

# **Fovea-styrd rendering: En analys av tekniken inom virtuell verklighet**

Lucas Liljekvist 1901595

Kandidatavhandling i datateknik

Handledare: Jan Westerholm

Fakulteten för naturvetenskaper och Åbo Akademi

2023

# Innehåll

<b>1</b>	<b>Introduktion-bör fixas</b>	<b>1</b>
1.1	Virtuell verklighet . . . . .	1
1.1.1	Utmaningar . . . . .	1
1.1.2	Lösningar . . . . .	2
1.2	Fovea-styrd rendering . . . . .	2
1.2.1	Nya utmaningar . . . . .	3
<b>2</b>	<b>Ögonspårning och tillämpningar</b>	<b>4</b>
2.0.1	Ögats anatomi . . . . .	4
2.0.2	Ögonspårning . . . . .	5
2.0.3	Stereoskopi . . . . .	5
2.0.4	Oculär parallax . . . . .	6
2.0.5	Latens . . . . .	6
<b>3</b>	<b>Varianter av foveastyrd rendering</b>	<b>7</b>
3.0.1	Guenter et al . . . . .	7
3.0.2	Problem med periferi-området . . . . .	8
3.0.3	DeepFovea . . . . .	8

# Kapitel 1

## Introduktion-bör fixas

### 1.1 Virtuellt verklighet

Virtuellt verklighet eller den vanligare benämningen VR som kommer från engelskans virtual reality har varit en teknologi som växt i popularitet under det senaste decenniet men trots dess komplexitet så fanns en VR-hjälm redan år 1968.<sup>1</sup> Ivan Sutherland hade byggt en hjälm som lät en se virtuella objekt i den fysiska världen, detta är dock inte virtuellt verklighet utan förstärkt verklighet (engelskan AR eller augmented reality). Skillnaden mellan virtuellt verklighet och förstärkt verklighet är att medan förstärkt verklighet lägger till virtuella objekt som man ser igenom glasögon eller en kamera till vår fysiska värld så ser man en helt virtuell som inte behöver ha någon koppling alls till vår värld. Det bör också tilläggas att det är tre dimensionella objekt eller världar som man kan se med VR och AR teknologier. Men vad Sutherland byggde var bara en början till VR-glasögon och det var först under 2010-talet som virtuellt verklighet teknologin blev kommersiellt möjligt med Oculus Rift. Det gjordes försök att producera VR-glasögon tidigare men diverse brister förhindrade att det skulle bli till något. Till dessa hör att skärmar med ett tillräckligt stort antal pixlar inte kunde reduceras till en tillräckligt liten storlek. Och att det inte fanns tillräckligt med datorkraft för att producera virtuella världar i realtid.

#### 1.1.1 Utmaningar

Även om Oculus Rift fungerade och gjorde framsteg så var teknologin fortfarande inte bra, för att göra virtuellt verklighet mer tilltalande så behöver VR-glasögonen fungera som glasögon. Det vill säga att rörelser man gör ska återskapas i det glasögonen visar. För det behövs spårning av glasögonens position kunna följas av en dator görs antingen

---

<sup>1</sup>Jay David Bolter, Maria Engberg och Blair MacIntyre. ”2 The History of Reality Media”. I: *Reality Media: Augmented and Virtual Reality*. 2021, s. 23–41.

med kameror eller basstationer som skickar signaler till sensorer i glasögonen som kan användas för att kalkylera var de befinner sig och vart de är på väg. Två andra områden som behöver förbättras är synfältet som glasögonen ger in i den virtuella världen och upplösningen. Människan har ett synfält på 210 grader horisontellt och 150 grader vertikalt. VR-glasögonen Valve Index har jämförelsevis ett horisontellt synfält på 120 grader och 100 grader vertikalt vilket innebär att man får ett mindre synfält med VR-glasögonen på. Upplösningen behöver även höjas till en nivå där man inte längre kan se pixlarna som utgör skärmarna som används. Om då både upplösningen och den del av synfältet som skärmarna täcker bidrar till ett behov av kraftfullare datorer. Utöver det så behöver VR-glasögon rendera två olika bilder, ett för varje öga. Vilket är ett ineffektivt sätt att hantera resurser eftersom fovean som är området var synen i mänskliga ögon har mest detaljer bara täcker 4% av pixlarna i ett par VR-glasögon från 2016<sup>2</sup>. Vilket gör att mer effektiv rendering bidrar till att minska på datorkrafterna som behövs.

### 1.1.2 Lösningar

Ett sätt att öka effektivitet är att rendera vad som ses i en lägre upplösning och antingen med hjälp av algoritmer eller artificiell intelligens öka upplösningen vilket ger en lite sämre bildkvalité än vad det vore om man renderar på den högre upplösningen men arbetet minskar. En annan metod är rendering med fixerad fovea eller engelskans fixed foveated rendering. Här tas ögats fovea i beaktande och en cirkel i mitten renderas med högre upplösningen medan allt runtomkring har en betydligt lägre upplösning<sup>3</sup>. Detta kan ge en prestandaökning mellan 10 och 25 procent. Bristerna med den här metoden är att området inte kan röra på sig och måste var större för att inte göra det svårt att se saker och sedan är periferiområdet alltid i en lägre upplösningen och det kräver då att huvudet vrids så att objekt kommer i området med högre upplösning.

## 1.2 Fovea-styrd rendering

Vidareutvecklingen av rendering med fixerad fovea är då fovea-styrd rendering eller engelskans foveated rendering. Fovea-styrd rendering använder sig av ögonspårning för att veta vart en användare riktar sina ögon och då kan datorn flytta på området med hög upplösning för att matcha vart på skärmen ögonen är riktade. Med hjälp av detta så kan de yttersta delarna av ögats synfält renderas i en mycket lägre upplösning än

---

<sup>2</sup>Susmija Jabbireddy m. fl. *Foveated Rendering: Motivation, Taxonomy, and Research Directions*. 2022. arXiv: 2205.04529 [cs.GR].

<sup>3</sup>Kyle Orland. *How valve got passable VR running on a four-year-old graphics card*. 2016. URL: <https://arstechnica.com/gaming/2016/03/how-valve-got-passable-vr-running-on-a-four-year-old-graphics-card/>.

vad som man kunde göras med fast fovea rendering i och med att detta område i ögat är minst detaljrikt.

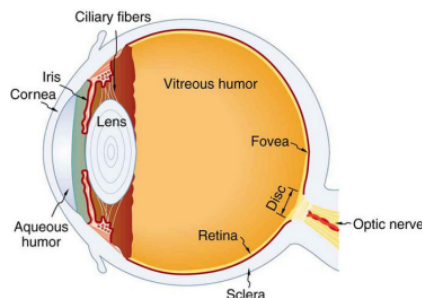
### **1.2.1 Nya utmaningar**

Däremot så finns det några bra anledningar till varför teknologin inte har blivit mer använd än i nuläget. Det största problemet ligger i ögonspårningen, det är svårt att exakt beräkna vart ögonen är riktade. Pupillernas storlek varierar beroende på ljusstyrka och storleken behöver inte vara den samma för båda ögonen. Utöver det så är människor olika och det bidrar till svårigheten i att göra teknologin funktionell för så många människor som möjligt. Ögonspårningssensorerna behöver då alltså även rymmas i VR-glasögonen så att de kan se ögonens riktning utan att ta upp utrymme som behövs för att ge ett stort synfält. Vikt, formfaktor och bekvämlighet behöver det med tas i beaktande. Så teknologin är komplex och det blir en balansgång för att få alla saker att fungera utan att priset blir för högt.

# Kapitel 2

## Ögonspårning och tillämpningar

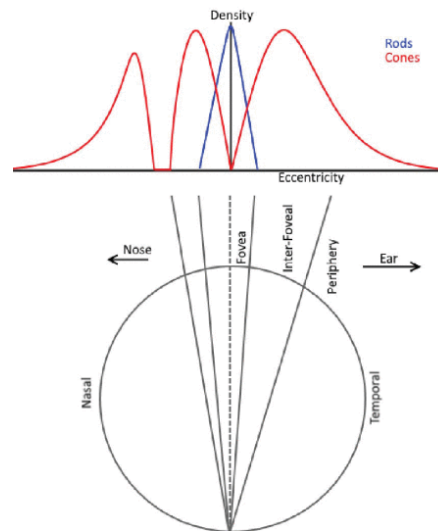
### 2.0.1 Ögats anatomi



Figur 2.1: Överblick av ögats anatomi, Foveated Rendering: Motivation, Taxonomy, and Research Directions

För en lite grundligare förklaring av vad som utnyttjas av ögat för foveastyrd rendering så hjälper en lite grundligare granskning av hur ögat fungerar. När ljuset träffar ögat kommer det igenom pupillen som reglerar mängden ljus som kommer in beroende på ljusnivån genom att bli större eller mindre. Ljuset bryts i hornhinnan och linsen och återger en bild på ögats näthinna med fokus vid fovean eller gula fläcken som den också kallas. Näthinnan innehåller synceller som skickar elektriska signaler via synnerven till hjärnans syncentrum. Det är två olika synceller i näthinnan som kallas för tappar och stavar. Tappar är känsliga för färg och fungerar bäst under bra belysning, det finns tre typer av tappar och de varierar i vilken färg de är känsliga för: rött, blått och grönt. De andra syncellerna som finns är stavarna som är bäst för mörkerseende i och med att de är mer ljuskänsliga än vad tapparna är även känsliga för rörelse. Fördelningen av tappar och stavar är inte jämn, det finns nästan enbart tappar vid fovean men desto längre ut vid näthinnan desto större mängd stavar än tappar. Fördelningen av tappar minskar mest direkt utanför fovean. Det finns inga tappar eller stavar vid synnerven

och den delen kallas även för blinda fläcken.



Figur 2.2: Hur syncellerna fördelas, Rendering Optimizations for Virtual Reality Using Eye-Tracking

## 2.0.2 Ögonspårning

Ögonspårning är ett måste för att uppnå foveastyrd rendering och påverkar i stor grad kvalitén av upplevelsen i VR när det kommer till just foveastyrd rendering. Tekniken bakom ögonspårningen när det gäller VR-glasögon är två infraröda kameror som får sitt videoflöde av ögonen genom linsen på glasögonen som beskrivs i artikeln av Hans-Martin Lutz et al.<sup>1</sup> De två kamerorna undviker att synnas genom linsen genom att använda en "hot mirror" som är en spegel som reflekterar infrarött ljus men släpper igenom den delen av ljusets spektrum som våra ögon kan se. Spegeln placeras mellan linsen och skärmen. Runt lenserna lägger man små infraröda LED-lampor som reflekteras på hornhinnan och hjälper med precision av kartläggningen över vart blicken är riktad. Avståndet av dessa till pupillen räknas och pupillen känns igen med maskin inlärning.

## 2.0.3 Stereoskopi

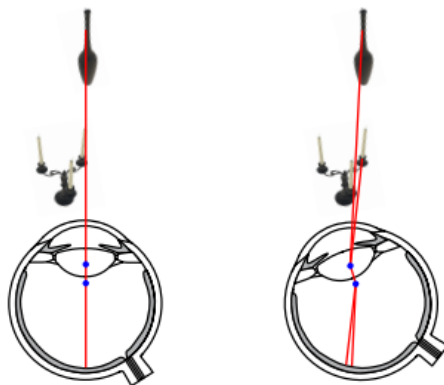
En del av vad som gör VR så övertygande är illusionen av djupseende, de två skärmarna som används visar enbart tvådimensionella bilder. Illusionen kommer ifrån att de två bilderna visar samma sak men från lite olika synvinklar. Synvinklarna matchar hur det

<sup>1</sup>Otto Hans-Martin Lutz m. fl. I: *Current Directions in Biomedical Engineering* 3.1 (2017), s. 53–56. DOI: doi:10.1515/cdbme-2017-0012. URL: <https://doi.org/10.1515/cdbme-2017-0012>.

skulle se ut för ögonen när de enskilt skulle betrakta något objekt. När de två skärmarna i VR-glasögonen betraktas så har man djupseende. Stereoskopi-tekniken används även med 3D-filmer men då visas de två perspektiven på samma skärm och diverse tekniker med glasögon, den minst komplicerade av dem är då anaglyf-glasögon. Pappersglasögonen som har en röd lins och en cyan lins för att skilja de två färglagren på anaglyfbilder.

Stereoskopi-tekniken har brister när det kommer till VR och det är när föremål betraktas på nära håll. Avståndet mellan pupillerna (english IPD = Interpupillary distance) minskar då matchar inte synvinklarna i skärmarna med ögonens och allt tappar fokus. När föremål observeras på längre håll sker ingen betydelsefull förändring i pupillernas avstånd till varandra. Bekymret är då som mest besvärligt när man försöker läsa texter som är små.

## 2.0.4 Oculär parallax



Figur 2.3: Beskrivning av oculär parallax, Gaze-Contingent Ocular Parallax Rendering for Virtual Reality

Det finns många detaljer som bidrar till en mer övertygande upplevelse i VR och en av de mindre detaljerna är oculär parallax Konrad et al.<sup>2</sup> Även om huvudet hålls stilla så flyttar den bild som projiceras på näthinnan en aning när man flyttar på blicken som kan ses i figur 2.3. Effekten beror på att ögats center för rotation inte befinner sig på samma ställe som ögats center för projektion utan de skiljer sig från varandra med ett avstånd som varierar mellan 7 till 8 mm. Normalt är den virtuella miljön statisk men med ögonspårning och oculär parallax rendering rör miljön på sig enligt ögats rörelser.

## 2.0.5 Latens

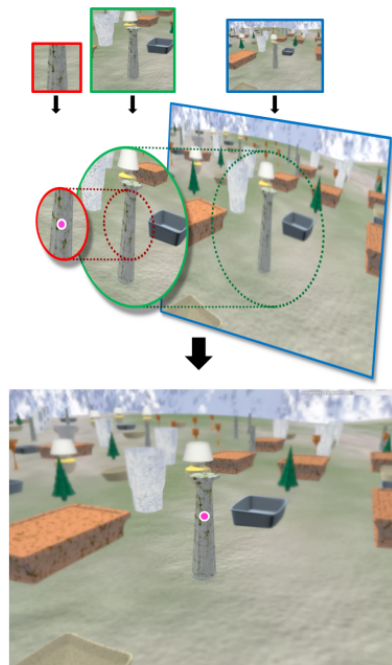
<sup>2</sup>Robert Konrad, Anastasios Angelopoulos och Gordon Wetzstein. "Gaze-Contingent Ocular Parallax Rendering for Virtual Reality". I: *ACM Trans. Graph.* 39 (2 2020).



# Kapitel 3

## Varianter av foveastyrd rendering

### 3.0.1 Guenter et al



Figur 3.1: Behövs mer information här, Foveated 3D Graphics

En av de tidigaste moderna implementeringarna av foveastyrd rendering har gjorts av Guenter et al.<sup>1</sup> Implementation gjordes inte med några VR-glasögon utan det var en datorskärm med upplösningen 1920x1080 och en uppdateringsfrekvens på 120Hz. För ögonspårning användes en Tobii TX300 med en uppdateringsfrekvens på 300Hz med en latens under 10ms. Metoden Guenter et al. använde sig av var att rendera tre lager med olika upplösningar som interpoleras till skärmens inbyggda upplösning enligt figur 3.1. Det minsta och översta lagret använder sig av skärmens samplingsfrekvens,

<sup>1</sup>Brian Guenter m. fl. "Foveated 3D Graphics". I: *ACM Trans. Graph.* 31.6 (2012). ISSN: 0730-0301. DOI: 10.1145/2366145.2366183. URL: <https://doi.org/10.1145/2366145.2366183>.

det mellersta har en lägre samplingsfrekvens och det yttre och lägsta lagret täcker hela skärmen och är hela tiden stationär medan de två andras lagrens position beror på data från ögonspåraren och har den lägsta samplingsfrekvensen och upplösning. Upplösningen i periferiområdet kan inte bara sänkas för det märks att området är väldigt kantigt isåfall så därför behövs kantutjämning eller anti-aliasing. Det finns saker som behöver förbättras med den här metoden, delar av skärmen kan renderas upp till tre gånger på grund av de olika lagren överlappas vilket är något som har en stor betydelse. Sedan när man tänker på VR så är inte utrustningen som användes något man kan använda i VR-glasögon men ögonspårare var fortfarande en ny teknik år 2012 och VR var inte ändamålet.

### 3.0.2 Problem med periferi-området

Det viktigaste med fovea-styrd rendering är att det inte märks att tekniken blir använd och att den är såpass effektiv att det är värt att använda sig av den. Problemen med att rendera periferi-området som nämndes tidigare är att det lätt märks. Flimmer och kanter är saker drar till sig uppmärksamhet och om man suddar ut omgivningen för mycket känns omgivningen till fovean som en tunnel. Latensen behöver även vara under 50 ms för att det inte skall vara suddigt för en kort stund när blicken rör på sig.

### 3.0.3 DeepFovea

Meta, tidigare Facebook, använder sig av en annan metod.<sup>2</sup> Alla pixlar renderas inte och fovea-området renderas helt men ju längre bort från fovean desto glesare finns det renderade pixlar. Istället för att rendera periferin i en lägre upplösningen så återuppbyggs bilden av djup maskininlärning med de glesare renderade pixlarna som referens och data som nätverket är tränat på. Det här ger en kompression på uppemot 14 gånger av RGB-video utan märkbar minskning av kvalitet.

---

<sup>2</sup>Anton S. Kaplanyan m. fl. *DeepFovea: Neural reconstruction for foveated rendering and video compression using learned statistics of natural videos - meta research*. 2019. URL: <https://research.facebook.com/publications/deepfovea-neural-reconstruction-for-foveated-rendering-and-video-compression-using-learned-statistics-of-natural-videos/>.

# Källförteckning

- Bolter, Jay David, Maria Engberg och Blair MacIntyre. "2 The History of Reality Media". I: *Reality Media: Augmented and Virtual Reality*. 2021, s. 23–41.
- Guenter, Brian m. fl. "Foveated 3D Graphics". I: *ACM Trans. Graph.* 31.6 (2012). ISSN: 0730-0301. DOI: 10.1145/2366145.2366183. URL: <https://doi.org/10.1145/2366145.2366183>.
- Jabbireddy, Susmija m. fl. *Foveated Rendering: Motivation, Taxonomy, and Research Directions*. 2022. arXiv: 2205.04529 [cs.GR].
- Kaplanyan, Anton S. m. fl. *DeepFovea: Neural reconstruction for foveated rendering and video compression using learned statistics of natural videos - meta research*. 2019. URL: <https://research.facebook.com/publications/deepfovea-neural-reconstruction-for-foveated-rendering-and-video-compression-using-learned-statistics-of-natural-videos/>.
- Konrad, Robert, Anastasios Angelopoulos och Gordon Wetzstein. "Gaze-Contingent Ocular Parallax Rendering for Virtual Reality". I: *ACM Trans. Graph.* 39 (2 2020).
- Lutz, Otto Hans-Martin m. fl. I: *Current Directions in Biomedical Engineering* 3.1 (2017), s. 53–56. DOI: doi:10.1515/cdbme-2017-0012. URL: <https://doi.org/10.1515/cdbme-2017-0012>.
- Orland, Kyle. *How valve got passable VR running on a four-year-old graphics card*. 2016. URL: <https://arstechnica.com/gaming/2016/03/how-valve-got-passable-vr-running-on-a-four-year-old-graphics-card/>.