

Institut für Visualisierung und Interaktive Systeme

Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Bachelorarbeit

Wahrnehmungsorientiertes Volumen-Rendering

Ruben Bauer

Studiengang:	Informatik
Prüfer/in:	Prof. Dr. Thomas Ertl
Betreuer/in:	Valentin Bruder M.Sc., Dipl.-Inf. Christoph Schulz, Dr. Steffen Frey
Beginn am:	12. Mai 2018
Beendet am:	12. Oktober 2018

Kurzfassung

Diese Arbeit handelt um Wahrnehmungsorientiertes Volumen-Rendering. Das menschliche Auge ermöglicht dem Menschen seine visuelle Wahrnehmung, welche in foveales und peripheres Sehen unterteilt werden kann. Das foveale Sehen ist detailliert, scharf und farbig und befindet sich im Zentrum des visuellen Wahrnehmungsbereiches, während das periphere Sehen sich außerhalb des Zentrums befindet und unschärfer und weniger farbig ist. In dieser Arbeit wird speziell diese Eigenschaft des menschlichen Sehapparates im Zusammenhang mit Volumen-Rendering untersucht. Dafür werden unterschiedliche Ansätze betrachtet, die Bildqualität im peripheren Bereich zu senken, um die Performanz des Volumen-Renderings zu erhöhen und gleichzeitig die Qualität der Darstellung zu erhalten oder zu verbessern. Die Daten zur Ermittlung des fovealen beziehungsweise peripheren Bereichs werden mit einem Eye-Tracking Gerät gemessen und fließen direkt in die Darstellung ein.

Inhaltsverzeichnis

1	Einleitung	7
2	Grundlagen	9
2.1	Related Work	9
2.2	Sehapparat	12
2.3	Raytracing	13
2.4	Volumenrendering und Transferfunktion	13
2.5	Eyetracking	14
2.6	GPU Architektur	14
3	Entwurf	15
3.1	Projekt	15
3.2	Arbeitspakete und Integration	15
4	Implementation	17
5	Ergebnisse	19
6	Zusammenfassung und Ausblick	21
	Literaturverzeichnis	23

1 Einleitung

Einleitung zu dieser Arbeit schreiben.

Gliederung der Arbeit kurz beschreiben.

2 Grundlagen

Dieses Kapitel beschäftigt sich mit den Grundlagen für diese Arbeit. Der erste Abschnitt des Kapitels beschäftigt sich mit verwandten Arbeiten zum Thema Wahrnehmungsorientiertes Volumen-Rendering. Der zweite Abschnitt handelt über die Grundlagen des menschlichen Sehapparates. Hier werden die Fähigkeiten und Limitierungen der visuellen Wahrnehmung des Menschen diskutiert. In Abschnitt drei, wird die Funktionsweise von Raytracing erläutert, welches ein grundlegender Algorithmus, der für diese Arbeit zugrunde liegender Implementierung ist. Zusammenhängend mit Raytracing wird im Abschnitt vier, die Verwendung des Raytracers für das Volumenrendering erläutert. Abschnitt fünf diskutiert die Auswahl des für diese Arbeit zugrunde liegenden Eyetrackers und dessen Verwendung für die Erfassung des fovealen und peripheren Bereichs. Aus Performanzgründen kann es hilfreich sein für Berechnungen auf einer GPU, die Architektur der GPU zu betrachten und unter Umständen Algorithmen für eine bessere Effizienz anzupassen. Diese Thematik wird in Abschnitt sechs behandelt.

Beschreibung der Grundlagen überarbeiten

2.1 Related Work

Wahrnehmungsorientiertes Volumenrendering ist kein absolut neues Arbeitsgebiet und ist schon Teil einiger wissenschaftlicher Arbeiten gewesen

Beispiele

. Da diese Arbeit auf zwei trennbare Aspekte beruht, unterteile ich den Related Work Abschnitt in zwei Teile. Der erste Abschnitt bezieht sich auf Arbeiten im Bereich des wahrnehmungsorientierten Renderings mit dem Ziel, die Performanz einer Anwendung zu steigern. Die Ansätze hier beziehen sich meist darauf, dass die Qualität der Darstellung im peripheren Bereich der visuellen Wahrnehmung, gesenkt wird und so für die Berechnung eines Bildes weniger Rechenleistung aufgewendet werden muss.

Was genau ist mit Performanz gemeint?

Der zweite Abschnitt bezieht sich auf Arbeiten zu wahrnehmungsorientiertem Volumenrendering, mit dem Ziel, die Qualität der Darstellung zu erhöhen. Dabei werden vor allem Ansätze zur geschickten Anpassung von Parametern einer Transferfunktion vorgestellt, die dem Betrachter ein insgesamt besseres Verständnis der Volumendaten ermöglichen soll.

2.1.1 Performanz- und Wahrnehmungsorientiertes Volumenrendering

In einem Paper von Marc Levoy und Ross Whitaker „Gaze-Directed Volume Rendering“ [LW90] erforschten diese Methoden, wie Eyetracking-Daten in Rendering-Algorithmen eingesetzt werden können. Das Ziel ihrer Forschung war es, dem Nutzer einen Arbeitsplatz für Echtzeit-Volumen-Rendering, mit der Illusion eines hochauflösenden Bildes über den gesamten Bildschirm, zu präsentieren, welches durch Eyetracking unterstützten Verfahren einen geringeren Berechnungsaufwand als herkömmliche Volumen Renderer hat. Dafür präsentierten sie eine Implementierung eines Ray-Tracers für Volumen Daten, in welcher mit Hilfe der Eyetracking-Daten der Blickfokus auf dem Bildschirm berechnet wurde und um diese Position herum, die Anzahl der Strahlen und die Samples pro Strahl, abhängig von der Distanz zum Blickfokus, angepasst wurden. Die Implementierung basiert auf darauf, dass der Detaillgrad der visuellen Wahrnehmung des Menschlichen Auge nur in einem kleinen, zentralen Bereich, der Fovea, am höchsten ist und zu den Rändern des Blickfelds hin, der periphere Bereich, stark abnimmt. Ausgehend davon, berechneten sie in dem von dem Nutzer fokussierten Bereich, das Bild mit der vollen Auflösung und einer hohen Abtastung der Strahlen. In dem restlichen Bereich reduzierten sie die Auflösung des Bildes und abhängig von dem Abstand eines Pixels zum Blickfokus, die Abtastung eines Strahls. In ihrer Implementierung verwendeten sie 2D und 3D mip maps, einen Eye Tracker und die Pixel-Planes 5 rendering engine, ein hoch paralleles Raster Display System. Um die geringere Abtastung in der Peripherie gut nutzen zu können, wurde aus den 3D Volumendaten eine 3D mip map erstellt. Die 3D mip map wird durch den Ray-Tracing Algorithmus abgetastet und durch trilineare Interpolation eine 2D mip map erstellt. Die 2D mip map ist Grundlage für die Erstellung des endgültigen Bildes. In ihren Ergebnissen konnten sie die Rendering Kosten für ein teilweise hochauflösendes Bild im Vergleich zu einem vollständig hochauflösenden Bildes, um bis den Faktor fünf, senken.

In einer Arbeit von Guenter et. Al. „Foveated 3D Graphics“ [GFD+12], wurde ähnlich zu dem Paper von Levoy und Whitaker, der Abfall der visuellen Auflösung des Auges außerhalb des visuellen Zentrums, der Fovea, ausgenutzt, um eine beschleunigte Berechnung des Bildes zu erhalten. Im Gegensatz zu der davor genannten Arbeit, ist das darunterliegende System hier kein Ray-Tracer für Volumendaten, sondern eine Grafikpipeline für 3D Szenen. Das Bild wird hier aus drei Teilbilder zusammengesetzt, welche jeweils mit unterschiedlichen Auflösungen berechnet wurden und wie Schichten übereinander gelegt werden. Die drei Schichten sind: innere Schicht, mittlere Schicht und äußere Schicht. Die innere Schicht hat ungefähr die Größe des fovealen Bereichs auf dem Bildschirm und wird in der maximalen Auflösung berechnet. Ihr Mittelpunkt ist der Blickfokus des Betrachters. Die mittlere Schicht ist ein bisschen größer als die innere Schicht und wird mit einer niedrigeren Auflösung berechnet. Sie wird ebenfalls auf den Blickfokus des Betrachters zentriert. Die äußerste Schicht überdeckt den gesamten Bildbereich und wird in der niedrigsten Auflösung berechnet. Um die Schichten zusammensetzen zu können, werden die mittlere und äußere Schicht jeweils zur nativen Auflösung des Bildschirms, die gleiche Auflösung wie die innere Schicht, interpoliert. Scharfe Kanten zwischen den Schichten werden dadurch vermieden, dass diese sich leicht überlappen und spezielle Blend-Masken verwendet werden, um die verschiedenen Schichten glatt übereinander zu blenden. In ihrer Arbeit nahmen sie sich auch das Problem an, dass das starke Unterabtasten, um bis zu dem Faktor sechs in jede Dimension, in der mittleren und äußeren Schicht zu störenden und sich bewegenden Artefakten führen kann. Um dies entgegen zu wirken, verwendeten sie drei Antialiasing Techniken: Hardware Multi-Sample Antialiasing (MSAA), „Temporal Reprojection“ und „whole frame jitter sampling“. Um die Blickposition zu erfassen nutzten sie den Tobii Tx 300 Eye Tracker mit einer Worst-Case Latency von 10 ms. Das Bild wurde auf einem Computer mit Intel Xeon CPU

(E5640 mit 2.67 Ghz) und einer NVidia GeForce GTX 580 GPU berechnet. Dargestellt wurde es auf einem 23"1920x1080 LCD Monitor mit einer Bildwiederholungsrate von 120 Hz. Mit ihrem System und Techniken erlangten sie eine Performanz-Verbesserung von einem Faktor zwischen fünf und sechs auf einem Monitor mit HD Auflösung. Dabei erreichten sie eine Darstellungsqualität, die vergleichbar mit dem Rendern eines hochauflösenden Bildes über den gesamten Bildschirm ist.

2.1.2 Wahrnehmungsorientiertes Volumenrendering zur Qualitätssteigerung

R. Englund und T. Ropinski stellten in ihrem Paper „Quantitative and Qualitative Analysis of the Perception of Semi-Transparent Structure in Direct Volume Rendering“ [ER] verschiedene Techniken, zur Verbesserung der Wahrnehmung von komplexen volumetrischen Daten, vor. In einer Studie mit über 300 Teilnehmern untersuchten sie, wie diese Techniken zur verbesserten Wahrnehmung von Volumen beitragen und verglichen die verschiedenen Ansätze miteinander. Dabei mussten die Teilnehmer jeweils kleine Aufgaben absolvieren, so dass Rückschlüsse darauf geschlossen werden können, wie sehr eine gewisse Technik dem Teilnehmer bei der Erkennung von Form und Tiefe eines Objekts in einem Volumen beiträgt. Um eine bessere direkte Erforschung der Volumen-Daten später ermöglichen zu können, wurden Techniken, die automatisch Rendering-Parameter angepasst haben hier ausgelassen. Nur wenn der Nutzer selbst interaktiv sein kann, also die Parameter des Volumen-Rendings, wie die Transferfunktion und Kamera selbst anpassen kann, ermöglicht dies eine direkte Erforschung. In ihrer Studie haben sie sechs Techniken ausgewertet. Darunter „Direct Volume Raytracing“ (DVR) als grundlegende Technik und „Depth Darkening“ wobei Tiefeneffekte ähnlich zu „Ambient Occlusion“ dadurch hervorgerufen werden, dass tiefere Objekte dunkler gezeichnet werden. In ihrer Auswertung kamen sie unter Anderem dazu, dass Techniken, die natürliche Lichteffekte in den Volumendaten simulieren, deutliche Vorteile gegenüber die anderen getesteten Ansätze gezeigt haben.

Anders als zu den untersuchten Techniken von Englund und Ropinski [ER] präsentieren Aidong Lu et. Al. in ihrem Paper „Volume Composition Using Eye Tracking Data“ [LME06] eine Methode für automatisierte Parameterauswahl bei der Betrachtung von Volumen Daten mit Hilfe eines Eye Tracking Gerätes. Das Ziel, welches sie mit dieser Methode verfolgen ist es, die mühsame Nutzerinteraktionen bei der Auswahl der Parameter zu vereinfachen und damit die Nutzbarkeit des Darstellungssystems zu verbessern. Volumen Daten können sehr komplex sein und daher ist es oft schwierig herauszufinden, was der Nutzer in dem Volumen betrachten möchte und was dahingehend hervorgehoben werden soll. Trotzdem ermöglichen Eigenschaften des Volumen Renderings, wie konstante Größen, Formen und Positionen der Objekte eine automatische Anpassung der Parameter. Um die Bereiche, die für den Nutzer von Interesse sind, zu bestimmen, wird ein Eyetracker zur Hilfe genommen. Dieser misst die Augenbewegungen und die Blickposition auf dem Bildschirm. Es wird zwischen zwei Hauptsächlichen Augenbewegungen unterschieden: Sakkade und Fixation. Eine Sakkade ist eine schnelle Augenbewegung von einem Punkt zu einem anderen. Bei einer Fixation ruht das Auge auf einem Punkt. Die Punkte, die fixiert werden, sind meist für den Nutzer von Interesse. Da die Blickposition durch den Eye Tracker nur auf einer 2D-Ebene bestimmt werden kann, wird durch ein konstantes Rotieren des Volumen Objekts und der parallelen Aufzeichnung der Augenbewegung versucht, die 3D Position des fixierten Objektes zu rekonstruieren. Aus den Eyetracking und Volumendaten bestimmen sie mit Hilfe mehrerer Clustering-Methoden gewichte für die einzelnen Voxel des Volumens und berechnen so die Wichtigkeit der Objekte innerhalb des Volumens für den Nutzer. Entsprechend dieser Ergebnisse wurden die Render Parameter angepasst,



Abbildung 2.1: Modell des visuellen Wahrnehmungsapparates des Menschen aus [doi:10.1111/cfg.13150]

um die für den Nutzer am interessantesten Objekte hervorzuheben und anzuzeigen. Aidong Lu et. Al. kamen zu dem Schluss, dass die präsentierte Methode den Aufwand für den Nutzer, die Render Parameter anzupassen, signifikant reduzieren kann. Trotzdem kann ein solcher regelbasierter Ansatz nicht mit einer manuellen Einstellung der Render-Parameter mithalten.

Related Work nochmal anschauen. Zum Bsp. ob aus der richtigen Perspektive geschrieben wurde.

2.2 Sehapparat

Das Auge ist der visuelle Sensor des Menschen und ermöglicht ihm das Sehen. Wahrnehmungsorientiertes Rendering nutzt gezielt Eigenschaften der visuellen Wahrnehmung des Menschen aus. Dafür ist es notwendig, ein gutes Verständnis des menschlichen visuellen Wahrnehmungsapparates zu besitzen. In diesem Abschnitt stelle ich einige Eigenschaften des visuellen Wahrnehmungsapparates vor, wie sie in [doi:10.1111/cgf.13150] vorgeführt werden.

2.1 ist ein Modell des visuellen Wahrnehmungsapparates des Menschen. Das Modell unterteilt den Sehapparat des Menschen in Optik, Sensorik, Motorik, Verarbeitung, Speicherung und Aufmerksamkeit. Licht trifft auf die Augen und wird durch die Optik auf die Retina, die Sensorik, weitergeleitet. Hier wird der visuelle Input abgetastet und gefiltert. Dabei entstehen zwei Datenströme welche die Verarbeitung stereoskopischer Bilder über einen großen Blickwinkel mit unterschiedlichen räumlich unterschiedlicher Auflösungen ermöglichen. Die Retina (oder auch Netzhaut) ist mit dem visuellen Kortex verbunden. Die Signale werden über die visuellen Nerven komprimiert und zum visuellen Kortex transportiert. Dort werden sie von verschiedenen Bereichen im Gehirn verarbeitet. Speicherung und Aufmerksamkeit spielen dabei eine wesentliche Rolle.

Das visuelle Wahrnehmungssystem des Menschen hat wesentliche Limitierungen, die bei der Darstellung von Bildern gezielt genutzt werden können. Die Sehschärfe des ungleich auf der Retina verteilt und nur in in einem kleinen, zentralen Bereich ist die Sehschärfe maximal. Dieser Bereich wird Fovea oder auch Gelber Fleck genannt. Je weiter man sich von der Fovea nach außen hin entfernt, desto mehr nimmt die Sehschärfe ab. Der Bereich um die Fovea ist der periphere Bereich des Auges. Das Sichtfeld des visuellen Wahrnehmungsapparates des Menschen ist bei gerade

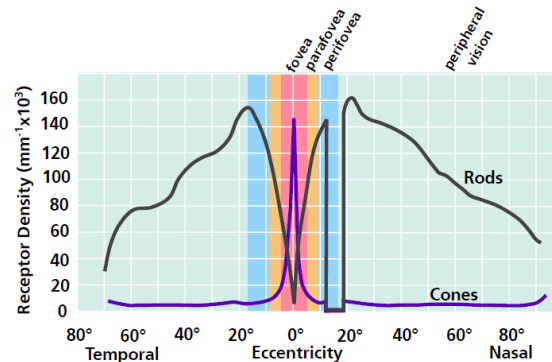


Abbildung 2.2: Verteilung der Photorezeptoren auf der Retina [doi:10.1111/cfg.13150]

gerichteten Blick horizontal bis circa 190° und mit Augenrotation bis zu 290° . Visuelle Reize werden über das gesamte Sichtfeld wahrgenommen. Abhängig von dem zuständigen Bereich auf der Retina gibt es starke Unterschiede, wie die visuellen Reize über das Sichtfeld verteilt, verarbeitet werden. Durch das Verkleinern (Miosis) und Vergrößern (Mydriasis) der Pupille wird die Menge des einfallenden Lichts in das Auge gesteuert. Die Pupille nimmt dabei Größen zwischen 2mm und 8mm Durchmesser an.

Die Netzhaut (Retina) ist die photosensitive Schicht des Auges und besteht aus zwei Typen von Photorezeptoren, aus Zapfen und Stäbchen. Es sind circa $6 \cdot 10^6$ Zapfen und ungefähr 20 mal so viele Stäbchen auf der Retina verteilt. Stäbchen sind für die Helligkeitswahrnehmung verantwortlich. Zapfen sind für die Farbwahrnehmung verantwortlich. Man unterscheidet zwischen drei Zapfentypen: L-Zapfen für lange Wellenlängen, M-Zapfen für mittlere Wellenlängen und S-Zapfen für kurze Wellenlängen.

Die Fovea ist der Bereich um circa $5,2^\circ$ um das Zentrum der Retina und besteht fast ausschließlich aus Zapfen. Die Anzahl der Zapfen nimmt aber nach außen hin stark ab. Der Bereich von circa $5,2^\circ$ bis 9° wird als Parafovea bezeichnet. Der Bereich zwischen circa 9° bis 17° heißt Perifovea. Fovea, Parafovea und Perifovea sind für die zentrale Sicht verantwortlich. Alles außerhalb ist die periphere Sicht.

Grundlagen zum menschlichen Sehapparat hier.

2.3 Raytracing

Grundlagen zu Raytracing hier.

2.4 Volumenrendering und Transferfunktion

Grundlagen zum Volumenrendering und der zugehörigen Transferfunktion hier.

2.5 Eyetracking

Grundlagen zum Eyetracking hier.

2.6 GPU Architektur

Grundlagen zur GPU Architektur hier.

3 Entwurf

Entwurfkapitel schreiben.

3.1 Projekt

Allgemeine Visual Studio Projektbeschreibung. Architektur und Funktionsweise des Projekts.

Raycaster

Beschreibung der Funktionsweise des Raycasters für das Volumenrendering in dem Projekt.

3.2 Arbeitspakete und Integration

Vorüberlegungen zur Umsetzung und daraus entstandene Arbeitspakete beschreiben.

4 Implementation

Implementation der drei Ansätze.

5 Ergebnisse

Messbare Ergebnisse der Arbeit. (Performanzmessungen..).

Diskussion

Diskussion und Bedeutung der Ergebnisse.

6 Zusammenfassung und Ausblick

Hier bitte einen kurzen Durchgang durch die Arbeit.

Ausblick

...und anschließend einen Ausblick

Literaturverzeichnis

- [ASF16] The Apache Software Foundation. *Apache ODE™ – The Orchestration Director Engine*. 2016. URL: <http://ode.apache.org>.
- [Dud01] Duden. *Die Deutsche Rechtschreibung*. 22., völlig neu bearbeitete und erweiterte Auflage. Aktualisierter Nachdruck. Dudenverlag, 2001. ISBN: 3-411-04012-2.
- [ER] R. Englund, T. Ropinski. „Quantitative and Qualitative Analysis of the Perception of Semi-Transparent Structures in Direct Volume Rendering“. In: *Computer Graphics Forum* 37.6 (), S. 174–187. DOI: [10.1111/cgf.13320](https://doi.org/10.1111/cgf.13320). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13320>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13320> (zitiert auf S. 11).
- [GFD+12] *Foveated 3D Graphics*. ACM SIGGRAPH Asia, Nov. 2012. URL: <https://www.microsoft.com/en-us/research/publication/foveated-3d-graphics/> (zitiert auf S. 10).
- [LME06] A. Lu, R. Maciejewski, D. S. Ebert. „Volume Composition Using Eye Tracking Data“. In: *Proceedings of the Eighth Joint Eurographics / IEEE VGTC Conference on Visualization*. EUROVIS’06. Lisbon, Portugal: Eurographics Association, 2006, S. 115–122. ISBN: 3-905673-31-2. DOI: [10.2312/VisSym/EuroVis06/115-122](https://doi.org/10.2312/VisSym/EuroVis06/115-122). URL: <http://dx.doi.org/10.2312/VisSym/EuroVis06/115-122> (zitiert auf S. 11).
- [LW90] M. Levoy, R. Whitaker. „Gaze-directed Volume Rendering“. In: *Proceedings of the 1990 Symposium on Interactive 3D Graphics*. I3D ’90. Snowbird, Utah, USA: ACM, 1990, S. 217–223. ISBN: 0-89791-351-5. DOI: [10.1145/91385.91449](https://doi.org/10.1145/91385.91449). URL: <http://doi.acm.org/10.1145/91385.91449> (zitiert auf S. 10).
- [RVA16] H. Reijers, I. Vanderfeesten, W. van der Aalst. „The effectiveness of workflow management systems: A longitudinal study“. In: *International Journal of Information Management* 36.1 (Feb. 2016), S. 126–141. DOI: [10.1016/j.ijinfomgt.2015.08.003](https://doi.org/10.1016/j.ijinfomgt.2015.08.003).
- [WCL+05] S. Weerawarana, F. Curbera, F. Leymann, T. Storey, D. F. Ferguson. *Web Services Platform Architecture : SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More*. Prentice Hall PTR, 2005. ISBN: 0131488740. DOI: [10.1.1/jpb001](https://doi.org/10.1.1/jpb001).

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Ort, Datum, Unterschrift