Eye Tracking and Virtual Reality

By

Ann McNamara and Eakta Jain

Course Notes
SIGGRAPH ASIA 2019

November 2019

ABSTRACT

Virtual Reality has the potential to transform the way we work, rest and play. We are seeing use cases as diverse as education and pain management, with new applications being imagined every day. VR technology comes with new challenges, and many obstacles need to be overcome to ensure good user experience. Recently many new Virtual Reality systems with integrated eye tracking have become available. This course presents timely, relevant information on how Virtual Reality (VR) can leverage eye-tracking data to optimize the user experience and to alleviate usability issues surrounding many challenges in immersive VEs. The integration of eye tracking allows us to determine where the viewer is focusing their attention. If we, as the content creators and world builders, need the user to focus on another area of the VE we can use techniques to attract attention to these regions and also we can confirm we are doing so successful as we continually track the users gaze. Advancing these approaches could make the VR experience more comfortable, safe and effective for the user.

TABLE OF CONTENTS

Page

INTRODUCTION

Virtual Reality has the potential to transform the way we work, rest and play. We are seeing use cases as diverse as education and pain management, with new applications being imagined every day. VR technology comes with new challenges, and many obstacles need to be overcome to ensure good user experience. Recently many new Virtual Reality systems with integrated eye tracking have become available. This course presents timely, relevant information on how Virtual Reality (VR) can leverage eye-tracking data to optimize the user experience and to alleviate usability issues surrounding many challenges in immersive VEs. The integration of eye tracking allows us to determine where the viewer is focusing their attention. If we, as the content creators and world builders, need the user to focus on another area of the VE we can use techniques to attract attention to these regions and also we can confirm we are doing so successful as we continually track the users gaze. Advancing these approaches could make the VR experience more comfortable, safe and effective for the user.

1.1 Motivation

Several Head Mounted Displays (HMDs) with integrated eye tracking have recently hit the market. For these HMDs to prove useful, it is essential that the community has an understanding of how these measurements should be recorded, analyzed, and reported. It is also critical that the community can take advantage of built-in eye tracking to advance IVEs using novel techniques such as gaze direction, foveated rendering, and beyond.

1.2 Target Audience

The target audience for this course is researchers interested in applying eye-tracking in Virtual Reality. This course represents a birds-eye view of the eye-movements, eye-tracking, data capture and analysis, and state-of-the-art applications of eye-tracking in VR. Those wishing to grasp the basics will all benefit from this course.

1.3 Course Overview

The integration of eye-tracking and VR reveals where the user focuses their attention. Content creators and world builders can exploit this information to make the VR experience comfortable, safe and effective for the user. This course provides the necessary background and overview to get up-and-running quickly with eye-tracking in VR

1.4 Course Outline

Welcome and Overview

Eye Movements

Virtual Environments and Eye Tracking (including an overview of hardware available)

Special Considerations for Eye Tracking in Virtual Environments

Collecting Your Own Data

Data Analysis & Reporting

A brief overview of existing studies including two recent case studies

Summary and Future Directions

Questions from the Audience

1.5 Course Speaker Bio: Ann McNamara

Ann McNamara is an Associate professor in the Department of Visualization at Texas A&M University. Her research focuses on novel approaches for optimizing an individual's experience when creating, viewing and interacting with virtual and augmented spaces. She is the recipient of an NSF CAREER AWARD entitled "Advancing Interaction Paradigms in Mobile Augmented Reality using Eye Tracking". This project investigates how mobile eye tracking, which monitors where a person is looking while on the go, can be used to determine what objects in a visual

scene a person is interested in, and thus might like to have annotated in their augmented reality view. In 2019, she was named as one of twenty-one Presidential Impact Fellows at Texas A&M University.

1.6 Course Notes Contributor: Eakta Jain

Eakta Jain is an Assistant Professor of Computer and Information Science and Engineering at the University of Florida. She received her PhD and MS degrees in Robotics from Carnegie Mellon University, working in the Graphics Lab. Her B.Tech. degree is in Electrical Engineering from IIT Kanpur. She has worked in industrial research at Texas Instruments RD labs, Disney Research Pittsburgh, and the Walt Disney Animation Studios. Eakta's research interests are in building human-centered computer graphics algorithms to create and manipulate artistic content, including traditional hand animation, comic art, and films. Her work has been presented at venues such as ACM SIGGRAPH, and has won multiple awards.

ANNOTATED COURSE DECK

Eye Tracking and Virtual Reality
Ann McNamara, Department of Visualization, Texas A&M University

sa2019.siggraph.org

Welcome and Overview
Ann McNamara

sa2019.siggraph.org

Course Outline

- Welcome and Overview
- Eye Movements
- Virtual Environments and Eye Tracking, including an overview of hardware
- Special Considerations for Eye tracking in VR
- Collecting Your Own Data
- Data Analysis Reporting
- A brief overview of existing studies including two case studies
- Conclusion and Future Directions

Eye Movements
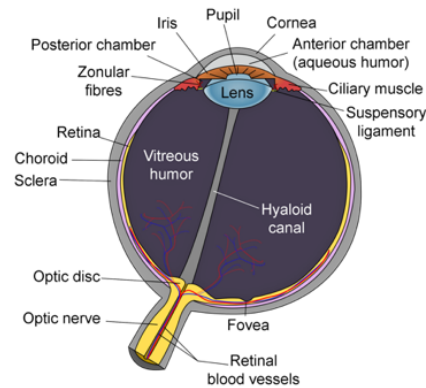
The Human Eye

Vision is a complex process that requires numerous components of the human eye and brain to work together. Vision is defined as the ability to see the features of objects we look at, such as color, shape, size, details, depth, and contrast. A sketch of the anatomical components of the human eye is shown here.The main structures are the iris, lens, pupil, cornea, retina, vitreous humor, optic disk and optic nerve.

ision is not uniform across our entire field of view. The area if highest acuity is called the \emph{fovea}. This is a small area at the back of the retina densely packed with cones. Humans constantly re-position our eyes so the image is focused on this foveal regions.

The Human Eye
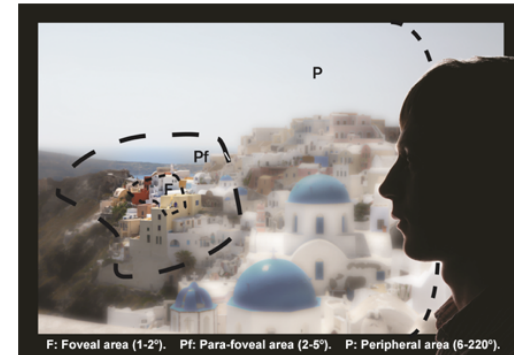
F: Foveal area (1-2°).   Pf: Para-foveal area (2-5°).   P: Peripheral area (6-220°).

https://featuredcontent.psychonomic.org/keeping-andromeda-peripheral-tracking-multiple-targets-out-of-the-corner-of-your-eye/

Vision is not uniform across our entire field of view. The area if highest acuity is called the FOVEA. This is a small area at the back of the retina densely packed with cones. Humans constantly re-position our eyes so the image is focused on this foveal regions. Areas of a scene projected outside the foveal region peripheral vision has much lower acuity. As can be seen above at any moment in time only a very small region of the viewing scene appears in high resolution, with gradual decrease in acuity as distance from the fovea increases. For this reason, we must keep moving our eyes to maintain focus in the scene.

Generally eye-tracking data is analyzed in terms of fixations and saccades. Saccades are rapid eye movements used to position the gaze. During each saccade, visual acuity is suppressed and the visual system is in effect blind. Only during fixations is clear vision possible. The brain virtually integrates the visual images that we acquire through successive fixations into a visual scene or object.

A Brief History of Eye Tracking
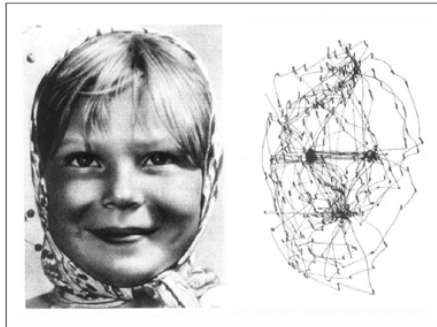
Eye tracking refers to techniques used to record and measure eye movements. Recent years have seen a rapid evolution of eye-tracking technology with systems becoming cheaper, easier to operate and less intrusive on the viewer.

 Eye-tracking systems first emerged in the early 1900, with early studies by Yarbus showing gaze patterns over still images. Until the 1980s, eye-trackers were primarily used to collect eye movement data during psychophysical experiments. This data was typically analyzed after the completion of the experiments. During the 1980s, the benefits of real-time analysis of eye movement data were realized as eye-trackers evolved as a channel for human-computer interaction. More recently, real-time eye-tracking has been used in interactive graphics applications and large scale display systems to improve computational efficiency and perceived quality. With the (re) advent of VR systems, the market has seen many commercially available VR HMDs with integrated eye tracking, including the VIVE PRO.

Types of Eye Movements

Saccades

Rapid Eye Movements to quickly reposition the eye

Saccades are rapid eye movements that allow the observer to quickly re-position the fovea (region of highest visual acuity) to the next region of interest. During saccades the observer is essentially blind due to the poor quality of perceived information. While most saccades are triggered automatically, they can be voluntarily triggered. The latency of the saccade is task dependent and can vary from 100ms to 1000ms, while the range of saccade duration lies between 20ms and 40 ms. There is a linear relationship between the duration and amplitude of a saccade, this is intuitive because larger movements take longer to complete. Once a saccade is involved the end point cannot be changed. Most often, but not always, the trajectory of a saccade is linear.

Fixations
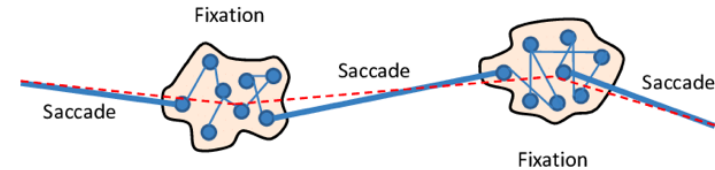
Fixations occur when the viewer rests their gaze on a specific scene location or region of interest. Essentially fixations happen when the viewer stops scanning the scene and holds their gaze at a certain point of regard. Holding the eye steady allows the visual system to receive more information about the area of interest, and glean more detail about a location or object. Eye trackers typically record a spatial (x, y) gaze location. Fixations have a \emph{duration} in addition to a spatial component.

Fixations are used to indicate attention, visibility and cognitive processing. For example, higher fixation duration may correspond with an area of high saliency that is considered more engaging by the user, or conversely an area that takes more processing to make sense of.

A fixation comprises smaller, slower movements, called microsaccades, tremor and drift, that allow the viewers eye to align with the target areal. Fixation duration can vary between 150 to 600 milliseconds.

https://www.researchgate.net/publication/313805602_SaccadeLenses_Interactive_Exploratory_Fltering_of_Eye_Tracking_Trajectories

Smooth Pursuit
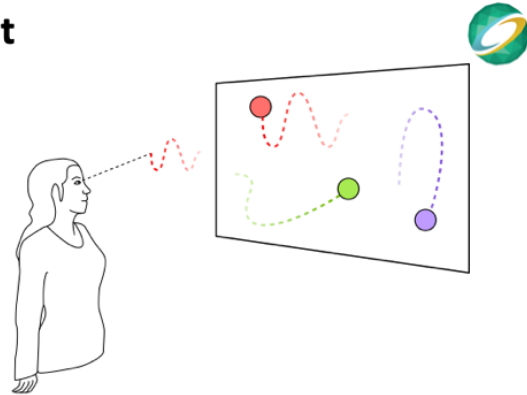
In contrast to saccades, smooth pursuits are much slower movements of the eyes designed to track the fovea to a moving target stimulus. Smooth pursuits are voluntary eye movements in that the viewer can decide whether or not to track a moving object. It is very rare for smooth pursuit to be triggered voluntarily, in absence of a moving target. Typical eye velocity during smooth pursuit is 30 degrees per second, although velocities as high as 100 degrees per second are achievable by some individuals. Once the target moves faster than is possible to engage in smooth pursuit i.e. greater than 30 degrees per second, humans will use saccades to maintain focus on the target.

Vergence
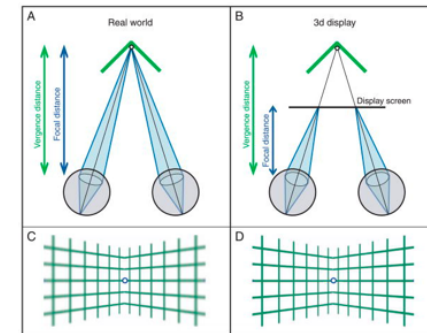
Vergence
> Eyes converge so that object is on the fovea for each eye
> May be initiated by disparity cues (object not in fovea for one of the eyes) or accommodation cues (presence of blur in one of the eyes)

The image above and following text is excerpted from "Hoffman DM, Girshick AR, Akeley K, Banks MS. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J Vis*. 2008;8(3):1–30. Published 2008"

Vergence and focal distance with real stimuli and stimuli presented on conventional 3D displays. (A) The viewer is fixated and focused on the vertex of a hinge. Vergence distance is the distance to the vertex. Vergence response is the distance to the intersection of the eyes' lines of sight. Focal distance is the distance to which the eye would have to focus to create a sharply focused image. Accommodative response is the distance to which the eye is accommodated. (B) The viewer is fixated on the simulated hinge vertex on a computer display screen. Vergence distance is the same as in panel A. Focal distance is now the distance to the display. (C) The appearance of the stimulus when the viewer is accommodated to the vertex of a real hinge. In the retinal image, the joined planes (the sides) of the hinge are blurred relative to the vertex. (D) Appearance when the viewer is accommodated to the vertex of a simulated hinge. The sides and the vertex are equally sharp in the retinal image.

Immersive Virtual Environments

How Immersive Eye Trackers Work

Available Hardware

Available Hardware

Looxid with EEG

There are many other examples in the eye tracking space including Varjos Bionic Display, SMI…

Available Hardware

Device	Eye Image Resolution	Sample Rate (Hz)	Cost (USD)
7invensun	-	120	$200
FOVE VR HMD	320 x 240	120	$599
aGlass and aSee	-	120-380	-
Pupil Labs VR (VIVE USB)	320 x 240	30	$1,572*
Pupil Labs VR (Dedicated USB)	640 x 480	120	$1,572*
Pupil Labs AR (Hololens)	640 x 480	120	$1,965*
Pupil Pro Glasses	800 x 600	200	$2,066?*
Pupil Pro Glasses	800 x 600	200	$2,066?*
Looxid Labs	-	-	$2999
Hololens v2	-	-	$3500
Tobii Pro Glasses 2	240 x 960	100	$10,000

Special Considerations for Eye tracking in VR

sa2019.siggraph.org

VR Relevant Parameters

Rotation within socket (Gaze In Head)

3D Point of Regard (Gaze in World)

3D Point of Regard (Gaze In World)

Rotating head (Head In World)

(Eyes rotated within the head's coordinate frame)

(Head rotated within the global coordinate frame)
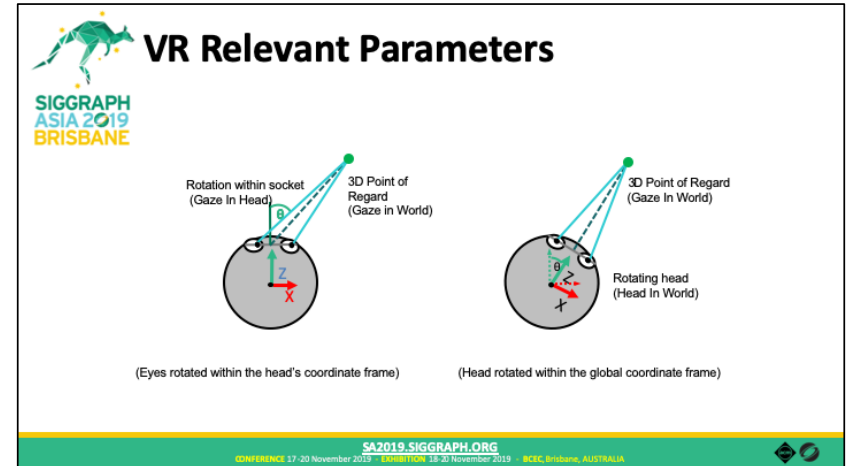
Special Considerations for VR

Koulieris et al (2017) SIGGRAPH. Accommodation and Comfort in Head Mounted Displays

Collecting your own Data
An overview of hardware available

sa2019.siggraph.org
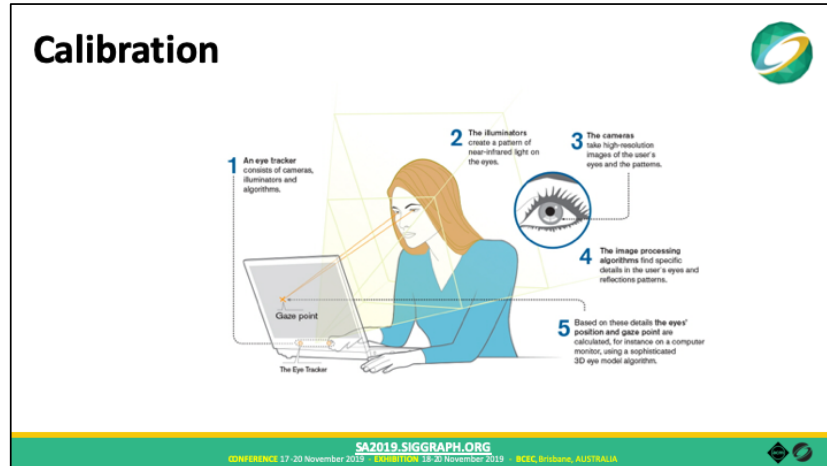
During the calibration the user is asked to look at specific points on the screen, also known as calibration dots. This enables the eye tracker to build a correspondence between where the user is looking in the scene and the position of the eye. A good calibration is crucial for robust, accurate data collection. In VR eye tracking calibration is conducted in the headset. The user is asked to look at certain 3D objects so a similar correspondence is established.

IMAGE FROM TOBII.COM

https://www.youtube.com/watch?v=eBINRku82hA

Most commercial eye trackers available today provide a mechanism to export the data. Raw gaze data enables the experimenter to perform independent, tailored analysis. While some eye trackers will export common measures such as number and duration of fixations, most in depth studies rely on raw data. An example of raw data is shown here.
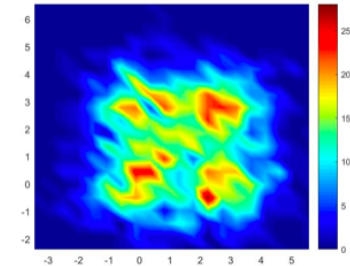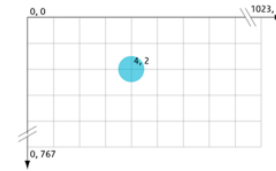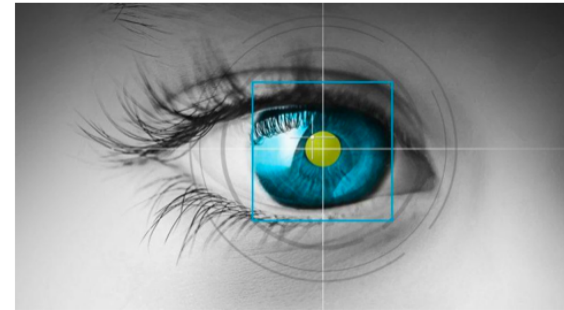
It's a good idea to do a sanity check on your data and graph it roughly to ensure the captured eye movement data maps to the space you expect. For example – its always good to check the image, screen dimensions and resolution. You also want to check screen origin to ensure the image is not flipped.

Data Measures

While eye tracking data is often straightforward to capture, making sense of the data is often not so simple. There are many ways to analyze and report the data.

Visualizing the distribution of the gaze points results in images such as the one shown here. These are called **Heatmaps**. Heatmaps are shown as a color gradient overlaid on the image or scene observed. The colors (red, yellow, green) indicate the density of the gaze points in decending order with RED showing the most attended areas (hotspots) and green the lowest.

VR Heatmap
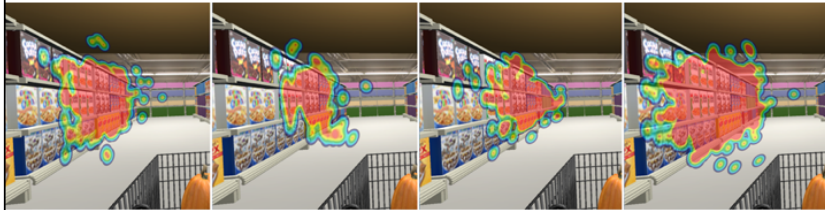
VR Heatmap
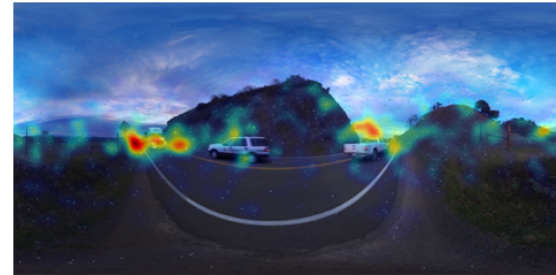
John, Raiturkar, Le Meur, Jain. A Benchmark of Four Methods for Generating 360 degree Saliency Maps from Eye Tracking Data, AIVR 2018

Hard to visualize – equirectangular projection

Fixation Analysis

- Time to first Fixation
- Fixation Duration
- Number of Fixations

Areas of Interest

The Time to First Fixation (TTFF) records time that it takes a respondent (or all respondents on average) to first fixate. This can provide information on how viewer prioritizes scene understanding.

Absolute Fixation Duration can indicate interest, or confusions.

Number of Fixations tells how how many fixations occurred at a specific point in a scene.

Fixation order

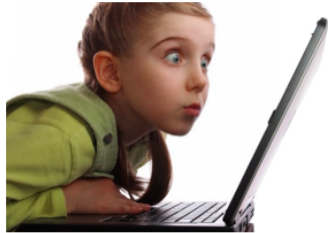http://neurorelay.com/2016/09/11/eye-tracking-in-neuromarketing-research/

Others

- Number of times gaze returned
- Number of Respondents (as a %) that were drawn to an AOI

Existing Studies

sa2019.siggraph.org

https://research.nvidia.com/publication/perceptually-based-foveated-virtual-reality.
SIGGRAPH 2019 ET
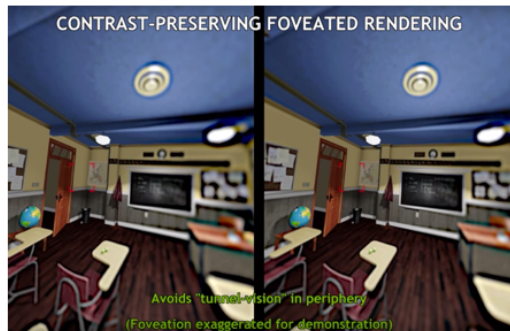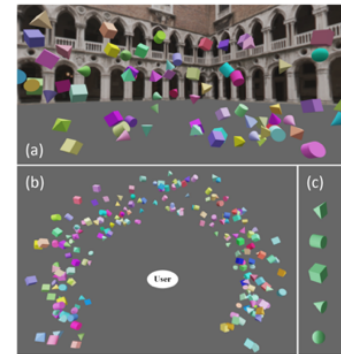
Steve Grogorick, Michael Stengel, Elmar Eisemann, and Marcus Magnor. 2017. Subtle gaze guidance for immersive environments. In Proceedings of the ACM Symposium on Applied Perception (SAP '17), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, Article 4, 7 pages. DOI: https://doi.org/10.1145/3119881.3119890

Trends and a Look to the Future

According to a Forbes Article (https://www.forbes.com/sites/solrogers/2019/02/05/seven-reasons-why-eye-tracking-will-fundamentally-change-vr/#7551c02e3459) Eye tracking will lead to more flexible HMDS, Enhanced Interactions, Increased Comfort (IPD), increased security, simplified profile management, deeper insights and more natural interaction.

Questions

sa2019.siggraph.org

This material is based upon work supported by the National Science Foundation CAREER Award under Grant No. IIS-1253432