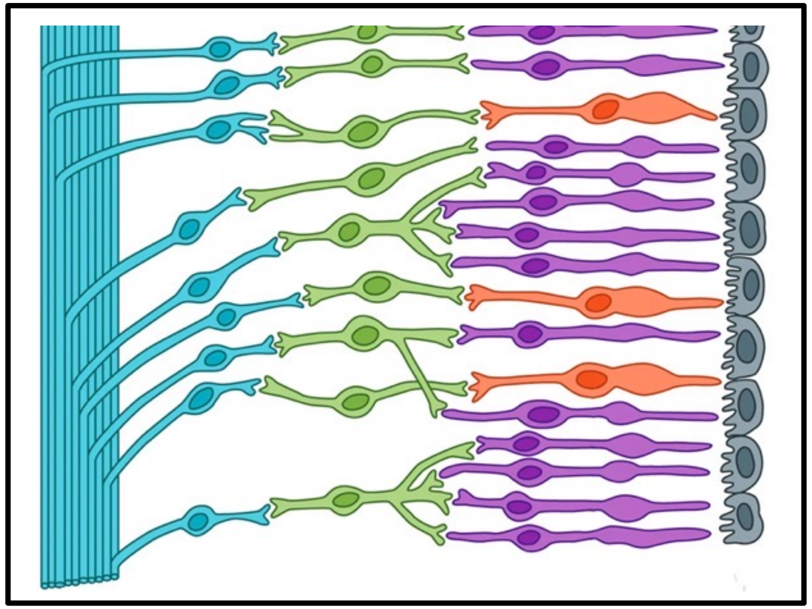
Beyond the Metaverse

Towards Human-centric XR

Gordon Wetzstein
Stanford University, Assoc. Professor
Raxium, Co-founder
Zinn Labs, Co-founder & Chief Scientist

SCI STANFORD COMPUTATIONAL IMAGING LAB
WWW.COMPUTATIONALIMAGING.ORG



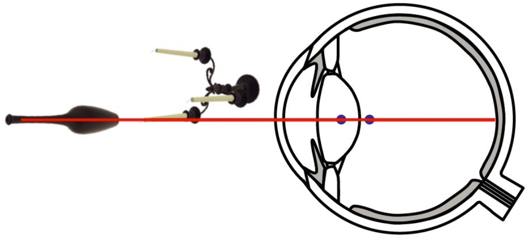


Perceptual realism
Visual comfort

Low-power & robust sensing

VR/AR Displays

Perceptually-driven Rendering

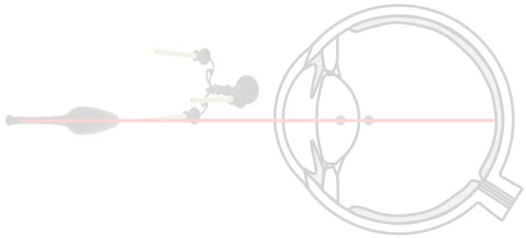


Event-based Eye Tracking

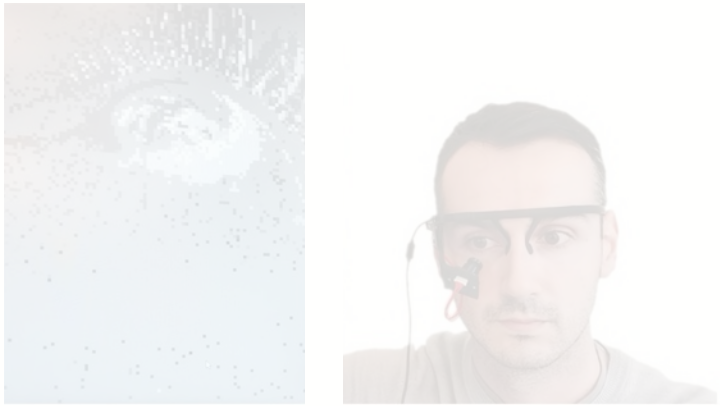
(Some) Emerging Technologies in XR

VR/AR Displays

Perceptually-driven Rendering



Event-based Eye Tracking

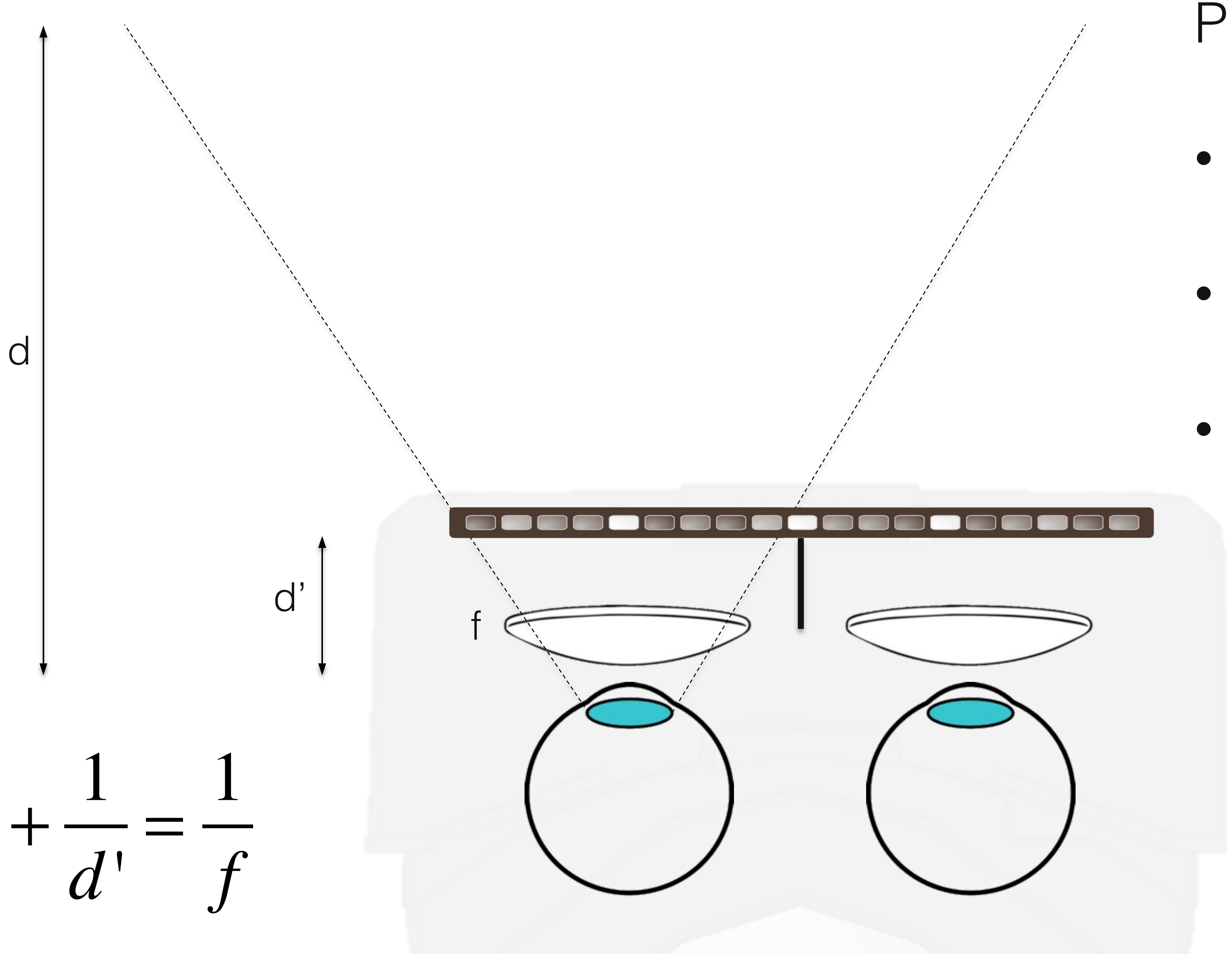


(Some) Emerging Technologies in XR

Sir Charles Wheatstone, 1838

NAE Grand Challenge!

Virtual Image



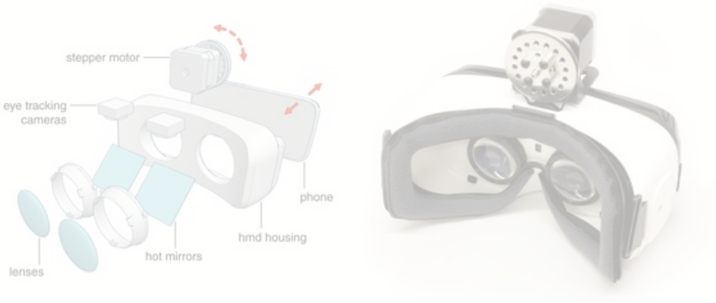
$$\frac{1}{d}+\frac{1}{d'}=\frac{1}{f}$$

Problems:

- fixed focal plane
- no focus cues ☹
- vergence-accommodation conflict (nausea)

Computational Near-eye Displays with Focus Cues

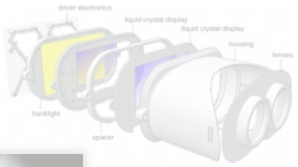
Gaze-contingent Varifocal Displays

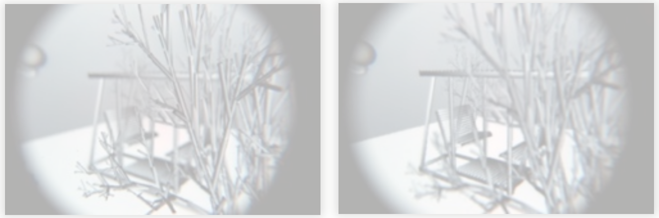


Konrad et al., SIGCHI 2016;
Padmanaban et al., SIGGRAPH 2016, PNAS 2017

Facebook/Oculus 2018

Near-eye Light Field Displays

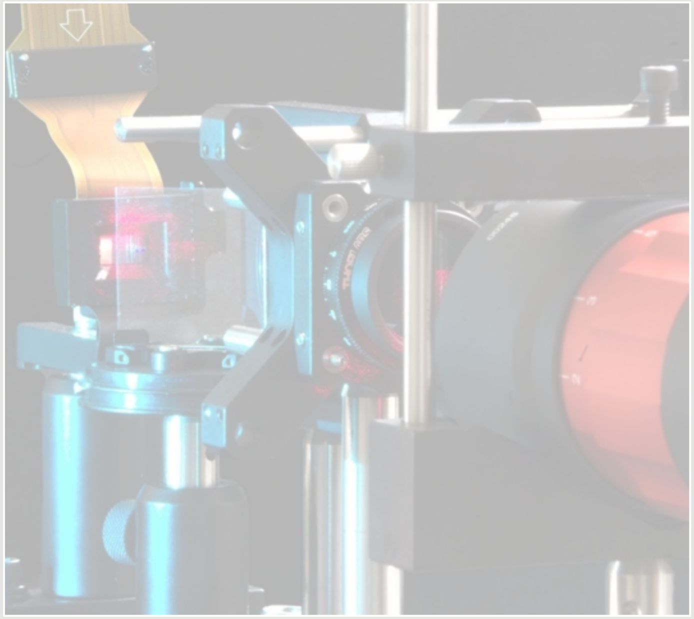




Huang et al., SIGGRAPH 2015;
Wetzstein et al., SIGGRAPH 2011, 2012

magic leap

Holographic Near-eye Displays

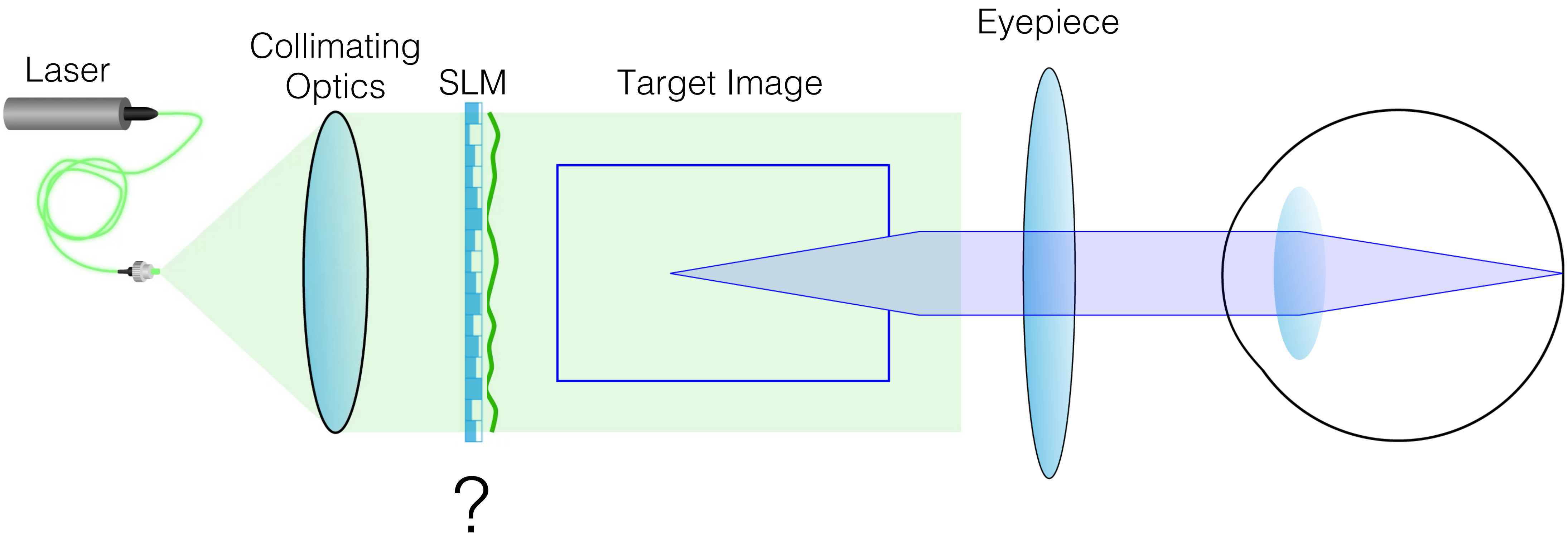


Padmanaban et al., SIGGRAPH Asia 2019;
Peng et al., SIGGRAPH Asia 2020;
Choi et al. SIGGRAPH Asia 2021; SIGGRAPH 2022, …

Holographic Displays

https://www.youtube.com/watch?v=sQS6_DbYIAw

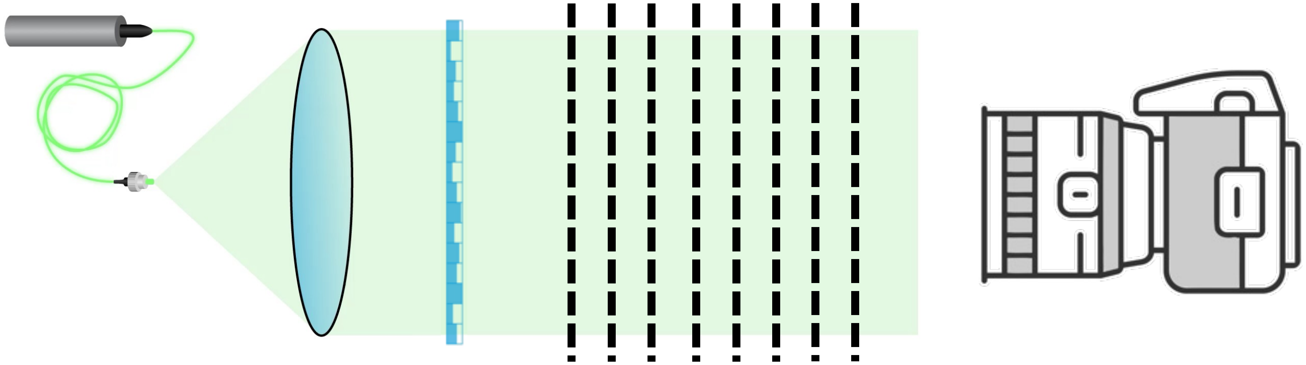
Holographic Near-eye Displays

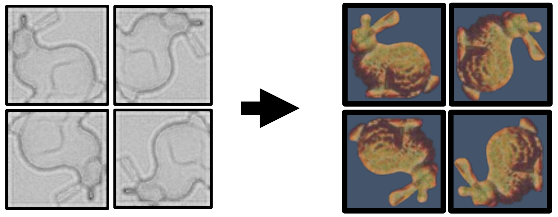


Challenge: low image quality due to mismatch between physical optics and simulation

Neural Holography

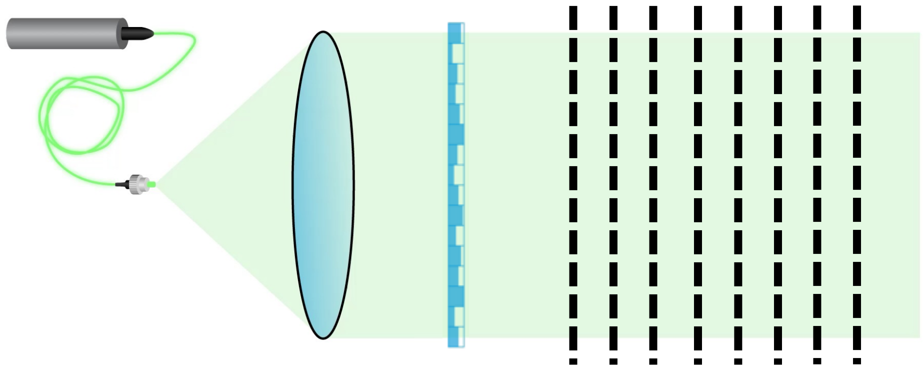
Physical Optics



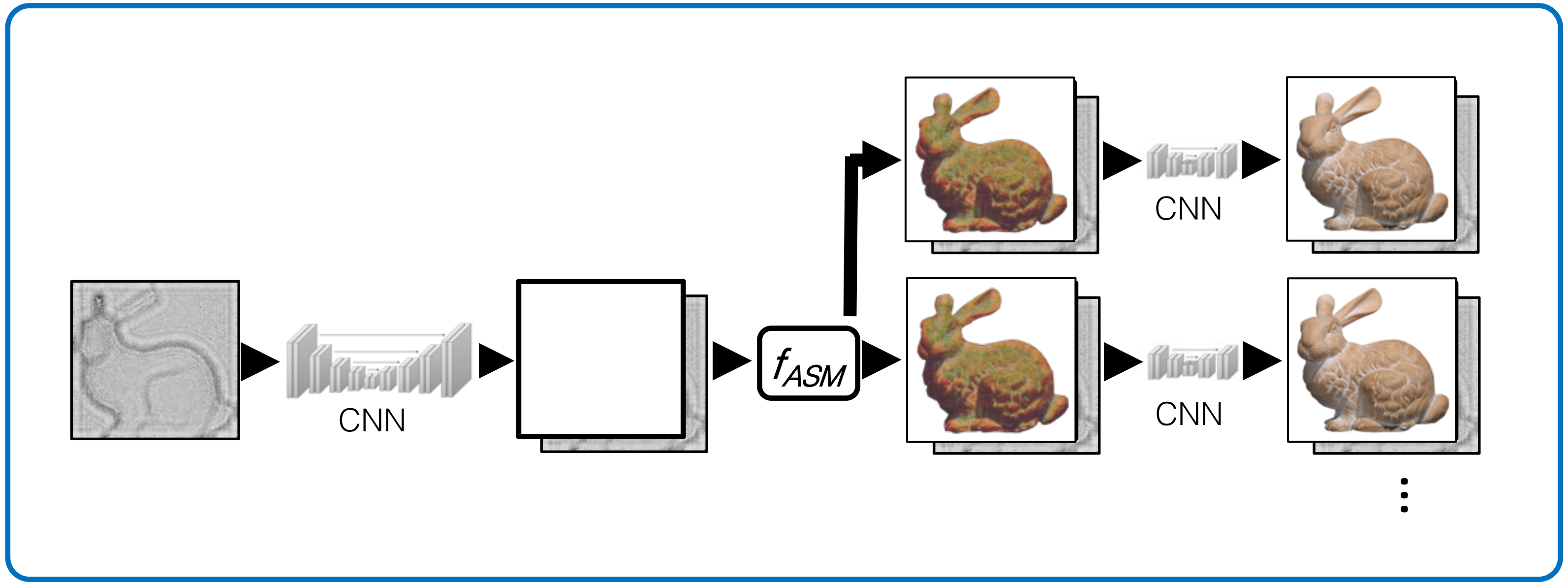


[Peng et al., SIGGRAPH Asia 2020; Choi et al., SIGGRAPH ASIA 2021; SIGGRAPH 2022; …]

Neural Holography



Physical Optics

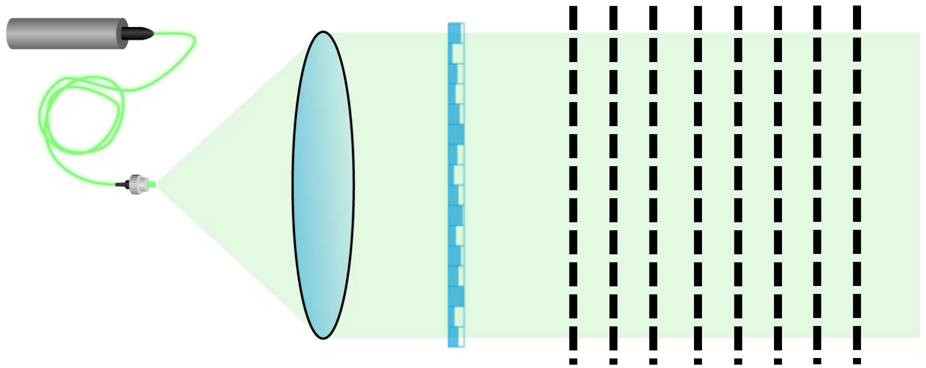


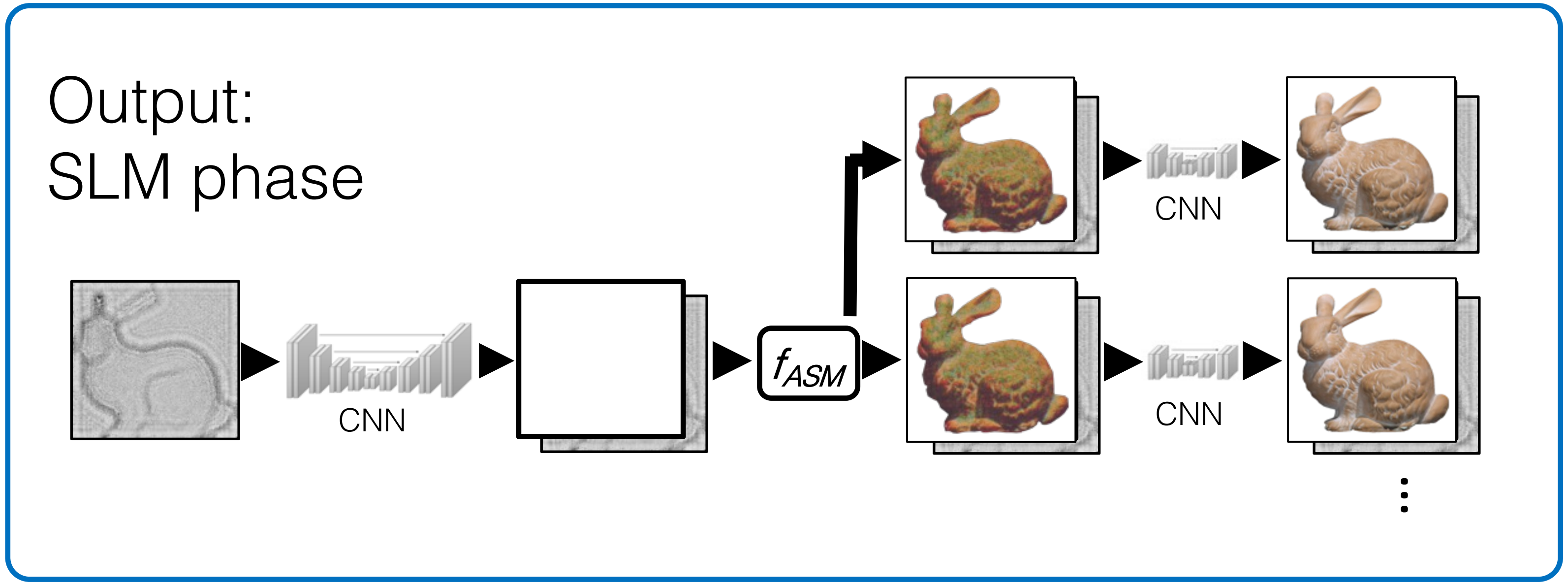
Camera-calibrated Wave Propagation Model

[Peng et al., SIGGRAPH Asia 2020; Choi et al., SIGGRAPH ASIA 2021; SIGGRAPH 2022; …]

Neural Holography

Physical Optics



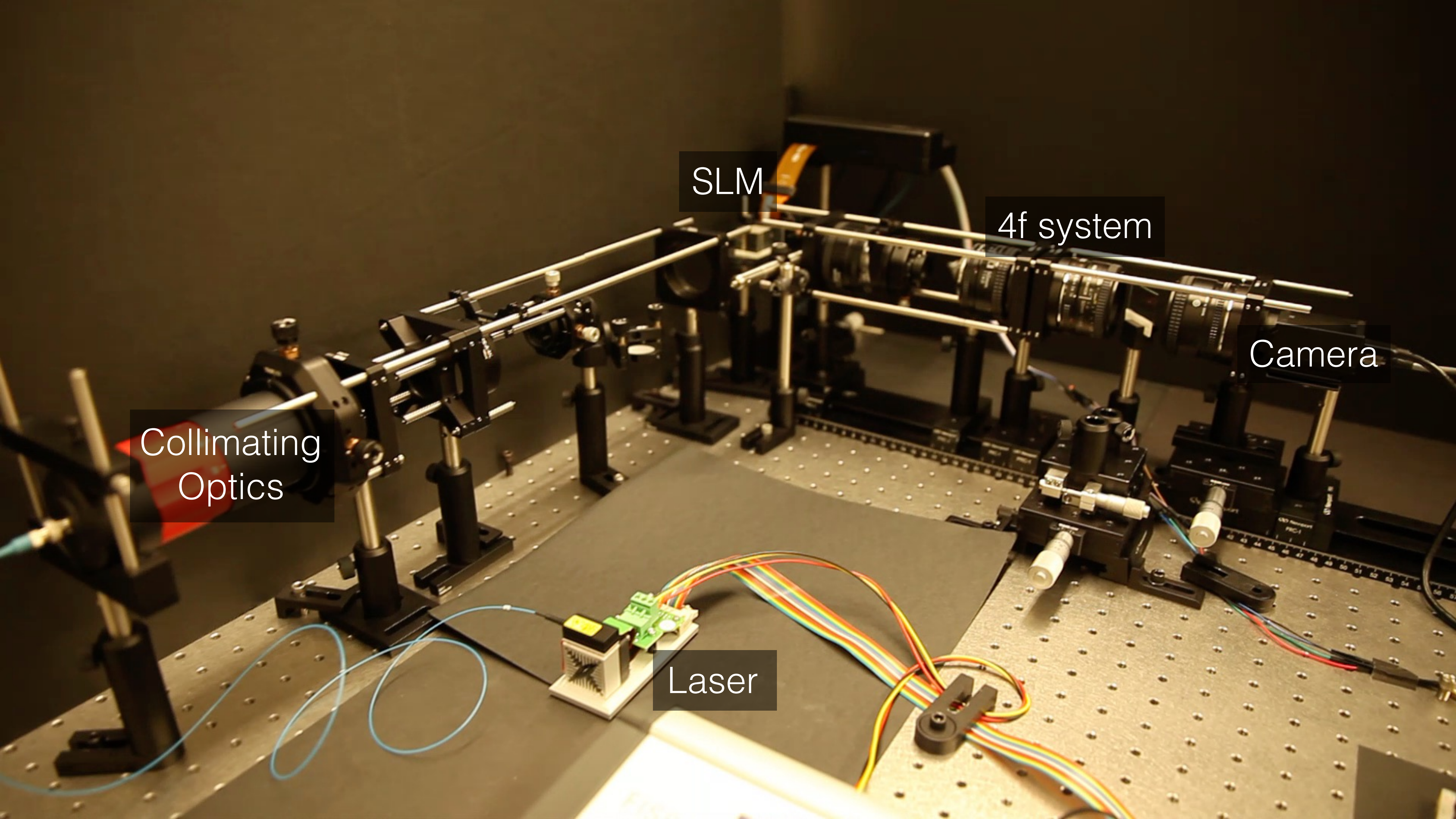


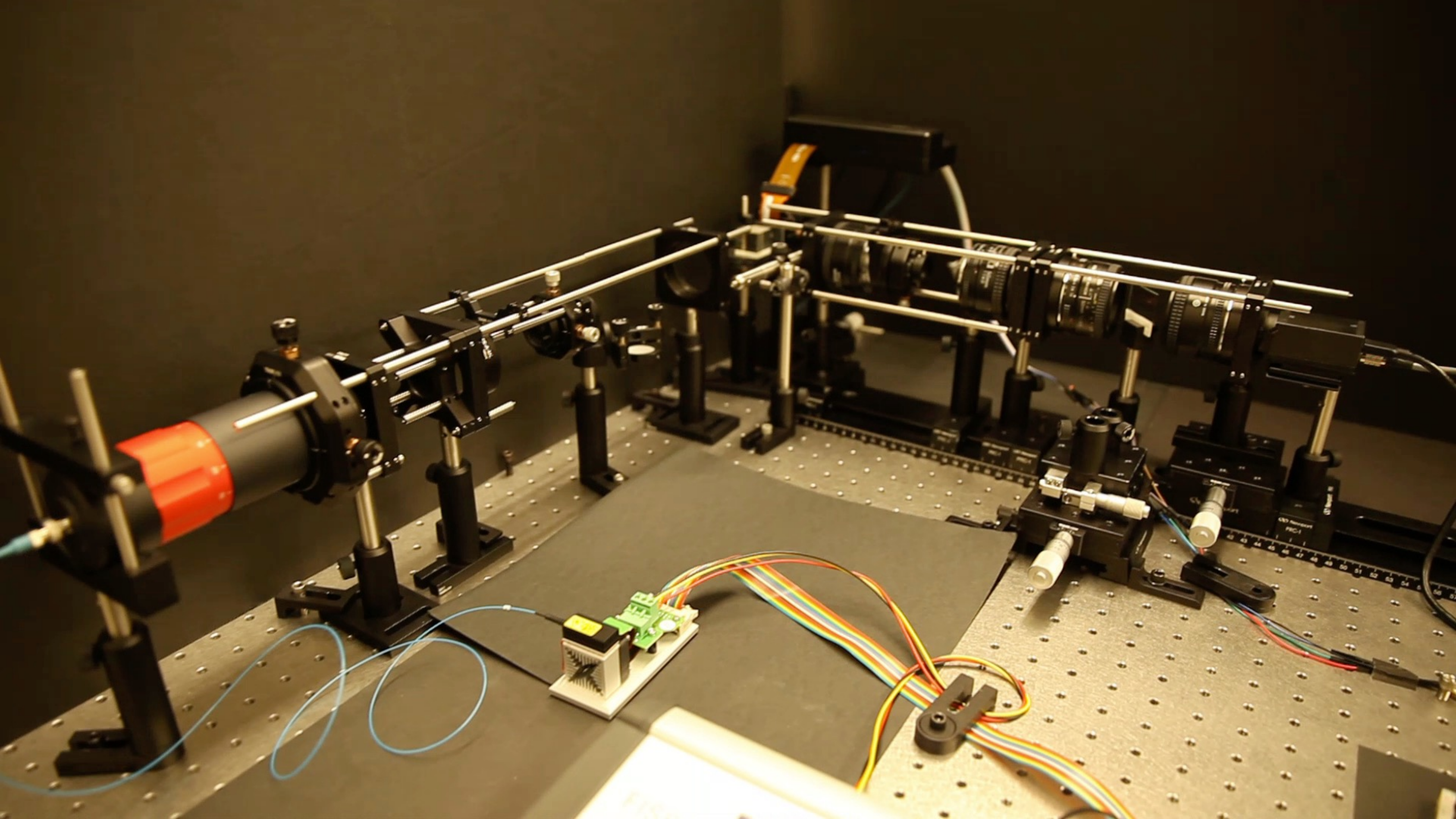
Output:
SLM phase

Camera-calibrated Wave Propagation Model

Input:
2D, 2.5D RGBD, 3D focal stack, 4D light field, …

[Peng et al., SIGGRAPH Asia 2020; Choi et al., SIGGRAPH ASIA 2021; SIGGRAPH 2022; …]


SLM
4f system
Camera
Collimating Optics
Laser

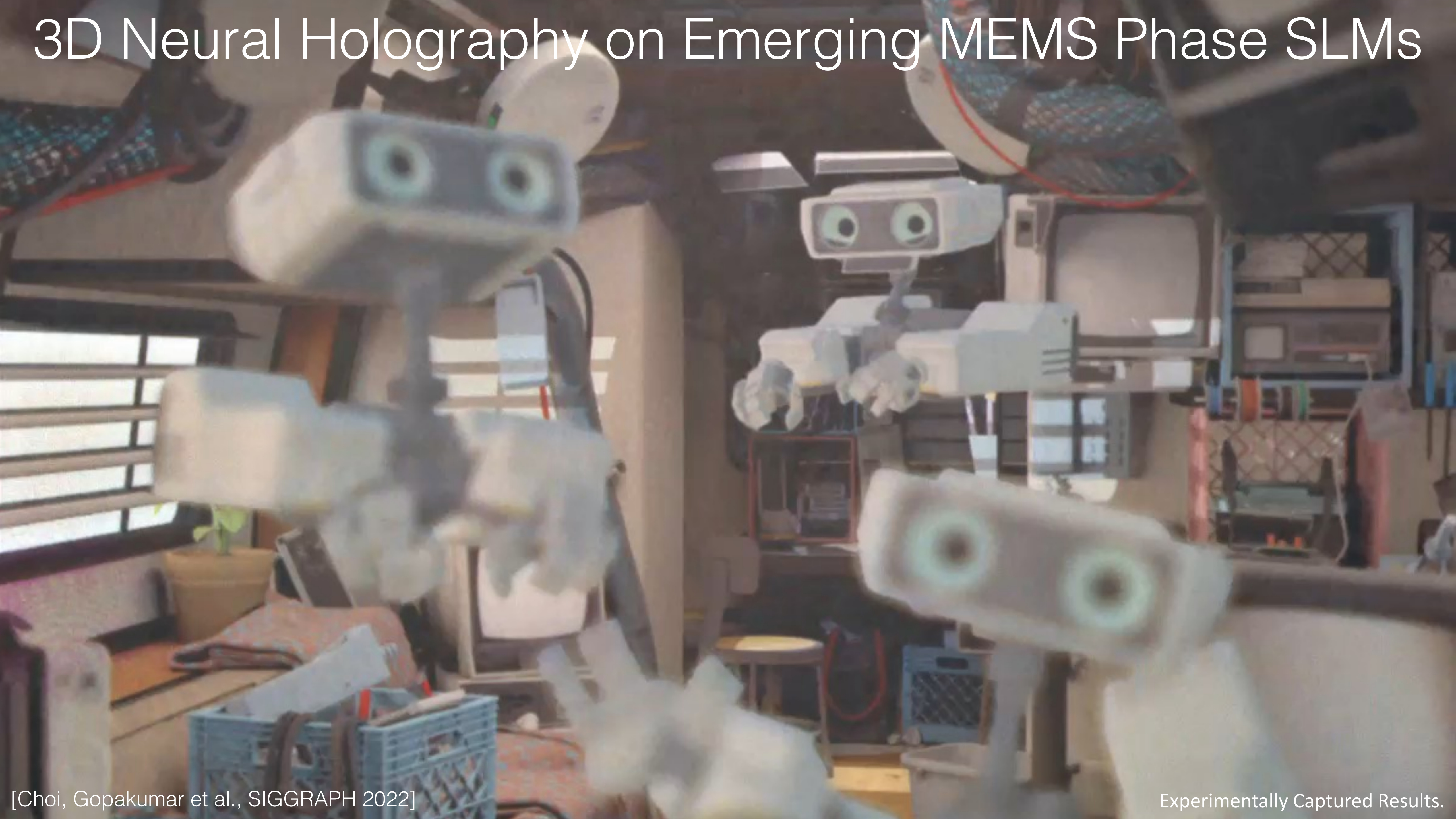

PRC-1

Gerchberg–Saxton

Neural Holography 2020 Results

[Peng et al., SIGGRAPH & SIGGRAPH Asia 2020]

3D Neural Holography on Emerging MEMS Phase SLMs



[Choi, Gopakumar et al., SIGGRAPH 2022]

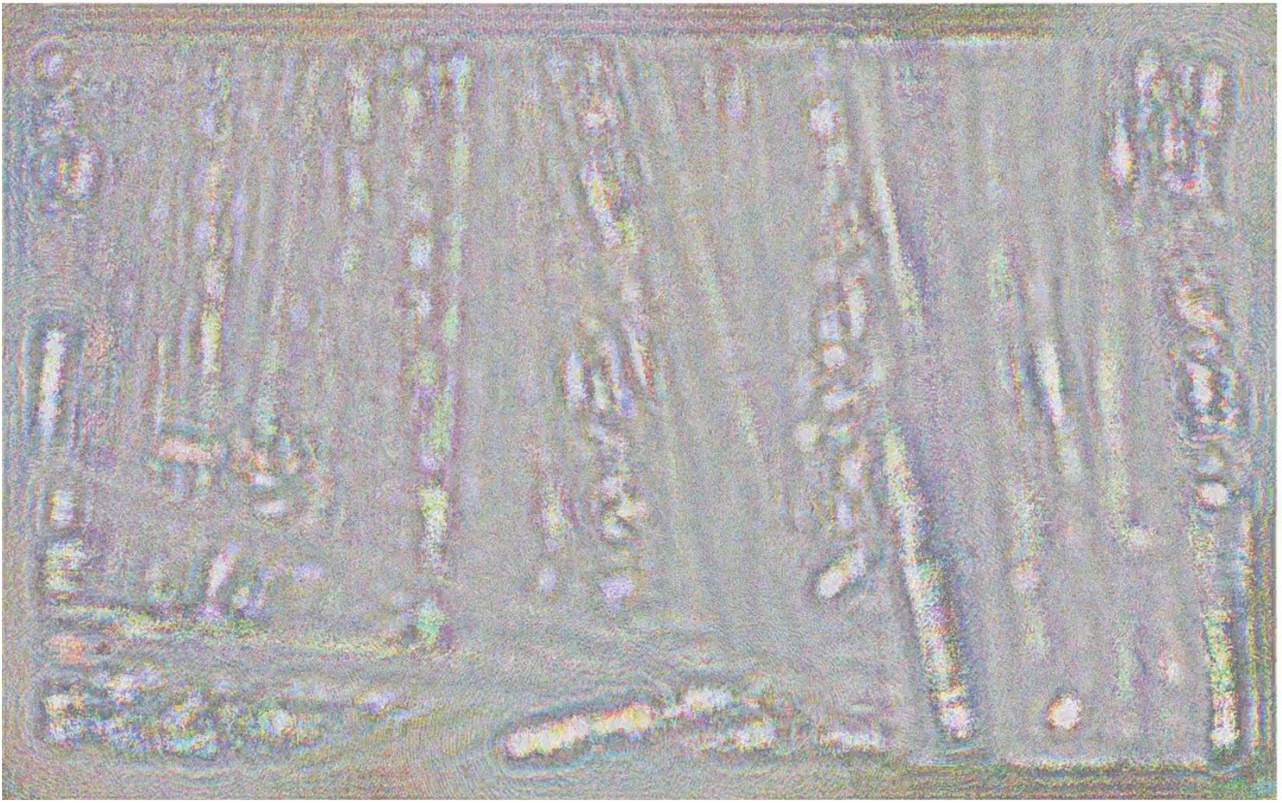
Experimentally Captured Results.

3D Neural Holography on Emerging MEMS Phase SLMs

[Choi, Gopakumar et al., SIGGRAPH 2022]

Experimentally Captured Results.

3D Neural Holography on Emerging MEMS Phase SLMs



Displayed patterns on phase SLM

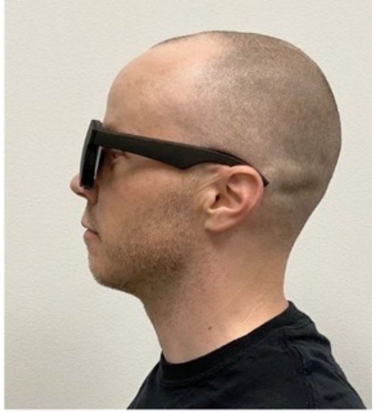
Holograms captured with our prototype

[Choi, Gopakumar et al., SIGGRAPH 2022]

Experimentally Captured Results.

Additional Benefits of Holographic Near-eye Displays

Thin VR Display Form Factors



Maimone et al., SIGGRAPH 2020

Kim et al., SIGGRAPH 2022

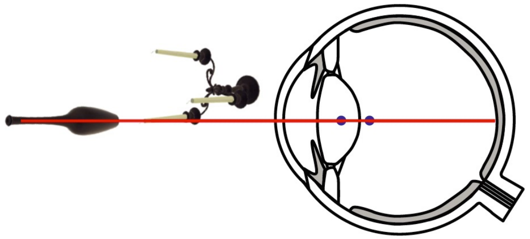
Other:

- Light-efficient AR Displays
- Prescription correction (including astigmatism and higher-orders)
- Correcting optical aberrations

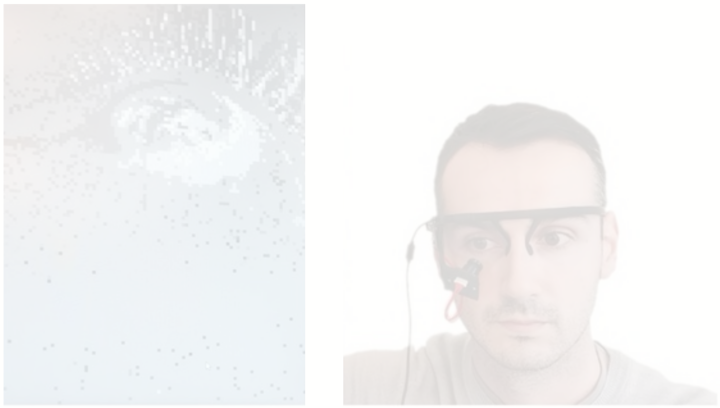
…

VR/AR Displays

Perceptually-driven Rendering

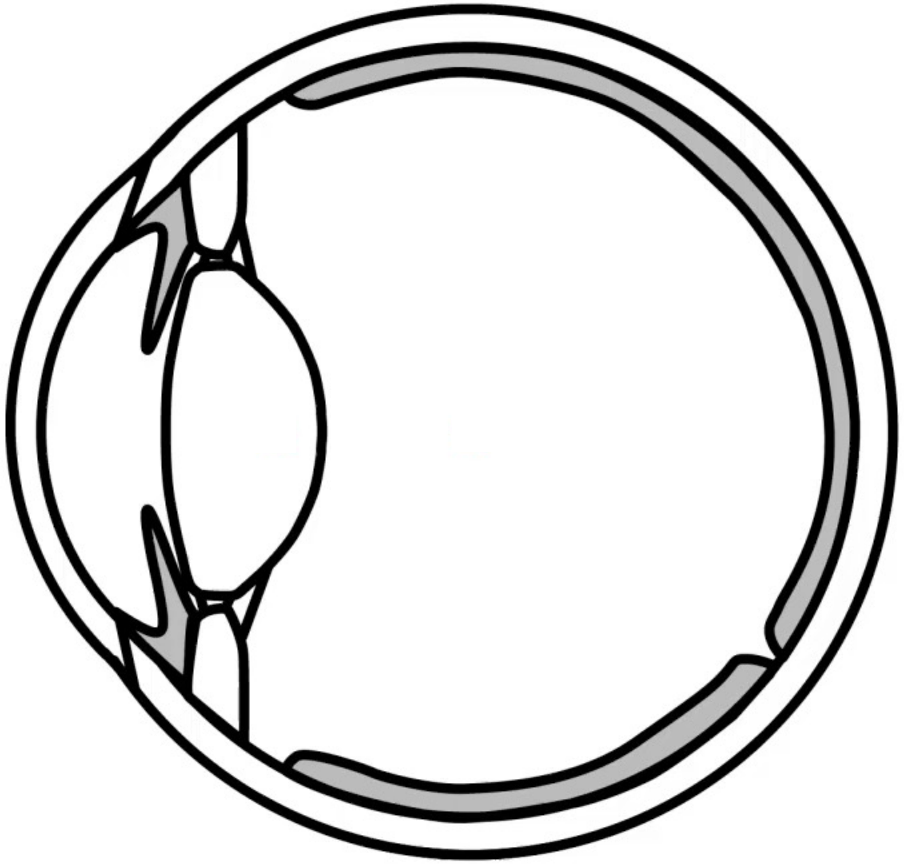


Event-based Eye Tracking

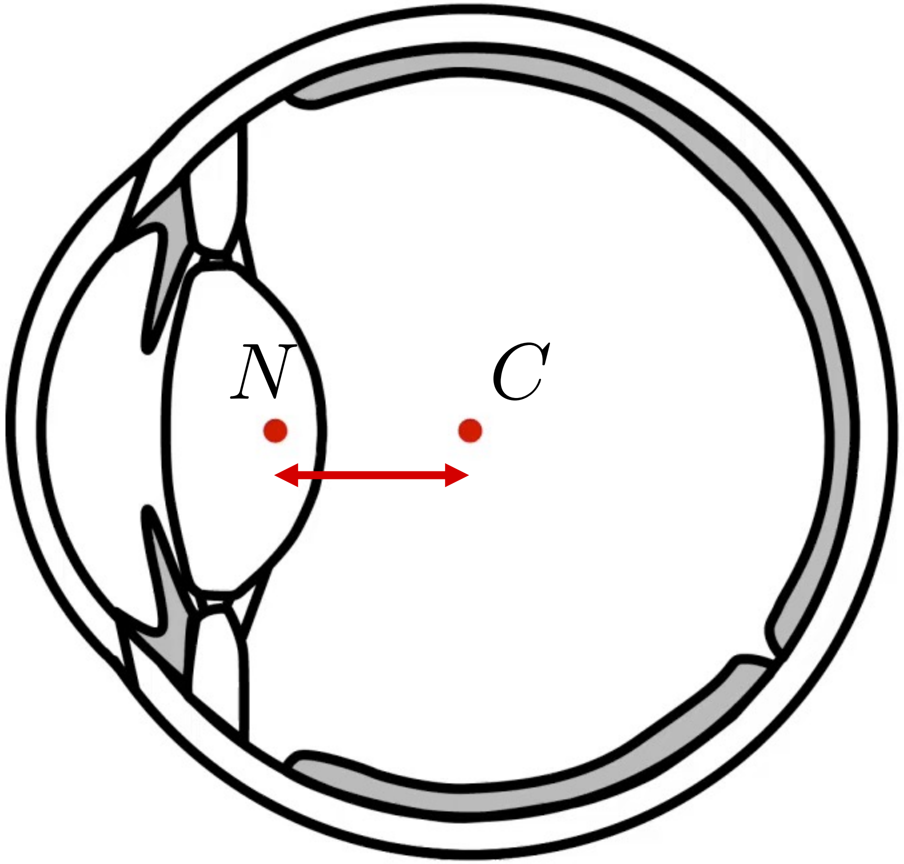


(Some) Emerging Technologies in XR

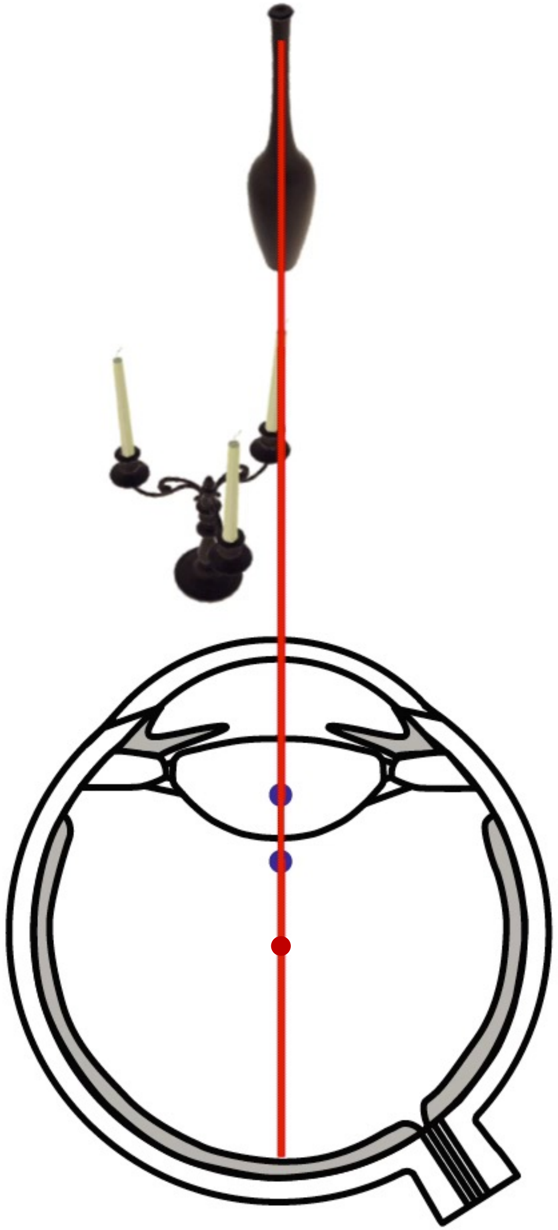
The Human Eye



The Human Eye

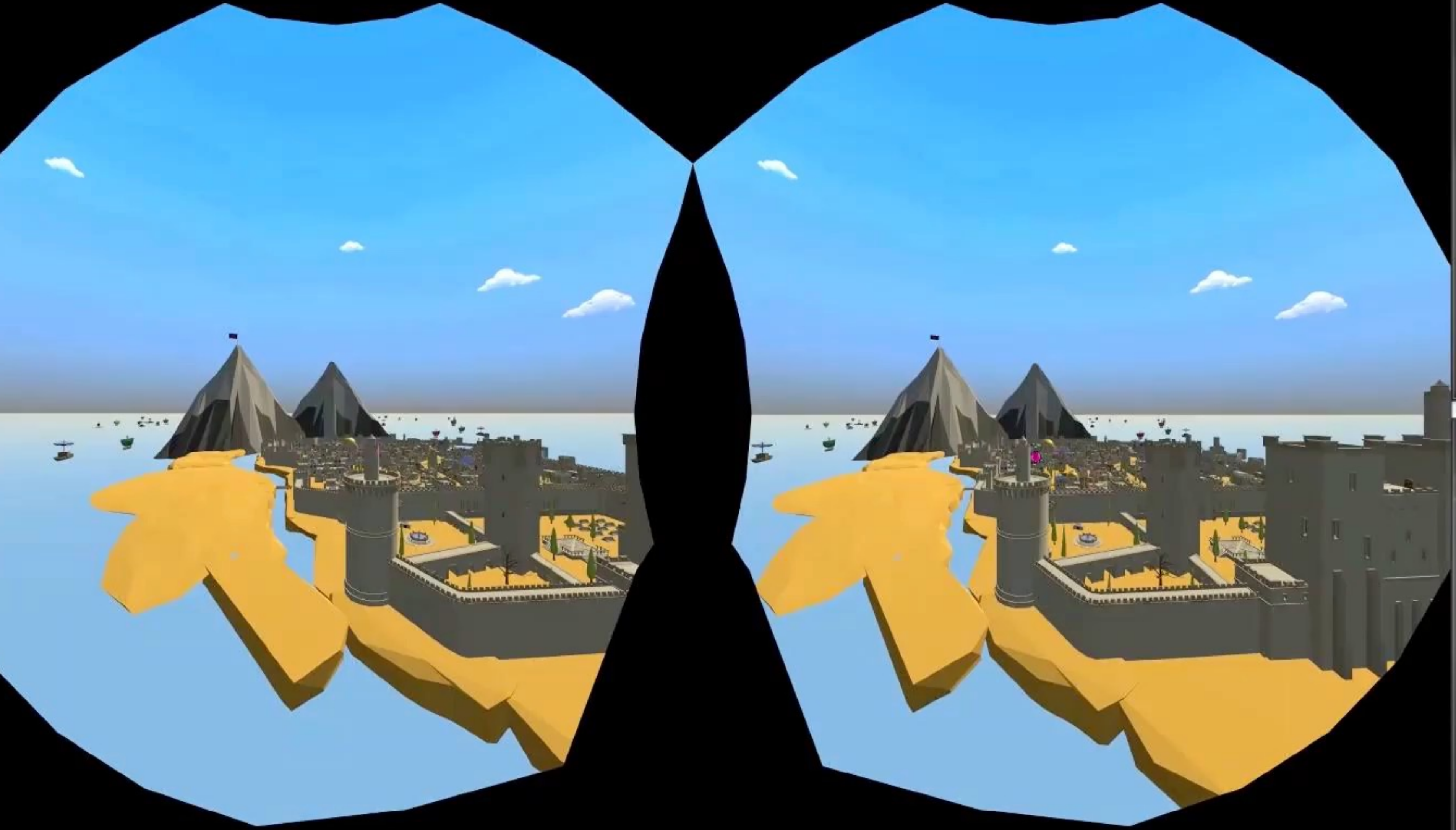


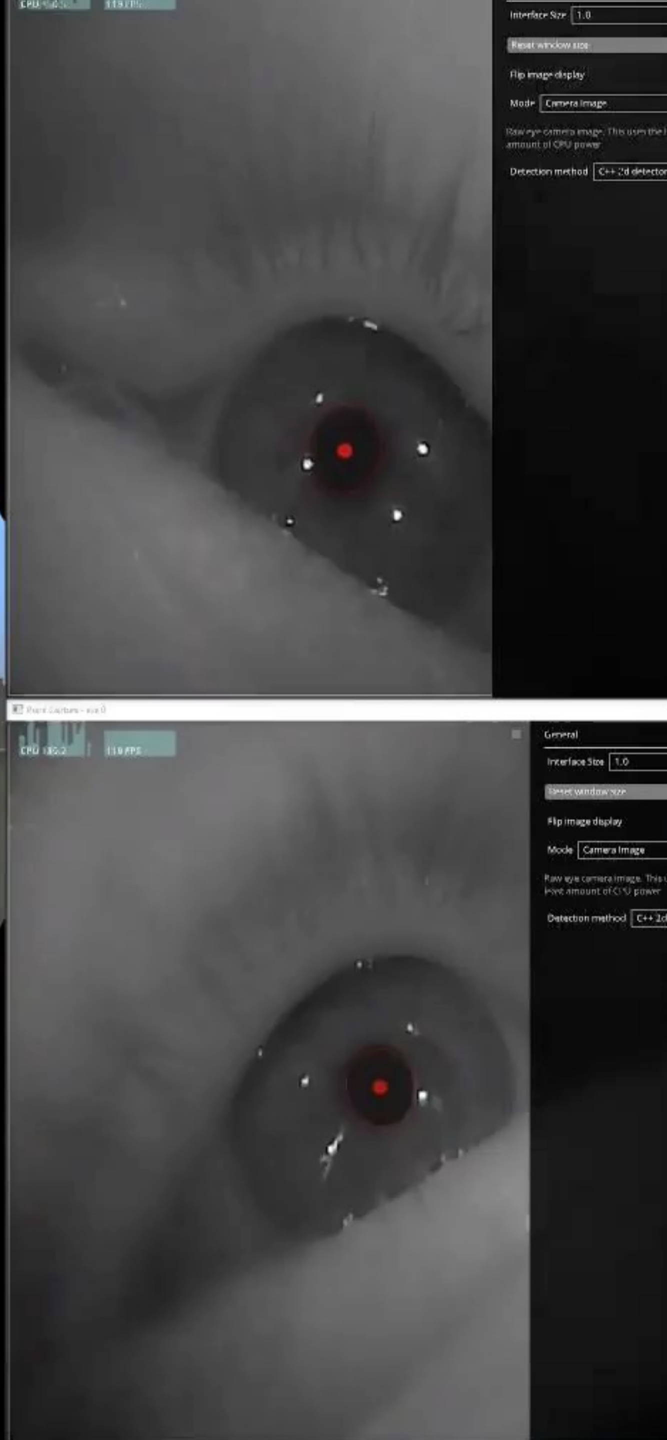
Ocular Parallax



Parallax effect enhanced for visualization

Ocular Parallax Disabled





Interface Size 1.0

Reset window size

Flip image display

Mode Camera image

Raw eye camera image. This uses the
amount of CPU power

Detection method C++ 2d detector

General

Interface Size 1.0

Reset window size

Flip image display

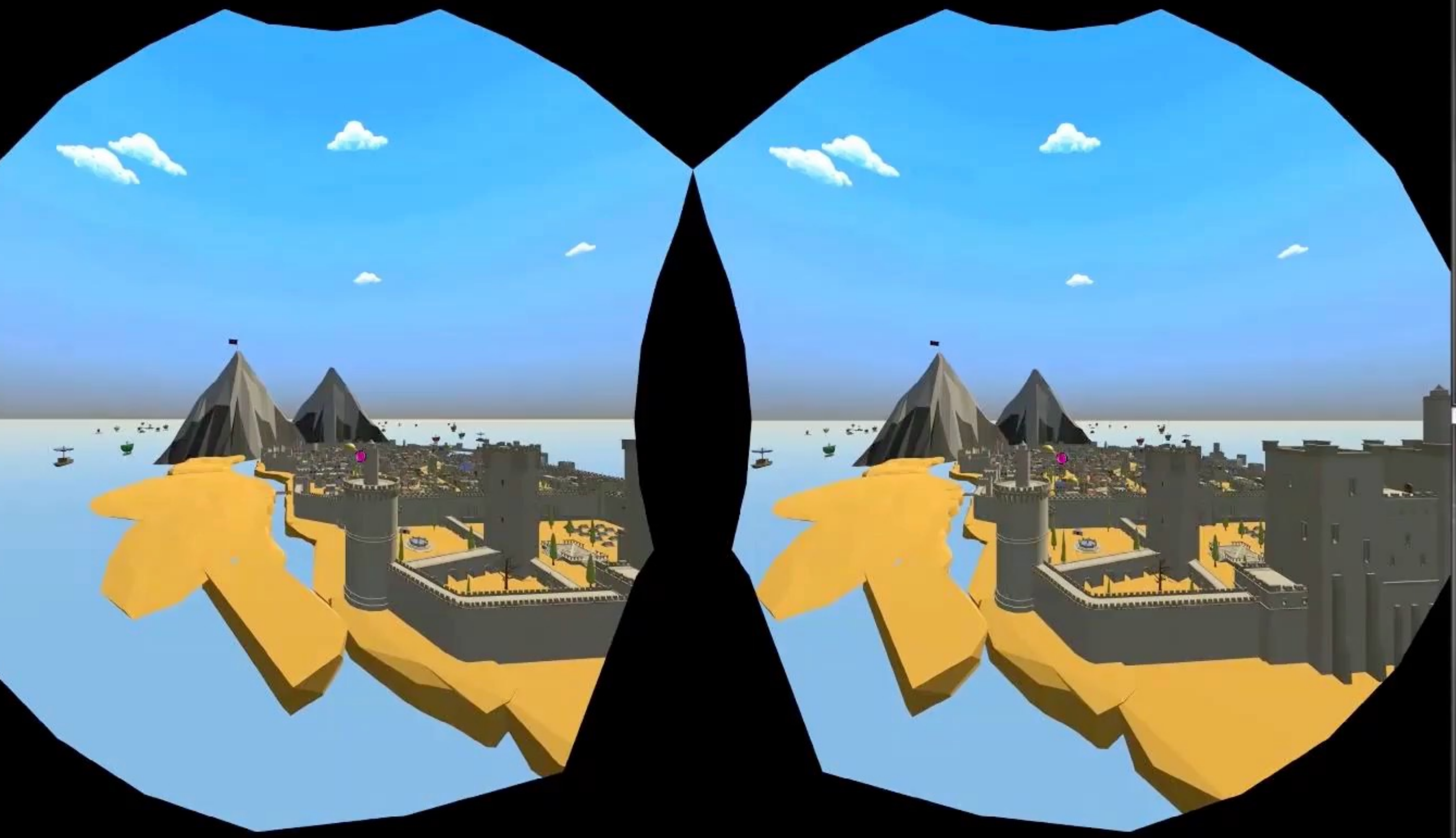
Mode Camera Image

Raw eye camera image. This
least amount of CPU power

Detection method C++ 2d

Konrad et al., ACM TOG 2020

Ocular Parallax Enabled



Interface Size 1.0

Mode Camera Image

Raw eye camera image. This uses the l
amount of CPU power

Detection method C++ 2d detector

General

Interface Size 1.0

Mode Camera Image

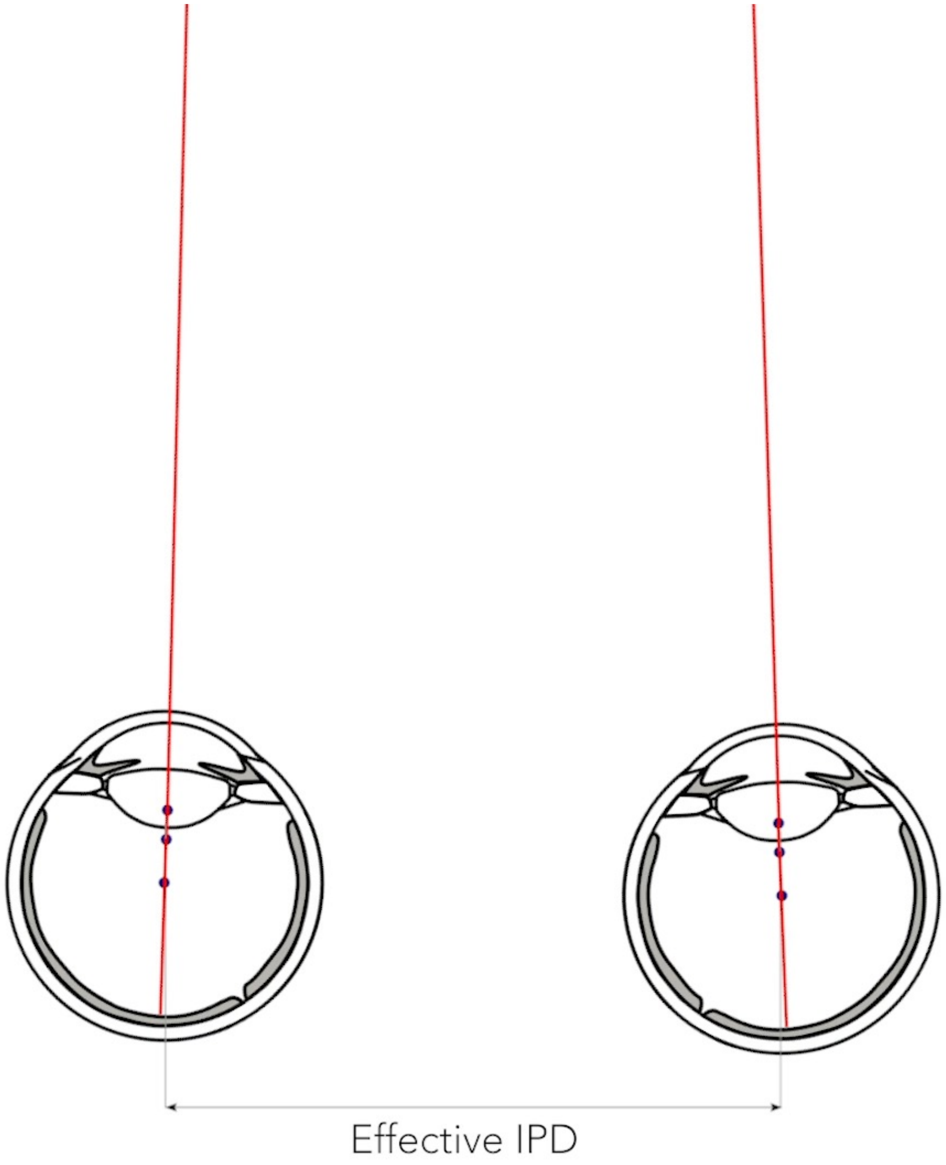
Detection method C++ 2d

Konrad et al., ACM TOG 2020

Summary of Insights

Is ocular parallax visible in VR?	Yes! Same effect size as retinal blur!
Does ocular parallax increase perceptual realism?	Yes!

Ocular Parallax Affects IPD!

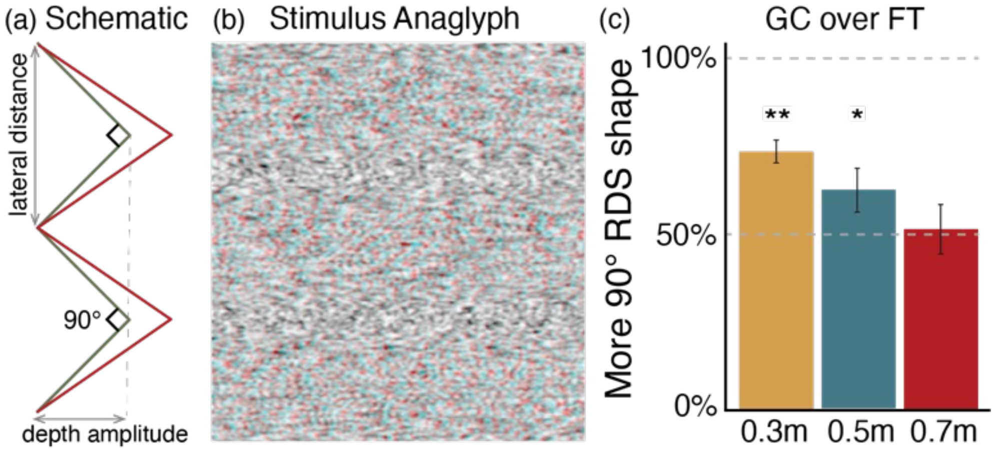


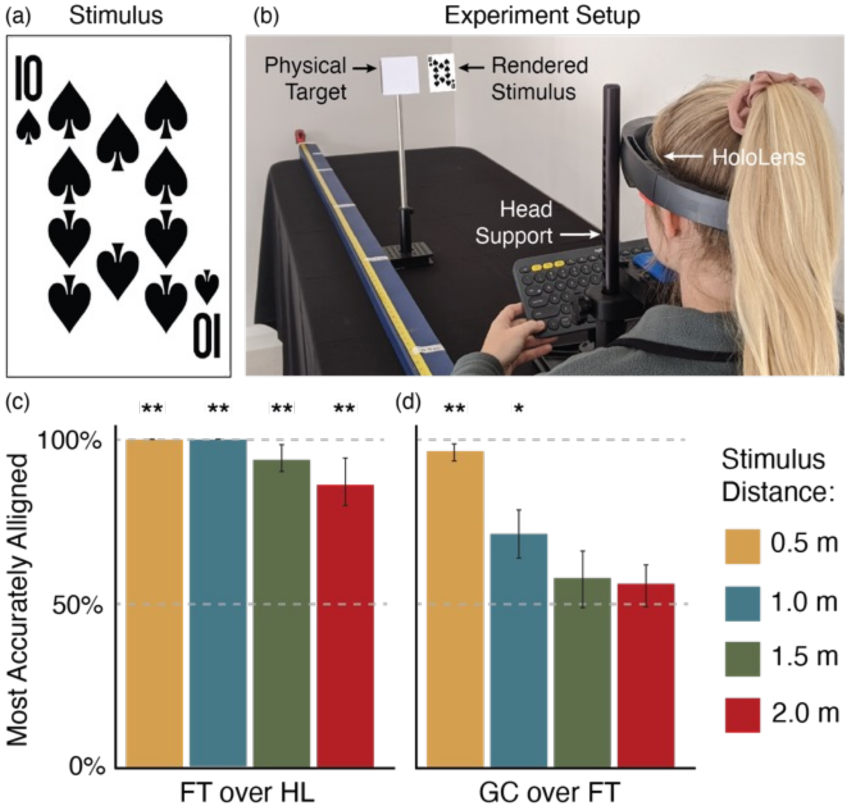
Krajancich et al., ACM SIGGRAPH Asia 2020

Conventional Stereo Rendering Distorts Disparities!

Krajancich et al., ACM SIGGRAPH Asia 2020

Gaze-contingent Stereo Rendering





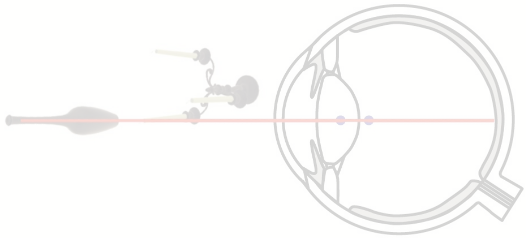
GC stereo improves depth perception in VR and optical see-through AR!

FT … "fine-tuned" (user-adjusted) IPD
HL … "Hololens" built-in stereo rendering
GC … gaze-contingent IPD

Krajancich et al., ACM SIGGRAPH Asia 2020

VR/AR Displays

Perceptually-driven Rendering



Event-based Eye Tracking

(Some) Emerging Technologies in XR

Eye Tracking

Near-eye Systems:

- Tobii, Pupil Labs, …
- 120-200 Hz
- Accuracy: 0.5-1°

Pupil Labs

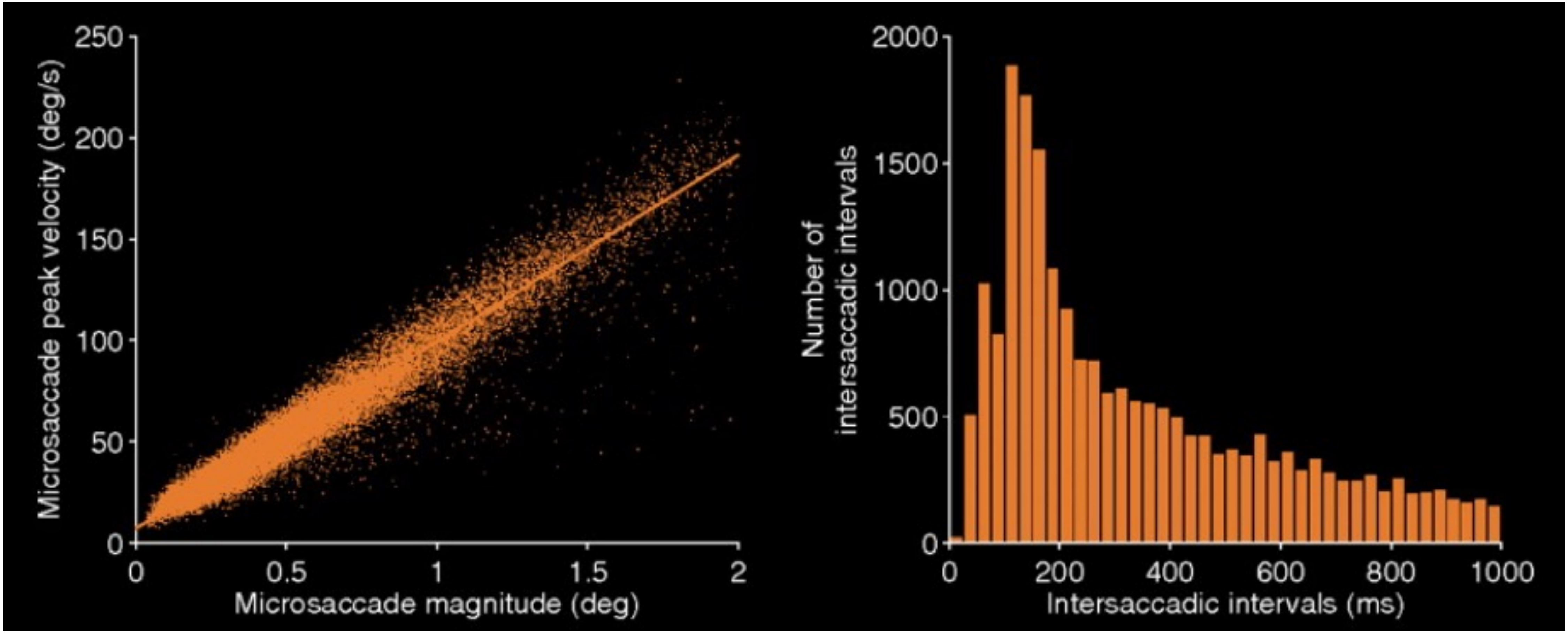
Desk-mounted Systems

- EyeLink, …
- 1,000 Hz
- Accuracy: ~0.5°

EyeLink 1000

Fast tracking is needed...

for capturing extremely fast eye movements, like microsaccades, and
for new technologies in VR/AR, like foveated rendering

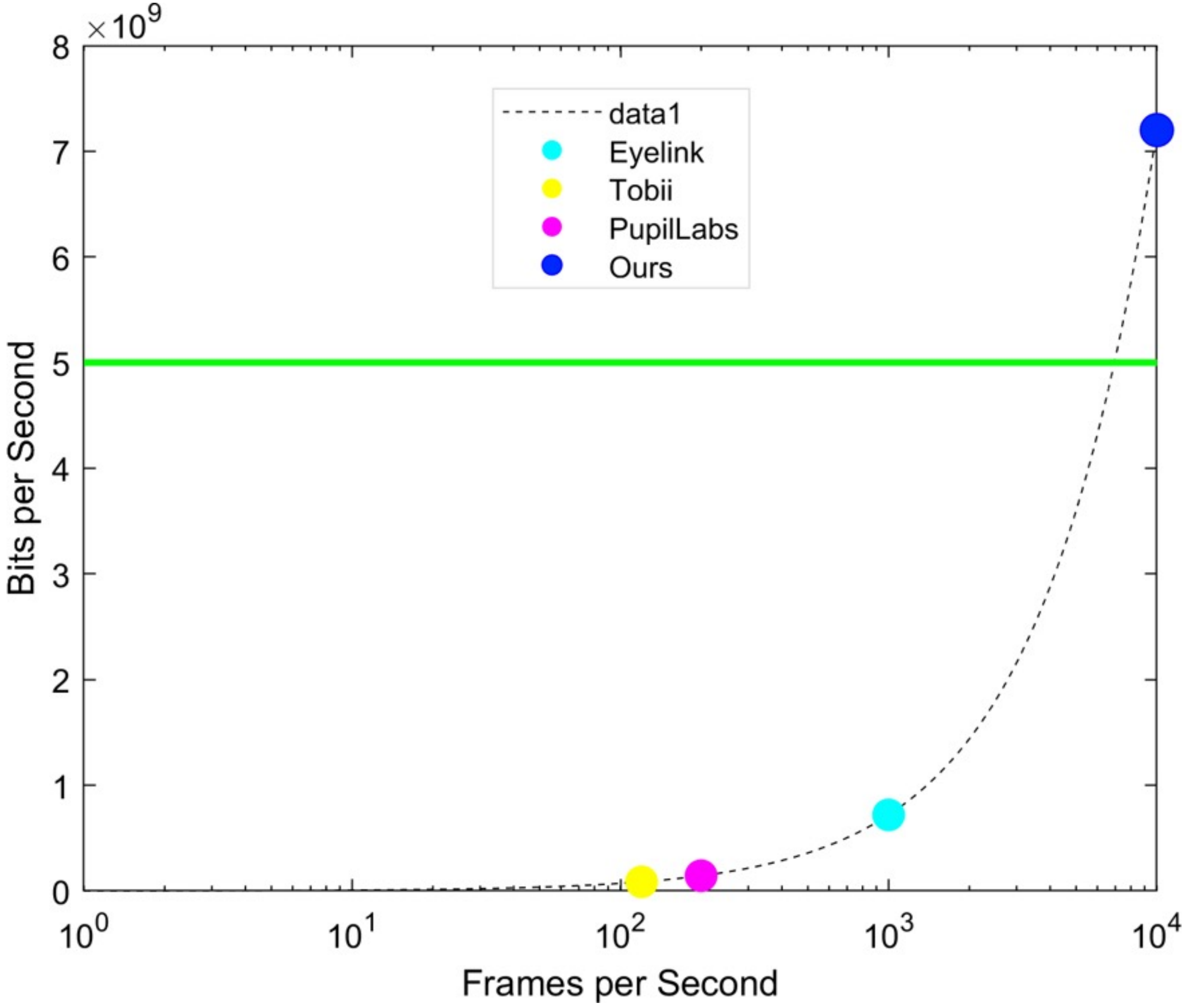


from psychophysics, miccrosaccades reach up to 200°/s

… but consumes too much bandwidth

- Imagine a 10,000 FPS system
- Each frame is 300x300
- Each pixel is 8 bits

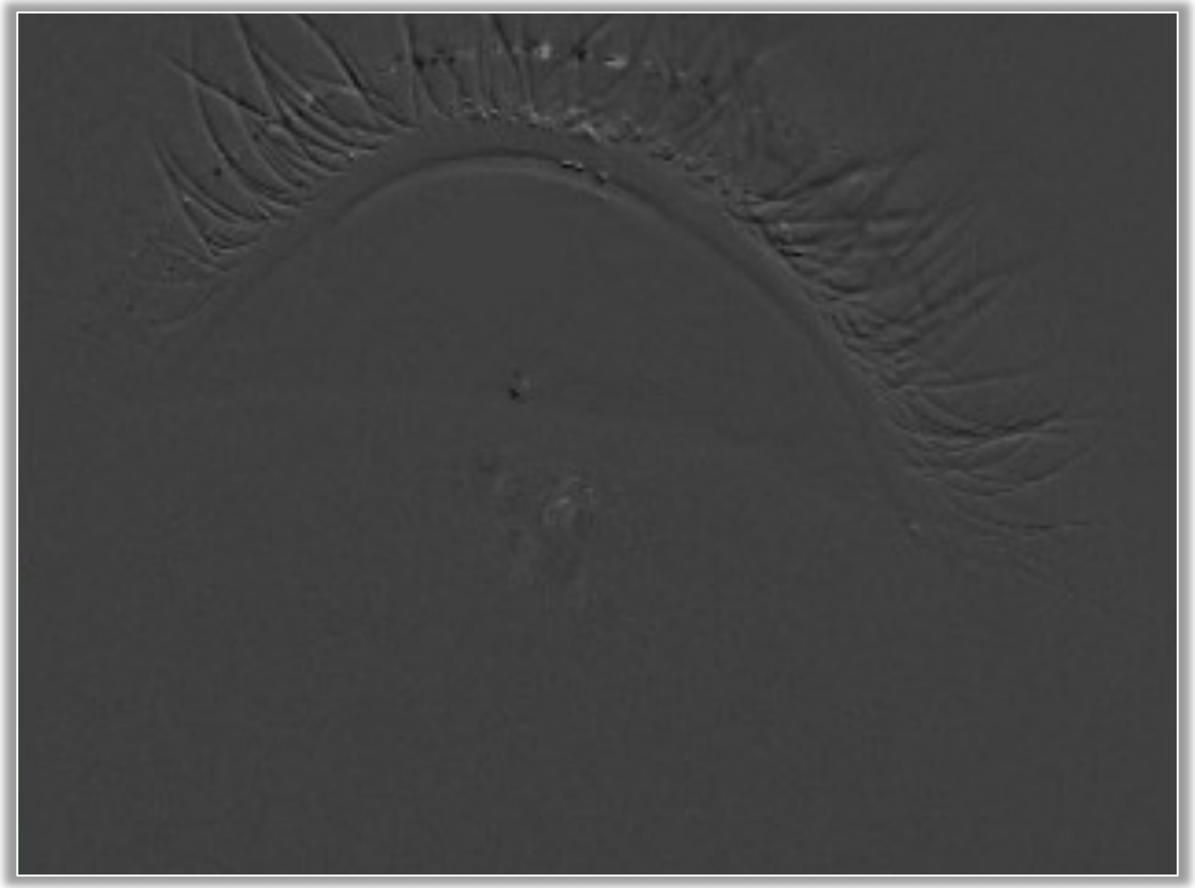
Bandwidth is 10,000 x 300 x 300 x 8 =
7.2 Gbps!



But near-eye gaze tracking is sparse!

An intensity image

$$I(x, y, t)$$



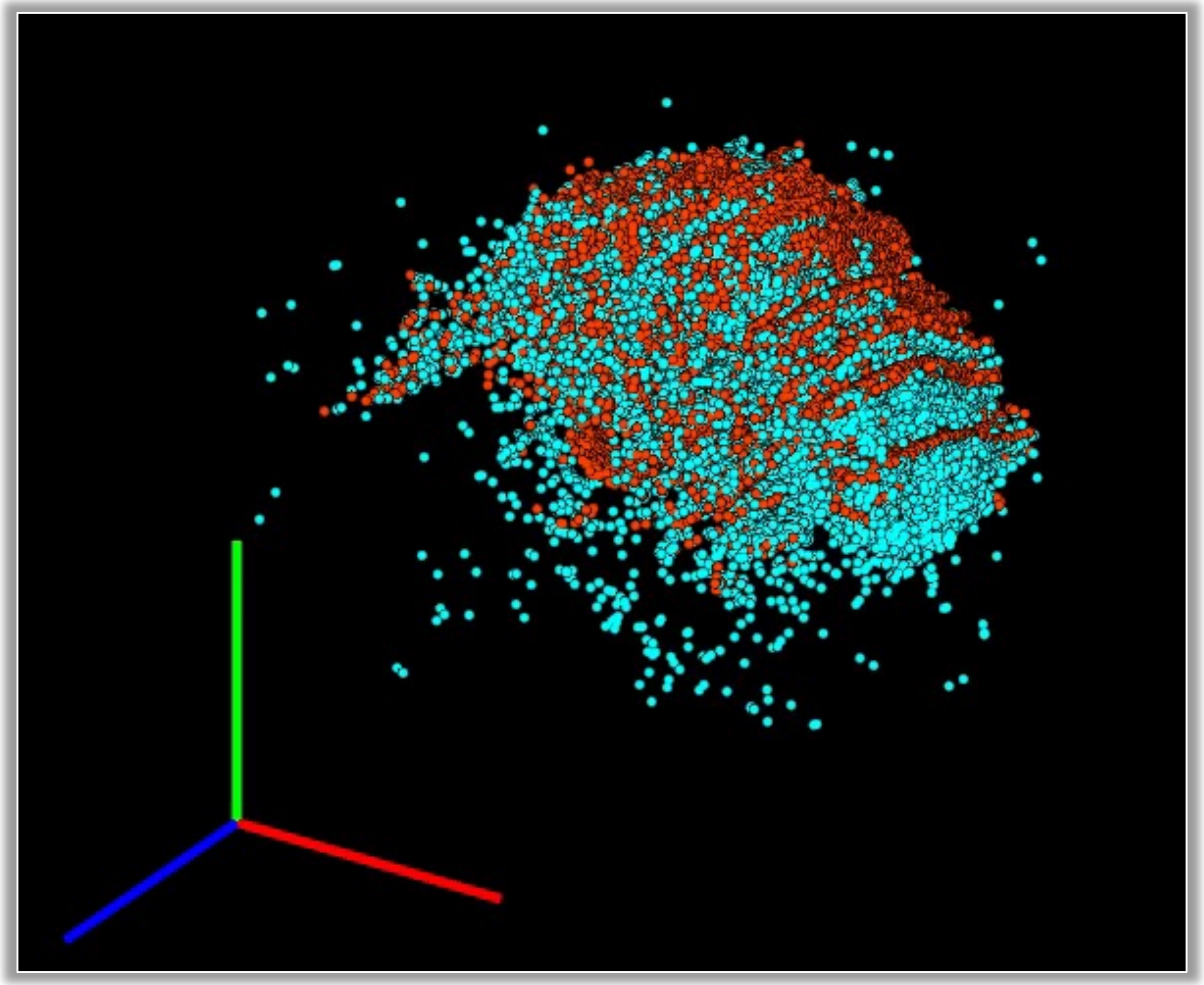
Difference between two consecutive images

$$\frac{\partial}{\partial t} I(x, y, t)$$

Event-based sensors trigger events (at high speed) when the brightness changes

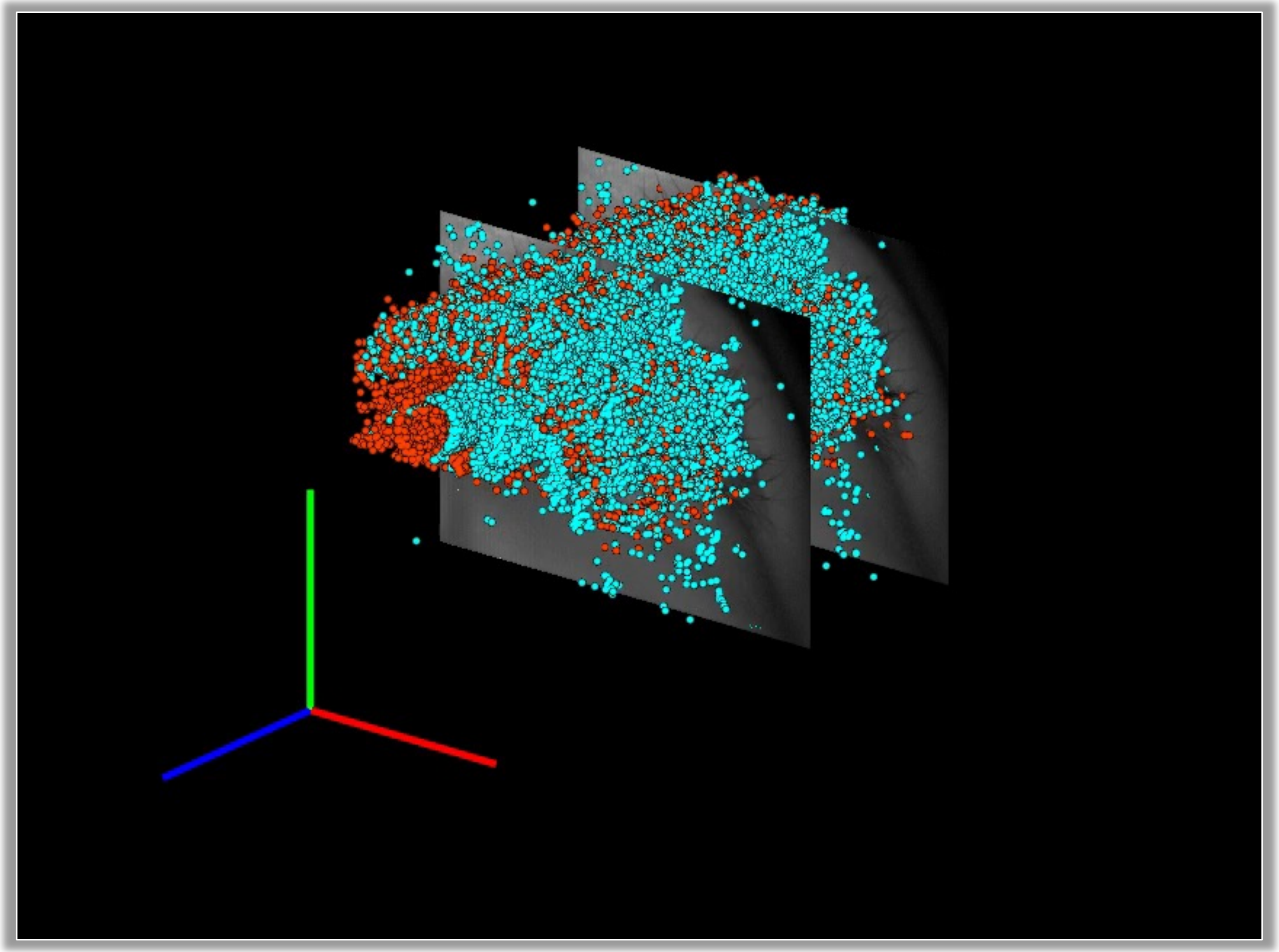
Davis 346B

outputs



Angelopolous, Martel et al., IEEE VR 2021

We use a sensor outputting both frames and events

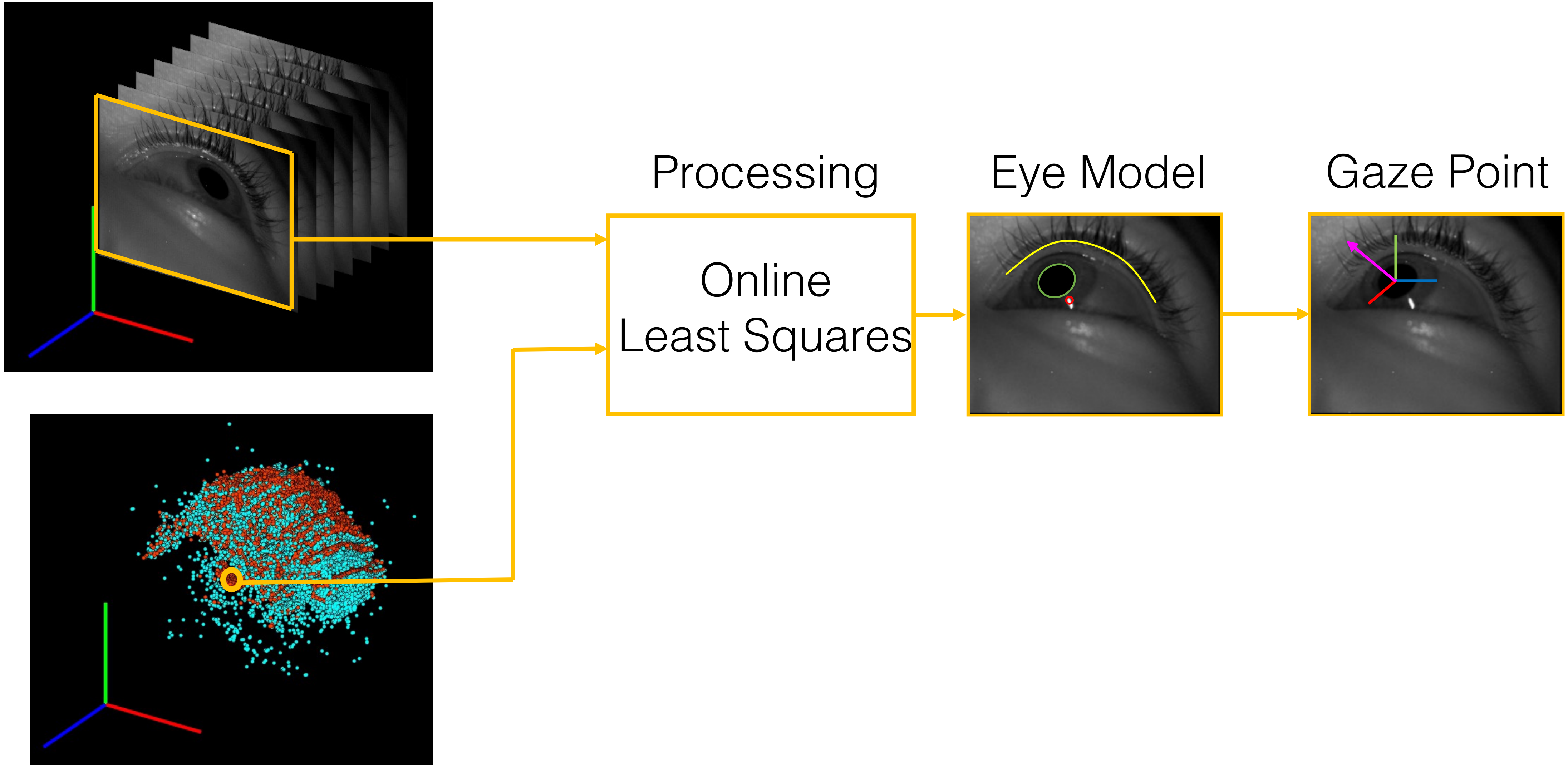


Angelopolous, Martel et al., IEEE VR 2021

And can capture what

happens between frames!

An overview of our system



Processing

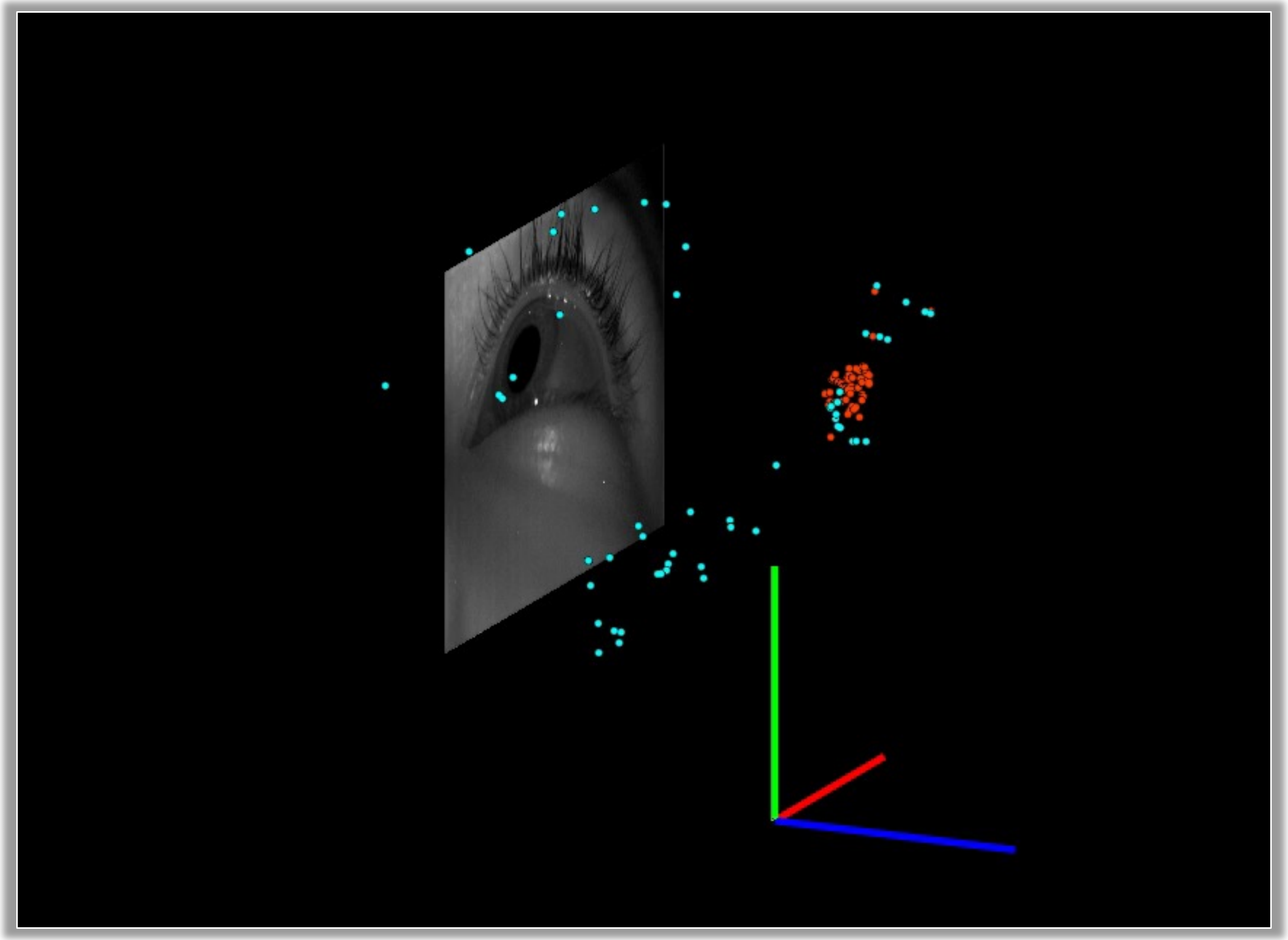
Online Least Squares

Eye Model

Gaze Point

Angelopolous, Martel et al., IEEE VR 2021

Fitting the pupil at high update rate



Angelopolous, Martel et al., IEEE VR 2021

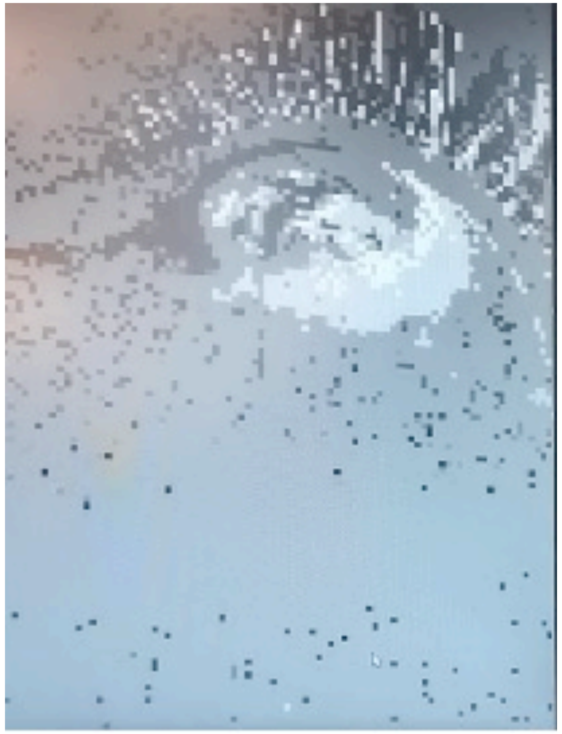
Estimating Gaze Vectors at 10kHz

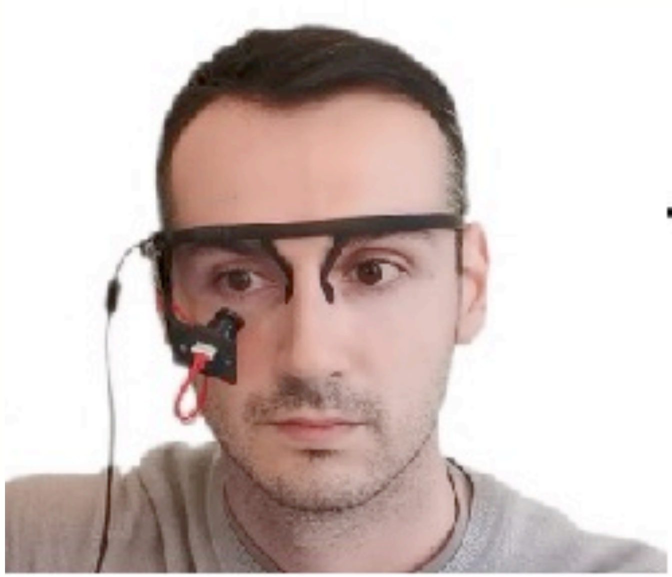
Angelopolous, Martel et al., IEEE VR 2021

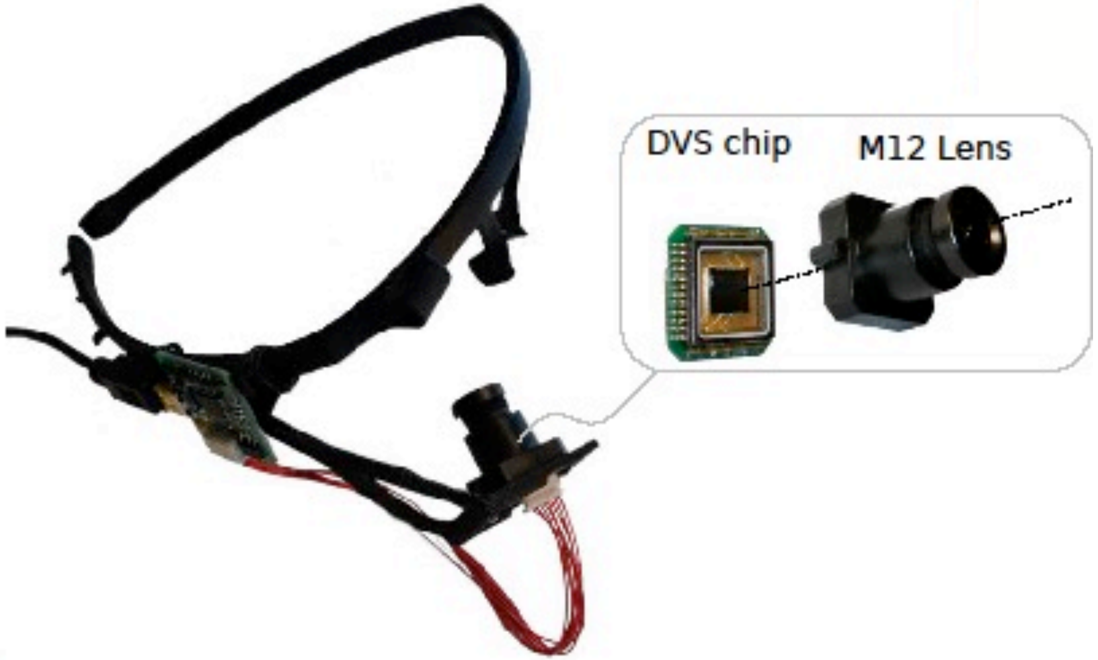
Real-time system

Angelopolous, Martel et al., IEEE VR 2021

Wearable prototype







Angelopolous, Martel et al., IEEE VR 2021

Augmenting Human Performance

Warby Parker

Augmented Vision

Bose Frames

Augmented Hearing

Snap Spectacles

Ray-Ban Stories

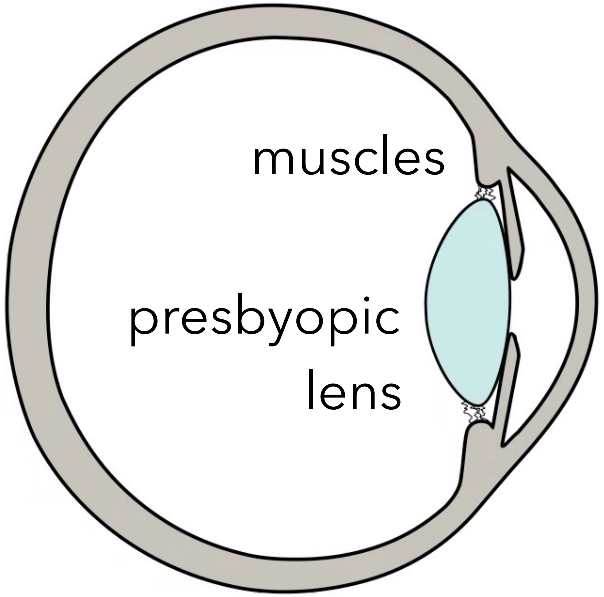
Augmented Memory

People don’t need technologies, so let’s engineer **experiences**!

Prebyopic Vision

Distance Near ∞

Normal



also used by schoolchildren to give classroom reports on planets, penguins, and poets. Microsoft rightly boasts of 1.2 billion copies of PowerPoint at large—one copy for every seven people on earth. In any given month, approximately 200 million of these copies are used, and although nobody's really counting, our cumulative generation of PowerPoint slides surely reaches well into the billions. So profound is PowerPoint's influence that prominent figures have decried the software's effects on thinking itself. Edward Tufte, the guru of information visualization, has famously railed against the "cognitive style" of PowerPoint, which he characterizes as having a "foreshortening of evidence and thought" and a "deeply hierarchical single-path structure."

PowerPoint is so ingrained in modern life that the notion of it having a history at all may seem odd. But it does have a very definite lifetime as a commercial product that came onto the scene 30 years ago, in 1987. Remarkably, the founders of the Silicon Valley firm that created PowerPoint did not set out to make presentation software, et alone build a tool that would transform roup communication throughout the orld. Rather, PowerPoint was a recovery om dashed hopes that pulled a strug- ing startup back from the brink of fail- e—and succeeded beyond anything its eators could have imagined.

"**Evasion and Deception**": In a 2003 speech to the United Nations Security Council, then U.S. Secretary of State Colin Powell made the case for going to war with Iraq. The accompanying PowerPoint slides included satellite imagery [top] that Powell said showed secret work on chemical and biological weapons.

* * *

owerPoint was not the first software for creating presentations on personal computers. Starting in roughly a half-dozen other pro- s came on the market before -Point's 1987 debut. Its eventual domination was not the result of first mover advantage. What's more, some of its most familiar features—the central motif of a slide containing text and graphics; bulleted lists; the slideshow; the slide sorter; and even the animated transitions between slides—did not originate with PowerPoint. And yet it's become the Kleenex or Scotch Tape of presentation software, as a "PowerPoint" has come to mean any presentation created with software.

With PowerPoint as well as its predecessors, the motif of the slide was, of course, lifted directly from the world of photography. Some presentation programs actually generated 35-mm slides for display with a slide projector. In most cases, though, the early programs created slides that were printed on paper for incorporation into reports, transferred to transparencies for use on overhead projectors, or saved as digital files to be displayed on computer monitors.

The upshot was that personal computer users of the 1980s, especially business users, had many options, and the market for business software was undergoing hypergrowth, with programs for generating spreadsheets, documents, databases, and business graphics each constituting a multimillion-dollar category. At the time, commentators saw the proliferation of business software as a new phase in office automation, in which computer use was spreading beyond the accounting department and the typing pool to the office elites. Both the imagined and actual users of the new business software were white-collar workers, from midlevel managers to Mahogany Row executives.

PowerPoint thus emerged during a period in which personal computing was taking over the American office. A major accelerant was the IBM Personal Computer, which Big Blue unveiled in 1981. By then, bureaucratic America—corporate and government alike—was well habituated to buying its computers from IBM. This new breed of machine, soon known simply as the PC, spread through offices like wildfire.

The groundwork for that invasion had been laid the previous decade, in the 1970s technosocial vision of the "office of the future." It started, like so much of what we now take for granted in our contemporary world of networked personal computing, at Xerox's legendary Palo Alto Research Center (PARC). The site was established in 1970 to invent the computing systems that would equip the future's white-collar office, an arena the company hoped to dominate in the same way it did photocopying. Many of the bright young computer scientists and engineers recruited to work at PARC knew one another from the major computer science programs funded by the Department of Defense's Advanced Research Projects Agency (ARPA) at MIT, Carnegie Mellon, Stanford, UC Berkeley, the University of Utah, and SRI.

In 1972, PARC researchers began to focus on a new personal computer they called the Alto. Led by Alan Kay, Butler Lampson, Bob Taylor, and Chuck Thacker, they were captivated by an extraordinary idea: that in the office of the future, every individual would have a dedicated computer like the Alto. Moreover, these computers would be networked to one another and to other, larger computers, both locally and far away. This networking would form a web of communication and computing resources well beyond the capacity of any single personal computer. In the pursuit of this vision, Ethernet emerged, as did the PARC Universal Packet protocol, or PUP, an important predecessor of the TCP/IP standard of today's Internet.

The Alto's creators emphasized the machine's graphics capabilities, dedicating much of the computer's hardware and software to rendering high-resolution imagery onscreen, including typography, drawings, digital photographs, and animations. It was a huge step up from the mainstream computers of the day, which still used punch cards, paper printouts, teletypes, and "dumb" terminals. Alto users interacted with it through a graphical interface to access, generate, and manipulate information. Even the text was treated as an image. The computer was controlled through a standard keyboard and the then-novel mouse that had emerged from Doug Engelbart's SRI laboratory.

This graphical turn in computing was perhaps most pronounced in one of the Alto's programming languages, called Smalltalk. Developed by Kay, Dan Ingalls, Adele Goldberg, and other collaborators, Smalltalk wasn't just a programming language; it was also a programming and user environment. It introduced the graphical user interface, or GUI, to personal computing, including a metaphorical desktop with overlapping windows, contextual and pop-up menus, file browsers, scroll bars, selection by mouse clicks, and even cut, copy, and paste.

While such innovations were ostensibly proprietary, by the end of the 1970s, Xerox managers and PARC staff were routinely discussing their findings with outsiders and publishing details of the Alto system in journals. PARC researchers were, after all, still part of the broader ARPA community of computer scientists and engineers. Many visitors who saw the Alto system considered it transformative.

Rob Campbell quit Apple to cofound Forethought with the ambitious goal of creating a graphical-software environment for the IBM PC.

One such visitor was Apple cofounder Steve Jobs. Following Xerox's investment in Apple in 1979, PARC researchers gave Apple engineers and management detailed demonstrations of Smalltalk and other programs previously reserved for Xerox insiders. Jobs was so enthralled by what he saw that he decided to reorient the Lisa, a business computer Apple was developing at the time, to fully embrace the PARC idiom. A few years later, when Jobs was transferred out of the Lisa project, he seized control of another effort aimed at creating a low-cost computer and pushed it, too, toward the PARC idiom. That computer became the Macintosh.

Taylor Pohlman cofounded Forethought with Rob Campbell. When their initial plan failed, they pivoted to other projects, including presentation software called Presenter.

What does all this have to do with PowerPoint? Apple lavished resources—people and cash alike—to embrace the PARC paradigm with the Lisa and the Macintosh, but not everyone at Apple was happy about that. In particular, those working to maintain the existing Apple II and III lines felt that their efforts were being shortchanged. By 1982, the product marketing manager for the Apple III, Taylor Pohlman, and the software marketing manager for the Apple II and III, Rob Campbell, had had enough. They quit and went into business together, founding the company that would create PowerPoint.

But PowerPoint was not at all in their original plan.

* * *

ne thing that united Pohlman and Campbell—but alienated them at Apple—was that they were cut from a different cloth than the computer-science types working on the Lisa and the Macintosh. Though both Pohlman and Campbell were technically minded, they were also oriented toward marketing and sales. Before Apple, Pohlman had worked in marketing at Hewlett-Packard, and Campbell had run a small accounting software company.

The pair left Apple late in 1982, and by early 1983, they had secured US $600,000 in venture capital to create a software company, which they called Forethought. Ironically, the startup's aim was to bring the PARC idiom to the IBM PC and its clones—in essence, to outplay Apple at its own game. That year, the Apple Lisa appeared, priced at nearly $10,000 (more than $25,000 in today's dollars).

TOP: U.S. DEPARTMENT OF STATE/[illegible] IMAGES; BOTTOM: [illegible]
TOP: [illegible] AUSTIN; BOTTOM: TAYLOR POHLMAN

Presbyopic Vision

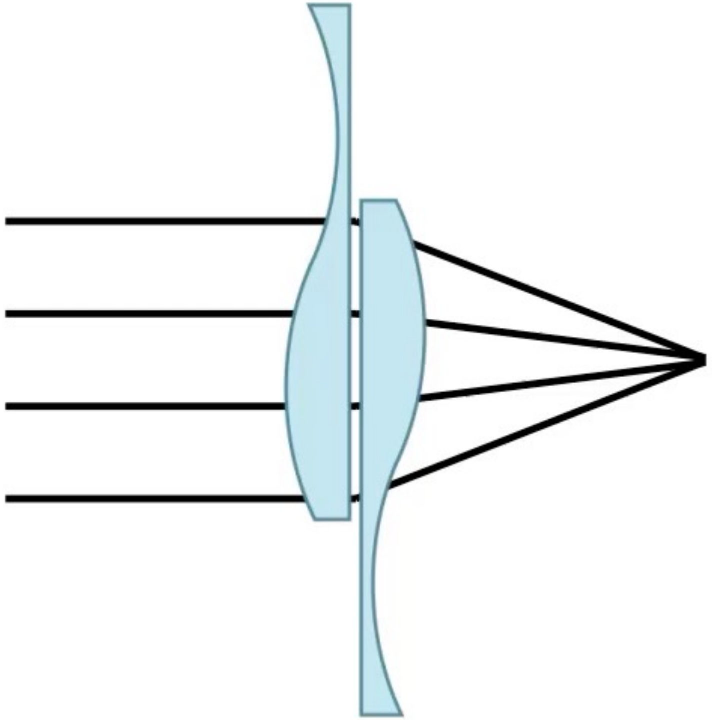
Reading Glasses

Progressive Glasses

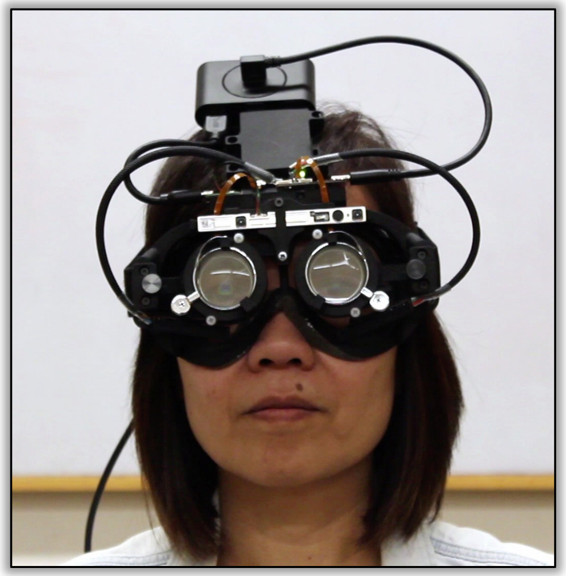
Natural Solution

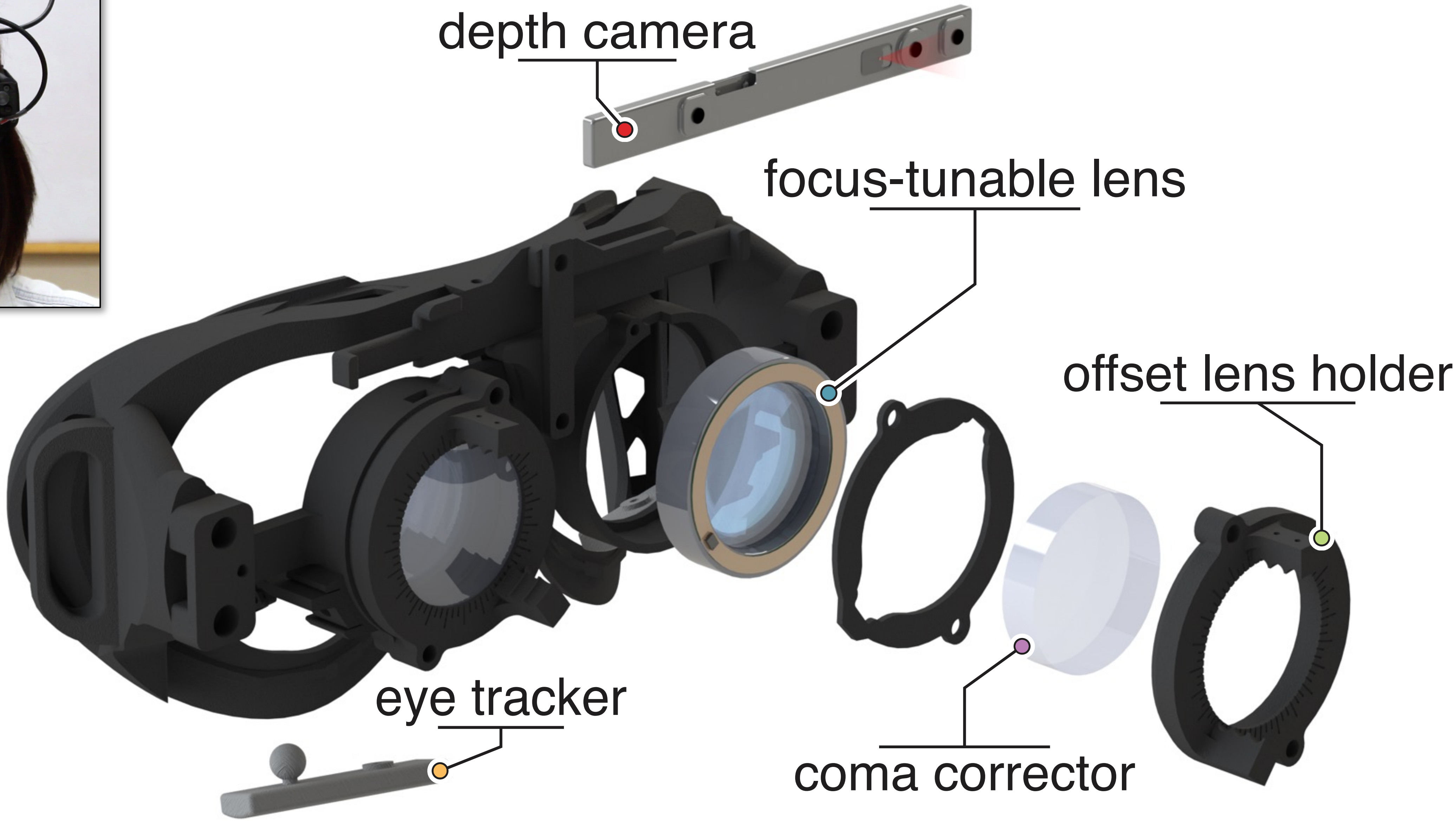
Focus-tunable Lens Technology

Alvarez



Autofocals





Padmanaban et al., Science Advances 2019

Autofocals

Padmanaban et al., Science Advances 2019

XR Device Form Factors / Go-to Market

Launch with full-field display features

Miniaturization and integration over time →

Mass-market form factor

Classic frame shapes and looks

New features and tech innovation →

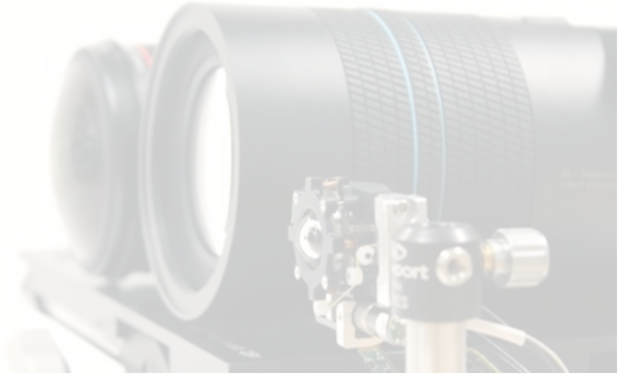
Full-featured, competitive smart frames

ZINN labs

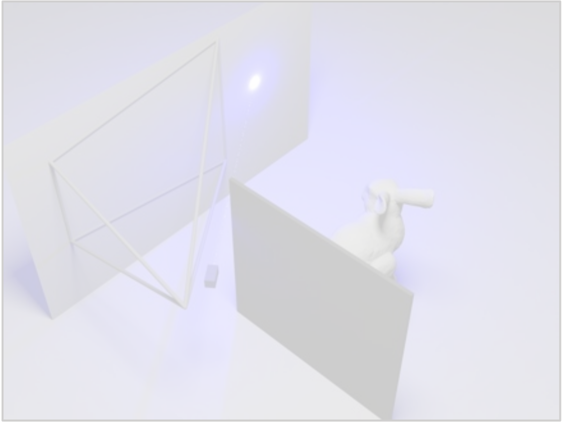
Eye tracking–enabled AI

Stanford Computational Imaging Lab

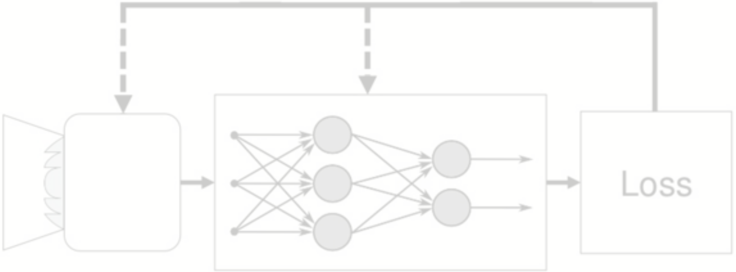
Computational Cameras



Single-photon Imaging

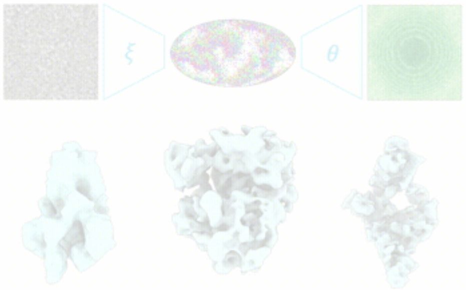


Deep Optics



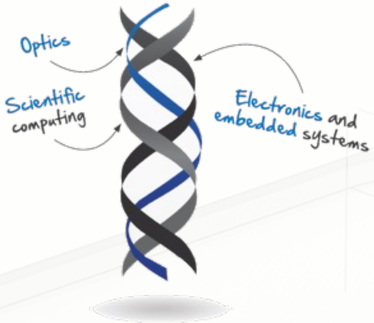
VR/AR & Wearable Computing

Computational Microscopy



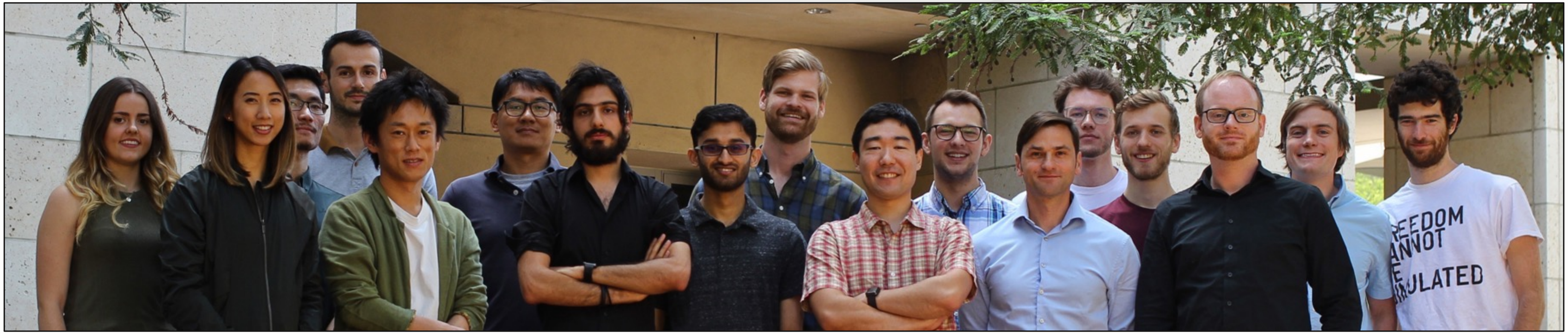
Neural Rendering

Gordon Wetzstein
stanford.edu/~gordonwz



Computational Imaging Lab
Stanford University EE & CS

computationalimaging.org



SCI STANFORD COMPUTATIONAL IMAGING LAB · NSF · DARPA · intel · THE OKAWA FOUNDATION for Information and Telecommunications · SAMSUNG · SONY