Wearable Displays with Focus Cues

EE367/CS448I: Computational Imaging and Display
stanford.edu/class/ee367
Lecture 16

Gordon Wetzstein
Stanford University
remote control of vehicles, e.g. drones

architecture walkthroughs

simulation & training

visualization & entertainment

robotic surgery

gaming

virtual travel

education

a trip down the rabbit hole
VR at Stanford’s Medical School

- Lucile Packard Children’s Hospital: used to alleviate pain, anxiety for pediatric patients

- VR Technology Clinic: applications in psychotherapy, mental health, for people with phantom pain, …

- help train residents, assist surgeons planning operations, …
Exciting Engineering Aspects of VR/AR

- cloud computing
- shared experiences
- compression, streaming
- VR cameras
- CPU, GPU
- IPU, DPU?
- sensors & imaging
- computer vision
- scene understanding
- photonics / waveguides
- human perception
- displays: visual, auditory, vestibular, haptic, …
- HCI
- applications

Images by microsoft, facebook
Exciting Engineering Aspects of VR/AR

- cloud computing
- shared experiences

- compression, streaming

- VR cameras

- CPU, GPU
- IPU, DPU?

- sensors & imaging
- computer vision
- scene understanding

- photonics / waveguides
- human perception
- displays: visual, auditory, vestibular, haptic, …

- HCI
- applications

Images by microsoft, facebook
Where We Want It To Be
Personal Computer
e.g. Commodore PET 1983

Laptop
e.g. Apple MacBook

Smartphone
e.g. Google Pixel

AR/VR
e.g. Microsoft Hololens
A Brief History of Virtual Reality

- **1838**: Stereoscopes
  - Wheatstone, Brewster, ...

- **1968**: VR & AR
  - Ivan Sutherland

- **1995**: Nintendo Virtual Boy
  - Virtual Boy

- **2012-2017**: VR explosion
  - Oculus, Sony, HTC, MS, ...

- ???
Ivan Sutherland’s HMD

- optical see-through AR, including:
  - displays (2x 1" CRTs)
  - rendering
  - head tracking
  - interaction
  - model generation

- computer graphics
- human-computer interaction

I. Sutherland “A head-mounted three-dimensional display”, Fall Joint Computer Conference 1968
Nintendo Virtual Boy

- computer graphics & GPUs were not ready yet!

Game: Red Alarm
Virtual Image

Problems:

- fixed focal plane
- no focus cues 😞
- cannot drive accommodation with rendering!

\[ \frac{1}{d} + \frac{1}{d'} = \frac{1}{f} \]
Focus Cues – An Important Depth Cue
Importance of Focus Cues Decreases with Age - Presbyopia

Duane, 1912
Relative Importance of Depth Cues

Cutting & Vishton, 1995
Stereopsis (Binocular)

Oculomotor Cue

Vergence

extraocular muscles

Visual Cue

Binocular Disparity

Focus Cues (Monocular)

Accommodation

relaxed

contracted

ciliary muscles

Retinal Blur
Stereopsis (Binocular)

Oculomotor Cue

Vergence

Visual Cue

Binocular Disparity

Focus Cues (Monocular)

Extraocular muscles

Ciliary muscles

Relaxed

Contracted

Accommodation

Retinal Blur
Stereopsis (Binocular)

Oculomotor Cue

Vergence

extraocular muscles

Visual Cue

Binocular Disparity

Focus Cues (Monocular)

Accommodation

relaxed

ciliary muscles

contracted

Retinal Blur
Blur Gradient Driven Accommodation

Conventional Display

<table>
<thead>
<tr>
<th>Distance (m)</th>
<th>Accommodation (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>(4D)</td>
</tr>
<tr>
<td>0.3</td>
<td>(3.33D)</td>
</tr>
<tr>
<td>0.35</td>
<td>(2.88D)</td>
</tr>
<tr>
<td>0.5</td>
<td>(2D)</td>
</tr>
<tr>
<td>0.7</td>
<td>(1.43D)</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(0.5D)</td>
</tr>
<tr>
<td>∞</td>
<td>(0D)</td>
</tr>
</tbody>
</table>

virtual image of screen
Blur Gradient Driven Accommodation

Conventional Display

- 0.25m (4D)
- 0.3m (3.33D)
- 0.35m (2.86D)
- 0.5m (2D)
- 0.7m (1.43D)
- 1m
- 2m (0.5D)
- ∞ (0D)
Blur Gradient Driven Accommodation

Conventional Display

virtual image of screen

0.25m (4D)  0.3m (3.33D)  0.35m (2.86D)  0.5m (2D)  0.7m (1.43D)  1 m  2m (0.5D)  ∞ (0D)
Blur Gradient Driven Accommodation

Conventional Display

0.25m (4D)  0.3m (3.33D)  0.35m (2.86D)  0.5m (2D)  0.7m (1.43D)  1m  2m (0.5D)  ∞ (0D)
Blur Gradient Driven Accommodation

Conventional Display

- 0.25m (4D)
- 0.3m (3.33D)
- 0.35m (2.86D)
- 0.5m (2D)
- 0.7m (1.43D)

Virtual image of screen

1 m
2 m (0.5D)
∞ (0D)
Blur Gradient Driven Accommodation

Conventional Display

virtual image of screen

0.25m (4D) 0.3m (3.33D) 0.35m (2.86D) 0.5m (2D) 0.7m (1.43D) 1m 2m (0.5D) ∞ (0D)
Blur Gradient Driven Accommodation

Conventional Display

Accommodation-dependent Point Spread Functions

virtual image of screen

0.25m (4D)  0.3m (3.33D)  0.35m (2.86D)  0.5m (2D)  0.7m (1.43D)  1m  2m (0.5D)  ∞ (0D)
The Vergence-Accommodation Conflict (VAC)
Real World: 
Vergence & Accommodation 
Match!
Current VR Displays: Vergence & Accommodation Mismatch
Consequences of Vergence-Accommodation Conflict

- Visual discomfort (eye tiredness & eyestrain) after ~20 minutes of stereoscopic depth judgments (Hoffman et al. 2008; Shibata et al. 2011)

- Degrades visual performance in terms of reaction times and acuity for stereoscopic vision (Hoffman et al. 2008; Konrad et al. 2016; Johnson et al. 2016)

- also: double vision (diplopia), reduced visual clarity, possibly nausea
• **Q1**: How to address the vergence-accommodation conflict for users of different ages?

• **Q2**: Can computational displays effectively replace glasses in VR/AR?

• **Q3**: What are (in)effective near-eye display technologies?

  possible solutions: gaze-contingent focus, light fields, ...
1. Gaze-contingent Focus
Fixed Focus

\[
\frac{1}{d} + \frac{1}{d'} = \frac{1}{f}
\]
Adaptive Focus

Magnified Display

Lens

Display

\[ \frac{1}{d} + \frac{1}{d'} = \frac{1}{f} \]
Adaptive Focus

Magnified Display

\[ \frac{1}{d} + \frac{1}{d'} = \frac{1}{f} \]
Adaptive Focus - History

- M. Heilig “Sensorama”, 1962 (US Patent #3,050,870)
- S. Shiwa, K. Omura, F. Kishino “Proposal for a 3-D display with accommodative compensation: 3DDAC”, JSID 1996
- S. McQuaide, E. Seibel, J. Kelly, B. Schowengerdt, T. Furness “A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror”, Displays 2003
- S. Liu, D. Cheng, H. Hua “An optical see-through head mounted display with addressable focal planes”, Proc. ISMAR 2008
Conventional Stereo / VR Display

Conventional

virtual image distance

stereoscopic distance

stereoscopic distance

virtual image distance

stereoscopic distance

vergence

accommodation
Removing VAC with Adaptive Focus

With Focus Cues

virtual image distance

vergence
accommodation

stereoscopic distance

stereoscopic distance
Follow the target with your eyes
Accommodative Response

Padmanaban et al., PNAS 2017
Accommodative Response

Conventional

Relative Distance [D]

Time [s]

n = 59, mean gain = 0.29

Stimulus
Accommodation

Padmanaban et al., PNAS 2017
Accommodative Response

Padmanaban et al., PNAS 2017
Accommodative Response

Padmanaban et al., PNAS 2017

n = 24, mean gain = 0.77
Do Presbyopes Benefit from Dynamic Focus?
Do Presbyopes Benefit from Dynamic Focus?

Gain vs. Age

Gain

Age

Conventional

Padmanaban et al., PNAS 2017
Do Presbyopes Benefit from Dynamic Focus?

Gain vs. Age plot showing data points for conventional and dynamic focus. The graph indicates a trend where dynamic focus may benefit presbyopes as it shows lower gain values compared to conventional focus with age.
Do Presbyopes Benefit from Dynamic Focus?

Gain vs Age

- Conventional
- Dynamic

Response for Physical Stimulus
Heron & Charman 2004

Padmanaban et al., PNAS 2017
Gaze-contingent Focus

- non-presbyopes: adaptive focus is like real world, but needs eye tracking!

Padmanaban et al., PNAS 2017
Gaze-contingent Focus

Padmanaban et al., PNAS 2017
Gaze-contingent Focus

Padmanaban et al., PNAS 2017
Gaze-contingent Focus

Padmanaban et al., PNAS 2017
Oculus announces gaze-contingent varifocal display at F8, 05/2018
Oculus Half Dome Prototype

Video courtesy of Facebook/Oculus
Summary

- adaptive focus drives accommodation and can correct for refractive errors (myopia, hyperopia)

- gaze-contingent focus gives natural focus cues for non-presbyopes, but require eyes tracking

- presbyopes require fixed focal plane with correction
2. Light Field Displays
Near-eye Light Field Displays

Idea: project multiple different perspectives into different parts of the pupil!
Light Field Stereoscope

Huang et al., SIGGRAPH 2015
Input: 4D light field for each eye

Model Courtesy of Bushmills Irish Whiskey
Multiplicative Two-layer Modulation

Input: 4D light field for each eye

Model Courtesy of Bushmills Irish Whiskey
Multiplicative Two-layer Modulation

Input: 4D light field for each eye

Model Courtesy of Bushmills Irish Whiskey
Multiplicative Two-layer Modulation

Input: 4D light field for each eye

Model Courtesy of Bushmills Irish Whiskey
Reconstruction:

\[
\begin{align*}
t_1 &\leftarrow t_1 \circ \frac{\phi_1^T(\beta_1 \circ (\phi_2 t_2))}{\phi_1^T(\beta_1 \circ (\phi_2 t_2)) + \epsilon} \\
\end{align*}
\]

\(\text{for layer } t_1\)

Tensor Displays, Wetzstein et al. 2012

\[
\begin{align*}
\text{minimize } & \| \beta_1 - (\phi_1 t_1) \circ (\phi_2 t_2) \|^2 \\
\text{s.t. } & 0 \leq t_1, t_2 \leq 1
\end{align*}
\]
Traditional HMDs
- No Focus Cues

The Light Field HMD Stereoscope

Huang et al., SIGGRAPH 2015
Light Field Stereoscope

Traditional HMDs - No Focus Cues

The Light Field HMD Stereoscope

Huang et al., SIGGRAPH 2015
Light Field Stereoscope

Traditional HMDs - No Focus Cues

The Light Field HMD Stereoscope

Huang et al., SIGGRAPH 2015

Model Courtesy of Paul H. Manning
Light Field Stereoscope

Traditional HMDs - No Focus Cues

The Light Field HMD Stereoscope

Huang et al., SIGGRAPH 2015

Model Courtesy of Paul H. Manning
Tensor Displays

Wetzstein et al., SIGGRAPH 2012
300 dpi or higher

prototype
Diffraction in Multilayer Light Field Displays

Wetzstein et al., SIGGRAPH 2011
Lanman et al., SIGGRAPH Asia 2011
Wetzstein et al., SIGGRAPH 2012
Maimone et al., Trans. Graph. 2013
...

Less diffraction artifacts with LCoS

Hirsch et al, SIGGRAPH 2014
3. Accommodation-invariant Near-eye Displays
Blur Gradient Driven Accommodation

Conventional Display

Accommodation-dependent Point Spread Functions

虚像位置的屏幕

距离

0.25m (4D) 0.3m (3.33D) 0.35m (2.86D) 0.5m (2D) 0.7m (1.43D) 1m 2m (0.5D) + (0D)
PSF Engineering

Conventional Display

Accommodation-invariant Display

Accommodation-dependent Point Spread Functions

virtual image of screen

0.25m (4D) 0.3m (3.33D) 0.35m (2.86D) 0.5m (2D) 0.7m (1.43D) 1 m 2m (0.5D) \(\infty\) (0D)
Q: can we drive accommodation with stereoscopic cues by optically removing the retinal blur cue?
Vergence

extraocular muscles

Accommodation

relaxed
contracted

ciliary muscles

Binocular Disparity

Retinal Blur
How do we remove the blur cue?
Aperture Controls Depth of Field

Image courtesy of Concept One Studios
Aperture Controls Depth of Field

Image courtesy of Concept One Studios
Aperture Controls Depth of Field

Image courtesy of Concept One Studios
Maxwellian-type (pinhole) Near-eye Displays

Point Light Source
Maxwellian-type (pinhole) Near-eye Displays

Severely reduces eyebox; requires dynamic steering of exit pupil
Focal Sweep

EDOF Cameras:
Nagahara et al., ECCV 2008
Cossairt et al., SIGGRAPH 2010
Convolution
Convolution
Deconvolution
Stimulus

Distance (D)

Time (s)

Accommodation Invariant

0 5 10 15 20

Gain: 0.61

Dynamic

0 5 10 15 20

Gain: 0.85

Gain: 0.85

Gain: 0.61
Measured User Response

![Graph showing user response over time with Gain: 0.35 for conventional method.](image)

**Distance (D)**

- Gain: 0.61
- Dynamic
- Conventional

**Time (s)**

0 5 10 15 20

**Stimulus**

**Average**
Measured User Response

**Distance (D)**

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Gain: 0.61</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

**Conventional**

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Gain: 0.35</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

**Stimulus**

Gain: 0.85
Now: benchtop

Future: multifocal lenses
Photonics Challenges for Getting Here
Thin Beam Combiner?
Thin Beam Combiner!
Pepper’s Ghost 1862
Case Studies
Google Glass
Meta 2

- larger field of view (90 deg) than Glass
- also larger device form factor
Microsoft HoloLens

- diffraction grating-based waveguide
- LCoS microdisplay
- field of view: 34° diagonally, 16:9 aspect, 47 pixels per visual degree
Microsoft HoloLens 2

- laser-scanned waveguide display
- claimed 2K resolution per eye (2560x1440), probably via “interlaced” scanning
- field of view: 52° diagonally (3:2 aspect, 47 pixels per visual degree)

Zeiss Smart Optics

- great device form factor
- polycarbonate light guide – easy to manufacture and robust
- smaller field of view (17 deg)
Challenges: Eye Box vs Field of View
Challenges: Eye Box vs Field of View

• need small entrance pupil (small device) and large exit pupil (large eye box) - pupil needs to be magnified
Challenges: Eye Box vs Field of View

- need small display (small device) but large field of view – image needs to be magnified
Challenges: Eye Box vs Field of View

- pupil needs to be magnified
- image needs to be magnified

Can’t get both at the same time – etendue!
Challenges: Eye Box vs Field of View

- possible solutions: exit pupil replication (loss of light), live with small FOV (not great), dynamically steer eye box (mechanically difficult), ..
Challenges: Chromatic Aberrations

- thin grating couplers create chromatic aberrations
Challenges: Chromatic Aberrations

- all solutions have their own problems: ease of manufacturing, yield, robustness, cost, …

volume holographic couplers, e.g. TruLife Optics

stacked waveguides
Occlusions

Case 1:
virtual in front of real

→ difficult: need to block real light!

Case 2:
real in front of virtual

→ easy: don’t render virtual object everywhere
Next quarter: EE 267
The End