Lecture 25: Virtual Reality and Advanced Rendering Topics

Yuhaoo Zhu
http://yuhaozhu.com

CSC 292/572, Fall 2020
Mobile Visual Computing
Logistics

- Programming assignment 2 released. Due 12/04.
- Three guest lectures in a row!
  - Michael J. Murdoch 12/2 on Color in Augmented Reality
  - Robert LiKamWa 12/4 on Mobile Augmented Reality
  - Adam Sefkow 12/9 on Physics Simulation
Virtual Reality

immersive display; everything you see is emitted from the display.

Oculus Quest

https://wccftech.com/oculus-quest-2-leaked-almost-4k-resolution-snapdragon-xr-2-6gb-ram/
Oculus Rift S

https://uploadvr.com/rift-s-hardware-review
https://www.youtube.com/watch?v=-1izYqKyJ80
Samsung Gear VR


Google Cardboard

https://www.youtube.com/watch?v=dVdQdT5sN8
Augmented Reality

what you see = real world light + light projected from display.
Google Glass

https://www.techradar.com/reviews/gadgets/google-glass-1152283/review

Google Glass

Couple-in optics

Wave Guide Plate (Glass)

Couple-out optics

HOE

Lens

Micro-display

Eye

https://www.techradar.com/reviews/gadgets/google-glass-1152283/review

https://www.intechopen.com/books/state-of-the-art-virtual-reality-and-augmented-reality-know-how/waveguide-type-head-mounted-display-system-for-ar-application
Microsoft Hololens

AR on Smartphone

Massive Pixels to Render?

~200°

~5°
Photoreceptor Density

- Uneven distribution. The area on the retina with the highest cone density is called fovea.
- Fovea angle is about 2°. Peripheral vision has very low acuity.

Foveated Rendering
Towards Foveated Rendering for Gaze-Tracked Virtual Reality, SIGGRAPH Asia 2016

Anjul Patney

Marco Salvi
Joohwan Kim
Anton Kaplanyan
Chris Wyman
Nir Bental

Abstract

Foveated rendering could drive future wide field-of-view displays. As a result, algorithms that imperceptibly reduce cost are becoming more important. Interestingly, human visual acuity radially decreases between the retina center (the fovea) and the eye's periphery, and for HMDs and large desktop displays a significant percentage of pixels lie in regions viewed with lower visual acuity. Given our perceptual target, we designed a practical foveated rendering system that reduces number of shades by up to 70% and allows adoption of realistic lighting and physically based shading using pre-filtered shading terms, contrast preservation, and applying a new temporal antialiasing algorithm, and use contrast enhancement in the periphery, introduce a novel multiresolution- and saccade-aware temporal antialiasing algorithm, 

coarsened shading up to 30× with credit is permitted. To copy otherwise, or republish, to post on servers, to redistribute to lists, or to use any par...
Towards Foveated Rendering for Gaze-Tracked Virtual Reality

Anjul Patney

Marco Salvi Joohwan Kim Anton Kaplanyan Chris Wyman Nir Benty

Towards Foveated Rendering for Gaze-Tracked Virtual Reality, SIGGRAPH Asia 2016

Foveated Rendering

Foveated rendering

We validate our system by performing another user study. Frequency to study and improve potential gains, we designed a foveated render-

vision when viewing today's displays. We determined that filtering

this notice and the full citation on the first page. Copyrights for components made or distributed for profit or commercial advantage and that copies bear

Permission to make digital or hard copies of all or part of this work for

measurements of temporal stability show we obtain quality similar

degraded by filtering.

aware temporal antialiasing algorithm, and use contrast enhancement

in the periphery, introduce a novel multiresolution- and saccade-

both pre- and post-shading to address aliasing from undersampling

Given our perceptual target, we designed a practical foveated render-

from a non-foveated ground truth. After verifying these insights on

peripheral regions reduces contrast, inducing a sense of tunnel vi-

than the performance of traditional real-time renderers.

displays, where target framerate and resolution is increasing faster

speedups for wide field-of-view displays, such as head mounted

outside the eye fixation region, potentially unlocking significant

coarsened shading up to

ing system that reduces number of shades by up to 70% and allows

of magnitude (providing stability similar to a temporally antialiased non-foveated renderer). The original version of the classroom scene is

using pre-filtered shading terms, contrast preservation, and applying a new temporal antialiasing that improves temporal stability by an order

Figure 1:

Periphery

Fovea
How?

- Render with different resolutions.
- Stack and blend them together.
- Patch boundaries will have clear edge artifacts. Use low-pass filter.
- Enhance contrast as a final pass.
Foveated Ray Tracing

- Idea: trace more rays for pixels in fovea.
- Rendering can also employ a sparse foveated pattern [Stengel et al. 2016; Weier et al. 2016], which becomes a practical option with the recent advances in real-time ray tracing hardware, such as NVIDIA RTX, and fine resolution control for rasterization.
FOVEATED RENDERING

- We can only see clearly where we are looking at
- Shading at full rate everywhere is a waste of computation

Steps
- Create a density map
- Ray trace 1 sample for each area
- Reconstruct full resolution image
FOVEATED RENDERING

- We can only see clearly where we are looking at
- Shading at full rate everywhere is a waste of computation

Steps
- Create a density map
- Ray trace 1 sample for each area
- Reconstruct full resolution image
FOVEATED RENDERING

- We can only see clearly where we are looking at
- Shading at full rate everywhere is a waste of computation

Steps
- Create a density map
- Ray trace 1 sample for each area
- Reconstruct full resolution image
DeepFovea: Neural Reconstruction

DeepFovea: Neural Reconstruction for Foveated Rendering and Video Compression using Learned Statistics of Natural Videos, SIGGRAPH Asia 2019
Gaze Tracking
Gaze Tracking Hardware

Gaze Tracking Hardware


Pico Neo 2 Eye headset

A ring of cameras

An IR LED?
Gaze Tracking Algorithm

Segment the pupil from the eye

Estimating the orientation (gaze)

Lens in VR Headsets

› Create a wide field of view.

Lens in VR Headsets

- Create a wide field of view.

Vive

Rift

Lens in VR Headsets

- Placing the display far away, although the display is incredibly close to your eyes.

Virtual display that users think they are looking at (meters ~ ∞)

Actual display in VR headset (~10cm)

Eye

https://xinreality.com/wiki/Virtual_Reality_Headset_Lenses
Lens in VR Headsets

- Placing the display far away, although the display is incredibly close to your eyes.

Virtual display that users think they are looking at (meters $\sim \infty$)

Actual display in VR headset ($\sim 10$ cm)

Eye

1.4 m in Oculus Rift DK2/

https://xinreality.com/wiki/Virtual_Reality_Headset_Lenses
Lens in VR Headsets

- Placing the display far away, although the display is incredibly close to your eyes.

Virtual display that users think they are looking at (meters ~ ∞)

Actual display in VR headset (~10cm)

Eye

Oculus Rift DK1, virtual display is ∞

1.4 m in Oculus Rift DK2/

https://xinreality.com/wiki/Virtual_Reality_Headset_Lenses
Accommodation and Vergence

- Normally when you eyes focus on objects at different distances, both accommodation and vergence are activated and are *coordinated*.

**Accommodation**: changing eye lens’s focal length to focus on objects at different distances.

**Vergence**: rotating eye balls so that images are formed on the retina (highest visual acuity).
Accommodation Vergence Conflict (VAC)

- Problem with VR: the lights from objects at difference distances are actually emitted at the same fixed distance!
  - The physical display doesn’t move, and so the virtual display distance is fixed!

https://www.businessinsider.com/virtual-reality-gave-me-depth-perception-2016-4
Accommodation Vergence Conflict (VAC)

- Problem with VR: the lights from objects at difference distances are actually emitted at the same fixed distance!
  - The physical display doesn’t move, and so the virtual display distance is fixed!

If you were to look at the two objects in the real-world, your eyes should have a different set of A and V, but since the virtual display is the fixed, the A will be fixed. Your brain received signals that it’s not used to (fatigue, nausea, etc.).
How to Resolve VAC?

- **Requirement 1: Mimic variable focal length**
  - Farther objects should require longer focus length to focus; closer objects should require shorter focus length to focus.

- **Three basic techniques:**
  - Varifocal display
  - Multifocal display
  - Lightfield display
Varifocal Display

- Idea: using a focus-tunable lens to mimic different focal lengths of eyes.
- One single display, but multiple focal planes.
- Require precise eye/gaze tracking.
  - Not the eye movement, but the gaze: the focus could change even if the eye is not moving (more later).
- Work flow:
  - Render one image —> Track Gaze —> Adjust lens
- Need to explicitly render defocus blur
  - Otherwise everything will be in focus, even out-of-focus area
Multifocal Display

- Idea: using multiple focal planes, each displaying an image; multiple images are overlapped on the retina.
- Does not require eye/gaze tracking
  - Images at different focal planes are specifically rendered in a way that when user focuses at a particular distance (e.g., one object), the object is in focus and the rest is blurred.
Multifocal Display

- Could have multiple physical implementations
  - Multiple transparent displays (Jannick Rolland, 1999)
  - A focus-tunable lens with one display

- **Multiple displays work flow:**
  - Render one image for each display —> That’s it!

- **Focus-tunable lens flow:**
  - Adjust the focal length to a particular length (e.g., apply voltage) —> wait for the lens to settle —> *(in parallel)* render the image at that particular length —> adjust the focal length again…
  - Different images have huge redundancies. Rendering-optics co-design to simplify rendering?
Multifocal Display

- Quality depends on the number of “effective” focal planes.
  - Intuitively the more the better
- But limited by the physical space (for a multi-display setup) or by the settling time of a focus-tunable lens (for a tunable-lens setup).
  - Typically 5ms to settle, which means 200 focus planes in 1 second. If the display refresh rate is 60 Hz (i.e., requires 60 frames a second), each frame can use at most 3 focal planes.
Multifocal Display: Dense Focal Stacks

- Key idea: don’t wait for the lens to settle, but track the length and then render accordingly. Sweep the lens through its entire range of focal lengths at high speed.

- Key design parameters:
  - Tracking time
  - Tracking accuracy
  - Rendering time
  - Rendering&tracking sync.

Work flow: **Track focal length → render according to that focal length.** Note the dependency!

Towards Multifocal Displays with Dense Focal Stacks, SIGGRAPH Asia 2018
Multifocal Display: Dense Focal Stacks

- In their design:
  - 40 sweeps per second, 40 tracks per sweep, so 40 different focal planes per frame
  - But needs to render 1600 frames per second!
  - Also the display needs display 1600 frames per second too. Commercial Digital Micromirror Displays (DMDs) could do that.
  - They are all similar with huge redundancies; opportunities for optimizations.

- Synchronization
  - Rendering needs to happen after focal length detection. So when a frame is rendered and displayed, the focal length has changed.
  - If not perfect synchronization, why bother tracking anyways?
  - Why don’t we assume uniform speed, and render 40 frames at an even focal length gap? This could be done in parallel, too.

How to Resolve VAC?

- Requirement 2: Mimic retinal defocus cue
  - Smithline [1974] identified retinal defocus blur, chromatic aberration, and looming as potentially involved in accommodative control.
  - Burge and Geisler [2011] reported reliable depth estimates solely from defocus blur.
  - Synthesizing accurate defocus blur has also been shown to result in the correct perception of depth and scale [Held et al. 2010, 2012].
DeepFocus

Everything is in-focus

In-focus objects are rendered in-focus

Out-of-focus objects are actually rendered out-of-focus

Input RGB  Unity (0.1D)  Nuke (0.1D)  Nalbach et al. (0.1D)  DeepFocus (0.1D)  Reference (0.1D)
DeepFocus: Learned Image Synthesis for Computational Displays, SIGGRAPH Asia, 2018
Reference Materials

- Two gentle articles to VR displays:
  - Head-mounted Displays and Lenses
  - Accommodation and Vergence in Head-mounted Displays

- Resolving VAC:
  - Towards Multifocal Displays with Dense Focal Stacks, SIGGRAPH Asia 2018
  - DeepFocus: Learned Image Synthesis for Computational Displays, SIGGRAPH Asia, 2018

- Foveated rendering and eye tracking:
  - Towards Foveated Rendering for Gaze-Tracker Virtual Reality, SIGGRAPH Asia 2016
  - DeepFovea: Neural Reconstruction for Foveated Rendering and Video Compression using Learned Statistics of Natural Videos, SIGGRAPH Asia 2019
  - RITnet: Real-time Semantic Segmentation of the Eye for Gaze Tracking, 2019