

Blickpunkt-adaptive Verschiebung der Stereo-3D-Konvergenzebene

Stefan Eickelberg, Lehrstuhl für Kommunikationstechnik, TU Dortmund, Dortmund, Deutschland, stefan.eickelberg@tu-dortmund.de

Kurzfassung

Die Diskrepanz zwischen Akkomodation und Konvergenz beim Betrachten von Stereo-3D-Inhalten führt bei vielen Menschen zu Kopfschmerzen und weiteren Symptomen der Erschöpfung des visuellen Systems. In der Stereo-3D-Produktion werden verschiedene Maßnahmen ergriffen, diesen Effekt möglichst gering zu halten. Dies führt jedoch nicht immer zum gewünschten Erfolg oder bringt Nachteile mit sich. In diesem Dokument wird daher ein Verfahren vorgeschlagen und untersucht, welches den Konflikt zwischen Akkomodation und Konvergenz in eine andere Domäne transformiert und die Diskrepanz auf diesem Wege eliminiert. Das Verfahren funktioniert in der Art, dass die Konvergenzebene immer zu der Tiefe des momentan betrachteten virtuellen Objekts verschoben wird. Dabei ist die Konvergenzebene die virtuelle Tiefenebene, die keine Disparität hat. Das Verfahren basiert auf Blickpunktdetektion und Signalverarbeitung in Kombination mit Disparitätskarten. Dieser Artikel beschreibt die ersten Schritte in diesem Projekt.

1. Einleitung

Abgesehen von einigen eher experimentellen Geräten zeigen alle heutigen 3D-Displays einen Effekt, der als *Akkommodation-Konvergenz-Diskrepanz* (AKD) bezeichnet wird [1]. Dabei handelt es sich um einen Konflikt zwischen zwei grundlegenden Mechanismen der menschlichen visuellen Wahrnehmung: Der Anpassung der Brennweite (*Akkommodation*) und der Auslenkung der Augen (*Konvergenz*). Der Konflikt liegt darin begründet, dass die Augen eines Stereo-3D-Betrachters auf die Displayebene fokussiert sein müssen, während sie gleichzeitig auf einen stereoskopischen 3D-Stimulus konvergieren. Liegt dieser vor oder hinter der Displayebene, so entsteht der besagte Konflikt, wie **Bild 1** verdeutlicht. Der Konflikt ist unnatürlich. Er ist aber bis zu einem gewissen Grad vom visuellen System tolerierbar. In starker Ausprägung wird die AKD allerdings in vielen Quellen als einer der Hauptgründe für visuelles Unbehagen (Engl. *Visual Discomfort*) und die Erschöpfung des visuellen Systems (Engl. *Visual Fatigue*) genannt [2]. Um diese subjektiv respektive objektiv messbare negative Beeinflussung des visuellen Systems zu vermeiden und eine hohe *Quality of Experience* zu gewährleisten, müssen bei der Stereo-3D-Produktion Maßnahmen ergriffen werden, die AKD gering zu halten. Dies ist allerdings nicht immer möglich. Noch dazu wird das Tiefenbudget einer Szene durch diese Maßnahmen deutlich reduziert, was von Konsumenten häufig kritisiert wird. Es besteht daher großes Interesse an Signalverarbeitungsmethoden und Displays, mit denen kongruente Akkomodation und Konvergenz erreicht werden kann. Eine solche Signalverarbeitungsmethode wird im Rahmen dieser Arbeit vorgestellt und untersucht. Dazu wird in **Abschnitt 2** zunächst auf Maßnahmen zur Reduktion der AKD und deren Wirksamkeit eingegangen. Dabei werden auch einige Grundlagen vermittelt, die nötig sind, um das entwickelte Verfahren in **Abschnitt 3** zu erläutern. Das Verfahren wird in **Abschnitt 4** diskutiert und in **Abschnitt 5** wird diese Arbeit zusammengefasst und ein Ausblick auf weitere Forschung gegeben.

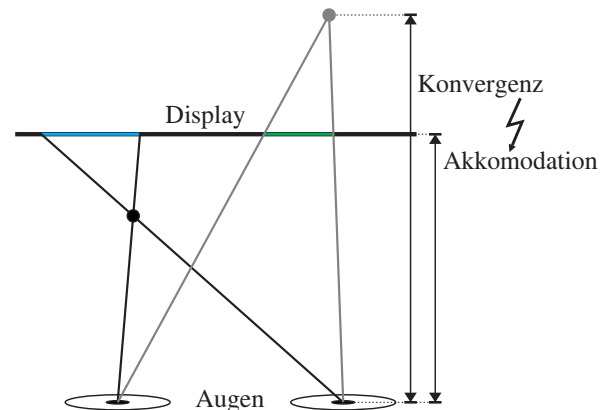


Bild 1 Verdeutlichung der Akkomodation-Konvergenz-Diskrepanz an zwei Beispielstimuli mit hervorgehobener gekreuzter (blau) und ungekreuzter Disparität (grün).

2. Maßnahmen zur Reduktion der Akkomodation-Konvergenz-Diskrepanz

2.1. Wahl der Aufnahmeparameter

Die Reduzierung der AKD durch geringeres Tiefenbudget ist gleichbedeutend mit einer geringen Disparität. Diese wird bei der Aufnahme durch die Szenenstruktur und die Distanz zwischen den Stereokameras bestimmt. Um diese Distanz zu dimensionieren haben Stereographen pragmatische Faustregeln formuliert (zum Beispiel *1/30th Rule of Thumb of 3D* [3] oder *Percentage Rule* [4]), die aber in der Regel wenig allgemeingültig und daher eher ineffektiv sind. Auch die Wissenschaft hat sich dieser Problematik zugewandt. Eine generell anerkannte Empfehlung lautet, die präsentierten Disparitäten auf komfortable Dimensionen – die sogenannte *Zone of Comfort* – zu beschränken [2]. Die Zone of Comfort wurde durch verschiedene Auto-

ren (z.B. Lamboij, Percival, Sheard, Shibata [2][4]) mit Hilfe von subjektiven Testreihen untersucht und definiert. Doch auch bei diesen wissenschaftlichen Empfehlungen ist die Umsetzbarkeit nicht zwingend gegeben. Die Parameter Displaygröße und Betrachtungsdistanz sind zum Produktionszeitpunkt nicht bekannt, sie skalieren jedoch maßgeblich die Wiedergabedisparität. Zudem ist es in Live-Übertragungen oder Echtzeit-Rendering-Anwendungen aufgrund der unvorhersehbaren Tiefenverteilung einer Szene nicht möglich, die maximale Disparität zu bestimmen und das Aufnahmesystem entsprechend zu dimensionieren.

2.2. Verschiebung der Konvergenzebene

Bilder, die mit einer parallelen Stereokamera-Konfiguration aufgezeichnet wurden, sind nur in der Unendlichkeit parallaxefrei. Werden diese Bilddaten unbearbeitet auf einem stereoskopischen Display dargestellt, so befindet sich die „Unendlichkeit“ aufgrund ihrer Nulldisparität in der Displayebene. Die gesamte virtuelle 3D-Szene befindet sich demnach in einem näherungsweise pyramidenförmigen Volumen zwischen Display und Betrachter. Dabei entsteht meist nicht nur eine unrealistische Tiefendarstellung (z.B. *Puppettheater Effect* [5] und pyramidenförmige Szenenverzerrung), sondern auch eine starke AKD. Ein praktisch überall eingesetzter Ansatz um dies zu vermeiden ist, aufmerksamkeitsregende Szeneninhalte in einem Nachverarbeitungsschritt in die Displayebene zu verschieben. Dies wird als Konvergenzebenenverschiebung (KEV) (Engl. *Reconvergence* oder *Depth Grading*) bezeichnet [1][3]. Ein zu verschiebendes Objekt habe die Disparität d zwischen den Stereoansichten. Um die Verschiebung nun durchzuführen, muss diese Disparität kompensiert werden. Also wird die rechte Ansicht um $d/2$ nach rechts und die Linke Ansicht um $d/2$ nach links verschoben. Das Objekt hat nun keine Disparität und liegt damit in der Displayebene, wie **Bild 2** verdeutlicht. In anderen Worten: Die Konvergenzebene wurde zu einer anderen virtuellen Tiefe verschoben. Aus diesem Prinzip ist ersichtlich, dass nach Anwendung dieser Prozedur sowohl negative als auch positive Disparitäten existieren. In **Bild 2** äußert sich dies in der vertauschten Reihenfolge der linken und rechten Ansicht. Die Disparitäten werden abhängig davon, ob die Blickrichtungsvektoren der Augen vor oder hinter der Displayebene aufeinander treffen, auch als gekreuzt oder ungekreuzt klassifiziert (vgl. **Bild 1**) [2]. Dabei erzeugen ungekreuzte Disparitäten 3D-Stimuli hinter der Displayebene. Für die ungekreuzten Disparitäten existiert eine einfache Einschränkung: Die Augen des Betrachters dürfen nicht divergieren [4], da dies zu Visual Fatigue führt. Die dargestellten ungekreuzten Disparitäten dürfen daher nie größer als der Abstand zwischen den Augen des Betrachters sein. Disparitäten in dieser Größenordnung liegen aber ohnehin außerhalb der in **Abschnitt 2.1** erläuterten Zone of Comfort und sollten daher vermieden werden.

Das Prinzip der KEV findet außerdem in dynamischer Form Anwendung, um Tiefensprünge bei Kameraschnitten

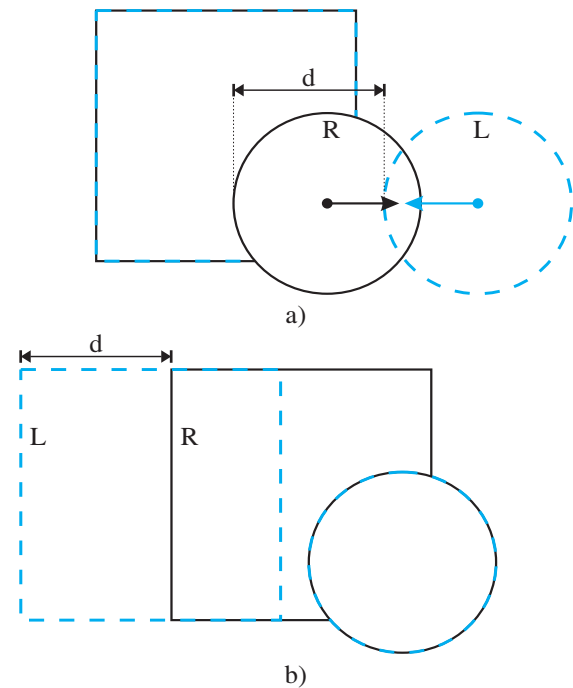


Bild 2 Prinzip der Konvergenzebenenverschiebung. a) Konvergenzebene verschoben zu quadratischem Hintergrundobjekt und b) zu kreisförmigem Vordergrundobjekt.

zu glätten (Engl. *Active Depth Cuts* [3]). Der Grund dafür ist, dass häufige Tiefensprünge das okulomotorische System stark belasten und daher ebenfalls zu Visual Fatigue führen [1][3][6]. Das Glätten funktioniert dabei wie folgt. Die Konvergenzebene wird vor und nach Kameraschnitten langsam so verschoben, dass die Aufmerksamkeitspunkte zweier Aufnahmen zum Zeitpunkt des Schnitts in der gleichen Tiefenebene liegen. Diese Verschiebung wird vom Betrachter bei geringen Verschiebungsgeschwindigkeiten kaum wahrgenommen [1][6]. Die Unsensitivität für KEV ist die Grundlage für das im Folgenden beschriebene Single-User Stereo-3D-Darstellungsverfahren.

3. Beschreibung des Verfahrens

Das Konzept der KEV aus **Abschnitt 2.2** kann dahingehend erweitert werden, dass die AKD vollständig eliminiert wird. Die Grundidee ist hier, die Konvergenzebene unter Verwendung von Disparitätskarten immer zum momentanen betrachteten Punkt (Blickpunkt) zu verschieben. Dabei wird die oben beschriebene Unsensitivität ausgenutzt. Dies führt dazu, dass der Blickpunkt immer disparitätsfrei ist und somit in der Displayebene liegt. Das Verfahren besteht aus fünf Schritten, wie sie in **Bild 3** dargestellt sind. Zunächst erfolgt eine Blickpunkterkennung. Für den gefundenen Blickpunkt wird die Disparität ausgelesen. Dabei müssen Okklusionsbereiche detektiert und gesondert behandelt werden. Da die KEV laut [6] vom Betrachter nur für langsame Verschiebungsgeschwindigkeiten nicht wahrgenommen wird, wird eine zeitliche Glättung der Folge der Disparitäten angewendet, bevor die Ver-

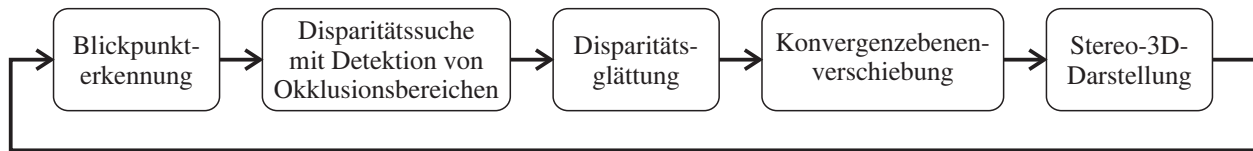


Bild 3 Verarbeitungsschritte des Verfahrens.

schiebung durchgeführt wird. Schlussendlich wird ein Stereo-3D-Frame für die Darstellung generiert und die Verarbeitung des nächsten Bilds beginnt. Diese Schritte werden nun im Detail erläutert.

3.1. Blickpunkterkennung

Eine Blickpunkterkennung kam in dieser Anfangsphase des Projekts noch nicht zum Einsatz, da das Ziel war, schnell erste Ergebnisse zu erhalten. Stattdessen wird zunächst nur ein bewegter Zeiger genutzt, um die Aufmerksamkeit des Betrachters zu lenken. Im späteren Projektverlauf soll dieses Modul dann durch die Blickpunkterkennung ersetzt werden. Dazu kommt zum Beispiel ein System in Frage, wie es in [8] beschrieben wurde.

3.2. Disparitätssuche, Okklusionsbehandlung und Disparitätsglättung

Bei der Disparitätssuche sind zwei Informationen zu berücksichtigen:

1. Ein intaktes menschliches visuelles System erzeugt bei der Wahrnehmung einer beliebigen dreidimensionalen Szene eine zentrale 3D-Ansicht, die sich genau zwischen den Augen befindet [2].
2. Es muss in der i -ten Iteration des Algorithmus davon ausgegangen werden, dass schon in der $(i-1)$ -ten Iteration eine KEV durchgeführt wurde.

Der Zeiger aus **Abschnitt 3.1** bzw. die gefundenen Blickpunktkoordinaten (x, y) befinden sich also demnach in der zentralen Ansicht und müssen in die Koordinaten der linken bzw. rechten Ansicht transformiert werden. Bei dieser Transformation handelt es sich um eine Umkehrung der Konvergenzebenenverschiebung. Es wird also eine Subtraktion bzw. Addition der letzten halben Konvergenzdisparität $d_{\text{Konv},i-1}/2$ auf die x -Koordinate des Blickpunkts angewendet. Dies wird auch in **Bild 2 a)** deutlich, wenn man diese Abbildung so interpretiert, dass keine KEV gegeben ist: Der Kreis im Vordergrund wird zentral wahrgenommen, sodass die Korrektur um $d/2$ nötig ist. Mit den transformierten Koordinaten kann die neue Konvergenzdisparität $d_{\text{Konv},i}$ dann aus der linken oder der rechten Disparitätskarte ausgelesen werden¹:

$$d_{\text{Konv},i} = d_{L,i} = d_{R,i} \quad (1)$$

¹ Die Abbildungsrichtung von der linken auf die rechte Ansicht bzw. umgekehrt wird hier aufgrund einer einfacheren Notation nicht im Vorzeichen der Disparitätskarte kodiert.

mit

$$d_{L,i} = d_L(x + d_{\text{Konv},i-1}/2, y), \quad (2)$$

$$d_{R,i} = d_R(x - d_{\text{Konv},i-1}/2, y). \quad (3)$$

Formel (1) ist allerdings in Okklusionsbereichen nicht gültig. In **Bild 2 a)** ist die Darstellung auf das Hintergrundobjekt mit der Disparität $d_{\text{Konv},i-1}$ konvergiert und das (flache) Vordergrundobjekt hat relativ zu dieser Konvergenzdisparität eine Differenzdisparität

$$d = d_v - d_{\text{Konv},i-1}, \quad (4)$$

mit der Vordergrundsdisparität d_v . Der gesamte mit d markierte Bereich stellt eine linksseitige Okklusion dar. Ebenso existiert auf der anderen Seite eine rechtsseitige Okklusion der gleichen Größe, die in dem Bild jedoch nicht markiert ist. Befindet sich nun ein Zeiger in einem dieser Okklusionsbereiche, zum Beispiel am Mittelpunkt des Kreises R in **Bild 2 a)**, so ergeben sich aus Formel (2) und (3) unterschiedliche Werte. Dabei entspricht ein Wert der Hintergrunddisparität und ein anderer der Vordergrundsdisparität. Diese Werte entsprechen gerade d_v und $d_{\text{Konv},i-1}$, wenn folgende zwei Annahmen zutreffend sind.

1. Die Disparität des Vordergrundobjekts ist im Okklusionsbereich näherungsweise konstant.
2. Die gefundene Hintergrunddisparität entspricht ungefähr der Konvergenzdisparität.

Diese Annahmen sind plausibel, da der Okklusionsbereich nur relativ klein ist und innerhalb von Objekten meist kontinuierliche flache Tiefenverläufe gegeben sind. Eine unmittelbare Folgerung aus diesen Annahmen ist, dass die Disparitätskarten keine Okklusionsartefakte enthalten dürfen. Mit Hilfe der gefundenen Disparitäten kann d nun über Formel (4) auf einfachem Wege bestimmt werden. Dieser Wert ist nötig, um zu entscheiden, ob Vorder- oder Hintergrunddisparität für die Konvergenzverschiebung verwendet werden soll. Das Umschalten auf den Vordergrund sollte vollzogen werden, wenn die entsprechende Objektgrenze in der zentralen 3D-Ansicht überschritten wurde. Dies ist am Beispiel der linken Ansicht äquivalent mit dem Prinzip, dass umgeschaltet wird, wenn die Blickpunktkoordinaten näher als $d/2$ am Vordergrundobjekt liegen. Beim Mittelpunkt des Kreises R in **Bild 2 a)** müsste demnach auf Vordergrund umgeschaltet werden.

Algorithmisch lässt sich dies, unter abermaliger Ausnutzung der obigen Annahmen, effizient und ohne konditionale Ausführungsblöcke umsetzen. Die Blickpunktkoordinaten werden lediglich um $d/2$ korrigiert, bevor die neue Konvergenzdisparität ausgelesen wird.

$$d_{\text{Konv},i} = d_L \left(x + d_{\text{Konv},i-1} / 2 + d/2, y \right) \quad (5)$$

Durch diese Maßnahme wird automatisch die Disparität des Vordergrundobjekts ausgelesen, wenn sich der Zeiger näher als $d/2$ an der Grenze dieses Objekts befindet. Außerhalb von Okklusionsbereichen gilt $d \approx 0$, sodass Formel (2) und (5) näherungsweise äquivalent sind.

Bisher wurde die zeitliche Glättung der Disparitätsfolgen nicht berücksichtigt. Hierzu wird ein einfaches rekursives Filter mit einem Gewichtungsfaktor α eingesetzt, sodass die i -te Konvergenzdisparität wie folgt korrigiert wird:

$$d_{\text{Konv},i} = \alpha \cdot d_L \left(x + d_{\text{Konv},i-1} / 2 + d/2, y \right) + (1 - \alpha) \cdot d_{\text{Konv},i-1} \quad (6)$$

3.3. Konvergenzebenenverschiebung und Stereo-3D-Darstellung

Das Prinzip der KEV wurde schon in **Abschnitt 2.2** erläutert: Die Ansichten werden gegenüber dem Ausgangszustand um $\pm d_{\text{Konv},i} / 2$ verschoben. Die Konvergenzdisparität $d_{\text{Konv},i}$ kann, insbesondere nach der zeitlichen Glättung, nicht-ganzzahlige Werte annehmen. Daher wird eine horizontale kubische Interpolation bei der KEV eingesetzt. Durch die Verschiebung würden in der Darstellungsphase links und rechts außen undefinierte (schwarze) Randbereiche mit Bildinhalten der jeweils anderen Ansicht überlagert sein. Um dies zu vermeiden werden diese Randbereiche der Breite $b = \max(|d_L(x,y)|, |d_R(x,y)|) / 2$ mit $(x,y) \in \mathbb{N}^2$ von der Darstellung ausgeschlossen, also abgeschnitten.

Schlussendlich werden die Ansichten nach entsprechender Antialias-Filterung und Unterabtastung in eines der gängigen Stereo-3D-Formate gebracht (z.B. *Side-by-Side*, *Over-Under*, *Row-interleaved* [3]) und auf einem Display ausgegeben.

4. Diskussion und Evaluierung

In diesem Abschnitt sollen einige Aspekte des oben beschriebenen Verfahrens diskutiert und evaluiert werden.

4.1. Betrachtung eines neuen Konflikts

Das Verfahren löst die AKD, da die Konvergenzdistanz immer der Displaydistanz entspricht. Bei genauer Betrachtung eines Blickpunktwechsels hin zu einem tiefenversetzten Objekt, wird jedoch Folgendes deutlich: Die AKD wird nicht einfach nur eliminiert, sondern in einen neuen Konflikt transformiert. Normalerweise geht mit einer Änderung der Aufmerksamkeit zu einem Objekt anderer Tiefe auch eine Änderung der Konvergenzdistanz der Augen einher. Diese Beziehung ist hier jedoch gebrochen, da das neue Objekt in die Displayebene verschoben wird. Eine Änderung der Konvergenzdistanz ist also nicht mehr erforderlich. Zur einfachen Notation wird dieser Konflikt im Folgenden mit dem Begriff *Tiefenänderung-Konvergenzdistanzänderung-Entkopplung* (TKE) referenziert.

Durch die Fixierung der Konvergenzdistanz ist die okulomotorische Beanspruchung des visuellen Systems im Prinzip gleich der Beanspruchung bei der Betrachtung eines 2D-Films. Dies begründet die hypothetische Verbesserung im Bezug auf Visual Comfort und Visual Fatigue. Welche Einflüsse die TKE allerdings auf diese Qualitätsparameter hat, ist unklar und soll im Rahmen dieses Projekts näher untersucht werden. In diesem Kontext ist insbesondere die Fragestellung interessant, ob trotz der prinzipiellen TKE Konvergenzimpulse generiert werden. Dies zu untersuchen ist mit diesem System und einer sowohl zeitlich als auch örtlich hochauflösenden Blickpunkterkennung für beide Augen prinzipiell möglich. Dazu muss gemessen werden, ob die Blickpunkte der Augen kurzzeitig horizontale Parallaxen aufweisen.

4.2. Zeiger vs. Blickpunkterkennung

Die Verwendung eines Zeigers anstelle einer echten Blickpunkterkennung hat neben der einfachen Umsetzbarkeit den Vorteil, dass etwaige Schätzfehler der Blickpunkterkennung umgangen werden. Ein Nachteil ist, dass keine sprunghaften Blickpunktwechsel (*Sakkaden*) analysiert werden können, da die Signalverarbeitung in Mensch und Maschine synchronisiert ablaufen muss und die endliche visuelle Reaktionsgeschwindigkeit des Menschen nicht deterministisch ist. Es können also nur relativ langsame Folgebewegungen des Auges untersucht werden.

Weiterhin stellt ein animierter Zeiger neben der eigentlichen 3D-Szene ebenfalls einen 3D-Stimulus dar, der vom Betrachter verfolgt wird und sich (in der Regel) nur in der Bildebene bewegt. Es ist daher nicht klar, ob ein Proband bei Einsatz einer Blickpunkterkennung die gleichen Ergebnisse liefern würde, wie bei dem hier verwendeten Zeiger.

In Okklusionsbereichen und durch die langsame KEV kann es außerdem zu einem unangenehmen Konflikt zwischen Okklusion und Disparität kommen. Die Zeigerdisparität suggeriert, dass der Zeiger im Hintergrund liegt, obwohl er ein Vordergrundobjekt ganz oder nur partiell überragt. Um dies zu vermeiden, wird zum einen neben der Szene auch der Zeiger mit einer stereoskopischen Tiefe dargestellt. Zum anderen wird pro Ansicht eine Sichtbarkeitsberechnung für jeden Bildpunkt des Zeigers durchgeführt. Dies ist in **Bild 4** zu erkennen.

4.3. Allgemeine Systemdiskussion

Aufgrund der Blickpunktadaptivität handelt es sich bei diesem System um eine Single-User-Anwendung. Ein Nebeneffekt der Blickpunktadaptivität und der TKE ist außerdem, dass die in **Abschnitt 2.2** erwähnte Beschränkung der ungekreuzten Disparitäten entfällt: Die Divergenz der Augen kann niemals eintreten, da sie bei diesem System immer auf die Displaydistanz konvergiert sind. Die „Zone of Comfort“-Beschränkung der Disparität wird gemäß **Abschnitt 2.1** durch die AKD bestimmt, die aber unter Verwendung dieses Systems nicht existiert. Eine Hypothese ist daher, dass mit Hilfe dieses Verfahrens ein größeres Tiefenbudget innerhalb der Fusionsgrenzen angenehm dar-

stellbar ist. Unterstützt wird diese Hypothese zusätzlich dadurch, dass bei peripherem Sehen größere Disparitäten fusioniert werden können [2] und bei diesem Verfahren auf dem Display meist nur im peripheren Sichtfeld Disparitäten ungleich Null dargestellt werden.

Neben der AKD sind *Window Violations* eine Quelle für Visual Discomfort [6]. Dabei handelt es sich um am rechten oder linken Displayrand abgeschnittene Objekte, die sich zwischen Display und Betrachter befinden. Die Tiefenhinweise Okklusion und Disparität stehen hier in einem unangenehmen Konflikt. Da Okklusion dominant ist, wird hier außerdem die Tiefe verzerrt [3]. In der Stereo-3D-Produktion wird dieser Konflikt mit verschiedenen Methoden vermieden [3]. Bei diesem System kann der Konflikt ersichtlich nur bei peripherem Sehen auftreten, da der Blickpunkt immer disparitätsfrei ist. Ob bzw. wie stark diese peripheren Window Violations die Quality of Experience negativ beeinflussen, wird in **Abschnitt 4.5** näher betrachtet. Ein weiteres Darstellungsartefakt am Displayrand ist die Aufdeckung von neuen Bildbereichen, was ebenfalls in dem referenzierten Abschnitt untersucht wird.

4.4. Testumgebung und Testmaterial

Wie in **Abschnitt 3.2** im Kontext der Okklusionsbehandlung erwähnt, dürfen die verwendeten Disparitätskarten keine Okklusionsartefakte enthalten. Echte Kameraaufnahmen als Testmaterial zu verwenden, ist daher kritisch, da dabei in der Regel Disparitätsschätzverfahren eingesetzt werden müssen. Diese Schätzverfahren basieren auf Korrespondenzanalyse. Da für den Hintergrund in Okklusionsbereichen keine Korrespondenzen in der jeweils anderen Ansicht gegeben sind, entstehen die unerwünschten Okklusionsartefakte. Im Rahmen dieser Arbeit wurden daher Testbilder mit perfekt konditionierten Disparitätskarten mit Hilfe der Rendering-Software *Blender* [9] synthetisiert. Die erzeugte virtuelle Testszene gemäß **Bild 4** weist sowohl flache Tiefengradienten als auch Tiefensprünge auf. Die Stereoansichten wurden mit verschiedenen Kameradistanzen synthetisiert, um auch die Auswirkungen von großen und kleinen Disparitäten untersuchen zu können. Die Testumgebung wurde so realisiert, dass Kameradistanz, Blickpunktzeigergeschwindigkeit und der Grad der zeitlichen Disparitätsglättung zur Laufzeit variiert werden können. So konnten die verschiedenen oben genannten Aspekte untersucht werden.

4.5. Evaluierung und Schlussfolgerungen

Im weiteren Verlauf dieses Projekts sind umfangreiche subjektive Tests geplant. Erste qualitative Testergebnisse werden jedoch schon in dieser Arbeit präsentiert.

Eine Hypothese ist, dass mit Hilfe dieses Systems ein größeres Tiefenbudget innerhalb der Fusionsgrenzen darstellbar ist. Ein zu großes Tiefenbudget führt im Regelfall zu unmittelbarem Visual Discomfort und/oder zu langsam akkumuliertem Visual Fatigue. Bei kleinen Tiefenbudgets kommt es sowohl mit als auch ohne das vorgestellte System nicht zu Visual Discomfort. Erst bei großen Disparitäten zeigen sich Unterschiede in der Wahrnehmung. Ohne

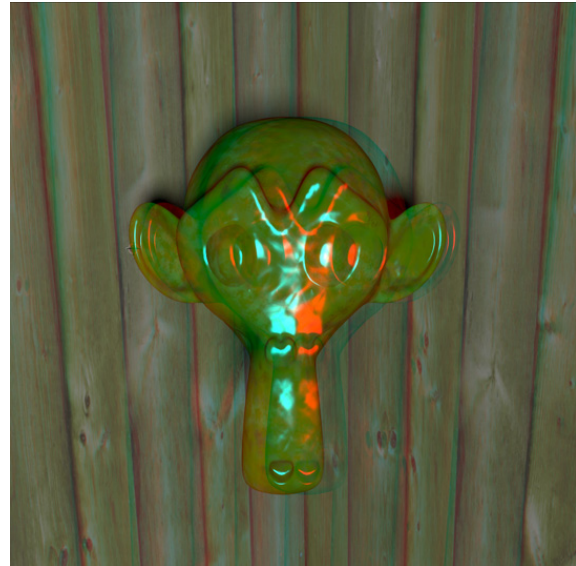


Bild 4 Anaglyphendarstellung der Testszene bei Konvergenz im Hintergrund und partieller Okklusion des Zeigers.

das System sind große Disparitäten äußerst kritisch und annähernd schmerzhaft, da große AKDen gegeben sind. Die Fusion von Objekten mit großer Disparität setzt komplett aus oder braucht sehr lange, was zu unangenehmem Doppelsehen führt. Dies ist in Okklusionsbereichen am schlimmsten, denn hier ergeben sich binokulare Rivalitäten, die Unbehagen hervorrufen. Mit dem vorgestellten System ist sichergestellt, dass der Blickpunkt immer fusioniert werden kann, sodass nur noch im peripheren Sichtfeld Probleme dieser Art entstehen können. Dort kommt es bei großen Differenzdisparitäten ebenfalls zu Doppelsehen, was nicht ganz so kritisch wie vorher, aber dennoch sehr störend ist. Abhilfe könnte an dieser Stelle eine künstliche Verringerung der Schärfentiefe schaffen. Es geht dann mit größerer Differenzdisparität eine stärkere Weichzeichnung einher und Unschärfe erleichtert die Fusion [2]. Es kommt ferner zu dem Phänomen, dass der Umschaltmoment von Hintergrund auf Vordergrund willkürlich erscheint. Dies liegt darin begründet, dass das Vordergrundobjekt aufgrund mangelnder Fusion nicht in der zentralen 3D-Ansicht des Betrachters vorhanden ist.

Ein wichtiger Aspekt des betrachteten Verfahrens ist die Wahrnehmbarkeit der dynamischen KEV. Bei kontinuierlichen, relativ flachen Tiefenverläufen ist die dynamische KEV so langsam, dass sie nicht detektiert wird. Bei Tiefensprüngen hingegen wird eine große KEV vollzogen. Diese ist bei hohen Verschiebungsgeschwindigkeiten (z.B. direktes Umschalten ohne zeitliche Glättung) stark wahrnehmbar sowie ungewohnt und daher störend. Dem lässt sich allerdings durch eine zeitliche Disparitätsglättung, also eine Verlangsamung der KEV, entgegenwirken. Dabei muss die Glättung umso stärker sein, je größer der Tiefensprung ist. Eine sinnvolle Erweiterung des Systems wäre daher, die zeitliche Glättung adaptiv zum Tiefensprung zu gestalten. Die Forderung großer zeitlicher Glättung ist jedoch bei sehr großen Tiefensprüngen problematisch. Die

Zielferparität sei hier so groß, dass sie die Fusionsgrenze überschreitet. Um eine schnelle Fusion am Blickpunkt sicherzustellen, sollte die KEV äußerst schnell durchgeführt werden, was im Widerspruch zu der Forderung einer nicht wahrnehmbaren langsamen KEV steht. Demnach ist das Tiefenbudget also auch unter Einsatz dieses Verfahrens auf die Fusionsgrenzen beschränkt. Eine weitere Beobachtung war, dass ein Betrachter bei uneglätteter KEV (direktes Umschalten) länger braucht, um die visuellen Informationen zu verarbeiten und einen neuen 3D-Eindruck zu erlangen. Dies könnte ein Indiz dafür sein, dass Nachbarschaftsinformationen bei der Erzeugung motorischer Konvergenzimpulse berücksichtigt werden. Genauso wäre es aber auch denkbar, dass sich diese Symptome zeigen, weil das visuelle System möglicherweise auch zeitlich zurückliegende Informationen ausnutzt, die aufgrund der ruckartigen Bildänderung nicht mehr korreliert werden können. Hier sind genauere Untersuchungen von Nöten. Als weitere Einflussgröße der KEV-Sensibilität stellte sich die Blickpunktzeigergeschwindigkeit heraus. Eine höhere Geschwindigkeit führte dabei zu einer geringeren Sensibilität gegenüber der KEV. Es kann also mit schwächerer zeitlicher Glättung gearbeitet werden. Aus dieser Beobachtung ergibt sich eine Schlussfolgerung, die für die Stereo-3D-Produktion von Interesse ist. Schnelle Blickpunktbewegungen resultieren aus schnellen Objektbewegungen. In Szenen mit vielen schnellen Bewegungen kann daher auch eine schnellere KEV durchgeführt werden, als in ruhigen Szenen. Durch eine gedankliche Extrapolation der obigen Beobachtung lässt sich außerdem die Hypothese formulieren, dass auch während und/oder kurz nach Sakkaden eine geringere Sensibilität gegenüber KEV gegeben ist. Sakkaden stellen schließlich lediglich sehr schnelle Blickpunktbewegungen dar. Um diese Hypothese zu überprüfen, bedarf es aber einer Blickpunktdetektion, wie in **Abschnitt 4.2** erwähnt wurde.

Die Abhängigkeit zur KEV-Geschwindigkeit trifft auch auf die Aufdeckung neuer Bildinhalte am linken und rechten Rand des Bildschirms zu. Der Unterschied ist hier, dass schnelle Randaufdeckungen zwar auffällig, aber nicht störend sind. Sehr wohl störend sind hingegen die häufig dabei entstehenden Window Violations. Ohne die Blickpunkt-adaptive KEV führen diese Konflikte in großer Ausprägung zu äußerst starkem Visual Discomfort. Unter Verwendung des thematisierten Verfahrens können zumindest im fovealen Sehen keine Window Violations auftreten, wohl aber im peripheren Sichtfeld. Dort werden sie deutlich weniger störend wahrgenommen. Um die Window Violations gänzlich zu eliminieren, wäre eine Erweiterung des Verfahrens um *Dynamic Floating Windows* denkbar [3][6]. Dabei wird das „stereoskopische Fenster“, durch das die 3D-Szene betrachtet wird, vor der Displayebene platziert, um den Konflikt zu beheben.

5. Zusammenfassung und Ausblick

In dieser Arbeit wurde ein Verfahren vorgestellt, welches die Diskrepanz zwischen Akkomodation und Konvergenz mittels Blickpunkt-adaptiver Konvergenzebenenverschiebung löst. Die dabei entstehende Entkopplung von Tiefenänderung und Konvergenzdistanzänderung der Augen sowie weitere Eigenschaften und Einschränkungen des Gesamtsystems wurden diskutiert und evaluiert. Es wurden außerdem einige potentielle Verbesserungsmöglichkeiten genannt, die in Zukunft näher untersucht werden. Zu den wichtigsten bisherigen Ergebnissen dieser Arbeit zählt die Erkenntnis, dass hohe Bildinhaltsgeschwindigkeiten zu einer geringeren Sensibilität gegenüber der Konvergenzebenenverschiebung führen. Im weiteren Verlauf des Projekts soll das System um eine hochauflösende Blickpunktdetektion erweitert und umfangreiche subjektive Tests durchgeführt werden.

6. Literatur

- [1] H. Tauer (2010): Stereo-3D. 1. Auflage. Berlin: Fachverlag Schiele & Schön.
- [2] M. Lambooy, W. IJsselsteijn, M. Fortuin, I. Heynderickx (2009): Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. In: Journal of Imaging Science and Technology, Vol. 53, Nr. 3.
- [3] B. Mendiburu: 3D Movie Making - Stereoscopic digital Cinema from Script to Screen. Focal Press, 2009.
- [4] T. Shibata, J. Kim, D. Hoffman, M. Banks (2011): The zone of comfort : Predicting visual discomfort with stereo displays. In: Journal of Vision, Vol. 11, Nr. 8.
- [5] L. Meesters, W. IJsselsteijn, P. Seuntiëns (2004): A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV. In: IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, Nr. 3, S. 381–391.
- [6] F. Dvernay, P. Beardsley (2010): Stereoscopic Cinema. In: R. Ronfard (Hg.), G. Taubin (Hg.): Image and Geometry Processing for 3-D Cinematography. 5. Auflage. Heidelberg: Springer Verlag, S. 11–52.
- [7] M. S. Banks, J. C. A. Read, R. S. Allison, S. J. Watt (2012): Stereoscopia and the Human Visual System. In: SMPTE Motion Imaging Journal, Vol. 121, Nr. 4, S. 24–43
- [8] M. Y. Al Nahlaoui (2010): Methoden und Konzepte der Blickrichtungserkennung. Dissertation Technische Universität Dortmund. Aachen: Verlag Shaker.
- [9] Blender Foundation: Blender. Download im Internet unter <http://www.blender.org>