



8

Herausgeber: Nikolaus Hottong

Stereoskope HD-Produktion

Messungen statt Mythen

ISBN-10 3-9810384-9-5
ISBN-13 978-3-9810384-9-1
EAN 9783981038491

Eine Schriftenreihe der Fakultät Digitale Medien

Herausgeber: Nikolaus Hottong

Stereoskope HD-Produktion



Nikolaus Hottong (Hrsg.)

Fakultät Digitale Medien
Arbeitspapier Nr. 8

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung, sowie der Übersetzung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der Fakultät reproduziert oder unter Verwendung elektronischer Systeme (inkl. Online-Netzwerken) gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

Copyright © 2016 Fakultät Digitale Medien | Hochschule Furtwangen

ISBN-10 3-9810384-9-5
ISBN-13 978-3-9810384-9-1
EAN 9783981038491

REDAKTION: Prof. Nikolaus Hottong,
COVER: Sandra Beuck
SATZ: Silvia Becker
DRUCK: Druckerei Hochschule Furtwangen

Nikolaus Hottong (Hrsg.)

Stereoskope HD Produktionen
Messungen statt Mythen

Schriftenreihe Fakultät Digitale Medien
HFU Furtwangen University
Dezember 2016

Inhalt

Vorwort des Herausgebers	15
--------------------------------	----

Teil 1

Empirische Untersuchung zur Auswirkung differierender Objektivverzeichnungen in Stereo-3D-Produktionen – Test.....	19
1. Einleitung.....	27
1.1 Hintergrund.....	27
1.2 Einfluss der Verzeichnung auf die stereoptische Wahrnehmung.....	28
1.3 Einordnung der Fragestellung.....	30
2. Item-Auswahl	32
2.1 Strukturelle Klassifizierung der Testitems	33
3. Verzeichnung.....	35
3.1 Synthetische Verzeichnung.....	39
3.1.1 Verzeichnungsparameter	40
3.2 Praxisorientierte Klassifizierung der verwendeten Verzeichnungs-differenzen.....	45
3.2.1 Globale Störungsbetrachtung.....	48
3.2.2 Feldbezogene Störungsbetrachtung	49
3.2.3 Vertikale Disparitäten	50
3.2.4 Aniseikonie-Effekte	52
4. Setup für die empirische Untersuchung.....	54
4.1 Ziel der empirischen Untersuchung.....	54
4.2 Verwendete Hard- und Software	54
4.2.1 3D-TV Gerät	55
4.3 Pretest mit Experten	56
4.4 Testaufbau	56
4.5 Konzipierter Testablauf	59
4.6 Produktion der verzeichneten Items	60
4.6.1 Aufbau einer Testsequenz.....	72
4.7 Probandenpanel	73

Teil 2

Empirische Untersuchung zur Auswirkung differierender Objektivverzeichnungen in Stereo-3D-Produktionen – Auswertung.....79

1. Einleitung..... 85
2. Auswertung Befindlichkeitstest..... 86
3. Kategorisierung der Probandenantworten 88
4. Globale Analyse der Probandenaussagen 93
5. Min/Max-Analyse des Probandenpanels..... 94
6. Globalanalyse entlang stellvertretender Praxiskategorien 95
7. Globalanalyse nach Medientyp der Items..... 99
8. Globalanalyse nach Strukturtyp der Items 99
9. Detailanalyse „vertikale Disparitäten“ 104
10. Detailanalyse „Aniseikonie-ähnliche Effekte“ 115
11. Detailanalyse Einfluss von Verzeichnungszentrum-Verschiebungen 118
12. Zusammenfassung 119

Teil 3

Untersuchungen zur Montage in Stereo 3D-Realfilmen125

1. Einleitung..... 129
 - 1.1 Abgrenzung..... 130
 - 1.2 Methodik..... 130
 - 1.3 Begriffe 131
 - 1.3.1 Editing 131
 - 1.3.2 Cutting..... 131
 - 1.3.3 Szene 131
 - 1.3.4 Einstellung 131
 - 1.3.5 Tiefensprung..... 131

1.3.6	Tiefenposition.....	132
1.3.7	Tiefenausdehnung.....	132
2.	Montage.....	132
2.1	Die Geschichte der Montage	133
2.1.1	(Handlungs)-Achsenschema	134
2.1.2	Shot/Reverse-Shot Schema	134
2.1.3	Kontinuität.....	135
2.1.4	Cross-Cutting	135
2.1.5	Montage-Ellipsen	136
2.1.6	Match-Cut.....	136
2.1.7	Zwischenschnitt	136
2.2	Die Einstellung	137
2.3	Transition	137
2.3.1	Cut.....	138
2.3.2	Überblendung.....	138
2.3.3	Irisblende	138
3.	Tiefenwahrnehmung	138
3.1	Monokulare Depth Cues.....	139
3.1.1	Perspektive & relative Größe	140
3.1.2	Vertraute Größe	140
3.1.3	Texturdichtegradient	140
3.1.4	Bewegungsparallaxe	140
3.1.5	Verdeckung	140
3.1.6	Relative Höhe.....	140
3.1.7	Atmosphärische Perspektive	140
3.1.8	Schatten	141
3.1.9	Perspektivische Konvergenz	141
3.1.10	Okulomotorik.....	141
3.1.11	Patente Stereopsis.....	141
4.	Montage in der Stereo-3D-Produktion	142
4.1	Technischer Bildwechsel	143
4.2	Blickanschluss.....	144
4.3	Größenverhältnisse.....	144
4.4	Perspektivkontrast	145
4.5	Schnittrhythmus.....	145

4.6	Zwischenschnitte und Einstellungsgrößen.....	145
4.7	Match-Cut.....	146
4.8	(Handlungs)- Achsenschema	146
5.	Tiefensprünge.....	146
5.1	Horizontal Image Translation	147
5.2	Tiefenscript	148
5.3	Relevanz	148
6.	Empirische Studie	149
6.1	Produktion der Test-Items.....	150
6.1.1	Technische Umsetzung	150
6.1.2	Postproduktion.....	152
6.2	Testaufbau	152
6.3	Expertentest	153
6.4	Optimierung der Studie	154
6.5	Testablauf	154
7.	Ergebnisse	157
8.	Fazit.....	159
	Literaturverzeichnis:	162

Teil 4

	Unschärfe in Stereo-3D-Produktionen.....	167
1.	Einleitung.....	171
2.	Unschärfe im Film	171
3.	Abgrenzung und Methodik unserer Untersuchung	173
4.	Literaturrecherche.....	173
5.	Eigene Untersuchungen mit Randot	174
6.	Eigene Untersuchungen mit Stereo-3D-Realfilm	176
6.1	Vorgehen	177
6.2	Auswertung.....	179
7.	Rücktransfer der Realfilm-Ergebnisse auf RDS	181
8.	Fazit.....	182

Teil 5

Cardboarding Impacts: Long Lenses vs. HIT	185
1. Einleitung.....	191
1.1 Wahrnehmung stereoskopischer Bilder.....	191
1.2 Betrachtung von Körpervolumina durch Augenpaare vs. Stereokamera	194
1.3 Wahrnehmungspsychologie von Aufnahmen mit langer Objektivbrennweite.....	198
2. Cardboarding messen (Teil 1)	203
3. Horizontal Image Translation.....	204
3.1 Allgemeine Funktionsweise und Auswirkung von HIT ...	204
3.2 Funktionsweise und Auswirkung positiver HIT	205
3.3 Funktionsweise und Auswirkung negativer HIT	206
3.4 Generelle HIT Nebenwirkungen	206
4. Cardboarding messen (Teil 2)	208
5. Diskussion.....	211
6. Zusammenfassung	212
Literaturverzeichnis	214

Teil 6

Entwicklungstendenzen stereoskopischer Computer- spiele	217
1. Einleitung.....	223
2. Entwicklung stereoskopischer Spiele.....	225
2.1 Stereoskopie zur Verbesserung der Spieltiefe	225
2.2 Stereoskopie als Vorteil für den Spieler	226
2.3 Stereoskopie als Hauptbestandteil der Spielmechanik	227
2.3.1 Eliminieren der monokularen und bewegungsinduzierten Tiefenreize	228
2.4 Eingabegeräte und Navigation im Raum	233
2.5 Unterschiede der stereoskopische Rendermethoden.....	235
2.5.1 Duales Rendern.....	235

2.5.2 Post Stereo	238
2.5.3 Bewusste Wahl der Rendermethode	239
3. Gestalterische Herausforderungen und Probleme	240
3.1 Immer alles tief?.....	241
3.2 Rahmenverletzungen und Schwebefenster.....	241
3.3 Negative Parallaxen / Pop-Outs	243
3.4 Seiteneffekte überproportionaler Stereobasen	245
3.5 Visuelle Effekte	247
3.6 Spielinformationen und Benutzer-Menüs	252
3.7 Präventive Maßnahmen gegen Übersprechen / Geisterbilder	256
4. Ansätze für eine automatische, kontextadaptive Tiefendarstellung.....	258
Literaturverzeichnis	262

Teil 7

Stereoskopische Einflussmöglichkeiten in 3D-Spieleumgebungen für Machinimas.....267

1. Einleitung.....	273
2. Marktanalyse.....	273
2.1 NVIDIA 3D Vision, 3D-PC, 3D-Display und Fraps.....	274
2.1.1 Funktionsweise der 3D-Vision Technologie	275
2.1.2 NVIDIA-Kompatibilitätsliste	275
2.1.3 Stereo-Parametrisierung	275
2.1.4 Workflow-Aspekte	277
2.2 Valve Source Filmmaker	278
2.2.1 Stereo-Parametrisierung.....	279
2.2.2 Workflow-Analyse	282
2.3 iClone.....	283
2.3.1 Stereo-Parametrisierung.....	283
2.3.2 Workflow-Analyse	285
3. Prototypische Stereo-3D-Machinima Produktion	286
3.1 Benötigte Zusatztools	286

3.2 iClone-Workflow.....	287
3.2.1 Schritt 1 – Grundbewegung	287
3.2.2 Schritt 2 – szenenspezifische Bewegung.....	288
3.2.3 Schritt 3 – Mimik.....	288
4. Fazit.....	290

Teil 8

Web-Based-Training für stereoskopische Bildgestaltung.....	293
---	-----

1. Berufsbilder	298
1.1 Der Kinematograf.....	298
1.2 Der Stereograf.....	299
2. Typische Anfängerfehler und deren Vermeidung	300
2.1 Gestalterische Unschärfe.....	300
2.2 Nahpunkt und Fernpunkt.....	302
2.3 Window-Violation	302
2.4 Gigantismus	303
2.5 Liliputismus	304
2.6 Kulisseneffekt.....	304
2.7 Teilbildkonflikte	304
2.8 Schnitt.....	306
3. Konzeption eines Web-Based-Trainings.....	306
3.1 Zielgruppenanalyse.....	306
3.1.1 Situative Disposition.....	307
3.1.2 Affektive Disposition.....	307
3.1.3 Kognitive Disposition.....	308
3.2 Lehr-/Lernziele.....	308
3.3 Verbreitungsmedium.....	310
3.4 Lehrstrategie	310
3.5 Lernwegstruktur.....	311
3.6 Darstellung des Lerninhalts.....	312
3.7 Motivationsdesign.....	313
3.8 Interaktionsdesign	314

3.9 Aufgaben	315
3.9.1 Aufgabentyp.....	315
3.9.2 Feedback.....	315
3.10 Didaktisches Modell	316
3.11 Realisierung des Web-Based-Trainings	318
3.11.1 Überlegungen zum Verbreitungsmedium.....	318
3.11.2 Die Videoproduktion	318
3.11.3 Post-Produktion.....	320
3.11.4 Die Webseite	321
4. Fazit.....	324
Quellenverzeichnis	328

Vorwort des Herausgebers

Dies ist der achte Arbeitsbericht der Fakultät Digitale Medien an der Hochschule Furtwangen. Wie alle vorherigen stellt auch dieser Bericht eine Querschnittsaufnahme unserer praxisbezogenen Medien-Forschung dar. Nach vier Jahren haben wir wieder die Gelegenheit ergriffen, aktuelle Projektstände im Forschungsgebiet Stereoskope HD-Produktion zu dokumentieren und in Druckform für interessierte Leser zugänglich zu machen.

Studierende der Hochschule Furtwangen haben dafür u.a. Teile ihrer erfolgreich abgeschlossenen Thesis- und Projektarbeiten beigetragen. Diese Teile wurden von uns überarbeitet, zusammengeführt und um den aktuellen Stand der laufenden Forschungsarbeiten ergänzt. Das Ergebnis halten Sie hiermit in Händen.

Der Untertitel dieser Publikation weist dezent darauf hin, mit welchen Schwerpunkten wir uns diesmal beschäftigt haben: Oft haben wir uns in den zurückliegenden Jahren gefragt, was von den „Mythen“ der S3D-Produktion zu halten ist, von denen man in den Gängen der Produktionshäuser und auf Fachkonferenzen immer wieder hört, deren gezielte Literaturrecherche aber oft keine belastbare Treffer ergab. Vier dieser Mythen haben wir uns konsequent wissenschaftlich genähert und die Ergebnisse ausführlich dokumentiert. Dabei haben wir großen Wert auf die Methodik gelegt, so daß die Ergebnisse jederzeit von Dritten nachvollzogen werden kann.

Um die Lust auf das Lesen anzufachen möchten wir – wie immer – kurz aufzeigen wie dieser Bericht strukturiert ist:

In den Forschungsprojekten ADEKKA¹ und BinoQ3D², welche sich beide mit kamerabedingten Bildfehlern in der S3D-Produktion auseinandersetzen, stellte sich oft die Frage, ab welchem technischen Fehlergrad S3D-Zuschauer von visuellen Irritationen heimgesucht werden. Speziell für objektivbedingte Verzerrungsdifferenzen zwischen linken und rechten Teilbild hat Tolga Sipahi einen Testaufbau entwickelt, mit der dieser Frage erstmals gezielt empirisch für S3D-Film nachgegangen werden konnte. Silvia Becker, Sandra Beuck und Arthur Wittmann haben die Ergebnisse seines ersten, umfassenden Probandentests statistisch ausgewertet und kommen diesbezüglich zu sehr interessanten Schlussfolgerungen.

Mit dem Mythos des „Tiefensprungs“ in S3D-Produktionen und seinen irritierenden Wahrnehmungsauswirkungen befasst sich Benjamin Hauser. Schon der Ansatz etwas Messen zu wollen, was eigentlich nicht exakt spezifiziert und definiert ist, zeigt wieviel Mut dazugehört, sich den stereoskopischen Mythen zu nähern. Tatsächlich produzierte er sehr sprunghafte S3D-Montagebeispiele, welche nachweislich kaum einen Probanden irritierten. Warum? Dies können sie in Kapitel 3 nachlesen.

Benjamin Hauser und Robert Lichtl versuchten gemeinsam mit Nadine Leonhardt schrittweise und methodisch der Frage nachzugehen, ob selektive Unschärfe im S3D-Film vom Zuschauer eher irritierend oder als unproblematisch wahrgenommen wird und präsentieren dazu interessante Analysen und eigene Messergebnisse, welche in Kapitel 4 beschrieben sind.

Wer hat sich nicht schon über die „Pappkameraden“ in S3D-Filmen gewundert? Über die Ursachen des sog. cardboard-effect wird in der S3D-Community oft wild spekuliert, dabei sind die Gründe offensichtlich. Lange Brennweiten bei der Aufnahme und nachträgliche horizontal image translation (HIT) können sich schnell zu diesem unerwünschten Bildeffekt aufschaukeln. Welcher von beiden Faktoren der ausschlaggebendere ist und wie sich beide auch gegenseitig neutralisieren können wird in Kapitel 5 dargelegt.

Thomas Mendelssohn befasst sich im darauffolgenden Kapitel mit der Frage, welche grundsätzlichen Vorteile stereoskopische 3D-Games dem Spieler bringen

-
- 1 „Autodetektion und Korrektur von Kameraobjektivfehlern im Hinblick auf stereoskope Videoproduktionsverfahren“, gefördert durch das Land Baden-Württemberg im Rahmen des Programms Innovative Projekte / Kooperationsprojekte.
 - 2 „Binokulares Blickerfassungssystem zur Qualitätsbeurteilung von Stereo-3D-Darstellungen“, gefördert vom BMWi im Rahmen des Programms ZIM.

könnten. Das verbesserte Schätzen von Distanzen scheint dabei eine große Rolle zu spielen. Sein spezielle Fragestellung lautete: Wie muss ein Spiel konzipiert sein, welches nur stereoskopisch spielbar ist. Seine ausführliche Analyse finden sie in Kapitel 6.

Stereoskopische Machimina, also in Game-Engines gedrehte S3D-Filme, wären sicherlich ein spektakuläres Genre. Aber welche stereoskopischen Parameter lassen sich in gängigen Engines des Jahres 2014 überhaupt wahrnehmungsgerecht manipulieren? Eine eher ernüchternde Analyse dazu finden sie in Kapitel 7 von Robin Grabowski. Trotzdem gelang es ihm mit der Software iClone einen wahrnehmungsgerechten Machimina-S3D-Kurzfilm erstellen.

Warum fällt es selbst erfahrenen Kameramännern und -frauen schwer, ihren ersten S3D-Film fehlerfrei zu drehen und wie kann man die schwierige Lernkurve verkürzen und den Lerneffekt zielgruppengerecht beschleunigen? Mit diesen Fragen setzte sich Nadine Leonhardt in ihrer preisgekrönten³ Abschlussarbeit auseinander und entwickelte aus ihren Erkenntnissen ein Web-based Training, welches als letzter Beitrag in Kapitel 7 vorgestellt wird.

Schlussendlich sind es mal wieder über 300 Seiten Informationen zum Thema geworden.

Wir möchten uns mit dieser Druckausgabe ganz herzlich bei allen Beteiligten für das Engagement und die Zusammenarbeit bedanken. Die Arbeiten werden natürlich fortgesetzt.

Furtwangen, im Dezember 2016

Nikolaus Hottong

Teil 1
Empirische Untersuchung zur Auswirkung
differierender Objektivverzeichnungen in
Stereo-3D-Produktionen – Test

Tolga Sipahi, Nikolaus Hottong

Abbildungsverzeichnis

Abbildung 1: Prinzipielle Verdeutlichung der stereoptischen Auswirkungen von geometrischen Teilbildkonflikten, ausgelöst durch unterschiedliche Objektiv-Verzeichnungen („Verzeichnungsdifferenz“)	29
Abbildung 2: Auswirkung der Verzeichnung auf Tiefenwahrnehmung Item 4	30
Abbildung 3: Screenshots der verwendeten Testitems	33
Abbildung 4: Verzeichnung in Abhängigkeit von der Blendenposition	36
Abbildung 5: Beschreibung einer starken Kissenverzeichnung durch ein Polynom 2. Ordnung (Dreiecke = Meßpunkte Soll/Ist)	38
Abbildung 6: Anonymisierter Auszug aus dem Datenblatt einer HD-Industriekamera mit Angaben zur Montagepräzision (man beachte zusätzlich die „Note“)	39
Abbildung 7: Verschiebungen der Verzeichnungszentren im Probandentest	40
Abbildung 8: Polynom K0 - Kissenverzeichnung sehr schwach	41
Abbildung 9: Polynom K1 - Kissenverzeichnung schwach	42
Abbildung 10: Polynom K2 - Kissenverzeichnung stark	42
Abbildung 11: Polynom K3 - Kissenverzeichnung sehr stark	43
Abbildung 12: Polynom T0 - Tonnenverzeichnung sehr schwach	43
Abbildung 13: Polynom T1 - Tonnenverzeichnung schwach	44
Abbildung 14: Polynom T2 – Tonnenverzeichnung stark	44
Abbildung 15: Polynom T3 - Tonnenverzeichnung sehr stark	45
Abbildung 16: Bildfelder für die meßtechnische Analyse und die Probandenbefragung	50

Abbildung 17: Augenaufnahmevorrichtung der Masterprojektgruppe.....	55
Abbildung 18: Aufsicht des Testaufbaus	57
Abbildung 19: Testaufbau - Foto 1	58
Abbildung 20: Testaufbau - Foto 2	58
Abbildung 21: Prozesskette Verzeichnung in NukeX.....	61
Abbildung 22: LensDistortion Node.....	61
Abbildung 23: StreetDance Video 1 unverzeichnet.....	62
Abbildung 24: StreetDance Video 1 maximal kissenverzeichnet (Radial Distortion 1: +0,3).....	63
Abbildung 25: Streetdance Video 1 maximal tonnenverzeichnet Screenshot in Nuke (Radial Distortion 1: -0,3).....	63
Abbildung 26: Crop-Node	64
Abbildung 27: Streetdance Video 1 zugeschnitten	64
Abbildung 28: Vergrößert Randot Video 1 (Frame 30) nach Verzeichnung	66
Abbildung 29: Colorlookup Node - Filter von Grauwerten.....	67
Abbildung 30: Vergrößert Randot Video 1 (Frame 30) nach ColorLookup Korrektur.....	67
Abbildung 31: Nuke Screenshot vom Ablauf zur Erzeugung der Parallaxen- Histogramme.....	68
Abbildung 32: Erhöhung der Querdisparitäten durch die Verzeichnungs- differenzen (in Winkelminuten aus Betrachtungsposition).....	72
Abbildung 33: Grundmuster - Probandenvideo.....	73
Abbildung 34: Altersverteilung der Probanden in Bezug auf Geschlecht und Sehhilfe	74
Abbildung 35: Verteilung der Stereosehfähigkeit (x-Achse) der Probanden gemäß Randot Visual 3D-Test. Ein höherer Wert bedeutet eine bessere binokulare Tiefenauflösungsfähigkeit.....	75
Abbildung 36: Probanden und ihre Vertrautheit mit Medien.....	76
Abbildung 37: Mediennutzung der Probanden.....	76

Tabellenverzeichnis

Tabelle 1: Klassifizierung der Items gemäß der geometrischen Eigenschaften ihrer Inhaltsbestandteile	34
Tabelle 2: Verwendete Testitems, Verzeichnungsparameter und deren praxisorientierte Klassifizierung	47
Tabelle 3: Verzeichnungsparameter und globaler Grad der verursachten Bildveränderungen	49
Tabelle 4: Verzeichnungsparameter und globale vertikale Disparitäten in den verwendeten Testitems	51
Tabelle 5: Verzeichnungsparameter und bedingte Feldvergrößerungen in den verwendeten Testitems	53
Tabelle 6: Tonneverzeichnete Items und ihre Cropwerte	65
Tabelle 7: Finale Items, Verzeichnungsparameter und resultierende Querveränderungen	72

1. Einleitung

Die Produktion von S3D-Filmen ist derzeit immer noch aufwendig, u.a. weil sie von vielen potenziellen Fehlerquellen begleitet wird. Insbesondere bei der Verwendung konventioneller Kameras und Objektive in side-by-side oder mirror-rig-Anordnung kommt es zu verschiedensten optischen Störungen, die sich negativ auf die stereoptische Wahrnehmung der Zuschauer auswirken. Eine dieser Störungen ist die unterschiedliche Bildfeldwölbung im linken und rechten Stereoteilbild – im Folgenden Verzeichnungsdifferenz genannt. Dieser Beitrag (und der nachfolgende) untersucht die Auswirkung dieses Effekts auf die Wahrnehmung von stereoskopischen Bildern im HD-Format.

Die Untersuchung beruht auf einem eigenständig konzipierten, empirischen Test welcher in Kooperation mit dem Forschungsprojekts ADEKKA¹ an der Fakultät Digitale Medien der Hochschule Furtwangen durchgeführt wurde. Im hier vorliegenden Beitragsteil 1 wird der Test bezüglich seiner Planung und Durchführung vorgestellt. Teil 1 ist folgendermaßen strukturiert: Zunächst wird der untersuchte Fehler eingehend definiert und seine mögliche Auswirkung auf die stereoptische Wahrnehmung diskutiert. Danach werden sowohl die Konzeption als auch die Umsetzung eines Probandentests sowie die Erzeugung der dafür verwendeten Video- und Bildmaterialien beschrieben. Im nachfolgenden Teil 2 werden die Testergebnisse methodisch ausgewertet und statistisch interpretiert.

Folgende zentrale Fragestellung soll dabei versucht werden zu beantworten: ab welchem Maß von Verzeichnungsdifferenz kommt es zu signifikanten Störungen in der stereoptischen Wahrnehmung von S3D-Zuschauern?

1.1 Hintergrund

Für einen Laien ist ein Stereo-3D Film im Grunde ein Film, der aus zwei leicht ver-

¹ Das Projekt ADEKKA (Autodetektion und Korrektur von Kameraobjektivfehlern im Hinblick auf stereoskope Videoproduktionsverfahren) wurde 2011 – 2013 an der Fakultät Digitale Medien der Hochschule Furtwangen (HFU) durchgeführt. Es wurde durch das Land Baden-Württemberg im Rahmen des Programms Innovative Projekte / Kooperationsprojekte gefördert. Projektinitiatoren Prof. Dr. Thomas Schneider und Prof. Nikolaus Hottong.

setzten Blickwinkeln aufgenommen wurde. Die Tiefenwahrnehmung erfolgt später beim Betrachten durch gezielte stereoptische Stimulation des visuellen Systems, im Wesentlichen durch die im Bild vorhandenen Querdisparitäten. Obgleich die stereoptische Wahrnehmung des Menschen gewisse Toleranzen duldet, ist der Grat zwischen unmerklichen Fehlern und irritierenden/ermüdenden/schmerzenden Fehlern ungleich schmaler als in der 2D-Produktion.

Um den gewünschten stereoskopischen Tiefeneindruck zu erzielen, spielen viele Faktoren eine Rolle und müssen von Anbeginn an mitbedacht werden. Einer dieser Faktoren ist beispielsweise die Ausrichtung beider Kameras zueinander. Abhängig davon, ob diese parallel oder konvergent ausgerichtet sind, entstehen große Unterschiede im Bildaufbau und in der Bildwirkung. In dieser Arbeit soll ausschließlich die parallele Anordnung der Kameras angenommen werden². Innerhalb dieser Anordnung können optische Abbildungsfehler der Objektive ungewollte negative Einflüsse auf die stereoptische Bildwahrnehmung haben. Einer dieser Fehler ist die sogenannte Verzeichnungsdifferenz.

An diesem Punkt setzt diese Arbeit ihre zentrale Fragestellung an, die folgendermaßen lautet: »Wie und ab welchem Grad wirken sich Verzeichnungsdifferenzen in den Stereoteilbildern auf die stereoskopische Wahrnehmung der Zuschauer aus?« Es sei an dieser Stelle nochmals darauf hingewiesen, dass nicht die (häufig als Stilmittel) verwendete Bildfeldwölbung an sich als problematisch angesehen wird, sondern die Verzeichnungsdifferenzen L vs. R. Es ist bei heutigen S3D-Kinoproduktionen gängige Postproduktion-Praxis die Verzeichnung zunächst aus beiden Teilbildern zu entfernen (z.B. um eine perfekte Ausrichtung der Bilder sicherzustellen), um später *einen* definierten Verzeichnungstyp künstlich als Stilmittel auf beide Bilder wieder aufzuprägen. Der Aufwand ist allerdings hoch und steht derzeit für Live-Bilder nur eingeschränkt zur Verfügung. Des Weiteren kann bei der Doppelprojektion von S3D-Bildern im Kino dieser Effekt erneut auftreten.

1.2 Einfluss der Verzeichnung auf die stereoptische Wahrnehmung

Die grundlegenden Bedingungen für die Wahrnehmung von Tiefe durch Menschen werden an dieser Stelle nicht mehr ausgeführt (siehe Arbeitsberichte 5 und 7) sondern es werden nur die Auswirkungen von Verzeichnungsdifferenzen in S3D-Bildern diskutiert. In nachfolgender Abbildung sind diese nochmals verdeutlicht:

² Vgl. Tauer, Holger (2010), S. 363.

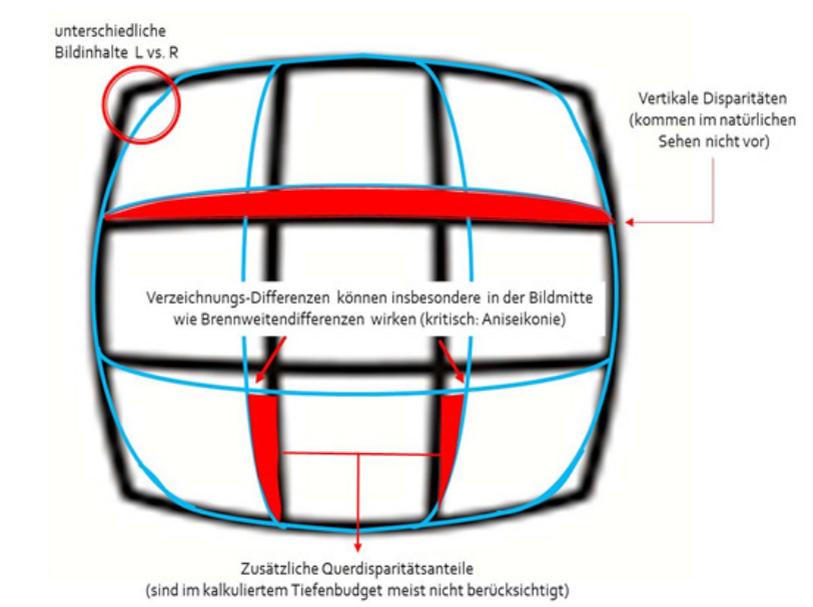


Abbildung 1: Prinzipielle Verdeutlichung der stereoptischen Auswirkungen von geometrischen Teilbildkonflikten, ausgelöst durch unterschiedliche Objektiv-Verzeichnungen³ („Verzeichnungs-Differenz“)

Um die Auswirkungen konkret sichtbar zu machen, wird eine beispielhafte Konstellation in S3D/HD/16:9 betrachtet. Es handelt es sich um das spätere Testitem Nr. 4, das links unverzerrt und rechts mit Polynom K3 sehr stark kissenverzerrt ist. Mit Adobe Photoshop wurden im linken Teilbild die Kanten des roten Würfels auf der linken Seite, die Kanten des grauen Objekts unter dem Flugobjekt und die Umrisse des Donut-förmigen Gebildes in der rechten oberen Bildecke mit einer grünen Farbe nachgezeichnet. Die korrespondierenden Kanten derselben Objekte wurden im Bild für das rechte Auge in lila Farbe nachgezeichnet.

³ Anmerkung: findet für die Bildmitte im Rahmen eines Stereobild-Alignments ein Brennweitenausgleich statt, wandern die Aniseikonie-Effekte entsprechend an die Bildränder. Im weiteren Verlauf der Untersuchung gehen wir von diesem eher praxisorientierten Fall aus.

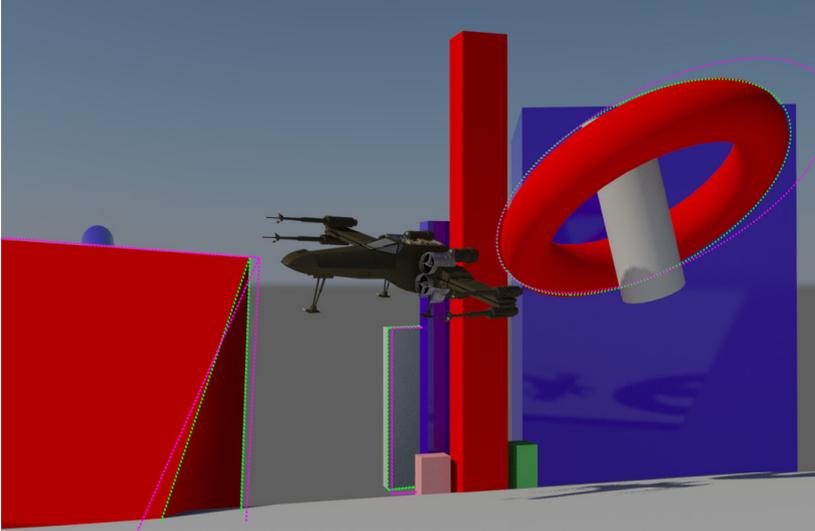


Abbildung 2: Auswirkung der Verzeichnung auf Tiefenwahrnehmung Item 4

Die bildlichen Auswirkungen der Verzeichnungs-differenzen sind hier sehr gut sichtbar: so bekommt das rote Objekt auf der linken Seite eine etwas verschobene Außenkante. Die Verschiebung verursacht eine Skalierung in Richtung Bildmitte. Das Objekt, das sich in der Bildmitte direkt unter dem Flugobjekt befindet, scheint von der Verzeichnung kaum betroffen zu sein, was dem Charakter realer Verzeichnung absolut entspricht. Denn diese wird stärker, je weiter ein Punkt vom Verzeichnungszentrum entfernt ist. Das Verzeichnungszentrum dieses Bildes wurde nicht verändert, es befindet sich somit exakt am Bildmittelpunkt.

Sehr starke Auswirkungen zeigt der Eingriff auch am Donut-Objekt in der rechten oberen Ecke. Die Größe des Objekts wird vor allem zum Rand hin stark verändert. Hier sind starke horizontale wie vertikale Disparitäten zu erwarten. Die neue Objektgröße führt zudem links-unten zu einer spontanen Scheinfensterverletzung (window violation).

1.3 Einordnung der Fragestellung

Recherchen zur Thematik ergaben, dass es sich in der Vergangenheit oft um Einzelstudien handelte, die sich durch ihre unterschiedlichen Methoden nur schwer vergleichen lassen. Dennoch soll hier auf einige vorab Bezug genommen werden.

Bereits im Jahr 1993 forschte Andrew Woods in einer Studie⁴ über die Unterschiede bei konvergenten und parallelen Kamerastellungen bei Stereoproduktionen. Es wurden verschiedene Parallaxenkombinationen (konvergente steigende/sinkende und parallele steigende/sinkende) simuliert. Die durch konvergente Stellung ausgelöste Trapezverzerrung wurde als störend, jedoch vermeidbar befunden. Eine parallele Kamerastellung sei allein aufgrund dieser Aussagen einer konvergenten vorzuziehen. Über die Linsenverzeichnung wurde ausgesagt, dass diese durchaus zu vertikalen Disparitäten führen und damit die Tiefenwahrnehmung stören kann. Durch eine verbesserte Auswahl von Objektiven oder durch die Vermeidung von Weitwinkelobjektiven könne dieses Problem - laut der Studie - minimiert werden.

In einer weiteren Studie⁵ wurde die Augenermüdung von 20 Probanden untersucht. Hierzu heißt es in der Studie: „In this letter, we propose a visual fatigue prediction metric for stereoscopic images by measuring excessive horizontal and vertical disparities“.⁶ Es wurde versucht, eine Korrelation zwischen den vorhergesagten und den tatsächlichen Werten der Probanden herzustellen, um die „prediction metrics“⁷ zu validieren. Zur Betrachtung wurde ein vier Zoll großes, polarisiertes Stereodisplay verwendet, das aufgrund seiner Größe jedoch wenig repräsentativ für Consumer-Fernsehgeräte ist.

Ein interessantes Testergebnis brachte eine weitere Studie⁸ zu Tage. Die Autoren fanden heraus, dass Randot-Sehtests, wie sie auch in unseren Probandentests verwendet werden, nicht perfekt geeignet scheinen, die Stereosehfähigkeit von Probanden zu bestimmen. Eine alternative Methode, in Form eines umfangreichen praktischen Tests, wurde für aussagekräftiger befunden.

Über die Wichtigkeit von gleicher Bildgröße beider Teilbilder berichtete eine weitere Studie.⁹ Hierbei wurde erkannt, dass es für die Stereowahrnehmung am besten ist, wenn die Abbildungsgröße in den Teilbildern jeweils gleich beziehungsweise annähernd gleich ist. Ihre Grenzwertbetrachtungen sind hierbei besonders interessant, da es bei Verzeichnungsunterschieden *lokal* zu einem geometrisch ähnlichen Effekt kommt.

John Vincent Lovasik und Mary Szymkiw zeigten in ihrer Studie¹⁰ ebenfalls einen Verlust von Stereowahrnehmung durch Erhöhung von Aniseikonie.

In einer anderen Studie¹¹ über die Toleranz der Stereopsis bez. vertikaler Teilbildabweichungen ergaben sich maximale Werte zwischen 44 und 88 Winkelminuten.

4 Vgl. Woods, Doherty, Koch (1993)

5 Vgl. Donghyun, Kwanghoon (2011)

6 ebd., S. 231.

7 ebd., S. 231.

8 Vgl. Long, Sui (2005)

9 Vgl. Vlaskamp, Filippini, Banks (2009)

10 Vgl. Vincent, Szymkiw (1985)

11 Vgl. Fukuda, Wilcox, Allison, Howard (2009)

Allerdings hatte die Studie lediglich vier Probanden, wovon zwei die Autoren selbst waren.

Um die Auswirkung von Tonnenverzeichnung auf die Wahrnehmung zu untersuchen, veranlasste Lachlan Pockett eine weitere Studie.¹² Dort wurden speziell texturierte Bilder genutzt, um die Verzeichnung nicht allzu offensichtlich zu machen. Dies wurde mit der Tatsache begründet, dass Verzeichnung generell an geraden Linien besser erkennbar sei. Die elf Probanden bekamen die Bilder auf einem Tablet-PC zu sehen. Bei der Kombination unterschiedlich verzeichneter Bilder gab es deutlich mehr Probleme bei der Wahrnehmung, als bei gleich verzeichneten Kombinationen. Die Autoren zogen unter anderem den Schluss, dass die Effekte auch inhaltsbedingt Differenzen hervorrufen können.

In einer Studie¹³ des Department of Computer Science der Durham University wurde die Auswirkung von Kompression auf die stereoskopische Wahrnehmung untersucht. Dieser Studie nach ist die Stereowahrnehmung stark durch die Kompression beeinflussbar. Daher wurde in unserer empirischen Untersuchung - zum Ausschluss dieser Beeinflussung - ausschließlich unkomprimierte Video- und Bildmaterialien verwendet.

2. Item-Auswahl

Um verschiedenen Kombinationen von Verzeichnungsparametern zu untersuchen, wurden den Probanden sowohl S3D-Standbilder als auch -Videos in Form von maximal zehn Sekunden langen Clips gezeigt („Items“). Alle Items lagen unkomprimiert mit einer Teilbildauflösung von jeweils 1920x1080 Pixeln vor. Auf die Verwendung von Ton wurde verzichtet, um die Probanden in ihrer visuellen Wahrnehmung nicht abzulenken.

Zur Verfügung standen folgende Bildinhalte:

1. typisches S3D-Filmmaterial (Kinofilme Street Dance und Hybrid, zur Verfügung gestellt von der Universum Film GmbH + interne Produktionen der Hochschule Furtwangen),
2. computergenerierte S3D-Animation mit einfachen geometrischen Körpern (selbst erstellt mit 3ds Max)
3. Random-Dot Animationen (interne Produktionen an der Hochschule Furtwangen),

alle jeweils als Videosequenz oder Standbild.

¹² Vgl. Pockett, Salmimaa, Pölonen, Häkkinen (2010)

¹³ Vgl. Gorley, Holliman (2008)

Auf Basis dieser Definition lassen sich die verwendeten 38 Items jeweils einem Struktur-Typen zuordnen.

Struktur-Typen der Items		
Typ A	Typ B	Typ C
Item 4 – CGI Standbild 1	Item 1 – Street Dance Video 1	Item 2 – Randot Video 4
Item 5 – Louise Video 4	Item 10 – Street Dance Video 6	Item 3 – Street Dance Video 5
Item 7 – CGI Video 2	Item 14 – Randot Video 3	Item 6 – Street Dance Video 4
Item 9 – CGI Standbild 2	Item 18 – Street Dance Standbild 1	Item 8 – Randot Video 1
Item 15 – Louise Standbild 2	Item 20 – Randot Standbild 3	Item 11 – Street Dance Standbild 2
Item 19 – Louise Video 1	Item 25 – Street Dance Video 3	Item 12 – Street Dance Standbild 5
Item 22 – Street Dance Standbild 6	Item 27 – Street Dance Video 2	Item 13 – Randot Standbild 2
Item 23 – Louise Standbild 1	Item 33 – Louise Video 3	Item 16 – Street Dance Video 7
Item 26 – Louise Video 2	Item 34 – Louise Standbild 3	Item 17 – Street Dance Standbild 7
Item 29 – CGI Video 1		Item 21 – Street Dance Standbild 4
Item 31 – Louise Standbild 4		Item 24 – Randot Video 2
Item 35 – Sauber1		Item 28 – CGI Video 3
		Item 30 – Randot Standbild 1
		Item 32 – CGI Standbild 3
		Item 36 – Sauber2
		Item 37 – Sauber3
		Item 38 – Sauber4

Tabelle 1: Klassifizierung der Items gemäß der geometrischen Eigenschaften ihrer Inhaltsbestandteile

3. Verzeichnung

Die Thematik der Verzeichnung wäre nicht denkbar ohne kurz auf die Objektive einzugehen, die sie durch ihre optischen Eigenschaften verursachen. Objektive werden den optischen Systemen zugeordnet. Für diese lieferte F. Pedrotti im Jahr 2005 folgende Beschreibung:

»[...] der Bereich, der mit »optisches System« bezeichnet wird, eine beliebige Anzahl von reflektierenden und/oder brechenden Oberflächen beliebiger Krümmung und Brechzahl, die die Richtung von Strahlen, die den Objektpunkt O verlassen, ändern.«¹⁴

Aufgrund von Lichtstreuung, Beugung und Abbildungsfehlern können optische Systeme in der Praxis nicht immer ideale Ergebnisse liefern.¹⁵ Die Bildfeldwölbung ist einer dieser Abbildungsfehler. Bereits 1956 schrieb Hans-Martin Brandt in seinem Werk über das Fotoobjektiv:

„Ein weiterer Fehler, die sog. Verzeichnung, liegt vor, wenn das von einem optischen System entworfene Bild eines Gegenstandes diesem nicht in allen Teilen geometrisch ähnlich ist.“¹⁶

Objektive existieren in verschiedensten Ausführungen, um diverse Anforderungen erfüllen zu können. Ihre verschiedenen optischen Eigenschaften hängen mit verschiedenen Bauweisen zusammen. Dabei ist die Zahl verwendeter optischer Instrumente wie Linsen, Spiegel und Teilsysteme mit mehreren Linsen sehr unterschiedlich.¹⁷ Durch eine planvolle Kombination verschiedener Linsentypen wird versucht, die Abbildungsqualität zu optimieren.¹⁸ Die Verwendung mehrerer Linsen erschwert jedoch die Berechnung der resultierenden Abbildungsverhältnisse, weshalb es üblich ist, ein oder zwei repräsentative Hauptebenen einzuführen, um die optische Abbildung zu beschreiben.¹⁹ Für die besten optischen Eigenschaften sollten alle Krümmungsmittelpunkte der Teilsysteme auf einer Geraden liegen. Dies lässt sich in der Realität allerdings nicht optimal umsetzen, wodurch es zu verschiedenen Fehlerscheinungen kommt.²⁰

Die sog. Verzeichnung beschreibt im Wesentlichen eine sich radialsymmetrisch ausbreitende und zum Rand hin stärker werdende Änderung des Abbildungsmaßstabes. Es existieren zwei grundlegende Varianten der Verzeichnung, die Tonnen- und die Kissenverzeichnung. Abhängig sind diese hauptsächlich durch die Position der Linsengruppen eines Objektivs zur Blende. Die Lage der Blende entscheidet maßgeb-

14 Pedrotti F (2005), S. 53

15 Vgl. ebd., S. 54

16 Brandt Hans Martin (1956), S. 42

17 Vgl. Schröder, Treiber (2007), S. 43

18 Vgl. Demtröder Wolfgang (2013), S. 27

19 Vgl. Schmidt Ulrich (2013), S. 385

20 Vgl. Schröder, Treiber (2007), S. 43

lich darüber, welche Strahlen beeinflusst werden und zu welcher Verzeichnungsart es kommt. Abbildung 4 illustriert diesen Zusammenhang anhand eines Beispiels. Nimmt die Vergrößerung des Maßstabs von der Bildmitte zum Rand hin zu, handelt es sich um eine Kissenverzeichnung. Nimmt er ab, spricht man von Tonnenverzeichnung.²¹ Kurt Michel bezeichnet die Kissen als positive und die Tonnen als negative Verzeichnung.²²

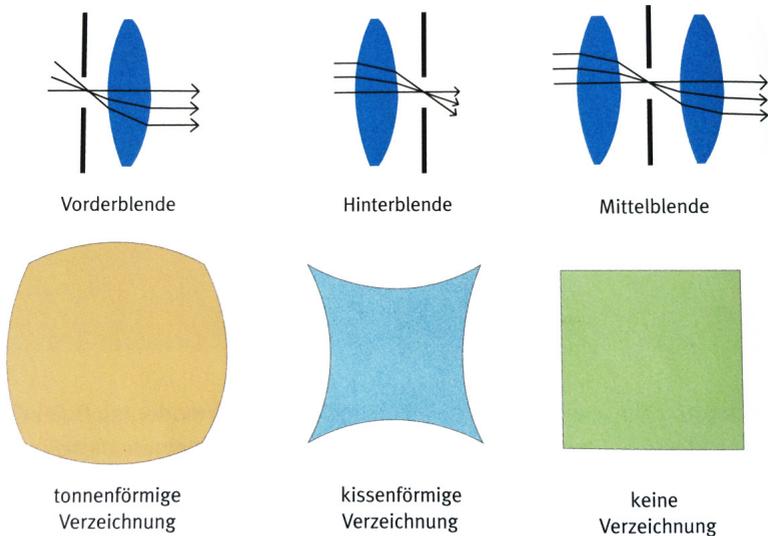


Abbildung 4: Verzeichnung in Abhängigkeit von der Blendenposition²³

Abbildung 4 deutet auf der rechten Seite an, dass Objektive mit streng symmetrischem Aufbau verzeichnungsfrei abbilden können. Solche Objektive werden orthoskopisch genannt.²⁴ Dieser Aufbau ist aber bei typischen Film- und TV-Objektiven kaum realisierbar. Derzeit machbar sind Objektive mit quasi-symmetrischem Aufbau, welche theoretisch wenig Verzeichnung aufweisen. Doch selbst in der quasi-symmetrischen Anordnung kann es plötzlich doch wieder zu Verzeichnungsphänomenen kommen, beispielsweise bei Weitwinkelobjektiven in Retrofokusbauweise (zerstreuende Elemente vor der Blende + sammelnde Elemente dahinter). Hier addieren sich dann beide Beiträge zu negativer Verzeichnung.

²¹ Vgl. Brandt Hans Martin (1956), S. 42

²² Vgl. Michel Kurt (1955), S. 66

²³ Tauer Holger (2010), S. 124

²⁴ Vgl. Brandt Hans-Martin, S. 43

Die Verzeichnung hat nur Einfluss auf die geometrischen Eigenschaften des Bildes und nicht etwa auf die Bildschärfe oder andere Parameter.²⁵ Umgekehrt kann aber die Einstellentfernung (Focus) eines Objektivs einen deutlichen Einfluss auf den Verzeichnungsgrad ausüben: Je nach Fokussiermethode (Gesamtverstellung, Hintergliedfokussierung, Innenfokussierung) ändern sich durch die Drehung am Fokusring die Abstände einzelner Linsengruppen untereinander und damit wiederum ihre relative Lage zur Blende. Interessanterweise hat die gewählte Blendenöffnung keinen Einfluss auf die Verzeichnung - zumindest nicht bei üblichen Film- und TV-Objektiven. Ein interessanter Effekt ist noch für Zoomobjektive zu vermelden. Bei ihnen treten um die Normalbrennweite herum oft Vorzeichenwechsel in der der Verzeichnung auf.

Die Verzeichnungsvarianten und die Stärke ihrer radialsymmetrisch wirkenden Abbildungsverzerrung werden heute häufig durch Polynome n-ter Ordnung beschrieben. Dabei wird die radiale Position $P_V(r)$ der realen Bildpunkte (wie sie von einem Objektiv auf dem Kamerasensor erzeugt werden) in Abhängigkeit von der eigentlichen Sollposition der idealen Bildpunkte (wie sie sich ergäben, wenn die Objektivbrennweite bzw. der Abbildungsmaßstab des Objektivs über das ganze Bildfeld konstant wäre) beschrieben.

$$\tilde{r} = P_V(r) \qquad P_V(r) = a_1 \cdot r + a_2 \cdot r^2 + a_3 \cdot r^3 + a_4 \cdot r^4$$

$$\tilde{r} = \sqrt{(\tilde{x} - x_0)^2 + (\tilde{y} - y_0)^2}$$

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

Hierbei sind $(x_0; y_0)$ die Koordinaten des sog. Verzeichnungszentrums.

²⁵ Vgl. Litfin Gerd (2005), S. 78

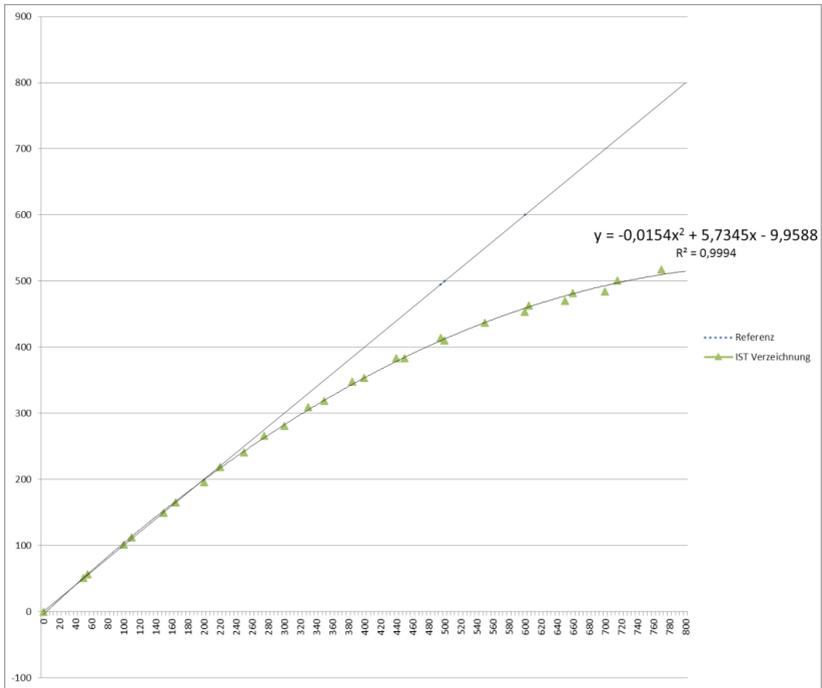
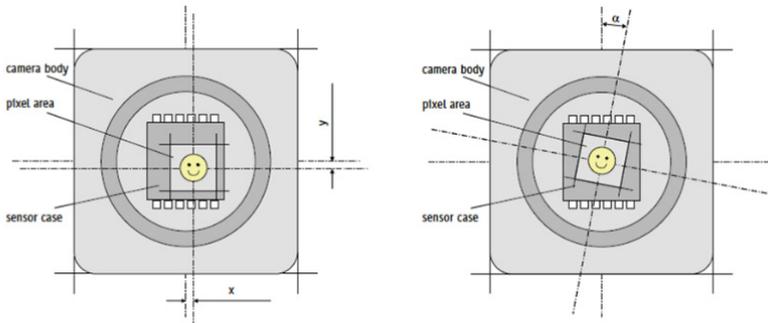


Abbildung 5: Beschreibung einer starken Kissenverzeichnung durch ein Polynom 2. Ordnung (Dreiecke = Meßpunkte Soll/Ist)

Die obige Formel macht deutlich, wie wichtig die exakte Bestimmung des optischen Verzeichnungszentrums eines Objektivs zur Beschreibung von Verzeichnung ist. Nur von diesem Zentrum aus breitet sich die Verzeichnung exakt radialsymmetrisch aus. Das Verzeichnungszentrum ist selten exakt-identisch mit der Bildmitte im aufgezzeichneten Bild, da die optische Achse des Objektivs und die Bildsensormitte sowohl produktionsbedingten wie auch mechanischen (und manchmal sogar thermischen) Schwankungen unterliegen. Bei Wechselobjektiven kommt noch die mechanische Präzision des Mounts hinzu.



Beispiel: Datenblatt einer HD-Industriekamera

Method of Positioning:	Optical alignment of photo sensitive sensor area into camera front module. (lens mount front flange)	
Reference points:	Sensor: Center of pixel area (photo sensitive cells). Camera: Center of camera front flange (outer case edges).	
Accuracy:	x/y: +/- 150 μm	(Sensor shift)
	z: +0 / -150 μm	(Optical back focal length)
	α : +/- 0.5°	(Sensor rotation)

Note: _____ x/y - tolerances between C-Mount hole and pixel area may be higher.

Abbildung 6: Anonymisierter Auszug aus dem Datenblatt einer HD-Industriekamera mit Angaben zur Montagepräzision (man beachte zusätzlich die „Note“)

3.1 Synthetische Verzeichnung

In unserem Probandentest wurde keine von realen Objektiven verursachte Verzeichnung verwendet. Eine bildbasierte - synthetisch erzeugte - Verzeichnung bietet eine bessere, leicht umsetzbare und reproduzierbare Alternative. Denn: kein reales Objektiv besitzt exakt dieselben optischen Eigenschaften wie ein zweites derselben Baureihe/Charge.

Zur Analyse und Aufbringung der Verzeichnungsparameter wurde die Postproduction-Software NukeX von The Foundry verwendet. Die Software bietet verschiedene Methoden der Verzeichnungskorrektur sowie eine automatische Bestimmung aller Verzeichnungsgrößen, insbesondere des optischen Verzeichnungszentrums.

3.1.1 Verzeichnungsparameter

Die Verzeichnung ist in NukeX definiert durch den radialen Verzeichnungsverlauf (im Sinne eines Polynoms 4. Ordnung mit rein geradzahligem Exponenten, bezeichnet mit RadialDistortion1 und RadialDistortion2) sowie dem Verzeichnungszentrum x_0 ; y_0 , von dem sich die Verzeichnung rotationsymmetrisch ausbreitet. Die in diesem Test konkret genutzten Verzeichnungswerte werden daher auch genau in dieser Notation dokumentiert.

Mangels belastbarer Recherche-Ergebnisse zum Einfluss von Verzeichnungsabweichungen auf die Stereopsis (s. Kap. 1.3) sind die im Rahmen dieses Tests genutzten Verzeichnungsparameter zunächst eher aggressiv gewählt, um eine große Spannweite zu erfassen. Dies betrifft sowohl die Parameter RD1 und RD2 wie auch x_0 und y_0 .

Die Verschiebung des Verzeichnungszentrums (x_0 , y_0) wurden in drei verschiedenen Varianten vorgenommen.

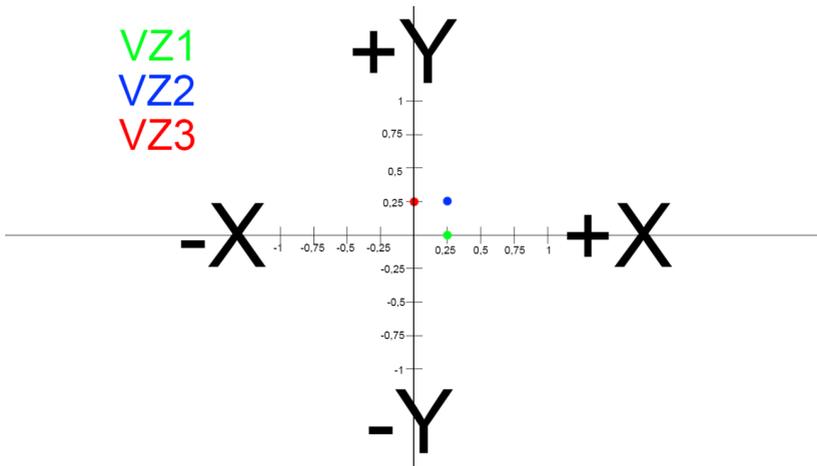


Abbildung 7: Verschiebungen der Verzeichnungszentren im Probandentest

Die drei genutzten Zentrumsverschiebungen sind nicht in Pixeln angegeben sondern als relative Einheiten. Eine Verschiebung in positiver x- und y-Richtung bewirkt (siehe VZ2) am Bild starke Effekte in der unteren, linken Ecke, weil sich dort die Pixel befinden, die am weitesten vom verschobenen Zentrum entfernt liegen. Negative Verschiebungen der genutzten Werte wurden nicht berücksichtigt, da es dadurch nur zu spiegelbildlichen Auswirkungen kommt. Da sich die Bildinhalte innerhalb des Bildes stark unterscheiden können, ist dennoch klar, dass ein anderes Verzeichnungszentrum durchaus andere Irritationsmeldungen ergeben kann.

In den folgenden Beispieldarstellungen befinden sich zwei identische Grids. Das blaue Grid stellt das unverzeichnete Original dar, wobei das rote die angegebene Verzeichnung repräsentiert. Um den Umgang mit den Kombinationen im Verlauf des Tests zu vereinfachen, wurden Bezeichnungen wie K0, K1, etc. eingeführt, die auf die Verzeichnungsart und den Grad der Verzeichnung schließen lassen.

Bezeichnung	Radial Distortion 1	Radial Distortion 2
Polynom K0	0.01	0.005

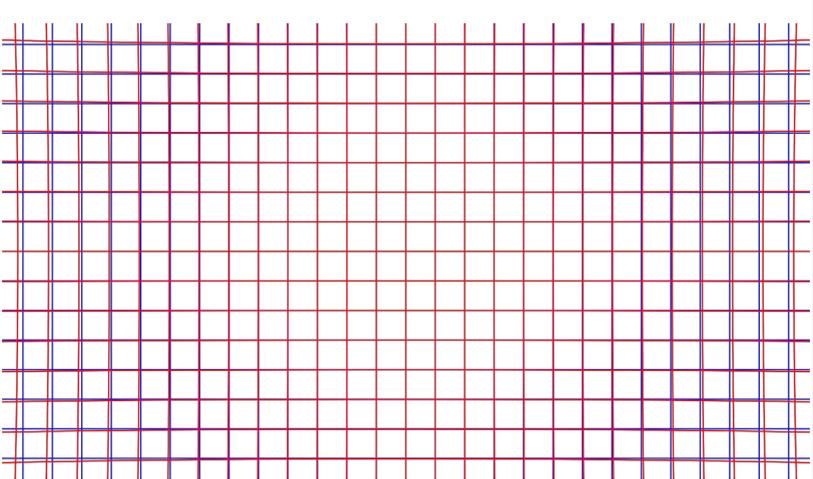
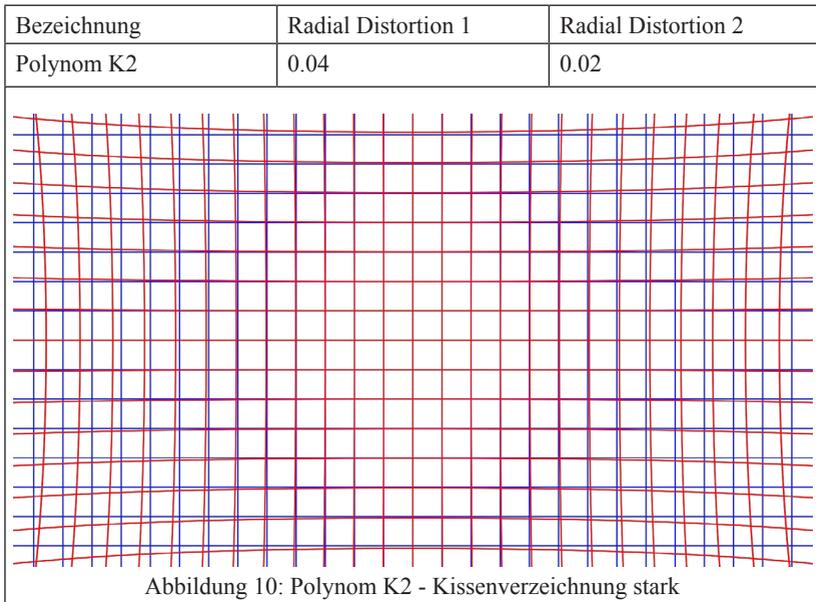
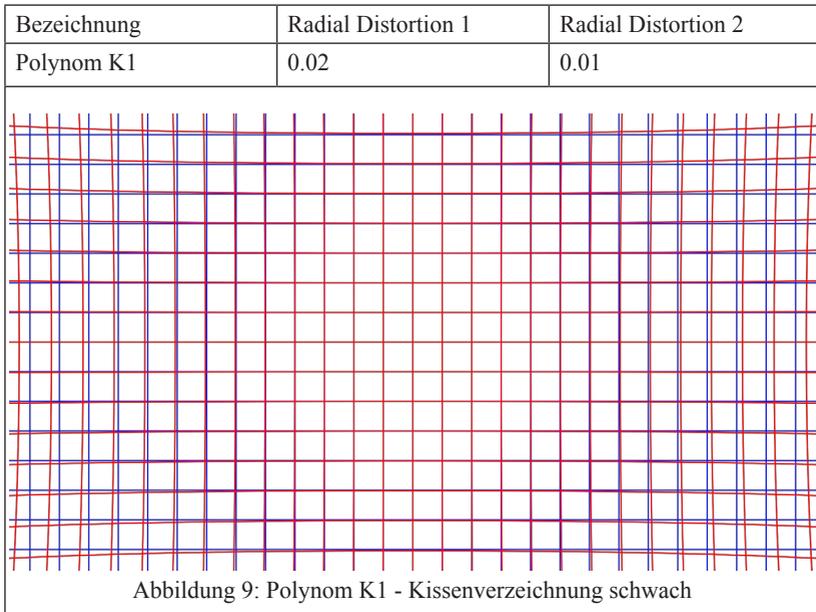


Abbildung 8: Polynom K0 - Kissenverzeichnung sehr schwach



Bezeichnung	Radial Distortion 1	Radial Distortion 2
Polynom K3	0.08	0.03

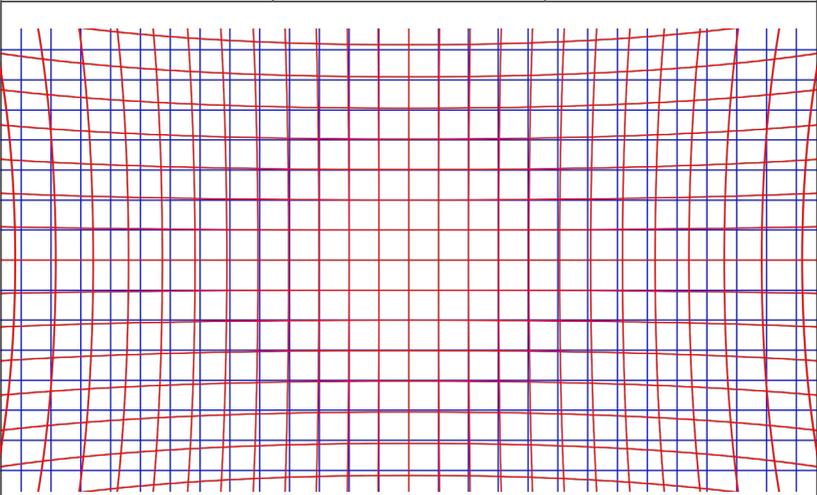


Abbildung 11: Polynom K3 - Kissenverzeichnung sehr stark

Bezeichnung	Radial Distortion 1	Radial Distortion 2
Polynom T0	-0.01	-0.005

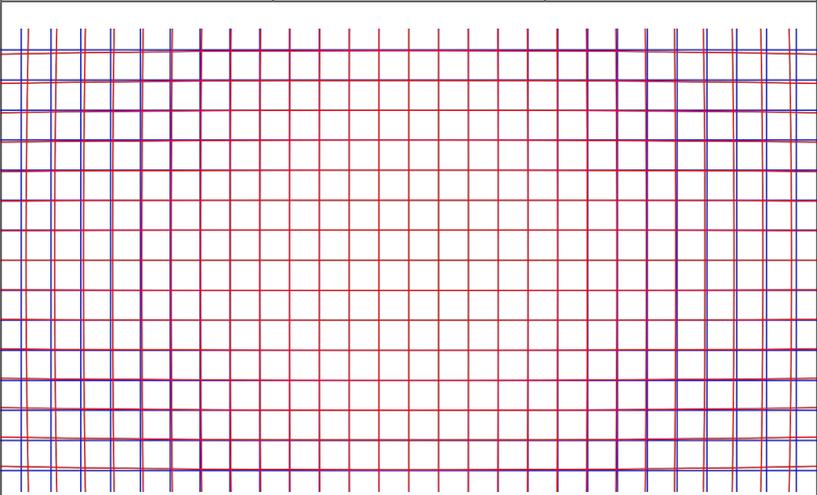
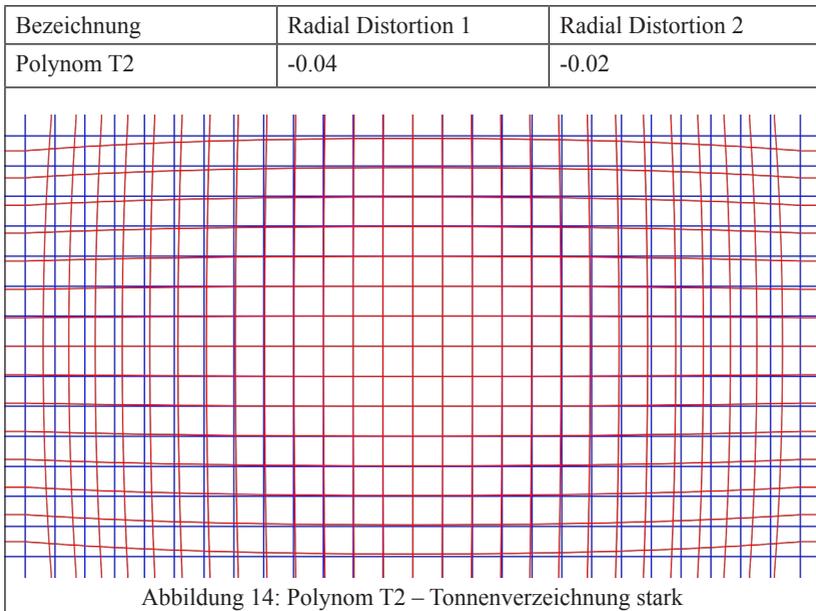
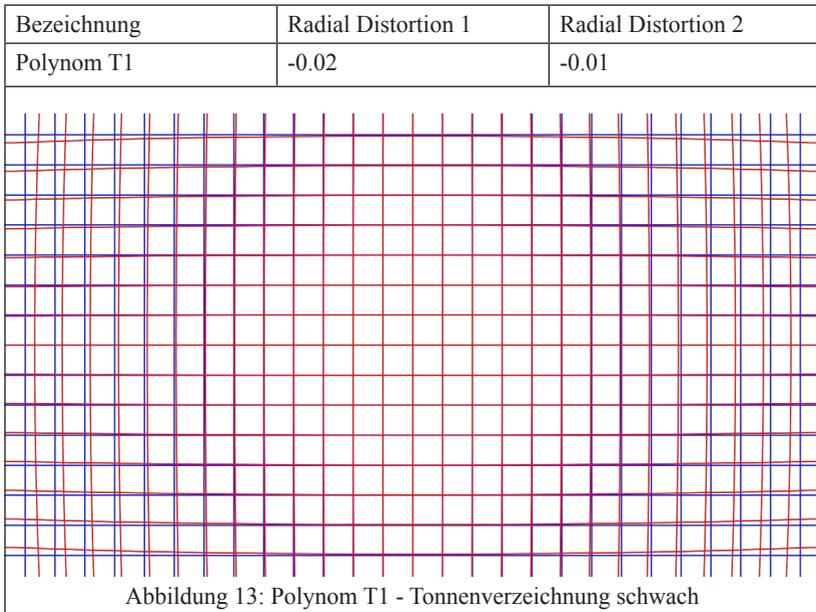
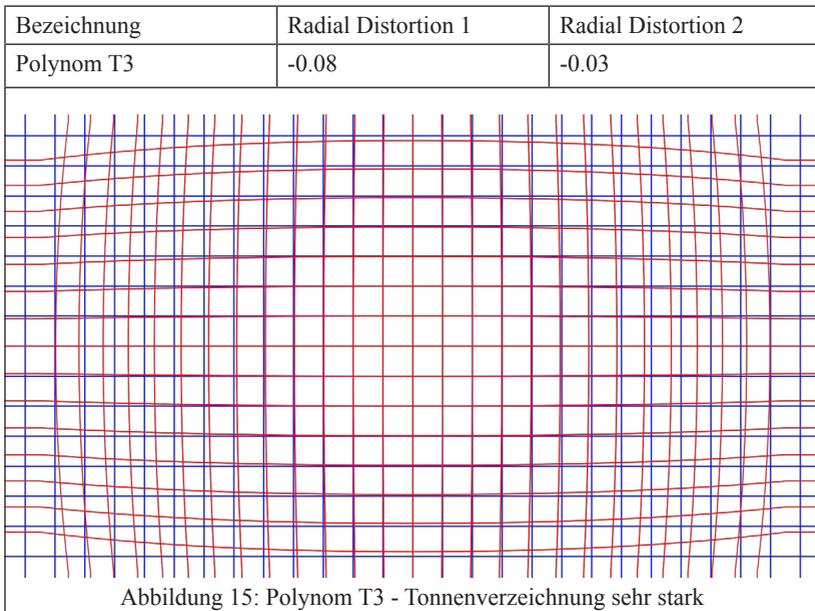


Abbildung 12: Polynom T0 - Tonnenverzeichnung sehr schwach





3.2 Praxisorientierte Klassifizierung der verwendeten Verzeichnungs-differenzen

Ein Test, wie er hier durchgeführt wurde, konnte - trotz umfassender Recherche - nicht in dieser Form und in diesem Umfang an Videosequenzen und Probanden identifiziert werden. Die in unserem Test verwendeten Verzeichnungs-differenzen stellen daher aus Praxissicht zunächst eher worst-case Betrachtungen dar, um eine grobe Orientierung über den Einfluss von Verzeichnungs-differenzen in S3D-Bildern zu erhalten.

Item	Inhaltstyp	Verzeichnungstyp [L]+[R]	repräsentiert
35	Hybrid Video1	unverzeichnet	Referenzpunkt1 (best case)
36	Street Dance Video8	unverzeichnet	
37	Street Dance Video9	unverzeichnet	
38	Street Dance Video10	unverzeichnet	

1	Street Dance Video 1	unverzeichnet + K0	typ. Verzeichnungen in szenischen S3D-Produktionen („Primes“)
18	Street Dance Standbild 1	unverzeichnet + T0	
5	Louise Video 4	K0 + K1	
13	Randot Standbild 2	K0 + K1	
22	Street Dance Standbild 6	T0 + T1	
30	Randot Standbild 1	T0 + T1	
8	Randot Video 1	K0 + K1 + x-Verschiebung	typ. Verzeichnungen in S3D-Liveproduktionen („Zooms“)
25	Street Dance Video 3	T0 + T1 + x-Verschiebung	
10	Street Dance Video 6	K0 + K1 + xy-Verschiebung	
28	CGI Video 3	T0 + T1 + xy-Verschiebung	
6	Street Dance Video 4	K1 + K2	typ. Verzeichnungen in Prosumer-Geräten bei kleinen Brennweiten und geringen Motivabständen
14	Randot Video 3	K1 + K2	
23	Louise Standbild 1	T1 + T2	
31	Louise Standbild 4	T1 + T2	
11	Street Dance Standbild 2	K1 + K2 + y-Verschiebung	
26	Louise Video 2	T1 + T2 + y-Verschiebung	

7	CGI Video 2	K2 + K3	typ. Verzeichnungen in sog. „Actioncams“
15	Louise Stand- bild 2	K2 + K3	
24	Randot Video 2	T2 + T3	
32	CGI Standbild 3	T2 + T3	
12	Street Dance Standbild 5	K2 + K3 + x-Verschie- bung	
29	CGI Video 1	T2 + T3 + x-Verschie- bung	
9	CGI Standbild 2	K0 + K3 + xy-Verschie- bung	Referenzpunkt2 (worst case)
27	Street Dance Video 2	T2 + T3 + xy-Verschie- bung	

2	Randot Video 4	unverzeichnet + K1	zur methodischen Voll- ständigkeit
16	Street Dance Video 7	unverzeichnet + K1	
19	Louise Video 1	unverzeichnet + T1	
33	Louise Video 3	unverzeichnet + T1	
3	Street Dance Video 5	unverzeichnet + K2	
20	Randot Stand- bild 3	unverzeichnet + T2	
4	CGI Standbild 1	unverzeichnet + K3	
17	Street Dance Standbild 7	unverzeichnet + K3	
21	Street Dance Standbild 4	unverzeichnet + T3	
34	Louise Stand- bild 3	unverzeichnet + T3	

Tabelle 2: Verwendete Testitems, Verzeichnungsparameter und deren praxisorientierte Klassifizierung

3.2.1 Globale Störungsbetrachtung

Den Items wurden zunächst globale Grade an Störung zugeordnet, anhand derer erkennbar ist, wie stark die verursachte Gesamtabweichung der Teilbilder durch die Verzeichnungsdifferenzen war. Da es sich immer um zwei Bilder mit unterschiedlichen Werten handelte, wurde immer die stärkere Verzeichnung als Hauptbezugswert genommen. Eine Verschiebung des Verzeichnungszentrums hat den Grad jeweils um eine Stufe erhöht. Eine Verschiebung in x- und y-Richtung erhöht den globalen Störungsgrad entsprechend um zwei Stufen. Aus dieser Festlegung ergeben sich die Werte in Tabelle 3.

zugewiesenes Item	Verzeichnungstyp [L]+[R]	Grad der global verursachten Störung im Bild
1	unverzeichnet + K0	1
2	unverzeichnet + K1	2
3	unverzeichnet + K2	3
4	unverzeichnet + K3	4
5	K0 + K1	2
6	K1 + K2	3
7	K2 + K3	4
8	K0 + K1 + x-Verschiebung	3
9	K0 + K3 + xy-Verschiebung	6
10	K0 + K1 + xy-Verschiebung	4
11	K1 + K2 + y-Verschiebung	4
12	K2 + K3 + x-Verschiebung	5
13	K0 + K1	2
14	K1 + K2	3
15	K2 + K3	4
16	unverzeichnet + K1	2
17	unverzeichnet + K3	4
18	unverzeichnet + T0	1
19	unverzeichnet + T1	2
20	unverzeichnet + T2	3
21	unverzeichnet + T3	4
22	T0 + T1	2
23	T1 + T2	3

24	T2 + T3	4
25	T0 + T1 + x-Verschiebung	3
26	T1 + T2 + y-Verschiebung	4
27	T2 + T3 + xy-Verschiebung	6
28	T0 + T1 + xy-Verschiebung	4
29	T2 + T3 + x-Verschiebung	5
30	T0 + T1	2
31	T1 + T2	3
32	T2 + T3	4
33	unverzeichnet + T1	2
34	unverzeichnet + T3	4

Tabelle 3: Verzeichnungsparameter und globaler Grad der verursachten Bildveränderungen

Anm.: Diese globale Kategorisierung berücksichtigt noch nicht, dass verzeichnungsbedingt auftretende *vertikale* Disparitäten evtl. gravierende Einflüsse auf die stereopistische Wahrnehmung haben als verzeichnungsbedingte zusätzliche *horizontale Querdisparitäten*.

3.2.2 Feldbezogene Störungsbetrachtung

Die Stärke der Verzeichnung ist, wie bereits geschildert, abhängig von der radialen Position der Bildpunkte vom Verzeichnungszentrum und dessen absoluter Lage. Deshalb wurden die einzelnen Verzeichnungskombinationen innerhalb von neun Bildfeldern à 640 x 360 px detailliert analysiert und in ihrem feldspezifischen Fehlergrad quantifiziert.

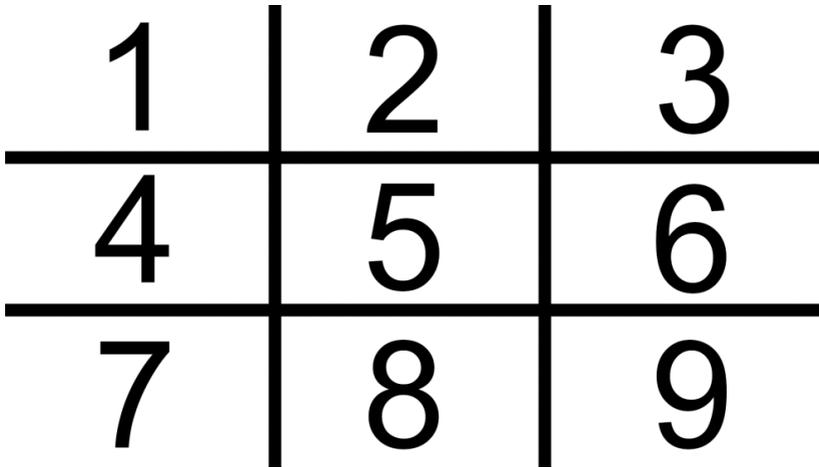


Abbildung 16: Bildfelder für die meßtechnische Analyse und die Probandenbefragung

3.2.3 Vertikale Disparitäten

Zu welchen vertikalen Disparitäten kommt es in den Items mit den genutzten Verzeichnungswerten? Laut Holger Tauer²⁶ führen vertikale Disparitäten bereits ab einem globalen Höhenversatz von 2,5 - 3% zu visueller Überforderung. In diesem Test wurde für alle Items eine Bildhöhe von 1080 Pixeln verwendet. Demnach ergibt sich für 3% ein Höhenversatz von 32 Pixeln. Vereinzelte Items wiesen in den Feldern 1, 3, 7 und 9 bis zu 57Pixel Höhenversatz auf. Item 9 („worst case“) wies im Randfeld 7 einmalig sogar 155px Höhenversatz auf.

Item	Feld 1	Feld 2	Feld 3	Feld 4	Feld 5	Feld 6	Feld 7	Feld 8	Feld 9
1	9,5	1	9,5	0	0	0	9,5	1	9,5
2	19,6	2,6	19,6	0	0	0	19,6	2,6	19,6
3	40,3	6	40,3	0	0	0	40,3	6	40,3
4	57,7	12,3	57,7	5,6	0	5,6	57,7	12,3	57,7
5	9,8	1,6	9,8	0	0	0	9,8	1,6	9,8
6	21,7	2,7	21,7	1,7	0	1,7	21,7	2,7	21,7
7	22,2	5,6	22,2	6,7	1	6,7	22,2	5,6	22,2
8	25	1,6	25	2,5	0	2,5	25	1,6	25

²⁶ Vgl. Tauer Holger (2010), S. 75

9	10,5	3,5	26,3	27,6	4,2	14,5	155	32,7	54
10	4,5	1	5	12	0	2,5	46	12	17
11	3,8	1,6	3,8	11,5	1	11,5	55	15	55
12	62	7,5	2,6	8,5	0,5	0,5	62	6,4	0,5
13	9,8	1,6	9,8	0	0	0	9,8	1,6	9,8
14	21,7	2,7	21,7	1,7	0	1,7	21,7	2,7	21,7
15	22,2	5,6	22,2	6,7	1	6,7	22,2	5,6	22,2
16	19,6	2,6	19,6	0	0	0	19,6	2,6	19,6
17	57,7	12,3	57,7	5,6	0	5,6	57,7	12,3	57,7
18	8,6	1,4	8,6	0,5	0	0,5	8,6	1,4	8,6
19	14	2,5	14	3,6	0	3,6	14	2,5	14
20	26,8	5,6	26,8	3,6	0	3,6	26,8	5,6	26,8
21	47,3	9,6	47,3	6,3	0	6,3	47,3	9,6	47,3
22	7,5	1,5	7,5	0,6	0	0,6	7,5	1,5	7,5
23	15,3	3,3	15,3	1,7	0	1,7	15,3	3,3	15,3
24	22,6	4,6	22,6	1,7	0	1,7	22,6	4,6	22,6
25	18,5	1,2	1	4,6	0	0,5	18,5	1,2	1
26	3,6	1,6	3,5	9,6	0	13,5	45,7	16,6	44,5
27	8,4	3,6	19,5	20	3,6	11,7	96	29	34
28	4	1	4,6	9,3	1	3,3	37,6	10,6	13,6
29	48,5	5,5	1	6,6	1	0	45	6,6	0,5
30	7,5	1,5	7,5	0,6	0	0,6	7,5	1,5	7,5
31	15,3	3,3	15,3	1,7	0	1,7	15,3	3,3	15,3
32	22,6	4,6	22,6	1,7	0	1,7	22,6	4,6	22,6
33	14	2,5	14	3,6	0	3,6	14	2,5	14
34	47,3	9,6	47,3	6,3	0	6,3	47,3	9,6	47,3
35	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0

Tabelle 4: Verzeichnungs-differenz-bedingte vertikale Disparitäten in den verwendeten Testitems

3.2.4 Aniseikonie-Effekte

Aniseikonie beschreibt laut Gerhard K. Lang einen Bildgrößenunterschied zwischen linker und rechter Netzhaut.²⁷ Eine Studie²⁸ von Lovasik und Szymkiw behandelt die Auswirkungen von unterschiedlichen Größen der beiden Teilbilder. Dabei beziehen sich die Größenangaben im Gegensatz zu unserem Test auf das gesamte Bild. Verzeichnungs-differenzen verändern die Größen in den Teilbildern je nach Bildfeld unterschiedlich, was darauf zurückzuführen ist, dass ein Bildpunkt, der nah am Verzeichnungszentrum liegt, wesentlich weniger beeinflusst wird, als ein Bildpunkt am Rand des Bildes, der weit entfernt liegt. Somit ist der direkte Vergleich mit der erwähnten Studie nicht so einfach zu ziehen.

Zur feldbezogenen Bestimmung der Vergrößerung wurden reine Gitter (Grids) genutzt, welche 14x26 gleichgroße Quadrate über den gesamten Bildinhalt besaßen. Nach der Verzeichnung wurde die Größe der verzeichneten Quadrate in jedem der 9 Bewertungsfelder einzeln gemessen. Die Messung erfolgte über die Länge und Höhe der verzeichneten Quadrate. Ein Standardfeld, welches unverzeichnet ist, hatte die Höhe und Länge von 66 Pixeln und damit eine Fläche von 4356px^2 ($66 \text{ Pixel} * 66 \text{ Pixel}$). Dieser gilt als Grundwert und hat damit eine Vergrößerung/Verkleinerung von 0%. Die Vergrößerungen werden immer zusätzlich in Prozent angegeben und beziehen sich auf den genannten Grundwert von 4356px^2 .

Vereinzelte Items wiesen in den Feldern 1, 3, 7 und 9 Vergrößerungswerte bis zu 70% auf. Item12 wies im Randfeld1 einmalig sogar eine 175%ige Vergrößerung vom Grundwert auf.

²⁷ Vgl. Lang Gerhard (2014), S.326

²⁸ Vgl. Vincent, Szymkiw (1985)

4. Setup für die empirische Untersuchung

Zur Beantwortung der zentralen Fragestellung wurde als Methode eine empirische Untersuchung in Form eines Probandentests durchgeführt. Ein solcher Test ist umso aussagekräftiger, desto optimaler die Rahmenbedingungen für die Probanden sind. Ebenfalls wichtig sind stereoskopisch einwandfreie Basis-Items. Die Probanden waren zu keinem Zeitpunkt über die wahre Intention des Tests informiert worden. Offiziell wurde der Test als „Stereo-3D-Wahrnehmungstest“ angekündigt, bei dem es um die Wahrnehmung von S3D-Bildern im Allgemeinen geht. Vor und nach den Bewertungen der Testitems mussten sämtliche Probanden Formulare ausfüllen. Neben einem Mediennutzungsteil gab es einen allgemeinen Teil, in dem nach Sehhilfe, Sehschwächen, Alter und Geschlecht gefragt wurden. Außerdem gab es sowohl vor als auch nach den Itembewertungen einen Befindlichkeitstest. Zur Bestimmung der Stereosehfähigkeit der Probanden wurde der Randot-Sehtest SO-002 der us-amerikanischen Stereo Optical Co., Inc. verwendet. Der Test umfasst ein Tiefenbudget von 20 bis 400 Winkelsekunden bei einem Betrachtungsabstand von ungefähr 40 cm.²⁹ Am Ende des Tests wurden die Probanden zudem noch gefragt, ob es besondere Auffälligkeiten gab und ob sie noch eine anschließende freie Anmerkung machen möchten.

4.1 Ziel der empirischen Untersuchung

Das Ziel dieses Tests war es, herauszufinden, ob und ab welchem Grad Probanden Verzeichnungs-differenzen in stereoskopischen Videosequenzen als störend empfinden. Ebenfalls sollte geklärt werden, ob der Einsatz verschiedenster Bildmaterialien Auswirkung auf die Wahrnehmung dieses Fehlers hat.

4.2 Verwendete Hard- und Software

Um eine fehlerfreie technische Wiedergabe und damit eine unbeeinflusste Bewertung der Videosequenzen zu gewährleisten, galt es im Vorfeld des Tests sehr genau zu überlegen, welche Hardware verwendet werden soll.

Die für die Zuspaltung verwendete Hardware war ein Apple Mac mini, Produktionsjahr 2010. Der Mac mini besaß einen Intel Core 2 Duo Prozessor mit einem Takt von 2,4 Ghz. Er verfügte zudem über 2GB Arbeitsspeicher mit einem Takt von 1066 Mhz. Als Festplatte diente ein SSD-RAID-0 von Samsung mit einer Gesamtspeicherkapazität von 1TB. An dieser Stelle ist zu erwähnen, dass die Verwendung einer SSD („Solid State Disk“) zwingend notwendig ist, da es sich bei den Clips durchweg um unkomprimierte Daten handelt. Bei den dabei aufkommenden Datenmengen ist es nicht möglich, diese flüssig über herkömmliche, magnetische Festplatten abzuspielen.

²⁹ Vgl. <http://eiwebassets.s3.amazonaws.com/s/stereoptical/pdf/stereotests/Randot%20manual.pdf> (zuletzt überprüft am 24.3.2014)

Das Abspielen der Items wurde zum Testzeitpunkt mit dem VLC-Player³⁰ in Version 2.1.4 umgesetzt.

Im Rahmen des Masterstudiengangs Medieninformatik (MIM) der Hochschule Furtwangen forschten drei Studierenden innerhalb des Tests über die Augenbewegungen der Probanden bei der Betrachtung von stereoskopischem Video- und Bildmaterial. Zu diesem Zweck wurden zusätzlich eine spezielle Kinnstütze und eine mit Infrarot-LEDs bestückte Brille entwickelt, die ebenfalls im Test mit verwendet wurden. Die Kinnstütze ermöglicht eine stabile Kopfposition der Probanden und dadurch bessere Bedingungen zum Abfilmen der Augen.

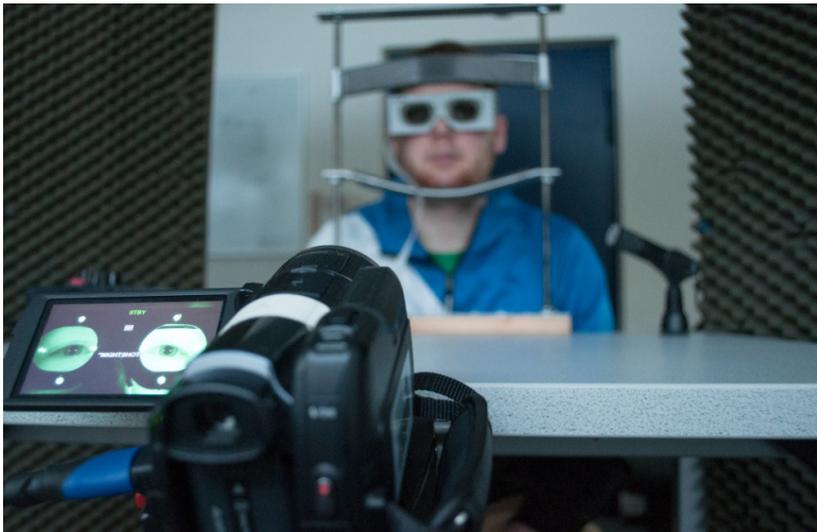


Abbildung 17: Augenaufnahmevorrichtung der Masterprojektgruppe

4.2.1 3D-TV Gerät

Beim verwendeten 3DTV-Gerät handelte es sich um das Modell 50PFL5008K/12 des Elektronikherstellers Philips. Der LCD-Fernseher hat eine Bildschirmdiagonale von 50 Zoll / 127cm. Durch die verwendete Interlace-Polfiltertechnik wurden die Originalvideos in ihrer vertikalen Auflösung jeweils auf 540 Pixel halbiert. Dies wurde durch die Zuspiegelung vorab berücksichtigt (over/under interlaced).

30 * <http://www.videolan.org/vlc/> (zuletzt überprüft am 1.3.2014)

4.3 Pretest mit Experten

Im Januar 2014, circa zwei Monate vor dem eigentlichen Probandentest, wurden Experten probeweise mit unterschiedlich verzeichnetem S3D-Material konfrontiert, um die Eignung der Items und der Testumgebung zu evaluieren. Die Items wurden Professoren, Studenten und Mitarbeitern der Hochschule Furtwangen sowie Kooperationspartnern der Technischen Universität Ilmenau vorgeführt und diese um eine visuelle Bewertung gebeten. Dabei wurden ähnliche Verzeichnungswerte verwendet, wie sie später von Polynom K2 und K3 (siehe Abbildung 9 und 10) verwendet wurden. Es handelte sich bei dem gezeigten Material um verschiedene Szenen aus dem S3D-Film Street Dance.

Die Experten bemerkten - trotz starker Verzeichnungsabweichungen - kaum Störungen im Bild. Es stellte sich schnell heraus, dass dies mit hoher Wahrscheinlichkeit mit dem Bildinhalt der Szenen zusammenhing. Die für diesen Vortest produzierten Clips hatten an den Bildrändern kaum geometrische Details. Grund waren zum Teil sehr dunkle Hintergründe oder große selektive Unschärfen, die dem visuellen System kaum Korrespondenzansätze boten. Diese Erkenntnis wurde daraufhin im späteren Haupttest durch eine geeignetere Wahl der Items und durch die Definition der „Strukturtypen A, B, C“ berücksichtigt.

4.4 Testaufbau

Der Test fand vom 24. März bis einschließlich 5. April 2014 im N-Bau der Hochschule Furtwangen statt. Da der Raum keine Fenster besitzt, konnte die Helligkeit im Raum optimal mit zwei „Kinoflos“ (im Hochfrequenzbetrieb) angepasst werden. Sie beleuchteten die weiße Wand hinter dem TV-Gerät. Die Helligkeit hinter dem Fernsehgerät wurde mit einem Konica Minolta XY-DC Chroma Meter gemessen und betrug konstant 175 lux mit einer Farbtemperatur von 6400 Kelvin. Der Abstand der Probanden zum Fernsehgerät betrug, gemäß der in Rec ITU-BT.500-13³¹ definierten PVD (preferred viewing distance), 3,3m.

31 Vgl. https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf, S.4 (zuletzt geprüft am 22.01.2014).

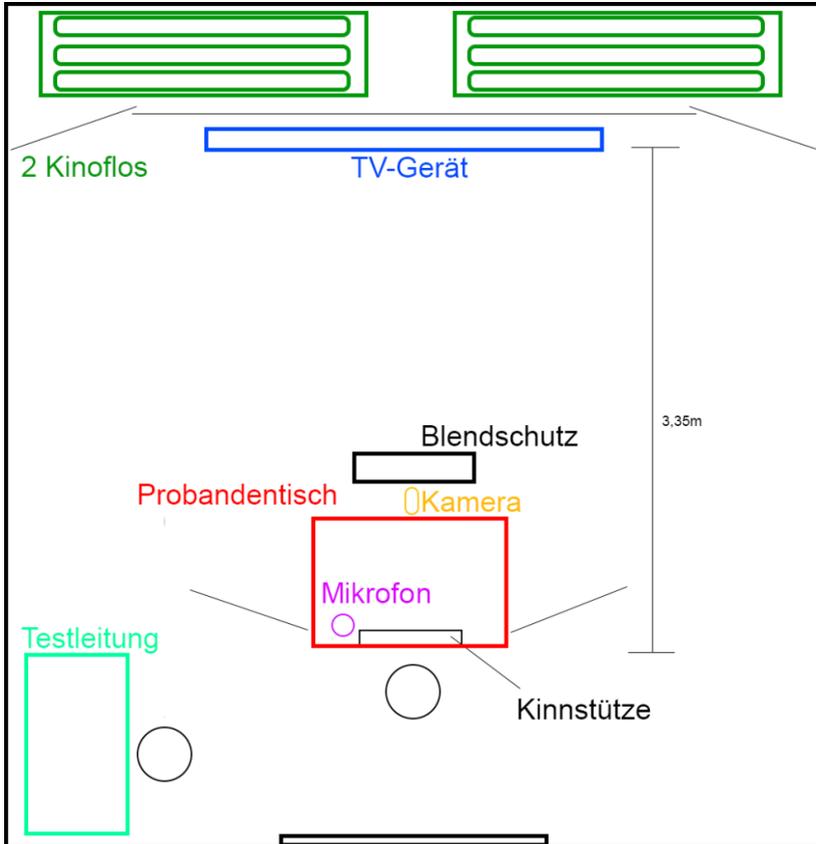


Abbildung 18: Aufsicht des Testaufbaus

Außerdem galt es, spiegelnde Gegenstände im Raum von den Wänden zu hängen oder bei Bedarf abzukleben, um Reflexionen zu vermeiden. Hierfür kamen Molltonstoffe zum Einsatz. Vor jedem Testbeginn wurden die Kinoflos und der Fernseher aufgewärmt, um deren volle und gleichmäßige Beleuchtungsleistung sicherzustellen. Die Bewertungen der Probanden wurden mit einem Mikrofon, das am Rand des Probandentischs stand, aufgezeichnet. Eine Aufsicht des Testaufbaus ist in Abbildung 18 zu sehen.

In den Abbildungen 19 und 20 ist der Aufbau nochmals photographisch dokumentiert.



Abbildung 19: Testaufbau - Foto 1



Abbildung 20: Testaufbau - Foto 2

4.5 Konzipierter Testablauf

01. Der Proband wird begrüßt. Für die aufgebrauchte Zeit wird gedankt. Es wird darauf hingewiesen, dass der Test freiwillig ist und jederzeit abgebrochen werden kann. Der Ablauf des Tests wird kurz erläutert. Es wird gefragt, ob die Person an Epilepsie und/oder ähnlichen Krankheiten leidet. Sollte dies der Fall sein, wird der Test mit dieser Person nicht durchgeführt.
02. Die Einverständniserklärung wird zum Lesen vorgelegt. Der Proband soll diese lesen und nur bei Zustimmung unterschreiben.
03. Der Fragebogen zur Mediennutzung soll ausgefüllt werden.
04. Der Randot-Sehtest wird erklärt und durchgeführt. Schafft der Proband mindestens Stufe 5 oder mehr, darf fortgefahren werden.
05. Der Befindlichkeitstest 1 wird ausgehändigt.
06. Die Funktionsweise der Kinnstütze wird erklärt. Es muss dafür gesorgt werden, dass der Proband bequem sitzt, da die Stützeinrichtung wenig Bewegungsfreiheit bietet.
07. Die Ton- und Augenaufnahme wird gestartet.
08. Das Kalibrierungsvideo für das Gaze-Tracking wird gestartet.
09. Die Eingewöhnungsphase wird gestartet.
10. Das Bewertungssystem und die Bildschirmfelder werden erklärt. Die Items folgen und damit der beginnt die Testreihe.
11. Der Befindlichkeitstest 2 wird ausgehändigt.
12. Es folgt der Abschlussfragebogen. Diese Fragen werden von der Testleitung mündlich gestellt und dementsprechend notiert.
13. Der Proband kann zum Abschluss beliebige Fragen stellen oder Anmerkungen machen. Im Anschluss wird dem Probanden für die Zeit gedankt.

4.6 Produktion der verzeichneten Items

Die synthetische Verzeichnung von Bild- und Videomaterial ist keine alltägliche Praxis, weshalb sich die Suche nach geeigneter Software als schwierig erwies. Da die Verzeichnung häufig im Kontext der Verzeichnungskorrektur eine Rolle für Nutzer spielt, lag es nahe, von dieser Perspektive aus die Recherche zu beginnen. Mehrere Hersteller bieten Software zur Korrektur von Verzeichnung an.

Eine Marktrecherche ergab folgende in Frage kommende Programme (Stand: 2014):

1. Photoshop / Lightroom
2. NukeX
3. PTLens
4. ShiftN
5. Dxo Viewpoint 2
6. Cyberlink Photodirector 5
7. Capture NX2

Alle Programme bieten mehr oder weniger ähnliche Instrumente zur Korrektur von Tonnen- und Kissenverzeichnung in Bildern. Für die Arbeit mit Videoformaten ist jedoch nur NukeX vom Hersteller The Foundry geeignet. Im Rahmen dieser Arbeit wurde die Versionsnummer NukeX 7.0v9 verwendet. Das Programm ermöglicht durch den „LensDistortion“-Node eine präzise Verzeichnungskorrektur. Umgekehrt lässt sich eine vorgegebene Verzeichnung auf nahezu jedes beliebige Bild- und Videomaterial auch direkt anwenden. Die Prozesskette wird in nachfolgender Abbildung ersichtlich.

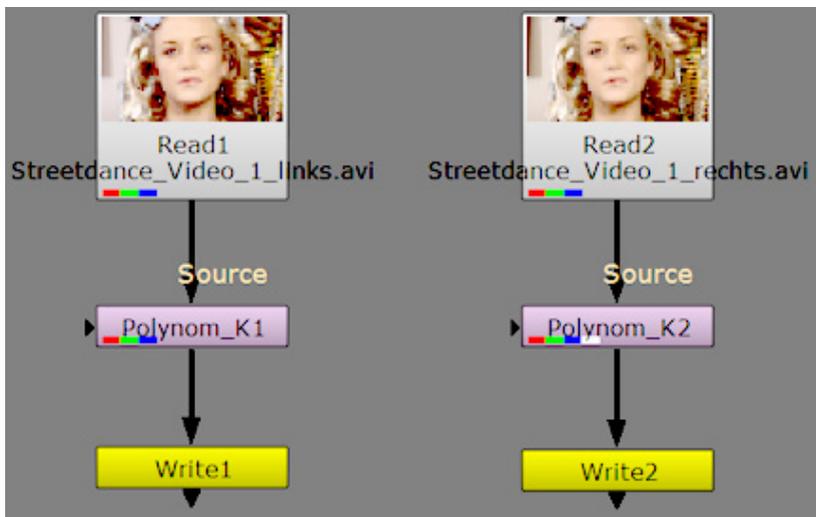


Abbildung 21: Prozesskette Verzeichnung in NukeX

Die Lens-Distortion Nodes, die je nach linkem oder rechtem Bild unterschiedlich angewendet werden mussten, wurden in „Polynom_K1“ und so weiter umbenannt, um eine bessere Wiederverwendbarkeit für die anderen Items zu ermöglichen.

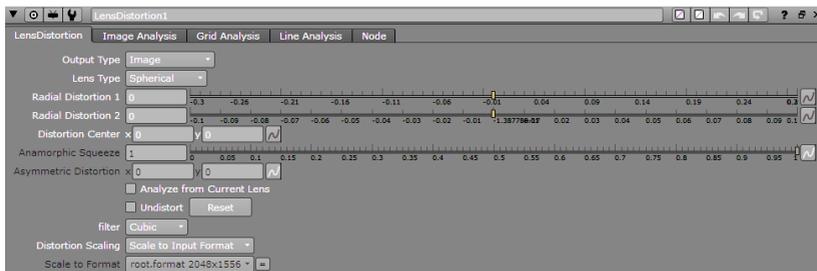


Abbildung 22: LensDistortion Node

Die LensDistortion Node bietet die Möglichkeit, die Verzeichnung im Wesentlichen über zwei Parameter zu steuern, Radial Distortion 1 und 2. Laut Handbuch³² definiert Radial Distortion 1 den Koeffizienten proportional zu r^2 , wobei r die Distanz vom Verzeichnungszentrum darstellt. Bei Radial Distortion 2 ist der Koeffizient proportional zu r^4 . Die beiden Werte müssen immer im selben Vorzeichenbereich liegen, da man ansonsten eine Mischverzeichnung aus Tonnen- und Kissenverzeichnung erzeugt.

32 Vgl. <http://thefoundry.s3.amazonaws.com/products/nuke/releases/7.0v10/Nuke7.0v10-User-Guide.pdf>, S.644 (zuletzt überprüft am 20.01.2014)

Der Wertebereich für Radial Distortion 1 liegt zwischen $-0,3$ und $+0,3$ und für Radial Distortion 2 zwischen $-0,1$ und $+0,1$. Positive Werte erzeugen eine Kissenverzerrung und negative Werte eine Tonnenverzerrung. Wichtig ist, dass der Haken bei „Undistort“ nicht aktiviert sein darf, denn sonst wird die Verzerrung invertiert. Das Verzerrungszentrum kann mit den Distortion Center Werten für x und y verschoben werden. Der x -Wert verschiebt das Verzerrungszentrum horizontal und der y -Wert vertikal.

Das Write-Objekt wurde genutzt, um die verzeichneten Objekte zu speichern. Das Speichern erfolgte als Einzelbildersequenz im DPX-Format (Digital Picture Exchange). So konnte sichergestellt werden, dass keinerlei ungewollte Einflüsse, wie zum Beispiel Kompressionsartefakte, entstehen. Im späteren Verlauf mussten die Einzelbilder durch den Import als „Bildsequenz“ in Adobe Premiere wieder zu einer Videodatei zusammengefügt werden. Beim Import wird standardmäßig eine Bildrate von 25 Bildern pro Sekunde angenommen, die dann auf den Originalwert des Videos von 24 Bildern pro Sekunde korrigiert werden musste.

Besondere Aufmerksamkeit soll an dieser Stelle der synthetischen Tonnenverzerrung gelten. Folgender Vergleich soll deren Problem verdeutlichen:



Abbildung 23: StreetDance Video 1 unverzeichnet



Abbildung 24: StreetDance Video 1 maximal kissenverzeichnet (Radial Distortion 1: +0,3)

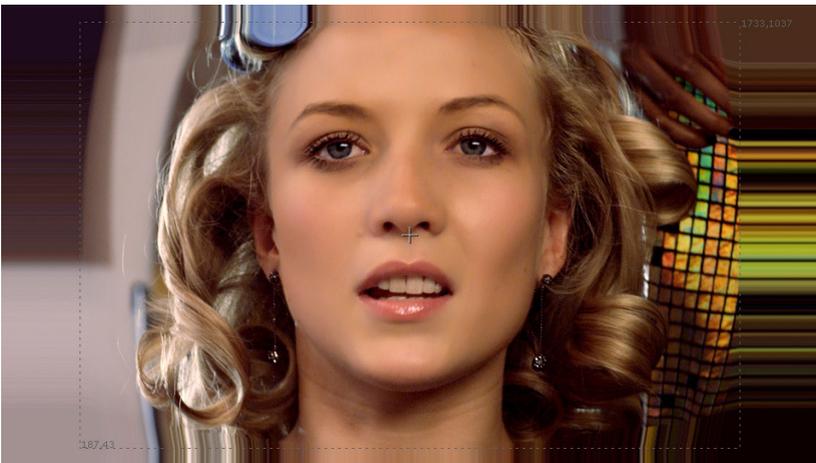


Abbildung 25: Streetdance Video 1 maximal tonnenverzeichnet Screenshot in Nuke (Radial Distortion 1: -0,3)

Beim Betrachten der letzten Abbildung fällt sofort der störende Randbereich auf. Durch die tonnenförmige Stauchung des Bildes können die Randbereiche nur noch durch sich wiederholende Pixel gefüllt werden.

NukeX erstellt bei der Tonnenverzeichnung automatisch einen Rahmen um das Bild mit genauen Pixelangaben, an welchen Stellen der Rahmen positioniert ist. Da die

radiale Verzeichnung immer rotationssymmetrisch ist und somit das Bild beim Stauchen an den Rändern abrundet, ist der Teil des Originalbildes in der Mitte nicht mehr mit geraden Kanten ausgestattet. Der Hilfsrahmen zeigt jeweils die Kante des äußersten Pixels in Richtung des Bildrandes. Der gravierende Eingriff in die Randbereiche des Bildes durch die Tonnenverzeichnung ist visuell nicht zu ignorieren. Für den Test schien es notwendig, zumindest die äußersten Ränder des Bildes möglichst so zu bearbeiten, dass die Wahrnehmung der Probanden nicht stärker gestört und abgelenkt wird als notwendig. Durch die pixelgenauen Angaben durch NukeX wurde eine Crop-Node genutzt, um die Ränder am bereits erläuterten äußersten Pixel aller Ränder gerade abzuschneiden. Sämtliche Bilder, die der Tonnenverzeichnung unterzogen wurden, wurden nach dieser Methode optimiert.

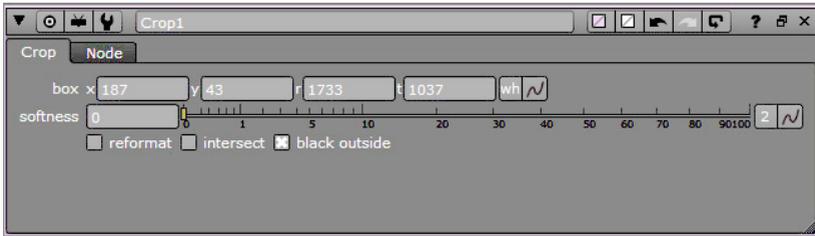


Abbildung 26: Crop-Node

In Abbildung 26 sind die Crop-Node sowie die vom LensDistortion Node angezeigten äußersten Pixelangaben zu sehen. Diese wurden direkt übernommen und beschneiden das Bild auf die gewünschte Weise, was zu folgendem Ergebnis führt:

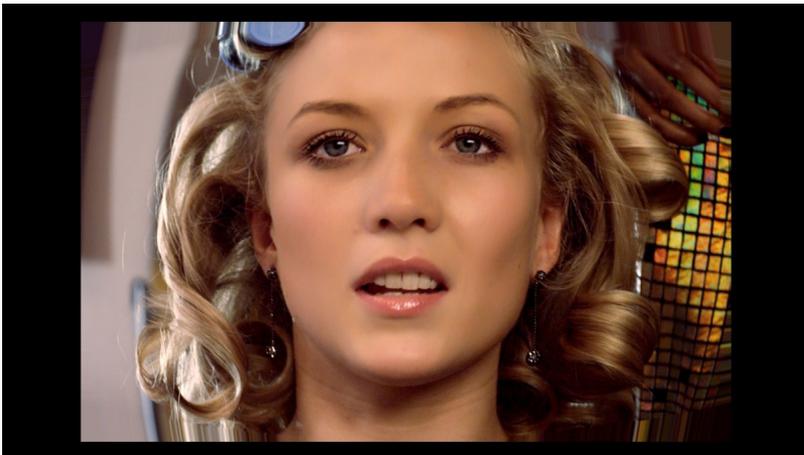


Abbildung 27: Streetdance Video 1 zugeschnitten

Die Beschneidung wurde auf beide Teilbilder gleichermaßen angewendet. Bei Verzeichnungskombinationen, bei denen beide Bilder tonnenverzeichnet wurden, wurde stets die größere Beschneidung an beiden Bildern vorgenommen, um auszuschließen, dass es Bildbereiche gibt, die nur für das eine Auge vorhanden sind. Bei dem genutzten Maximalwert der Tonnenverzeichnung entsteht nach der dazugehörigen Beschneidung durch die Crop-Node ein sehr großer Rand, der nicht mehr mit Bildinformationen gefüllt ist. Obiges Extrembeispiel wurde im Test nicht verwendet und dient nur der prinzipiellen Illustration des Problems.

Item	Linkes Bild	Rechtes Bild	Angewandt auf	Crop-Werte
18	Unverzeichnet	Polynom T0	Street Dance Standbild 1	1906 x 1078
19	Unverzeichnet	Polynom T1	Louise Video 1	1893 x 1076
20	Unverzeichnet	Polynom T2	Randot Standbild 3	1870 x 1072
21	Unverzeichnet	Polynom T3	Street Dance Standbild 4	1836 x 1066
22	Polynom T0	Polynom T1	Street Dance Standbild 6	1893 x 1076
23	Polynom T1	Polynom T2	Louise Standbild 1	1870 x 1072
24	Polynom T2	Polynom T3	Randot Video 2	1836 x 1066
25	Polynom T0	Polynom T1 $x_0 = 0.25$	Street Dance Video 3	1910 x 1075
26	Polynom T1	Polynom T2 $y_0 = 0.25$	Louise Video 2	1866 x 1079
27	Polynom T2	Polynom T3 $x_0, y_0 = 0.25$	Street Dance Video 2	1881 x 1076
28	Polynom T0	Polynom T1 $x_0 = 0.25, y_0 = 0.25$	CGI Video 3	1909 x 1079
29	Polynom T2	Polynom T3 $x_0 = 0.25$	CGI Video 1	1885 x 1063
30	Polynom T0	Polynom T1	Randot Standbild 1	1893 x 1076
31	Polynom T1	Polynom T2	Louise Standbild 4	1870 x 1072
32	Polynom T2	Polynom T3	CGI Standbild 3	1836 x 1066
33	Unverzeichnet	Polynom T1	Louise Video 3	1893 x 1076
34	Unverzeichnet	Polynom T3	Louise Standbild 3	1836 x 1066

Tabelle 6: Tonnenverzeichnete Items und ihre Cropwerte

Die in der Tabelle genannten Pixelangaben bedeuten zum Beispiel für Item 18, dass es am rechten Rand am Pixel 1906 den letzten originalen Pixel hatte, wodurch es also 14 Pixel zum Rand hin verliert. Aus diesem Grund muss es auf der gegenüberliegenden Seite um genau die gleiche Differenz abgeschnitten werden. Die zwei Pixel, welche am oberen Rand abgeschnitten werden, müssen ebenfalls am unteren Rand entfernt werden. Obwohl alle tonnenverzeichneten Items durch dieses Vorgehen ihren verzerrten Rand zum Großteil verlieren, existiert aufgrund der Abrundung des Originalbildes ein Bereich, der weiterhin störende Auswirkungen auf die Wahrnehmung haben könnte.

Auch bei der synthetischen Verzeichnung von Randot-Videos ergibt sich ein grundsätzliches Problem. Die Ursache kann durch eine Ausschnittsvergrößerung deutlich gemacht werden.

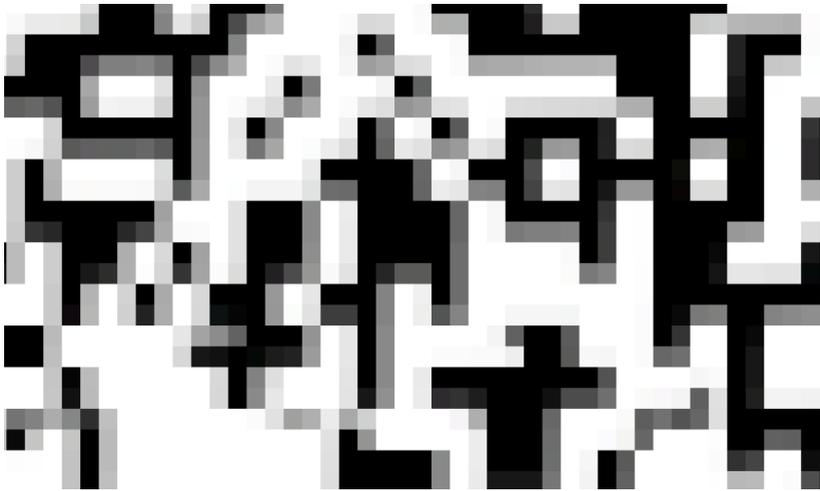


Abbildung 28: Vergrößert Randot Video 1 (Frame 30) nach Verzeichnung

In einem Random-Dot Bild dürfen nur schwarze oder weiße Pixel vorkommen. Durch die synthetische Verzeichnung und deren Interpolation entstehen Zwischenwerte, die als Grauwerte abgelegt werden. Diese erzeugen ein störendes Muster über das gesamte Bild. Unsere Lösung: NukeX bietet für die Farbkorrektur einen Node namens ColorLookup. Mit diesem ist es möglich, Farbwerte zu filtern.

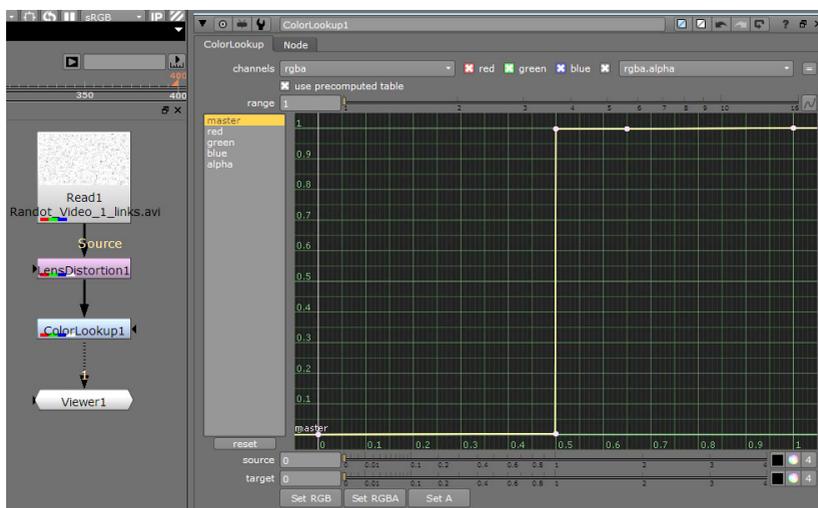


Abbildung 29: Colorlookup Node - Filter von Grauwerten

Die horizontale Achse zeigt den Inputwert und die vertikale den zugeordneten Outputwert.³³ So ist es möglich, alle Grauwerte ab- oder aufzurunden. Es wurde ein Schwellenwert von 0,5 gewählt, wobei 0 für schwarz und 1 für weiß steht. Alle Werte, die heller als der Schwellenwert sind, werden weiß. Umgekehrt werden darunterliegende, also dunklere Punkte zu schwarz. Der ausgewählte Kanal ist der Masterkanal, der sich auf alle Farbkanäle auswirkt. Nach der Anwendung sieht der korrigierte Ausschnitt wie folgt aus:

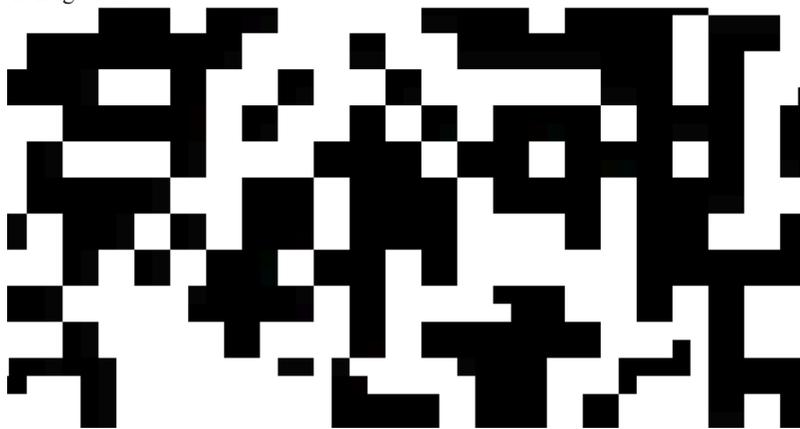


Abbildung 30: Vergrößert Randot Video 1 (Frame 30) nach ColorLookup Korrektur

33 Vgl. <http://thefoundry.s3.amazonaws.com/products/nuke/releases/7.0v10/Nuke7.0v10-User-Guide.pdf>, S.70 (zuletzt überprüft am 20.01.2014)

Sämtliche Mittelwerte sind entfernt und das Gesamtbild kann genutzt werden. Die Eignung von Random-Dot-Material für eine nachträgliche Verzeichnung ist sicher nicht optimal, dennoch wurde diese bewusst eingesetzt, um den Einfluss von Verzeichnungs-differenzen möglichst isoliert nur auf die patente Stereopsis zu testen.

Um das verfügbare Bildmaterial zu erweitern, wurden zusätzlich zwei CGI-Szenen in 3ds Max 2014 von Autodesk entworfen. Zur Umsetzung in stereoskopischem 3D wurde das Plugin 3DHIPPIE Stereocam³⁴ verwendet, das die einfache Einstellung von Nahpunkt, Fernpunkt, Scheinfester, Brennweiten und so weiter ermöglicht. Die beiden erstellten CGI-Szenen zeigen einfache geometrische Körper, die sich in unterschiedlichen Abständen befinden. CGI-Items sind für einen solchen Test besonders gut geeignet, da sie keinerlei optischen Abbildungsfehlern von Objektiven unterliegen. Zudem sind sämtliche stereoskopische Parameter optimal über die Software konfigurierbar und unterliegen der vollen Kontrolle des Nutzers.³⁵

Abschließend wurden über Nuke die maximalen relativen Parallaxen aller verzeichneten Items ermittelt, um stereoptische Grenzerfahrungen der Probanden auszuschließen.

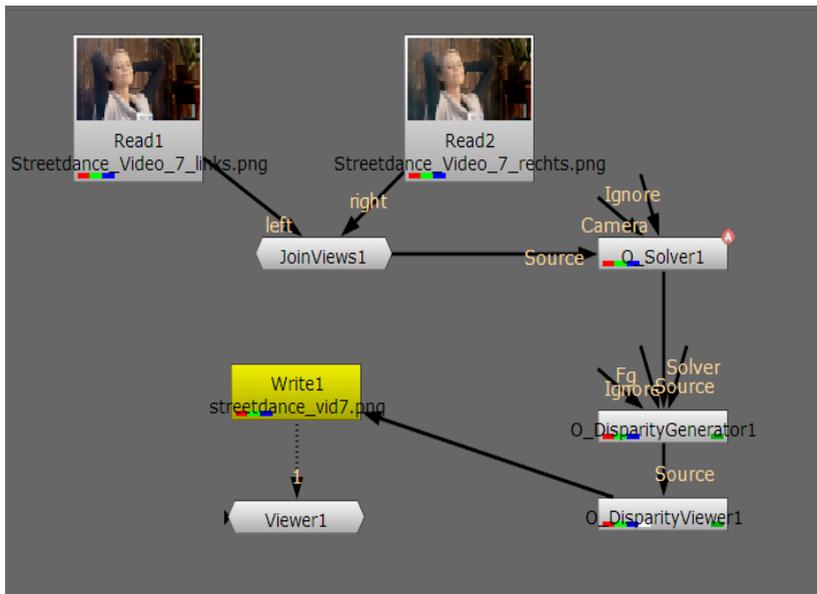


Abbildung 31: Nuke Screenshot vom Ablauf zur Erzeugung der Parallaxen-Histogramme

34 Vgl. <http://davidshelton.de/blog/?p=431> (zuletzt geprüft am 10.04.2014)

35 Vgl. Tauer Holger (2010), S. 282 f.

Durch ein Tool von Studenten der Hochschule Furtwangen konnten die ausgelösten Querdisparitäten in Winkelminuten errechnet werden. Da es sich bei vielen Items um Videos und nicht um Einzelbilder handelt, wurden stellvertretende Screenshots gewählt.

Item	Linkes Bild	Rechtes Bild	Medientyp	Tiefenbudget / rel. Disparität in Winkelminuten (unverzeichnetes Item)	Tiefenbudget / rel. Disparität in Winkelminuten (verzeichnetes Item)
	Kissenverzeichnung				
1	Unverzeichnet	Polynom K0	Street Dance Video 1	19,5°	24,3°
2	Unverzeichnet	Polynom K1	Randot Video 4	-	-
3	Unverzeichnet	Polynom K2	Street Dance Video 5	24,3°	46,7°
4	Unverzeichnet	Polynom K3	CGI Standbild 1	25,4°	37,2°
5	Polynom K0	Polynom K1	Louise Video 4	11,2°	39,7°
6	Polynom K1	Polynom K2	Street Dance Video 4	20,1°	32°
7	Polynom K2	Polynom K3	CGI Video 2	18,9°	41,4°
8	Polynom K0	Polynom K1 $x_0 = +0.25$	Randot Video 1	-	-
9	Polynom K2	Polynom K3 $x_0 = +0.25,$ $y_0 = +0.25$	CGI Standbild 2	18,3°	37,2°

10	Polynom K0	Polynom K1 $x_0 = 0.25,$ $y_0 = 0.25$	Street Dance Video 6	16°	$34,3^\circ$
11	Polynom K1	Polynom K2 $y_0 = +0.25$	Street Dance Standbild 2	$17,6^\circ$	$32,8^\circ$
12	Polynom K2	Polynom K3 $x_0 = +0.25$	Street Dance Standbild 5	$24,9^\circ$	26°
13	Polynom K0	Polynom K1	Randot Standbild 2	-	-
14	Polynom K1	Polynom K2	Randot Video 3	-	-
15	Polynom K2	Polynom K3	Louise Standbild 2	$27,2^\circ$	$47,9^\circ$
16	unverzeichnet	Polynom K1	Street Dance Video 7	$10,1^\circ$	$39,7^\circ$
17	unverzeichnet	Polynom K3	Street Dance Standbild 7	$10,4^\circ$	71°
	Tonnenverzeichnung				
18	Unverzeichnet	Polynom T0	Street Dance Standbild 1	$29,6^\circ$	39°
19	Unverzeichnet	Polynom T1	Louise Video 1	$47,3^\circ$	$35,3^\circ$
20	Unverzeichnet	Polynom T2	Randot Standbild 3	-	-
21	Unverzeichnet	Polynom T3	Street Dance Standbild 4	$21,3^\circ$	$59,1^\circ$

Setup für die empirische Untersuchung

22	Polynom T0	Polynom T1	Street Dance Standbild 6	16°	37,2°
23	Polynom T1	Polynom T2	Louise Standbild 1	31,3°	31,3°
24	Polynom T2	Polynom T3	Randot Video 2	-	-
25	Polynom T0	Polynom T1 $x_0 = +0.25$	Street Dance Video 3	2,9°	29°
26	Polynom T1	Polynom T2 $y_0 = +0.25$	Louise Video 2	36,6°	40,2°
27	Polynom T2	Polynom T3 $x_0 = +0.25,$ $y_0 = +0.25$	Street Dance Video 2	20,1°	46,1°
28	Polynom T0	Polynom T1 $x_0 = +0.25,$ $y_0 = +0.25$	CGI Video 3	30,8°	30,8°
29	Polynom T2	Polynom T3 $x_0 = +0.25$	CGI Video 1	28,4°	53,2°
30	Polynom T0	Polynom T1	Randot Standbild 1	-	-
31	Polynom T1	Polynom T2	Louise Standbild 4	11,8°	37,8°
32	Polynom T2	Polynom T3	CGI Standbild 3	30,8°	32,8°
33	Unverzeichnet	Polynom T1	Louise Video 3	20,8°	27,2°
34	Unverzeichnet	Polynom T3	Louise Standbild 3	22,5°	51,5°
35	Unverzeichnet	Unverzeichnet	Sauber1	18,9°	18,9°

36	Unverzeichnet	Unverzeichnet	Sauber2	14,8°	14,8°
37	Unverzeichnet	Unverzeichnet	Sauber3	26°	26°
38	unverzeichnet	Unverzeichnet	Sauber4	25,4°	25,4°

Tabelle 7: Finale Items, Verzeichnungskombinationen und resultierende Querdisparitäten

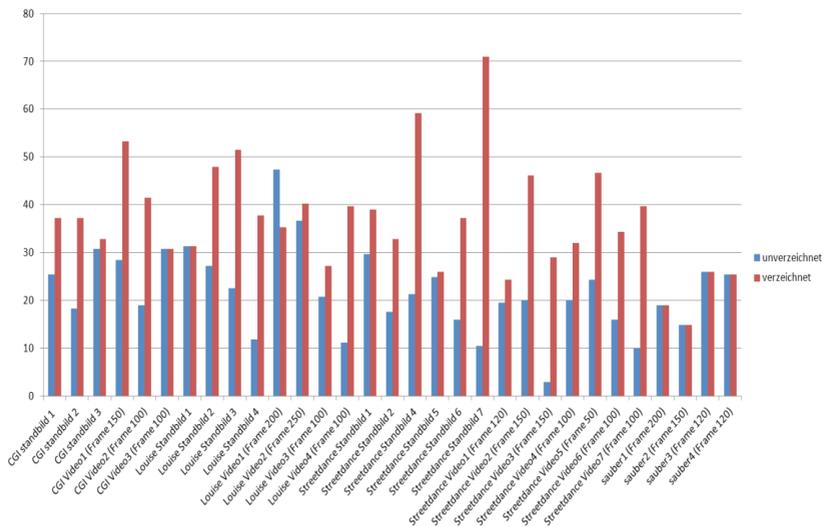


Abbildung 32: Erhöhung der Querdisparitäten durch die Verzeichnungsdifferenzen (in Winkelminuten aus Betrachtungsposition)

4.6.1 Aufbau einer Testsequenz

Die Probandenvideos wurden in ihrem Ablauf so an die Gegebenheiten angepasst, dass der Ablauf des Tests möglichst flüssig absolviert werden konnte. Das Grundmuster des gesamten Clips, den jeder Proband zu sehen bekam, sah - angelehnt an subjektive Tests der EBU - wie folgt aus:

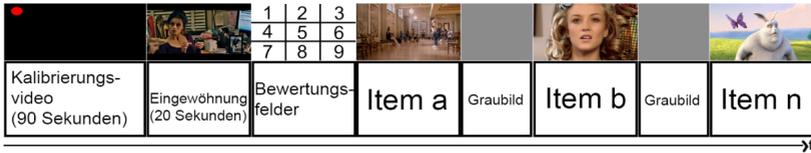


Abbildung 33: Grundmuster - Probandenvideo

Jedes Probandenvideo begann mit dem Kalibrierungsvideo für die Augendetektion. Anschließend sahen die Probanden einen 20-sekündigen unbearbeiteten Clip aus dem S3D-Film „Hybrid“, der zur Eingewöhnung an die Tiefenwahrnehmung dienen sollte. Darauf folgte das beschriebene Bewertungsfeld. Danach kamen die zu bewertenden Items, jeweils durch ein 50%-Graubild getrennt. Sowohl das Bewertungsbild als auch die Graubilder hatten zwar im Clip eine feste Dauer eingestellt, wurden jedoch während des Tests von der Testleitung immer pausiert, um dem Probanden die notwendige Zeit für Bewertungen beziehungsweise die Erklärung des Bewertungsfeldes zu geben. Erst wenn ein Proband mit seiner Bewertung abgeschlossen hatte, wurde das pausierte Video fortgesetzt.

Jeder Proband bekam seine eigene, randomisierte Itemliste. Zur Erstellung der Listen wurde der „List Randomizer“ von random.org genutzt.³⁶

4.7 Probandenpanel

Die Anzahl der für den Test ausgewählten Probanden betrug letztendlich $N=27$ (Anm.: Einer der 28 angesprochenen Probanden gab an, unter einer Farbsehschwäche zu leiden. Er konnte ergänzend selbst Stufe 1 des Randot-Sehtests nicht richtig auflösen. Trotzdem entschied die Testleitung, dem Probanden noch die Eingewöhnungs-Sequenz zu zeigen. In dieser konnte der Proband ebenfalls keinerlei Tiefeninformati-
onen wahrnehmen. Aus diesem Grund musste der Test für diesen Probanden endgültig abgebrochen werden).

Das Alter der Probanden erstreckte sich von 19 bis 28 Jahren. Mit einem Durchschnittsalter von 23,5 Jahren waren die Probanden vergleichsweise jung. Prinzipiell nimmt die Stereosehfähigkeit mit zunehmendem Alter ab, allein schon aufgrund der Tatsache, dass die Schärfeleistung der Augen mit zunehmendem Alter zurückgeht.³⁷

36 Vgl. <http://www.random.org/lists/> (zuletzt überprüft am 4.04.2014)

37 Vgl. http://www.gbe-bund.de/oowa921-install/servlet/oowa/aw92/dboowasys921.xwdevkit/xwd_init?gbe.isgbetol/xs_start_neu/&p_aid=3&p_aid=15888621&nummer=758&p_sprache=D&p_indsp=-&p_aid=61617984 (zuletzt überprüft am 1.05.2014).

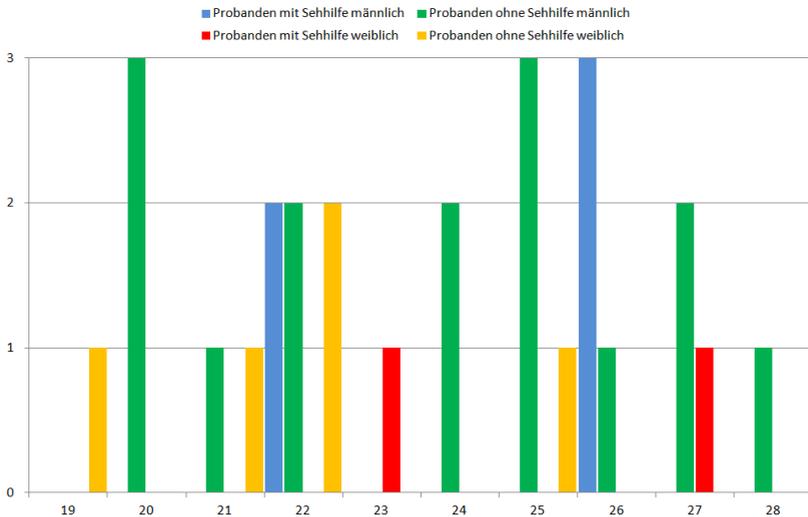


Abbildung 34: Altersverteilung der Probanden in Bezug auf Geschlecht und Sehhilfe

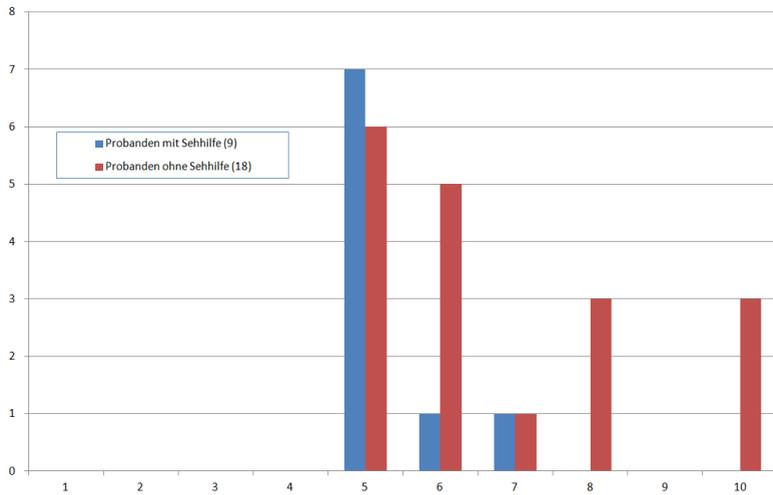


Abbildung 35: Verteilung der Stereosehfähigkeit (x-Achse) der Probanden gemäß Randot Visual 3D-Test. Ein höherer Wert bedeutet eine bessere binokulare Tiefenauflösungsfähigkeit

In Abbildung 35 zur stereoptischen Sehfähigkeit der Probanden fällt bei der höchsten Tiefenauflösungsfähigkeit 10 sofort auf, dass es dort nur Probanden gibt, die keine Sehhilfe nutzen. Die Stufen 5 und 6 wurden im Verhältnis sehr häufig erreicht. Letztendlich hatten alle 27 Probanden grundsätzlich eine höhere stereoptische Auflösungsfähigkeit, als es der verwendete Monitor (bezogen auf ihre Sitzposition) überhaupt bieten konnte.

Bereits am Anfang des Tests bekamen die Probanden Fragen bezüglich ihrer Medienvertrautheit und Mediennutzung gestellt. Die Probanden gaben ihre Vertrautheit mit Medien auf einer Skala von 0-6 (0 für nicht vertraut und 6 für sehr vertraut) nicht unter 3 an. Ein Großteil verwendete sogar den Wert 4 oder höher. Diese Zuordnung könnte bedeuten, dass die Testteilnehmer aufgrund ihrer Medienvertrautheit möglicherweise zu einer strengeren, beziehungsweise einer detaillierteren Bewertung neigten.

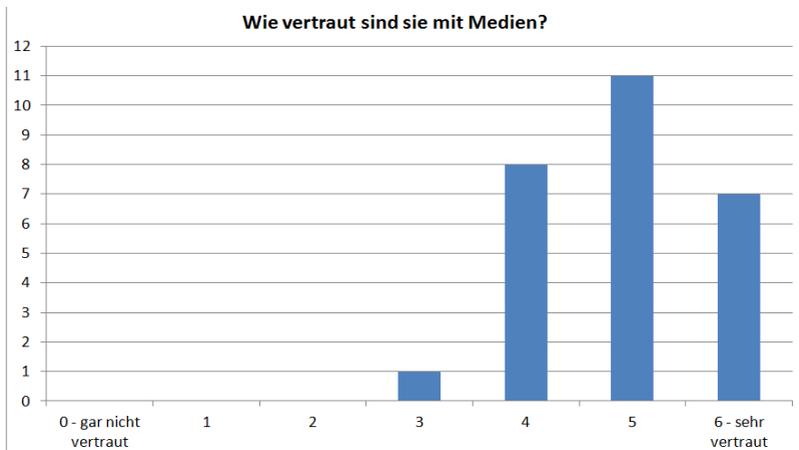


Abbildung 36: Probanden und ihre Vertrautheit mit Medien

Einen etwas anderen Eindruck vermittelt hingegen die Mediennutzung der Probanden. Dabei scheinen die Probanden einen signifikant niedrigeren Teil im S3D-Bereich zu verorten, der aus 3D-Kino, 3D-TV und S3D-Spielen besteht. So gaben 17 Probanden (62,9%) an, nie S3D-Spiele zu konsumieren.

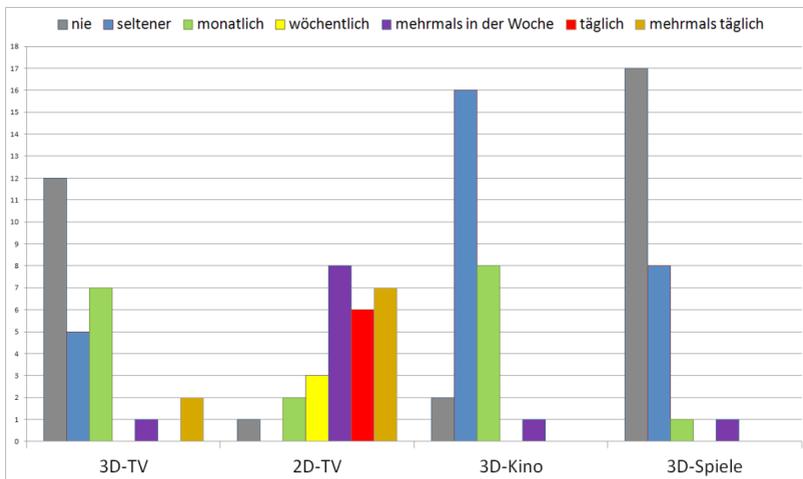


Abbildung 37: Mediennutzung der Probanden

Zwölf Probanden (44,4%) gaben an, kein 3D-TV zu nutzen. Deutlich anders hingegen sah es mit 3D-Kino aus, das den größten Teil der S3D-Unterhaltung unter allen Probanden repräsentiert. Nur zwei Personen (7,4%) gaben an, nie ein 3D-Kino besucht zu haben. Der Großteil von insgesamt 16 Probanden (59,2%) gab hingegen an, selten ins 3D-Kino zu gehen, monatlich immerhin noch acht Probanden (29,6%). Einer der Probanden gab an, mehrmals in der Woche 3D-Kinos zu besuchen.

An der Spitze aller Mediennutzung unserer Probanden stand also eindeutig das 2D-TV, das von sieben (25,9%) Probanden mehrmals täglich, von 6 (22,2%) täglich und von 8 (29,6%) mehrmals in der Woche rezipiert wird. Die restlichen 6 (22,2%) gaben an, weniger als wöchentlich 2D-TV zu konsumieren.

Teil 2
Empirische Untersuchung zur Auswirkung
differierender Objektivverzeichnungen in
Stereo-3D-Produktionen – Auswertung

Silvia Becker, Sandra Beuck, Tolga Sipahi, Arthur Wittmann

Abbildungsverzeichnis

Abbildung 1: Befindlichkeitstest 1 und 2. 100% stehen für alle 16 Antworten der 27 Probanden (16x27 = 432 Antworten).....	86
Abbildung 2: Angaben der Probanden zu ihrer Befindlichkeit vor dem Testdurchlauf (Befindlichkeitsstest 1)	87
Abbildung 3: Angaben der Probanden zu ihrer Befindlichkeit nach dem Testdurchlauf (Befindlichkeitsstest2)	88
Abbildung 4: Anzahl der Probanden, die angaben, <u>nicht gestört</u> worden zu sein, bezogen auf den globalen Verzeichnungsgrad des jeweiligen Items .	93
Abbildung 5: Anzahl der Probanden, die angaben, <u>nicht gestört</u> worden zu sein, sortiert nach dem globalen messtechnischen Grad der Störung	94
Abbildung 6: Probanden OHNE Irritationen bei Items vom Strukturtyp A.....	100
Abbildung 7: Probanden OHNE Irritationen bei Items vom Strukturtyp B.....	101
Abbildung 8: Probanden OHNE Irritationen bei Items vom Strukturtyp C.....	101
Abbildung 9: Messtechnische Analyse, in welchen Bildfeldern mit dem verstärkten Auftreten von vertikalen Disparitäten gerechnet werden muss	104
Abbildung 10: Legende zum schnellen Verständnis der Probandenaussagen	105
Abbildung 11: Messtechnische Analyse, in welchen Bildfeldern mit welchem maximalen, verzeichnungsbedingten Flächen-Vergrößerungsfaktor (L vs. R) zu rechnen ist.	115

Tabellenverzeichnis

Tabelle 1: Erste Kategorisierung der Probandenaussagen	90
Tabelle 2: Finale Kategorisierung der Probandenaussagen	91
Tabelle 3: Feldbezogene Irritationsaussagen der Probanden (bereinigt um Aussagen, welche definitiv nicht auf die eingebrachten Verzeichnungen zurückzuführen sind z.B. „Flimmern“, etc.)	92
Tabelle 4: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie in szenischen S3D-Produktionen vorkommen können.	96
Tabelle 5: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie S3D-Liveproduktionen vorkommen können.	97
Tabelle 6: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie S3D-Consumerkameras vorkommen können.	97
Tabelle 7: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie S3D-„Actioncams“ vorkommen könnten.	98
Tabelle 8: Statistische Analyse der Probandenaussagen nach Strukturtyp und globalem Fehlergrad.....	102
Tabelle 9: Verdichtete Analyse der Probandenaussagen nach Strukturtyp und globalem Fehlergrad.....	103
Tabelle 10: Vergleich von Probandenaussagen zu Testitems mit und ohne Verschiebungen des Verzeichnungszentrums	119

1. Einleitung

Die nachfolgende Datenanalyse beruht auf den Probandenaussagen aus den in Teil I beschriebenen Testsessions. Die erhobenen Daten wurden aus verschiedenen Blickwinkeln methodisch ausgewertet und statistisch interpretiert. Dabei stand der Versuch der Beantwortung folgender Frage im Mittelpunkt: ab welchem Maß von Verzeichnungs-differenz in Stereoteilbildern kommt es zu signifikanten Störungen in der visuellen Wahrnehmung von S3D-Zuschauern?

Die Probandenzahl N betrug insgesamt 27. Die Probanden waren relativ jung (19-28J), hatten volle stereoptische Sehfähigkeiten und eine hohe Medienaffinität (diese allerdings mehrheitlich im 2D-Bereich). 20 Probanden waren männlich, 7 weiblich. Neun Probanden trugen eine Sehhilfe.

Um die Unvoreingenommenheit der Probanden sicherzustellen wurden diese zu keinem Zeitpunkt über die konkrete Intention des Tests informiert.

Offiziell wurde der Test als „Stereo-3D-Wahrnehmungstest“ angekündigt, bei dem es um die Wahrnehmung von stereoskopischen Bildern im Allgemeinen geht.

Die konkrete Fragestellung an die Probanden nach jeder Item-Präsentation lautete:

1. Hat etwas bei der Tiefenwahrnehmung gestört?
2. Wo im Bild hat es gestört? (Feld 1 -9)
3. Was hat gestört?

Wurde die erste Frage bereits verneint, brauchten die beiden Folgefragen nicht beantwortet werden. Wurde sie hingegen bejaht, sollten die nachfolgenden Fragen ebenfalls beantwortet werden. Die Fragen wurden nach einer gewissen Routine der Probanden nicht zwingend nach jedem Item wiederholt. Wurde von der Testleitung allerdings eine gewisse Unsicherheit beim Probanden erkannt, so wurde die Frage erneut gestellt, um sicherzustellen, dass dieser die Frage inhaltlich richtig verstanden hat.

Die Antworten der Probanden sind aufgrund der offenen Frage Nr.1 qualitative Antworten, die dementsprechend behandelt werden müssen. Viele der Antworten ent-

halten sehr ähnliche Muster und besitzen damit auch einen quantitativen Charakter, den es ebenfalls zu berücksichtigen gilt. Das bereits im vorangegangenen Kapitel vorgestellte Bildraster gab den Probanden bei Beantwortung einer Frage die Möglichkeit, vorgegebene Felder zu nutzen, um eine Störung besser lokalisieren zu können. Dennoch verwendeten einige ihre eigenen Bezeichnungen und verzichteten völlig auf die Teilung in neun Einzelfelder. 25,9% der Probanden haben über den gesamten Testzeitraum vollständig auf die Verwendung der Feldbezeichnungen verzichtet.

Ein Testdurchlauf dauerte im Schnitt 30 Minuten. Da die Bewertungen durch die Probanden manchmal sehr kurz und manchmal sehr ausführlich ausfielen, gab es in der Durchführungspraxis starke zeitliche Schwankungen. Alle 27 Probanden wurden zu Beginn und zum Ende der Untersuchung zu ihrem körperlichen Befinden befragt.

2. Auswertung Befindlichkeitstest

Wie bereits in Teil I erläutert, können Verzeichnungsdifferenzen L vs. R zu starken zusätzlichen Disparitäten im S3D-Bild führen, sowohl horizontal als auch vertikal. Des Weiteren wurden Aniseikonie-ähnliche Wahrnehmungseffekte bei den Probanden vermutet, wenn sie in Bildbereiche schauen, in denen stereoskopisch abgebildete Objekte verzeichnungsbedingt deutliche Größendifferenzen L vs. R aufweisen.

Über den gesamten Testverlauf stiegen negative Aussagen der Probanden zu ihrer Befindlichkeit nur leicht von 15% (vor dem Test) auf 21% (nach dem Test) an. Der Testdurchlauf löste – wie im Voraus vermutet – keine visuelle Überforderung oder körperliche Probleme bei den Probanden aus.

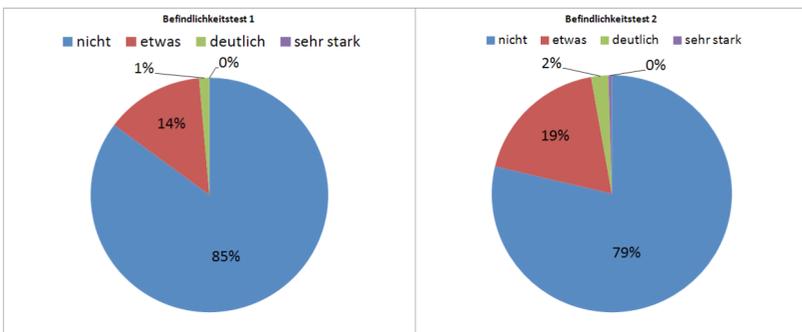


Abbildung 1: Befindlichkeitstest 1 und 2. 100% stehen für alle 16 Antworten der 27 Probanden (16x27 = 432 Antworten)

Die nachfolgenden Abbildungen zeigen alle Befindlichkeitsangaben der Probanden vor und nach dem Test im Detail. Auf den ersten Blick fällt beim vorangegangenen Test auf, dass die Probanden schon vor dem Test angaben, von einigen Merkmalen

betroffen zu sein, da sie sich vermutlich der Testsituation gegenüber sahen und ihr S3D-Mediennutzungsverhalten gering ausgeprägt war (s. vorangegangenes Kapitel). Zusammenfassend zeigt sich, dass die Probanden nach dem Test tendenziell häufiger angaben, in ihrer Befindlichkeit verändert zu sein. Nach dem Test waren alle Probanden von mindestens einer Befindlichkeitsstörung betroffen. Den auffälligsten Auschlag erkennt man bei der Angabe „Überanstrengte Augen“. Während in der vorgelegerten Befragung acht Probanden diese Angaben machten, sind es im zweiten Test schon 15 Probanden, sofern man alle Ausprägungsstufen berücksichtigt.

Auffällig ist, dass die hohen Ausprägungen wie „deutlich“ und „sehr stark“ von den Probanden in beiden Befragungsrunden sehr selten genannt wurden, bzw. es wurde die Ausprägung „nicht“ in beiden Tests mit 85% und 79% aller Antworten am häufigsten genannt.

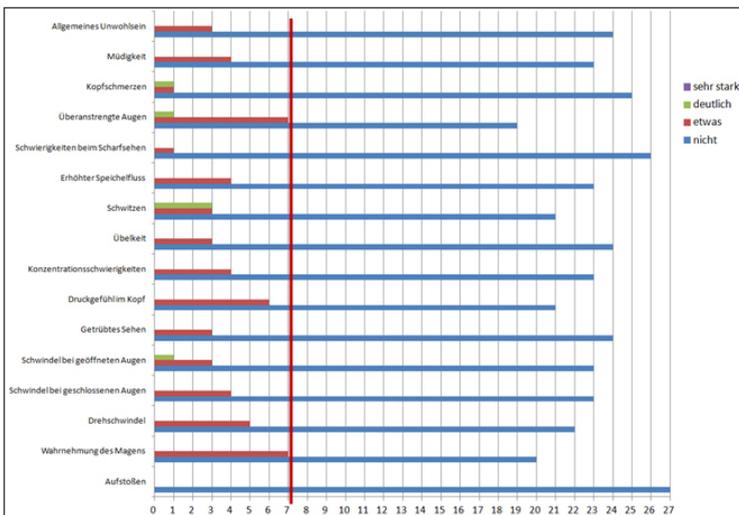


Abbildung 2: Angaben der Probanden zu ihrer Befindlichkeit vor dem Testdurchlauf (Befindlichkeitsstest 1)

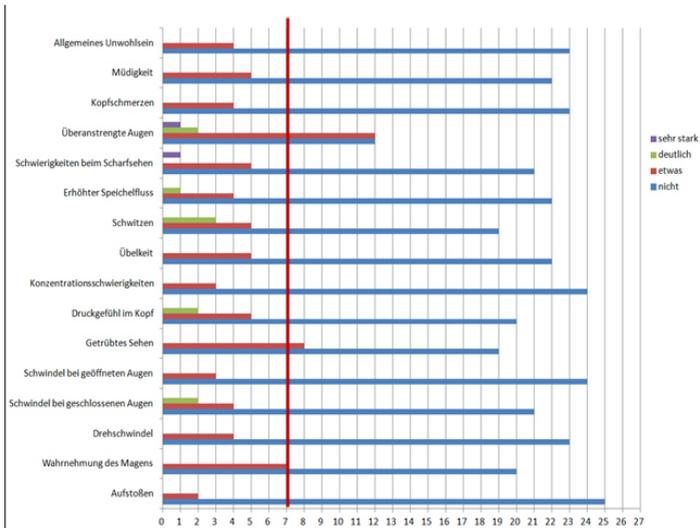


Abbildung 3: Angaben der Probanden zu ihrer Befindlichkeit nach dem Testdurchlauf (Befindlichkeitsstest2)

3. Kategorisierung der Probandenantworten

Nach jedem betrachteten Item hatten die Probanden beliebig viel Zeit, um ihren persönlichen Eindruck zu äußern. Zunächst lässt sich feststellen, dass die Antworten in den meisten Fällen recht knapp waren. Dies mag daran liegen, dass die gestellten Fragen sehr präzise Fragen waren.

Eine häufig verwendete Praxis in der qualitativen Forschung ist die Entwicklung eines Kategoriensystems.¹ Beim Durcharbeiten der Texte werden die Informationen entweder bereits bestehenden Kategorien zugeordnet oder es werden dazu passende, neue Kategorien entwickelt. Die für unsere Auswertung genutzte Methode orientiert sich an der qualitativen Inhaltsanalyse von Philipp Mayring².

Nach einer ersten Sichtung der Antworten ergaben sich die in Tabelle 1 dargestellten Kategorien, die untereinander häufig große Schnittmengen besaßen. Der inhaltliche Schwerpunkt für die Kategorienbildung bezog sich auf die Beschreibungen der Störfaktoren, die von den Probanden festgestellt werden konnten. Für eine bessere Nachvollziehbarkeit der Ergebnisse wurden jeder Kategorie die zur Entstehung der Störung genannten Begriffe zugeordnet.

¹ Vgl. Mayring, Philipp (2010), S. 59.

² ebd. (2010)

Kategorie	Beispiele der Wortwahl der Probanden (getrennt durch Semikolon)
Positionierung	Links, mittig, rechts (horizontal/vertikal); Bewertungsfelder (1-9)
Schärfe	Unschärf; scharf
Verschwommenheit	verschwommen
An einer Geraden verlaufende Störung	Striche; Linien; Streifen
Tiefe	vor/hinter dem Bildschirm; Tiefeneffekt; 3D;
Klarheit	Unklar; klar; sauber; deutlich
Flackern	Pulsieren; zittern; schimmern
Helligkeit	Dunkel; hell;
Objekt im Bild	Fahrzeug; Star Wars Flugschiff; Stuhl
Randbereich	Rand; Rahmen; Außenbereich
Bewegung	Bewegendes Objekt; ins Bild kommendes Objekt
Versatz	Versatz; versetzt
Nicht kategorisierbare Störung	<Probanden gaben Irritation an, jedoch nicht wovon>
Doppelbild	Doppelbild; zwei separate Bilder erkennbar, Überschneidungen
Artefakt	Artefakt
Verzerrung	Fisheye; Bogen; Tonnenform, Kissenform; Stauchung; Verzeichnung;
Anstrengende Bildinhalt	Übelkeit; Augenschmerzen; Betrachtung unangenehm
Eigene Bezeichnung	Kaputt; zerfetzt
Zusammenbruch der Tiefenwahrnehmung	3D-Effekt ist zusammengebrochen; Tiefe nicht wahrnehmbar; kein Tiefeneffekt
Spiegelungen	Spiegelungen am Boden
Irritationen	Objekt hat irritiert; Lichtquelle hat mich irritiert
Größenverhältnisse	Objekt x ist größer als Objekt y
Window Violation	Window Violation
Fokussierung	Fokussieren fällt schwer; kann nicht richtig fokussieren
Fusion	Kein Fusion möglich

Vignettierung	Es gibt eine sichtbare Vignettierung
Lichter	Lichter an der Decke
Tiefenunterschiede	Person auf der Treppe ist flach, der Rest hat Tiefe

Tabelle 1: Erste Kategorisierung der Probandenaussagen

Die im ersten Schritt erarbeiteten Kategorien sind sehr inhaltsnah, was dazu führt, dass es viele flüssige Übergänge zwischen den Kategorien gibt. Diese gilt es in einem zweiten Schritt zu größeren, übergeordneten Kategorien zusammenzufassen. Dadurch ergeben sich größere, jedoch umfassendere Kategorien. Durch diesen Prozess soll dem Phänomen entgegengewirkt werden, dass jedes Schlagwort eine eigene Kategorie bekommt. Trotzdem soll eine zu starke Verallgemeinerung ebenfalls vermieden werden.

Bereits bei der ersten Betrachtung der Probandenaussagen war es auffällig, dass die Probanden häufig ähnliche Phänomene mit verschiedenen Begriffen beschrieben. Unter Berücksichtigung dieser Bedingungen ergeben sich nun die folgenden finalen Kategorien der Störfaktoren:

Kategoriename	Beschreibung	Beispiele
Bildqualität	Generelle optische Eigenschaften von Bildmaterial	Schärfe, Verschwommenheit, Klarheit, Flackern, Zittern, Pulsieren, Farben, Kontraste,
Objekte im Bild	Im Bildinhalt auftretende Objekte der Szenen	Flugzeug, Stuhl, Lampe
Geometrische Störungen	Verzerrungen jeglicher Art	Wölbungen an Kanten, Versatz, Doppelbilder, Fisheye
Gesamtbildstörungen	Zusammenbruch der Wahrnehmung, gesamter Bildinhalt als störend empfunden	Tiefenwahrnehmung ist nicht möglich, der Tiefeneindruck ging verloren
Eigene Bezeichnungen	Besondere Wortwahl der Probanden, die keine besondere Störung bezeichnet	Kaputt, komisch, verwirrend
Tiefe	Auffälligkeiten bei Störungen der Tiefe	Position ist zu flach, ein Objekt schaut zu weit aus dem Bildschirm heraus

<p>Randproblem bei Tonnenverzeichnung</p>	<p>Mit dieser Kategorie sollen eindeutige Hinweise auf das produktionsbedingte Randproblem herausgefiltert werden, um diese möglichst wenig in die Auswertung miteinfließen zu lassen.</p>	<p>Streifenbildung am Rand. Zerbrochenes Glas am unteren Rand. Wie ein zerbrochener Spiegel wirkender Effekt.</p>
<p>Nicht identifizierte Störungen</p>	<p>Unter diese Kategorie fallen Störungen, die entweder nur als Störung bezeichnet werden oder keine spezifische Bezeichnung beinhalten.</p>	<p>Rechts oben hat mich etwas gestört</p>

Tabelle 2: Finale Kategorisierung der Probandenaussagen

Die Feldpositionierung der Irritationen wurde sowohl durch eigene Formulierungen der Probanden als auch durch die Nennung der Feldnummer dokumentiert – bereinigt um Aussagen, welche definitiv nicht auf die eingebrachten Verzeichnungen zurückzuführen sind z.B. „Flimmern“, etc..

4. Globale Analyse der Probandenaussagen

Einen ersten Überblick über die von den Probanden gemeldeten Irritationen - global bezogen auf die Items - zeigen die nachfolgenden Abbildungen.

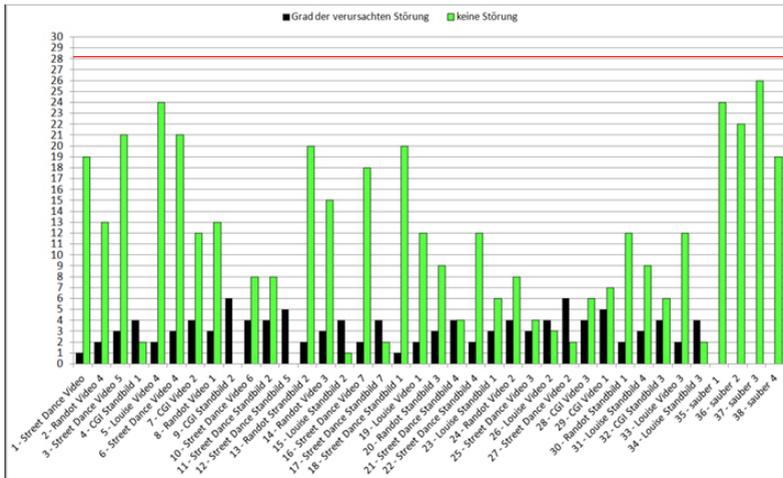


Abbildung 4: Anzahl der Probanden, die angaben, nicht gestört worden zu sein, bezogen auf den globalen Verzeichnungsgrad des jeweiligen Items

Betrachtet man die letzten vier Items, welche völlig unverzeichnet waren, fällt die korrekterweise hohe Anzahl an Probanden auf, die angaben, durch nichts gestört worden zu sein. Einzige Ausnahme bildet Item 38, welches nur 19 Probanden nicht irritiert hat, 8-mal hingegen als störend empfunden wurde. Dazu später mehr.

Vergleicht man die Bereiche der Items 1-17 (Kissenverzeichnungen) und 18-34 (Tonnenverzeichnungen), so fällt auf, dass die erste Gruppe deutlich störungsfreier bewertet wurde als die zweite Gruppe. Dies liegt mit Sicherheit daran, dass die zweite Gruppe die produktionsbedingt problematischeren Tonnenverzeichnungs-Items beinhaltet (siehe vorangegangener Artikel, Kap. „Produktion der Items“) und dies - trotz kategorischer Bereinigung - nicht vollständig herausgefiltert werden konnte.

Die verzeichneten Items 1, 3, 5, 6, 13 und 18 wurde deutlich seltener als irritierend bezeichnet. Diese Items besitzen einen maximalen Grad der globalen Störung bis 3. Dies könnte also evtl. ein möglicher Grenzwert für deutlich ausgelöste Probandenirritationen sein.

Betrachtet man die Angaben über Störungen, sortiert nach dem globalen messtechnischen Grad der Störung, ist die Irritationsfähigkeit der Items mit dem Grad der vorhandenen Verzeichnungsunterschieden deutlich korreliert. Es sind aber auch Ausnah-

men zu finden, welche darauf hinzuweisen scheinen, dass bestimmte Bildinhalte das Irritationspotenzial bei den Probanden verringern konnten.

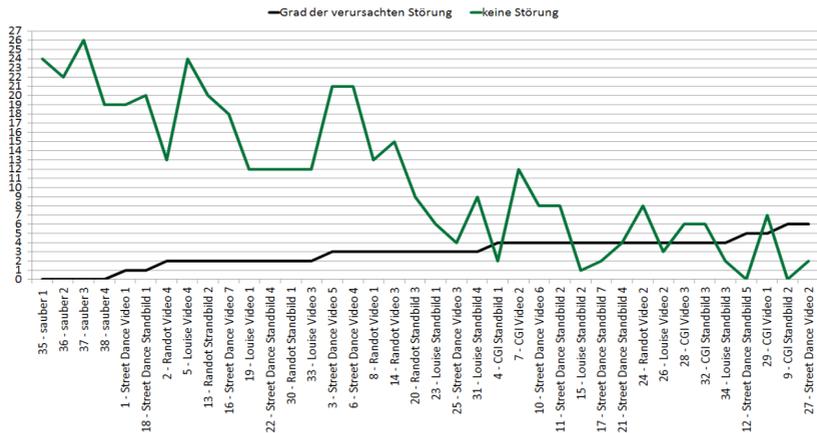


Abbildung 5: Anzahl der Probanden, die angaben, nicht gestört worden zu sein, sortiert nach dem globalen messtechnischen Grad der Störung

5. Min/Max-Analyse des Probandenpanels

Zwei Probanden (Nr. 20 und 24) weisen nach der Kategorisierung und Bereinigung die geringste Anzahl an wahrgenommenen Irritationen auf. Lediglich bei vier (von de-facto 34 gestörten) Items fühlten sie sich im Betrachten des Bildes irritiert. Dabei stimmen sie aber nur in einem Item überein (Item 9). Beide haben in zwei sehr stark gestörten S3D-Bildern keinerlei Irritationen gesehen. Proband 24 hat - zur Überraschung der Testleitung - in einem sauberen Item alle Felder bis auf die Mitte als störend bezeichnet.

Beide Probanden sind männlich, 20 und 24 Jahre alt, haben keine Sehhilfe, keine Farbensehchwäche und keine offensichtliche stereoptische Sehschwäche. Beim Randot-Test haben beide die fünfte Stufe erreicht und damit das Mindestmaß erreicht um an dem Test teilnehmen zu können. Diese Stufe wurde allerdings insgesamt von 13 der 27 Probanden erreicht, somit sind die Probanden 20 und 24 diesbezüglich nicht ungewöhnlich. Beide Probanden gaben eine recht hohe Vertrautheit mit Medien an (Stufe 5 von 6), sie sehen allerdings selten S3D Filme, sowohl im Kino als auch zuhause.

Als „sensitivster“ stellte sich der Proband Nummer 15 heraus. Dieser gab bei 29 Items eine Irritation an. Besonders der linke Bildrand wird von ihm (korrekterweise) oft genannt. Allerdings werden von ihm auch in den sauberen Items vereinzelt Irritationen gemeldet. Proband 15 ist männlich, 24 Jahre alt und hat keinerlei Schwächen im Seh-

vermögen. Er erreichte beim Randot-Test die höchste Stufe 10 und damit eine hervorragende Stereosehfähigkeit attestiert. Mit Medien ist er sehr vertraut, nutzt monatlich einen 3D-TV, etwas seltener das 3D-Kino, aber mehrmals täglich den 2D-TV. Mit S3D-Spielen hat er nur geringe Erfahrung. 3DTV-Vertrautheit und hohe Auflösungsfähigkeit in der patenten Stereopsis scheinen also mit einer hohen Empfindlichkeit für Verzeichnungsdifferenzen in S3D-Bildern einherzugehen. Zwei weitere Probanden (14, 27) mit dergleichen hohen Stufe beim Randot-Sehtest wiesen unbereinigt ein ähnlich hohes Störempfinden auf³.

6. Globalanalyse entlang stellvertretender Praxiskategorien

Die im Test verwendeten Verzeichnungsdifferenzen stellen aus Praxissicht zunächst eher worst-case Betrachtungen dar, um überhaupt ein grobe Orientierung über deren Einfluss in S3D-Bildern zu erhalten.

Im Folgenden werden zu jedem Item die Probanden angezeigt, die eine Störung erkannt haben. Nicht irritierte Probanden sind farblich entsättigt. Das gleiche gilt für die farbigen Felder: je höher die Deckkraft, desto mehr Probanden verorteten in diesem Feld eine Irritation (Mehrfachnennungen möglich).

³ nach Kategorisierung und Bereinigung allerdings nur noch Proband 14.

Kategorie	Item	Inhaltstyp	Strukturtyp	Verzeichnungstyp (L+R)	globaler Störungs-faktor L vs. R	Probandenverteilung	Anzahl irritierter Probanden
typ. Verzeichnungen in szenischen S3D-Produktionen ("Primar")	1	Street Dance Video 1	B	unverzeichnet + K0	1		33,3%
	12	Street Dance Standbild 1	B	unverzeichnet + T0	1		18,5%
	3	Lounge Video 4	A	K0+K1	1		3,7%
	22	Street Dance Standbild 6	A	T0+T1	1		37,0%
	13	RandortStandbild 1	C	K0+K1	1		18,5%
	30	RandortStandbild 1	C	T0+T1	1		44,4%

Tabelle 4: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie in szenischen S3D-Produktionen vorkommen können⁴.

Zwischenfazit „typisch szenisch“:

3,7% der Probanden äußerten bei einem Video Irritationen. Bei Standbildern äußerten 18% bis 44% der Probanden Irritationen (ohne diese zu qualifizieren = stark/schwach).

4 Anmerkung: Item1 wurde auskommentiert, da im mittleren Bildfeld 5 messtechnisch nachweislich keine Verzeichnung auftrat. (Beispielhafte Probandenaussagen zu diesem Feld: „Zwischendurch unscharf“, „Wagen hatte eine komische Tiefe“.)

Globalanalyse entlang stellvertretender Praxiskategorien

Kategorie	Item	Inhaltstyp	Strukturtyp	Verzerrungstyp (L H R)	gebäuer- Störungs- faktor L H R	Probandenverhalten	Anzahl Irritierter Probanden
top. Verzerrungen in S3D-Liveproduktionen (Zooem)	25	StreetDance Video 3	B	T0+T1+xy-Verzerrung	3		81,5%
	7	Bandet Video 1	C	K0+K1+xy-Verzerrung	3		11,1%
	10	StreetDance Video 6	B	K0+K1+xy-Verzerrung	4		63,0%
	22	OD Video 3	C	T0+T1+xy-Verzerrung	4		40,7%

Tabelle 5: Probandenaussagen zu Testitems mit typ. Verzerrungen wie sie S3D-Liveproduktionen vorkommen können.

Zwischenfazit „typisch live“:

Alle Items waren Videos. Je nach Bildinhalt/Strukturtyp gaben 11% - 81% der Probanden Irritationen an den Bildrändern an.

Kategorie	Item	Inhaltstyp	Strukturtyp	Verzerrungstyp (L H R)	gebäuer- Störungs- faktor L H R	Probandenverhalten	Anzahl Irritierter Probanden
top. Verzerrungen in Produkt-Cameras bei Kameradreharbeiten gezeigt Hintergrund	23	Louise Standbild 1	A	T1+T2	3		51,9%
	21	Louise Standbild 4	A	T1+T2	3		63,0%
	24	Bandet Video 3	B	K1+K2	3		22,2%
	6	StreetDance Video 4	C	K1+K2	3		11,1%
	26	Louise Video 1	A	T1+T2+xy-Verzerrung	4		63,0%
	22	StreetDance Standbild 1	C	K1+K2+xy-Verzerrung	4		51,9%

Tabelle 6: Probandenaussagen zu Testitems mit typ. Verzerrungen wie sie S3D-Consumerkameras vorkommen können.

Zwischenfazit „typisch Consumer“:

Bei Kissenverzeichnungen gaben, abhängig vom Bildinhalt/Strukturtyp, 11% - 51% der Probanden Irritationen vornehmlich an den Bildrändern an. Bei Tonnenverzeichnungen erhöhte sich die Quote, je nach Bildinhalt/Strukturtyp auf 51% - 61% der Probanden. Aufgrund der Produktionsproblematik der tonnenverzeichneten Items, sollten letztere Prozentangaben aber evtl. geringer gewichtet werden.

Kategorie	Item	Inhaltstyp	Strukturtyp	Verzeichnungstyp (S3D-M)	globaler Störungs- faktor s. Tab. 6	Probandenverhalten	Anzahl irritierter Probanden
typ. Verzeichnungen in sog. „Actioncams“	7	Video 2	A	I2+I3	4		51,9%
	15	Levante Standbild 2	A	I2+I3	4		55,6%
	24	Paradisi Video 2	C	I2+I3	4		59,3%
	32	Standbild 5	C	I2+I3	4		66,7%
	27	Video 1	A	I2+I3+V+Verfälschung	5		55,6%
	12	StreetDance Standbild 5	C	I2+I3+V+Verfälschung	5		70,4%

Tabelle 7: Probandenaussagen zu Testitems mit typ. Verzeichnungen wie sie S3D-„Actioncams“ vorkommen könnten.

Zwischenfazit „typisch Actioncam“:

Egal welcher Verzeichnungstyp, Inhaltstyp oder Strukturtyp: Zwischen 51% bis 70% der Probanden meldeten Irritationen - vornehmlich in den Randbereichen der Bilder/Videos.

7. Globalanalyse nach Medientyp der Items

Zunächst wurden die Probandenaussagen zu allen Items vom Typ **Standbild** genauer betrachtet. Für diesen Medientyp lässt sich eine deutliche Kausalität feststellen: je geringer der technische Fehlergrad, desto weniger Probanden empfanden eine Störung. Dies bezieht alle Typen von Standbildern mit ein (real, CGI, Randot). Item 12, welches zu den am stärksten verzeichneten Items gehört (Fehlergrad 5), wird jedoch weniger häufig als andere Standbilder als störend empfunden, trotz des hohen Fehlergrades. Es ist allerdings vom Strukturtyp C und enthält keine geraden Kanten oder geometrische Objekte.

Bei **Video**-Items ist der Trend ähnlich. Mit steigendem Fehlergrad, sinkt die Angabe einer fehlerfreien Wahrnehmung. Item 25 sticht dabei deutlich aus der Gruppe der realen Videos heraus. Zweiundzwanzig mal wurde dieses Item als störend empfunden. Es ist zwar dem Strukturtyp B zugeordnet, das Video weist aber in einigen Einstellungen deutliche Elemente von Strukturtyp A auf (Hochhäuser, extreme Detaillierung), die die Irritationen ausgelöst haben könnten. Zwei Videos hatten die höchsten Verzeichnungsgrade aller Items, einmal Realfilm und einmal eine CGI-Animation. Hierbei ist das CGI-Video 10 mal weniger oft als störend empfunden worden als der Realfilm. Beide haben eine Tonnenverzeichnung, es gibt lediglich einen Unterschied im Verzeichnungszentrum: das CGI Video weist eine Verschiebung der x-Achse, der Realfilm auf beiden Achsen auf. Daher resultiert auch der Unterschied im globalen Fehlergrad von 5 (CGI) und 6 (Realfilm).

Die unterschiedliche Anzahl an Irritationsmeldungen bei Standbildern und Videos lässt sich evtl. mit der Zeit erklären, die die Probanden bei der Betrachtung eines Standbildes im Vergleich zu einem Video haben. Hier können sie sich auf ein unveränderliches Bild konzentrieren und somit deutlich genauer auf vorhandene Störungen achten.

Interessant wäre für spätere Tests evtl. auch ein Vergleich CGI-Animation vs. Realvideo mit gleichem Fehlergrad und derselben Verzeichnung. Daran ließe sich festmachen, ob bestimmte Medientypen sensitivere Probandenaussagen aufweisen.

8. Globalanalyse nach Strukturtyp der Items

Wie sich die einzelnen Strukturtypen in Zusammenhang mit dem technischen Fehlergrad auf die Irritations-Wahrnehmung der Probanden auswirken, wird im Folgenden diskutiert. Dazu wurden alle Items eines Strukturtyps und ihre jeweilige Störungsanfälligkeit in einem Streudiagramm festgehalten.

Bei Strukturtyp A, den hoch strukturierten Items, gab es nur wenige Items, welche störungsfreie Aussagen bekamen und diese auch nur beim technischen Fehlergrad 0 oder 2. Ab Fehlergrad 3 wurden die hoch strukturierten Items von mindestens der Hälfte aller Probanden als störend empfunden.

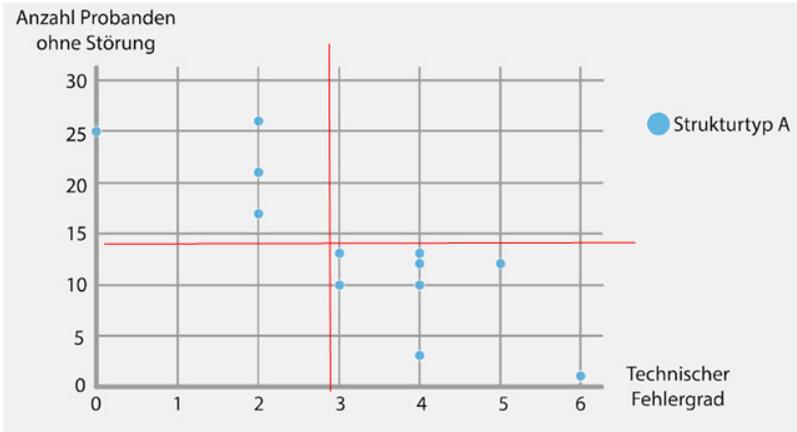


Abbildung 6: Probanden OHNE Irritationen bei Items vom Strukturtyp A

Die Grafik für Strukturtyp B dagegen hat eher einen Schwerpunkt auf der linken Seite. Die Items mit geringerem technischen Fehlergrad erhalten wenige störende Nennungen durch den Probanden. Die Verteilung zeigt im Gesamten gesehen, wie bei Strukturtyp A, eine Verringerung der nicht gestörten Probanden mit zunehmendem technischen Fehlergrad.

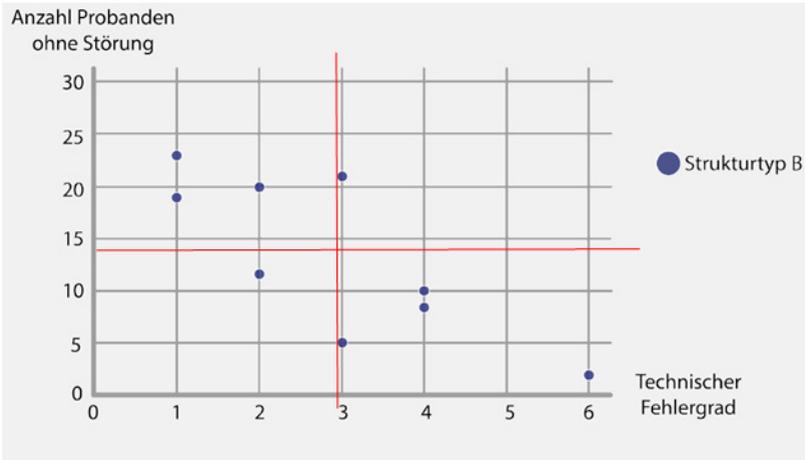


Abbildung 7: Probanden OHNE Irritationen bei Items vom Strukturtyp B

Die letzte Gruppe bilden die Items mit Strukturtyp C, den wenig strukturierten Items. Typ C war mit insgesamt 17 Items der häufigste Strukturtyp im Test. Die Zahl der ungestörten Probanden ist, bis zum Fehlergrade 3 sehr hoch angesiedelt. Ab Fehlergrad 4 sinkt diese Zahl jedoch wieder auf unter die Hälfte der Probanden.

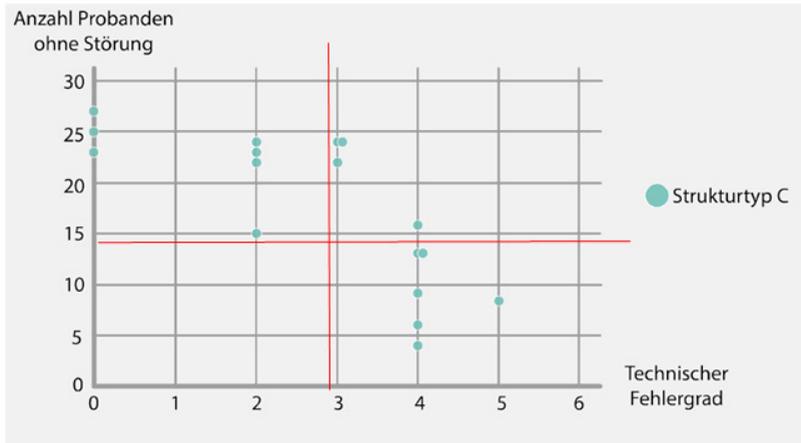


Abbildung 8: Probanden OHNE Irritationen bei Items vom Strukturtyp C

Die nachfolgende Tabelle dokumentiert die vollständige statistische Auswertung:

Strukturtyp / globaler Fehlergrad	Median	Arithme- tisches Mittel	Spann- weite	Standard- abwei- chung	Variations- koeffizient
A0	-	-	-	-	-
A1	-	-	-	-	-
A2	21	21,3333	9	3,6818	17,26%
A3	-	11,5	3	1,5811	13,75%
A4	11	9,5	10	3,9051	41,11%
A5	-	-	-	-	-
A6	-	-	-	-	-
B0	-	-	-	-	-
B1	-	21	4	2	9,52%
B2	-	-	-	-	-
B3	12	12,6666	16	6,549	51,70%
B4	-	9	2	1	11,11%
B5	-	-	-	-	-
B6	-	-	-	-	-
C0	-	25	4	1,633	6,53%
C1	-	-	-	-	-
C2	22,5	21	9	3,5355	16,84%
C3	24	23,3333	2	0,9428	4,04%
C4	11	11,8333	12	5,2615	44,46%
C5	-	-	-	-	-
C6	-	-	-	-	-

Tabelle 8: Statistische Analyse der Probandenaussagen nach Strukturtyp und globalem Fehlergrad

Verdichtet man die Ergebnisse nochmals, ergibt sich folgende Lesart:

Strukturtyp / globaler Fehlergrad	Median Unge­störte	Arithm. Mittel Unge­störte	Ø Probanden mit eindeutigen Irritationen	Spannweite Unge­störte	Standardabwei­chung Unge­störte	Variationskoeffizient Unge­störte
A2	21	21,3333	21,0%	9	3,6818	17,26%
C2	22,5	21	22,2%	9	3,5355	16,84%
A3	-	11,5	57,4%	3	1,5811	13,75%
B3	12	12,6666	53,1%	16	6,549	51,70%
C3	24	23,3333	13,6%	2	0,9428	4,04%
A4	11	9,5	64,8%	10	3,9051	41,11%
B4	-	9	66,7%	2	1	11,11%
C4	11	11,8333	56,2%	12	5,2615	44,46%

Tabelle 9: Verdichtete Analyse der Probandenaussagen nach Strukturtyp und globalem Fehlergrad

Zwischenfazit:

Durchschnittlich gaben - ab einem globalen Fehlergrad >Stufe 2 - über 50% der Probanden signifikante Irritationen bei der Betrachtung von S3D-Items des Strukturtyps A (hoch strukturiert) an. Bis zum Fehlergrad 3 *konnten* Items vom Strukturtyp C vereinzelt die (messtechnisch eindeutig vorhandenen) Störungen in der Wahrnehmung maskieren. Items gleich welchen Strukturtyps lösten ab dem globalen Fehlergrad Stufe 4 jederzeit bei über 50% der Probanden Irritationen aus.

Schon beim globalen Fehlergrad Stufe 2 gaben - unabhängig vom Strukturtyp - durchschnittlich 20% der Probanden an, verzeichnungsbedingte Irritationen in den S3D-Items zu bemerken, insbesondere bei Standbildern bzw. langen Einstellungen (bis max. 10 sek.).

Durchschnittlich 35% der Probanden schienen andererseits sehr wahrnehmungsresistent gegenüber jeglichen, eingebrachten Verzeichnungs­differenzen in den präsentierten S3D-Items zu sein.

Für zukünftige Tests sollte evtl. überlegt werden mit feldbezogenen Strukturklassen zu arbeiten, um den Einfluss bzw. die Maskierungsfähigkeit bestimmter Bildstruk-

turen auf die Wahrnehmung von Verzeichnungs-differenzen in S3D-Bildern präziser zu dokumentieren. Andererseits könnte auch argumentiert werden, dass Items vom Strukturtyp A generell die beste Indikation für das untersuchte Problem aufweisen, insbesondere wenn es um präzisere Analysen der globalen Fehlergrade 1 bis 2 geht.

9. Detailanalyse „vertikale Disparitäten“

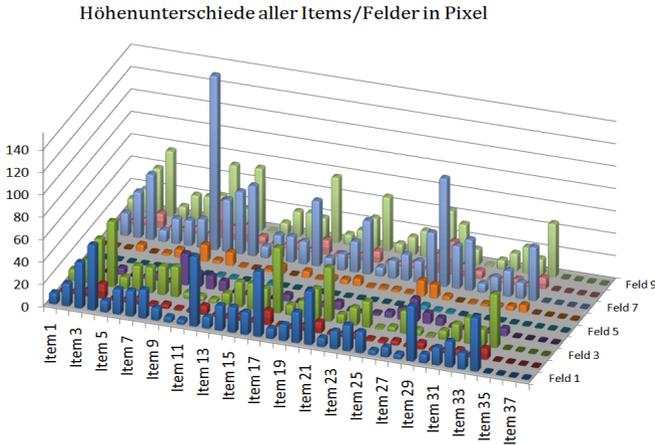


Abbildung 9: Messtechnische Analyse, in welchen Bildfeldern mit dem verstärkten Auftreten von vertikalen Disparitäten gerechnet werden muss

Die nachfolgenden grafischen Auswertungen beziehen die Probanden und ihre Fähigkeiten nach folgender Legende mit ein:

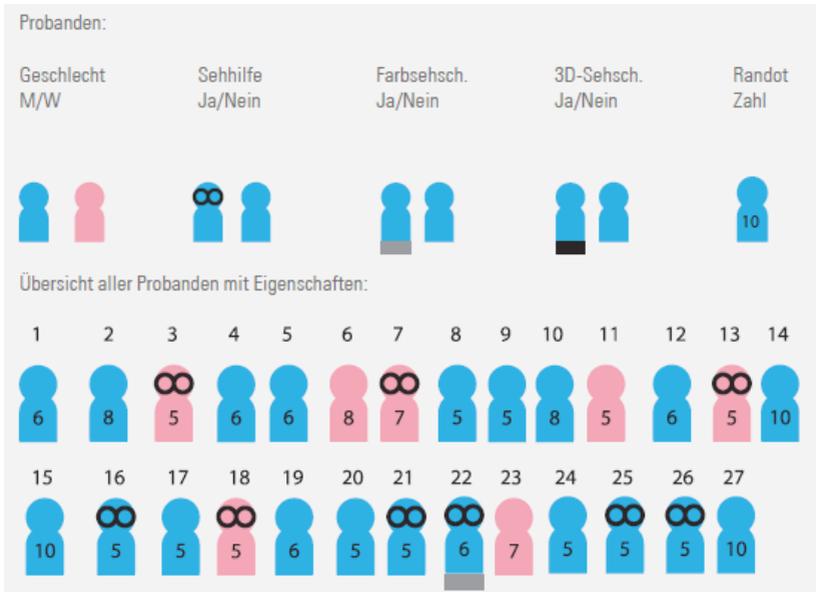
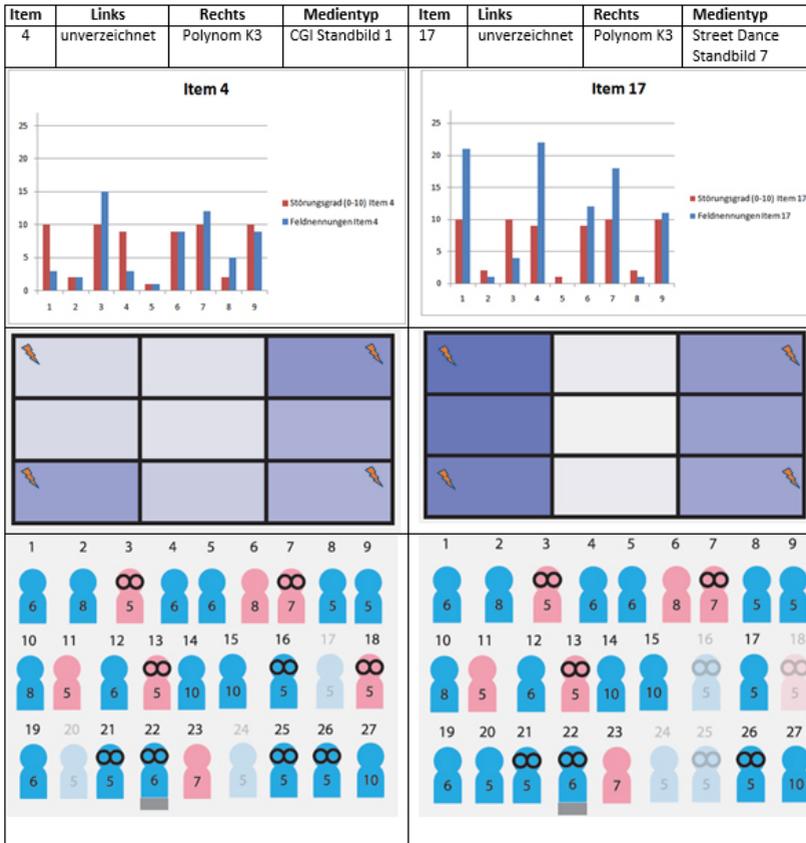


Abbildung 10: Legende zum schnellen Verständnis der Probandenaussagen

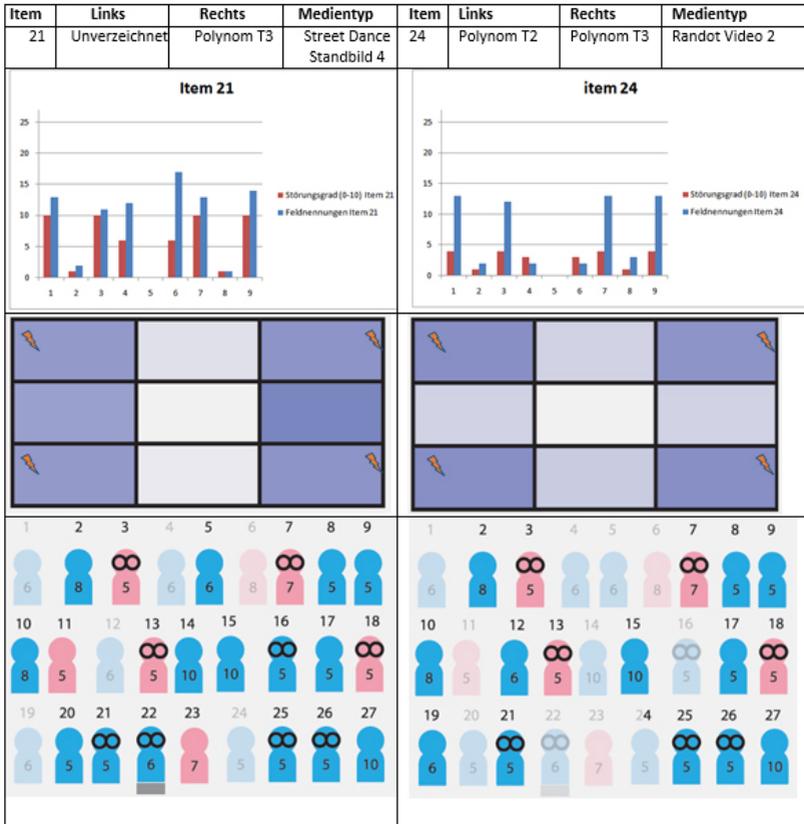
Probanden, welche keine Störungen erkannten werden farblich entsättigt dargestellt.

Lokale vertikale Disparitäten > 40px waren messtechnisch am **häufigsten** in den Items 4, 17, 21 und 24 jeweils in den Randfeldern 1,3, 7 und 9 zu finden. Daher wurden alle Probandenaussagen zu diesen Items nochmals einer detaillierteren Analyse unterworfen:

Alle vier Items besitzen den Fehlergrad 4 und sind (bis auf Item4 = A) vom Strukturtyp C – also eher gering strukturiert. Es handelt sich aber mehrheitlich um Standbilder, welche statistisch eine erhöhte Sensitivität bei den Probanden erzeugen.



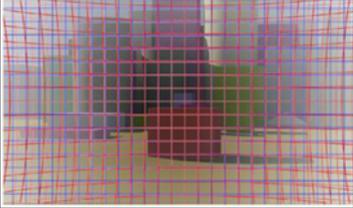
Die Probandenaussagen zu Item 24 (Randot Video3) bestätigen die Korrelation zwischen Verzeichnungsgrad, resultierenden vertikalen Disparitäten und Irritationspotenzial bei den Probanden eindeutig. Item 17 bestätigt die Korrelation ebenfalls, wobei eine Mehrheit der Probanden Bildfehler eher im (zum Feld1 benachbarten) Feld4 vertorten. Dies lässt sich auf den gezeigten Bildinhalt zurückführen. Der Arm der Frau stellt eine relativ gerade Linie im linken Bildrand dar während am rechten Bildrand nur der Hintergrund zu sehen ist.



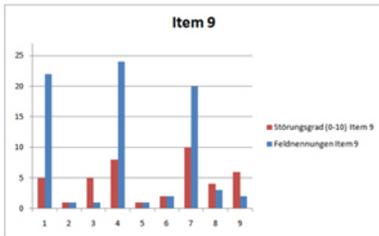
Ähnliche Spill-off-Effekte sind auch in den Aussagen zu Item 4 zu beobachten. Die Probanden verorteten hier starke Irritationen zusätzlich auch im direkt benachbarten Feld mit dem Sondereffekt, dass die starken vertikalen Disparitäten im Feld1 kaum wahrgenommen wurden. Kein Wunder: das verwendete CGI-Item bot hier mit einer einheitlichen, texturlosen blaugrauen Fläche den Probanden keinerlei Anhaltspunkte für eine stereoptische Erfahrung.

Messtechnisch am **stärksten** vertreten waren lokale, vertikale Disparitäten > 60px insbesondere in Item 9/Feld7, Item 27/Feld7 sowie in Item 12/Feld1+Feld7. Bezeichnenderweise alles Items, in denen zusätzlich zur starken radialen Verzeichnung auch noch das Verzeichnungszentrum aus der Bildmitte verschoben ist. Item 9 und 27 gelten aus Praxissicht als eher unrealistische worst-cases, geben aber gute Hinweise bez. der Sensitivität der Probanden bez. vertikaler Disparitäten.

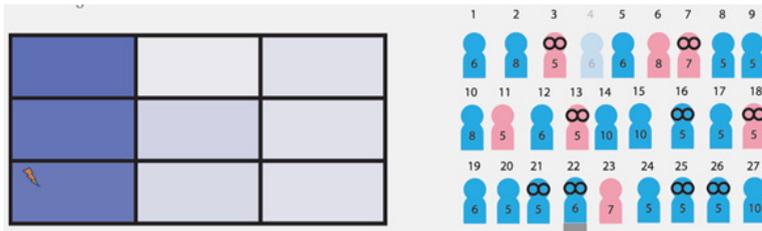
Item	Medien-Typ	Struktur-Typ	Links	Rechts
9	CGI Standbild 2	A	Polynom K2	Polynom K3 $x_0 = +0.25$, $y_0 = +0.25$



Analyse Probandenaussagen:



 = MESSTECHNISCH stärkste vertikale Disparitäten

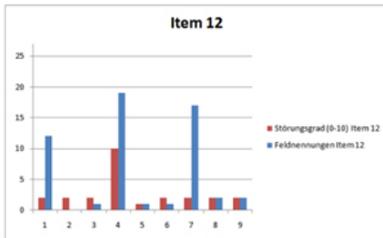


Bei Item 9 wurden die linken Randfelder von etwa zwei Dritteln der Probanden als störend notiert. Während die vertikale Disparität in den Feldern 1 und 4 nur zwischen 10 und 30 Pixeln liegt, beträgt sie in Feld 7 extreme 155 Pixeln. Kurioserweise wurde Feld 7 nur 20-mal „irritierend“ genannt. Feld 1 hingegen 22-mal und Feld 4 sogar 24-mal - sie wurden also als störender empfunden. Dies kann ebenfalls auf den dargestellten Bildinhalt zurückgeführt werden. Item 9 gehört zum Strukturtyp A und besitzt viele geometrische Formen und Linien. Diese sind in den oberen zwei Dritteln des Bildes angesiedelt, daher wurde die untere linke Ecke und somit Feld 7 seltener genannt, da hier hauptsächlich der bräunliche Boden der Szene zu sehen ist. Die große vertikale Disparität wirkt sich somit nur auf eine farbige Fläche aus und nicht auf eine geometrische Form.

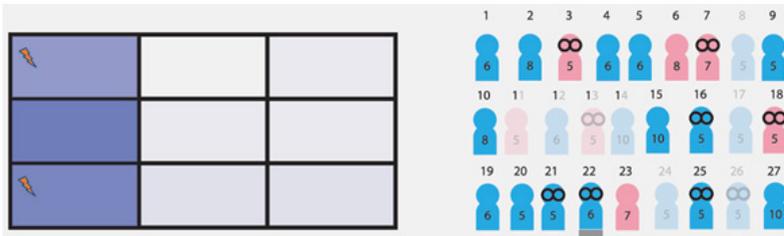
Item	Medien-Typ	Struktur-Typ	Links	Rechts
12	Street Dance Standbild 5	C	Polynom K2	Polynom K3 $x_0 = +0.25$



Analyse Probandenaussagen:



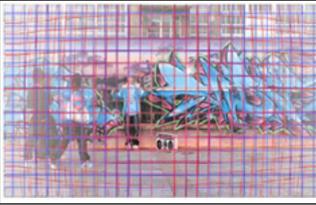
 = MESSTECHNISCH stärkste vertikale Disparitäten



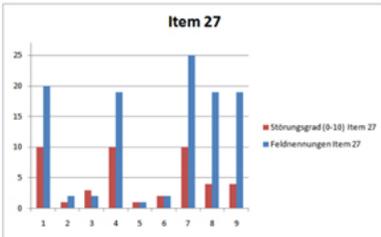
Bei Item 12 dominieren Irritationsmeldungen zu den linken Randfeldern 1, 4 und 7 aufgrund der Verzeichnungszentren-Verschiebung in x-Richtung. Die messtechnisch stärksten vertikalen Disparitäten liegen in den Feldern 1 und 7 mit jeweils 62 Pixeln. Allerdings wurde das mittlere Randfeld 4 (mit nur 8,5 Pixeln vertikaler Disparität) häufiger genannt als die andern beiden⁵. Das Item gehört zum Strukturtyp C, inhaltlich ist eine Tanzgruppe auf einer Bühne zu sehen. Feld 1 ist einheitlich gräulich gefärbt, daher ist für die Probanden die Verzeichnung nicht so stark ersichtlich wie in den beiden anderen Randfeldern.

Bei Item 9 und 12 handelt es sich um zwei Standbilder mit Verzeichnungsdifferenzen. Item 12 wurde von den Probanden jedoch deutlich weniger als störend empfunden. Die beiden Items unterscheiden sich zwar leicht im globalen Fehlergrad (6 zu 5), aber deutlich im Strukturtyp. Item 9 ist vom Strukturtyp A. Bei Item 12 handelt es sich um Strukturtyp C.

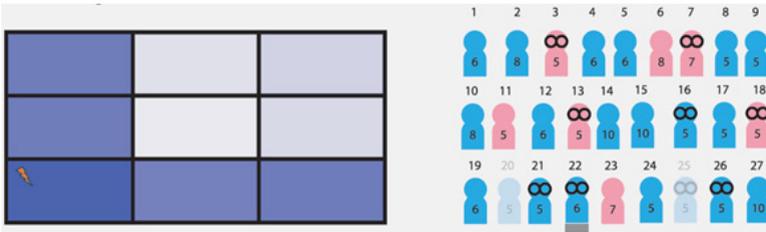
⁵ Wie sich später zeigt, sind in diesem Feld auch hohe Aniseikonie-ähnliche Effekte aufgetreten.

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
27	Street Dance Video 2	B	Polynom T2	Polynom T3 $x_0 = +0.25$, $y_0 = +0.25$	

Analyse Probandenaussagen:



 = MESSTECHNISCH stärkste vertikale Disparitäten



Da dieses Item eine Verschiebung des Verzeichnungs-zentrums in x- und y-Richtung aufwies, wurden hier häufig der linke und der untere Rand genannt. Die lokale vertikale Disparität Δv ist messtechnisch in Feld 7 mit 96 Pixeln tatsächlich auch am höchsten. Das Item gehört zum Strukturtyp B, zu sehen ist eine Gruppe Menschen, die vor einer bemalten Mauer zusammen stehen. Das Rechteck-Muster am Boden wird durch die Verzerrung verzerrt, was vielen Probanden aufgefallen ist. Feld 7, mit der stärksten vertikalen Disparität, wird mit 25 Anmerkungen am meisten genannt. Durch die durchgezogene Mauer über die gesamte Bildmitte ist die vertikale Disparität für die Probanden relativ schnell ersichtlich, daher werden die Felder 1 ($\Delta v = 18,5\text{px}$) und 4 ($\Delta v = 20\text{px}$) ebenfalls häufig genannt.

Zwischenfazit:

Die von Verzerrungsdifferenzen ausgelösten vertikalen Disparitäten irritierten die Probanden ab dem globalen Fehlergrad 4 eindeutig und nachweisbar. Bildfelder

mit den stärksten oder häufigsten Fehlern wurden von der Mehrheit der Probanden eindeutig als irritierend identifiziert. Leichte Schwankungen ergaben sich durch den Strukturtyp der Items bzw. den Strukturtyp der Felder, in denen die vertikalen Disparitäten auftraten. Am eindeutigsten reagierten die Probanden beim Medientyp Standbild, vermutlich aufgrund der zur Verfügung stehenden Betrachtungsdauer (10s).

Auf der anderen Seite legt die Datenanalyse nahe, dass bei manchen Items - insbesondere vom Strukturtyp C - bis zu 30% der Probanden vollkommen immun gegen deutliche vertikale Disparitäten zu sein scheinen.

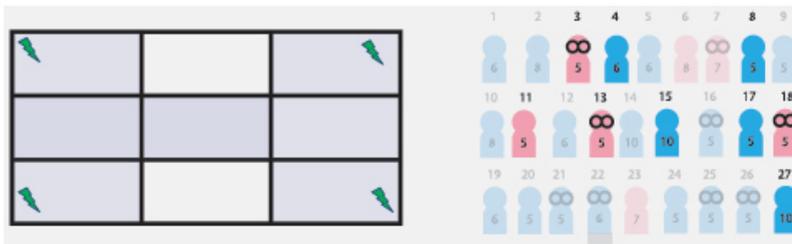
Erweiterte Grenzbetrachtung zu schwachen vertikalen Disparitäten:

Die Items 1,5, 13,18,22 und 30 wiesen nur geringe vertikale Disparitäten (ca. 1% der Bildhöhe) in den Randfeldern (1, 3, 7, 9) auf. Da die Items eine Bildgröße von 1920x1080 Pixel besitzen, bewegen sich die Werte der maximal auftretenden vertikalen Disparitäten - in besagten Randfeldern - um die 10 Pixel.

Bei den Items handelt es sich um zwei Videos und vier Standbilder. Es kommen die gleichen Verzeichnungsdifferenzen vor, jedoch einmal vom Kissen- und einmal mit Tonnen-Typ. Der globale Fehlergrad wechselt zwischen 1 und 2.

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
1	Street Dance Video 1	B	unverzerrt	Polynom KO	

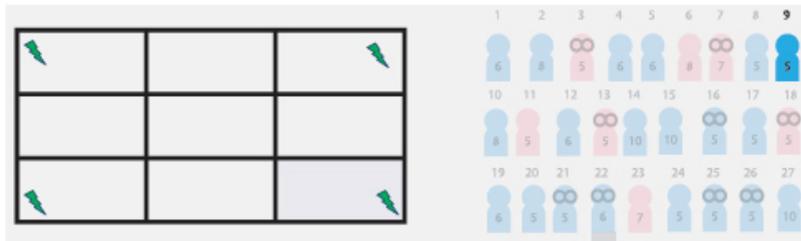
 = Minimale vertikale Disparitäten ≤ 10 px



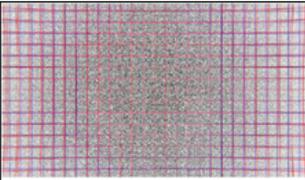
Probanden mit Irritationen: $9/27 = 33,3\%$

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
5	Louise Video 4	A	Polynom K0	Polynom K1	

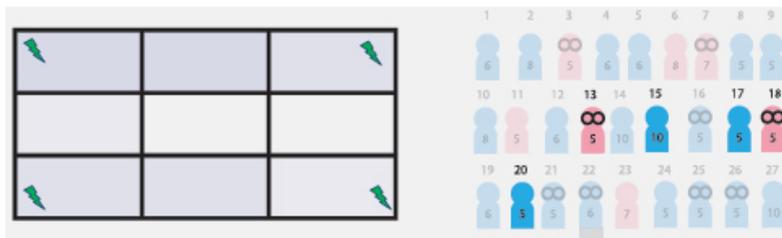
 = Minimale vertikale Disparitäten ≤ 10 px



Probanden mit Irritationen: $1/27 = 3,7\%$

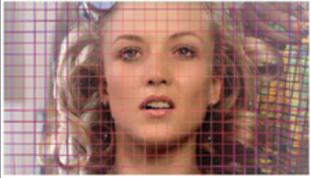
Item	Medien-Typ	Struktur-Typ	Links	Rechts	
13	Randot Standbild 2	C	Polynom K0	Polynom K1	

 = Minimale vertikale Disparitäten ≤ 10 px

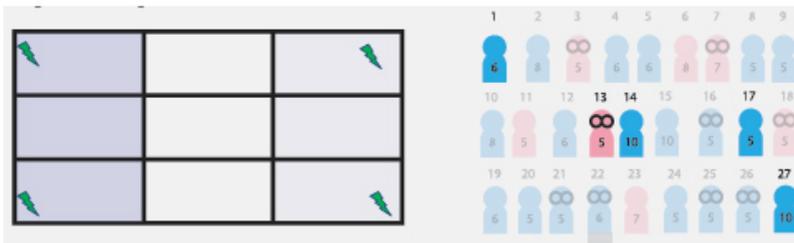


Probanden mit Irritationen: $5/27 = 18,5\%$

Item	Medien-Typ	Struktur-Typ	Links	Rechts
18	Street Dance Standbild 1	B	unverzeichnet	Polynom T0



 = Minimale vertikale Disparitäten ≤ 10 px

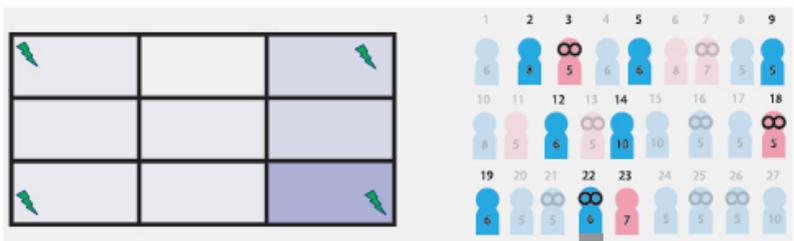


Probanden mit Irritationen: $5/27 = 18,5\%$

Item	Medien-Typ	Struktur-Typ	Links	Rechts
22	Street Dance Standbild 6	A	Polynom T0	Polynom T1

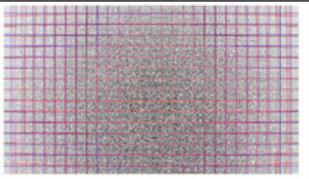


 = Minimale vertikale Disparitäten ≤ 10 px

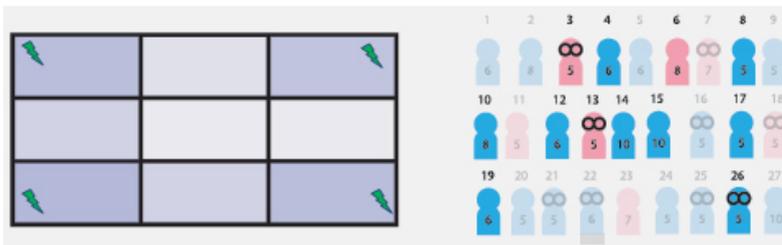


Probanden mit Irritationen: $10/27 = 37,0\%$

Item	Medien-Typ	Struktur-Typ	Links	Rechts
30	Randot Standbild 1	C	Polynom T0	Polynom T1



 = Minimale vertikale Disparitäten ≤ 10 px



Probanden mit Irritationen: $12/27 = 44,4\%$

Zwischenfazit:

Bei den analysierten *Standbild*items mit schwachen vertikalen Disparitäten ≤ 10 px (=1% der Bildschirmhöhe) ergaben sich feldexakte Irritationsmeldungen durch 18,5% – 44,4% der Probanden, ohne spezifische Geschlechts-, Sehhilfe- oder Randot-Sehtestergebnis-Korrelation.

Bei zwei *Video*items ist das Ergebnis nicht so eindeutig. Bei Item 5 - obwohl vom Strukturtyp A - fühlte sich kaum ein Proband irritiert. Im zweiten Fall (Item1 vom Strukturtyp B) waren 33,3% der Probanden irritiert, gepaart mit einer hohen Treffsicherheit bez. der verursachenden Felder. Hier sind weitergehende Untersuchungen in späteren Tests zu empfehlen.

Irritationsmeldungen zu schwach tonnenverzeichneten Items traten tendenziell öfters auf als zu schwach kissenverzeichneten. Für eine Interpretation ist aber evtl. die generelle Produktionsproblematik der tonnenverzeichneten Items in diesem Tests zu berücksichtigen (s. Teil1).

Für alle Aussagen in dieser Detailanalyse muss zusätzlich berücksichtigt werden, dass es aufgrund der verwendeten S3D-Bildschirmtechnologie womöglich zur einer generell reduzierten vertikalen Wahrnehmung kam, da durch die Darstellung der Items auf dem interlace-arbeitenden S3D-TV-Gerät (zeilenweise Polarisation) die vertikale örtliche Auflösung von originalen 1080px effektiv auf 540px halbiert wurde.

Die unverzeichneten „Placebo-Items“ 35, 37 und 38 erhielten korrekterweise keine *signifikanten* Irritationsmeldungen, was auf eine grundsätzlich solide Validität der Probandenaussagen schließen lässt.

10. Detailanalyse „Aniseikonie-ähnliche Effekte“

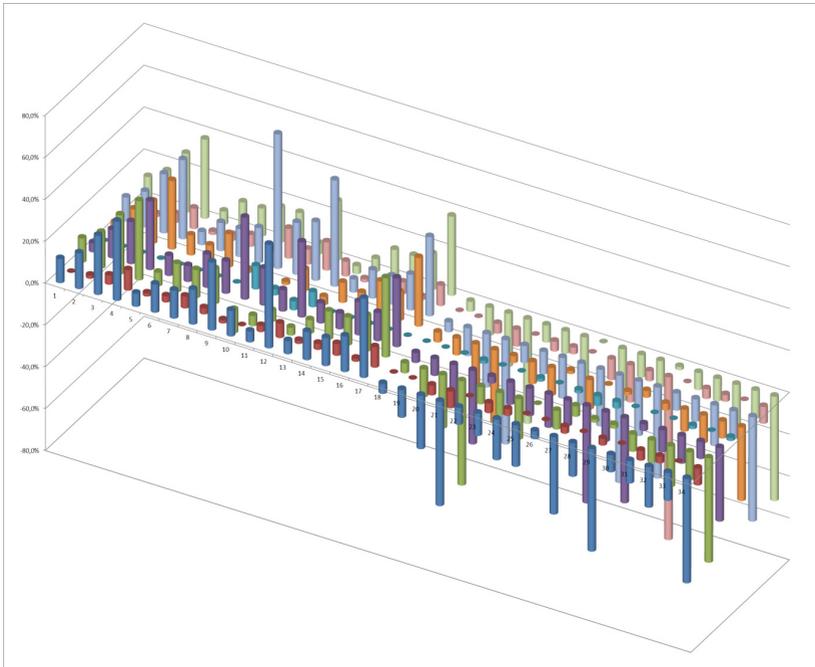


Abbildung 11: Messtechnische Analyse, in welchen Bildfeldern mit welchem maximalen, verzeichnungsbedingten Flächen-Vergrößerungsfaktor (L vs. R) zu rechnen ist.

Aniseikonie (Bildgrößenunterschied auf der Netzhaut)⁶ ist die Bezeichnung einer chronischen Krankheit, bei der optische Fehler im Auge oder funktionale Fehler auf der Netzhaut zu unterschiedlichen Abbildungsgrößen L/R führen. Bei starker Ausprägung führt dies für die Betroffenen zu extremen Problemen in der Lebensbewältigung. Der medizinische Begriff bezieht sich auf einen permanenten, globalen Abbildungsunterschied zwischen rechtem und linkem Auge. Wie unsere Beispielgitter in Teil1 deutlich zeigen, kann in Stereobildern mit Verzeichnungsunterschieden lokal ein ähnlicher Effekt lokal auftreten. Die Größenunterschiede L/R weisen hier allerdings

⁶ Vgl. Lang, Gerhard K., Augenheilkunde, 2014, Georg Thieme Verlag, S.326

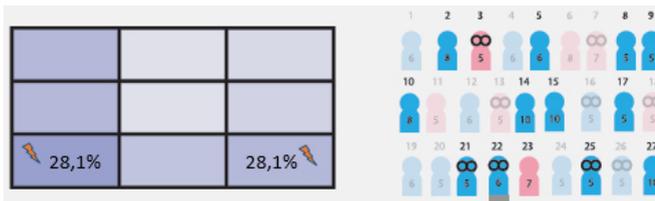
stetige, radiale Gradienten mit absoluten Maxima an den Bildrändern auf. Wir sprechen daher im Weiteren von Aniseikonie-ähnlichen Effekten.

Dieser Effekt tritt besonders stark in den Randbereichen der Testitems auf und ist zunächst vom Effekt der vertikalen Disparitäten schlecht zu isolieren. Es gab aber Items, welche Aniseikonie-ähnliche Effekte nur in einzelnen Feldern aufwiesen, z.T. deutlich entkoppelt von vertikalen Disparitäten. Dies waren in Item 11 die Felder 7+9, in Item 12 das Feld4 und in Item 26 das Feld7.

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
11	Street Dance Standbild 2	C	Polynom K1	Polynom K2 + $y_0 = +0,25$	

Analyse Probandenaussagen:

 = MESSTECHNISCH höchste Aniseikonie-ähnliche Werte L. vs. R. (bei geringer vert. Disparität)



Probanden, welche Irritationen meldeten : 14/27 = 52%
 davon Irritations-Nennungen zu Aniseikoniefeld 7: 12/14 = **86%**
 davon Irritations-Nennungen zu Aniseikoniefeld 9: 8/14 = 57%
 Probanden ohne Irritationen: 13/27 = 48%

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
12	Street Dance Standbild 5	C	Polynom K2	Polynom K3 + $x_0 = +0.25$	

Analyse Probendenaussagen:

 = MESSTECHNISCH höchster Aniseikonie-ähnlicher Wert L. vs. R. (bei geringer vert. Disparität)

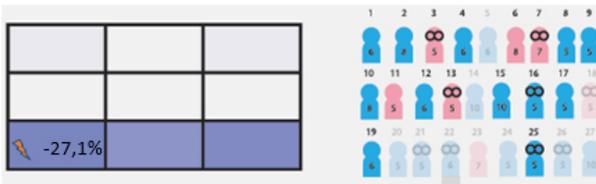


Probanden, welche Irritationen meldeten : $19/27 = 71\%$
 davon Irritations-Nennungen zu Aniseikoniefeld 4: $19/19 = 100\%$
 Probanden ohne Irritationen: $8/27 = 29\%$

Item	Medien-Typ	Struktur-Typ	Links	Rechts	
26	Louise Video 2	A	Polynom T1	Polynom T2 + $y_0 = +0.25$	

Analyse Probendenaussagen:

 = MESSTECHNISCH höchster Aniseikonie-ähnlicher Wert L. vs. R. (bei geringer vert. Disparität)



Probanden, welche Irritationen meldeten : $17/27 = 63\%$
 davon Irritations-Nennungen zu Aniseikoniefeld 7: $17/17 = 100\%$
 Probanden ohne Irritationen: $10/27 = 37,0\%$

Zwischenfazit:

Es waren in unserer Testanordnung drei Items zu finden, bei denen sich aniseikonie-ähnliche Bildeffekte *eindeutig* von vertikalen Disparitäten isolieren ließen. Der induzierte Vergrößerungsradius lag zwischen 27% und 36%.

Bei diesen Items ergaben sich feldexakte Irritationsmeldungen durch 52% – 71% der Probanden, ohne spezifische Geschlechts- oder Randot-Testergebnis-Korrelation. Der Anteil der Nennungen von Probanden mit Sehhilfe ist bei zwei Items (12 und 26)

leicht erhöht, aber aufgrund der geringen Probenzahl natürlich nicht grundsätzlich repräsentativ.

Andererseits scheinen je nach Bildinhalt 29% - 48% der Probanden vollkommen immun gegen offensichtliche aniseikonie-ähnliche Effekte in den drei Items zu sein, selbst wenn diese stark ausgeprägt sind.

Leider lassen sich für eine Gegenprobe kaum Items/Felder mit geringen aniseikonie-ähnlichen Vergrößerungswerten von solchen mit vertikalen Disparitäten unterscheiden. Für spätere Tests muss hier evtl. auf ein anderes Item-Konzept gesetzt werden. Bedenkenswert wären evtl. Items mit schwacher Verzeichnung aber globalen Brennweiten-Differenzen, so dass der Effekt eindeutiger zu lokalisieren ist.

11. Detailanalyse Einfluss von Verzeichnungszentrum-Verschiebungen

Für diese Analyse wurden Probandenaussagen zu Paaren von Video-Items verglichen, die ähnliche Verzeichnungspolynome aufwiesen, jedoch einmal mit und einmal ohne Verschiebung des Verzeichnungszentrums. Verschiebungen in x-Richtung führten messtechnisch zu erhöhten horizontalen und vertikalen Disparitäten am linken Bildrand des Items. Verschiebungen in x- und y-Richtung führten messtechnisch zu überproportional erhöhten vertikalen Disparitäten am linken Bildrand.

Item	Inhaltstyp	Strukturtyp	Verzeichnungstyp (L/R)	globaler Brennwei- tefaktor L vs. R.	Probandenurteilen	Anzahl/Inzidenz Probanden
14	Random Video 3	B	K1+K2	3		22,2%
8	Random Video 1	C	K0+K1+x-Verschiebung	3		11,1%
6	Street Dance Video 4	C	K1+K2	3		11,1%
23	Street Dance Video 3	B	T0+T1+x-Verschiebung	3		81,5%

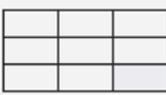
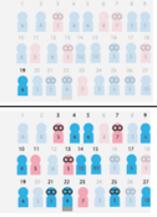
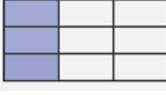
3	Louise Video 4	A	KO+KI	2			3,7%
10	Street Dance Video 6	B	KO+KI+x y-Verschiebung	4			63,0%

Tabelle 10: Vergleich von Probandenaussagen zu Testitems mit und ohne Verschiebungen des Verzeichnungszentrums

Zwischenfazit.

Unter Berücksichtigung der generellen Sensitivitätssteigerung der Probanden für bestimmte Strukturtypen, zeigt die Datenanalyse, dass Verschiebungen des Verzeichnungszentrums das Irritationspotenzial von S3D-Videos deutlich anheben, bei hoher Trefferquote der Probanden bez. der Lokalisation des Fehlers. Die in den Items verwendeten Verzeichnungszentrum-Verschiebungen wurden in diesem ersten Test relativ hoch gewählt. Trotzdem legen die Ergebnisse nahe, dass in der S3D-Produktion bei der Analyse und Korrektur von Objektivverzeichnungen die präzise Bestimmung des Verzeichnungszentrums äußerst wichtig ist.

12. Zusammenfassung

Die hier durchgeführte empirische Untersuchung zum Einfluss von Verzeichnungsdifferenzen in S3D-Bildern fand unter den technisch und organisatorisch bestmöglich herzustellenden Bedingungen statt, welche umfassend in Teil1 dokumentiert sind. Eine Testsession umfasste pro Proband 38 randomisiert präsentierte S3D-Testitems, von denen 4 keinerlei Verzeichnungsdifferenzen aufwiesen.

Die Probandenzahl N betrug insgesamt 27. Die Probanden waren relativ jung (19-28J), hatten volle stereoptische Schfähigkeiten und eine hohe Medienaffinität (diese allerdings mehrheitlich im 2D-Bereich). 20 Probanden waren männlich, 7 weiblich. Neun Probanden trugen eine Sehhilfe. Die Probanden wurden zu keinem Zeitpunkt über die konkrete Intention des Tests informiert, um ihre Unvoreingenommenheit sicherzustellen. Offiziell wurde der Test als „Stereo-3D-Wahrnehmungstest“ angekündigt, bei dem es um die Wahrnehmung von stereoskopischen Bildern im Allgemeinen geht.

Alle 27 Probanden wurden zu Beginn und zum Ende ihrer Sitzung zu ihrem körperlichen Befinden befragt. Über den gesamten Testverlauf stiegen negative Aussagen der Probanden zu ihrer Befindlichkeit nur leicht von 15% (vor dem Test) auf 21% (nach dem Test) an. Der Testdurchlauf löste - wie vorausgesagt - keine visuelle Überforderung oder körperliche Probleme bei den Probanden aus.

Das Probandenpanel wurde in seinem Aussageverhalten statistisch durchleuchtet. Dabei ergaben sich folgende Erkenntnisse: 3DTV-Vertrautheit und hohe Auflösungs-fähigkeit in der patenten Stereopsis (Stufe 10 im Randot-Sehtest) schienen mit einer hohen Empfindlichkeit für Verzeichnungsdifferenzen in S3D-Bildern einherzugehen. Auf der anderen Seite zeigten sich ca. 35% der Probanden sehr wahrnehmungsresistent gegenüber jeglichen, eingebrachten Verzeichnungsdifferenzen in den präsentierten S3D-Items. Die unverzeichneten „Placebo-Items“ erhielten korrekterweise sehr wenige Irritationsmeldungen, was auf eine grundsätzlich solide Validität der Probandenaussagen schließen lässt.

Die konkrete Fragestellung an die Probanden lautete nach jeder Item-Präsentation:

1. Hat etwas bei der Tiefenwahrnehmung gestört?
2. Wo im Bild hat es gestört? (Feld 1 -9)
3. Was hat gestört?

Durchschnittlich gaben - ab einem globalen Fehlergrad >Stufe 2 - über 50% der Probanden signifikante Irritationen bei der Betrachtung von S3D-Items des Strukturtyps A (hoch strukturiert) an. Bis zum Fehlergrad 3 *konnten* Items vom Strukturtyp C vereinzelt die (messtechnisch eindeutig vorhandenen) Störungen in der Wahrnehmung maskieren. Ab dem globalen Fehlergrad Stufe 4 lösten Items gleich welchen Strukturtyps jederzeit bei über 50% der Probanden Irritationen aus.

Schon beim globalen Fehlergrad Stufe 2 gaben - unabhängig vom Strukturtyp - durchschnittlich 20% der Probanden an, verzeichnungsbedingte Irritationen in den S3D-Items zu bemerken, insbesondere bei Standbildern bzw. langen Einstellungen (bis max. 10 sek.). Die unterschiedliche Anzahl an Irritationsmeldungen bei Standbildern und Videos lässt sich möglicherweise mit der Zeit erklären, die die Probanden bei der Betrachtung eines Standbildes im Vergleich zu einem Video haben.

Die im Test verwendeten Verzeichnungsdifferenzen stellten aus Praxissicht zunächst eher worst-case Betrachtungen dar, um überhaupt eine grobe Orientierung über deren Einfluss in S3D-Bildern zu erhalten. In der Globalanalyse ergaben sich dazu folgende Ergebnisse:

Verzeichnungsdifferenzen „typisch szenisch“:

Bei (einem) Video äußerten nur 3,7% der Probanden Irritationen. Bei Standbildern (bzw. langen Einstellungen) äußerten 18% bis 44% der Probanden Irritationen.

Verzeichnungsdifferenzen „typisch live“:

Je nach Bildinhalt/Strukturtyp vermeldeten 11% - 81% der Probanden Irritationen an den Bildrändern.

Verzeichnungs-differenzen „typisch Consumer“:

Bei Kissenverzeichnung gaben, abhängig von Bildinhalt/Strukturtyp, 11% - 51% der Probanden Irritationen vornehmlich an den Bildrändern an. Bei Tonnenverzeichnung erhöhte sich die Quote, je nach Bildinhalt/Strukturtyp auf 51% - 61% der Probanden. Aufgrund der Produktions-problematik der tonnenverzeichneten Items (s. Teil1), sollten letztere Prozentangaben aber evtl. geringer gewichtet werden.

Verzeichnungs-differenzen „typisch Actioncam“:

Egal welcher Verzeichnungstyp, Inhaltstyp oder Strukturtyp: Zwischen 51% bis 70% der Probanden meldeten Irritationen - vornehmlich in den Randbereichen der Bilder/Videos.

Die feldgenauen Detailanalysen ergaben folgende Erkenntnisse:

Die von Verzeichnungs-differenzen ausgelösten vertikalen Disparitäten irritierten die Probanden ab dem globalen Fehlergrad 4 eindeutig und nachweisbar. Bildfelder mit den stärksten oder häufigsten Fehlern wurden von der Mehrheit der Probanden treffsicher identifiziert. Leichte Schwankungen ergaben sich durch den Strukturtyp des Items bzw. den Strukturtyp der Felder, in denen die vertikalen Disparitäten auftraten. Am sensitivsten reagierten die Probanden beim Medientyp Standbild, vermutlich aufgrund der zur Verfügung stehenden Betrachtungsdauer (10s).

Bei den analysierten *Standbilditems* mit schwachen vertikalen Disparitäten $\leq 10\text{px}$ (=1% der Bildschirmhöhe) ergaben sich feldexakte Irritationsmeldungen durch 18,5% – 44,4% der Probanden, ohne spezifische Geschlechts-, Sehhilfe- oder Randot-Sehtestergebnis-Korrelation. Irritationsmeldungen zu schwach tonnenverzeichneten Items traten tendenziell öfters auf als zu schwach kissenverzeichneten. Für eine Interpretation ist aber wieder die generelle Produktionsproblematik der tonnenverzeichneten Items zu berücksichtigen.

Für alle Probandenaussagen muss generell berücksichtigt werden, dass es aufgrund der im Test verwendeten S3D-Bildschirmtechnologie zur einer generell reduzierten vertikalen Wahrnehmung kam, da durch die Darstellung der Items auf dem interlaced-arbeitenden S3D-TV-Gerät (zeilenweise Polarisation) die vertikale örtliche Auflösung von originalen 1080px effektiv auf 540px halbiert wurde.

Es waren in unserer Testanordnung drei Items zu finden, bei denen sich aniseikonie-ähnliche Bildeffekte *eindeutig* von vertikalen Disparitäten isolieren ließen. Bei diesen Items ergaben sich feldexakte Irritationsmeldungen durch 52% – 71% der Probanden, ohne spezifische Geschlechts- oder Randot-Testergebnis-Korrelation. Der Anteil der Nennungen von Probanden mit Sehhilfe war leicht erhöht, ist aber aufgrund der geringen Probenzahl nicht grundsätzlich repräsentativ. Auch hier schienen wieder 29% - 48% der Probanden vollkommen immun gegen diesen Effekt in den Items zu sein, selbst wenn er stark ausgeprägt war.

Unter Berücksichtigung der generellen Sensitivitätssteigerung der Probanden für bestimmte Strukturtypen, zeigt die Datenanalyse, dass Verschiebungen des Verzeichnungszentrums das Irritationspotenzial von S3D-Videos deutlich anhaben, bei hoher Trefferquote der Probanden bez. der Lokalisation des Fehlers. Die in den Testitems verwendeten Verschiebungen wurden relativ hoch gewählt. Trotzdem legen die Ergebnisse nahe, dass in der S3D-Produktion bei der Analyse und Korrektur von Objektivverzeichnungen die präzise Bestimmung des Verzeichnungszentrums äußerst wichtig ist.

Ausblick:

Leider ließen sich für eine Gegenprobe kaum Items/Felder mit geringen aniseikonie-ähnlichen Vergrößerungswerten finden. Für spätere Tests könnte hier evtl. auf ein anderes Produktionskonzept gesetzt werden. Bedenkenswert wären evtl. Items mit schwacher Verzeichnung aber globalen Brennweiten-Differenzen, so dass der Effekt eindeutiger zu lokalisieren ist.

Des Weiteren ergab die statistische Globalanalyse einen möglichen Fehler in der Item-Systematik gegenüber den Probanden. Da fast alle verzeichneten Items ihre *stärksten* lokalisierbaren Fehler links im Bild hatten, könnte dies zu einer Schema-Bildung bei den Probanden geführt haben, so dass insbesondere diese Bereiche einer besonderen Beobachtung durch die Probanden unterzogen wurden. Dieser Effekt konnte durch die 4 im Testablauf eingestreuten unverzeichneten Items (von 38) anscheinend nicht gemildert werden. Ihr Anteil müsste evtl. erhöht sowie zukünftig auch Items mit rechts-lastigem Fehler in den Testverlauf eingebaut werden.

Für eine weitere Testreihe gäbe es sinnvolle Erweiterungen der vorliegenden Untersuchung. Die Verwendung einer anderen S3D-Technologie könnte zeigen, ob die Erhöhung der vertikalen Auflösung auf die originalen 1080 Pixel (anstatt hier 540) eine Auswirkung auf Wahrnehmung von Verzeichnungsdifferenzen haben könnte.

Interessant wäre für spätere Tests evtl. auch ein Vergleich CGI-Animation vs. Realvideo mit gleichem Fehlergrad und derselben Verzeichnung. Daran ließe sich festmachen, ob bestimmte Medientypen sensitivere Probandenaussagen aufweisen.

Die verwendeten Werte der Verzeichnungszentrumsverschiebungen könnten in weiteren Untersuchungen feiner abgestuft und in weiteren Kombinationen untersucht werden.

Ausgehend von der hohen Medienaffinität der Probanden wäre es sicher interessant herauszufinden, ob eine Gruppe gleichaltriger Probanden, die keine hohe Medienvertrautheit angeben, zu gleichen Aussagen neigen würde.

Für zukünftige Tests sollte evtl. auch überlegt werden mit feldbezogenen Strukturklassen zu arbeiten, um den Einfluss bzw. die Maskierungsfähigkeit bestimmter Bildstrukturen auf die Wahrnehmung von Verzeichnungsdifferenzen in S3D-Bildern präziser zu dokumentieren. Andererseits könnte auch argumentiert werden, dass Items

vom Strukturtyp A generell die beste Indikation für das untersuchte Problem aufweisen, insbesondere wenn es um präzisere Analysen der globalen Fehlergrade 1 bis 2 geht.

Abschließend sei darauf hingewiesen, dass ein konsequentes binokulares Gazetracking der Probanden während des Testdurchlaufs zu einer präziseren Interpretation ihrer Wahrnehmungsbeeinträchtigungen beitragen könnte. Eine solche Untersuchung soll im Rahmen des Folgeprojekts BinoQ3D stattfinden.

Nachtrag: Die Ergebnisse einer finalen Testreihe, in der alle o.g. Analyseoptimierungen eingeflossen sind, wurden im Dezember 2015 auf der IEEE/IC3D-Konferenz in Liège präsentiert. Nachzulesen in:

Hottong, N.; Becker, S.; Beuck, S.: Evaluation of observers' sensitivity to lens distortion discrepancies in stereoscopic moving pictures, 4rd IEEE International Conference on 3D Imaging (IC3D) - Liège, 2015 (Tagungsband DOI: 10.1109/IC3D.2015.7391832).

Teil 3
Untersuchungen zur Montage in
Stereo 3D-Realfilmen

Benjamin H. Hauser

Abbildungsverzeichnis

Abbildung 1: Studiosituation	149
Abbildung 2: Eigenschaften der Probanden	156
Abbildung 3: Parallaxen in den 22 Test-Items	157
Abbildung 4: Irritationen im Testverlauf	158

1. Einleitung

Im Laufe der Filmgeschichte haben sich einige Montagetechniken für die monoskopische Filmproduktion entwickelt. Seitdem der Stereo-3D-Film Avatar einen finanziellen Megaerfolg verbuchen konnte, werden zunehmend mehr stereoskopische Filme produziert [1]. Die Stereoskopie wurde jedoch schon früher durch die Fotografie geprägt. [2] Die Kombination aus Stereoskopie und bewegtem Bild ist dabei nicht neu, jedoch sind die technischen Möglichkeiten seither stetig verbessert worden. [3] Dadurch kann sich erst heute die Gestaltung vom 2D-Film zur Gestaltung im stereoskopischen Film entwickeln.

Im monoskopischen Film gilt eine unauffällige Montage als Qualitätsmerkmal. [4] Bei den frühen Stereo-3D-Filmen galt dem Tiefeneffekt eine große Aufmerksamkeit. Seit einigen Jahren entwickeln sich Stereo-3D-Produktionen wieder dahingehend, die Handlung in den Vordergrund zu stellen, weshalb der Tiefeneffekt in neuen Produktionen oft dezenter als dramaturgische Unterstützung eingesetzt wird. [3]

Aufgrund fehlender Erfahrungen wird in der S3D-Produktion die Montage häufig nach den bekannten 2D-M Montagetechniken umgesetzt. Dabei wird oft vergessen, dass der stereoskopischen Montage zum Teil andere Rahmenbedingungen zugrunde liegen. So kann es z.B. dazu kommen, dass die Schnitte im Stereo-3D-Film für den Zuschauer nicht mehr unsichtbar werden, sondern diesen irritieren und aus dem Geschehen reißen. [5]

Das Ziel dieser Beitrags ist die Sensibilisierung für etwaige Differenzen zwischen der monoskopischen Filmmontage und einer möglichen optimierten stereoskopischen Filmmontage, sowie die Analyse möglicher Einflüsse und deren Auswirkungen. Zudem wird der in der Fachliteratur erwähnte Begriff »Tiefensprung« analysiert und definiert. Darauf basierend werden erste empirische Untersuchungen durchgeführt.

Für die zukünftige Produktion stereoskopischer Realfilme will dieser Bericht eine Orientierung schaffen, und einige Denkanstöße liefern. Besonderes Augenmerk liegt diesbezüglich auf der Notwendigkeit eines Tiefenscriptes, welches bereits vor den Dreharbeiten erstellt wird und die stereoskopischen Eigenschaften über den Filmverlauf bestimmt. Daraus ergibt sich konkret folgende Fragestellung:

»Welche Relevanz haben Tiefeneigenschaften stereoskopischer 3D-Realfilmaufnahmen auf die Montage und inwieweit kann ein Tiefenskript Irritationen zum Zeitpunkt des Schnittes beim Betrachter vermeiden?«

Hierbei dienen aktuelle Thesen aus der Fachliteratur als Grundlage:

„Wird zwischen Bildern sehr unterschiedlicher → Tiefenausdehnung oder → Tiefenposition geschnitten, kann es zu einem Tiefensprung kommen. Der Zuschauer wird dann gezwungen, blitzschnell die Augenkonvergenz zu verändern. Zu → Visueller Überforderung kommt es dabei vor allem bei Tiefensprüngen aus der Ferne in die Nähe.“ (Tauer, 2010, S. 530)

Da es sich hierbei um eine Grundlagenfrage handelt, sollen zunächst mögliche Einflussgrößen recherchiert und untersucht werden.

1.1 Abgrenzung

Bei der Produktion stereoskopischer Realfilm-Aufnahmen kann die Festlegung der absoluten Tiefenpositionierung durch Konvergieren der zwei benötigten Kameras erfolgen. Es besteht aber auch die Möglichkeit der nachträglichen Positionierung in der Postproduktion. Hierbei werden beide Kameras parallel ausgerichtet. In dieser Arbeit wird ausschließlich davon ausgegangen, dass die Produktion mit zwei parallelen Kameras erfolgt.

Bei der Produktion von Animationsfilmen bestehen umfangreichere Einflussmöglichkeiten auf stereoskopische Parameter, als es in einer Realfilmproduktion möglich ist. Umgekehrt ist es aber ebenfalls bei der Produktion von Animationsfilmen möglich, die möglichen Parameter aus der Realfilmproduktion zu kontrollieren. Daher sind die Ergebnisse dieser Arbeit, wenn auch für Stereo-3D-Realfilmproduktion ausgelegt, auch auf die Animationsfilm-Produktion anwendbar. Auch wenn es vielerlei Varianten von Bildübergängen gibt, welche sich bei der Montage zweier Einstellungen anwenden lassen, werden hier neben dem harten Schnitt (Cut) lediglich Überblendungen und Irisblenden thematisiert.

1.2 Methodik

Zunächst wird die Montage des monoskopischen Films und deren Entstehung untersucht, wobei hier nur die grundlegenden Methoden, welche eine häufige Anwendung finden, behandelt werden. Darauf aufbauend werden die unterschiedlichen Arten von Einstellungen und die drei wichtigsten Bildübergänge untersucht. Hiervon getrennt wird die Entwicklung der Stereoskopie beleuchtet und die wichtigsten Eigenschaften der menschlichen Tiefenwahrnehmung betrachtet. Dabei dient dieser Abschnitt vor allem dazu, ein einheitliches Verständnis über die Merkmale der stereoskopischen Filmproduktion und Filmbetrachtung herzustellen.

Darauf aufbauend wird im weiteren Verlauf die Relevanz stereoskopischer Eigenschaften von Stereo-3D-Realfilmaufnahmen für die Montage abstrahiert. Die dabei entwickelten Eigenschaften werden auf ihre Kontrollierbarkeit durch ein Tiefenscript erörtert. Auf dieser Grundlage wird eine empirische Untersuchung die Relevanz dieser Parameter für die Montage genauer untersuchen. Dabei dient die zuvor erfolgte Abstraktion als Grundlage für den Aufbau der Studie. Die Studie beschränkt sich dabei auf wenige Parameter, damit sie im Rahmen dieser Arbeit durchführbar bleibt. Die Studie liefert auch einen Ausblick für die weitere Forschung.

1.3 Begriffe

Um ein einheitliches Verständnis zu schaffen, werden in diesem Kapitel zunächst einige Begriffe und deren Bedeutung betrachtet. Insbesondere werden einige zentrale Bezeichnungen, welche oft umgangssprachlich in einer anderen Form Verwendung finden oder für die es noch keine klare Definition gibt, definiert.

1.3.1 Editing

Der Begriff Editing bezeichnet den handwerklich-kreativen Prozess des Zusammenfügens unterschiedlicher Bewegtbild-Sequenzen. Dabei ist die Bedeutung äquivalent mit dem Begriff der Montage. Editing wird oft auch als Sammelbegriff verwendet und beinhaltet dabei die Auswahl, Begrenzung und Anordnung von visuellen und akustischen Elementen eines Filmes (vgl. Beller, 2002, S. 9).

1.3.2 Cutting

Bezeichnet das Schneiden der bereits produzierten Aufnahmen in Einstellungen einer bestimmten Länge.

1.3.3 Szene

Eine Szene ist ein Abschnitt eines Filmes, der sich durch räumliche und/oder zeitliche Sprünge von weiteren Szenen abgrenzt.

1.3.4 Einstellung

Eine Einstellung ist eine Abfolge von einzelnen Bildern, die einen kontinuierlichen Verlauf aufweist. Eine Einstellung wird ebenso als Sequenz bezeichnet und besteht dabei mindestens aus einer Einstellung.

1.3.5 Tiefensprung

Der Begriff Tiefensprung wird im Folgenden verwendet, um Problematiken bei der Montage zu verdeutlichen. Jedoch erfolgt die Verwendung ausschließlich als Überbe-

griff für Unterschiede in den stereoskopischen Tiefeneigenschaften. Als Tiefensprung wird im Folgenden jeder Parallax-Unterschied zwischen zwei aufeinanderfolgenden stereoskopischen Bildern bezeichnet. Eine umfassendere Definition dieses Begriffes ist im Kapitel "Tiefensprünge" zu finden.

1.3.6 Tiefenposition

Die Tiefenposition bezeichnet im Folgenden die absolute Parallaxe eines Elementes im Bild.

1.3.7 Tiefenausdehnung

Der Begriff der Tiefenausdehnung bezeichnet im Folgenden die relative Parallaxe zwischen dem vordersten Nahpunkt und dem hintersten Fernpunkt in einem S3D-Bild.

2. Montage

Bei der Montage geht es um das ästhetische Prinzip, bei dem geschnittene Einstellungen und Inhalte aneinandergesetzt werden. Hierdurch wird eine freie Bewegung durch Raum und Zeit ermöglicht (vgl. Steber, 2007, S. 13). Wird hierbei der Aufbau der Einstellungen, innerhalb einer Szene, dem menschlichen Blickrhythmus angenähert, kann erreicht werden, dass der Schnitt vom Zuschauer nur unbewusst wahrgenommen wird (vgl. Steber, 2007, S. 14). Auf dieser Basis gibt es unterschiedliche Methoden, um eine flüssige Montage, mit für den Zuschauer nur unbewusst wahrnehmbaren Schnitten, zu ermöglichen. Bei Anwendung des Shot/Reverse-Shot Schemas wird zwischen zwei Einstellungen geschnitten, wovon die Zweite die Reaktion auf die Aktion der ersten Einstellung liefert. Hierbei ist es zudem möglich, eine Reaktion auf eine Aktion bereits während der Aktion zu zeigen, wenn die Aktion über die Tonspur vermittelt wird (vgl. Beller, 2002, S. 17). Die Anschlüsse sorgen für einen flüssigen Übergang, indem diese eine logische, zeitliche oder räumliche Kontinuität schaffen (vgl. Katz, 2004, S. 201). Bei der Montage kommt es also darauf an, dass die Schnitte vom Zuschauer nur unterbewusst wahrgenommen werden. Daher gilt es im Allgemeinen eine Kontinuität zu schaffen, welche einzelne Einstellungen innerhalb einer Szene miteinander verbindet. Der Breakdown ist dabei ein wichtiger Bestandteil, um eine flüssige Montage zu ermöglichen. Hierbei müssen Einstellungsgrößen und Perspektiven zum logischen Inhalt der Szene passen. Ansonsten kann es passieren, dass Inhalte gegenüber der Konzeption verfälscht wahrgenommen werden. Zudem kann es auch zu ungewollten Irritationen der Zuschauer führen, wenn die räumliche Vorstellung des Motivs gestört wird oder Schnitte keine Logik aufweisen. Die Montage hat also einen großen Einfluss auf die Aussage eines Filmes.

2.1 Die Geschichte der Montage

Mit dem Erfinder Thomas Alva Edison beginnt an dieser Stelle die Filmgeschichte, da durch seine Erfindung die ersten Bewegtbilder hergestellt und gegen Bezahlung betrachtet werden konnten. Das handbetriebene Kinetoskop wurde 1893 von Edison patentiert und ab 1894 am Broadway und in größeren Städten gezeigt. Der Münzfilmkasten konnte eine Filmschleife von 25 bis 50 Fußlängen fassen, wodurch die Filme bis zu einer Minute dauern konnten, bevor sie wieder am Anfang starteten. Hierfür entstanden Vorführräume, in denen jeder an seinem Guckkasten kurbeln musste. Dabei bestand jeder Film aus einer einzelnen Szene, welche aus nur einer Einstellung bestand. Um die Filmrollen am Anfang und Ende besser zusammenfügen zu können, oder diese besser in das Vorführgerät zu bekommen, wurden lediglich die beiden Filmmenden beschnitten (vgl. Beller, 2002, S. 12).

Im Jahre 1895 wurden zum ersten Mal mehrere solcher Filme mit dem Bioscope auf eine Leinwand projiziert. Zu dieser Zeit war die Filmvorführung Teil des Varietés, weshalb ein größeres zahlendes Publikum erreicht werden konnte. Edison patentierte daher 1896 das Vitascope, womit er auch ein größeres Publikum bei öffentlichen Filmvorführungen bedienen konnte (vgl. Beller, 2002, S. 12f).

Erst 1896 entstand durch einen Zufall die Montage. Beim Filmen des Pariser Straßenverkehrs verfiel sich der Film des französischen Filmkünstlers George Méliès innerhalb seiner Kamera, weshalb er erst nach einer kurzen Pause den Film weiter kurbeln konnte. Dabei hatte sich die aufgenommene Szenerie jedoch völlig verändert:

„Als ich den Streifen, den ich dort, wo er gerissen war, wieder zusammengefügt hatte, projizierte, sah ich plötzlich, daß [sic!] aus dem Omnibus Madeleine-Bastille ein Leichenwagen und aus Männern Frauen geworden waren.“ (Ebert, 1979, S. 558)

Dieses Prinzip verwendete Méliès daraufhin häufiger und tauschte beim Anhalten der Kamera Gegenstände aus oder ließ Menschen verschwinden. Dabei blieben es aber Szenen aus einer Einstellung. Seine Filme bestanden jedoch lediglich aus mehreren Single-Shot Szenen. Dabei wurde diese Art von Filmen in Frankreich und Amerika schon oft mit Irisblenden oder Überblendungen umgesetzt, wodurch harte Schnitte vermieden wurden (vgl. Beller, 2002, S. 13).

Kameramänner, die das aktuelle Geschehen aufnahmen, begannen hingegen ab 1897 die Kamera an neuen Standorten aufzubauen, um einem Ereignis zu folgen oder suchten sich zwischendurch eine bessere Kameraperspektive, da sie den Filmdreh schlecht planen konnten. Daraus entstanden also bereits Filme mit unterschiedlichen Einstellungen, obwohl Filme noch nicht geschnitten wurden. Dies dauerte bis ca. 1900 an, ehe auch zur Lenkung der Aufmerksamkeit, die Kamera den Schauspielern und deren Spiel zu folgen begann. Dadurch entwickelte sich die Filmkunst und die Gestaltung erlangte neue Möglichkeiten, z.B. die Ereignisse in ihrer Dauer und Chronologie zu beeinflussen. Die Montage und der Schnitt wurden genutzt, um Einstellungen abzuwechseln oder neu anzuordnen (vgl. Beller, 2002, S. 13f).

So entwickelte sich bis 1905 der dokumentarische Film zu einem Medium, zur Erzählung von 13-16 Minuten dauernden fiktiven Geschichten. Zu dieser Zeit entstand in Amerika das Nickelodeon als Ort zur Aufführung der Filme. Ab dem Jahr 1906 wurde die Arbeit bei der Filmproduktion stärker aufgeteilt, Regisseure aus dem Theater wurden eingestellt und nahmen dem Kameramann die Regie ab (vgl. Beller, 2002, S. 14f).

Im Jahr 1910 kam das Starsystem auf und es erfolgte die Nennung der Schauspieler. Das Starsystem brachte die Kamera dazu, sich im Raum zu bewegen, um mit den Augen des Helden zu sehen. Die Kamera war nicht mehr nur der außenstehende Beobachter. Somit waren nun andere Einstellungsgrößen und abwechslungsreichere Kameraperspektiven gefordert, wodurch die notwendige Montage Komplexität erlangte (vgl. Beller, 2002, S. 15).

Bis 1914 entwickelten sich Normen, Standards, eine Arbeitsteilung und es setzte sich eine Zentralisierung in der Filmbranche durch. Es entstand die industrielle Filmproduktion. Um den Wünschen des Publikums zu entsprechen, wurden die Filme länger. Nachdem die Filme ab 1911 nicht mehr auf eine Filmspule passten, entstanden die ersten Mehrspuler und der Beruf des Filmeditors entwickelte sich. Ab 1913 setzte sich eine abendfüllende Spielfilmlänge von über einer Stunde durch. Dies sorgte nun auch für das Aufkommen eines Produzenten, welcher die Produktion aus ökonomischer Sicht organisierte, weshalb Dreharbeiten nicht mehr in chronologischer Reihenfolge erfolgten. Daher wurde auch ein Breakdown in der Planung von Einstellungen und Szenen notwendig. Daraufhin entstand das (Handlungs)-Achsenschema (auch 180 Grad-Prinzip genannt), um für den Zuschauer eine Orientierung in Raum und Zeit zu schaffen (vgl. Beller, 2002, S. 15).

2.1.1 (Handlungs)-Achsenschema

Nach diesem Schema wird ein Handlungsraum unterteilt, wobei der Zuschauer (die Kamera) immer auf der gleichen Seite der Handlung bleibt. Demnach kann der Zuschauer nur auf die andere Seite der Handlung gelangen, wenn die Kamera innerhalb einer Einstellung die Seite durch eine Bewegung über die Handlungsachse wechselt. Ansonsten kommt es zur Desorientierung beim Betrachten des Filmes. Es bleibt also ein Radius von 180 Grad um die Handlung für die Kamera. Bei Überschreitung dieser Grenze durch einen Schnitt entsteht ansonsten ein Achsensprung. Dabei etablierten sich bestimmte Kameraeinstellungen, wie der Master-Shot, die Over-Shoulder oder Cut-In Einstellung, etablierte sich laut Beller eine gute Montage (vgl. Beller, 2002, S. 16).

2.1.2 Shot/Reverse-Shot Schema

Darauf aufbauend entstand zwischen 1911 und 1914 das Shot/Reverse-Shot Schema (SRS-Schema). Dabei wird zwischen zwei Einstellungen geschnitten, welche jeweils im Wechsel Aktion und die darauf folgende Reaktion zeigen. Nach diesem Prinzip kann beliebig oft hintereinander geschnitten werden, wobei die Blickanschlüsse be-

achtet werden. Dieses Prinzip kann sowohl für Over-Shoulder als auch für Point-of-View Einstellungen angewandt werden.

So etablierte sich ein analytisches System, nach dem gedreht und montiert wird. Durch einen Establishing-Shot wird eine Szene eingeleitet. Darauf folgend werden ein Cut-In und eine Reihe Einstellungen nach dem SRS-Schema geschnitten, ehe ein Cut-Back die Szene abschließt. Dabei kann das SRS-Schema durch einen Cut-Away unterbrochen werden, welcher z.B. durch eine Kopfdrehung motiviert wird. Dabei ist die Kontinuität, die Blickführung und die räumliche Orientierung durchgängig ein Bestandteil der Montage und sorgt für eine logische Montage (vgl. Beller, 2002, S. 17f).

2.1.3 Kontinuität

Ab ca. 1910 entstand der Begriff der Kontinuität. Dabei galt es sowohl räumliche, als auch zeitliche Kontinuität über unterschiedliche Zeiten und Räume hinweg zu schaffen. Es muss hierbei nicht nur innerhalb einer Szene, sondern auch über die Szene hinaus eine Kontinuität geschaffen und gelenkt werden. Die Notwendigkeit dient dabei nicht nur der Orientierung für die Zuschauer, sondern auch ökonomischen Aspekten. So war die Planung ab 1910 Bestandteil des Skriptes, um Richtlinien für die Dreharbeiten festlegen zu können. Es galt dabei vor allem eine Geschichte flüssig zu erzählen, ohne zeitliche Sprünge zu erzeugen, und um beim Zuschauer für eine Raumordnung zu sorgen. Gleichzeitig entstand daher das Continuity-Script, wohingegen sich erst zwischen 1917 und 1920 die Berufe Script-Girl und Continuity-Clerk entwickelten (vgl. Beller, 2002, S. 18f).

Um Kontinuität im Schnitt zu ermöglichen, müssen Bewegungen überlappend gedreht werden. Dadurch kann eine genaue Positionierung des Schnittes in der Bewegung erfolgen, ohne dass die Bewegung ins Stocken gerät, Lücken oder doppelte Abläufe enthält. Diese Art der Montage erforderte jedoch zunächst einen Lernprozess von Filmemachern und Zuschauern. Zuvor wurden gleichzeitige Abläufe durchaus hintereinander gezeigt, ehe zur Schaffung von Kontinuität der Cross-Cut entstand. Auch in Bewegungen über Szenen hinweg werden Bewegungen zur Überleitung eingesetzt. Dabei kann beispielsweise eine Türe durch ihre Bewegung eine Innen- und Außenszene miteinander verbinden. Jedoch kann auch beispielsweise eine Bewegungsrichtung der Kamera zweierlei Szenen verbinden. Dies ermöglicht dem Zuschauer das Eintauchen in die Erzählung und verhindert das bewusste Wahrnehmen von Schnitten, die durch ihre Härte auf sich aufmerksam machen. Vor allem in amerikanischen Filmen sollte der Blick des Zuschauers sanft geführt werden und die Schnitte vergessen lassen (vgl. Beller, 2002, S. 19f).

2.1.4 Cross-Cutting

Diese Art des Hin- und Herschneidens wurde insbesondere von Griffith weiterentwickelt und unter anderem in seinen Filmen *The Lonely Villa* (1909) und *The Londale*

Operator (1911) angewendet. Das besagte Montagemuster spricht die Bedürfnisse nach Abwechslung und Wiederholung an. Dabei wird chronologisch zwischen simultanen Ereignissen im Wechsel geschnitten. Hierdurch kann die Spannung eines Filmes deutlich gesteigert werden, weshalb diese Art der Montage oft gegen Filmente eingesetzt wird, um vor dem Finale nochmals die Spannung beim Zuschauer zu steigern. Alternativ ist jedoch auch die Parallelmontage möglich, bei der zwischen unterschiedlichen Ereignissen hin und her gewechselt wird. Diese sind untereinander vergleichbar, ohne dass sie dabei chronologisch oder simultan sein müssen. Die Parallelmontage ist folglich eine Art des Cross-Cuttings. Dabei können z.B. auch Kontraste zwischen zwei Handlungssträngen oder die gleiche Bedeutung der Handlungsstränge für eine Verbindung sorgen. Durch Cross-Cutting sind Sprünge durch Raum und Zeit, wie in der schriftlichen Literatur, möglich (vgl. Beller, 2002, S. 23f).

2.1.5 Montage-Ellipsen

Nicht alles, was in einer Handlung logisch passiert, ist für die Erzählung der Handlung notwendig. Montage-Ellipsen werden verwendet, um die Geschichte kompakter zu machen und Unwesentliches heraus zu schneiden. Dabei ist es die Kunst, die Handlung so wieder zusammensetzen, dass der Zuschauer die Erzählung als kontinuierlich fortlaufende Bewegung wahrnimmt. Die Aufgabe im Schneiderraum besteht also auch hier im Schaffen von sanften Übergängen, welche weich und für den Zuschauer unsichtbar werden sollen (vgl. Beller, 2002, S. 25f).

2.1.6 Match-Cut

Beim Match-Cut handelt es sich um eine kontinuierliche Bewegungsmontage, bei der zumeist eine Figur zwischen mehreren Räumen wechselt. Dabei übernimmt die gleiche kontinuierliche Bewegung der unterschiedlichen Einstellungen einen fließenden Szenenwechsel. Hierdurch wird für den Zuschauer die Kontinuität bei großen Sprüngen durch Raum und/oder Zeit möglich (vgl. Beller, 2002, S. 26).

2.1.7 Zwischenschnitt

Der Zwischenschnitt bietet vielerlei Möglichkeiten, wie beispielsweise in eine kontinuierliche Handlung Informationen wie Landkarten, Briefe oder sogar Zwischentitel für den Zuschauer einzufügen. Oft dienen Zwischenschnitte ausschließlich dem Zweck, die kontinuierliche Erzählung beizubehalten, obwohl die Handlung unterbrochen oder ein Teil herausgeschnitten wurde. Insbesondere bei Interviews mit einer Kamera wird durch einen Zwischenschnitt eine Kürzung möglich, ohne dass es dem Zuschauer auffällt, wenn der Interviewte dabei weiter spricht (vgl. Beller, 2002, S. 27).

2.2 Die Einstellung

Wie im letzten Kapitel erläutert, haben sich in der Filmgeschichte eine Reihe unterschiedlicher Montagetechniken etabliert. Um die Montage nach diesen Methoden zu ermöglichen, haben sich unterschiedliche Einstellungsgrößen und Kameraperspektiven entwickelt.

Die typische Kameraeinstellung zur Einführung von Motiven oder Szenen ist der Master-Shot. Bei dieser Einstellung wird der Handlungsbereich weitwinklig abgebildet, wobei die Kameraposition bereits die Handlungsachse definiert. Die Handlung spielt folglich vor der Kamera, wobei die Kamera über den Handlungsverlauf nie in jene Richtung filmt, aus der der Master-Shot aufgenommen wurde.

Im weiteren Verlauf einer Szene stellt die Over-Shoulder-Einstellung einen Klassiker dar, wenn nicht sogar den Klassiker schlechthin. Durch deren Einsatz wird eine gleichzeitige Reaktion und Aktion zweier Figuren möglich. Die Kamera kann in dieser Position aber auch von etwas außerhalb auf das Geschehen schauen, wodurch die Distanz zum Geschehen für den Zuschauer vergrößert wird. Wird dagegen die Kamera dem Blick der Figur angenähert, wird der Zuschauer stärker in die Situation der Figur einbezogen, bis die Kamera vollständig die Perspektive der Figur übernimmt.

Zu den Blickwinkeln gibt es unterschiedliche Einstellungsgrößen, die üblicherweise eingesetzt werden. Diesbezüglich ist zu beachten, dass Einstellungsgrößen mit gleichem Inhalt, zwischen dem geschnitten werden soll, einen deutlichen Kontrast erfordern. Auch der Blickwinkel ist dabei zu beachten. Insbesondere leichte Winkeländerungen können für Verwirrungen sorgen. Allgemein sind bei der Montage zweier Einstellungen sowohl deutliche Bildunterschiede als auch verbindende Elemente in den Einstellungen notwendig (vgl. Reisz, 1991, S. 217-221). Gibt es in der ersten Einstellung eine Bewegungsrichtung, muss diese in der zu montierenden Folgeeinstellung fortgeführt werden, um einen flüssigen Schnitt zu erhalten. Gegenläufige Bewegungen sind dabei unbedingt zu vermeiden (vgl. Reisz, 1991, S. 224f).

2.3 Transition

Neben den verschiedenen Methoden der Montage, womit unterschiedliche Handlungsabläufe in Einstellungen aufgelöst und montiert werden können, gibt es neben dem harten Schnitt noch weitere Transitions-Varianten. Es gibt so vielfältige Bildübergänge, dass hier lediglich die Wichtigsten Beachtung finden werden. Da der harte Schnitt mit Abstand am häufigsten in Kino und TV-Produktionen Verwendung findet, wird dieser im Weiteren priorisiert behandelt. Allerdings haben sich in der Geschichte der Montage noch weitere Bildübergänge sehr früh entwickelt und etabliert, weshalb diese in diesem Kapitel erläutert werden.

2.3.1 Cut

Der harte Schnitt kann durch die richtige Montage, wie in den vorangegangenen Kapiteln beschrieben, einen flüssigen unauffälligen Wechsel zwischen Einstellungen ermöglichen. Er wird in Filmproduktionen zum größten Teil eingesetzt. Als einziger Übergang hat er keine Länge und es kann durch gute Montage erreicht werden, dass er vom Zuschauer nicht bewusst wahrgenommen wird.

2.3.2 Überblendung

Die Überblendung war als weicherer Bildwechsel schon früh möglich und diente der Milderung des harten Schnittes; insbesondere zwischen Szenen oder bei Montage von Einstellungen, die nicht optimal zusammenpassen. Bei Verwendung einer Überblendung kann dann der Schnitt abgemildert werden. Die Überblendung wird mittlerweile als Gestaltungsmittel eingesetzt, um Sprünge in Raum oder Zeit anzudeuten. Sie findet beispielsweise bei der Abtrennung von Traumsequenzen oder Rückblenden Verwendung (vgl. Ast, 2002, S. 18, 19, 27, 35, 65).

2.3.3 Irisblende

Ebenso wie die Überblendung wurde auch die Irisblende schon früh verwendet, um sehr harte Schnitte zu vermeiden und einen fließenderen Übergang zu schaffen. Die Irisblende wird zur Kaschierung von nicht zueinander passenden Übergängen verwendet. Zudem wird sie auch als Stilmittel zur Separation von Szenen verwendet, um das Verstreichen von Zeit oder einen Ortswechsel zu verdeutlichen. Die Irisblende wirkt dabei wie die Abdunkelung eines Raumes. Dies kommt dem natürlichen Sehen zugute, da dieser Effekt auch dort beim Blinzeln und Ausschalten des Lichtes auftritt.

3. Tiefenwahrnehmung

In diesem Kapitel wird nun das natürliche räumliche Sehen erläutert. Dabei wird insbesondere auf die Unterschiede zwischen dem natürlichen Sehen und der Stereo-3D-Filmbetrachtung, wie sie heute eingesetzt wird, eingegangen.

Um sehen zu können, nutzt der Mensch zwei Augen. Diese werden dabei gleichzeitig verwendet und ermöglichen die Wahrnehmung der räumlichen Umgebung. Dabei kommt es im natürlichen Sehen üblicherweise nicht zu einer doppelten Wahrnehmung der Umgebung trotz der Nutzung zweier Augen. Dies wird durch das Gehirn ermöglicht, welches die Wahrnehmungen der beiden Augen fusioniert und dem Menschen das räumliche Sehen ermöglicht. Für das räumliche Sehen sind vor allem drei Effekte verantwortlich. Monokulare Tiefeninformationen und die patente Stereopsis liefern dabei die wichtigsten Tiefeninformationen. Sie ermöglichen vor allem die detaillierte Wahrnehmung von Tiefe innerhalb eines Objektes und von Objekten zueinander. Die Okulomotorik ist dagegen ein schwacher Tiefenreiz, der nur eine grobe absolute

Einschätzung von Entfernungen ermöglicht. Okulomotorik und die patente Stereopsis wirken insbesondere bei kleinen Distanzen stärker. Bei größeren Entfernungen werden kaum noch stereoskopische Tiefenreize hervorgerufen und der Mensch muss sich auf monokulare Tiefeninformationen verlassen.

Der Mensch kann mit gesunden Augen seine gesamte Umgebung scharf sehen. Dies wird möglich, da die Augen bei der Fixation eines Punktes auf dessen Entfernung akkomodieren und somit dieser Punkt scharf wahrgenommen werden kann. Allerdings sind die menschlichen Augen lediglich in der Lage, einen sehr kleinen Bereich um den Fixationspunkt scharf zu sehen. Da sich die Augen aber kontinuierlich bewegen und akkomodieren, sobald ein Punkt fixiert wird, erscheint die Welt insgesamt scharf. Gleichzeitig findet bei Fixation eines Punktes ein Konvergieren der Augen statt. Die Augen drehen sich also so weit ein, dass im linken und rechten Auge die fixierten Punkte auf aneinander zugehörige Stellen im Auge auf die Netzhaut treffen. Dadurch wird an dieser Stelle eine Fusionierung der Bilder möglich und es kommt nicht zur doppelten Wahrnehmung. Dabei kann aber nur ein kleiner Bereich um den Fixationspunkt fusioniert werden, weshalb die Umgebung für Doppelbilder sorgt. Diese werden vom Gehirn unterdrückt und für gewöhnlich nicht wahrgenommen. Möglich wird dies unter anderem durch die enorme Unschärfe, welche bei Objekten anderer Tiefe, wie die akkomodierte, einsetzt (Kuhn, 1999, S. 21).

Was beim natürlichen Sehen passiert, ist bei der Betrachtung eines Stereo-3D-Films in dieser Art nicht möglich. Die Leinwand oder der Fernseher hat einen festen Abstand zum Betrachter. Egal, in welche Tiefe der Zuschauer des Stereo-3D-Films schaut, um das Bild scharf sehen zu können, muss auf den Fernseher akkomodiert werden. Hier werden dann lediglich die gleichen Inhalte für das linke und rechte Auge versetzt dargestellt, wodurch eine Positionierung in der Tiefe erfolgt. Dabei wandern Objekte, welche im rechten Bild weiter rechts und im linken Bild weiter links dargestellt werden hinter den Bildschirm, wohingegen umgekehrt die Objekte vor dem Bildschirm schweben. Durch unterschiedliche Blickwinkel auf die Objekte bei der Aufnahme erhalten auch die Objekte an sich Tiefe. Es entstehen Parallaxen. Sind Bildinformationen durch Bewegungsunschärfe oder geringe Schärfentiefe nicht scharf verfügbar, kann kein Scharfstellen durch Akkommodation erfolgen. Umgekehrt wird die Umgebung, welche sich in einer anderen Entfernung befindet um Fixationspunkte weniger unscharf, da lediglich durch die schlechtere Wahrnehmung der Augen um den Fixationspunkt die Unschärfe entsteht, nicht aber durch die fehlende Akkommodation in diese Tiefe. Daher besteht hier leicht die Gefahr, dass es zur Wahrnehmung von Doppelbildern kommt (Kuhn, 1999, S. 21f).

3.1 Monokulare Depth Cues

Monokulare Tiefeninformationen werden seit langem bereits in der 2D-Filmproduktion bewusst eingesetzt. Dadurch ist es dem Zuschauer auch beim Betrachten eines klassischen Filmes möglich, eine Räumlichkeit wahrzunehmen, auch wenn stereosko-

pische Informationen nicht vorhanden sind (vgl. Mendiburu, 2009, S. 11). Monokulare Depth Cues können in folgende Kategorien aufgeteilt werden:

3.1.1 Perspektive & relative Größe

Bei gleichen Objekten kann durch den Größenunterschied der Objekte die Entfernung zueinander bestimmt werden. Dabei spielen die Perspektive und insbesondere die Brennweite eine wichtige Rolle (vgl. Mendiburu, 2009, S. 11).

3.1.2 Vertraute Größe

Mittels Größeninformation bekannter Objekte wie Menschen, Häuser oder Autos lässt sich deren Entfernung einschätzen (vgl. Hofmeyer, 2009, S. 32).

3.1.3 Texturdichtegradient

Aus der Natur hat der Mensch gelernt, dass Texturen die sehr fein sind, sich viel weiter entfernt befinden als grobe Texturen (vgl. Mendiburu, 2009, S. 12).

3.1.4 Bewegungsparallaxe

Bei einer Kamerafahrt bewegen sich nahe Objekte viel schneller als weit entfernte Objekte. Daher wird im 2D-Film oft eine Dolly-Fahrt zur Erzeugung von Bewegungsparallaxen eingesetzt, wodurch ein Effekt von Räumlichkeit ermöglicht wird (vgl. Tauer, 2010, S. 413-414).

3.1.5 Verdeckung

Bei Objekten, die sich gegenseitig verdecken, wird das verdeckte Objekt als weiter entfernt wahrgenommen als das verdeckende Objekt (vgl. Hofmeyer, 2009, S. 30).

3.1.6 Relative Höhe

Je höher der tiefste Punkt eines Objektes im Gesichtsfeld liegt, desto weiter weg wird er wahrgenommen. Hierbei hat auch die Blickrichtung einen wesentlichen Einfluss auf die wahrgenommene Tiefe, wobei der Horizont als Indikator dient, um die Perspektive zu bestimmen (vgl. Hofmeyer, 2009, S. 30).

3.1.7 Atmosphärische Perspektive

Bei sehr großen Entfernungen kommt es vor, dass durch Partikel in der Luft die Farbe verfälscht wird, eine leichte Unschärfe auftritt und der Kontrast sowie die Sättigung sinken. Daher kann die atmosphärische Perspektive ein wichtiger Tiefenhinweis für

große Entfernungen sein (vgl. Mendiburu, 2009, S. 13). Dieser Effekt kann auch in der Filmproduktion eingesetzt werden, beispielsweise durch die Verwendung von Nebel (vgl. Schumacher, 2009, S. 125).

3.1.8 Schatten

Zumeist gibt es eine klar erkennbare Lichtrichtung. Die dabei entstehenden Schatten und Lichtreflexionen sorgen für die Erkennung von Tiefe. Dieser Tiefenreiz gilt oft als das wichtigste Gestaltungsmittel einer klassischen Filmproduktion. Erst durch das Licht und insbesondere die Schatten, die Objekte plastisch machen, wird ein Bild spannend. Daher wird oft das Backlight verwendet, um Personen vom Hintergrund abzuheben und sie plastischer wirken zu lassen (vgl. Mendiburu, 2009, S. 14).

3.1.9 Perspektivische Konvergenz

Linien, die parallel verlaufen, scheinen mit zunehmender Entfernung zu konvergieren. Das bedeutet, dass durch die optische Annäherung von Linien in der Tiefe, eine räumliche Achse geschaffen werden kann. Ein klassisches Beispiel ist der Blick entlang einer Straße. Dabei scheinen die Häuser immer näher beieinanderzustehen und die Straße schmaler zu werden, je größer die Entfernung ist (vgl. Hofmeyer, 2009, S. 31).

3.1.10 Okulomotorik

Die Okulomotorik wird durch ein Konvergieren der Augen, zur Fixation eines Punktes im Raum ermöglicht. Dabei werden durch Muskeln die Augen eingedreht und gleichzeitig erfolgt die Wahrnehmung von Tiefe. Hierbei wird wahrgenommen, wie weit die Augen eingedreht werden müssen, um einen Punkt zu fixieren. Diese Tiefeninformation funktioniert insbesondere im Nahbereich sehr gut, da sich hier die Winkeländerung und die damit verbundene Muskelanspannung stark ändern. Beim Sehen in größere Entfernungen werden die Blickachsen der Augen zunehmend parallel und es erfolgen nur noch sehr geringe Winkeländerungen. Kritisch ist dies insbesondere bei der Betrachtung naher Stereo-3D-Displays, da die Akkommodation im natürlichen Sehen mit der Konvergenz gekoppelt ist. Da bei der Stereo-3D Betrachtung aber Akkommodation und Konvergenz auf unterschiedliche Distanzen erfolgen muss, können im Nahbereich schnell zu große Abweichungen auftreten (vgl. Tauer, 2010, S. 52f).

3.1.11 Patente Stereopsis

Die patente Stereopsis liefert den wichtigsten Hinweis auf Tiefe im natürlichen Sehen. Sie ermöglicht bei Stereo-3D-Filmen erst die Tiefenwahrnehmung. Sie sorgt dafür, dass der Mensch feinste Tiefenunterschiede wahrnehmen kann. Bei der Fixation eines Punktes wird dieser mit beiden Augen durch Konvergenz aufeinander zugehörigen

Stellen der Netzhaut wahrgenommen. Gleichzeitig gibt es aber bei gleicher Konvergenz weitere Punkte, die auf der Netzhaut auf zueinander gehörende Stellen treffen. Die Summe all dieser Punkte bildet eine Kreisform, welche für jede Fixationsentfernung unterschiedlich ist. Dieser Kreis wird Horopter genannt und grenzt den Raum in Vorder- und Hintergrund, in Relation zum aktuellen Fixationspunkt, ab. Dabei ermöglicht das Gehirn aber noch eine Fusionierung eines engen Toleranzbereiches in der Tiefenausdehnung um den Fixationspunkt. Hierbei werden folglich disparate Netzhautstellen durch denselben nicht fixierten Punkt im Raum angesprochen. Der Bereich, in dem die Fusionierung möglich ist, wird als Panumraum bezeichnet. Die Fusionierung im Panumraum ermöglicht das Sehen von sehr fein abgestuften Tiefenunterschieden um den aktuellen Fixationspunkt. Folglich entsteht der 3D-Effekt durch das Auftreffen von Punkten um den Fixationspunkt auf disparate Netzhautstellen in einem kleinen Bereich um den Fixationspunkt. Die auftretende Reizung, einander nicht zugehöriger Netzhautpunkte, wird als Querdisparität bezeichnet und wird als abweichender Winkel zum Konvergenzwinkel in Winkelminuten gemessen. Wird die Tiefenausdehnung dabei überschritten, kann keine Fusionierung ermöglicht werden und es kommt zur Wahrnehmung von Doppelbildern (Kuhn, 1999, S. 41f).

Bei der Produktion stereoskopischer Aufnahmen ist folglich zu beachten, dass keine zu großen Querdisparitäten auftreten. Diese werden dabei durch Parallaxunterschiede zwischen nahen und fernen Objekten im Bild erzeugt. Zugleich sollten absolute Parallaxen nicht zu groß werden, da die Augen des Betrachters in jedem Fall auf die Projektionsebene akkomodieren müssen und somit jede Parallaxe eine Abweichung zwischen Akkommodation und Konvergenz darstellt. Werden Parallaxen absolut zu groß, weicht der notwendige Konvergenzwinkel von dem bei der Akkommodation üblichen Konvergenzwinkel zu stark ab und es werden Doppelbilder wahrgenommen. Dabei ist zu beachten, dass die Skalierung stereoskopischer Produktionen auf größere Projektionsflächen als die bei der Konzeption bestimmte, für die Vergrößerung von relativen und absoluten Parallaxen sorgt. Dadurch kann es schnell passieren, dass absolute Parallaxen im Hintergrund den Abstand der Augen überschreiten und somit eine Divergenz der Augen notwendig wird, um den Hintergrund zu fixieren. Beim natürlichen Sehen kommt dies nicht vor, weshalb es zunächst zu großen Anstrengungen kommt, bevor die Fusionierung nicht mehr möglich ist. Zudem besteht die Gefahr zu großer relativer Parallaxen durch Verringerung des Betrachtungsabstandes gegenüber der Kalkulation bei der Stereo-3D-Produktion, wodurch erhöhte Querdisparitäten auftreten.

4. Montage in der Stereo-3D-Produktion

Erst mit dem Stereo-3D-Hype, der seit der Jahrtausendwende entstanden ist, sind stereoskopische Produktionen in hoher Präzision möglich. Durch die digitale Produktion ist sowohl eine Bildkontrolle am Set, wie auch eine umfassende Korrektur in der Postproduktion möglich. Probleme durch Asynchronität oder Misalignment können

zunehmend verhindert werden [6]. Die Feinheiten, in denen sich die Gestaltung von Stereo-3D-Produktionen von monoskopischen Produktionen unterscheidet, sollen zunehmend berücksichtigt werden. Bei der Gestaltung stereoskopischer Standbilder haben sich in der Geschichte bereits wichtige Gestaltungsmerkmale entwickelt [7]. Der Zeitpunkt des Schnittes war für die Fotografie und deren Gestaltung nicht von Bedeutung, weshalb hier keinerlei Forschung betrieben wurde.

4.1 Technischer Bildwechsel

Bei der Montage von stereoskopischen Produktionen muss im Gegensatz zur Montage in der monoskopischen Filmproduktion ein Bildwechsel bei zwei Bildmaterialien erfolgen. Wichtig ist hierbei, dass sowohl der Wechsel zweier Einstellungen für beide Augen simultan erfolgt. Bei einem harten Cut, der sowohl bei monoskopischen Produktionen als auch bei Stereo-3D-Produktionen am häufigsten verwendet wird, ist der technische Wechsel sehr einfach möglich, indem zweierlei Videos simultan geschnitten und montiert werden.

Bei einer Irisblende oder auch Schwarzblende ist die Verwendung in stereoskopischen Produktionen ebenso einfach zu realisieren, indem beide Filmsequenzen synchron abgeblendet und aufgeblendet werden. Dabei sind auch Blenden auf andere Farben wie schwarz auf gleiche Weise möglich. Allerdings ist darauf zu achten, dass auch bei relativ gut angepassten Kameras schnell ein Helligkeitsunterschied durch den Spiegel entstehen kann. Dies passiert durch den verwendeten Spiegel im Spiegelrig, welcher nicht exakt die Hälfte des einfallenden Lichtes auf beide Kameras aufteilt. Wird nun eine Weißblende am Computer generiert, kann es vorkommen, dass leichte Helligkeitsunterschiede im originalen Bildmaterial dafür sorgen, dass Lichter nun zu unterschiedlichen Zeitpunkten beginnen auszubrennen. Dadurch können leichte Helligkeitsunterschiede, welche im Originalmaterial nicht störend wirkten, nun für Zuschauer störend werden.

Die Überblendung von Videosequenzen ist technisch ebenso möglich wie die Schwarzblende. Dabei kann eine Überblendung trotz richtiger technischer Umsetzung für den Zuschauer nun störend sein. Wie beschrieben, wird eine Überblendung in der 2D-Produktion oft für Sprünge durch Zeit und Raum verwendet. Dabei ist dem Publikum die Bedeutung durch ihre Seherfahrung bekannt. Allerdings sorgt die Überblendung in einem stereoskopischen Film für einen ganz anderen Effekt. Daher hat eine Überblendung für die Zuschauer auch eine andere Bedeutung, auch wenn die gewollte Bedeutung der Filmemacher durch die 2D-Sehgewohnheiten errahnt werden kann. Wird bei einer stereoskopischen Produktion eine Überblendung verwendet, bedeutet dies, dass der Zuschauer den dreidimensionalen Raum nicht einfach wechselt, sondern dass sich der alte Raum in dem sich der Zuschauer befindet, auflöst. Dies kann natürlich gewollt eingesetzt werden, jedoch sollte für den Filmemacher die Verwendung einer 2D und einer Stereo-3D Überblendung klar differenziert werden. Der Effekt des sich auflösenden Raumes um den Zuschauer und das Erscheinen eines neu-

en Raumes könnte durch geringe Tiefen eventuell abgemildert werden. Zum Beispiel könnte kurz vor Beginn der Überblendung die Stereobasis für die Überblendung verringert werden, ehe nach der Überblendung die Stereobasis in der neuen Einstellung langsam wieder erhöht wird, bis der gewünschte 3D-Effekt erreicht ist.

Auch wenn es noch viele weitere Bildübergänge gibt, sind diese drei die mit Abstand am häufigsten verwendeten. Ein Großteil der Filmproduktionen kommt mit diesen drei Arten von Bildwechsellern aus, auch wenn sich in der 2D-Montage durch die umfassenden Möglichkeiten weitere Bildübergänge entwickelt haben, welche als Gestaltungsmittel oder Effekte Verwendung finden. Da die Verwendung von Überblendungen eher spartanisch eingesetzt wird, befassen sich die folgenden Untersuchungen ausschließlich mit dem harten Cut.

4.2 Blickanschluss

In der Filmgeschichte hat sich die Montage derart ausgeprägt, dass der Handlung und dem Schauspiel mit der Kamera gefolgt wird. Dadurch erhält der Blickanschluss einen hohen Stellenwert in der Montage [8]. Bei einer monoskopischen Filmproduktion waren Blickanschlüsse unkritischer als in einer Stereo-3D-Produktion, in welcher auch die Positionierung in der Tiefe eine logische Folgerung erhalten muss. Die Tiefen-Platzierung erfolgt sowohl am Set als auch in der Postproduktion. Jedoch werden die Möglichkeiten in der Postproduktion durch die gewählten stereoskopischen Eigenschaften bei der Aufnahme stark eingeschränkt. Daher ist eine Planung von elementarer Bedeutung.

4.3 Größenverhältnisse

Werden zwei Einstellungen, welche die gleichen Elemente zeigen, hintereinander montiert, so kann es dazu kommen, dass die Größenverhältnisse nicht zueinander passen. Durch die Möglichkeit, die Tiefenausdehnung zu gestalten, entsteht ein Einfluss auf die Größenwirkung durch die Tiefe. Wird eine Figur weit entfernt platziert und dabei zugleich mit einem hohen Tiefeneindruck versehen, wirkt sie vergrößert. Hierbei kann es darüber hinaus auch zu Verzerrungen von Geometrien kommen [9]. Bei Einstellungen, welche die gleichen Elemente enthalten, sollten deren Tiefenausdehnungen zueinander passen. Auch bei der Montage von Einstellungen, die vollständig unterschiedliche Elemente aufweisen, ist auf die Größeneindrücke durch die Tiefenwirkungen zu achten. Insbesondere bei dem Zuschauer bekannten Elementen, wie Autos oder Menschen, kann der Größeneindruck durch unterschiedliche Tiefen von der Handlung ablenken.

4.4 Perspektivkontrast

Bei der Montage in monoskopischen Produktionen soll ein deutlicher Perspektivkontrast zwischen den zu montierenden Einstellungen bestehen, um nicht den Eindruck eines holprigen Sprunges beim Zuschauer hervorzurufen. Dabei gilt ein deutlicher Unterschied der Einstellungsgrößen zweier Einstellungen gleichen Inhaltes oder ein Perspektivwechsel, durch welchen wesentliche Inhalte verändert werden, als Ziel. Inwieweit dies auch auf stereoskopische Produktionen zutrifft, sollte noch untersucht werden. Grundsätzlich gilt auch hier, dass Einstellungen mit sehr geringem Perspektivkontrast nicht erwünscht sind, da sonst unschöne Sprünge im Raum entstehen. Allerdings ist in bestehender Literatur oft davon die Rede, dass Probleme durch Tiefensprünge bei der Montage auftreten, wenn mit dem aus der 2D-Produktion üblichen Perspektivkontrast gearbeitet wird (vgl. Schumacher, 2009, S. 129). Jedoch erfordert ein hoher Perspektivkontrast noch nicht zwingend eine deutliche Änderung von Nah- und Fernpunkt. Hinzu kommt, dass auch durch Brennweite und Stereobasis eine vollständige Kompensation von relativen und in der Postproduktion auch absoluten Parallaxen möglich ist. Hierbei ist selbstverständlich auf die weiteren Nebeneffekte zu achten.

4.5 Schnittrhythmus

Der Schnittrhythmus vieler monoskopischer Produktionen ist sehr hoch. Oft haben Einstellungen eine Länge von einer Sekunde oder weniger. In aktuellen stereoskopischen Produktionen wird oft mit ähnlichem Schnittrhythmus geschnitten. Allerdings wird häufig von einem deutlich langsameren Schnittrhythmus berichtet, welcher bei einer Stereo-3D-Produktion notwendig werden soll [10]. Dabei soll hauptsächlich die notwendige Zeit zur erneuten Konvergierung der Augen nach dem Schnitt einen langsamen Schnittrhythmus notwendig machen. Ein neues Konvergieren auf eine andere Einstellung ist bei Vermeidung von Tiefensprüngen jedoch überflüssig, da keine Änderung des Konvergenzwinkels notwendig ist, solange der Zuschauer in beiden Einstellungen die gleiche Entfernung fixiert.

4.6 Zwischenschnitte und Einstellungsgrößen

Zwischenschnitte verfolgen das Ziel, Details oder Emotionen zu verdeutlichen [8]. In stereoskopische Produktionen, welche auf eine Kinoleinwand projiziert werden, sollten Detailaufnahmen jedoch vermieden werden. Beim natürlichen Sehen werden stereoskopische Tiefenreize, welche durch derart weit entfernte Objekte, wie sie bei der Leinwandbetrachtung erzeugt werden, viel dezenter angesprochen. Wenn nun also nahe Einstellungen gezeigt werden sollen, können diese nicht mit deren Größe und der damit implizierten Nähe wiedergegeben werden, ohne die Szenen mit großen Parallaxen im Raum schweben zu lassen. Bei einer Platzierung auf der Leinwand kann der Stereo-3D-Effekt kaum Verwendung finden, ohne dass der Gigantismus-Effekt

entsteht [9]. Somit bleiben Zwischenschnitte bei stereoskopischen Produktionen möglich, solange die Einstellungsgröße beachtet wird. Stereo-3D-Produktionen sollten weitwinklig produziert werden, wodurch die neue Bildsprache entsteht. Ein Beispiel hierfür sind die Filme »The Hobbit« von Peter Jackson, in welchen Unschärfen gänzlich vermieden werden und dem Zuschauer ein sehr weitläufiger Blick geboten wird. Weitwinklige durchgängig tiefenscharfe Aufnahmen ermöglichen das Betrachten des ganzen Raumes. Die Augen des Zuschauers werden weniger gelenkt, und er wird aufgefordert, durch den Raum zu wandern [11].

4.7 Match-Cut

Ein guter Match-Cut wirkt schon in einer 2D-Produktion beeindruckend und wird daher gerne verwendet. Bei einer Stereo-3D-Produktion kann ebenso ein Match-Cut verwendet werden, wobei hier der Effekt noch beeindruckender werden kann, da sich nun beim Schnitt ganze Räume verwandeln. Allerdings ist ein Match-Cut in einem Stereo-3D-Film deutlich schwieriger zu planen. Tiefenplatzierungen der Elemente, welche gematcht werden sollen, müssen exakt übereinstimmen. Ansonsten wird ein Match-Cut nicht beeindrucken, sondern lediglich für Irritationen sorgen. Daher müssen für einen Stereo-3D-Match-Cut Stereobasis, Brennweite und vor allem die Abstände zwischen Kamera und den einzelnen Objekten genauestens geplant werden. Bei einem monoskopischen Film reicht es, wenn hier die Platzierung im Bild, die Lichtsituation und die Größenverhältnisse zueinander passen. Schwierig wird es dabei auch, wenn zwischen zweierlei ähnlichen Geometrien geschnitten werden soll, welche allerdings deutlich unterschiedliche Größenverhältnisse aufweisen. Die gleiche Erscheinungsgröße im Bild und die identische Entfernung zum Zeitpunkt des Schnittes limitieren deren Größenunterschiede.

4.8 (Handlungs)-Achsenschema

Auch bei der Stereo-3D-Produktion sollte das (Handlungs)-Achsenschema berücksichtigt werden [8]. Die typische Over-Shoulder-Einstellung sollte jedoch vermieden werden, da dabei Irritationen durch Rahmenverletzungen entstehen können. Alternativ kann die Szene so weit hinten platziert werden, dass die Schulter bereits hinter der Scheinfensterlage erscheint, wodurch allerdings eine erhöhte Distanz zum Betrachter entsteht.

5. Tiefensprünge

Wie erwähnt, können Tiefensprünge bei der Montage auftreten. Der Begriff »Tiefensprung« bezeichnet den Unterschied zwischen Parallaxen zweier Einstellungen, welche hintereinander montiert werden. Einerseits kann eine Einstellung mit gleicher relativer Tiefe nach einem Schnitt folgen, wobei die Parallaxänderung durch die abso-

lute Platzierung entsteht, andererseits kann der Tiefensprung durch einen Unterschied in der relativen Parallaxe entstehen. Dabei entsteht auch automatisch ein absoluter Tiefensprung an Nah- und/oder Fernpunkt. Bei diesen Arten von Tiefensprüngen ist jeweils eine positive oder negative Parallaxänderung möglich. Hinzu kommt, dass diese Tiefensprünge auch kombiniert auftreten können.

Bei einer S3D-Aufnahme verhält sich die Parallaxe zur Brennweite an jeder Stelle im Bild proportional. Da sich hierbei der Bildausschnitt ändert, und üblicherweise eine bestimmte Einstellungsgröße gewünscht ist, kann durch den Abstand von Kamera und Motiv ein Ausgleich erfolgen. Dabei wird der Raum visuell, wie aus der 2D-Produktion bekannt, gestaucht oder gestreckt. Um die Parallaxen wieder an die Rahmenbedingungen der Betrachtungssituation anzupassen, sollte nun die Stereobasis angepasst werden. Es gilt, dass bei einer kurzen Brennweite und dem dadurch resultierenden nötigen Nahpunkt, zur Beibehaltung des Bildausschnittes, eine kleinere Stereobasis notwendig ist. Wird hingegen eine lange Brennweite verwendet und die Distanz zwischen Kamera und Nahpunkt vergrößert, kann eine größere Stereobasis gewählt werden. Dabei wirken sich diese Parameter jedoch nicht in identischer Form auf den Bildeindruck aus. Wird die gleiche Einstellungsgröße mit einer längeren Brennweite gedreht, so werden Objekte, die weit von der Kamera entfernt platziert sind, stärker vergrößert, als wenn die gleiche Einstellung mit einer kurzen Brennweite aufgenommen wird. Erfolgt nun durch Veränderung der Stereobasis ein Ausgleich der Parallaxen, wodurch zwei Einstellungen mit unterschiedlichen Brennweiten ohne Parallaxunterschiede an Nah- und Fernpunkt möglich werden, so kann es zu einem Tiefensprung kommen. Dieser besteht nun jedoch nicht in einem absoluten Parallaxunterschied am Nah- und/oder Fernpunkt, sondern in der relativen Tiefe von gleich tiefen Elementen, welche gleich groß im Bild erscheinen. Eine lange Brennweite lässt Objekte im Vordergrund flacher erscheinen, kurze Brennweiten geben Objekten im Vordergrund mehr Tiefe [9].

5.1 Horizontal Image Translation

Bei der Produktion von Stereo-3D-Aufnahmen mit parallelen Kameras ist meist die nachträgliche Positionierung der Szenen in der absoluten Tiefe in der Postproduktion notwendig. Wird keine Nachkonvergierung oder Horizontal Image Translation (H.I.T.) (s.a. Kapitel 5 in diesem Arbeitsbericht) vorgenommen, so erscheint die Aufnahme vollständig vor der Scheinfensterlage. Da die Szene später zumeist hinter der Scheinfensterlage platziert werden soll, ist durch H.I.T. eine Verschiebung nach hinten notwendig.

Dabei wird die Szene in ihrer stereoskopischen Tiefenwirkung gestreckt. Es ist also auch schon beim Dreh stereoskopischer Aufnahmen eine Planung der absoluten Tiefe notwendig, wenn in der Montage keine Tiefensprünge auftreten sollen.

5.2 Tiefenscript

Wie das Kapitel "Tiefensprünge" bisher verdeutlicht hat, ist der Tiefeneindruck von sehr vielen Parametern abhängig. Hierbei sind die meisten zwingend am Set festzulegen. Die nachträglichen Möglichkeiten in der Postproduktion sind sehr beschränkt. Um diese Vielzahl von Parametern kontrollieren zu können, ist eine Planung absolut notwendig. Durch die Verwendung eines Tiefenscriptes ist eine Planung der Tiefenverläufe möglich. Dabei wird aus dramaturgischer Sicht die Ausnutzung des Tiefenbudgets über die Handlung des Filmes geplant (vgl. Gardner, 2009).

Da im Tiefenscript die gewünschten Parallaxen von Nah- und Fernpunkt über den zeitlichen Verlauf des Filmes geplant und festgehalten werden, sind Sprünge schnell zu erkennen. Ein Tiefensprung, aufgrund unterschiedlicher Parallaxen am Nah- oder Fernpunkt ist hier als gerade vertikale Linie verzeichnet.

Allerdings können die weiteren Tiefensprünge, welche durch Parallaxunterschiede innerhalb der Szene auftreten können, hierbei nicht berücksichtigt werden. Diese können auftreten, wenn Brennweitenunterschiede und/oder deutliche Unterschiede in der nachträglichen Tiefenplatzierung durch H.I.T. vorgenommen werden. Hinzu kommt, dass durch unterschiedliche Stereobasen ebenfalls Tiefensprünge entstehen können. Soll hierbei jedoch die Parallaxe für Nah- und Fernpunkt keinen Sprung enthalten, muss zwangsweise eine Änderung der Brennweite und/oder der Abstände zur Kamera erfolgen.

Zur vollständigen Kontrolle über die beschriebenen Tiefensprünge reicht demnach ein Tiefenscript wie es von Brian Gardner beschrieben wird nicht aus. Zu der Planung von Parallaxen an Nah- und Fernpunkten wird auch eine Planung über Verläufe der Brennweite notwendig. Hinzu kommt, dass auch der notwendige H.I.T.-Wert geplant werden sollte. Hierbei ist zwar bereits in dem beschriebenen Tiefenscript eine Tiefenplatzierung enthalten, jedoch fehlt der wichtige H.I.T.-Wert, welcher dabei aus zwei Anteilen zusammengesetzt wird. Der erste Teil ist immer dann genau null, wenn der Fernpunkt der Szenen unendlich entspricht. Wird also eine Einstellung mit näherem Fernpunkt gedreht, steigt dieser Wert an und spiegelt die notwendige Verschiebung der zwei Bilder wieder, um den Fernpunkt auf der Scheinfensterlage zu platzieren. Der zweite Teil besteht dann schließlich in der Tiefenpositionierung. Nun ist noch ein zusätzlicher Wert notwendig, um eine volle Kontrolle zu erhalten. Hierbei stehen die drei Parameter Stereobasis, Nahpunkt und Fernpunkt zur Auswahl.

5.3 Relevanz

Der Begriff Tiefensprung ist somit noch weitläufig und beschreibt eine Vielzahl sprunghafter Änderungen stereoskopischer Tiefeneindrücke. Inwieweit welche Tiefensprünge für Irritationen sorgen, wurde noch nicht untersucht.

Für die weitere Untersuchung wird angenommen, dass irritierende Tiefensprünge aufgrund unterschiedlicher Parallaxen auftreten, welche eine Konvergenzänderung der Augen notwendig machen [12]. Anhaltspunkte hierfür sind in der Literatur zu finden: „Wird zwischen Bildern sehr unterschiedlicher → Tiefenausdehnung oder → Tiefenposition geschnitten, kann es zu einem Tiefensprung kommen. Der Zuschauer wird dann gezwungen, blitzschnell die Augenkonvergenz zu verändern. Zu → Visueller Überforderung kommt es dabei vor allem bei Tiefensprüngen aus der Ferne in die Nähe.“ [13].

6. Empirische Studie



Abbildung 1: Studiosituation

Durch eine empirische Studie sollen reproduzierbare Ergebnisse gesammelt werden, welche auf die Filmproduktion übertragbar sind. Die Forschung soll in Form von Probandentests durchgeführt werden, wobei Testsequenzen betrachtet und evaluiert werden sollen. Die Tiefensprünge werden durch die Variation der Stereobasis und/oder Nachkonvergenz erzeugt.

Die Probanden sollten, wie bei einer Fernsehsendung, einer logischen Handlung folgen. Die Einstellungsgröße soll konstant bleiben. Daher fiel die Wahl auf zwei Moderatoren, welche in einer klassischen Fernsehsituation positioniert wurden [4]. Um beim Wechsel zwischen den Einstellungen eine gute Montage zu ermöglichen, wurde der Schnitt durch eine Kopfdrehung der Moderatoren motiviert.

6.1 Produktion der Test-Items

Da in der Studie Irritationen aufgrund wechselnder stereoskopischer Tiefeninformationen untersucht werden, sollten andere mögliche Einflussgrößen minimiert werden. Zu diesem Zweck wurden die Items im Studio der Hochschule Furtwangen produziert, weshalb der Hintergrund des Senders HD-Campus-TV genutzt werden konnte. Dadurch wurde es möglich, einen zweidimensionalen Hintergrund zu erhalten, welcher später ausschließlich durch stereoskopische Tiefeninformationen für die Probanden in der Entfernung bestimmbar ist. Auch bei der Einstellungsgröße wurde zwischen den jeweiligen Einstellungen keine Veränderung vorgenommen, um eventuelle Abhängigkeiten zu vermeiden. Daher fiel die Wahl auf zwei Moderatoren, welche in einer klassischen Fernsehsituation positioniert wurden. Um für die Probanden einen abwechslungsreichen Inhalt zu gestalten, wurden von den Moderatoren unterschiedliche Spielfilme der letzten Jahre vorgestellt. Die Texte der Moderatoren haben dabei eine Länge von ca. 10-15 Sekunden, wodurch insgesamt eine Item-Länge von 20-30 Sekunden entstand. Um beim Wechsel zwischen den zwei Einstellungen eine gute Montage zu ermöglichen, wurde der Schnitt zwischen den zwei Einstellungen durch eine Kopfdrehung des ersten Moderators eingeleitet und durch das Weiterführen der Kopfdrehung des zweiten Moderators abgeschlossen.

6.1.1 Technische Umsetzung

Bei der Erstellung der Test-Items wurden zwei Canon EOS C300 Kameras verwendet, wodurch eine Synchronisierung mittels Genlock-Signals ermöglicht wurde. Die synchrone Aufzeichnung erfolgte über eine Capture-Karte von Blackmagic Design über HD-SDI. Dabei wurde zur Datenreduktion komprimiert mit einer Datenrate von ca. 58000 kb/s aufgezeichnet, wodurch der Speicherbedarf für das Rohmaterial ohne Outtakes von ca. 10 TB auf ca. 80 GB beschränkt werden konnte. Die Aufzeichnung erfolgte in einem AVI-Container mit einer mjpeg Komprimierung im Format yuvj422p 1920x1080@25p. Die Kameras wurden auf einem unmotorisierten Spiegelrig der Firma SwissRIG montiert und mit Hilfe von Einsteinium und der Software Uranium manuell justiert. Hierbei wurden die Kameras zunächst ohne Stereobasis auf 2 Pixel genau deckungsgleich ausgerichtet und danach die Stereobasis für den Dreh an der Präzisionsspindel eingestellt. Bei den verwendeten Objektiven handelte es sich um zwei Festbrennweiten mit einer Brennweite von 50 mm vom Hersteller Canon (Canon EF 50mm f/1.4 USM 2515A003). Hierbei diente die Betrachtungssituation nach ITU-R BT.500-13 als Grundlage (vgl. ITU, 2009, S. 4). Für den 50 Zoll großen Fernseher wurde ein Betrachtungsabstand von 3,35 m berechnet, welcher wiederum in die Kalkulation der Stereobasen einbezogen wurde. Alle Einstellungen wurden dabei mit fünf unterschiedlichen relativen Parallaxen aufgezeichnet. Dabei sollten die Einstellungen bei den Probanden im Test Querdisparitäten von 10, 20, 30, 40 und 50 Winkelminuten hervorrufen. Dadurch ergaben sich die fünf Stereobasen von rund 13, 27, 40, 53 und 67mm für die gewählte Einstellungsgröße. Somit sollten relative Par-

allaxen von ca. 13, 26, 39, 52 und 65 Pixeln in den Aufnahmen entstehen. Die Kameraposition wurde in die Mitte zwischen die zwei Moderatoren im Abstand von 2,25 m zu den Moderatoren gelegt. Der Studiohintergrund, welcher zugleich den Fernpunkt darstellt, hatte einen Abstand von 4,8 m zur Sensorebene der Kameras.

Um den Augen der Probanden vor der Betrachtung der unterschiedlichen Items eine Eingewöhnung zu ermöglichen, wurde zu den Items noch eine weitwinkligere Einstellung mit einer Dollyfahrt produziert. Durch die langsame Zufahrt auf die Moderatoren wird dabei der stereoskopische Tiefeneindruck allmählich gesteigert. Zudem wird durch diese Sequenz erreicht, dass die Probanden zunächst einen räumlichen Gesamteindruck über die Platzierung der beiden Moderatoren zueinander und zum Hintergrund erhalten.

Um die Dollyfahrt im Studio der Hochschule realisieren zu können, wurden aus Platzgründen weitwinkligere Objektive benötigt. Die Wahl fiel auf Zoom Objektive von Sigma, um die Brennweite optimal an die Studiogröße anpassen zu können. Die verwendeten 24-70mm F2.8 EX DG HSM verfügen laut Hersteller über einen Brennweitenbereich von 24 bis 70 mm und wurden auf 34 mm eingestellt. Dabei wurde die Brennweite allerdings nicht gemessen, sondern über die Schnittstelle vom EF-Mount des Objektivs ausgelesen. Um Rauschen durch Verstärkung der Sensorsignale zu vermeiden, und dennoch eine möglichst große Schärfentiefe zu erreichen, wurde die Blendenzahl 5 und ISO 1600 an beiden Kameras eingestellt. Der Fokus wurde auf die Gesichter der Moderatoren zum Ende der Fahrt eingestellt. Somit konnten rechnerisch maximale Unschärfekreise kleiner 8 Pixel während der ganzen Sequenz entstehen, wobei die Unschärfen während der Fahrt insgesamt kleiner wurden. Die Stereobasis wurde so gewählt, dass für die Probanden im Test keine größeren Querdisparitäten als 30 Winkelminuten auftreten. Berechnet wurde die Stereobasis auf gleiche Weise wie die Stereobasen der Items. Für die Fahrt wurden dann auf Grundlage der berechneten Stereobasis von 50 mm, die auftretenden Querdisparitäten für die Probanden am Anfang und am Ende der Einstellung berechnet und zusammen mit dem Verlauf der Unschärfen dokumentiert.

Die Ausleuchtung des Studios wurde mit vier Fresnel Scheinwerfern realisiert. Dabei wurde ein 200W HMI von ARRI mit flickerfreiem Ballast für die Hintergrundbeleuchtung verwendet. Dieser diente gleichzeitig auch als Aufheller der rechten Gesichtshälften und wurde mit Frostfolie versehen, um etwas weicherer Licht zu erhalten. Das Führungslicht wurde ebenfalls mit einem HMI-Scheinwerfer von ARRI realisiert, wobei hier ein etwas stärkerer Scheinwerfer mit 1200W Leistung Verwendung fand. Um harte Schatten im Bild zu vermeiden, wurde mit dem Scheinwerfer indirekt über einen großen Reflektor ausgeleuchtet. Durch die indirekte Ausleuchtung war allerdings weniger Licht in der Szene vorhanden, weshalb alle Aufnahmen mit ISO 1600 gedreht werden mussten, um die Blende so weit schließen zu können, dass bei den Items maximale Unschärfekreise auf dem Hintergrund von unter 8 Pixeln entstanden. Die beiden anderen Scheinwerfer dienten als Kanten zur Trennung der Moderatoren vom Hintergrund.

Hierfür wurden zwei ARRI Junior 650W Scheinwerfer mit 300W Brennern eingesetzt. Die Scheinwerfer wurden dabei nicht an die Farbtemperatur der HMI-Scheinwerfer angepasst.

6.1.2 Postproduktion

Neben der eigentlichen Montage wurde im Rahmen der Postproduktion die Tiefenpositionierung der einzelnen Einstellungen vorgenommen. Hierfür wurde die Software Final Cut Pro 7 von Apple eingesetzt. Dazu war allerdings eine Konvertierung des aufgezeichneten Rohmaterials notwendig. Die Konvertierung erledigte dabei das Programm ffnpeg, wodurch sowohl Konvertierung und Archivierung aller Rohdaten mittels einer Tool-Datei möglich waren. Der Audiostream wurde dabei ohne Konvertierung durchgereicht und im neuen MOV-Container zusammen mit dem Bildmaterial gespeichert. Durch den verwendeten ProRes-Codec von Apple wurde eine performante Postproduktion sichergestellt. Verwendet wurden dabei die ProRes 422 Proxy Variante mit 1920x1080@25p. Dabei ergab sich eine Datenrate von 38Mbit/s für den einfachen Video-Stream. Somit wurde das Material mit minimal geringerer Datenrate ausgegeben, als das aufgezeichnete Rohmaterial. Mit nur sehr geringem Qualitätsverlust konnte dann weiter gearbeitet werden. In der Postproduktion wurde die Auflösung dann allerdings halbiert, um die Auflösung des Fernsehers zu erreichen. Dieser kann ausschließlich 2 x 1920 x 540 im Stereomodus wiedergeben, da jede zweite Zeile für das Linke und die anderen Bildzeilen für die Wiedergabe des rechten Bildes genutzt wird. Um das Videomaterial zuspitzen zu können, wurde daher ein Over-Under-Stream aus Final Cut exportiert, welches bei einer Auflösung von 1920x1080 Pixeln das linke und rechte Bildmaterial enthält. Dabei werden die 540 ersten Zeilen für das linke Bildmaterial und die restlichen 540 Zeilen für das rechte Bildmaterial verwendet. Zudem wurden Farbunterschiede zwischen den zwei Kameras, welche durch den verwendeten Spiegel des Spiegelrigs entstanden sind, minimiert. Hierzu wurde eine globale 3-Wege-Farbkorrektur in Final Cut verwendet. Dabei wurden Farbunterschiede optisch und mittels RGB-Histogramm beurteilt.

Für jeden Probanden wurde eine eigene Testreihe aus Final Cut Pro mit zufälliger Reihenfolge exportiert. Die zufällige Reihenfolge wurde dabei mit dem Unix-Befehl `echo $[RANDOM%22+1]` festgelegt und dokumentiert. Die Testsequenzen aller Probanden wurden für die weitere Analyse archiviert.

6.2 Testaufbau

Für den Testaufbau wurde ein 50Zoll Fernseher mit Polfiltertechnik nach ITU-R BT.500-13 aufgestellt. Dabei ergab sich ein Betrachtungsabstand von 3,35 m als PVD [14]. Der Raum wurde für den Test vollständig abgedunkelt. Dunkle Wände verhinderten Reflexionen. Die Wand hinter dem Fernsehgerät wurde wie in ITU-R BT.500-13 definiert in Tageslicht (6500K) ausgeleuchtet.

Um den Probanden eine freie Äußerung über störende oder irritierende Elemente innerhalb der Test-Items zu ermöglichen, wurde die Think-Aloud-Methode angewandt. Dadurch wurde auch eine Evaluierung der Test-Items zu jedem Zeitpunkt gewährleistet. Die Aufzeichnung der Stimme erfolgte dabei über den gesamten Testverlauf. Die Aufzeichnung und Datenverwaltung erfolgte mit der Software Media Express vom Hersteller Blackmagic Design und der zugehörigen Capture-Karte des gleichen Herstellers. Über die HDMI-Schnittstelle angeschlossen, wurde zu diesem Zweck das integrierte Mikrofon einer HDR-CX550V Kamera von Sony eingesetzt. Dabei wurde die Kamera nicht nur zur Tonaufzeichnung, sondern auch zur Aufzeichnung der Augenbewegungen der Probanden verwendet. Die Kamera wurde dafür auf einem möglichst niedrigen Stativ vor den Probanden platziert, damit diese nicht zu sehr davon abgelenkt wurden.

Um die Augenbewegungen gut analysieren zu können, wurde zu diesem Zweck eine 3D-Polfilter-Brille mit Infrarot-LEDs ausgestattet, welche ohne Irritation der Probanden die Augen für die Aufzeichnung mit der Kamera beleuchtete. Um Kopfbewegungen der Probanden so weit minimieren zu können, dass sich die Augen während des gesamten Tests innerhalb des Aufzeichnungsbereichs der Kamera befanden, wurde eine Kinnstütze entwickelt. Durch die Verwendung der Kinnstütze wurde zudem ein konstanter Betrachtungsabstand über alle Probanden hinweg sichergestellt. Aus Hygienegründen wurden sowohl die Kinnstütze als auch die zwei 3D-Brillen, für den Sehtest und die Studie, nach jedem Testdurchlauf desinfiziert.

Zur Ausleuchtung der Augen, um die Augenbewegungen aufzeichnen zu können, wurden SMD-LEDs mit einer Leistung von bis zu 190mW pro LED in der Brille verbaut. Die Augen wurden dabei mit je einer LED vom Hersteller HARVATEK beleuchtet. Der verwendete Typ: HT-170IRPJ-XXXX strahlt dabei die maximale Leistung im Wellenlängenbereich von 850nm ab (vgl. HARVATEK, 2009, S. 4). Um die Ausleuchtung der Augen anpassen zu können, wurden für die zwei in der Brille verbauten LEDs separate Dimmer entwickelt, wodurch die Ströme beider LEDs mit zwei Potentiometern reguliert werden können. Als Spannungsquelle diente dabei ein Universal-Netzteil, welches bei einer Spannung von 6V mindestens mit einem Strom von 300mA belastet werden kann.

6.3 Expertentest

Vor dem eigentlichen Testdurchlauf sollte überprüft werden, ob sich Testaufbau und Items überhaupt für die Ziele der Untersuchung eignen. Zu diesem Zweck wurde ein Expertentest durch Professoren und Mitarbeiter der Hochschule Furtwangen University sowie der Technischen Universität Ilmenau durchgeführt. Dabei sind insbesondere die Proportionen der Moderatoren bei Items mit großer relativer Tiefe und bei Items mit großer absoluter Parallaxe irritierend aufgefallen. Hierbei wurde insbesondere die Länge der Arme und Nase als zu groß empfunden. Zudem wurde festgestellt, dass die sehr kurze Pause mit zwei Sekunden zwischen den verschiedenen Items dazu führen

kann, dass die Tiefenunterschiede zwischen den jeweiligen Items sofort auffallen. Dabei wurde auch festgestellt, dass die Reihenfolge, in welcher der Proband die Items sieht, einen großen Einfluss auf dessen Eindruck hat.

Die für die Wiedergabe verwendete Software VLC-Media-Player sorgte dafür, dass in den Pausen ein Pausenzeichen eingeblendet wurde. Zudem wurde festgestellt, dass der Timecode während der Wiedergabe nur auf dem Fernseher, nicht auf dem Computer angezeigt werden kann, und somit kein Überblick über den Testverlauf besteht.

6.4 Optimierung der Studie

Die Expertentests brachten einige Erkenntnisse, die für eine Optimierung der Studie sprachen. Diese wurden vor dem ersten Testdurchlauf mit Probanden durchgeführt. Somit sollte eine optimale Testsituation sichergestellt werden. Um Störungen oder Irritationen zu vermeiden, welche durch Verwendung der großen Stereobasen im Zusammenhang mit der extremen Tiefenpositionierung entstanden sind, wurde das Bildmaterial gecroppt. Dadurch wurde erreicht, dass die Arme, die zuvor den Nahpunkt im Bild darstellten, keine Irritationen bei den Probanden durch ihre ungewohnt hohe relative Tiefe auslösen. Dabei wurde das Seitenverhältnis verändert und die Auflösung des Bildmaterials beibehalten, weshalb die Darstellung mit schwarzen Balken über und unterhalb des Bildes gewählt wurde.

Um den Zeitaufwand der Probandentests für den Probanden auf 30-45 Minuten zu verringern, wurden einige Items, welche nur geringe Tiefensprünge enthielten, aus der Studie ausgeschlossen. Alle extremen Tiefensprünge wurden beibehalten.

6.5 Testablauf

Nachdem der Testdurchlauf optimiert war, wurde im Internet ein Kalender erstellt, in welchem sich Interessenten für einen Termin eintragen konnten. Dadurch konnte eine Terminierung der einzelnen Probanden erfolgen. Um die Vorbereitung des nächsten Testes nach einem Testdurchlauf zu ermöglichen, konnten sich die Probanden im Abstand von einer Stunde eintragen. Da es trotz guter Planung zu Wartezeiten kommen kann, indem ein Proband mehr Zeit in Anspruch nimmt als geplant, wurde ein Wartebereich eingerichtet. Um die Teilnahme für die Probanden möglichst angenehm zu gestalten, wurden ihnen Getränke und Schokoriegel angeboten. Im Wartebereich standen zudem Zeitschriften zur Verfügung.

Um Störungen während den laufenden Probandentests zu vermeiden, wurde die Labortür mit einem Schild versehen, womit die Probanden aufgefordert wurden, im Wartebereich Platz zu nehmen. Durch eine rote Lampe an der Labortür wurden Außenstehende über die laufenden Tests informiert.

Nach dem Eintreten der Probanden wurden diese zunächst freundlich begrüßt. Im Anschluss wurde ihnen Dank ausgesprochen für die Zeit, die sie in die Studie inves-

tieren. Im Anschluss wurden sie über den Inhalt und den Ablauf der Studie informiert. Um bei allen Probanden einen einheitlichen Kenntnisstand über die Studie sicherzustellen, wurden alle Probanden dabei identisch begrüßt, und sie wurden auch mit dem identischen Text über die Studie informiert.

Nun wurden die Probanden gebeten, die bereitgelegten Unterlagen (Seite 1-4 der Probandenunterlagen) durchzulesen und auszufüllen. Dabei war auf der ersten Seite nochmals der Versuchsablauf beschrieben, um den Probanden einen Überblick zu bieten. Daraufhin musste eine Einverständniserklärung ausgefüllt werden, um sicherstellen zu können, dass die Probanden mit der Datenerfassung, Archivierung und Veröffentlichung einverstanden sind. Durch den Mediennutzungsfragebogen wurde die Mediennutzung der Probanden insbesondere auf den stereoskopischen Medienkonsum bezogen festgehalten. Ein Befindlichkeitsfragebogen wurde ausgefüllt, um mithilfe des zweiten Befindlichkeitsfragebogens, diesen am Ende der Studie, vergleichen zu können. Im Weiteren wurde ein Stereo-3D-Sehtest durchgeführt, und die Ergebnisse vom Versuchsleiter auf der Seite 5 der Probandenunterlagen festgehalten. Da der Fernseher in einer Entfernung von 3,35 m zum Probanden aufgestellt wurde, konnte dieser mit einem Pixel Tiefenunterschiede von 35 Winkelsekunden darstellen. Daher entspricht dies dem Sehtest zwischen Stufe 7 (40arcsec) und 8 (30arcsec). Allerdings durften auch Probanden mit schlechteren Sehergebnissen an der Studie teilnehmen, da alle Items deutlich größere Tiefenunterschiede enthielten, als Stufe eins im Sehtest (400arcsec). Zudem wurden die Probanden nach einer Farbsehschwäche und einer Sehhilfe gefragt. Dies wurde ebenfalls dokumentiert.

Für die eigentlichen Tests wurden die Probanden gebeten, die beleuchtete 3D-Brille aufzusetzen, und den Kopf in der Kinnstütze zu platzieren. Beides wurde im Voraus bereits desinfiziert. Dabei wurde dem Probanden geholfen, die Kinnstütze und den Stuhl bequem einzustellen. Nachdem die Kamera auf die Augen justiert wurde, wurde die Deckenbeleuchtung ausgeschaltet.

Bevor der Test starten konnte, wurde die Bildseparation des Polfiltersystems überprüft. Hierzu zeigte der Fernseher für das linke Auge ein blaues L und für das rechte Auge ein blaues R (RGB = 0, 0, 255) auf einem 50% grauen Untergrund. Das L und das R waren dabei an unterschiedlichen Stellen des Bildschirms platziert. Die Probanden wurden gebeten, dabei jeweils ein Auge zuzuhalten. Nun sollten sie nur noch jeweils den entsprechenden Buchstaben sehen können. Bei diesem Test kam es zu keinen Geisterbildern.

Für eine Kalibrierungssequenz, bei welcher sich ein roter Punkt über den schwarzen Fernseher bewegt, wurde der Proband gebeten, dem roten Punkt mit den Augen zu folgen. Dadurch soll ein späteres Gaze-Tracking bei den Tests möglich werden. Während der Kalibrierungssequenz wurden die Augen über die Kamera aufgenommen, und über die HDMI-Schnittstelle unkomprimiert mit 1920x1080@50p aufgezeichnet. Diese Daten werden von dem Forschungsprojekt BinoQ3D an der Hochschule Furtwangen verwaltet.

Um das Datenaufkommen zu reduzieren, erfolgte die weitere Aufzeichnung der Augen und der Stimme der Probanden komprimiert. Dabei wurde der gleiche Codec wie bei der Aufzeichnung der Items verwendet. Zur Eingewöhnung an die S3D-Betrachtung wurde das Einführungsvideo abgespielt. Dabei wurden die Probanden gebeten, sich zu äußern, sobald der Stereo-3D-Effekt für sie wahrnehmbar ist. Durch die Aufzeichnung der Stimme kann somit im Nachhinein eine Auswertung der Zeiten erfolgen.

Vor dem ersten Test-Item wurden die Probanden nochmals aufgefordert, jegliche Irritationen oder Störungen bei der Betrachtung mitzuteilen. Daraufhin wurde die erste Testsequenz gestartet. Wurde vom Probanden keine Irritation oder Störung benannt, wurde in der mindestens zehn Sekunden dauernden Pause nach dem Video danach gefragt. Wurde vom Probanden etwas festgestellt, so wurde gefragt, ob dies irritierend oder störend wirkte. In jedem Fall wurde gefragt, ob es noch weitere Irritationen gab. Die Aufzeichnung erfolgte über den Testverlauf unterbrechungsfrei.

Nachdem das letzte Item abgespielt wurde und nach Irritationen gefragt wurde, wurde der Fernseher schwarz und die Aufzeichnung wurde beendet. Nun konnte der Proband die Kinnstütze verlassen und die Polfilterbrille absetzen. Die Raumbeleuchtung wurde wieder eingeschaltet.

Die Probandenunterlagen Seite sechs und sieben wurden während des Tests bereitgelegt und vom Probanden direkt im Anschluss ausgefüllt. Der Befindlichkeitsfragebogen 2 ermöglicht dabei spätere Vergleiche mit dem ersten Befindlichkeitsfragebogen. Erst im folgenden Abschlussfragebogen wurde explizit nach Irritationen oder Störungen durch Schnitte gefragt. Bis zu diesem Zeitpunkt wurden die Begriffe Schnitt und Montage nicht erwähnt.

Zum Schluss wurden die Probanden gefragt, ob sie noch offene Fragen zur Studie haben. Nachdem sich nochmals für die Teilnahme bedankt wurde, wurden die Probanden verabschiedet. Die ausgefüllten Probandenunterlagen wurden archiviert und sind Bestandteil dieser Arbeit.

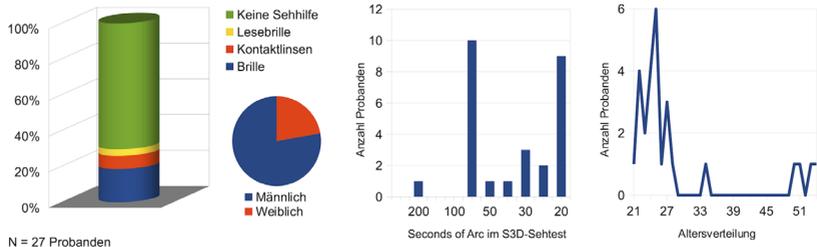


Abbildung 2: Eigenschaften der Probanden

7. Ergebnisse

Die Anmerkungen der Probanden mussten für die Auswertung kategorisiert werden. Die Kategorien entstanden auf Basis der Probandenaussagen. Dabei fiel auf, dass Schnitte nur selten als irritierend wahrgenommen wurden. Den 27 Probanden wurden jeweils 22 Items gezeigt. Drei enthielten keinen Tiefensprung, sechs Items enthielten einen maximalen Tiefensprung mit 52 Pixeln Parallaxunterschied. In Abbildung 3 sind die Parallaxen der 22 Items dargestellt. In den gesamten Tests wurde lediglich 18-mal der Schnitt als irritierend angegeben. Im Vergleich dazu wurde 422-mal keine Irritation festgestellt.

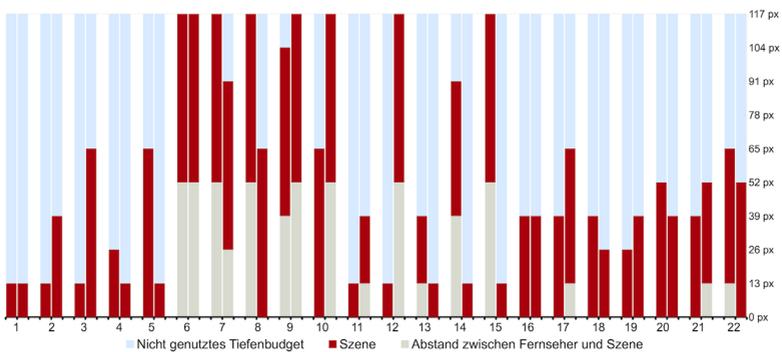


Abbildung 3: Parallaxen in den 22 Test-Items

Einige Irritationen, welche bei den Probanden entstanden sind, können durch das nicht perfekte Polfiltersystem des Fernsehers erklärt werden. Ghosting wurde 45-mal bemängelt. Einige Probanden empfanden Unschärfen als störend. Durch den üblichen 180° Shutterwinkel wurden die Aufnahmen mit einer Belichtungszeit von $1/50$ Sekunde aufgezeichnet [15]. Dabei sind Bewegungsunschärfen durch Lippenbewegungen und Kopfdrehungen aufgetreten. Auch durch die geringe Schärfentiefe kam es zu Irritationen. Insgesamt wurden 44-mal Unschärfen als irritierend genannt.

Die meisten Irritationen traten durch die Tiefenpositionierung oder Tiefenausdehnung auf. Um die großen Tiefensprünge zu ermöglichen, mussten die absoluten Parallaxen so groß werden, dass neben Ghosting noch weitere Irritationen auftraten. So wurde häufig die Positionierung des Vorder- oder Hintergrundes als störend oder irritierend bemängelt. Neben großen Anstrengungen beim Fusionieren der Bilder aufgrund großer Parallaxen empfanden einige Probanden auch eine nahe Platzierung an der Scheinfensterebene als störend. Hinzu kamen störende Effekte durch die enorme Nachkonvergenz. Durch H.I.T. wurden zur Erreichung der Tiefensprünge einige Einstellungen so platziert, dass Gigantismus oder übermäßige Proportionen von den

Probanden bemerkt wurden. Die Haare des Moderators wurden durch das Skalieren nach dem H.I.T. teilweise leicht angeschnitten.

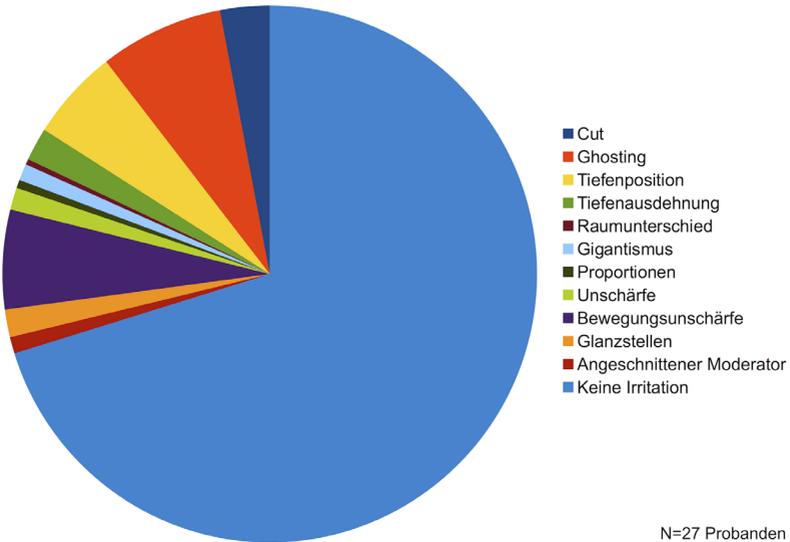


Abbildung 4: Irritationen im Testverlauf

Die Antworten, dass keine Irritationen aufgetreten sind, überwiegen deutlich. Auffällig ist, dass bei Items mit extremen Parallaxunterschieden nur von wenigen Probanden eine Irritation durch den Schnitt bemängelt wurde. Zumeist entstanden Irritationen viel eher durch die räumliche Situation an sich, indem das Tiefenbudget oder die absolute Parallaxe zu groß gewählt wurde. Die Probanden stellten zumeist beim Betrachten fest, dass sich die Tiefenposition verändert. Auf die Frage, ob dies für sie irritierend oder störend wirkte, kam meistens die Antwort »nein«. Teilweise fanden die Probanden den Unterschied der räumlichen Situation zwischen den zwei Einstellungen störend. Jedoch wurde hier nicht der Schnitt an sich als irritierend empfunden. Die Probanden gaben an, dass sie es ungewohnt finden, dass die Platzierung unterschiedlich erfolgt. Hier war auch die Rede von einer Hierarchie, welche durch die Platzierung zwischen den Moderatoren entsteht.

Auffällig waren zwei Items, bei welchen eine Anhäufung von als irritierend wahrgenommenen Schnitten verzeichnet wurde. Bei einem Item waren fünf und bei dem anderen Item vier Probanden durch den Schnitt irritiert. Trotz der geringen Zahl lässt sich eine klare Tendenz gegenüber den anderen Items erkennen. Bei beiden Einstellungen handelte es sich um den maximalen Tiefensprung mit einem Parallaxunterschied von 52 Pixeln. Beide Testsequenzen haben eine Gemeinsamkeit. Nachdem der Moderator direkt an der Scheinfensterebene platziert ist, erscheint die Moderatorin 52

Pixel tiefer platziert. Der Unterschied der Items besteht darin, dass diese Positionsänderung bei einem Item durch eine Änderung der Stereobasis und damit der Tiefenausdehnung und bei dem anderen Item durch H.I.T. erfolgt.

Durch die Blickführung, welche angewandt wurde, kam es zu Bewegungsunschärfen durch die Kopfdrehung, welche für einen motivierten Schnitt sorgte. Die unterschiedlichen Tiefenpositionierungen der zwei Moderatoren sorgten dafür, dass die Blickanschlüsse teilweise nicht stimmen. Auch wenn die Kopfdrehung für eine Blickführung sorgt, so erscheint dann der zweite Moderator in einer anderen Tiefe. Der Blick wird zwar in die richtige Richtung geleitet, nicht jedoch zur richtigen Tiefenposition.

Aufgrund der geringfügig wahrgenommenen Irritationen durch Schnitte könnte davon ausgegangen werden, dass die Relevanz von Tiefensprüngen in dieser Szenearrangierung eher gering ist. Insbesondere im Vergleich zu anderen Irritationen schienen Irritationen durch Schnitte eher eine untergeordnete Relevanz zu haben, obwohl die Items hierfür optimiert wurden. Dabei sollte berücksichtigt werden, dass hier explizit die Wirkung von Parallaxunterschieden an Nah- und Fernpunkten untersucht wurde. Daher wurden die einzelnen Einstellungen in der gleichen Studioumgebung in einer vergleichbaren Situation erstellt. Durch den Entfall räumlicher Elemente sollten zwei Ebenen geschaffen werden, auf welche die Probanden konvergieren können. Somit wurde die Studie dahingehend optimiert, dass eine Konvergenzänderung der Probandenaugen beim Schnitt erforderlich ist. Der zweidimensionale Hintergrund verhinderte dabei verzerrte Tiefeneindrücke durch unterschiedliche Tiefenausdehnungen. Jedoch hatten die Moderatoren im Vordergrund Tiefe, wodurch sich hier die Umpositionierung mithilfe der Stereobasis auch in der Dreidimensionalität der Moderatoren bemerkbar machte.

Es ist möglich, dass Tiefensprünge nicht dadurch irritieren, dass die Augen neu konvergieren müssen, sondern viel eher dadurch, dass sich räumliche Elemente in ihrem Tiefeneindruck verändern. In der Studie wurden räumlicher Elemente, welche unterschiedliche Tiefeneindrücke hervorrufen, weitestgehend vermieden. Dadurch und durch die gleichbleibende Studiosituation, wurde zudem das visuelle System nur gering gefordert, wodurch möglicherweise Tiefensprünge erst bei sehr großen Parallaxunterschieden für Irritationen sorgten.

Hier wurden nur wenige Möglichkeiten, einen Tiefensprung zu erzeugen, angewendet. Dabei wurden die Tiefensprünge durch Stereobasisänderungen und Nachkonvergenz erzeugt. Möglicherweise wirken Tiefensprünge durch Nah- oder Fernpunktänderungen sowie Brennweitenänderungen deutlich irritierender. Dies soll in einer Folgestudie weiter untersucht werden.

8. Fazit

Die Montage soll auch im stereoskopischen Film unsichtbar bleiben – das wurde in 2D-Filmen durch verschiedene Montagetechniken bereits erreicht. Bisher wurde auf-

grund fehlender Erfahrungen häufig auf die verschiedenen Montagetechniken, welche sich im Laufe der 2D-Filmmontage entwickelt haben, zurückgegriffen. Die dabei entstehende Wirkung auf den Zuschauer wurde lange nicht erforscht.

Die technischen Anforderungen, welche für das Schneiden und Montieren stereoskopischer Aufnahmen notwendig sind, sind einfach zu erfüllen. Viel schwieriger ist es, den gestalterischen Anforderungen gerecht zu werden. Ein Begriff, der in diesem Zusammenhang oft genannt wird, ist der Tiefensprung.

Bisher wurde der Begriff Tiefensprung nicht ausreichend definiert, weshalb dies in dieser Arbeit im Zusammenhang mit der Untersuchung der Anwendbarkeit von 2D-Montagetechniken auf die Stereo-3D-Produktion erfolgen musste. Es gibt verschiedene Arten von Tiefensprüngen, für welche es unterschiedliche Ursachen geben kann. Für die Studie wurden zwei Ursachen ausgewählt: Tiefensprünge mit absoluten Parallaxunterschieden an Nah- oder Fernpunkt sowie Tiefensprünge mit gleichen Parallaxunterschieden an Nah- und Fernpunkt. Diese lassen sich auch durch das bereits bekannte Tiefenskript planen und kontrollieren. Dadurch kann auch eine Übertragung der Ergebnisse auf die reale Filmproduktion erfolgen.

Mit der vorliegenden Arbeit wurden erste Untersuchungen angestellt, um diese Wirkung von Tiefensprüngen auf Zuschauer zu erforschen. Besonderen Wert wurde darauf gelegt, alle erhobenen Daten für eine weiterführende Forschung zu archivieren. Die Testsequenzen wurden unter Studiobedingungen produziert, wodurch eine Kontrolle und Dokumentation aller relevanten Parameter erfolgen konnte.

Untersucht wurde in der Studie ausschließlich der harte Schnitt. Der harte Schnitt wird sowohl bei Fernseh- als auch bei Kinoproduktionen mit Abstand am häufigsten eingesetzt. Die aus der 2D-Montage bekannte Blickführung, welche bei jeder Filmproduktion eingesetzt wird, wurde berücksichtigt. Somit haben die Ergebnisse eine hohe Relevanz für die Montage von Stereo-3D-Realfilmen.

Bei der Studie ergaben sich keine eindeutigen Ergebnisse. Jedoch waren klare Tendenzen erkennbar. Diese zeigten, dass Irritationen insbesondere bei Schnitten auftreten können, wenn die zweite Einstellung weiter entfernt positioniert wurde, als die erste Einstellung. Diese Erkenntnis kann schon eine Hilfestellung für Kinematografen, Stereografen und Cutter darstellen. Die Nutzung eines Tiefenskripts kann in Verbindung mit den Ergebnissen helfen, Irritationen durch Tiefensprünge zu vermeiden. In einem Tiefenskript kann der dramaturgische Verlauf des Filmes für die Tiefenausnutzung geplant werden. Dabei werden auch die Parallaxen an Nah- und Fernpunkten geplant. Änderungen der Tiefenplatzierung des Nahpunktes waren besonders dann relevant, wenn durch einen Schnitt eine Verschiebung des Nahpunktes in den Hintergrund erfolgte. Auch wenn es dafür noch unterschiedliche Gründe geben kann, welche weiter untersucht werden sollten, so ist es dennoch wahrscheinlich, dass derartige Tiefensprünge für ähnliche Auswirkungen sorgen. Da in der Studie eine Szene verwendet wurde, welche besonders wenig räumliche Elemente enthält, ist es möglich, dass Irritationen bereits durch schwächere Parallaxunterschiede entstehen können.

Dies sollte für die Stereo-3D-Produktion berücksichtigt werden. Ebenso wurde in allen getesteten Videos die gleiche Studioumgebung dargestellt, wodurch das visuelle System des Menschen nur gering gefordert wurde. Dabei sorgte die wiederholte Betrachtung der immer gleichen Studioumgebung möglicherweise für einen räumlichen Lerneffekt, wodurch das Gehirn zusätzlich entlastet werden konnte.

Die weiteren Arten von Tiefensprüngen sowie die weiteren Ursachen für deren Entstehung konnten an dieser Stelle nicht behandelt werden. Für weitere Tests sollten insbesondere die nichtbehandelten Ursachen für Tiefensprünge weiter erforscht werden. Hierbei sollten insbesondere Tiefensprünge durch Brennweitenänderungen sowie Tiefensprünge durch Änderung von Nah- und Fernpunkten hervorgerufen werden. Allerdings sind aller Voraussicht nach alleine hierfür zweierlei Studien notwendig. Eventuell kann auch eine deutliche Einschränkung der Test-Items erfolgen. Hierfür können die Testergebnisse dieser Studie eine Grundlage darstellen.

Die Raumgestaltung, welche in den Test-Items für eine nur geringe Ausnutzung der räumlichen Tiefenreize sorgte, könnte hingegen auch von größerer Bedeutung sein. Durch die Studioumgebung waren auch monokulare Tiefenhinweise ungewöhnlich schwach vertreten. Inwieweit gleiche Tiefensprünge durch eine detailreiche Raumgestaltung mit vielen monokularen Tiefenhinweisen und einer umfassenden Ansprache der patenten Stereopsis, andere Ergebnisse liefern, könnte ebenfalls eine weitere Studie klären. Dabei sollte allerdings berücksichtigt werden, dass die bereits in dieser Studie festgestellten Nebeneffekte wie Gigantismus oder ungewöhnliche Proportionen deutlich stärker auftreten werden.

Literaturverzeichnis:

- [1] Ast, Michaela (2002): Geschichte der narrativen Filmmontage. Theoretische Grundlagen und ausgewählte Beispiele. Marburg: Tectum Verlag.
- [2] Beller, Hans (2002): Handbuch der Filmmontage. 4. Aufl. Band 5. herausgegeben von Prof. Hans Beller. München: TR-Verlagsunion GmbH.
- [3] Ebert, Jürgen (1979): Montage Editing Schnitt. In: Filmkritik Nr. 276. München: Filmkritiker-Kooperative.
- [4] Frank, Anja (2009): Gestalterische Aspekte einer stereoskopischen HDProduktion. In: Hottong, Nikolaus/Lesik, Dominik (Hg.): Stereoskope HDProduktion. Grundlagen, Konzepte, Testergebnisse. S. 105-121
- [5] Hoffmann, Albrecht (2002): Das Stereoskop. Technikgeschichte, Modelle und Rekonstruktionen. 2. Aufl. 2002. München: Deutsches Museum.
- [6] Hofmeyer, Frank: Grundlagen der visuellen Wahrnehmung. In: Hottong, Nikolaus/Lesik, Dominik (Hg.): Stereoskope HD-Produktion. Grundlagen, Konzepte, Testergebnisse. 2009 (zit 2009a) S. 16-55
- [7] Hofmeyer, Frank: Grundlagen der stereoskopischen Theorie. In: Hottong, Nikolaus/Lesik, Dominik (Hg.): Stereoskope HD-Produktion. Grundlagen, Konzepte, Testergebnisse. 2009 (zit 2009a) S. 57-91
- [8] Katz, Steven D. (2004): SHOT BY SHOT, Die richtige Einstellung, Zur Bildsprache des Films. Aus dem Amerikanischen von Harald Utecht. 5. Aufl. Frankfurt: Zweitausendeins.
- [9] Kuhn, Gerhard (1999): Stereo-Fotografie und Raumbild-Projektion. Gilching: vfv.
- [10] Mendiburu, Bernard (2009): 3D movie making. Stereoscopic digital cinema from script to screen. Amsterdam, Boston: Focal Press/Elsevier.
- [11] Reisz, Karel/Millar, Gavin (1991): The Technique of Film Editing. Second Edition. Surrey: British Academy of Film and Television Arts.
- [12] Rollmann, W. (1853): Zwei neue stereoskopischen Methoden. In: Annalen der Physik und Chemie.

Herausgegeben von J. C. Proggendorff. Leipzig: Verlag von Johann Ambrosius Barth.

- [13] Sammons, Eddie (1992): The World of 3D Movies. Version 1B. Delphi.
- [14] Schumacher, Erik (2009): Planung einer stereoskopischen HD-Produktion. In: Hottong, Nikolaus/Lesik, Dominik (Hg.): Stereoskope HD-Produktion. Grundlagen, Konzepte, Testergebnisse. S. 123-143
- [15] Steber, Josef-Anton. Nowara, Thomas. Bonse, Thomas (2007): Bewegung in Video und Film. 1. Aufl. Berlin: Fachverlag Schiele & Schön GmbH.
- [16] Tauer, Holger (2010): Stereo 3D. Grundlagen, Technik und Bildgestaltung. 1. Aufl. 1 Band. Berlin: Schiele & Schön.

Journalistische Quellen

- [17] Bolliger, Matthias: Stereo 3D. In Film & TV Kameramann, Ausgabe 1 & 2 (2011)

Internetquellen

- [18] FFA - Filmförderungsanstalt (2012): Der Kinobesucher 2011 - Strukturen und Entwicklungen auf Basis des GfK Panels,
URL: http://www.ffa.de/downloads/publikationen/kinobesucher_2011.pdf (Zugriff am 06.02.2014)
- [19] FFA – Filmförderungsanstalt (2013): Aktuelle FFA Pressemitteilungen - Detail,
URL: http://www.ffa.de/index.php?page=presse_detail&ws=1055 (Zugriff am 06.02.2014)
- [20] Gardner, Brian (2009): Perception and The Art of 3D Storytelling,
URL: http://library.creativecow.net/gardner_brian/magazine_3d_storytelling/1 (Zugriff am 26.02.2014)
- [21] Harvatek (2009): Harvatek Surface Mount LED Data Sheet HT-170IR-PJXXXX. Version 1.0,
URL: http://www.produktinfo.conrad.com/datenblaetter/175000-199999/181697-da-01-en-SMD_IR_LED_805_HT_170IRPJ.pdf (Zugriff am 26.02.2014)

- [22] ITU-R, Radiocommunication Sector of ITU (2009): Methodology for the subjective assessment of the quality of television pictures,
URL: http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-RECBT.500-12-200909-S!!PDF-E.pdf (Zugriff am 26.02.2014)
- [23] Moser, Leo (2010): 3-D is Back again!. In: ray Filmmagazin.
URL: <http://www.ray-magazin.at/magazin/2010/03/3-d-is-back-again> (Zugriff am 17.02.2014)
- [24] Motion Picture Association of America (2012): Theatrical Market Statistics,
URL: <http://www.mpa.org/Resources/3037b7a4-58a2-4109-8012-58fca3abdf1b.pdf> (Zugriff am 26.02.2014)
- [25] Nielsen (2012): Kino Box Office - Insights 2011,
URL: http://www.weischer.net/fileadmin/user_upload/Werbeweischer/Dateien_ww/Nielsen_Kino_Box_Office_2011.pdf (Zugriff am 26.02.2014)
- [26] NPD Group Inc. - Display Search (2012): Shipments of 3D LCD TV Panels Reach 21M in 2011, Reaching 10% Penetration,
URL: http://www.displaysearch.com/pdf/120307_shipments_of_3d_lcd_tv_panels_reach_21m_in_2011.pdf (Zugriff am 26.02.2014)

Teil 4

Unschärfe in Stereo-3D-Produktionen

Benjamin Hauser, Nadine Leonhardt, Robert Lichtl, Nikolaus Hottong, et al.

Abbildungsverzeichnis

Abbildung 1: Objekt in Anaglyphen-Darstellung mit gaußscher Unschärfe von 2,9.....	175
Abbildung 2: Random-Dot-Diagramm scharf/unscharf.....	176
Abbildung 3: Versuchsaufbau zum Alignment der Kameras	177
Abbildung 4: Screenshot aus den Realfilmaufnahmen (Zählweise der Objekte aufsteigend von vorne nach hinten)	178
Abbildung 5: Ermittelte Tiefenwahrnehmungsgrenze in Abhängigkeit der Bildunschärfe	181
Abbildung 6: Tiefenwahrnehmungsgrenze in RDS	182

1. Einleitung

Weitverbreitet ist der Irrglaube, Schärfe wäre dasselbe wie die örtliche Auflösung. Ein einfacher Test belehrt uns aber schnell des Besseren: Defokussiert man eine 4K-Kamera, so hat das resultierende Bild immer noch über 8 Mio. Pixel – scharf ist es aber nicht.

Mit Schärfe wird korrekterweise ein Reiz bezeichnet, welcher in unserem visuellen System ausgelöst wird, wenn wir feine (Objekt-) Kanten noch in hohen Kontraststufen wahrnehmen können. Wird uns eine Kante oder eine Textur mit einem Kontrastgradienten $<30\%$ präsentiert, nehmen wir diese nur noch als weichen Übergang wahr. Wir nennen den dann eintretenden Effekt umgangssprachlich „Unschärfe“.

Ob eine Kamera eine Szene scharf abbildet oder ein Monitor ein Bild scharf darstellt, kann also nur in textur- oder kantenbehafteten Motiven ermittelt werden. Auch die Messung bzw. die Wahrnehmung von Unschärfe ist folglich an Texturen im Bildmotiv gebunden.¹

Interessanterweise wird die menschliche Tiefenwahrnehmung ebenfalls stark über Objekte und Texturen adressiert. In welcher Wechselwirkung stehen dann beide Effekte? Diese Thematik soll im Folgenden genauer untersucht werden.

2. Unschärfe im Film

In der Wahrnehmung unserer Realität tritt Unschärfe entweder:

- im Randbereich unseres Gesichtsfelds auf (und kann bei Bedarf durch augen-/kopfbasierte Verschiebung des scharfen Blickfeldes kompensiert werden), oder
- bei Objekten im Blickfeld auf, welche außerhalb der aktuellen Brennweite unserer Augenlinsen liegen (was bei Bedarf durch Akkommodation ausgeglichen werden kann), oder

¹ Wie es z.B. schon das einfache Testbild „Siemensstern“ demonstriert.

- in sehr kleinen/großen - nicht akkommodierbaren - Entfernungen auf. Im letzteren Fall zerstäubt zusätzlich die atmosphärische Perspektive alle Kontraste und verringert dadurch zusätzlich die monokulare Tiefenwirkung von potenziellen Texturdichtegradienten. Nur Verdeckung, relative Größe und eventuelle perspektivische Konvergenzen können in diesem Fall noch eindeutig zur Tiefenauswertung herangezogen werden².

Im Umkehrschluss: Wenn uns ein Objekt nahe ist (im sog. „Persönlichen Raum“ bzw. im „Handlungsraum“ bis maximal 50 m Entfernung), sind wir - bei gutem Licht, gesundem Visus und ausgeprägter Stereopsis - folglich immer in der Lage es scharf, tiefenverortet und (bis ca. 10 m Entfernung) mit eindeutigem Volumen wahrzunehmen.³

In planen Filmbildern können verschiedene Arten von Unschärfen auftreten. Technisch-verursachte Unschärfen können durch Objektivfehler, einen Fehlfokus, Trägheit von optischen Schaltelementen, überlangen Verarbeitungszeiten, Kompression etc. entstehen. Sie können sowohl global (über das gesamte Bild) oder lokal wirken.

Zusätzlich gibt es noch die Effekt-Unschärfen, welche aus einer gestalterischen Intention heraus bewusst erzeugt sind und die dramaturgische Wirkung eines Films deutlich erweitern können. Diese Effekt-Unschärfen lassen sich wiederum in zwei Kategorien unterteilen: Zum einen kann in bewegungsreichen Szenen über das gewählte Verhältnis Belichtungszeit/Framerate die *Bewegungsunschärfe* bewusst verstärkt werden (Motionblur). Zum anderen wird die, über Blende, Brennweite und Abstand manipulierbare, *Tiefenunschärfe* (Depth of Field) gezielt zur Andeutung von Räumlichkeit und insbesondere zur Lenkung der Aufmerksamkeit der Zuschauer eingesetzt (selektive Schärfe).

Beide Stilmittel sind heute auch in Stereo-3D-Produktionen zu finden. Im Fall von Motionblur klagen viele Filmschaffenden und Zuschauer allerdings, dass bei schnellen Action-Szenen leider der stereoskopische Tiefeneindruck zusammenbricht.

Im zweiten Fall, der Tiefenunschärfe, ist bei aktuellen Stereo-3D-Produktionen (des Jahres 2012) gerade zu beobachten, dass mit immer geringerer selektiver Schärfe gearbeitet wird, um angebliche, einschränkende Auswirkungen auf die Stereopsis der Zuschauer zu vermeiden⁴. Unsere nachfolgenden Ausführungen befassen sich genau mit dieser „Vermutung“.

2 Vgl. Cutting, James E. / Vishton, Peter M. (1995): Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth, In: W. Epstein & S. Rogers (eds.) Handbook of perception and cognition, Vol 5; Perception of space and motion, Academic Press, San Diego, S.69-117

3 Bewegungsunschärfe kommt im natürlichen Sehen des Menschen nur selten vor, u.a. da unser visuelles System reale Objekte bis zu Blickfeldgeschwindigkeiten von 100°/sek noch scharf verfolgen kann (im Zweifelsfall mit saccadischen Zwischensprüngen) s. Guski, Wahrnehmen - ein Lehrbuch, 1996, S. 47 & 85.

4 Prototypisch sei verwiesen auf Die Vermessung der Welt (Deutschland/Österreich, 2012, 119 Minuten, Regie Detlev Buck, Kamera Sławomir Idziak).

3. Abgrenzung und Methodik unserer Untersuchung

Im ersten Schritt wurden nach relevanten Fachartikeln bez. Unschärfe und Stereopsis recherchiert und diese ausgewertet, insbesondere um daraus Erkenntnisse für das Design einer eigenen, praxisorientierten Untersuchungsreihe zu gewinnen.

Diese eigene Untersuchungsreihe umfasste schließlich stereoskopische Random-Dot-Bilder (abgekürzt: Randots oder RDS), welchen künstliche Unschärfe hinzugefügt wurde sowie Stereo-3D-Realfilmaufnahmen mit definierter Tiefenschärfe.

Wir achteten bei unseren eigenen Untersuchungen auf eine möglichst praxisorientierte Umsetzung und gute Reproduzierbarkeit. Aus diesem Grund legten wir für die stichprobenartigen, subjektiven Probandentests zunächst auch einen praxisorientierten Betrachtungsabstand vom Typ PVD (Preferred Viewing Distance) nach ITU-R. BT 2022 fest, konkret: einen Betrachtungsabstand von 3.1 m bei der Betrachtung des Materials auf einem 50-Zoll-Bildschirm.

Die Anzahl N der Probanden betrug 6 (Randots) bis 8 (Realfilm). Unsere Ergebnisse besitzen folglich nur Indikations-Charakter und Erheben keinen Anspruch auf Allgemeingültigkeit.

Der Einfluss von – filmisch verstärkter – Bewegungsunschärfe wurde nicht untersucht. Der Effekt kam folglich im verwendeten Testmaterial bewusst nicht vor.

4. Literaturrecherche

Nach einer eher erfolglosen Fachartikel-Suche in den Archiven der Society of Motion Picture and Television Engineers (SMPTE), der Fernseh- und Kinotechnischen Gesellschaft (FKTG), der Fachzeitschrift Cinematographer und der Fachzeitschrift Der Kameramann, wurde die Recherche auf wissenschaftliche Datenbanken ausgeweitet. Aus dem Bereich der Ophthalmologie bzw. Vision Research konnten drei Fachartikel identifiziert werden:

George Mather beschreibt in seinem Artikel „The Use of Image Blur as Depth Cue“⁵, den Effekt, dass wenn ein Bildbereich eine scharfe Textur und ein benachbarter unscharfen Strukturen enthält, beiden Bereichen von Probanden unterschiedliche Tiefen zugeordnet werden, selbst in Abwesenheit jeglicher anderer Depth-Cues (getestet wurde mit randomisierten Fraktal-Bildern). In einem späteren Versuch (Depth cue integration: stereopsis and image blur, George Mather, David R.R. Smith)⁶ wurden den Probanden zwei Random Dot Stereogramme (RDS) im 2AFC (Two alternative forced choice) Auswahlverfahren vorgelegt. Die Probanden mussten entscheiden, welches die größere Tiefe vermittelt. In den Kontrolldurchgängen waren alle Dots der RDS scharf. In den Versuchsdurchgängen waren die Dots des einen RDS komplett scharf,

5 Perception, 1997, volume 26, pages 1147-1158

6 Vision Research 40 (2000), 3501–3506

die Dots des anderen RDS wurden gegeneinander verschoben und weich gezeichnet, um einen weiten Depth Cue einzuführen. Die Ergebnisse zeigten, dass das Beifügen von unterschiedlichen Weichzeichnungsarten nur einen geringen Einfluss auf die sichtbaren Tiefenabstufungen hatte. Die Kombination beider Depth Cues, also Blur und Stereopsis, wurde von den Probanden immer besonders intensiv wahrgenommen, insbesondere im Vergleich zu den rein-scharfen Stereogrammen, welche nur die Stereopsis adressierten.

Robert T. Held, Emily A. Cooper und Martin S. Banks kommen in ihrem Artikel „Blur and Disparity are Complementary Cues to Depth“⁷ zur Schlussfolgerung, dass im realen Sehen Querdiskontinuität gepaart mit Blur die gesamte Tiefenwahrnehmung eines Raums deutlich verbessert. Ihr Szenario geht davon aus, dass ein Mensch ein Objekt scharf fixiert und parallel in der Lage ist, auch die direkt benachbarten, geringer-auflösenden retinalen Randbereiche mit auszuwerten. In diesen Bereichen kann die Tiefe folglich nur noch über *unscharfe monokulare* Depth Cues bestimmt werden. Den Probanden wurden über ein eigenentwickeltes, volumetrisches Display monokulare Raumaufnahmen mit Tiefenunschärfe und stereoskopische Raumaufnahmen mit Tiefenunschärfe, sowie perfekt tiefenscharfe Raumaufnahmen gezeigt. Die Probanden sollten jeweils die Entfernung von Objekten einschätzen. Waren für die Probanden beide Cues (Depth + Disparity) verfügbar, verwendeten sie auch beide zur Tiefenbestimmung. Im Endergebnis ergaben sich dadurch deutlich bessere Trefferquoten als ohne Tiefenunschärfe.

5. Eigene Untersuchungen mit Randot

Zunächst führen wir erste Untersuchungen mit RDS im Anaglyphenverfahren durch. RDS besitzen mehrere Vorteile: Aufgrund der computergestützten Erstellung dieser Bilder kann zu jedem Zeitpunkt auf jeden Parameter und damit jedwede Beeinflussung individuell eingegangen werden. Zum anderen sprechen RDS ausschließlich die patente Stereopsis und somit die für den Tiefeneindruck bei Stereo-3D-Filmen maßgebliche Wahrnehmungskomponente an.

Um zunächst möglichst gut wahrnehmbare Tiefe in den RDS-Testbildern realisieren zu können, experimentierten wir mit unterschiedlichen Punktgrößen und verschiedenen Parallaxen. Die Dots, welche wir mittels Adobe Photoshop erstellten, bestanden aus schwarzen Pixeln auf weißem Hintergrund. Um die Unschärfe animieren zu können, haben wir die RDS mit Adobe After Effects weiterverarbeitet.

Im ersten Versuch verwendeten wir zunächst Dots mit einer Größe von 1x1 Pixel. Hierbei stellten wir allerdings fest, dass die Tiefenwahrnehmung trotz unterschiedlicher Parallaxen von 2-20 Pixeln nicht sonderlich gut funktioniert.

⁷ Current Biology 22, 426–431, March 6, 2012, Elsevier, DOI 10.1016/j.cub.2012.01.033

Es fiel unseren Probanden bei derart detaillierten RDS auch ohne Weichzeichnung schwer, aus der Betrachtungsentfernung (PVD) Tiefe wahrzunehmen. Minimale Besserung zeigten höhere Parallaxen von bis zu 20 px.

Bei unserem zweiten Versuch veränderten wir die Größe der Dots auf 2x2, 4x4, 8x8 und 16x16 Pixel. Bereits bei Dots mit einer Größe von 2x2 Pixeln konnten unsere Probanden (Stichprobe mit 3 Personen) deutlich Tiefe wahrnehmen, auch wenn wir unser Objekt nur um zwei Pixel in den Vordergrund versetzten. Zur besseren Trennung versetzten wir zudem auch den Hintergrund um 2 Pixel hinter die Displayebene. Allerdings wurden bei größeren Dots von 8x8 px und 16x16 px die Aufmerksamkeit der Probanden zu stark auf einzelne Dot-Anhäufungen gelenkt. Dots mit einer Größe von 2x2 px und 4x4 px wurden für unser Testsetup als ideal befunden.

Für den eigentlichen Probandentest entschieden wir uns für eine Dot-Größe von 2x2 Pixeln und einem Versatz des Objektes um 2 Pixel in den Vordergrund sowie einem Versatz des Hintergrundes um 2 Pixel hinter die Displayebene.

Um die Unschärfe zu erzeugen, haben wir zunächst den Gaußschen Weichzeichner verwendet, aber auch andere alternative Weichzeichner, welche über After Effects zur Verfügung standen, ausprobiert. Nachdem aber auch andere Weichzeichner wie z. B. der Schnelle Weichzeichner ähnliche Ergebnisse lieferten, entschlossen wir uns, es beim Gaußschen Weichzeichner zu belassen.



Abbildung 1: Objekt in Anaglyphen-Darstellung mit gaußscher Unschärfe von 2,9

In den Versuchen zeigten wir unseren Probanden zu Beginn immer ein Standbild mit einer Gaußschen Unschärfe von 40. Dieser Wert wurde in einer Animation über einem Zeitraum von einer Minute auf 0 reduziert. Währenddessen sollten die Probanden Objekte erkennen und diese verbal skizzieren (think aloud). Zudem sollten die Probanden ihren Tiefeneindruck laut wiedergeben.

Wir konnten feststellen, dass Unschärfe bei der Betrachtung von RDS nicht zwangsläufig den Tiefeneffekt zerstört. Geringe Unschärfen förderten bei unseren Probanden sogar die Tiefenwirkung. Eine Objekterkennung erfolgte erst, nachdem ein deutlicher Tiefeneindruck vorhanden war. Je größer die Unschärfe in den RDS war, desto län-

ger mussten die Probanden das Bild betrachten, ehe die Tiefenwahrnehmung eintrat. Dann jedoch wurde der Tiefeneindruck erst bei sehr großen Werten für die Weichzeichnung ernsthaft negativ beeinflusst.

Im nächsten Versuchsabschnitt stellten wir fest, dass die Wahrnehmung der Objekte sehr individuell ist. Insbesondere bei sehr filigranen Strukturen konnten Probanden im Ablauf der Animation teilweise Objekte überhaupt nicht identifizieren, obwohl die Unschärfe bereits gegen null lief. Ob hierfür ein nicht perfekter Visus oder fehlgeleitete Mustererkennungsprozesse der Probanden (oder beides in Kombination) verantwortlich waren, wurde nicht weiter untersucht.

Die Tiefenwahrnehmung von Randots lässt sich antrainieren. Bei wiederholten Versuchen setzte die Tiefenwahrnehmung der Probanden deutlich früher ein als bei den ersten Versuchen, obwohl alle übrigen Randbedingungen konstant blieben. Es scheint, dass auch bei RDS eine schnelle Schema-Bildung im visuellen System stattfindet.

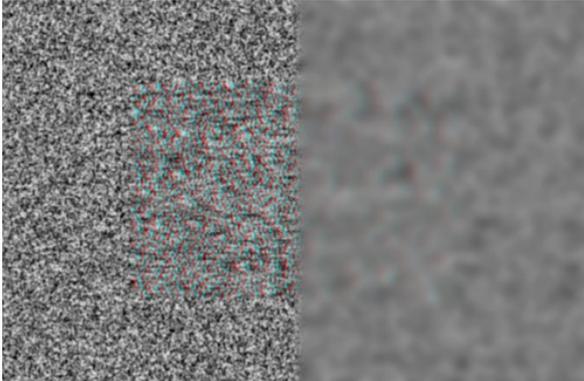


Abbildung 2: Random-Dot-Diagramm scharf/unscharf

6. Eigene Untersuchungen mit Stereo-3D-Realfilm

Da unsere Ergebnisse praxisorientiert evaluiert werden sollten, lautete der nächster Schritt Stereo-3D-Realfilmaufnahmen zu untersuchen. Dazu sollten die Versuchsbedingungen der Randots auf reale Aufnahmesituationen übertragen werden. Wir konzentrierten uns rein auf die Tiefenunschärfe, nicht auf die Bewegungsunschärfe.

6.1 Vorgehen



Abbildung 3: Versuchsaufbau zum Alignment der Kameras

Da zum Zeitpunkt der Aufnahmen kein Spiegelrig zur Verfügung stand, wichen wir auf ein Side-by-Side-Rig für parallele Kameras aus. Somit war unsere minimale (und auch verwendete) Stereobasis bereits auf 13 cm festgelegt. Um bei dieser großen Stereobasis, welche einen großen Mindestabstand zur Szene voraussetzt, noch sinnvolle Unschärfekreise realisieren zu können, sollte der Sensor unserer Kameras möglichst groß sein. Wir entschieden uns daher für die Canon EOS C300 mit einer Sensorbreite von 24.6 mm. Passend dazu wählten wir die EF 50mm f/1.4 – Objektive von Canon. Da der Blickwinkel groß genug ist, um unsere gewählten Objekte zu platzieren, können mit dieser Brennweite und dieser Maximalöffnung schnell große Unschärfekreise realisiert werden.

Über einen Stereokalkulator planten wir die übrigen Parameter so, dass sich in der Darstellung auf dem 50-Zoll-Stereo-3D-Bildschirm die Objekte dergestalt auf Vorder- und Hintergrund verteilen, dass der Betrachter im gesamten Bereich gleich große absolute Parallaxen erhält. Die maximale relative Parallaxe im Bild wurde mit maximal 42 Winkelminuten eingegeben.

Um die Kameras möglichst gut aufeinander abzustimmen, synchronisierten wir sie in ihrem Setup mittels einer geklonten SD-Karte und temporal mittels Genlock. Zudem verwendeten wir durchlaufenden Timecode, welcher von einer Kamera als Master generiert und an die zweite Kamera weitergegeben wurde. Um die Kameras möglichst perfekt auszurichten, benutzten wir ein großes Schachbrettmuster (Checkerboard), welches wir parallel vor den Kameras in einem Abstand von 4 m aufstellten. Die Ausrichtung der Bilder überprüften wir mittels eines Bildmischers, der die Möglichkeit bot, die Kamerabilder halbtransparent übereinander oder gesplittet darstellen zu können.

Damit die Unschärfe der Objekte später gut ausgewertet werden kann, berechneten wir für die HD-Teilbilder feste Unschärfekreise, nach denen wir unsere Objekte in der Szene platzierten wollten:

- 0 px: minimale Referenz

- 10 px: leichte bereits gut wahrnehmbare Unschärfe
- 20 px: mittlere Unschärfe
- 25 px: hohe Unschärfe
- 30 px: maximale Unschärfe
- 32 px: maximale Referenz (Wald im Hintergrund)

Als minimale Referenz sollte eine tiefenscharfe Aufnahme dienen, welche mittels einer hohen Blendenzahl realisiert wurde. Die Platzierung der Objekte erfolgte diagonal leicht versetzt. Mit einem Lasermessgerät wurden abschließend die Objektabstände bezogen auf die Sensorebene der Kameras nochmals exakt nachgemessen.



Abbildung 4: Screenshot aus den Realfilmaufnahmen (Zählweise der Objekte aufsteigend von vorne nach hinten)

Beim Dreh fokussierten wir jedes Objekt über die Einstellentfernung nacheinander an, wodurch die anderen Objekte die vorberechneten Unschärfekreise auf dem Sensor erzeugten.

Konkret ergab sich dabei unter anderem die Schwierigkeit, den Fokus in der linken und rechten Kamera absolut identisch zu justieren, da die Dreh-Übersetzungsverhältnisse bei den verwendeten Fotoobjektiven recht kurz waren und die Beschriftung aufgrund ihrer Ungenauigkeit in diesem Fall auch keine große Hilfe darstellte. Insbesondere bei den weit entfernten Objekten sind daher bezüglich der Unschärfekreise größere Toleranzen zu den errechneten Werten möglich.

6.2 Auswertung

Beim ersten Betrachten unserer Realfilmaufnahmen mussten wir bedauerlicherweise feststellen, dass die Aufnahmen aus einem Betrachtungsabstand von 3.1 m zum 50-Zoll-Fernseher keineswegs angenehm wirkten. Nach einer Überprüfung aller Aufnahmen, des ursprünglich verwendeten Stereokalkulators und einer manuellen Nachkalkulation wurde festgestellt, dass die maximale relative Parallaxe im Bild nicht den vorgegebenen 40 Winkelminuten, sondern immer dem jeweils Doppelten entsprach. Daraufhin wurde der Stereokalkulator durch eine selbst programmierte, verbesserte Version ersetzt.

Um trotzdem einen aufwendigen Nachdreh zu vermeiden und die Aufnahmen dennoch verwenden zu können, kompensieren wir diesen Umstand durch einen neu errechneten Betrachtungsabstand von 7 m.

Beim Betrachten der Aufnahmen stellten wir zudem fest, dass der Shutter-3D-Fernseher bei großen Parallaxen die Probanden durch auffällige Geisterbilder (ghosting) irritierten. Um einerseits die Irritation durch die Geisterbilder zu senken und andererseits eine angenehmere Betrachtung der Aufnahmen zu realisieren, beschlossen wir die Parallaxen auf dem Bildschirm mittels HIT (Horizontal Image Translation) zu reduzieren. Somit pushten wir die Aufnahmen um 64 Pixel nach hinten, wodurch wir für Vorder- und Hintergrund ähnliche Parallaxen erhielten (wohl wissend, dass wir damit die Shape-Ratio der Objekte auf ca. 1.5 erhöhten)⁸. Um auszuschließen, dass den Bildern keine weiteren Unschärfen durch HIT-Skalierungsprozesse hinzugefügt werden, kaschten wir nach der Bildverschiebung L zu R die unbrauchbaren Ränder links und rechts einfach nur ab.

Als wir unseren Probanden die Aufnahmen zeigten (s. beispielhaft Abbildung 4), stellten diese zunächst fest, dass auch über die unscharfen Bereiche hinweg eine gute Tiefenstaffelung der Objekte wahrnehmbar ist. Zusätzlich stellten wir fest, dass der Einfluss der Unschärfe auf die Tiefenwahrnehmung der Probanden – im Gegensatz zu den RDS – stetig verlief.

Allerdings zeigte sich im Vergleich zum tiefenscharfen Referenzbild auch, dass Objekte in der nach hinten laufenden Unschärfe, ab einem Unschärfekreis von ca. 20 px, deutlich an wahrgenommenen Volumen verloren. Dies bestätigte sich auch, als wir Personen im Bild von vorne nach hinten in die Unschärfe hineinlaufen ließen. Hierbei wirkten die Personen flacher, nachdem sie das dritte Objekt passierten. Unsere Probanden empfanden diesen Effekt allerdings nicht als störend oder irritierend.

Im Vergleich mit der entsprechenden tiefenscharfen Aufnahme fiel des Weiteren auf, dass die verschiedenen Objekte bei unterschiedlichen Unschärfekreisen scheinbar ihren Abstand zur Bildschirmenebene verringerten.

⁸ s. dazu den Artikel „long Lenses vs. HIT“ in diesem Buch.

Bei der Betrachtung von Objekt 1 (vorderster Karton) konnten wir selbst bei der maximalen Unschärfe von 30 Pixeln, keine signifikante Reduktion der Tiefenwahrnehmung innerhalb des Objektes beobachten. Allerdings beurteilten unsere Probanden den Abstand des Objektes bei einer Unschärfe von 30 Pixeln als minimal geringer zur Bildebene.

Bei Betrachtung von Objekt 2 trat die Reduktion des Abstands zur Bildebene deutlich früher bei einer Unschärfe von 10 Pixeln ein. Zudem konnten unsere Probanden bei einem Unschärfekreis von 20 Pixeln bereits eine minimale Reduktion der Tiefe des Objektes wahrnehmen. Aufgrund des Abstandes konnten keine größeren Unschärfekreise realisiert werden.

Da wir HIT einsetzten, befindet sich Objekt 3 (Stativ mit Kiste) auf der Bildebene, weshalb wir hier keine Veränderung der Wirkung der Tiefenstaffelung untersuchen konnten. Jedoch empfanden unsere Probanden eine deutliche Volumenreduktion der Kiste bereits bei Unschärfekreisen ab 10 Pixeln. Bei Unschärfekreisen von 20 Pixeln empfanden die Probanden die Kiste annähernd flach.

Das vierte Objekt (Checkerboard) wirkte auf unsere Probanden ab einem Unschärfekreis von 25 Pixeln vollständig zweidimensional, wobei es so schien, als wäre es noch minimal hinter der Displayebene platziert. Die Reduktion des Volumens erfolgte bereits ab einem Unschärfekreis von ca. 7 Pixeln, wobei der Abstand zwischen Bildschirm und Objekt erst ab einem Unschärfekreis von ca. 15 Pixeln geringer zu werden schien.

Das weiter entfernte Auto verlor schon bei einem Unschärfekreis von ca. 2 Pixeln an Volumen. Bei einem Unschärfekreis von 10 Pixeln wirkte das Auto für die Probanden komplett flach. Der Abstand zwischen Bildebene und Auto wirkte ab Unschärfekreisen von 5 Pixeln schon verringert.

Der Wald behielt sein Volumen bis zu einem Unschärfekreis von ca. 12 Pixeln, ehe dieses und auch der Abstand zur Bildebene sanken. Ab einem Unschärfekreis von 32 Pixeln war von den Probanden kein Volumen mehr wahrnehmbar.

Daher liegt die Vermutung nahe, dass es einen Zusammenhang zwischen Unschärfekreis und der wahrnehmbaren Tiefe im Bild besteht. Folgendes Diagramm stellt die ermittelte Grenze der Tiefenwahrnehmung der sechs Objekte dar. Die Tiefenangabe erfolgt in px-Parallaxe von gleichen Objekten zwischen den Teilbildern.

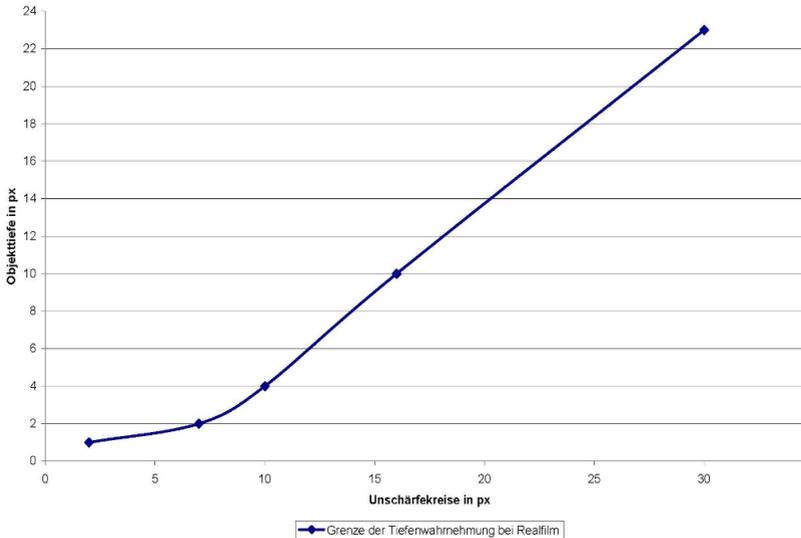


Abbildung 5: Ermittelte Tiefenwahrnehmungsgrenze in Abhängigkeit der Bildunschärfe

7. Rücktransfer der Realfilm-Ergebnisse auf RDS

Nachdem wir nun die Ergebnisse aus den Realfilmaufnahmen ausgewertet und interpretiert hatten, wendeten wir diese in einem dritten Versuch wieder auf RDS an. Hierbei sollten besonders die Zusammenhänge zwischen maximal zulässiger Unschärfe und Parallaxe herausgearbeitet werden. Gleiche Aufmerksamkeit sollte der Beziehung zwischen maximal zulässiger Unschärfe und dem Texturdichtegradienten gelten.

Die Tests wurden mit Parallaxen von 4, 8 und 20 px gefahren. Nachdem auch hier ein Zusammenhang mit der Dotgröße zu bestehen schien, führten wir alle Testreihen zusätzlich zu 2x2 px nochmals mit Dots in Größe 4x4 und 8x8 px durch.

In der ersten Testreihe mit 2x2px großen Dots stellten wir fest, dass größere Parallaxen auch bei größeren Gaußschen Unschärfen noch einen Tiefeneindruck hervorriefen. Dies bekräftigt unsere Theorie, dass größere Tiefen höhere Unschärfegrade zulassen. Des Weiteren sind auch größere Dots der Tiefenwahrnehmung zuträglich, weshalb bei großen Parallaxen und großen Dots extrem große Unschärfekreise möglich sind, bis der Tiefeneindruck komplett schwindet.

Im nachfolgenden Diagramm sind die ermittelten rauschbezogenen Tiefenwahrnehmungs-Grenzen für die unterschiedlichen Dot-Größen im Vergleich dargestellt.

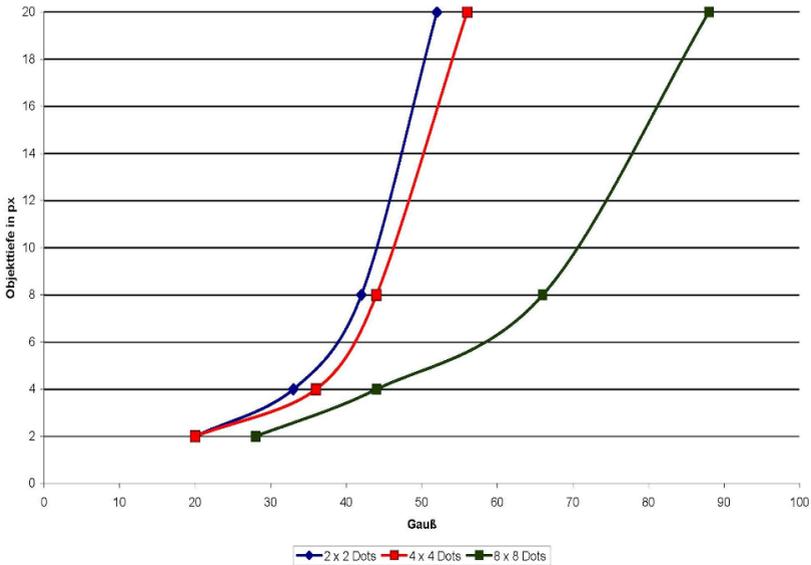


Abbildung 6: Tiefenwahrnehmungsgrenze in RDS

8. Fazit

Zusammenfassend haben wir in unseren eigenen Untersuchungen festgestellt, dass (Tiefen-) Unschärfen der räumlichen Wirkung von Stereo-3D-Filmen nicht so abträglich sind, wie oft behauptet wird. Platt gesagt: Sie müssen nur an der richtigen Stelle auftreten, dann wirken sie sogar verstärkend.

Unschärfen dürfen in einem Stereo-3D-Realfilm nur im Vorder- oder Hintergrund der eigentlichen Handlung auftreten. Fehlfokussierungen oder Blooming-/Blur-Effekte im Handlungsbereich können – wie im 2D-Film – den Zuschauer deutlich irritieren. Eine saubere Fokusebene im Bild ist ein Anker, den jeder Betrachter benötigt.

Diese Voraussetzung einhaltend ist die wichtigste Erkenntnis aus unseren Experimenten, dass dann Stereopsis und Blur gemeinsam eine deutlich stärkere Tiefenwirkung beim Betrachter auslösen können als einer von ihnen alleine. Eine Aussage, die auch in den recherchierten Veröffentlichungen zu finden ist. Beide Depth Cues ergänzen sich in ihrer Wirkung deutlich.

Grundsätzlich fanden wir weder bei Stereo-3D-Realaufnahmen noch bei stereoskopischen Randots pauschale Grenzen, ab wann Rauschen im ursprünglichen Schär-

febereich diskriminierend auf die Tiefenwahrnehmung wirkt. Dennoch lassen sich Tendenzen erkennen. So sind bei größerem Volumen der Objekte auch größere Unschärfen zulässig. Auch gröbere Texturen lassen größere Unschärfen zu.

In Produktionsberichten von Stereo-3D-Filmen ist manchmal zu lesen, dass Shots mit zu großen relativen Parallaxen (welche zu einer äußerst unangenehmen patenten Stereopsis beim Zuschauer führen können, Stichwort: Überschreitung der erträglichen Disparitätsgrenze) in der Postproduktion durch Weichzeichen abgemildert werden konnten. Dieser Trick kann für Objekte im Hintergrund tatsächlich funktionieren.

Interessanterweise bleibt bei sehr großen Unschärfen - trotz Verlust des wahrgenommenen Volumens - noch lange eine gut wahrnehmbare *Tiefenstaffelung* zwischen den weich gezeichneten Objekten bestehen.

Wie wir bei unseren Untersuchungen weiter feststellten, lassen sich Tiefeneindrücke, welche durch Unschärfe reduziert werden, punktuell auch wieder verstärken. Hierzu muss das z.B. dargestellte Volumen des unscharfen Objektes erhöht werden. Dies ist in reinen Stereo-3D-Animationsfilmen relativ problemlos möglich, indem man beispielsweise das tatsächliche Volumen oder die Stereobasis der virtuellen Kamera vergrößert.

Wird beim Stereo-3D-Filmen im vordergründigen Handlungsbereich dramaturgisch-gezielt mit Unschärfen gearbeitet, so sollte damit gerechnet werden, dass die Betrachter mehr Zeit benötigen, bis ein deutlicher Tiefeneindruck einsetzt.

In dieser Untersuchung wurde zunächst davon ausgegangen, dass Tiefenunschärfen vor oder hinter der Fokusebene *identischen* Einfluss auf die Tiefenwahrnehmung haben. Bei den stichprobenartigen Probandentests mit unseren (eher statischen) Stereo-3D-Realfilmaufnahmen *irritierten* eine Mehrheit der Probanden die Unschärfen im Vordergrund deutlicher als im Hintergrund. In der nachfolgenden Diskussion ergab sich als möglicher Grund: Stereo-3D-Realfilmaufnahmen besitzen gegenüber 2D-Filmen einen erhöhten Immersionsgrad. Folglich sind manche Probanden irritiert, dass die vordere Unschärfe – im Gegensatz zu Seherfahrungen in der Realität – nicht durch Akkommodation ausgeglichen werden konnte, während demgegenüber hintere Unschärfe problemlos in ihrer Nicht-Akkommodierbarkeit akzeptiert wurde.

Teil 5

Cardboarding Impacts: Long Lenses vs. HIT

Nikolaus Hottong, Andrej Bauseler, Arne Rosencrantz, Stefan Jeschke,
Thomas Keller, Carsten Rheindt

Abbildungsverzeichnis

Abbildung 1: Stereoskopisch wahrnehmbare Volumina um Fixationspunkte herum, hergeleitet aus den Parallaxen der umliegenden Teilbildpixel. Aufsicht! (Pfautz 2002: 84).....	193
Abbildung 2: Diskret wahrnehmbare Volumenelemente (Voxel) aufgrund der Pixelstruktur eines S3D-Displays. Dreidimensionale schematische Darstellung (Pfautz 2002: 84).....	193
Abbildung 3: Aufsicht der Betrachtung eines Würfels durch menschlichen Augen und durch ein S3D-Kamera Setup: p_d = pupil distance (Augenabstand des Betrachters; 65 mm), d = distance (Objektdistanz, Abstand der Augen zur Vorderseite des Würfels), w = cube width (Breite des Würfels), T = „sichtbarer“ Projektionsanteil der Würfeltiefe (Siragusano, Minuth 2012: 581).....	195
Abbildung 4: Veränderung des im Auge abgebildeten Tiefe-zu-Breite Verhältnisses eines Würfels mit zunehmender Objektdistanz d (Siragusano, Minuth 2012: 581).....	196
Abbildung 5: Veränderung des stereoskopisch aufgenommenen Tiefe-zu-Breite Verhältnisses eines Würfels mit zunehmender Objektdistanz d unter anschließender Variation der Stereobasis (Siragusano, Minuth 2012: 582).....	198
Abbildung 6: Aufnahme der gleichen Szene einmal mit Normalbrennweite aus der Distanz d (linkes Bild) und einmal „gezoomt“ mit der wesentlich längeren Brennweite f' aus der vergrößerten Distanz d' , so dass die weiße Gitarre messbar gleichgroß abgebildet ist (Siragusano, Minuth 2012: 583).	200

Abbildung 7: Veränderung des stereoskopisch aufgenommenen Tiefe-zu-Breite Verhältnisses des Würfels mit zunehmender Objektdistanz d unter Einkalkulation der brennweitenbasierten Fehlinterpretation durch die zerebrale Größenkonstanz bei der Betrachtung (Siragusano, Minuth 2012: 584).	202
Abbildung 8: Generelle Veränderung der Volumendicht als Funktion der gewählten Brennweite (links) und als wahrgenommene Verdichtung aller Objekte einer Szene nach Verdoppelung der Brennweite von normal=29mm auf neu=60mm (rechts). Eigene Grafik	203
Abbildung 9: HIT und absolute Parallaxe (Broberg 2011: 2)	205
Abbildung 10: Positive Parallaxe (Smith 2012: 3)	205
Abbildung 11: Horizontal Image Translation mit Re-Skalierung (Broberg 2011: 9)	207
Abbildung 12: S3D-Kettenabbildung von der ursprünglich aufgenommenen Szene (links) bis zur Betrachtung auf einem S3D-Display (rechts) (Smith 2012: 2 und 3)	208
Abbildung 13: Formen des Shape Ratio (Smith 2012: 5)	209
Abbildung 14: Veränderung der Shape Ratio durch HIT im orthostereoskopischen Aufnahme- und Betrachtungsfall für ein Objekt in 4m Entfernung von der Kamera. Eigene Grafik	211
Abbildung 15: Vergleich der Formeln von Volume Density und Shape Ratio für den konkreten Aufnahme- und Betrachtungsfall aus Tabelle 1	212

Im vorliegenden Beitrag werden die Einflüsse von langen Brennweiten während der S3D-Aufnahme (Long Lenses) einerseits und Horizontal Image Translation (HIT) in der S3D-Postproduktion andererseits auf den sog. Kulisseneffekt (Cardboarding) analysiert, gegenübergestellt und miteinander korreliert. Ziel ist es, die zentrale Praxisfrage zu beantworten, aufgrund welcher Produktionsparameter Szenenobjekte bei der Stereo-3D-Auspielung von einem Betrachter plötzlich als überproportional flacher (als in der Realität) wahrgenommen werden. Dazu werden insgesamt vier Fachartikel [Pfaltz 2002], [Yamanoue et. al. 2006], [Siragusano, Minuth 2012], [Smith, Collar 2012] ausgewertet und anhand eines konkreten Praxisfalls nachkalkuliert.

1. Einleitung

Bei der Produktion von stereoskopischen Inhalten kann es passieren, dass Objekte oder Personen irritierend mit übertrieben flachem Volumen wahrgenommen werden. Man bezeichnet diesen Effekt als Cardboard- oder Kulisseneffekt¹. Für ein grundsätzliches Verständnis dieses Effekts muss man zunächst das Setup für eine S3D-Aufnahme, für die S3D-Betrachtungssituation sowie die wirkenden wahrnehmungspsychologischen Effekte erläutern. Nachdem dies vollzogen ist, wird der Cardboard-Effekt als solcher definiert. Danach wird auf die Beeinflussungen des Effektes durch lange Brennweiten bei der Aufnahme und HIT eingegangen. Bezüglich des Einflusses langer Brennweiten orientieren wir uns an [Siragusano, Minuth, 2012], da hier konsequent der praktische Bezug zu den S3D-Produktionsgrößen gesucht wird (allerdings ohne HIT in der Postproduction mit einzubeziehen). Bezüglich des Einflusses von HIT orientieren wir uns an [Smith, Collar 2012].

1.1 Wahrnehmung stereoskopischer Bilder

Stereopsis entsteht bei der der Betrachtung auf S3D-Displays – vereinfacht dargestellt – über die Auswertung *korrespondierender Bildpunkte* zwischen dem linken und dem rechten Teilbild. Betrachtende richten ihren retinalen Schärfebereich (mittels Blickrichtungskonvergenz bei konstanter Akkommodation) interessen- oder aufmerksam-

¹ vgl. Tauer 2010: 105

keitsgeleitet auf ein Object of Interest. Umgehend greifen dann drei Effekte gleichzeitig ineinander:

- Wenn sich (Objekt-)korrespondierende Bildbereiche auf unterschiedlichen horizontalen Positionen in den Teilbildern befinden, ergibt sich bei deren Fixation durch die okkulomotorische Auswertung der Augenkonvergenz ein erster absoluter Positionseindruck (Objekt befindet sich scheinbar vor, auf oder hinter dem Bildschirm)
- Gleichzeitig verschmelzen beide retinalen Schärfebereiche zu einem zerebralen Seheindruck.
- Im Zuge dessen werden - mittels der patenten Stereopsis - alle Bildpunkte im Umfeld von ca. $\pm 6^\circ$ um die Fixation herum aufgrund ihrer retinalen Querdisparität / Deviation in einen relativen Raumbezug zum okkulomotorisch ermittelten Absolutpunkt gesetzt.

Abbildung 1 verdeutlicht diesen Aspekt aus der Sicht eines Betrachters, welcher (idealisiert) auf einem S3D-Display korrespondierende Teilbildpixel fixiert. Durch die resultierende Kreuzung der Blickachsen ergibt sich eine absolute Raumposition-Wahrnehmung der betrachteten Szene. Und durch die patente stereoptische Auswertung im $\pm 6^\circ$ Umfeld ergibt sich zeitgleich der relative Raumbezug von Umgebungsobjekten (zum Fixationspunkt). Dabei deckt jede L/R-Pixelkombination des Displays ein diskretes Elementarvolumen (Voxel) ab. Somit kann man sich plastisch recht einfach verdeutlichen, welche Kombination von L/R-Teilbildpixeln, welches dargestellte Raumsegment repräsentieren. Die absolute Größe dieser Elementarvolumina ist durch die Dotpitch (Pixelabstand) des Displays und den eingenommenen Betrachtungsabstand D_v vorgegeben.

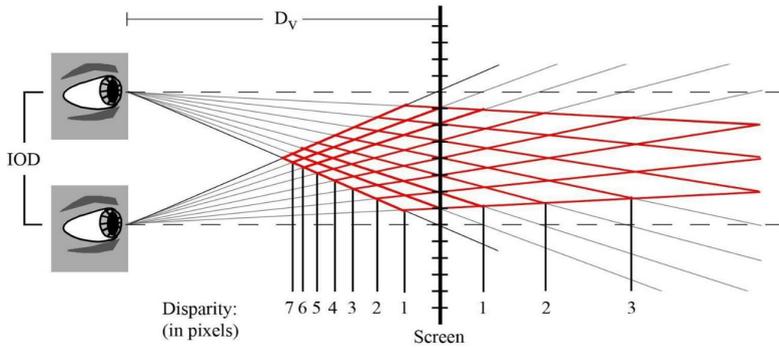


Abbildung 1: Stereoskopisch wahrnehmbare Volumina um Fixationspunkte herum, hergeleitet aus den Parallaxen der umliegenden Teilbildpixel. Aufsicht! (Pfautz 2002: 84)

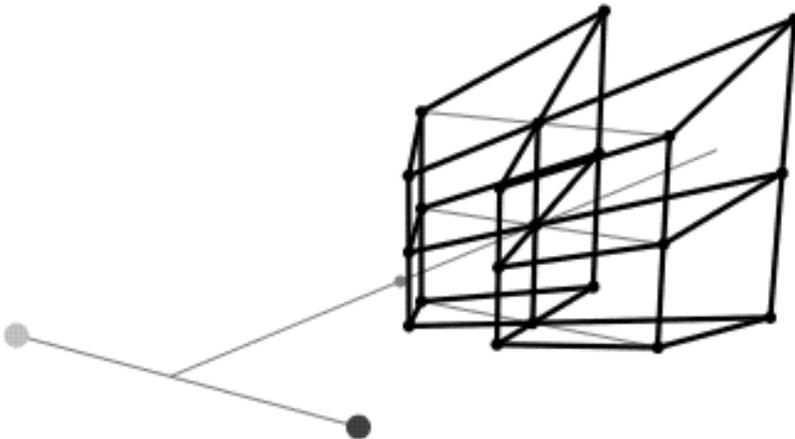


Abbildung 2: Diskret wahrnehmbare Volumenelemente (Voxel) aufgrund der Pixelstruktur eines S3D-Displays. Dreidimensionale schematische Darstellung (Pfautz 2002: 84)

Auf Basis der als Aufsicht zu verstehenden Abbildung 1 werden nun mehrere Praxisaspekte jeder S3D-Darstellung sofort verständlich:

- Alle darstellbaren Raumvolumina vom Screen zum Betrachter *hin* sind einer wachsenden *Stauchung* unterworfen.
- Alle darstellbaren Raumvolumina vom Screen zum Betrachter *weg* sind einer wachsenden *Dehnung* unterworfen.
- Jeder Medienwechsel zwischen unterschiedlich großen S3D-Darstellungsmedien gleicher örtlicher Auflösung (also unterschiedlicher Dotpitch) kann zu Raumstauchungen (z.B. 3D-Kinofilm auf kleinem Display gezeigt) oder Raumdehnungen (z.B. S3D-Game auf großer Leinwand gespielt) führen, welche durch Variation des Betrachtungsabstands kaum mehr kompensierbar sind.
- Jede bildbasierte horizontale Verschiebung der L/R-Teilbilder untereinander kann eine S3D-Szene mehr zum Zuschauer hin oder mehr vom Zuschauer weg positionieren (s. Kapitel „H.I.T.“)

1.2 Betrachtung von Körpervolumina durch Augenpaare vs. Stereokamera

Als erläuternden Anschauungsmodell von [Siragusano, Minuth 2012] wird zunächst ein Würfel im Raum mit den Maßen $1\text{m} \times 1\text{m} \times 1\text{m}$ angenommen (s. Abbildung 3/links, planare *Aufsicht!*). Dieser fiktive Würfel liegt genau auf der Blickachse des linken Auges eines Beobachters und wird nun auf dieser schrittweise bis in die Unendlichkeit entfernt. Im Verlaufe seiner zunehmenden Distanz d verändert sich auf der rechten Retina des fixierenden Beobachters – gekoppelt an die Winkel ω und δ – das Verhältnis von retinal abgebildeter Tiefe zu Breite. Dieses Verhältnis nimmt mit zunehmenden Abstand offensichtlich ab.

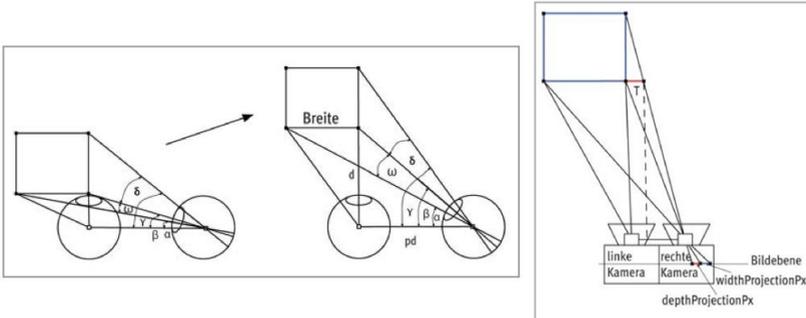


Abbildung 3: Aufsicht der Betrachtung eines Würfels durch menschlichen Augen und durch ein S3D-Kamera Setup: pd = pupil distance (Augenabstand des Betrachters; 65 mm), d = distance (Objektdistanz, Abstand der Augen zur Vorderseite des Würfels), $width$ = cube width (Breite des Würfels), T = „sichtbarer“ Projektionsanteil der Würfeltiefe (Siragusano, Minuth 2012: 581)

(Anm.: Wir haben in den nachfolgenden Formeln statt des Begriffs *pupil distance* die Angabe *Interocular Distance IO* verwendet, damit die Formeln der unterschiedlichen Autoren vergleichbar werden)

$$\alpha = \deg \left(\tan^{-1} \left(\frac{d}{width + IO} \right) \right)$$

$$\beta = \deg \left(\tan^{-1} \left(\frac{d}{10} \right) \right)$$

$$\gamma = \deg \left(\tan^{-1} \left(\frac{d + width}{IO} \right) \right)$$

$$\omega = \beta - \alpha$$

$$\delta = \gamma - \beta$$

$$\gamma = \frac{\delta}{\omega}$$

Der Funktionsverlauf des Quotienten aus retinal projizierter Tiefe-zu-Breite über die Objektdistanz kann für Abb. 3/linker Fall über das Winkelverhältnis δ/ω abgeleitet werden und ist in nachfolgender Abbildung 4 dargestellt.

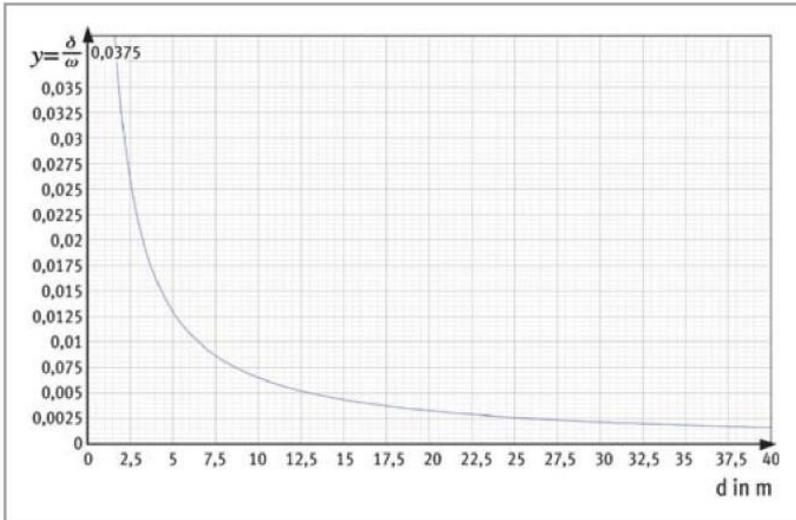


Abbildung 4: Veränderung des im Auge abgebildeten Tiefe-zu-Breite Verhältnisses eines Würfels mit zunehmender Objektdistanz d (Siragusano, Minuth 2012: 581)

Das menschliche Gehirn scheint diesen Effekt allerdings – erfahrungsbasiert – weitestgehend wieder auszugleichen, so dass menschliche Betrachter, trotz fortschreitender Entfernung immer das gleiche Volumenverhältnis vermuten. Betrachten Sie zum Beweis einfach ein sich von Ihnen entfernendes Auto. Zu keinem Zeitpunkt kommt Ihnen das Fahrzeug mit zunehmender Entfernung flacher vor. Auf diese „wahrnehmungspsychologische Korrektur“ der retinalen Abbildungsverhältnisse namens *Größenkonstanz* kommen wir später noch einmal im Detail zurück.

Im nächsten Gedankenschritt sollen im obigen Anschauungsmodell zunächst aber die Augen eines Betrachters durch zwei parallel ausgerichtete Kameras ersetzt werden. Bezüglich der Tiefe im Stereobild wird – analog zum vorherigen Fall – die Projektion der vorderen und hinteren rechten Ecke des Würfels auf den planen Sensor der rechten Kamera betrachtet. Von der linken Kamera wird in dieser Anordnung ebenfalls keine Tiefeninformation aufgenommen, da sie die hintere rechte Ecke nicht „sehen“ kann. Die Querdisparität korrespondierender Bildpunkte = relative Parallaxe zwischen beiden Würfel-Teilbildern, aus denen sich bei der S3D-Projektion das stereoskopisch wahrnehmbare Volumen ergibt, wird somit ausschließlich von der rechten Kamera generiert. Die beiden parallelen Kameras arbeiten zunächst im orthostereoskopischen Modus, d.h. sie bilden die binokulare menschliche Wahrnehmung quasi-identisch ab, in dem ihr Sensorabstand/Stereobasis (IA - Interaxial) dem durchschnittlichen menschlichen Augenabstand von 65mm entspricht. Die Interessen- und Aufmerksamkeits-gesteuerten (kontinuierlich bis saccadischen) Konvergenzbewegung der

menschlichen Augen werden folgendermaßen substituiert: das vom Menschen rein durch Augenbewegungen (ohne Kopfbewegungen) in hinreichender Schärfe abtastbare *maximale horizontale Blickfeld*² entspricht horizontal etwa 50°, also ±25° um die zentrale Blickachse Kopf-zu-Leinwand. Es wird daher von Siragusano konservativ angenommen, dass der durch Sensorbreite und Objektivbrennweite bedingte horizontale Öffnungswinkel der Kameras 45° entspricht und dass ein angenommener Betrachter des S3D-Bildes ebenfalls genauso zu Leinwand/Display positioniert wird, dass durch linken und rechten Bildrand gerade sein gesamtes horizontales 45°-Blickfeld ausgefüllt ist.

Da in der Testanordnung Abb.3/rechts davon ausgegangen werden kann, dass die Objektdistanz *d* schnell mehrere Größenordnungen größer ist als die Objektivbrennweite der Kamera, darf auf dieses Szenario die vereinfachte Abbildungsgleichung

$$\frac{\text{projizierte Bildgröße}}{\text{Gegenstandsgröße}} = \frac{\text{Brennweite}}{\text{Objektdistanz}} \quad \text{bzw.} \quad \text{projizierte Bildgröße} = f * \frac{\text{Gegenstandsgröße}}{\text{Objektdistanz}}$$

angewandt werden.

Setzt man nun als Gegenstandsgröße einmal die Würfelbreite *B* und ein weiteres Mal den „horizontal sichtbaren“ Anteil *T* der Würfeltiefe in obige Gleichung ein, erhält man für beide Würfelseiten das jeweilige *horizontale Projektionsergebnis*. Wird nun der gleiche Quotient gebildet, wie im diskutierten Fall der menschlichen Betrachtung, also projizierte Würfeltiefe zu projizierter Würfelbreite, so wird sich in diesem Quotienten der Einfluss der Brennweite offensichtlich herauskürzen.

Führt man nun ergänzend die „horizontal sichtbare“ Tiefe *T* des Würfels geometrisch auf die Produktionsparameter Stereobasis *IA* und Objektdistanz *d* zurück so ergibt sich:

$$T = IA - \frac{d * IA}{\text{Würfelbreite } B + d}$$

Setzt man diesen Term nun in die Verhältnisgleichung Tiefe-zu-Breite ein, so ergibt sich bei einer als konstant angenommenen Stereobasis *IA* wieder eine hyperbolische Funktion über den Würfelabstand *d*.

$$\frac{\text{projizierte Würfeltiefe}}{\text{projizierte Würfelbreite}} = \frac{T}{B} = \left[IA - \frac{d * IA}{\text{Würfelbreite } B + d} \right] \div B$$

Änderungen der Stereobasis *IA* an der S3D-Kamera verändern nachweislich nur die Steigung dieser Funktion, nicht aber ihren grundsätzlichen hyperbolischen Bezug zur Objektdistanz *d*, wie aus dem Plot in Abbildung 5 zu ersehen ist.

2 Nicht zu verwechseln mit dem ungleich größeren, bewegungssensitiven Gesichtsfeld von ca. 200°.

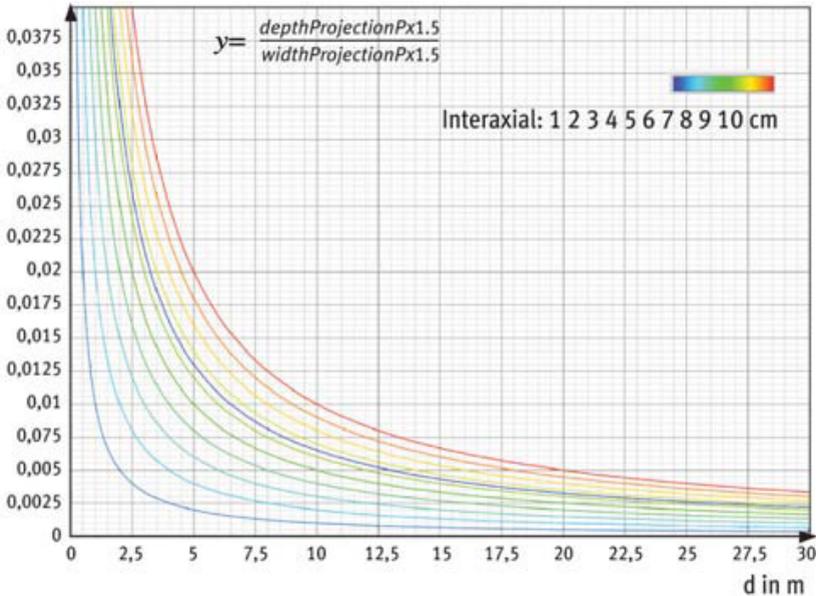


Abbildung 5: Veränderung des stereoskopisch aufgenommenen Tiefe-zu-Breite Verhältnisses eines Würfels mit zunehmender Objektdistanz d unter anschließender Variation der Stereobasis (Siragusano, Minuth 2012: 582)

Zwischenfazit:

Unter vergleichbaren Randbedingungen bilden S3D-Systeme das natürliche Sehverhalten des Menschen geometrisch exakt nach. Auch in einem S3D-Film nimmt das Tiefe-zu-Breite-Verhältnis eines sich entfernenden Objektes nachweislich hyperbolisch ab. Folglich kann dies durch einen S3D-Zuschauer zerebral genauso kompensiert werden wie im realen Leben. Vereinfacht gesagt: auch in einem (ortho-stereoskopischen) S3D-Film wird ein weggehendes Auto niemals als flacher-werdend wahrgenommen.

1.3 Wahrnehmungspsychologie von Aufnahmen mit langer Objektivbrennweite

Der im Zwischenfazit des zurückliegenden Kapitels beschriebene wahrnehmungspsychologische Korrektoreffekt wird als *Größenkonstanz* bezeichnet und sagt vereinfacht aus: „Eine wichtige evolutionären Leistungen der visuellen Wahrnehmung des Menschen ist es, die reale Größe von Objekte, losgelöst von der Distanz, schätzbar wiedergeben zu können. Dies erfordert eine spezifische Leistung des Gehirns, nämlich

aus einem Netzhautbild variabler Größe eine weitgehend konstant *wahrgenommene* Größe zu erzeugen. ... Hierzu scheinen zerebral unterschiedliche Kriterien herangezogen zu werden, z. B. die Konvergenzwinkel der Augen, vor allem aber Vergleichsobjekte bekannter Größe sowie der perspektive Rahmen³.

Wir kennen alle die optischen Täuschungsbilder mit denen wir uns diesen zerebralen Korrektoreffekt vergegenwärtigen können. Touristen, welche angeblich den Schiefen Turm von Pisa stützen oder ein scheinbar ganz naher, riesiger Mond, der über den Bergen am Horizont aufgeht weil wir automatisch die wahrgenommene Größe seiner Scheibe in Relation zu den uns bekannten Berghöhen setzen („Mondtäuschung“). Umgekehrt werden die (zufällig) scheinbar gleichen Größen von Sonne und Mond am Himmel häufig als gleiche relative Größen fehlinterpretiert⁴. Hier entsteht die Täuschung dadurch, dass aus nicht erkennbaren Entfernungsunterschieden zweier Objekte zerebral auf deren Äquidistanz geschlossen wird.

Bedeutet die bisherigen Betrachtungen folglich auch, dass die alte Photographen-Regel „Lange Brennweiten stauchen den Raum“⁵ falsch ist? Wer im TV die Teleobjektiv-Aufnahme eines Autorennens entlang der Zielgerade betrachtet, in der über 300 km/h fahrende Fahrzeuge lange Zeit kaum vorankommen zu scheinen, ist eher geneigt dieser Praxisregel der Photographie voll und ganz zuzustimmen. Und ist nicht Hitchcocks genialer Filmtrick des Vertigo-Effekts⁶ ebenfalls auf diese Faustregel zurückzuführen? JA! Denn das Anschauungsmodell berücksichtigt einen typischen Fall in der Aufnahmepraxis noch nicht. Dass eine Aufnahme statt mit Normalbrennweite f aus der Distanz d alternativ auch mit verlängerter Brennweite $f' > f$ aus einer wesentlich größeren Distanz $d' > d$ gemacht werden kann und trotzdem (zunächst scheinbar) das gleiche Bildergebnis erzielt wird. Betrachten wir dazu zwei Vergleichsaufnahmen der Autoren Siragusano & Minuth.

3 <https://de.wikipedia.org/wiki/Größenkonstanz>

4 Faktisch: Durchmesser Mond = 3476 km, Durchmesser Sonne = 1.392.684 km (wikipedia.de)

5 bzw. die Gegenregel „Weitwinkelaufnahmen dehnen den Raum“

6 vgl. <https://www.youtube.com/watch?v=sKJeTalEldM>



Abbildung 6: Aufnahme der gleichen Szene einmal mit Normalbrennweite aus der Distanz d (linkes Bild) und einmal „gezoomt“ mit der wesentlich längeren Brennweite f' aus der vergrößerten Distanz d' , so dass die weiße Gitarre messbar gleich groß abgebildet ist (Siragusano, Minuth 2012: 583).

Obwohl die weiße Gitarre – wie gewünscht – in beiden Aufnahmen gleich groß abgebildet ist, scheint bei genauerer Betrachtung im rechten Bild „die Welt aus den Fugen zu geraten“. Entgegen den bekannten/erlernten Regeln der Perspektive erscheint der weiter hinten gelegene Verstärker plötzlich größer, die weiter vorne stehende braune Gitarre plötzlich kleiner und die Abstände zwischen allen Geräten im Ganzen deutlich geschrumpft!?!

Warum ist dies so? Nun, vereinfacht gesagt, weil die Aufnahmekamera des rechten Bildes nicht mehr unsere Augen ersetzt hat. Das Bild ist jenseits der Normalbrennweite aufgenommen (und nur für *ein* bestimmtes Object of Interest konnte dies durch einen Erhöhung der Aufnahmedistanz ausgeglichen werden). Beide Bilder werden zum Schluss aber von einem Menschen mit Normalbrennweite aus dem gleichen Abstand Photo-zu-Kopf betrachtet. Und dabei kommt es im rechten Bild für den zerebralen Prozess der Größenkonstanz zu Fehlinterpretationen.

Wie kann diese Fehlinterpretation mathematisch erfasst werden, insbesondere da der Effekt relativ zwischen den Bildobjekten stattfindet? Siragusano schlägt dazu vor, für ein *object of interest* einen brennweitenbasierten Kompensationsfaktor abzuleiten:

$$\text{Kompensationsfaktor} = \left[\frac{\text{Öffnungswinkel der Kamera}}{\text{Blickfeldöffnung des Betrachters}} \right]$$

Daraus kann man dann für alle Fälle *diesseits und jenseits der Normalbrennweite* einen kompensierenden zusätzlichen Kameraabstand bilden

$$\text{kompens. Abstand } \mathbf{D} = \text{ tats. Objektdistanz } d - d * \left(\frac{\text{Öffnungswinkel der Kamera}}{\text{Blickfeldöffnung des Betrachters}} \right)$$

Wurde also der Öffnungswinkel der Kamera auf die Hälfte geschrumpft, da sich die Objektivbrennweite verdoppelt, hat ein 4m entferntes Objekt plötzlich die Anmutung, als wäre es mit Normalbrennweite aus 2m Entfernung aufgenommen worden und ein 2m entferntes Objekt, als wäre es mit Normalbrennweite aus 1m Entfernung aufgenommen worden. An dieser Stelle sei darauf hingewiesen, dass sich alle Kalkulationen auf *horizontale* Öffnungswinkel beziehen. Insbesondere in Breitbildformaten muss dies konsequent berücksichtigt werden, da klassische Brennweitenangaben in Kleinbildäquivalent (KB) sich z.B. auf die Bilddiagonale beziehen.

Wenden wir diesen Term nun auf unser Würfel-Experiment an und betrachten wiederum den Tiefe-zu-Breite Quotienten des Würfels – nun aber wahrnehmungskorrigiert, bezogen auf unterschiedliche (horizontalen) Aufnahmebrennweiten bei festem Interaxial IA:

$$\frac{\text{projizierte Würfeltiefe}}{\text{projizierte Würfelbreite}} = \left[\text{IA} - \frac{\mathbf{D} * \text{IA}}{\mathbf{B} + \mathbf{D}} \right] \div \mathbf{B}$$

Im orthostereoskopischen Aufnahmefall (also hor. Öffnungswinkel der S3D-Kamera = hor. Blickfeldöffnung des später Betrachtenden) bleibt das Ergebnis gleich. Für alle anderen Brennweiten ergeben sich durch $\mathbf{D} \neq d$ plötzlich deutlich andere Kurven, da nun die brennweitenbasierte Fehlinterpretation durch die zerebrale Größenkonstanz berücksichtigt ist.

Im nachfolgenden Plot (Abb. 7) entspricht z.B. die unterste Scharkurve einer 3D-Würfelaufnahme mit 100mm KB-Brennweite (statt ortho-stereoskopisch 29mm = gestrichelte Kurve). Über alle Objektdistanzen hinweg bleibt das Tiefe-zu Breite Verhältnis fast konstant in der Größenordnung 0.0025, als wäre der Würfel (ortho-stereoskopisch interpretiert) SEHR weit entfernt. Oder anders gesagt: schon bei einer Würfelaufnahme im Abstand 7,5m mit 100mm Brennweite, *schätzt* das Gehirn eines Betrachters aufgrund der Größenkonstanz, der Würfel könnte über 30m weit entfernt sein. Betrachtet man im Weiteren, dass im 100mm Aufnahmefall die Tiefe-zu-Breite Interpretation für alle Distanzen von 7,5m bis 30m fast konstant bleibt, ist die traditionelle Photographen-Regel „Lange Brennweiten stauchen (gefühl) den Raum“ mehr als verständlich. Oder noch praktischer: die Rennautos auf der Zielgeraden *scheinen*

sich der Übertragungskamera kaum zu nähern, da - bei der verwendeten Tele-Brennweite - ihr wahrgenommenes Tiefe-zu-Breite Verhältnis entgegen der „erlernten Regel“ nur minimalsten Veränderungen unterworfen ist.

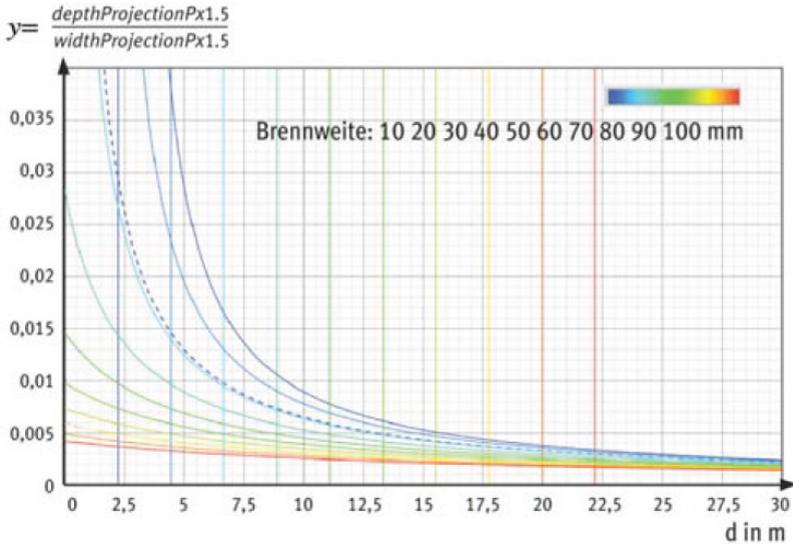


Abbildung 7: Veränderung des stereoskopisch aufgenommenen Tiefe-zu-Breite Verhältnisses des Würfels mit zunehmender Objektdistanz d unter Einkalkulation der brennweitenbasierten Fehlinterpretation durch die zerebrale Größenkonstanz bei der Betrachtung (Siragusano, Minuth 2012: 584).

Da wir in diesem Artikel aber speziell auf den *volumenorientierten* Kulisseneffekt (Cardboarding) abheben, muss nun noch eine letzte Betrachtung durchgeführt werden: wie verhält sich *das Produkt* aus wahrgenommener Tiefe mal wahrgenommener Breite?

Im orthostereoskopischen Fall:

$$\text{projiziertes Würfelvolumen (ortho)} = \text{projizierte Würfeltiefe} \cdot \text{projizierte Würfelbreite} \sim \frac{IA}{d}$$

Für all anderen Fälle mit längerer oder kürzerer Aufnahmebrennweite:

$$\text{projiziertes Würfelvolumen}_{\text{kompensiert}} \sim \frac{IA}{D}$$

Verdoppelt sich die Brennweite einer S3D-Kamera halbiert sich ihr Öffnungswinkel (in Bezug auf das natürliche Blickfeld des Betrachters). Folglich wird $D = d/2$. Unter Einbeziehung der zerebralen Größenkonstanz wird das Würfelvolumen also plötzlich wahrgenommen als wäre der Würfel dem Betrachter auf die Hälfte der tatsächlichen Distanz entgegengekommen. Er wird also als vergrößert empfunden, wobei sich – gem. Abbildung 7 – seine vom Betrachter *geschätzte* Tiefe-zu-Breite aber überproportional verringert. Er hat sich also in einen halben Dominostein verwandelt.

2. Cardboarding messen (Teil 1)

Das Phänomen des Cardboardings zeichnet sich dadurch aus, dass dargestellte Objekte (unter Einbeziehung des wahrnehmungspsychologischen Effekts der Größenkonstanz) durch *scheinbar mangelnde Tiefe* ein flaches Formgefühl auslösen und vom Betrachter plötzlich unnatürlich schmal eingeschätzt werden. Siragusano definiert als passende Messgröße für dieses Phänomen die wahrgenommene *Volumendichte* eines object of interest im Kameraabstand d und leitet diese direkt aus der Brennweitenkompensation ab:

$$\text{Volumendichte} = \frac{d - D}{d} = 1 - \frac{D}{d}$$

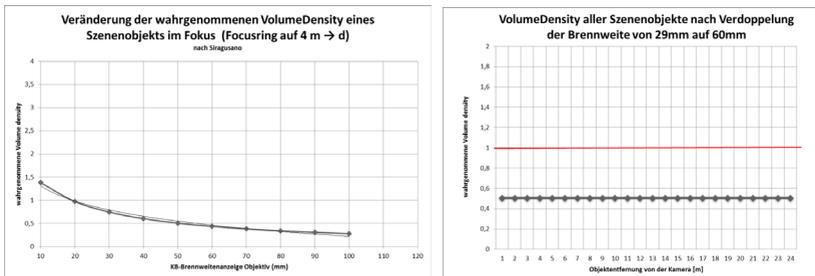


Abbildung 8: Generelle Veränderung der Volumendichte als Funktion der gewählten Brennweite (links) und als wahrgenommene Verdichtung aller Objekte einer Szene nach Verdoppelung der Brennweite von normal=29mm auf neu=60mm (rechts).

Eigene Grafik.

Es darf dabei aber niemals vergessen werden, dass die Volumendichte-Funktion von [Siragusano, Minuth 2012] immer nur die Aufnahmesituation für ein object of interest im Abstand d betrachtet und fest davon ausgeht, dass ein Zuschauer in einer orthostereoskopischen Betrachtungssituation zur Leinwand / Display verharrt. Des Weiteren gehen obige Funktionen beispielhaft von Super 35 Kameras mit einem quasi-orthostereoskopischen IA von 65mm aus. Die Brennweitenangaben sind KB-Äquivalente wie sie auf dem Objektivring abgelesen werden.

3. Horizontal Image Translation

Die *Horizontal Image Translation* (nachfolgend HIT) ist ein elektronisches Korrekturverfahren in der stereoskopischen Videoproduktion, um dreidimensionale Szenen in Bezug auf die Bildschirmenebene insgesamt mehr nach vorne oder mehr nach hinten zu verlagern, indem die Teilbilder für linkes und rechtes Auge einfach horizontal gegeneinander verschoben werden (vgl. Broberg 2011: 1).

Bei der technischen Anwendung des Verfahrens muss eigentlich zwischen einer horizontalen Bildsensor-Verschiebung in der Aufnahmesituation (Sensor Shift) oder einer nachträglichen horizontalen Teilbild-Verschiebung (HIT) in der Postproduktion unterschieden werden (vgl. Smith 2012: 11). Beide Varianten haben die gleiche stereoskopische Wirkung. Für ein besseres Verständnis von Cardboarding soll im Weiteren aber konsequent über den HIT-Ansatz argumentiert werden.

3.1 Allgemeine Funktionsweise und Auswirkung von HIT

Über die Manipulation der *absoluten Parallaxen* eines S3D-Bildes wird durch HIT Einfluss auf den Konvergenzpunkt der Augenachsen bei der Betrachtung eines Objekts genommen. Man unterscheidet bei der Betrachtung von 3D-Material drei grundsätzliche Tiefenebenen. Liegt (aus Zuschauersicht) der Konvergenzpunkt von fixierten Objekten *auf* der Screen- oder Leinwandebene, wird von der Zero-Parallax-Screen Ebene (kurz ZPS) gesprochen. Wenn der Konvergenzpunkt von Objekten *vor* der ZPS liegt, spricht man vom Bereich der negativen Parallaxe (retinal = gekreuzte Disparität). Umgekehrt, wenn der Konvergenzpunkt *hinter* der ZSP liegt, spricht man vom Bereich positiver Parallaxe (retinal = ungekreuzte Disparität).

HIT setzt genau an dieser Stelle an und ermöglicht eine nachträgliche absolute Positionsverlagerung der Szene durch eine horizontale Pixelverschiebung der Teilbilder für linkes und rechtes Auge. Folglich kann *HIT* Bildobjekte vor, auf oder hinter die ZPS verschieben, wobei die restlichen Objekte der Szene in gleicher Richtung mit wandern. Korreliert man dies mit den in Abb. 1 dargestellten wahrnehmbaren Volumensegmenten (Voxel) eines S3D-Displays, wird deutlich, dass jede Verschiebung grundsätzlich zu einer Stauchung oder Dehnung der Voxel führt. Zu klären ist aber, ob diese Stauchung/Dehnung evtl. ebenfalls der - als natürlich wahrgenommenen - hyperbolischen Volumenwahrnehmung entgegenarbeitet, so dass es aus menschlicher Sicht zu wahrnehmbaren Tiefenstauchungen oder Tiefendehnungen kommt. Smith & Collar definieren als Kontrollgröße die resultierende *Shape Ratio* von Objekten (s. Kapitel: *Cardboarding messen*)

Je nachdem in welche z-Richtung eine Gesamtszene durch HIT verschoben wird, spricht man in der Praxis entweder von *positivem HIT* oder *negativem HIT*. Anhand der Abbildung 9 soll dies erläutert werden.

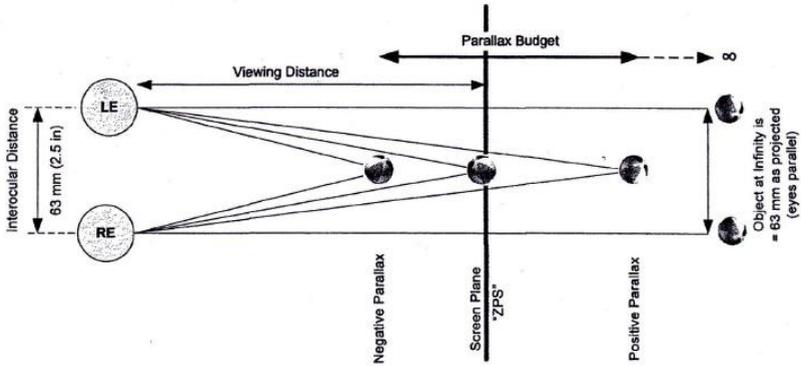


Abbildung 9: HIT und absolute Parallaxe (Broberg 2011: 2)

3.2 Funktionsweise und Auswirkung positiver HIT

Bei der Anwendung von *positiver HIT* wird die absolute L/R-Parallaxe aller Objekte vergrößert, indem eine Pixelverschiebung im rechten Teilbild nach rechts und im linken Teilbild nach links stattfindet. Dies hat zur Folge, dass der Konvergenzpunkt eines *Object of Interest* hinter den ZPS in den Raum der positiven Parallaxe geschoben wird.

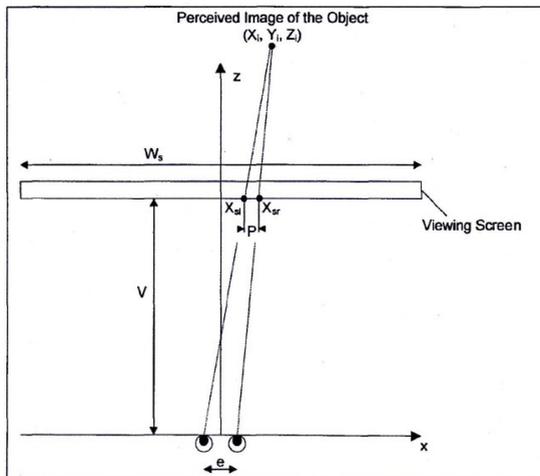


Abbildung 10: Positive Parallaxe (Smith 2012: 3)

Alle anderen Objekte der Szene werden ebenfalls nach hinten geschoben, was bei extremer Anwendung von HIT zu Divergenz der Augenachsen der Betrachter führen kann. Hervorgerufen werden diese dadurch, dass bereits sehr weit entfernt liegende Objekte noch weiter nach hinten verschoben werden und im Augenblick der stereoskopischer Fixation die Distanz P der korrespondierenden L/R-Pixel den Augenabstand des Betrachters überschreitet.

3.3 Funktionsweise und Auswirkung negativer HIT

Wird auf das *Object of Interest negative HIT* angewendet, wird die absolute Parallaxe verringert, indem eine Pixelverschiebung für das rechte Teilbild nach links und für das linke Teilbild nach rechts erfolgt. Dadurch wird der Konvergenzpunkt des *Objects of Interest* von der ZPS in den Raum der negativen Parallaxe vorgezogen und vom Betrachter näher wahrgenommen.

3.4 Generelle HIT Nebenwirkungen

Bevor die Auswirkung von HIT auf die *Shape Ratio* untersucht wird, soll noch festgehalten werden, dass nachträgliches HIT oft zu einem Auflösungsverlust in beiden Teilbildern führen kann. Durch die horizontale Pixelverschiebung entstehen in jedem Teilbild schwarze Ränder (vgl. Broberg 2011: 6). Falls diese nicht akzeptiert werden, müssen die Bilder anschließend hoch skaliert und überflüssige vertikale Pixel entfernt werden (vgl. Tauer 2010: 255f). Der komplette Anpassungsprozess ist in Abbildung 11 erläutert. Jede Hochskalierung kommt im Bildergebnis allerdings einer kleinen Brennweitenverlängerung gleich, so dass in diesem Falle gem. [Siragusano, Minuth 2012] immer mit leicht verringerter Volumendichte der Objekte zu rechnen ist.

Tauer empfiehlt deshalb dem Skalierungsproblem mit Überauflösung entgegen zu wirken (vgl. Tauer 2010: 256). Broberg kommt zu dem Ergebnis, dass sich z.B. ein 2k Kinoformat mit 2048x1080 Pixeln gut dafür anbietet, um Kinomaterial für den heimischen 3D-HDTV anzupassen (vgl. Broberg 2011: 6ff). In diesem Falle stehen für jedes Teilbild 64 Pixel für *HIT* zur Verfügung. Selbst bei einem maximalen Verlust von insgesamt 128 Pixeln horizontal resultiert daraus – ohne Re-Skalierung – gerade noch ein HD-Bild mit 1920x1080 Pixeln. (vgl. Broberg 2011: 10).

Smith und Collar scheinen bei ihren Kalkulationen von diesem Fall auszugehen, da sie in ihren Formeln eine Brennweitenverlängerung durch Re-Skalierungsprozesse nicht berücksichtigen.

Bei allen HIT-Operationen ist zusätzlich besonders darauf zu achten, dass es an den seitlichen Bildrändern nicht zu Window-Violations an Vordergrundobjekten kommt.

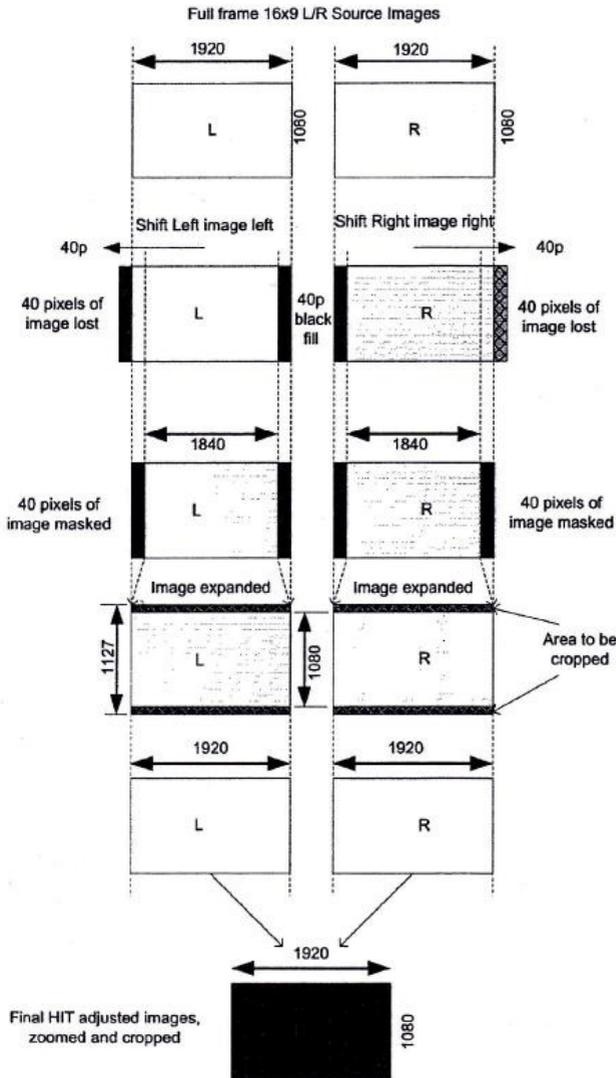


Abbildung 11: Horizontal Image Translation mit Re-Skalierung (Broberg 2011: 9)

4. Cardboarding messen (Teil 2)

Smith & Collar wählen für ihre Betrachtungen einen integralen Ansatz, indem Sie immer die komplette Kettenabbildung von der Szene über die S3D-Kamera, von dort weitergeführt über die S3D-Projektion bis auf die Retina des Zuschauers betrachten.

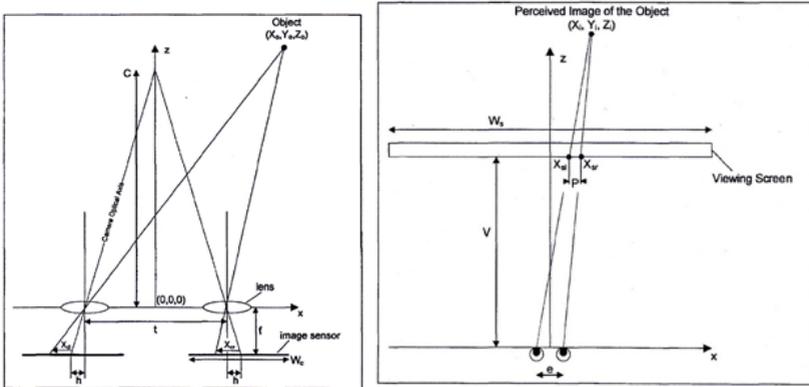


Abbildung 12: S3D-Kettenabbildung von der ursprünglich aufgenommenen Szene (links) bis zur Betrachtung auf einem S3D-Display (rechts) (Smith 2012: 2 und 3)

Dabei gehen Sie implizit von folgender Hypothese aus: Wenn eine orthostereoskopische Aufnahme unter nachfolgenden orthostereoskopischen Betrachtungsbedingungen wahrnehmungsbezogen zu keinerlei Dehnung oder Stauchung von Objekten führt⁷, dann bräuchte man ja nur noch die differentiellen Auswirkungen von nicht-orthostereoskopischen Einstellungen auf die wahrgenommenen Geometrie von Szenobjekten zu analysieren, als da wären:

1. Stereobasen (Interaxial) diesseits und jenseits des menschlichen Augenabstands
2. Kamerabrennweiten diesseits und jenseits der Normalbrennweite
3. Synthetisch eingefügte nachträgliche Bildverschiebungen zwischen rechtem und linkem Kamerabild (HIT)
4. Betrachtungsblickfelder diesseits und jenseits der Normalbrennweite (aufgrund veränderter Betrachtungsabstände oder vergrößerter bzw. verkleinerter Leinwände/Displays)

⁷ Wie im vorangegangenen Abschnitt von Siragusano & Minuth ebenfalls ausführlich erläutert

Hergeleitet wird dies über den Einfluss besagter 4 Parameter auf den Bildpunkt (X_i, Y_i, Z_i), der den ursprünglichen Szenepunkt (X_0, Y_0, Z_0) repräsentiert.

Smith & Collar bezeichnen ihren Indikator für HIT-basierte Dichteänderungen als Shape Ratio. Nach der S3D-Aufnahme und der anschließenden Projektion auf einem S3D-Display, sind drei grundsätzliche Wahrnehmungsformen des Objekts beim S3D-Betrachter möglich (nach Smith 2012: 4f):

1. Das Objekt wird in der originalen (Aufnahme-)Geometrie wahrgenommen, somit hat die Shape Ratio einen Wert von etwa 1.
2. Das Objekt wird – gegenüber der ursprünglichen Aufnahme – in der Tiefe gestaucht wahrgenommen, somit hat die Shape Ratio einen Wert kleiner als 1 → Cardboarding / Compaction.
3. Das Objekt wird – gegenüber der ursprünglichen Aufnahme – in der Tiefe gedehnt wahrgenommen, somit hat die Shape Ratio einen Wert größer als 1 → Elongation.

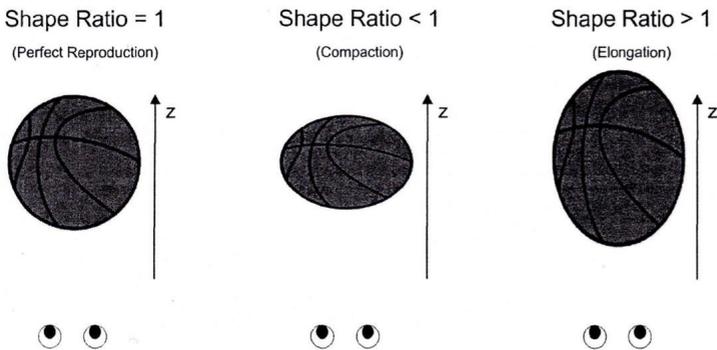


Abbildung 13: Formen des Shape Ratio (Smith 2012: 5)

Vereinfacht man vorab den horizontalen Vergrößerungsfaktor Sensor-zu-Display als Magnifikation M

$$M = \frac{\text{Displaybreite}}{\text{Sensorbreite}}$$

So ergibt sich (unter Berücksichtigung der aufnahmeseitigen Stereobasis IA , der ursprünglichen Aufnahmebrennweite f , des Objektabstands d und des Display-Betrachtungsabstands eines Zuschauers) die Shape Ratio S des Objekts als

$$S = \frac{\text{Displayabstand} * IA}{M*f*IA - d(2*M*h - IO)}$$

Die Variable h entspricht dabei der Horizontal Image Translation in Millimeter auf dem Sensor, bzw. der Term (M*h) der HIT in mm auf dem Display. Über die horizontale örtliche Auflösung des Displays lässt sich der HIT-Wert natürlich auch in Pixeln angeben.

Wir haben obige Gleichung für einen konkreten, wahrnehmungsgerecht-kalkulierten Praxisfall (quasi-orthostereoskopische S3D-HD-Aufnahme mit 2 x Canon C300, Darstellung auf einem S3D-HD-Display mit 51-Zoll Diagonale unter Standard HD-Betrachtungsbedingungen gem. ITU-R.BT.500) durchgerechnet

Stereobasis IA (mm):	65
Sensorbreite W_c (mm):	24,6
Sensorhöhe H_c (mm):	13,8
Sensordiagonale D_c (mm):	28,2
hor. Cropfaktor zum KB-Wert auf Einstellring:	1,463414634
Entfernung zulässigster Nahpunkt d NP (mm):	3000
Entfernung zulässigster Fernpunkt d FP (mm):	10000
gewünschte Brennweite für Shot [mm] = KB-Wert am Einstellring	44
resultierender hor. Öffnungswinkel Kamera [°]	31,2
gewünschter Focus [mm] = Entfernung Object of Interest d_{INTEREST} :	4000
angenommener Interokularabstand des Betrachters IO (mm):	65
hor. Blickfeld Betrachter auf Bildschirm (Grad): Default für 3DTV-Displays 1920x1080: 31° (gem ITU-R.BT500)	31,0
Betrachtungsabstand V (mm): Designed Viewing Distance gem. ITU-R.BT500	2070
hor. Pixel	1920
vert. Pixel	1080
Bildschirmdiagonale D_s (mm):	1319
Bildschirmbreite W_s (mm):	1150
Vergrößerungsfaktor Display-zu-Sensor M:	46,73203795

Tabelle 1 Praxisparameter einer typischen S3D-Produktion für-S3D-Displays in HD-Auflösung

und abschließend alleinig den HIT-Parameter h variiert um dessen Einfluss auf die Shape-Ratio eines 4m von der Kamera entfernten Szeneobjektes zu identifizieren.

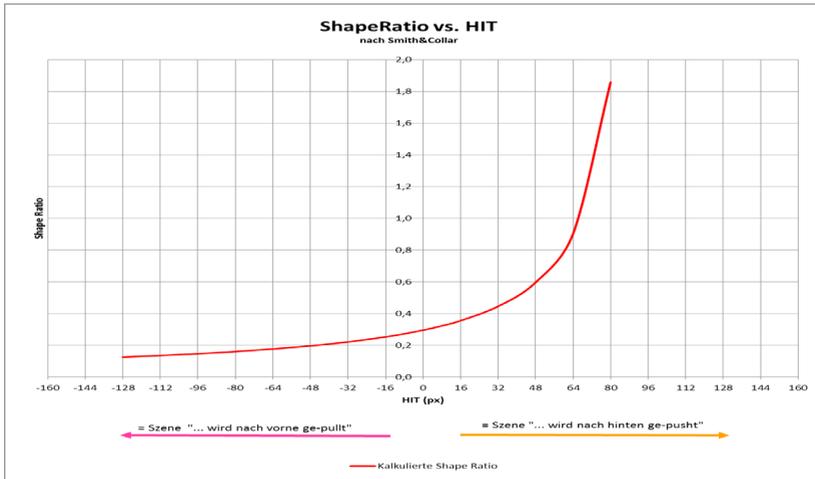


Abbildung 14: Veränderung der Shape Ratio durch HIT im orthostereoskopischen Aufnahme- und Betrachtungsfall für ein Objekt in 4m Entfernung von der Kamera. Eigene Grafik.

Es wird für den besprochenen Aufnahme- und Betrachtungsfall durch den Funktionsverlauf Abb.14 deutlich, daß durch eine positive HIT von 67 Pixeln für das Object of Interest eine neutrale Shape Ratio von 1 erhalten bleiben würde. Würde in der Postproduction kein HIT stattfinden, wäre das Object of Interest in der Tiefendarstellung generell auf 25% gestaucht.

5. Diskussion

Durch den integralen Ansatz von Smith&Collar kann nun auch eine interessante Gegenprobe zu Siragusano&Minuth gemacht werden. Hält man in der Gleichung für die Shape Ratio alle Parameter bis auf die aufnahmeseitige Brennweite orthostereoskopisch, dann müsste sich eigentlich die Funktion „Volume Density“ aus (Siragusano, Minuth 2012: 584) ergeben. Setzen wir den Praxisfall aus Tabelle 1 an, so ergibt sich tatsächlich ein ähnlicher Funktionsverlauf, allerdings mit einem leichten Versatz, dessen Ursache wir noch nicht identifizieren konnten. Insgesamt bestätigen Smith&Collar mit ihrer Shape-Ratio Formel aber die aufnahmeseitigen Einflüsse gem. Siragusano&Minuth zum Cardboard-Effekt.

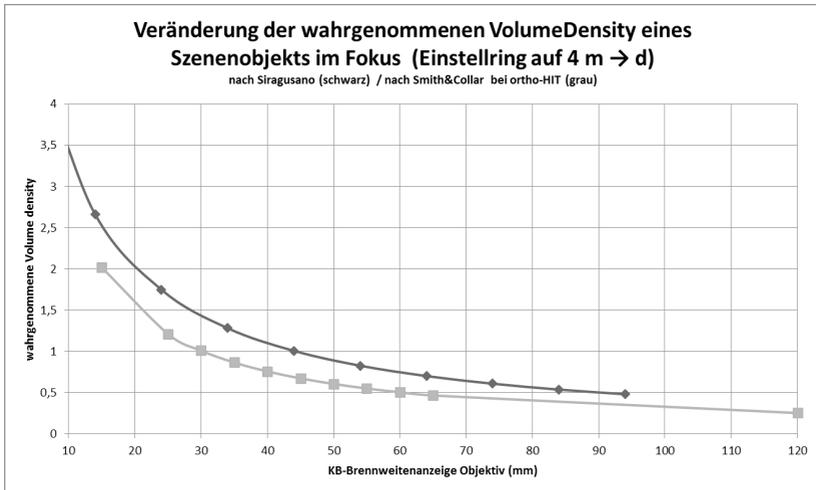


Abbildung 15: Vergleich der Formeln von Volume Density und Shape Ratio für den konkreten Aufnahmefall aus Tabelle 1

6. Zusammenfassung

Wie aus den obigen Abschnitten ersichtlich wurde, kann der Cardboard-Effekt

- in der S3D-Aufnahmephase durch die Verwendung überproportional langer Brennweiten,
- in der S3D-Postproduction bei starkem Einsatz von *negativer HIT*,
- in der S3D-Präsentation bei starker Verkleinerung des Betrachtungsmediums oder bei dramatischer Abstandsänderung der Beobachtungsposition

bedingt sein.

Alle drei Maßnahmen können sich im Laufe des Produktionsprozesses gegenseitig verstärken, in bestimmten Kombinationen aber auch gegenseitig kompensieren, z.B. in dem eine aufnahmeseitige Volume Density von 0,5 in der Postproduction durch HIT wieder auf eine Shape Ratio von 1,0 ausgeglichen wird.

Ein konkreter Grenzwert, ab wann der Kulisseneffekt vom S3D-Zuschauer bemerkt wird scheint empirisch noch nicht eindeutig bestimmt. In der S3D-Produktionspraxis sollten beim beteiligten Stereografen generell bei allen Aufnahmen mit langen Brennweiten die „Alarmglocken“ schrillen. In der nachfolgenden S3D-Postproduktion ist

ein zu starkes nach vorne „ziehen“ von Einstellungen ebenfalls mit hoher Vorsicht anzuwenden. Kommen beide Maßnahmen unbeabsichtigt zusammen ist ein starker Kulliseneffekt unvermeidlich. Auch im Falle, dass S3D-Material auf unterschiedlichen Bildschirmgrößen zum Einsatz kommt, sollte der entsprechende Konversionsschritt immer unter aufmerksamer Begutachtung der Objektvolumina durchgeführt werden.

Es sollte abschließend auch nicht unerwähnt bleiben, dass Cardboarding durchaus auch kreativ als Stilmittel eingesetzt werden kann und nicht zwangsläufig als negativer Effekt der S3D-Produktion angesehen werden muss. Durch eine geeignete Wahl entsprechender Brennweiten und dem gezielten Einsatz von HIT kann Cardboarding natürlich auch dramaturgisch erwünscht herbeigeführt werden.

Literaturverzeichnis

- Boev, Atanas / Hollosi, Danilo / Gotchev, Atanas (2011): Mobile 3DTV. Classification of stereoscopic artefacts. URL: http://sp.cs.tut.fi/mobile3dtv/results/tech/D5.1_Mobile3DTV_v1.0.pdf, abgerufen am 28.02.2013.
- Broberg, David K. (2011): Guidance for Horizontal Image Translation (HIT) of High Definition Stereoscopic Video Production. IN: SPIE 7863, Stereoscopic Displays and Applications XXIII, 82881O. SPIE-IS&T.
- Guski, Rainer (1996): Wahrnehmen. Ein Lehrbuch. Stuttgart: Kohlhammer. Online verfügbar unter <http://eco.psy.ruhr-uni-bochum.de/download/Guski-Lehrbuch/Inhaltsverzeichnis.html>.
- Hottong, Nikolaus/Walter, Paul (Hrsg.) (2011): Stereoskope HD Produktionen. Ausgewählte Prozessbetrachtungen von der Aufnahme bis zur Rezeption. Fakultät Digitale Medien - Arbeitspapier Nr. 7., Hochschule Furtwangen University.
- Hottong, Nikolaus/ Walter, Paul (2010): S3D-HD-Produktion. Teil 1: Akquise im Wandel. IN: FKT, 11/2010, S542 – 552.
- Kiening, Hans (2008): 4K+ Systems. Theory Basics for Motion Picture Imaging. URL: http://www.arri.de/fileadmin/media/arri.com/downloads/Camera/Tutorials/SystemsTechnologyB_rochure.pdf, abgerufen am 28.2.2013.
- Lipton, Lenny (1982): Foundations of the Stereoscopic Cinema. A Study in Depth. New York [u.a.]: Van Nostrand Reinhold.
- Mendiburu, Bernard (2009): 3D Movie Making. Stereoscopic Digital Cinema from Script to Screen. Amsterdam [u.a.]: Focal Press.
- Mendiburu, Bernard (2010): 3D-Cinema. Introduction to 3D Cinematography. URL: http://3dtv.fr/wordpress/wp-content/uploads/2010/08/BernardMendiburu_3D-Cinema_08-2010.pdf, abgerufen am 28.2.2013.
- Pfautz, Jonathan David (2002): Depth perception in computer graphics. Technical report. University of Cambridge, Cambridge, UK. Computer Laboratory. Online verfügbar unter <http://www.cl.cam.ac.uk/TechReports/>.
- Siragusano, Daniele/Minuth, Andreas (2012): Wahrnehmung des stereoskopischen

schen Volumens. IN: FKT, 11/2012, S.580 - 587.

- Siragusano, Daniele/Minuth, Andreas (2010): Die Rolle der Bildwandgröße in der Stereoskopie. IN: FKT, 11/2010, S.557-561.
- Smith, Michael D./Collar, Bradley T. (2012): Perception of Size and Shape in Stereoscopic 3D Imagery. IN: SPIE 8288, Stereoscopic Displays and Applications XXIII, 82881O. SPIE-IS&T.
- Tauer, Holger (2010): Stereo 3D. Grundlagen, Technik und Bildgestaltung. Berlin: Schiele&Schön GmbH.
- Yamanoue, Hirokazu/Okui, Makoto/Okano, Fumio (2006): Geometrical Analysis of Puppet-Theater and Cardboard Effects in Stereoscopic HDTV Images. In: IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, Vol. 16, No. 6, S.744-752.

Teil 6
Entwicklungstendenzen stereoskopischer
Computerspiele

Thomas Mendelssohn

Abbildungsverzeichnis

- Abbildung 1: Im Spiel Splinter Cell wird beim Werfen einer Granate die Flugbahn angezeigt, damit der Spieler die Tiefe besser schätzen kann.227
- Abbildung 2: zwei exakt gleich große Quader, die unterschiedlich im Raum angeordnet sind. In der perspektivischen Ansicht (a) sieht man durch die unterschiedliche Größe direkt, welcher Quader näher ist. In der orthogonalen Ansicht (b), die die gleiche Szene zeigt, sind beide Quader aufgrund der planaren Projektion gleich abgebildet. Erst die stereoskopische Ansicht (c) zeigt Unterschiede in der Parallaxe.....229
- Abbildung 3: zwei leicht durchsichtige Kugeln, die unterschiedlich weit entfernt sind. Auf den ersten Blick ist nicht eindeutig klar, welche Kugel die andere verdeckt.231
- Abbildung 4: unterschiedlich große und unterschiedlich weit entfernte Ringe. Betrachtet man Abbildung (a), nimmt man den größeren Ring als näher wahr, da man vermutet, dass alle Ringe gleich groß sind. Abbildung (b) zeigt, im Anaglyphen-Verfahren betrachtet, dass einige der Ringe weiter entfernt sind, dafür aber wesentlich größer sind.231
- Abbildung 5: Aufgrund der Textur (b) wirkt die nach hinten gekippte Fläche eher wie eine Trapez-förmige Mauer (a).232
- Abbildung 6: zwei unterschiedlich weit entfernte Ringe in der orthogonalen Ansicht. Bewegt sich die Kamera nach hinten (b), ändert sich lediglich die Parallaxe, die Größe der Objekte bleibt gleich.233
- Abbildung 7: Bei einer parallelen Anordnung der Kameras (a) sind die orange markierten Bereiche der Wiedergabefläche nur für ein Auge zu sehen. Bei konvergenten Kameras (b) kommt es bei größeren Parallaxen zu

starken Verzerrungen des Bildes.	236
Abbildung 8: Beim Frustum Shift werden die Kameras parallel positioniert und das Frustum wird so verschoben, dass gemeinsame Bildbereiche sichtbar sind.	237
Abbildung 9: Von den sechs möglichen Drehungen und Verschiebungen der beiden Kameras zueinander sind nur zwei wirklich anwendbar.	238
Abbildung 10: Am linken Bildrand wird beim Betrachten mit einer Anaglyphen-Brille deutlich, dass das rechte Auge bei Objekten mit negativer Parallaxe weniger sieht als das linke Auge. Es kommt zu Rahmenverzerrungen.	242
Abbildung 11: linkes maskiertes Teilbild (a), rechtes Teilbild (b) und resultierendes Bild mit Schwebefenster (c).	243
Abbildung 12: Trine 2 ermöglicht es, innerhalb des Spiels verschiedene Stereoskopie-bezogene Einstellungen vorzunehmen, sodass auch Objekte mit negativer Parallaxe auftreten.	244
Abbildung 13: Mit großer Stereobasis gerendert wirkt diese Stadt wie ein Spielzeugmodell.	246
Abbildung 14: Wird die Kamera etwas tiefer gesetzt und die Stereobasis sehr klein eingestellt, werden die Gebäude viel größer wahrgenommen.	247
Abbildung 15: Während die Flächen, aus denen die Blätter des Baumes bestehen, in der monoskopischen Ansicht (a) weniger zu erkennen sind, können sie in der stereoskopischen Ansicht (b) deutlich erkannt werden.	248
Abbildung 16: Im Spiel Flatout 2 werden öfter Blendeneffekte verwendet. Das Betrachten der Stelle rund um den Felsen führt dabei zu binokularer Rivalität.	249
Abbildung 17: Auf diesem Bild sieht man zwei Flächen, die nach hinten gekippt sind.	250
Abbildung 18: im Spiel Metro 2033, das eigentlich für die stereoskopische Ansicht optimiert ist, finden sich die bereits erwähnten Billboards. Weiterhin tritt ein Post-Effekt am Rand des Bildes störend in Erscheinung.	251
Abbildung 19: Im Spiel Left 4 Dead 2 wird verstärkt Film Grain verwendet. Ist dieser in der stereoskopischen Ansicht angeschaltet, ergibt sich ein recht unscharfes Bild.	252
Abbildung 20: Das Head-Up-Display von Crysis 2 ist etwas in die Tiefe versetzt und lässt sich daher ohne große Anstrengung während des Spielens betrachten.	253
Abbildung 21: Das Spiel Sonic Generations verwendet in der stereoskopischen Ansicht einen hohen Tiefenumfang. Der Wechsel beim Betrachten von Spielinhalt und grafischer Oberfläche strengt sehr an, da die Konvergenzpunkte zu verschieden sind.	254

Abbildung 22: Das Spiel Crysis 2 verwendet selbst im Menü den stereoskopischen Effekt. Hier wird bereits deutlich, dass das Spiel für die stereoskopische Betrachtung optimiert ist.....	254
Abbildung 23: im Spiel Starcraft II: Wings of Liberty werden die Lebensanzeigen der Einheiten und Gebäude in der richtigen Tiefe angezeigt. Einblendungen (z. B. die Anzeige des Spielers, dem eine Einheit gehört) werden allerdings auf der Nullebene ausgegeben und erzeugen damit schwer zu betrachtende Bildbereiche.....	256
Abbildung 24: Hier entsteht bewusst ein Geisterbild, da der Zylinder nur im Teilbild für das linke Auge vorhanden ist.....	257
Abbildung 25: Raumvolumen, in dem ein komfortables binokulares Betrachten stereoskopischer Inhalte theoretisch möglich ist. Darstellung angelehnt an Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 23.....	258
Abbildung 26: Je nach Abstand zur Wiedergabefläche ergibt sich eine andere Tiefenwirkung, obwohl die negative Parallaxe die gleiche bleibt.	259

1. Einleitung

Stereoskopie bietet einen zusätzlichen Grad an Immersion für den Zuschauer bei der Betrachtung von Filmen, wie im vorherigen Band (Nr. 7) dieser Reihe ausführlich dargelegt wurde. Gerne würde man diesen Nutzen - jenseits teurer CAVE-Installationen - auch in 3D-Anwendungen, insbesondere Spielen, nutzen. Hier stellen sich aber allein durch die Möglichkeit der freien Bewegung des Spielers in der 3D-Umgebung vollkommen neue Herausforderungen, welche bedacht und - womöglich mit erhöhtem Aufwand - berücksichtigt werden müssen. Die nachfolgende Analyse und Diskussion stützt sich auf die im Jahr 2012 erhältliche stereoskopische Spiele ab.

Zunächst stellt sich natürlich die grundsätzliche Frage, welchen Nutzen stereoskopische Spiele dem Anwender bringen, da dies eng mit dem Kaufanreiz für die Kunden verbunden ist. Wären keine Anreize zu erkennen, würden sich der höhere Aufwand und die damit verbundenen höheren Entwicklungskosten für den Entwickler nicht lohnen.

Der am häufigsten angeführte Vorteil der Stereoskopie, sowohl im Film als auch bei Spielen, ist die gesteigerte Spieltiefe. Dies bezieht sich vor allem darauf, dass die Wiedergabefläche als Fenster in eine andere Welt verschleiert werden kann, indem der Bildraum und der Zuschauerraum zusammenwachsen. Dies trifft auf die normale Bildschirmnutzung, im besonderen Maße aber auf die Spielenutzung mit VR-Brillen zu. Der Betrachter fühlt sich in beiden Nutzungsszenarien deutlich mehr „im Spiel“ statt nur als „Zuschauer“.

Auf wahrnehmungspsychologischer Ebene ergeben sich durch die Stereoskopie neue Hebelwirkungen. Spezial-Effekte mit Partikeln wie Wasser, Blut oder Schussbahnen wirken wesentlich überzeugender, wenn sie auf den Betrachter zukommen und lösen - quasi unterbewusst - starke körperliche Reflexe aus. Für Anwender, die selten oder noch nie stereoskopische Spiele gespielt haben, ist das Heraustreten von Objekten in den Zuschauerraum eine ganz neue Erfahrung. Dabei können die visuellen Neuerungen auch dazu verwendet werden, die Erzählweise der Handlung zu verändern, z.B. indem Objekte durch Herauskommen aus dem Wiedergabegerät die Aufmerksamkeit des Betrachters auf sich lenken, ihn erschrecken oder ihn unterhalten.

Dadurch können Botschaften übermittelt werden, die ohne den stereoskopischen Effekt nicht möglich wären.¹

Eine von Hubona et al. durchgeführte Evaluierung von Aufgaben zur Anpassung der Größe und Position von Gegenständen in stereoskopischen Anwendungen zeigte, dass das stereoskopische Sehen ein sehr dominanter Tiefenwahrnehmungsmechanismus ist und Fehler bei monokularen Tiefenreizen überschreiben kann. Die schnelle Verarbeitung binokularer Tiefenreize und das verbesserte Schätzen von Distanzen ermöglicht dabei eine bessere Raumwahrnehmung in der Spielumgebung, was dazu führt, dass der Spieler schneller in der Welt ist und Herausforderungen wie z. B. das Finden des richtigen Bremspunktes einfacher bewältigen kann. Außerdem ergibt sich daraus die Möglichkeit, neue Spielvarianten und Rätsel zu entwickeln, die eben diese Raumwahrnehmung zum Thema machen.²

Rajae-Joordens ermittelte in ihrer Studie mit Tests der elektrodermalen Aktivität³ und des Pulses die unterschiedliche Aufnahme von Filmen und Spielen in der monoskopischen und in der stereoskopischen Anzeige. Dabei wurde sowohl bei der monoskopischen als auch bei der stereoskopischen Anzeige der Spiele bei den Probanden ein Anstieg der elektrodermalen Aktivität festgestellt, allerdings fiel er bei der stereoskopischen Anzeige viel größer aus. Interessanterweise war dies bei der Betrachtung der stereoskopischen Filme nicht der Fall, was auf die unterschiedliche Einstellung der Betrachter zwischen passivem Zusehen (Film) und aktivem Spielen zurückzuführen sein könnte.⁴

Ebenfalls wurde ermittelt, dass die Punktzahlen von Spielern im stereoskopischen Modus zunächst höher waren als im monoskopischen Modus. Dies unterstützt die Aussage, dass Erstere als natürlicher empfunden wurden und eine schnellere Eingewöhnung in die Umgebung stattfindet. Da im monoskopischen Modus allerdings irgendwann auch eine Gewöhnung stattfindet, näherten sich die Punktzahlen mit der Zeit einander an, sodass nach einiger Zeit kein Unterschied mehr zwischen monoskopischer und stereoskopischer Spielweise bestand. Die Befragung der Probanden ergab, dass das stereoskopische Spielen signifikant höhere Emotionen und stärkere Gefühlsregungen als die monoskopische Anzeige erzeugte.⁵

Kurz zusammengefasst ergeben sich demnach folgende Vorteile der Stereoskopie in Spiele:

- 1 Vgl.: Crytek GmbH (2010): CryEngine 3: Stereoscopic 3D without Compromise.
- 2 Vgl.: Hubona et al. (1999): ACM Trans. Comput.-Hum. Interact., Nr. 3, Bd. 6, S. 238; Crytek GmbH (2010): CryEngine 3: Stereoscopic 3D without Compromise; Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation 3.
- 3 Messgröße für die Leitfähigkeit der Haut (s. <http://www.psychology48.com/deu/d/elektrodermale-aktivitaet/elektrodermale-aktivitaet.htm>).
- 4 Dabei wurden allerdings konvertierte Filme verwendet, die nicht ausdrücklich für die stereoskopische Wiedergabe optimiert sind und damit keine bestmögliche Tiefenwirkung erzeugen können.
- 5 Vgl.: Rajae-Joordens (2008): Measuring Experiences in Gaming and TV Applications, S. 88; Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation 3

- Gesteigerte Spieltiefe und emotionale Einbindung des Spielers
- Verbesserte visuelle Qualität des Spiels
- Neue Möglichkeiten der Erzählung von Handlungen
- Verbesserte Raumwahrnehmung und besseres Schätzen von Distanzen
- Möglichkeit neuer Spielprinzipien

2. Entwicklung stereoskopischer Spiele

Eine der wichtigsten Fragen für Entwickler stereoskopischer Spiele besteht darin, inwiefern sich das Spieledesign für diese Art von Spiel verändert. Einer der großen Vorteile ist es dabei, dass das Denken der Entwickler bereits generell im Raum stattfindet. Wurden bisher in 3D-Spielen alle Informationen über die Tiefe mittels monokularer oder bewegungsbasierter Tiefenreize vermittelt, erweitert die stereoskopische Darstellung die Tiefenwahrnehmung um die binokularen Tiefenreize (patente und okulomotorische Stereopsis).

Wie sehr sich das Spieledesign dadurch verändern muss, ist allerdings maßgeblich davon abhängig, wie und wofür der stereoskopische Effekt verwendet werden soll. Dabei lassen sich derzeit drei Varianten identifizieren:

- Verbesserung der Spieltiefe ohne direkten Nutzen für die Spielmechanik.
- Verbesserung der Tiefenwahrnehmung und damit verbundene Vorteile beim Spielen.
- Stereoskopische Tiefenwahrnehmung als Hauptbestandteil der Spielmechanik.

Diese drei Varianten werden in den folgenden Unterkapiteln näher erläutert.

2.1 Stereoskopie zur Verbesserung der Spieltiefe

Eine grundsätzliche Verbesserung der Spieltiefe lässt sich immer erzielen, wenn die „neuen“ Probleme und Herausforderungen durch die Adressierung der Stereopsis erkannt und bewältigt werden. Hierbei geht es hauptsächlich darum, die Grenzen zwischen Zuschauerraum und Bildraum wahrnehmungsgerecht zu eliminieren, um den Spieler mehr am Geschehen teilhaben zu lassen. Bei diesen Spielen findet meist nur eine globale Optimierung der monoskopischen Ansicht statt, in Einzelfällen werden zudem handlungsrelevante Gegenstände hervorgehoben. All diese Änderungen haben allerdings keine Auswirkung auf den grundsätzlichen Spielablauf und dienen nur als visu-

eller Zusatz. Diese Nutzung von Stereoskopie in Spielen findet bereits heute eine breite Anwendung, vor allem durch die automatische Umwandlung von bereits bestehenden Spielen durch Post-Stereoverfahren (s. nachfolgendes Kapitel). Litwiller und LaViola⁶ konnten zeigen, dass sich dabei in unterschiedlichen Genres keine Verbesserung der Leistung der Spieler im Vergleich zur monoskopischen Ansicht ergab. Sie weisen daher darauf hin, dass Spieleentwickler besser neue Wege für Interaktionen und die Erzählweise von Spielen entwickeln sollten, um einen echten Vorteil aus der Stereoskopie zu ziehen.⁷

2.2 Stereoskopie als Vorteil für den Spieler

Ein Schritt in diese Richtung ist die Verwendung der Stereoskopie zur besseren Raumwahrnehmung. Vorteile durch die verbesserte Tiefenwahrnehmung im Spiel kommen dann zustande, wenn das Spiel Aufgaben enthält, bei denen der stereoskopische Effekt ein präziseres Einschätzen von Distanzen ermöglicht. Dies ist vor allem bei Bewegungen auf der z-Achse der Fall, also Bewegungen vom Betrachter weg oder zu ihm hin, da diese monoskopisch sehr schwierig geschätzt werden können. Hier kann in der realen Welt nach Fällen gesucht werden, in denen das binokulare Sehen von großer Wichtigkeit ist. So ist z. B. das Fangen und Werfen von Bällen eine Aufgabe, die mit zwei Augen wesentlich einfacher ist. Innerhalb von monoskopischen Spielen wird dies oftmals ausgeglichen, indem dem Spieler Wurfbahnen angezeigt werden (s. Abb. 1). Da das Einschätzen der Entfernung in der stereoskopischen Ansicht wesentlich leichter fällt, hätte der Spieler hier einen Aktions-Vorteil, wodurch sich Aufgaben einfacher lösen lassen.

6 s. LaViola/Litwiller (2011): Evaluating the benefits of 3D stereo in modern video games.

7 Vgl. ebd., S. 2354.

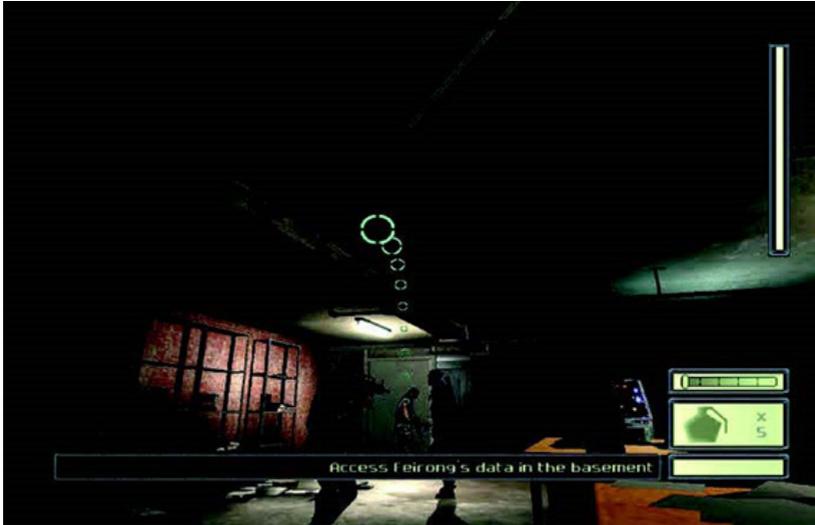


Abbildung 1: Im Spiel Splinter Cell wird beim Werfen einer Granate die Flugbahn angezeigt, damit der Spieler die Tiefe besser schätzen kann. Abbildung von <http://www.visualwalkthroughs.com/splintercell/mission8a/mission8a.htm>.

Während des Designs eines solchen Spiels müssen daher Spielsituationen identifiziert werden, bei denen die Stereoskopie von Vorteil ist. Grafische Hilfen für die Wahrnehmung von Abständen können dann für den stereoskopischen Modus angepasst oder sogar ganz abgeschaltet werden.

Diese Spiele eignen sich nicht für einen Multiplayer-Modus, denn hier geht es zu allererst um die gleichen Voraussetzungen der Spieler, die bei unterschiedlicher Betrachtungsweise nicht mehr gegeben ist.

2.3 Stereoskopie als Hauptbestandteil der Spielmechanik

Ist die stereoskopische Tiefenwahrnehmung der Hauptbestandteil der Spielmechanik, ändert sich das Spieledesign fundamental. Daher soll auf diese Variante ausführlich eingegangen werden.

Grundidee dieser Spiele ist nicht unbedingt, dass das Spielen dem Spieler in der stereoskopischen Darstellung einen Vorteil gegenüber der monoskopischen Ansicht bringt. Vielmehr geht es darum, das Spiel ausschließlich stereoskopisch spielbar zu machen, sodass bestimmte oder sogar alle Teilaufgaben des Spiels nicht mehr monoskopisch lösbar sind. Diese Variante stellt sicherlich einen Extremfall dar, da das Ausschließen der monoskopischen Nutzung umgehend eine Reihe von Einschränkungen

mit sich bringt, die das Spieledesign erheblich verändert. Andererseits ist ein nur rein stereoskopisch lösbares Spiel die beste Möglichkeit, eventuelle neue Spielpotenziale gezielt zu identifizieren.

Zur Präzisierung soll folgendes Entwicklungsszenario angenommen werden:

- Es findet eine Interaktion zwischen dem Spieler und stereoskopisch dargestellten Gegenständen statt.
- Hauptbestandteil der Interaktion ist die Tiefenverortung des Gegenstandes, vornehmlich basierend auf binokularen Tiefenreizen.
- Das Ermitteln der Tiefe rein über monokulare Tiefenreize soll kaum möglich sein.

Für die Interaktionen gibt es eine Reihe von Möglichkeiten wie: Auswahl von Objekten auf Basis der Tiefenanordnung, Objekten ausweichen oder diese gezielt treffen, Größe und Position von Objekten im Raum erkennen, etc. Hieraus lässt sich durchaus eine Reihe von neuen, interessanten Spielprinzipien entwickeln.

Die dritte Vorbedingung lässt allerdings erkennen, dass solche Spiele mit Beschränkungen verbunden sind, denn hier müssen monokulare und bewegungsinduzierte Tiefenreize gezielt eliminiert werden. Im Gegensatz zum Foto oder Filmen sollte dies bei computergenerierten Bildern allerdings wesentlich leichter sein, da dort alle Szenarien von Grund auf neu erstellt werden müssen und demnach sämtliche Parameter von Beginn an (theoretisch) noch variabel sind.

2.3.1 Eliminieren der monokularen und bewegungsinduzierten Tiefenreize

Im Folgenden werden Lösungsansätze für gezielt zu reduzierende monokulare Tiefenreize diskutiert:

- **Perspektive:**

Dieser Tiefenreiz gehört zu den kritischsten in diesem Zusammenhang, denn er kann nicht direkt verhindert werden, da bei monoskopischen Darstellungen dreidimensionaler Räume automatisch eine Perspektive vorhanden ist. Eine Möglichkeit, dies zu umgehen, stellt eine orthogonale Kamera-Einstellung dar (s. Abb. 2). Dabei entstehen keine perspektivischen Konvergenzen, da alle parallelen Linien gleicher Länge auch gleich lang dargestellt werden.

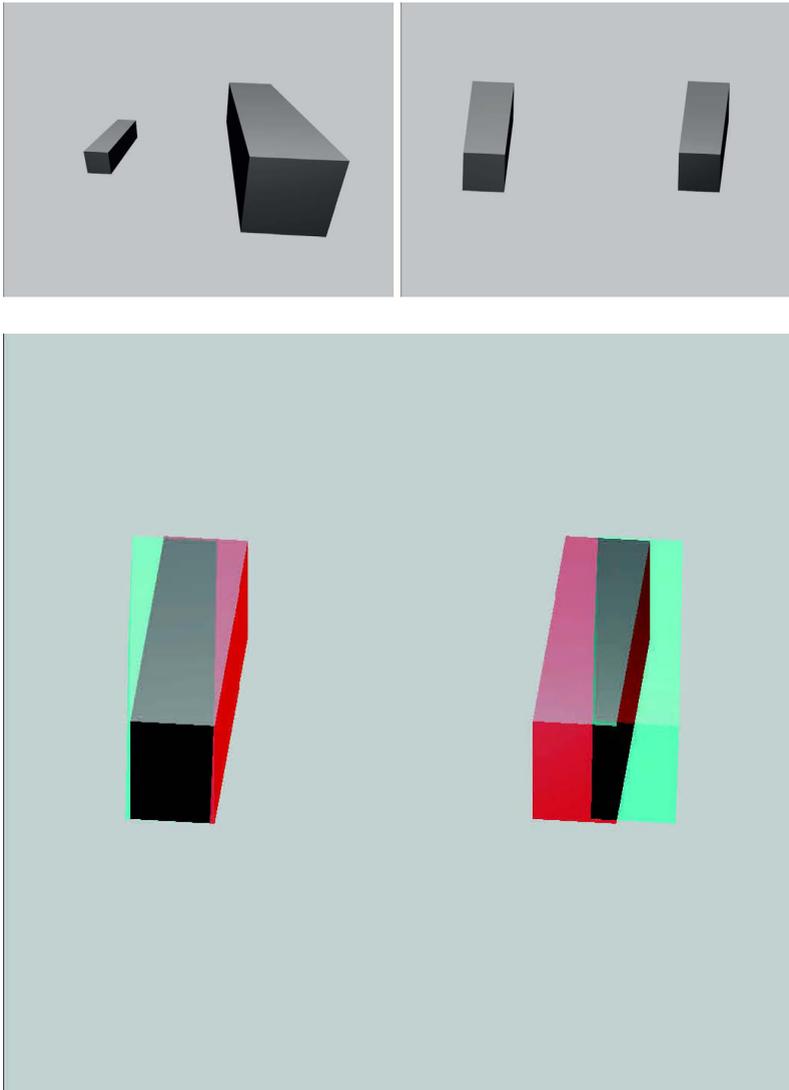


Abbildung 2: zwei exakt gleich große Quader, die unterschiedlich im Raum angeordnet sind. In der perspektivischen Ansicht (a) sieht man durch die unterschiedliche Größe direkt, welcher Quader näher ist. In der orthogonalen Ansicht (b), die die gleiche Szene zeigt, sind beide Quader aufgrund der planaren Projektion gleich abgebildet. Erst die stereoskopische Ansicht (c) zeigt Unterschiede in der Parallaxe.

- **Verdeckung:**

Sie stellt den zweiten kritischen Punkt der Tiefenreize dar. Da die Verdeckung bei ihrem Auftreten immer eine erkennbare Tiefenreihenfolge erzeugt, ist die Verhinderung dieses Reizes besonders wichtig. Grundsätzlich lassen sich Verdeckungen über zwei Eingriffe eliminieren:

Visuelle Einschränkungen: Werden Gegenstände z. B. nur als Drahtgittermodell oder durchsichtig angezeigt (s. Abb. 3), wird es schwierig zu erkennen, welches Objekt vorne und welches hinten ist.

Spielmechanische Einschränkungen: Verdeckungen während des Spiels können durch Bewegung oder durch Verschiebung der Objekte stattfinden. Beides kann durch eine Einschränkung im Spiel stattfinden. Kann sich der Spieler z. B. nicht bewegen, können Objekte nicht durch bewegungsinduzierte Verdeckung hinter einem anderen Gegenstand verschwinden. Basiert das Spiel darauf, Objekte zu bewegen, ist sicherzustellen, dass Objekte nicht hintereinander angeordnet werden können. Durch die Einschränkung der Freiheit des Spielers, indem bestimmte Aktionen einfach nicht möglich sind, sind allerdings auch die Möglichkeiten des Spieleentwicklers beschränkt.

- **Relative Höhe:**

Da die relative Höhe abhängig vom Horizont ist, kann dieser Tiefenreiz durch Ausblenden des Horizontes und Vermeidung von horizontalen Linien stattfinden.

- **Relative Größe:**

Hier geht es um die Relationen zwischen gleichen Objekten. Sind ähnliche Objekte nicht gleich groß, ist dieser Tiefenreiz beseitigt (s. Abb. 4).

- **Gewohnte Größe:**

Dieser Tiefenreiz kann durch unverhältnismäßige Größen der Objekte entfernt werden. Ein Mensch kann z. B. als Riese dargestellt werden, wodurch die Relation zu anderen Objekten verloren geht.

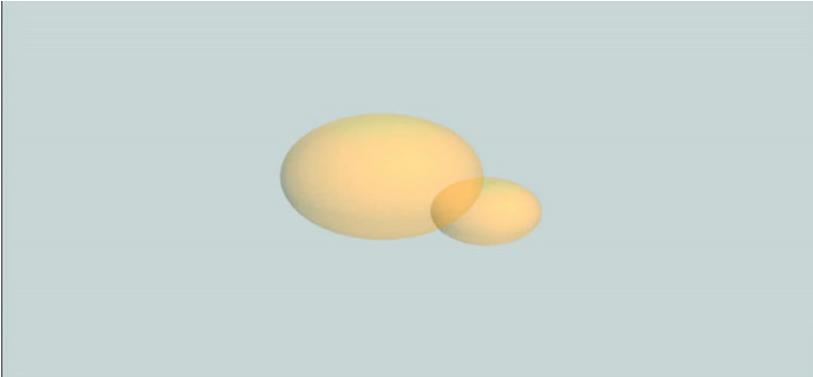


Abbildung 3: zwei leicht durchsichtige Kugeln, die unterschiedlich weit entfernt sind. Auf den ersten Blick ist nicht eindeutig klar, welche Kugel die andere verdeckt.

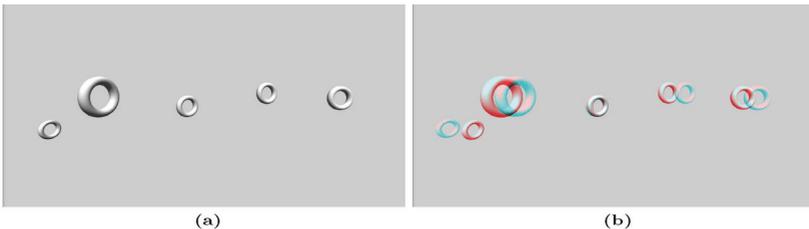


Abbildung 4: unterschiedlich große und unterschiedlich weit entfernte Ringe. Betrachtet man Abbildung (a), nimmt man den größeren Ring als näher wahr, da man vermutet, dass alle Ringe gleich groß sind. Abbildung (b) zeigt, im Anaglyphen-Verfahren betrachtet, dass einige der Ringe weiter entfernt sind, dafür aber wesentlich größer sind.

- **Texturgradient:**

Dieser Tiefenreiz kann durch unregelmäßige Texturen eliminiert werden, bei denen ähnliche Elemente der Textur nicht gleich groß sind. Dies ist auch der Fall, wenn die Elemente der Textur nach hinten größer statt kleiner werden (s. Abb. 4).

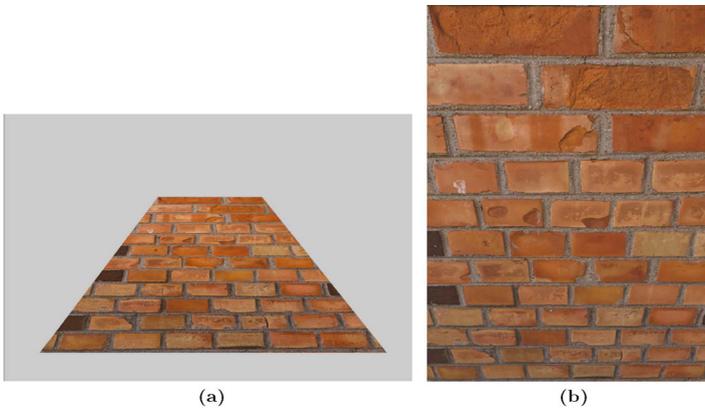


Abbildung 5: Aufgrund der Textur (b) wirkt die nach hinten gekippte Fläche eher wie eine Trapez-förmige Mauer (a).

- **Licht und Schatten:**

Der Schattenwurf von Lichtquellen kann in Spiel-Engines i.d.R. ausgeschaltet werden.

- **Atmosphärische Perspektive:**

Da dieser Reiz in Spielen wenn, dann nur als nachträglicher Effekt angewendet wird und nicht automatisch vorkommt, kann er ignoriert werden.

- **Bewegungsinduzierte Tiefenreize:**

Bewegung ist in solch einer Art von Spielen grundsätzlich problematisch, da dabei immer ein Tiefeneindruck entsteht. Daher sollte Bewegung nach Möglichkeit verhindert werden. Vor allem Rotationen der Kamera um einen bestimmten Punkt vermitteln einen sehr guten Raumeindruck und sollten daher nicht möglich sein. Ist die Kamera allerdings auf orthogonal gestellt, kann sie zumindest in gewissem Maße nach vorne und nach hinten bewegt werden, da durch die fehlende Perspektive kein monokularer Tiefeneindruck erzeugt wird (s. Abb. 6).

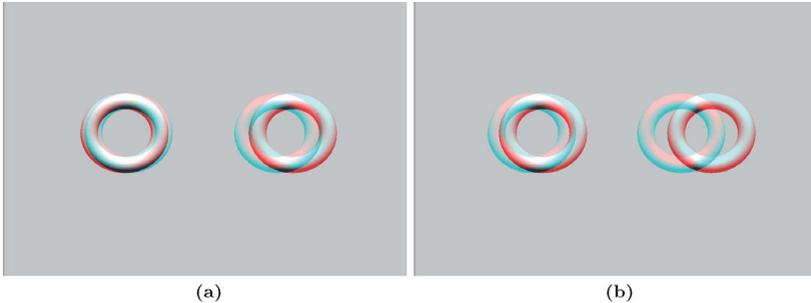


Abbildung 6: zwei unterschiedlich weit entfernte Ringe in der orthogonalen Ansicht. Bewegt sich die Kamera nach hinten (b), ändert sich lediglich die Parallaxe, die Größe der Objekte bleibt gleich.

An dieser Stelle soll allerdings auch darauf hingewiesen werden, dass nicht zwingend alle monokularen Tiefenreize zu jeder Zeit ausgeschaltet sein müssen. Ist es z.B. die Aufgabe des Spielers, den Abstand zwischen zwei Gegenständen zu schätzen, liefert die Verdeckung, abgesehen von der Tiefenreihenfolge, keine Hinweise. Genügen die Informationen aus einem monokularen Tiefenreiz oder aus der Verknüpfung verschiedener Reize nicht aus, um die Tiefen der Objekte in Bezug auf die Aufgabe zu schätzen, können sie vorhanden bleiben. Dies ermöglicht u.U. größere Freiheiten für den Spieler und gibt dem Spieleentwickler mehr Möglichkeiten. Die Kreativität des Spieldesigns besteht daher darin, ein Spielprinzip zu finden, mit dem möglichst viele Freiheiten erhalten werden können. Selbst ohne monokulare Tiefenreize besteht allerdings die Gefahr, dass dem Spieler mittels Ausprobieren die Lösung der Aufgabe auch in der monoskopischen Darstellung gelingt. Dies kann allerdings über verschiedene Mechanismen wie Punktabzüge beim Verfehlen von Objekten oder Zeitdruck verhindert werden.

2.4 Eingabegeräte und Navigation im Raum

Interessant im Zusammenhang mit stereoskopisch dargestellten Spielen ist die Frage der Eingabegeräte, denn weder Maus noch Tastatur sind eigentlich für die Bewegung im dreidimensionalen Raum ausgelegt. Während die Maus zwei Achsen abdecken kann, ist die Tastatur eigentlich nur eindimensional. Erst die geschickte Kombination aus beiden macht eine Bewegung im Raum möglich. In Spielen, in denen die Ego-Perspektive gewählt wurde, hat sich mittlerweile eine Steuerung durchgesetzt, bei der die Tastatur zur Bewegung des Charakters verwendet wird, während die Maus die Kamerabewegung steuert. Die rechte und linke Maustaste sowie die nicht verwendeten Tasten der Tastatur können dann genutzt werden, um Objekte in der Spielwelt zu manipulieren oder Aktionen auszulösen. Für andere Spieltypen müssen allerdings

besondere Eingabevarianten entwickelt werden, was unter Umständen viel Zeit im Rahmen des Spieldesigns in Anspruch nehmen kann und abhängig von der Menge der Aktionen ist, die der Spieler ausführen kann.

Sollen bereits weitverbreitete Eingabegeräte wie Maus, Tastatur, Touchscreens, Gamepads oder Joysticks verwendet werden, müssen daher neue Verarbeitungsvarianten der Eingaben entwickelt werden, um eine direkte Manipulation von Objekten im Raum zu ermöglichen. Im Zusammenhang mit stereoskopischen Anwendungen ist interessant, dass die Stereoskopie vor allem die Genauigkeit der Bestimmung von Positionen von Objekten und deren Lage im Raum und somit die Raumorientierung wesentlich verbessert.⁸ Dabei muss bedacht werden, dass neue Eingabemethoden vom Benutzer erst erlernt werden müssen und diese auch den Umgang mit dem Spiel vollkommen ändern.

Mauszeiger und Fadenkreuze werden i.d.R. auf der Bildschirmenebene angezeigt. Dies ist im Fall von stereoskopischen Inhalten problematisch, da sich der Mauszeiger z. B. beim Markieren eines Objektes auf der Nullebene befindet, während das zu markierende Objekt im Raum weiter hinten angeordnet ist. Sieht der Betrachter das zu markierende Objekt an, sieht er den Mauszeiger doppelt, da sich dieser auf einer anderen Tiefenebene befindet und umgekehrt. Diese Anzeigen sollten sich daher ebenfalls auf der Tiefenebene des Objektes befinden, auf das gezielt oder das markiert werden soll. Es ist daher sinnvoll, die Tiefe des Objektes zu ermitteln, auf das ein Fadenkreuz zeigt und eine Art Laser-Markierung auf der Oberfläche des Objektes anzuzeigen. Dies kann z. B. mithilfe eines Raycasts von der Kamera zu dem zu treffenden Ziel geschehen, dessen Länge dann die Tiefe des Objektes angibt.⁹

Wichtig ist die frühzeitige Auseinandersetzung mit dem Mauszeiger vor allem in Zusammenhang mit Selektionen im Raum, also z. B. bei Strategiespielen, in denen oft mehrere Einheiten zusammen ausgewählt werden müssen. Da der normale Mauszeiger nicht verwendet werden kann und zwei Teilbilder angezeigt werden, die jeweils eine andere Sichtweise der Selektion repräsentieren, wird ein neuer Mechanismus benötigt, um die korrekten Pixel zu finden. Details dazu finden sich in Gateau (25.03.2009): *The In and Out: Making Games Play Right with Stereoscopic 3D Technologies* und Gateau/ Neuman (16.12.2010): *Stereoscopy, From XY to Z*.

Mit raumorientierten Eingabegeräten, wie der Nintendo *Wii-Remote*, Sony *Playstation Move* oder Microsofts *Kinect* halten seit 2008 Eingabegeräte in den Haushalten Einzug, die alle drei Achsen des Raumes abdecken und damit neuen Ansätzen Platz bieten. Daher sind gerade diese Steuerungen für stereoskopische Spiele interessant. Einer der größten Vorteile dieser raumorientierten Eingabegeräte ist, dass sie eine recht genaue Positionierung von Objekten und eine gute Orientierung des Spielers ermöglichen, da er hier, im Gegensatz zu anderen Eingabegeräten, die gleichen Bewe-

8 s. Hubona et al. (1999): ACM Trans. Comput.-Hum. Interact., Nr. 3, Bd. 6., 9 Vgl. Moore (2011): Basics of game design, S. 325ff.

9 Vgl.: NVIDIA Corporation (2006): NVIDIA GPU Programming Guide, S. 70f und McDonald/ Skaljok (2011): NVIDIA 3D Vision Automatic: Best Practices Guide, S. 20.

gungen wie das Objekt auf den gleichen räumlichen Achsen durchführt.¹⁰

Theoretisch ist dabei sogar eine Interaktion zwischen physischen und virtuellen Gegenständen möglich. Aufgrund der unterschiedlichen Entfernungen müssen die Augen allerdings im Wechsel zwischen den beiden Objekten fokussieren, was anstrengend bis irritierend sein kann und u.U. sogar den Tiefeneindruck des virtuellen Objektes zerstören kann.

Aufgrund der besseren Tiefenwahrnehmung können sogar einige Mechanismen der Steuerung grundlegend verändert werden. So werden in Anwendungen, in denen Objekte im Raum positioniert werden sollen, oftmals Hilfsmechanismen verwendet, die Objekte entlang anderer Objekte positionieren, da die genaue Lage sonst nicht ersichtlich wäre. Dabei werden die Informationen eines zweidimensionalen Eingabegerätes auf eine dreidimensionale Welt projiziert.¹¹ Diese Hilfen könnten durch die bessere Tiefenwahrnehmung weggelassen werden und eine präzise Steuerung würde eine gute Positionierung ermöglichen.

2.5 Unterschiede der stereoskopische Rendermethoden

Bevor es an die eigentliche technische Umsetzung eines Spieles geht, ist die Wahl der Rendermethode¹² für die stereoskopische Darstellung sehr wichtig. Sie entscheidet maßgeblich darüber, wie stark sich der Entwickler mit der Stereoskopie auseinandersetzen muss.

Dabei gibt es zwei Varianten: das *Duale Rendern* und *Post Stereo*. Während das duale Rendern das aktive Erzeugen zweier Teilbilder durch zwei virtuelle Kameras der Spiel-Engine umfasst, wird Post Stereo nachträglich auf den monoskopischen Output einer virtuellen Kamera angewandt.¹³

2.5.1 Duales Rendern

Beim dualen Rendern wird jede Szene des Spiels zweimal gerendert, je einmal für das linke und das rechte Auge. Dazu wird die Szene von zwei virtuellen Kameras aufgenommen, die jeweils eine leicht andere Perspektive für jedes Auge aufnehmen. Im Fall computergenerierter stereoskopischer Inhalte bedeutet dies, dass jedes Bild einer virtuellen Szene zweimal gerendert werden muss, wodurch doppelt so viel Rechenleistung wie bei der monoskopischen Darstellung aufgewendet werden muss.

¹⁰ Vgl. Moore (2011): Basics of game design, S. 325ff.

¹¹ s. Hanson/Wernert (1997): Constrained 3D navigation with 2D controllers.

¹² Rendern bezeichnet den Vorgang, in dem die Grafikkarte aus Rohdaten über eine große Anzahl an Berechnungen (u.a. Oberflächenstruktur, Material, Farbe, Licht und Schatten) ein möglichst realitätsnahes Bild erzeugt (s. <http://www.itwissen.info/definition/lexikon/Renderingrendering.html>).

¹³ Vgl. Gouraud (2011): A Developer's Guide To Stereoscopic 3D In Games.

Die wichtigsten Parameter, die dabei eingestellt werden müssen, sind – wie beim Stereo-3D-Realfilm - neben der Stereobasis auch der Konvergenzwinkel der Kameras. Dabei lassen sich drei Varianten unterscheiden:

Parallele Kameras: Die Kameras sind nebeneinander angeordnet und die Sichtachsen liegen parallel zueinander (s. Abb. 7). Dabei gehen links und rechts Teile des Bildes verloren, da sich dort die Sichtbereiche der Kameras nicht decken.

Konvergente Kameras: Ähnlich wie das menschliche Auge konvergieren kann, werden hierbei die Kameras zueinander eingedreht (s. Abb. 7). Zwar ist der Sichtbereich beider Kameras dann fast deckungsgleich, allerdings wird das Bild zum Rand hin verzerrt. Um die Bilder richtig wiedergeben zu können, müssen diese in den meisten Fällen im Nachhinein entzerrt werden.

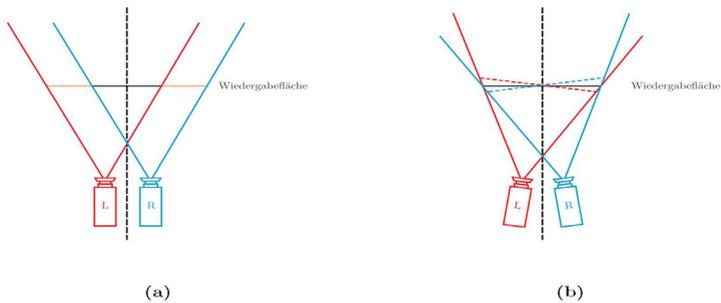


Abbildung 7: Bei einer parallelen Anordnung der Kameras (a) sind die orange markierten Bereiche der Wiedergabefläche nur für ein Auge zu sehen. Bei konvergenten Kameras (b) kommt es bei größeren Parallaxen zu starken Verzerrungen des Bildes.

Frustum Shift/HIT: Da die Bilderzeugung am Computer geschieht und somit virtuell stattfindet, ist dies ein weiteres mögliches Verfahren, das die Vorteile der beiden oben genannten Vorgehensweisen verbindet. Hierbei werden die Kameras parallel zueinander platziert, allerdings wird das Frustum¹⁴ beider Kameras verschoben (s. Abb. 8). Dies bedeutet, dass das optische Zentrum der Kameras so zueinander bewegt wird, dass beide Kameras den gleichen Bildbereich abdecken.¹⁵

14 Mathematische Abbildung einer 3D-Umgebung in computergenerierten Bildern (s. <http://wiki.delphigl.com/index.php/Frustum>)

15 Vgl.: Taube (2011): Stereoskopie für Spieleentwickler und Gateau/Neuman (16.12.2010): Stereoscopy, From XY to Z.

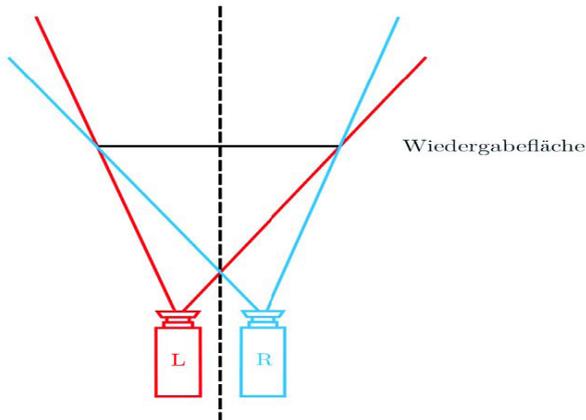


Abbildung 8: Beim Frustum Shift werden die Kameras parallel positioniert und das Frustum wird so verschoben, dass gemeinsame Bildbereiche sichtbar sind.

Bei allen drei Varianten ist zu beachten, dass die Kameras nur horizontal zueinander verschoben sein dürfen. Vertikale Verschiebungen sowie Rotationen um die y- oder z-Achsen der beiden Kameras dürfen nicht auftreten, da sie den stereoskopischen Effekt negativ beeinflussen (s. Abb. 9).¹⁶

¹⁶ Vgl. Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 51.

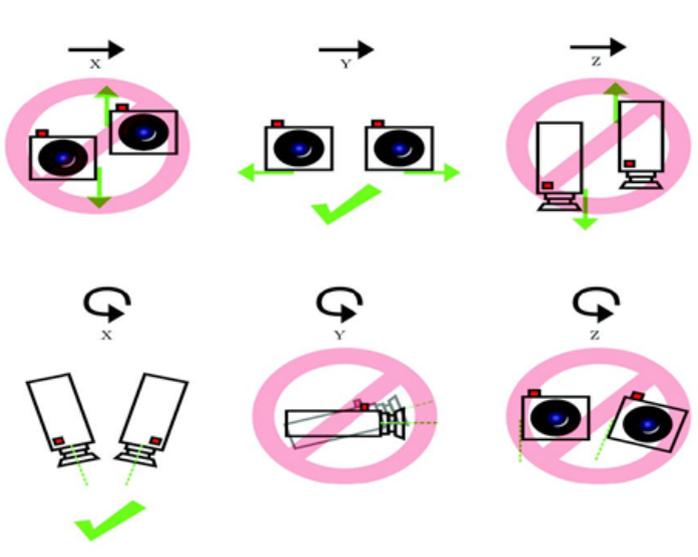


Abbildung 9: Von den sechs möglichen Drehungen und Verschiebungen der beiden Kameras zueinander sind nur zwei wirklich anwendbar.

2.5.2 Post Stereo

Die meisten virtuellen Welten, vor allem im Bereich der Spiele, liegen bereits als dreidimensionale Geometrie vor, die von der Grafikkarte gerendert wird. Die Tiefenordnung der Objekte ist daher bereits bekannt. Diesen Fakt nutzt das Post Stereo-Verfahren, indem die Erzeugung des stereoskopischen Bildes erst bei der Darstellung durch die Grafikkarte stattfindet. Dadurch muss der Programmierer diesem Entwicklungsprozess zunächst nicht berücksichtigen. Lediglich die richtige Platzierung der Objekte innerhalb der Geometrie muss sichergestellt werden.¹⁷

Bei Post Stereo wird das Bild einer virtuellen Kamera zunächst wie bisher gerendert und dann auf Basis der Tiefeninformationen des z-Buffers¹⁸ und eines angenommenen Augenabstands des Betrachters durch Verschiebung der Pixel das zweite Bild für das andere Auge erzeugt.

Einer der größten Vorteile dieses Verfahrens besteht darin, dass im Gegensatz zum zweifachen Rendern wesentlich weniger Rechenleistung verbraucht wird, die der Echtzeitfähigkeit der Engine im Sinne unverzögerter Reaktionen zugutekommt.

17 Vgl. Johnson (2012): The CryENGINE Sandbox Manual.

18 Beim Rendern von Objekten speichert die Grafikkarte die Tiefeninformationen der erzeugten Pixel im z-Buffer. Innerhalb des z-Buffers ist also für jeden Punkt auf der Anzeigefläche die Tiefe hinterlegt (s. http://www.at-mix.de/z_buffer.htm).

Die zurzeit am weitesten verbreitete Variante dieser Technik stammt von NVIDIA und nennt sich *3D Vision*. Dabei handelt es sich um einen Treiber, der diese Technik umsetzt und es dem Benutzer erlaubt, verschiedene Parameter wie absolute und relative Parallaxe nachträglich zu ändern. Dabei wird lediglich die Verschiebung der Pixel nach dem Rendern verstärkt oder abgeschwächt. Unter Verwendung verschiedener Algorithmen entscheidet der Treiber, welche Objekte für jedes Auge separat gerendert werden müssen und welche nicht, wodurch viele Ressourcen gespart werden können. Zusätzlich ist die Grafikkarte auch für die Umwandlung in verschiedene Formate und Wiedergabesysteme verantwortlich. Dabei werden neben den mitgelieferten Shutter-Brillen auch andere Wiedergabevarianten wie Anaglyphenbrillen unterstützt.¹⁹

Auch verschiedene Spiel-Engines, wie z. B. die *CryEngine*²⁰ und andere Treiber wie *TriDef* oder *iZ3D* setzten diese Technik ähnlich um, wodurch es dem Entwickler möglich wird, das Spiel stereoskopisch darstellen zu lassen, ohne eine Zeile Code zu schreiben.

2.5.3 Bewusste Wahl der Rendermethode

Die einfachste Variante ist sicherlich Post Stereo. Da hier das stereoskopische Bild erst nach dem Rendern erzeugt wird und die stereoskopischen Einstellungen außerhalb der Anwendung getätigt werden, muss sich der Spiele-Entwickler hauptsächlich mit der Optimierung der Szene für die stereoskopische Ansicht und den zulässigen Brennweiten im Viewing-Frustum beschäftigen.

Bei dem dualen Rendern ist dies anders. Hier kommen zwei Kameras zur Anwendung. Dadurch ergibt sich für den Entwickler neben der Möglichkeit zur Einstellung der Brennweite, der Stereobasis und des Konvergenzwinkels der Kameras auch eine wesentlich größere Kontrolle über den stereoskopischen Effekt, z. B. durch die starke Verwendung der negativen Parallaxe.

Eine weitere Entscheidung in diesem Zusammenhang ist die verwendete Spiel-Engine. Die Post Render-Methode von NVIDIAs *3D Vision* funktioniert z. B. nur für DirectX-basierte Anwendung²¹, ist daher allerdings für eine breite Anzahl an Spiel-Engines geeignet. Der Nachteil von *3D Vision* besteht aber darin, dass diese Technik ausschließlich auf dem PC mit entsprechender NVIDIA-Grafikkarte eingesetzt werden kann.

Möchte man das duale Rendern verwenden, ist sicherzustellen, dass die entsprechende Spiel-Engine die Möglichkeit besitzt, zwei Kameras zu verwenden. Die Bilder der beiden Kameras müssen dann separat gerendert und (je nach Ausgabeformat) miteinander verrechnet werden. Wird dies nicht ausdrücklich von der Spiel-Engine

19 Vgl. McDonald/Skaljak (2011): *NVIDIA 3D Vision Automatic: Best Practices Guide*, S. 10.

20 s. Johnson (2012): *The CryENGINE Sandbox Manual*.

21 OpenGL wird bisher nur bei Tegra-basierten Geräten unterstützt (s. McDonald/Skaljak (2011): *NVIDIA 3D Vision Automatic: Best Practices Guide*, S. 7).

unterstützt, müssen eigene Shader²² entwickelt werden, die über verschiedene Algorithmen die passende Ausgabe erzeugen. Da ein direkter Zugriff des Entwicklers aus der Spiel-Engine auf die Ausgabe der Grafikkarte i.d.R. nicht möglich ist, werden in diesem Fall meistens Anaglyphen als Wiedergabemethode verwendet, da hier die Farb-Kanäle der Kameras leicht über einen Shader in ein Bild zusammengefügt werden können. Das zweifache Rendern geht zulasten der Performance, vor allem dann, wenn die Spiele-Engine nicht dazu in der Lage ist zu entscheiden, ob Komponenten für beide Augen gleich sind (z. B. Texturen, Shadow-Maps etc.) oder nicht.²³

Interessante Ansätze bietet das *CryEngine 3 SDK*. Dabei sind beide vorgestellten Render-Methoden innerhalb der Spiel-Engine implementiert, sodass auch Konsolen, die nicht über einen Treiber zum Post Rendern verfügen, eine solche stereoskopische Wiedergabe erhalten. Zusätzlich werden in beiden Modi viele Wiedergabemethoden unterstützt.²⁴

3. Gestalterische Herausforderungen und Probleme

Das hauptsächliche Problem stereoskopischer Anwendungen am Computer sind gerenderte Teilbilder für linkes und rechtes Auge, welche in keiner Weise die stereoskopischen Wahrnehmungsgrenzen ihrer Betrachter berücksichtigen. Dabei kommt es im schlimmsten Fall zu schmerzhaften Betrachtungszuständen oder zu binokularen Rivalitäten.

Manche wahrnehmungsspezifische Lösungen und Erfahrungen aus der stereoskopischen Fotografie und Filmproduktion können sofort für stereoskopische Spiele übernommen werden, andere Probleme treten nur im Game-Kontext auf, wie z. B. die Wechselwirkungen in einer animierten Umgebung in der sich zusätzlich die Spieler frei bewegen können.

Besonders wichtig ist auch das Testen auf unterschiedlichen Wiedergabegeräten, da jedes seine technischen Spezifika und Rezeptions-Eigenheiten hat und das Spiel entsprechend angepasst werden muss.

Da sich diese Art von Spielen noch in der Entwicklung befindet, ist davon auszugehen, dass noch eine ganze Reihe an anderen Problemen auftreten kann. Es kristallisieren sich aber schon heute zentrale Randbedingungen des stereoskopischen Spieldesigns und der Spielmechanik heraus. Befolgt man diese, ist die grundlegende Stereo-3D-Funktionalität des Spiels erst einmal gewährleistet. Viele davon kommen dem erfahrenen Stereo-3D-Leser vielleicht bekannt vor.

22 Ein Shader ist ein Software-Modul, das Informationen darüber enthält, wie die Grafikkarte bestimmte Bildteile rendern soll (s.<http://compactiongames.about.com/od/glossary/g/Shader.htm>).

23 Vgl. McDonald/Skaljak (2011): NVIDIA 3D Vision Automatic: Best Practices Guide, S. 9.

24 Vgl. Johnson (2012): The CryENGINE Sandbox Manual.

3.1 Immer alles tief?

Im stereoskopischen Film ist es nicht üblich, durchgängig die maximal mögliche Tiefenausdehnung zu verwenden. Es wird nur an bestimmten Schlüsselstellen ein hoher Tiefenumfang verwendet, sonst bleibt die Tiefe eher gemäßigt. Daraus ergeben sich für den Zuschauer Sequenzen mit entspannenden „Augenpausen“.

Auch für stereoskopische Spiele ergibt sich die Erkenntnis, dass der stereoskopische Effekt nicht durchgängig im vollen Umfang verwendet werden sollte, sondern zeitweise auch so gering zum Einsatz kommt, dass sich der Betrachter entspannen kann. Der konkrete Vorteil liegt darin, dass Szenen mit stärkerer Tiefenintensivität dann wieder wesentlich intensiver zur Geltung kommen.²⁵

Des Weiteren erhält man damit zusätzlich die gestalterische Freiheit, dass Tiefe nicht nur unter immersiven, sondern auch unter dramaturgischen Gesichtspunkten gezielt als Stilmittel eingesetzt werden kann.²⁶

3.2 Rahmenverletzungen und Schwebefenster

Die Displayfläche wirkt bei Objekten mit positiver Parallaxe wie ein Fenster, durch das man blickt: Am linken Rand sieht dabei das rechte Auge mehr als das linke Auge. Bei Objekten mit negativer Parallaxe ist dies genau anders herum: Das rechte Auge sieht beim Betrachten des linken Displayrandes weniger des Bildes als das linke Auge (s. Abb. 10). Dies kommt beim natürlichen Sehen nicht vor und führt daher bei der stereoskopischen Wahrnehmung zu Irritationen.

Die Tiefenwahrnehmung, die dabei auftritt, ist widersprüchlich. Teile des Objektes werden vom vertikalen Rand des Wiedergabegerätes verdeckt, weshalb das visuelle System annimmt, der Gegenstand befände sich hinter der Wiedergabefläche. Der Teil, der von beiden Augen gesehen wird, vermittelt über die Parallaxe allerdings den Eindruck, sich vor der Wiedergabefläche zu befinden.

²⁵ Vgl. Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation 3.

²⁶ Vgl. Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S. 272f.



Abbildung 10: Am linken Bildrand wird beim Betrachten mit einer Anaglyphen-Brille deutlich, dass das rechte Auge bei Objekten mit negativer Parallaxe weniger sieht als das linke Auge. Es kommt zu Rahmenverletzungen.

Diese Problematik ist allerdings nur bei Objekten mit negativer Parallaxe störend, die sich längere Zeit im Bereich der vertikalen Ränder des Wiedergabegerätes befinden. Wenn sich die Objekte schnell durchs Bild bewegen, fällt dieser Effekt kaum auf. Dies liegt daran, dass das Gehirn bei neu auftauchenden Objekten eine längere Zeit benötigt, um sie in der Tiefe einzuordnen.

Unproblematisch sind dabei auch Objekte, die vom horizontalen Rahmen des Wiedergabegerätes teilweise verdeckt werden, da hier beide Augen die Parallaxe des Ausschnittes sehen. Das visuelle System vermittelt dabei allerdings den Eindruck, das Stereofenster sei zum Betrachter hin gebogen.²⁷

Eine Möglichkeit, Rahmenverletzungen im Bereich der vertikalen Ränder zu verhindern, ist das sog. *Schwebefenster*. Dabei wird jeweils für das Auge, das mehr des Gesamtbildes sieht, der Zusatz an Bildinformationen mithilfe eines schwarzen Balkens entfernt, sodass beide Augen wieder die gleichen Bereiche sehen (s. Abb. 11). Das Stereofenster wird damit nach vorne verlagert und scheint vor der Displayfläche zu schweben. Dieser Effekt wird vom Betrachter kaum bemerkt und kann sogar animiert oder asymmetrisch angewendet werden, ohne dass es störend ins Gewicht fällt.

Da es auch möglich ist, nur die eine Seite des Bildes zu beschneiden, auf der die Rahmenverletzung stattfindet, kann sich auch ein Stereofenster ergeben, welches schräg im Raum steht. Da bei dieser Technik das Stereofenster von der Nullebene entkop-

27 Vgl.: Lipton (1997): *StereoGraphics Developers' Handbook*, S. 18 und Mendiburu (2009): *3D movie making: Stereoscopic digital cinema from script to screen*, S. 80.

pelt wird, ist es zudem möglich, das Stereofenster weiter in den Betrachter-Raum zu ziehen, wodurch Objekte, die kurz hinter dem Stereofenster liegen trotzdem vor der Nullebene zu schweben scheinen.²⁸

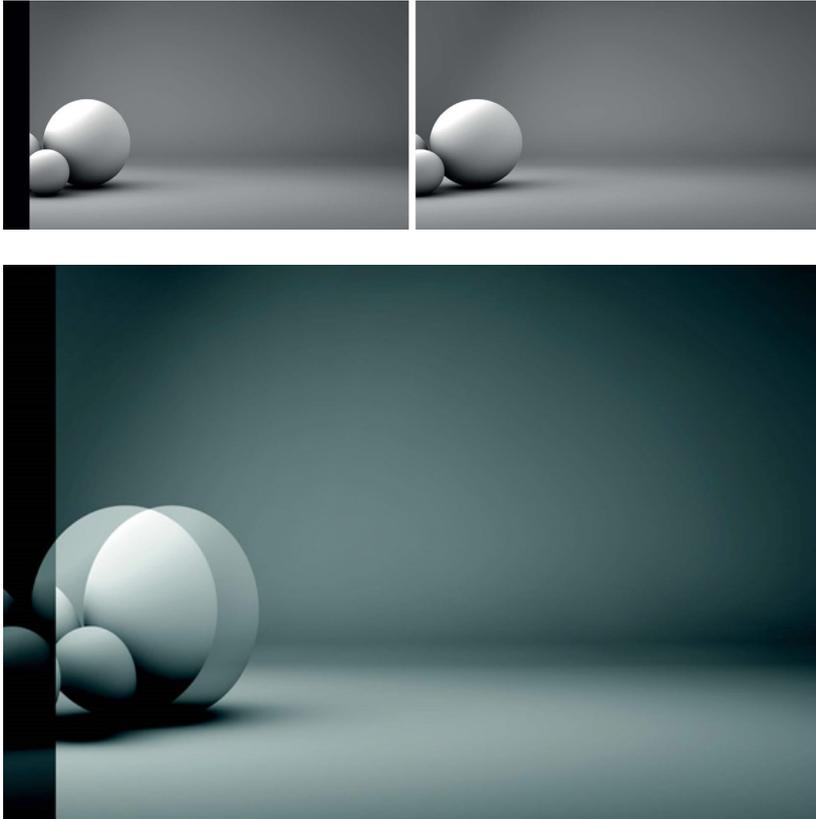


Abbildung 11: linkes maskiertes Teilbild (a), rechtes Teilbild (b) und resultierendes Bild mit Schwebefenster (c).

3.3 Negative Parallaxen / Pop-Outs

Eines der größten Probleme des Post Stereo-Verfahrens ist es, dass die beiden Teilbilder für die Augen nicht wirklich aus zwei Blickwinkeln aufgenommen wurden, sondern nur durch den Versatz von Pixeln nach dem Render-Prozess entstehen. Vor allem bei Objekten, die sich nah an der Kamera befinden, fehlt damit die korrekte Parallaxe.

28 Vgl.: Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S. 264ff und Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 80ff.

Dieser Effekt ist nur in Bereichen der negativen Parallaxe (Pop-Outs) wirklich sichtbar. Aus diesem Grund wird bei vielen heutigen Spielen, die stereoskopisch nutzbar sind, keine negative Parallaxe verwendet (z. B. *Crysis 2*) und die Entwickler von NVIDIA und Sony weisen darauf hin, Objekte möglichst „hinter“ der Wiedergabefläche zu platzieren. Dies führt allerdings dazu, dass ein wichtiger Bestandteil stereoskopischer Möglichkeiten in solchen Spielen keine oder nur geringe Anwendung findet.²⁹

Dass es allerdings durchaus möglich ist, Elemente mit negativer Parallaxe zu verwenden, zeigen Spiele wie *Trine 2* (s. Abb. 12). Hier hat der Spieler die Möglichkeit, die Stereobasis und die Tiefe des Stereofensters zu beeinflussen. Bei bestimmten Einstellungen werden dann hauptsächlich Partikel mit negativer Parallaxe angezeigt.³⁰



Abbildung 12: *Trine 2* ermöglicht es, innerhalb des Spiels verschiedene Stereoskopie-bezogene Einstellungen vorzunehmen, sodass auch Objekte mit negativer Parallaxe auftreten. Abbildung von http://www.mtbs3d.com/index.php?option=com_content&view=article&id=13120:trine2&catid=38:game-reviews&Itemid=76.

Eine Möglichkeit der Anpassung ist dabei das variable Setzen der Konvergenz. Dies kann durch Verschieben der Teilbilder zueinander geschehen, wobei die Parallaxe je nach Richtung verändert und damit der Konvergenzpunkt verschoben wird.

In einem Spiel, in dem die Kamera starr auf die Szene gerichtet ist, stellen Objekte mit negativer Parallaxe ein geringeres Problem dar als z. B. bei Action-Spielen, in

²⁹ Vgl.: Gateau/Neuman (16.12.2010): Stereoscopy, From XY to Z; Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation 3 und Taube (2011): Stereoskopie für Spieleentwickler

³⁰ Vgl. Hernandez (2012): *Trine 2* 3D Review.

denen der Spieler die Kamera selbst steuern kann. In diesem Zusammenhang kann es jederzeit zu Rahmenverletzungen durch Objekte im Vordergrund kommen, die vom Entwickler nicht direkt vorhersehbar sind. Auch dies ist ein Grund, warum viele Spiele die negative Parallaxe gar nicht oder nur in Zwischensequenzen einsetzen, die vom Spieler nicht steuerbar sind.

Das duale Rendern hat den Vorteil, dass die Teilbilder tatsächlich aus unterschiedlichen Blickwinkeln gerendert wurden und demnach die Parallaxe auch im Bereich vor dem Stereofenster korrekt ist. Dies verhindert allerdings nicht die Rahmenverletzungen, weshalb diese bei beiden Render-Varianten beachtet werden müssen.

Grundsätzliches Problem aller Stereo-3D-Darstellungsansätze ist, dass beim Betrachter stereoskopischer Inhalte Akkommodation und Konvergenz, anders als beim natürlichen Sehen, nicht aneinander gekoppelt sind. Wird der Unterschied zwischen Konvergenzpunkt und Akkommodationspunkt (Wiedergabefläche) zu groß, ist es nicht mehr möglich, das Bild komfortabel zu betrachten. Da dies bei Objekten mit großer negativer Parallaxe besonders der Fall ist, sollten Pop-Outs generell nur sehr kurzzeitig angewendet werden.

Sollen im Bildgeschehen Objekte mit negativer Parallaxe verwendet werden, dann sollte dies hauptsächlich für schnell durch das Bild fliegende Gegenstände und nur gelegentlich genutzt werden. Wenn sich ein Objekt aus dem Bildschirm heraus bewegen soll, sollte der Gegenstand am besten schon im Bereich der positiven Parallaxe sichtbar sein, damit sich das visuelle System darauf einrichten kann.³¹

Zu beachten ist auch, dass es bei der Game-Interaktion mit echten Gegenständen (z. B. Controller, Lenkrad, etc.) zu Akkommodationsproblemen im Sichtfeld kommen kann. Im Gegensatz zu den stereoskopisch dargestellten Objekten liegt bei ihnen der Akkommodation dort, wo sich der Controller befindet, also außerhalb des Monitors. Der echte Gegenstand und das stereoskopisch dargestellte Objekt können daher nicht gleichzeitig scharf fokussiert werden. Dies ist im Zusammenhang mit neuen Eingabegeräten wie der Kinect von Microsoft durchaus problematisch, da z. B. die Hände des Benutzers ein stereoskopisch dargestelltes Objekt kaum ohne visuelle Probleme manipulieren können.

3.4 Seiteneffekte überproportionaler Stereobasen

Die Stereobasis ist ein wichtiger Faktor der Stereoskopie, wobei der Augenabstand die Bezugsgröße darstellt. Der durchschnittliche Augenabstand beim Menschen beträgt 6,3 cm, ist aber bei jedem Menschen individuell unterschiedlich. Daher sollte die Stereobasis mithilfe eines Schiebereglers für den Benutzer einstellbar sein, um je nach Bildschirmgröße, Distanz zum Wiedergabegerät und persönlichen Vorlieben andere Werte einstellen zu können.

31 Vgl. McDonald/Skaljak (2011): NVIDIA 3D Vision Automatic: Best Practices Guide, S. 25.

Werden Aufnahmen mit sehr großer Stereobasis (Großbasis) gemacht, wirken Bilder meist wie Modellansichten (Modelleffekt, s. Abb. 13), während Aufnahmen mit kleiner Stereobasis (Kleinbasis) alles viel größer erscheinen lassen (Gigantismus, s. Abb. 14). Dies liegt daran, dass das visuelle System als Grundlage aller Berechnungen immer von dem tatsächlichen Augenabstand des Betrachters ausgeht. Bilder, die mit einer Großbasis aufgenommen wurden, täuschen somit den Augenabstand eines Riesen vor. Der Vorteil der Großbasis besteht darin, dass auch Objekte in weiter Entfernung, die sonst nicht mehr stereoskopisch wahrgenommen werden würden, nun einen stereoskopischen Effekt haben.

Die Stereobasis kann je nach gewünschtem Effekt auch während des Spiels immer wieder verändert werden, vor allem wenn man zwischen Szenen mit unterschiedlicher Tiefe wechselt. Wird die Stereobasis über die Zeit verringert, bewegen sich Objekte im Vordergrund weiter weg und werden größer, während Objekte im Hintergrund unverändert bleiben. Umgekehrt verhält es sich bei Vergrößerung der Stereobasis.³²

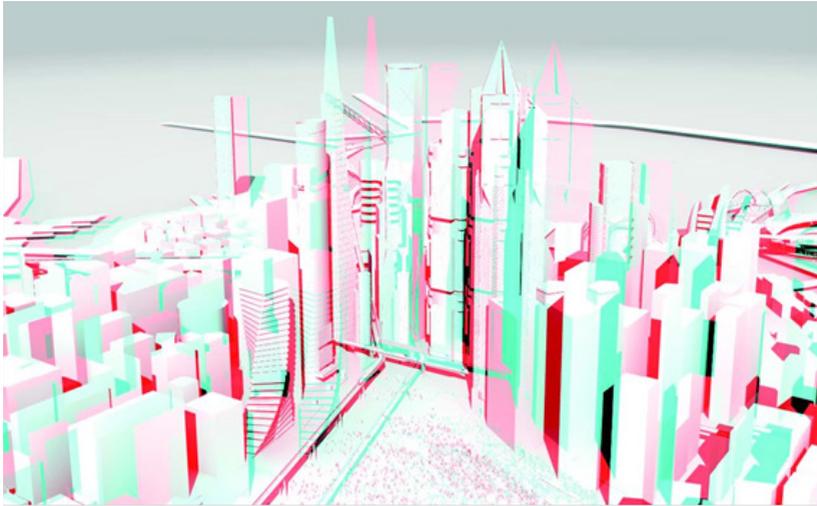


Abbildung 13: Mit großer Stereobasis gerendert wirkt diese Stadt wie ein Spielzeugmodell.

32 Vgl.: Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S. 106, 250, 350ff; Schertenleib (05.11.2010): Optimization for Making Stereoscopic 3D Games on PS3.

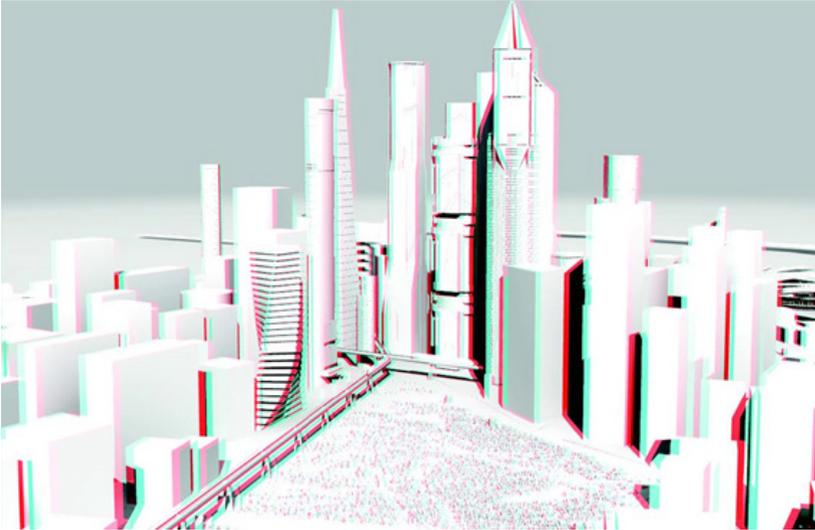


Abbildung 14: Wird die Kamera etwas tiefer gesetzt und die Stereobasis sehr klein eingestellt, werden die Gebäude viel größer wahrgenommen.

3.5 Visuelle Effekte

Über die Jahre wurden viele Tricks entwickelt, um Rechenleistung in Echtzeitspielen zu sparen, ohne die visuelle Qualität zu vermindern. Viele dieser Effekte funktionieren in der monoskopischen Darstellung problemlos, können aber bei stereoskopischer Ansicht deutliche Probleme bereiten.

Hintergrundbilder und Skyboxes sollten so weit wie möglich weg platziert werden. Dies heißt, dass die positive Parallaxe möglichst groß sein muss. Erstens sieht es bei Nichtbeachtung aus, als würde sich die ganze Szene in einer kleinen Box befinden.³³ Zweitens ist in der natürlichen Sehlogik auch erst ab großen Entfernungen kein Volumen mehr erkennbar.

³³ Vgl. NVIDIA Corporation (2006): NVIDIA GPU Programming Guide, S. 70f.

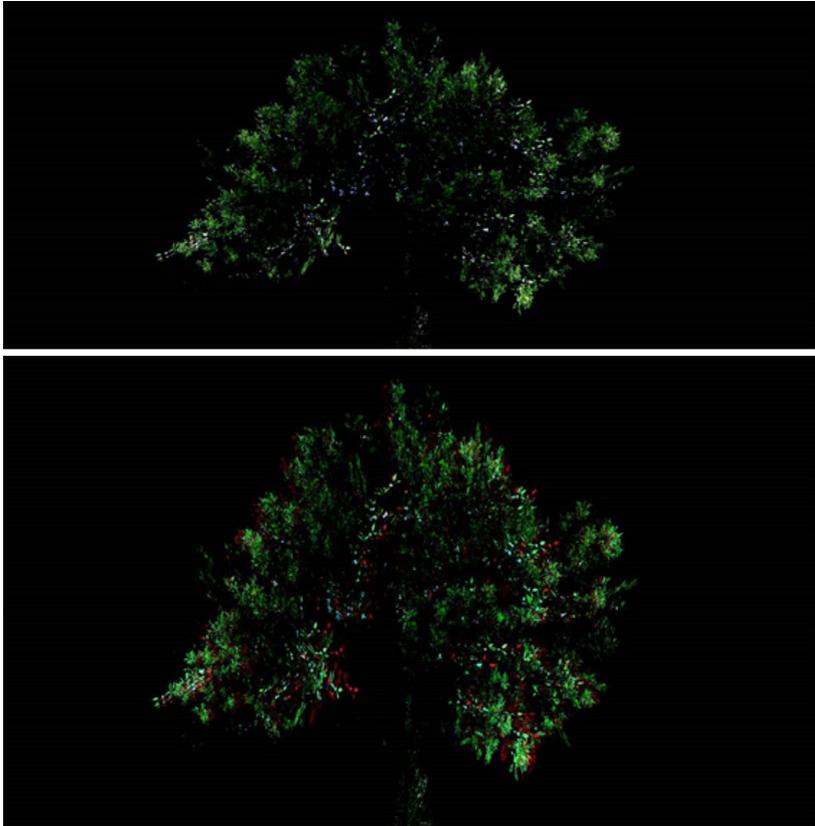


Abbildung 15: Während die Flächen, aus denen die Blätter des Baumes bestehen, in der monoskopischen Ansicht (a) weniger zu erkennen sind, können sie in der stereoskopischen Ansicht (b) deutlich erkannt werden.

Billboards³⁴ sehen, stereoskopisch betrachtet, komplett flach aus, da sie keinerlei Tiefeninformationen enthalten (s. Abb. 15). Da Billboards sich immer zur Kamera drehen und es bei der stereoskopischen Ansicht zwei Kameras gibt, kann es passieren, dass sich die Billboards beim Rendern erst zur Ersten und dann zur zweiten Kamera drehen. Im stereoskopisch betrachteten Spiel entsteht dann ein verzerrtes Bild des Billboards. Grundsätzlich ist es daher besser, echte Geometrie zu verwenden, da dieser Effekt dort nicht auftreten kann.³⁵

34 Planare Flächen, die sich immer einer Kamera zuwenden und so den Eindruck von Geometrie erzeugen, die nicht vorhanden ist.

35 Vgl. Lengyel (2011): Game engine gems 2, S. 173f.

Auch viele Partikel-Effekte (Feuer, Rauch, Staub) verwenden Billboards. Wenn sie entsprechend klein sind, fällt dies unter Umständen nicht ins Gewicht. Eine genaue Vorgabe kann hier nicht stattfinden. Stattdessen wird es zusätzliche Aufgabe der Entwickler, zu beurteilen, ob der Effekt gut genug aussieht.³⁶

Während Lichteffekte wie Highlights oder Glanzlichter am besten aussehen, wenn sie aus monoskopischer Sicht berechnet werden, sollten Reflexionen und Brechungen von Licht, aufgrund der feinen Unterschiede in beiden Augen, auch von zwei Blickwinkeln berechnet werden.³⁷

Problematisch sind auch Blendensterne, die in Spielen gerne eingesetzt werden (z. B. beim Blick gegen die Sonne). Hier sagt die Erfahrung, dass die Sonnenstrahlen aus einer großen Entfernung kommen, sie überdecken aber Objekte im Vordergrund (s. Abb. 16). Dabei kommt es zu Konflikten zwischen monoskopischer und stereoskopischer Wahrnehmung.³⁸

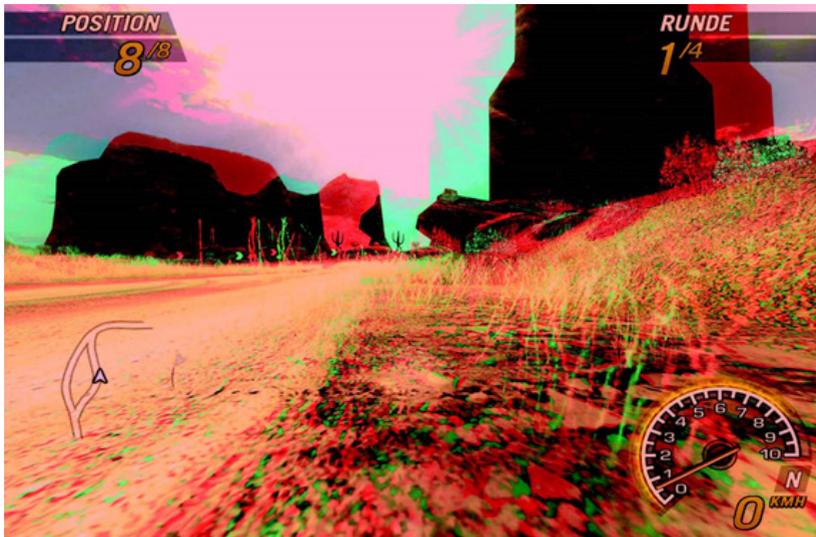


Abbildung 16: Im Spiel Flatout 2 werden öfter Blendeneffekte verwendet. Das Betrachten der Stelle rund um den Felsen führt dabei zu binokularer Rivalität.

36 Vgl. NVIDIA Corporation (2006): NVIDIA GPU Programming Guide, S. 71.

37 Vgl. Gateau/Nash (20.09.2010): Implementing Stereoscopic 3D in Your Applications.

38 Vgl. Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S. 49f.

Große untexturierte Flächen sollten generell vermieden werden, da dort keine Tiefenreize verortbar sind und diese Stellen eher flach erscheinen (s. Abb. 17). Dies liegt daran, dass Parallaxen aufgrund der Einfarbigkeit der Fläche kaum wahrnehmbar sind.

So sollte der Himmel im Spiel z. B. nicht nur blau sein, sondern viele Wolken enthalten.³⁹

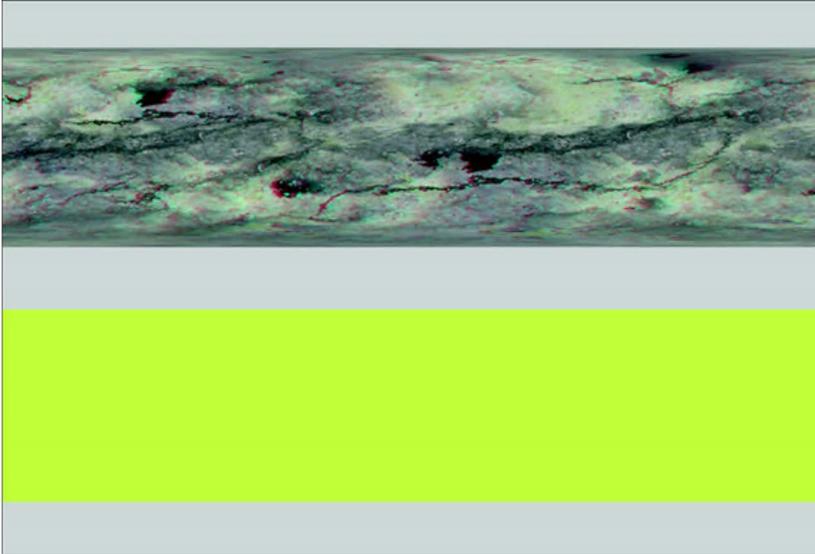


Abbildung 17: Auf diesem Bild sieht man zwei Flächen, die nach hinten gekippt sind.

Während bei der oberen Fläche die Textur Anhaltspunkte für die Lage gibt, ist die Neigung der unteren Fläche nicht zu erkennen.

Da einige der aufgezählten Veränderungen zulasten der Performance gehen, müssen andere Varianten zur Reduzierung der belegten Echtzeit-Ressourcen gefunden werden. Neben der Entfernung einiger „sinnloser“ Post-Effekte (s. folgender Abschnitt) sollte ein wesentlich aggressiveres Detaillevel⁴⁰ verwendet werden. Eine sehr einfache Maßnahme ist die Verringerung der Auflösung des Spiels im Vergleich zu der monoskopischen Varianten, da das visuelle System die Ähnlichkeiten zwischen den beiden Teilbildern dazu verwendet, die Auflösung des Gesehenen zu verbessern.⁴¹

Traditionelle grafische Post-Effekte wie (Bewegungs-)Unschärfe, Glow-Effekte und

39 Vgl.: Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 101f; Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation3.

40 eng.: Level of Detail: je nach Entfernung zum Objekt wird ein feineres oder gröberes Objektmodel geladen.

41 Vgl. Schertenleib (04.05.2011): Making Stereoscopic 3D Games for the Playstation 3.

Bloom-Filter können den stereoskopischen Tiefeneindruck unkontrolliert verändern, da sie meistens dem eigentlichen Renderergebnis als Überlagerungsoperation hinzugefügt werden. Post-Effekte werden in monoskopischen Computerspielen meist dazu verwendet, detailarme Bereiche im Hintergrund zu verdecken, die aufgrund mangelnder Ressourcen nicht besser angezeigt werden können. Post-Effekte sind daher Grundbestandteil jeder heutigen 3D-Spiele-Engine.⁴²

Im Stereo-Modus wirken sie eher kontraproduktiv, da die Änderungen durch die Effekte nicht aus der jeweils richtigen Perspektive erzeugt werden. Dabei kann es unter Umständen vorkommen, dass Pixel verschmiert wirken, was wiederum den stereoskopischen Effekt zerstören kann.⁴³

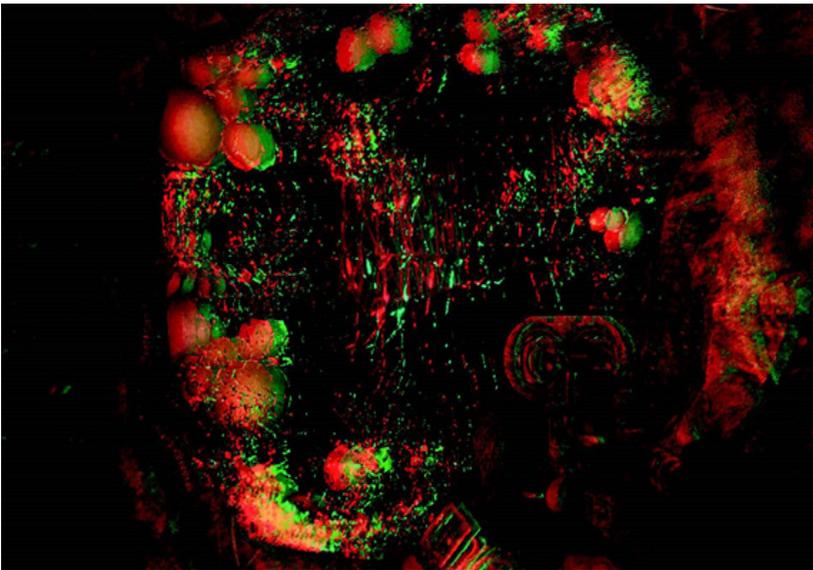


Abbildung 18: im Spiel Metro 2033, das eigentlich für die stereoskopische Ansicht optimiert ist, finden sich die bereits erwähnten Billboards. Weiterhin tritt ein Post-Effekt am Rand des Bildes störend in Erscheinung.

Auch der in vielen Spielen verwendete Film Grain funktioniert nur bedingt und sollte daher abschaltbar sein (s. Abb. 19). Dieser Effekt macht sich vor allem im Bereich von Schatten negativ bemerkbar, solange die Konturen allerdings noch gut sichtbar sind, können die Bilder fusioniert werden.⁴⁴

42 Vgl.: Lengyel (2011): Game engine gems 2, S. 139 und NVIDIA Corporation (2006): NVIDIA GPU Programming Guide, S. 71.

43 Vgl. Lengyel (2011): Game engine gems 2, S. 139.

44 Vgl.: Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S. 92f und 3D World (2011): Create stereo 3D games

Nebel, der oftmals als Post-Effekt angewendet wird, sollte mit Hilfe von Partikeln umgesetzt werden. Dies benötigt zwar wesentlich mehr Rechenleistung, sieht dann allerdings stereoskopisch richtig aus.⁴⁵

Problematisch kann auch eingeschaltetes Anti-Aliasing sein, da dieser Effekt oftmals für alle Tiefenebenen gleich angewendet wird. Dabei ist vor allem das Multisampling-Anti-Aliasing betroffen, welches störende Artefakte im stereoskopischen Bild injizieren kann.⁴⁶



Abbildung 19: Im Spiel Left 4 Dead 2 wird verstärkt Film Grain verwendet. Ist dieser in der stereoskopischen Ansicht angeschaltet, ergibt sich ein recht unscharfes Bild.

3.6 Spielinformationen und Benutzer-Menüs

Elemente wie Head-up-Displays sollten sich nicht, wie normalerweise üblich, auf Tiefe der Wiedergabefläche befinden (folglich flach wahrgenommen), sondern etwas tiefer dargestellt werden (s. Abb. 20).

Enthält die dargestellte Szene einen großen Tiefenumfang, muss der Betrachter immer wieder abwechselnd auf die Objekte im Spiel und die Head-up-Display-Elemente konvergieren. Je weiter diese in der Tiefe auseinanderliegen, desto anstrengender wird das dauernde Wechseln (s. Abb. 21). Dies gilt allerdings nur innerhalb des Spiels.

45 Vgl. 3D World (2011): Create stereo 3D games.

46 Vgl.: Schertenleib (05.11.2010): Optimization for Making Stereoscopic 3D Games on PS3 und 3D World (2011): Create stereo 3D games.

Menüs, in denen keine Spielinhalte im Hintergrund zu sehen sind, können sich auf der Nullebene befinden (s. Abb. 22).⁴⁷

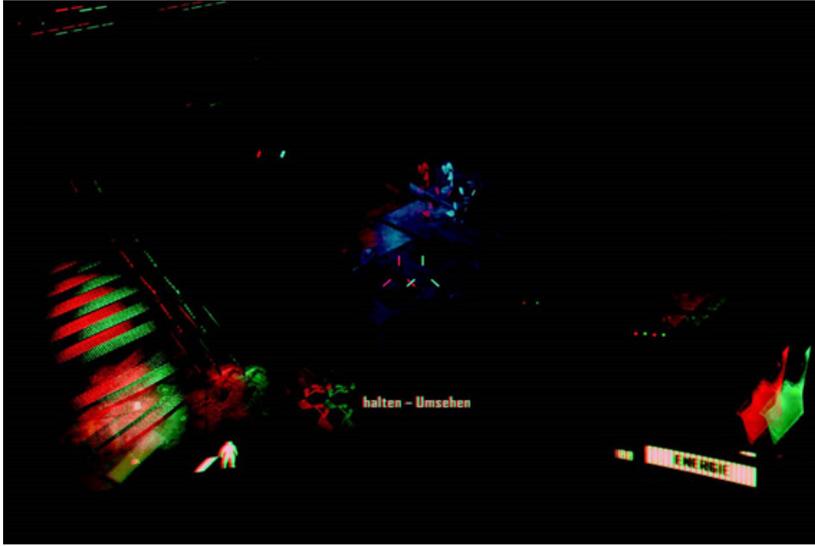


Abbildung 20: Das Head-Up-Display von Crysis 2 ist etwas in die Tiefe versetzt und lässt sich daher ohne große Anstrengung während des Spielens betrachten.

⁴⁷ Vgl.: NVIDIA Corporation (2006): NVIDIA GPU Programming Guide, S. 70; McDonald/Skaljak (2011): NVIDIA 3D Vision Automatic: Best Practices Guide, S. 24.

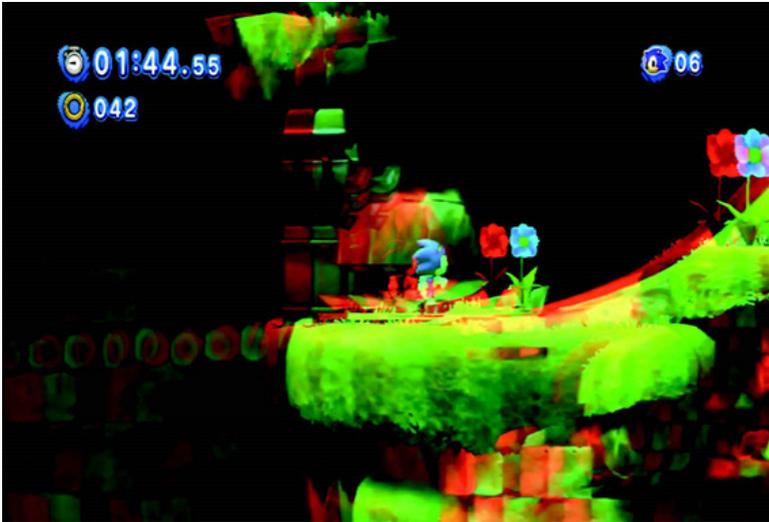


Abbildung 21: Das Spiel Sonic Generations verwendet in der stereoskopischen Ansicht einen hohen Tiefenumfang. Der Wechsel beim Betrachten von Spielinhalt und grafischer Oberfläche strengt sehr an, da die Konvergenzpunkte zu verschieden sind. Abbildung entnommen aus <http://www.youtube.com/watch?v=PWij0o5MBuI&feature=related>.

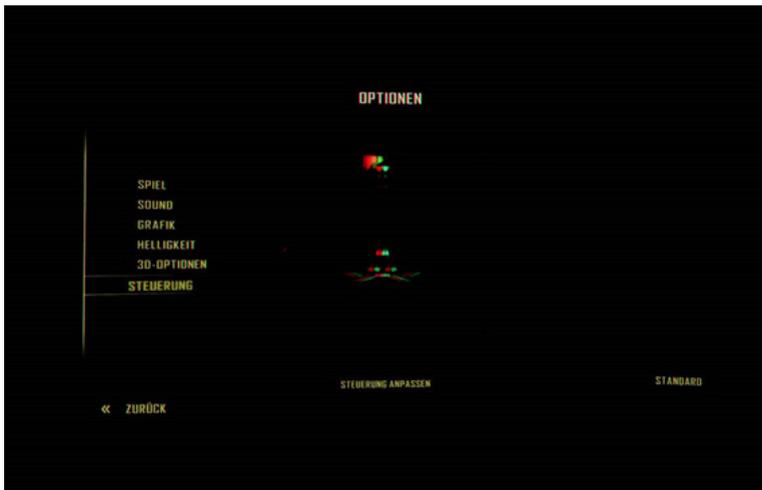


Abbildung 22: Das Spiel Crysis 2 verwendet selbst im Menü den stereoskopischen Effekt. Hier wird bereits deutlich, dass das Spiel für die stereoskopische Betrachtung optimiert ist.

Wenn Namen oder Informationen zu bestimmten Objekten (Lebensanzeigen, etc.) dargestellt werden, sollten diese ebenfalls in der Tiefe des Objekts sein und nicht auf der Nullebene, wie es für solche grafischen Elemente in monoskopischen Ansichten üblich ist (s. Abb. 23). Problematisch wird es dabei, wenn wichtige Informationen der Benutzeroberfläche im Raum durch andere Objekte verdeckt werden. Dies wird in monoskopischen Darstellungen dadurch verhindert, dass solche Elemente ausschließlich auf Bildschirmenebene angezeigt werden und nicht verdeckt werden können.⁴⁸

Oft werden in Elementen des Head-up-Displays und in grafischen Benutzeroberflächen auch Texturen verwendet, die einen dreidimensionalen Eindruck von Schaltflächen erzeugen sollen. Aufgrund der fehlenden Tiefeninformationen sehen diese Schaltflächen stereoskopisch betrachtet, ähnlich wie die bereits erwähnten Billboards, flach aus. Steuerungselemente dieser Art sollten daher aus echter Geometrie bestehen. Weiterhin besteht durch die räumliche Anordnung die Möglichkeit, einzelne Elemente durch unterschiedliche Tiefe hervorzuheben.⁴⁹

Interessant ist hierbei, dass die stereoskopische Darstellung auch ermöglicht, Elemente wie Entfernungen zum nächsten Wegpunkt wegzulassen, da die benötigte Information aufgrund der besser einzuschätzenden Tiefe schon vorhanden ist.⁵⁰

48 Vgl.: Lengyel (2011): *Game engine gems 2*, S. 139f; NVIDIA Corporation (2006): *NVIDIA GPU Programming Guide*, S. 70.

49 Vgl. NVIDIA Corporation (2006): *NVIDIA GPU Programming Guide*, S. 71.

50 Vgl. McDonald/Skajjak (2011): *NVIDIA 3D Vision Automatic: Best Practices Guide*, S. 19.



Abbildung 23: im Spiel Starcraft II: Wings of Liberty werden die Lebensanzeigen der Einheiten und Gebäude in der richtigen Tiefe angezeigt. Einblendungen (z. B. die Anzeige des Spielers, dem eine Einheit gehört) werden allerdings auf der Nullebene ausgegeben und erzeugen damit schwer zu betrachtende Bildbereiche.

3.7 Präventive Maßnahmen gegen Übersprechen / Geisterbilder

In stereoskopischen Betrachtungen kann es eigentlich nur zu sogenannten Doppelwahrnehmungen (Diplopia) kommen, wenn die Parallaxe von Objekten so groß ist, sodass die einzelnen Teilbilder vom visuellen System des Menschen nicht mehr zyklisch fusioniert werden können.⁵¹

Oftmals hat das Auftreten von Doppelbildern in stereoskopischen 3D-Spielen aber eine andere, technische Ursache: Keine Monitor/Bildschirm-Technologie kann derzeit eine perfekte Trennung der beiden Stereo-Teilbilder für die Augen garantieren (dies gewährleisten aktuell nur VR-Headsets mit vollständig getrennter L/R-Adressierung der Augen). Es kann daher bei der Monitor- oder Leinwandbetrachtung immer vorkommen, dass unsere Augen doch Teile des jeweils anderen Bildes wahrnehmen. Dieser Effekt wird *Übersprechen* (engl. *crosstalk*) genannt.

⁵¹ Vgl.: Tauer (2010): Stereo 3D: Grundlagen, Technik und Bildgestaltung, S95ff; Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 176f.

In diesem Falle wird eine zusätzliche *Geisterkontur* (eng. *ghosting*) um Gegenstände wahrgenommen und es entsteht binokulare Rivalität (s. Abb. 24). Je nach Art und Stärke können auch zusätzliche Schatten wahrgenommen werden. Dieser Effekt irritiert Spieler bereits in geringem Ausmaß sehr stark und verhindert das Zustandekommen eines immersiven stereoskopischen Tiefeneindrucks.

Es gibt je nach Wiedergabemethode unterschiedliche technische Ursachen, warum das Übersprechen auftreten kann. Bei Anaglyphen ist meist die Filtereigenschaft der Brille ungenügend. Dies ist z. B. auch der Fall, wenn das Bild Farben aufweist, die von dem Filter vor dem richtigen Auge eliminiert werden. Bei der linearen Polarisierung kommt es zum Übersprechen, wenn die Ausrichtung der Filter nicht mehr korrekt ist. Dies ist z. B. beim Neigen des Kopfes der Fall. Zirkuläre Polarisation ist in diesem Falle vorteilhafter. Bei Shutter-Brillen ist meist eine ungenügende Sperrung bzw. eine träge Shutter-Synchronisation. Zusätzlich können unterbrochene (Sicht-)Verbindungen zum IR-Sender auftreten. Autostereoskope Displays verursachen den Effekt, wenn ein Auge bereits den Teil des nächsten Linsenrasters zu sehen bekommt.

Bis es zu besseren Stereo-3D-Bildschirmtechnologien kommt, bleibt Stereo-3D-Spielentwicklern letztendlich zur Minimierung von potenziellen Ghostingeffekten auf den Displays ihrer Spieler nur ein probates Mittel übrig: sich in Szenen bezüglich des absoluten Kontrastumfangs zu beschränken und lieb gewonnene HDR-Effekte (High Dynamic Range) zu verbannen. Für VR-Headsets können diese ja wieder automatisch hinzugeschaltet werden.

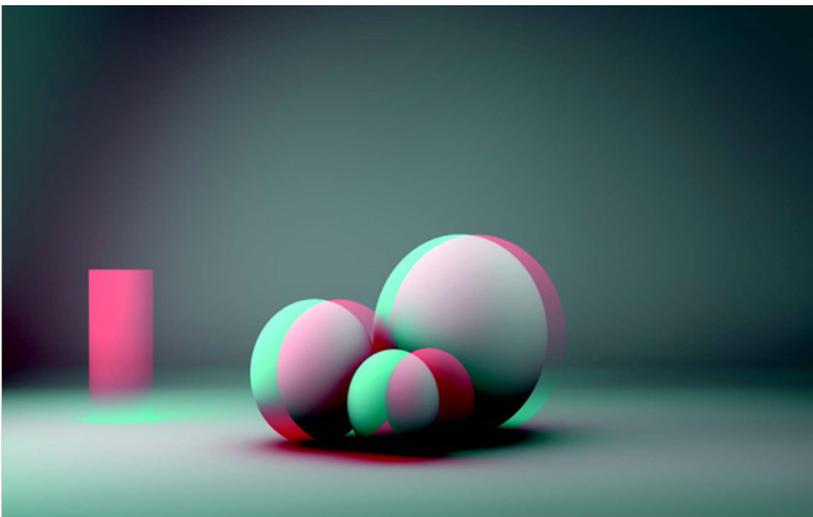


Abbildung 24: Hier entsteht bewusst ein Geisterbild, da der Zylinder nur im Teilbild für das linke Auge vorhanden ist.

4. Ansätze für eine automatische, kontextadaptive Tiefendarstellung

Das Sehen planarer, stereoskopischer Inhalte unterliegt, was die nutzbaren Bildparallaxen angeht, einigen wahrnehmungsphysiologischen Beschränkungen. Im nahen Vordergrund begrenzen Akkommodations-zu-Vergenz-Konflikte die Machbarkeit. Im weiten Hintergrund unnatürlichen divergenten Augenstellungen (also jenseits der Parallelstellung).

Diese Beschränkungen können durch die Displaygröße, den typischen (als bevorzugten angenommenen) Betrachtungsabstand des Nutzers zum Display und seinen ungefähren Augenabstand geometrisch bestimmt werden⁵² und ergeben ein binokulares Rahmenvolumen, in dem das stereoskopische Sehen überhaupt erst möglich ist (s. Abb. 25).

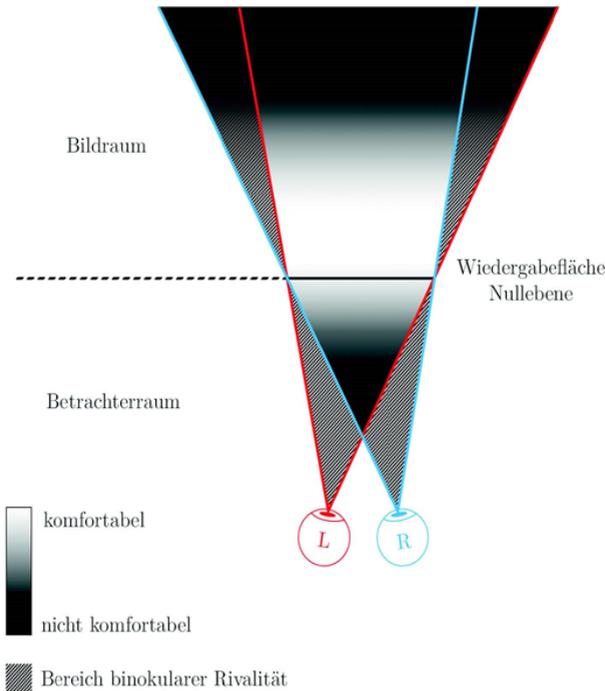


Abbildung 25: Raumvolumen, in dem ein komfortables binokulares Betrachten stereoskopischer Inhalte theoretisch möglich ist. Darstellung angelehnt an Mendiburu (2009): 3D movie making: Stereoscopic digital cinema from script to screen, S. 23.

⁵² Alles Parameter, die bisher von Spieleprogrammierern kaum beachtet werden mussten.

Ein paar kurze Erläuterungen sollen die geometrischen Zusammenhänge verdeutlichen:

Dasselbe Bild mit derselben Parallaxe wirkt aus unterschiedlichen Entfernungen unterschiedlich tief. Je weiter der Betrachter von der Wiedergabefläche weg sitzt, umso größer wird dabei die Tiefenwirkung, denn Objekte in der Entfernung sehen noch weiter entfernt aus und Objekte nah am Betrachter erscheinen noch näher zu sein (s. Abb. 26). Die Distanz zwischen den Objekten bleibt aber bei jedem Abstand proportional konstant.⁵³

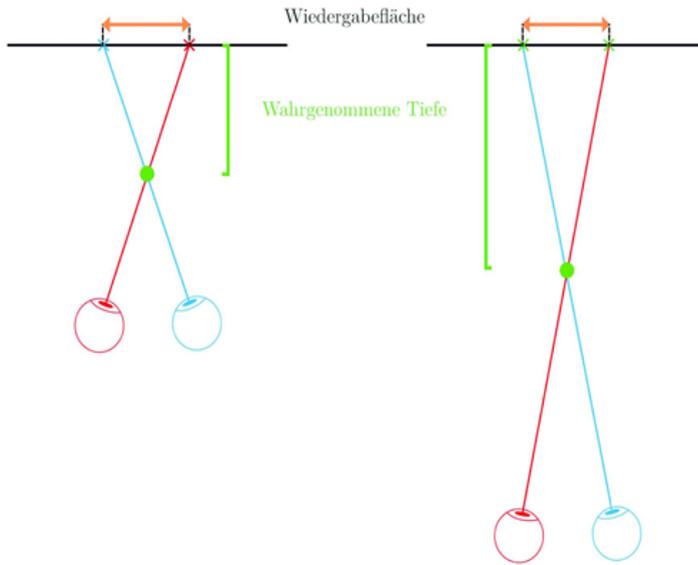


Abbildung 26: Je nach Abstand zur Wiedergabefläche ergibt sich eine andere Tiefenwirkung, obwohl die negative Parallaxe die gleiche bleibt.

Im Zusammenhang von unterschiedlich großen Wiedergabegeräten mit unterschiedlichen Dotpitches ist es zusätzlich von Bedeutung, dass sich - gleichbleibende Auflösung vorausgesetzt - die maximal zulässige Pixelparallaxe bei kleinerer Wiedergabefläche automatisch erhöht (z. B. Nintendo 3DS), bei einer Größeren (z. B. Stereo-3D-Beamer) dramatisch erniedrigt.⁵⁴

Alle o.g. Parameter müssen also unbedingt bei der Installation des stereoskopischen Spiels vorab abgefragt⁵⁵ und beim Rendern berücksichtigt werden.

53 Vgl. Lipton (1997): *StereoGraphics Developers' Handbook*, S. 22.

54 Vgl.: Gateau/Neuman (16.12.2010): *Stereoscopy, From XY to Z* und Mendiburu (2009): *3D movie making: Stereoscopic digital cinema from script to screen*, S. 21f .

55 oder wie der durchschnittliche Augenabstand von Erwachsenen mit 6cm als default fest ange-

Nach der Installation verbleiben innerhalb dieser geometrischen Grenzen letztendlich zwei Stellschrauben:

- Innerhalb des maximal zulässigen binokularen Rahmenvolumens gibt es das Untervolumen der komfortablen patenten Stereopsis. Hier wird zusätzlich berücksichtigt, dass die maximale Deviation in den Augen des Betrachters bei Fixierung benachbarter, tiefengestaffelter Objekte in der Szene ungefähr 40 Winkelminuten nicht überschreiten darf. Diese Stellschraube ist die Stereobasis der virtuellen Renderkamas⁵⁶.
- Dieses Volumen kann dann – innerhalb der zulässigen Grenzen - absolut zum Betrachter entweder nach vorne oder nach hinten verlagert werden. Diese Stellschraube ist die Konvergenz der virtuellen Renderkamas bzw. ihr Frustum Shift (auch Horizontal Image Translation/HIT genannt).

Diese beiden Stellschrauben im Verlauf eines Spiels permanent wahrnehmungsgerecht zu kalkulieren und zu adaptieren ist die große Herausforderung für zukünftige Entwickler von stereoskopischen Spielen.

nommen werden.

56 Theoretisch können natürlich auch die Objektabstände in der (interaktiven) Szene automatisch dieser Regel angepasst werden – es muss aber vom Betrachter unbemerkt erfolgen, was nicht ganz einfach zu bewerkstelligen ist. Alternativ wären theoretisch auch noch Abstands- und Brennweitenveränderungen der Renderkamas möglich; aber auch hier stellt sich die Frage der unmerklichen Anpassung.

Literaturverzeichnis

- Adams, Ernest/Rollings, Andrew: Fundamentals of game design. 2. Auflage. Berkeley and CA: New Riders, 2010, ISBN 978-0-321-64337-7.
- Blackman, Sue: Beginning 3D Game Development with Unity. Berkley: APRESS, 2012, ISBN 1430234229.
- Cutting, James E./Vishton, Peter M.: Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. In Epstein, William/Rogers, Sheena J. (Hrsg.): Perception of space and motion. San Diego: Academic Press, 1995, ISBN 0122405307, 69–117.
- Gillam, Barbara: The Perception of Spatial Layout from Static Optical Information. In Epstein, William/Rogers, Sheena J. (Hrsg.): Perception of space and motion. San Diego: Academic Press, 1995, ISBN 0122405307, 23–67.
- Goldstein, E. Bruce: Sensation and perception. 8. Auflage. Belmont and CA: Wadsworth, Cengage Learning, 2010, ISBN 0-495-60149-7.
- Hanson, Andrew J./Wernert, Eric A.: Constrained 3D navigation with 2D controllers. In Proceedings of the 8th conference on Visualization '97. Los Alamitos and CA and USA: IEEE Computer Society Press, 1997, VIS '97 – Online verfügbar unter: <http://dl.acm.org/citation.cfm?id=266989.267052i>, ISBN 1-58113-011-2, 175–ff.
- Hollwich, Fritz/Jaensch, P. A.: Augenheilkunde: ein kurzgefasstes Lehrbuch. 9. Auflage. Stuttgart: Thieme, 1979, ISBN 3133551094.
- Howard, Ian P./Rogers, Brian J.: Binocular vision and stereopsis. New York: Oxford University Press, 1995, ISBN 0-19-508476-4.
- Hubona, Geoffrey S. et al.: The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. ACM Trans. Comput.-Hum. Interact. 6 1999 Nr. 3, 214–242 – Online verfügbar unter: <http://doi.acm.org/10.1145/329693.329695i>, ISSN 1073-0516.
- Julesz, Bela: Binocular Depth Perception without Familiarity Cues: Random-dot stereo images with controlled spatial and temporal properties clarify problems in stereopsis. Science, 145 1964 Nr. 3630, 356–362.

- Kuhn, Gerhard: Stereofotografie und Raumbildprojektion: [Theorie und Praxis, Geräte, Materialien]. Gilching: Vfv, 1999, ISBN 3-88955-119-X. 133
- LaViola, Jr. Joseph J./Litwiller, Tad: Evaluating the benefits of 3d stereo in modern video games. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York and NY and USA: ACM, 2011, CHI '11 – Online verfügbar unter: <http://doi.acm.org/10.1145/1978942.1979286i>, ISBN 978-1-4503-0228-9, 2345-2354.
- Lengyel, Eric: Game engine gems 2. Natick and Mass: A K Peters Ltd, 2011, ISBN 1568814372.
- Lipton, Lenny: Foundations of the stereoscopic cinema: A study in depth. New York u.a: Van Nostrand Reinhold, 1982, ISBN 0-442-24724-9.
- Mendiburu, Bernard: 3D movie making: Stereoscopic digital cinema from script to screen. Amsterdam and Boston: Focal Press/Elsevier, 2009, ISBN 9780240811376.
- Moore, Michael E.: Basics of game design. Boca Raton: A K Peters/CRC Press, 2011, ISBN 978-1-4398-6776-1.
- Pernkopf, Eduard/Ferner, Helmut: Atlas der topographischen und angewandten Anatomie des Menschen. München: Urban & Schwarzenberg, 1963.
- Rajae-Joordens, Rosemarie J.E: Measuring Experiences in Gaming and TV Applications. In Westerink, Joyce H. D. M. et al. (Hrsg.): Probing Experience. Band 8, Springer Netherlands, 2008 – Online verfügbar unter: http://dx.doi.org/10.1007/978-1-4020-6593-4_7i, ISBN 978-1-4020-6593-4, 77-90.
- Richards, Whitman: Stereopsis and stereoblindness. Experimental Brain Research, 10 1970 Nr. 4, 380-388 – Online verfügbar unter: <http://link.springer.com/article/10.1007%2FBF02324765?LI=true#page-1i> – abgerufen am: 08.11.2012, ISSN 0014-4819.
- Schwartz, Steven H.: Visual perception: A clinical orientation. 4. Auflage. New York: McGraw-Hill Medical Pub. Division, 2010, ISBN 9780071604628.
- Siragusano, Daniele/Minuth, Andreas: Die Rolle der Bildwandgröße in der Stereoskopie. FKT, FKT64 2010 Nr. 11, 557-561.
- Tauer, Holger: Stereo 3D: Grundlagen, Technik und Bildgestaltung. 1. Auflage.

Berlin: Schiele & Schön, 2010, ISBN 9783794907915.

- Waldeyer, Anton: Anatomie des Menschen für Studierende und Ärzte. 13. Auflage. Berlin and New York: de Gruyter, 1975, ISBN 3110059738.
- Zachara, Matt/Zagal, José P.: Challenges for success in stereo gaming: a Virtual Boy case study. In Proceedings of the International Conference on Advances in Computer Entertainment Technology. New York and NY and USA: ACM, 2009, ACE '09, ISBN 978-1-60558-864-3, 99-106.

Teil 7
Stereoskopische Einflussmöglichkeiten in
3D-Spieleumgebungen für Machinimas

Robin Grabowski

Abbildungsverzeichnis

Abbildung 1: Battlefield Vietnam, Parallaxenmessung bei Einstellung „20% Tiefe“.....	276
Abbildung 2: Verhältnis von Objektparallaxen[px] zu Entfernung (NVIDIA).....	277
Abbildung 3: Parallaxenmessung Team Fortress2	279
Abbildung 4: Verhältnis von Objektparallaxen[px] zu Entfernung (Filmmaker) ..	281
Abbildung 5: Ergebnisse der vier Parallaxenmessungen - Source Filmmaker	281

Tabellenverzeichnis

Tabelle 1: Objektparallaxen vs. Einstellregler „Stereoskopische 3D-Funktion/ Tiefe“ in Nvidia / Battlefield Vietnam	276
Tabelle 2: Objektparallaxen vs. Einstellregler zeroPD (eyeSeparation=1,375)	280
Tabelle 3: Objektparallaxen vs. Einstellregler zeroPD (eyeSeparation=2,75)	280
Tabelle 4: Objektparallaxen vs. Einstellregler „Konvergenz“ (iClone)	284
Tabelle 5: Objektparallaxen vs. Brennweiten der virtuelln Kameras (iClone)	284

1. Einleitung

Machinima ist eine Wortschöpfung aus den Worten „machine“, „cinema“ und „animation“ und wird laut Paul Marino, Machinima-Künstler, als »animated film making within a real-time virtual 3D environment«¹ definiert. Machinima ist ein Mix aus verschiedenen Produktionstechniken und besteht aus Animation, einer 3D-Game-Engine bzw. einer 3D-Umgebung und dem Filmemachen darin. Das Zusammenwirken dieser drei Komponenten sieht dabei wie folgt aus: Mit Machinima werden in einer virtuellen Umgebung (3D-Umgebung) Charaktere gesteuert und Aktionen durchgeführt (Animation) und diese in Form von Geschichten aufgenommen (Film).²

Während Machinima eine innovative Form der Filmproduktion in virtuellen Welten darstellt, so ebnet der stereoskopische Film den Weg zu einer neuen Art von immersivem Filmeeerleben für den Zuschauer. Beide Komponenten zusammenzuführen wäre der nächste logische Schritt.

Doch welche technischen Einflussmaßnahmen sind in aktuellen Engines des Jahres 2013 gegeben, um wahrnehmungsgerechte³ stereoskopische Machinimas zu produzieren? Jene Frage soll im Verlauf dieses Artikels beantwortet werden, indem verschiedene 3D-Umgebungen auf ihre stereoskopischen Funktionen hin analysiert werden. Darüber hinaus wird beleuchtet, inwiefern der Faktor Stereoskopie den Workflow einer Machinima-Produktion verändert. Die dadurch gewonnenen Erkenntnisse wurden genutzt, um prototypisch einen Stereo-3D-Machinima zu realisieren.

2. Marktanalyse

Anhaltspunkte, ob mit aktuell vorhandenen 3D-Engines und tatsächlich schon wahrnehmungsgerechte Machinima produziert werden können, wurden in einer ersten Recherchephase nicht gefunden.

1 Marino (2004) S.1

2 Vgl.: Marino (2004) S.2

3 Unsere verwendete Definition: bei einer nutzungsgerechten Betrachtung kein Auftreten von Augendivergenz sowie eine maximale, relative Parallaxe von 40 Winkelminuten im Renderergebnis.

Zunächst wurden daher drei 3D-Umgebungen untersucht, welche Stereoskopie explizit unterstützen. Dadurch sollte herausgefunden werden, welche von ihnen die meisten Einstellungsmöglichkeiten, hinsichtlich der Stereoskopie bietet. Es dienen u.a. folgende Punkte als Orientierung:

- Veränderbarkeit der Stereobasis
- Veränderbarkeit der Brennweite / Point-of-View (PoV)
- Frei platzierbare Objekte und Personen
- Frei bewegliche Kamera(s)
- Ein- und abschaltbare Post-Effekte
- Erreichbare Framerate

Für die Thesis wurden folgende (im Jahr 2013 vorhandenen) Produktionsansätze untersucht und evaluiert: NVIDIA 3D Vision, Source Filmmaker und iClone.

Für jede der o.g. Umgebungen wurden verschiedene Messungen durchgeführt. Diese Tests sollten dazu dienen herauszufinden, wie sich der Stereo-3D-Effekt unter Veränderung der unterschiedlichen Systemparameter verhält. Die daraus gewonnenen Erkenntnisse sollten dann für eine prototypische Machinima-Produktion in einer der Umgebungen genutzt werden.

Laut ITU-R BT.2022⁴ liegt für HD-Bilder der minimale empfohlene Betrachtungsabstand (Design-Viewing Distance/DVD) beim 3,2 fachen der Bildschirmhöhe. Da der Bildschirm, mit dem getestet wurde, 17,3 Zoll Diagonale und FullHD-Auflösung hatte, lag dieser Betrachtungsabstand somit bei 68,8 cm. Die maximale relative Parallaxe sollte wie erwähnt 40 Winkelminuten nicht überschreiten, in diesem Fall also 40 Pixel zwischen vorderstem und hinterstem Objekt (bei Überdeckung).

2.1 NVIDIA 3D Vision, 3D-PC, 3D-Display und Fraps

Hierbei handelt es sich um eine Kombination aus mehreren Hard- und Software-Elementen. NVIDIA bietet mit dem 3D Vision Pack eine Brille inklusive Sensoren an, mit denen Computerspiele, die über eine entsprechende NVIDIA Grafikkarte laufen, in Stereo-3D gespielt werden können. Um diese für Machinimas nutzen zu können, verwendet man Fraps. Letzteres ist ein Tool, mit dem das jeweilige Spiel im Side-by-Side-Format aufgezeichnet werden kann. Zu erwähnen ist, dass in dieser Kombination nur Fraps verwendet werden kann, da es als einziges Programm NVIDIA 3D Vision unterstützt.

4 General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays, Vgl.: ITU-R (2014) Online (15.02.14)

2.1.1 Funktionsweise der 3D-Vision Technologie

Die Grundvoraussetzung für die 3D-Fähigkeit ist ein zweiter Prozessor (kompatibler GeForce-Grafikprozessor), der die Aufgabe des Renderns und Anzeigens von zwei Bildern (für das linke und das rechte Auge) auf dem Bildschirm übernimmt. Die Komponenten 3D-Display, 3D-Brille (Shutterbrille) und der 3D-Grafikprozessor sind so aufeinander abgestimmt, dass jedes Auge, das für ihn jeweilige Bild empfängt.

Die NVIDIA 3D Vision PCs arbeiten mit LCD-Bildschirmen, die 120 Hz unterstützen und eine Auflösung von 1920x1080 oder 1680x1050 besitzen. Da beim Shutterverfahren Bilder abwechselnd für das rechte und linke Auge angezeigt werden, halbiert sich bei 120-Hz-Displays die Bildwiederholrate für jedes Auge auf 60Hz, was der Standard bei heutigen LCD-Bildschirmen ist. Somit sieht jedes Auge die Bilder in voller Auflösung.

2.1.2 NVIDIA-Kompatibilitätsliste

In der NVIDIA-Kompatibilitätsliste lassen sich hunderte Spiele finden, die auf ihre 3D-Fähigkeit hin getestet wurden. Die Unterteilung der Kompatibilitätsgrade geht von „nicht empfohlen“ bis „hervorragend“ bzw. „3D-Vision ready“. Falls Fehler in der 3D-Darstellung auftreten, ist dies vermerkt. So findet sich u.a. zum Spiel Battlefield Vietnam, welches für die Stereo-3D-Tests verwendet wurde, folgender Kommentar: »Tough to hit targets with 2D sight; Gunsight or Pointer is 2D Object«.

2.1.3 Stereo-Parametrisierung

Die Tiefeneinstellung lässt sich in der NVIDIA-Systemsteuerung unter dem Punkt »Stereoskopische 3D-Funktion« finden. Dort ist es möglich, mithilfe eines Reglers die Tiefe von 1-100% (sehr schwach - sehr stark) einzustellen.

Die einzig andere Einstellungsmöglichkeit ist die Konvergenz der virtuellen Renderkameras. Wird diese Option aktiviert (in den Werkseinstellungen immer deaktiviert), kann während des Spiels eingestellt werden, ob alle Objekte vor, vor und hinter oder nur hinter dem Bildschirm liegen sollen. Es kann somit die Gewichtung von Vordergrund zu Hintergrund bestimmt werden. Anders als bei der Tiefe, die in Prozent angegeben wird, gibt es hier keine Einheit. Hier wird lediglich durch wiederholtes Drücken der festgelegten Tastenkombination die Konvergenz verändert.

Nachdem die Einstellmöglichkeiten bekannt waren, wurde untersucht, wie sich die Tiefe über beide Parameter hinweg verhält. Es sollte dabei festgestellt werden, ob das Ergebnis, mit dem einer wahrnehmungsgerechten Realfilm-Aufnahme zu vergleichen ist.

Der Test für diese Umgebung wurde mit dem Spiel Battlefield Vietnam⁵, das in der Kompatibilitätsliste ein »gut« erhielt, durchgeführt. Hierfür wurden vier Screenshots

⁵ Battlefield Vietnam, EA (2004), eigene Bearbeitung

mit unterschiedlichen Tiefen erstellt. Bei jedem Bild wurden sechs Objekte in unterschiedlicher Entfernung gewählt und ihre absoluten Parallaxen beider Halbbilder gemessen. Alle Objekte befanden sich hinter dem Bildschirm, da die Konvergenz standardgemäß so eingestellt war und für diesen Test nicht verändert wurde. So sollte erfahren werden, welche Werte die Parallaxen mit zunehmender Entfernung und mit verschiedenen Tiefeneinstellungen annehmen. Die Entfernung konnte zwar nicht gemessen, aber mithilfe der Bildobjekte (bekannte Größe) ungefähr abgeschätzt werden. Die Ergebnisse legen keine exakten Verhältnisse dar, sondern nur das grobe Verhalten. Getestet wurden Tiefenwerten von 10%, 20%, 40% und 80%.



Abbildung 1: Battlefield Vietnam, Parallaxenmessung bei Einstellung „20% Tiefe“

Ergebnisse (Objekte geordnet nach Entfernung von nah bis fern):

Objekte	Tiefe 10%	Tiefe 20%	Tiefe 40%	Tiefe 80%
Waffe Stelle A 30 cm	4 px	8 px	Fehlt	28 px
Waffe Stelle B 80 cm	8 px	16 px	30 px	60 px
Helikopter 1 4 m	20 px	40 px	78 px	154 px
Helikopter 2 15 m	22 px	44 px	84 px	168 px
Baumkrone 30 m	22 px	44 px	86 px	170 px
Baum am Horizont 200 m	24 px	44 px	86 px	172 px
Rel. Parallaxe	20 px	36 px	Fehlt	144 px

Tabelle 1: Objektparallaxen vs. Einstellregler „Stereoskopische 3D-Funktion/Tiefe“ in Nvidia / Battlefield Vietnam⁶

6 Anm.: Die Werte haben eine Abweichung von wenigen Pixeln, da von Hand gemessen wurde.

Es zeigt sich, dass sich die absoluten Parallaxen verdoppeln, wenn sich die Tiefe (in Prozent) verdoppelt. Setzt man die geschätzten Entfernungen mit den Parallaxenwerten in Relation ergeben sich folgende Kurven:

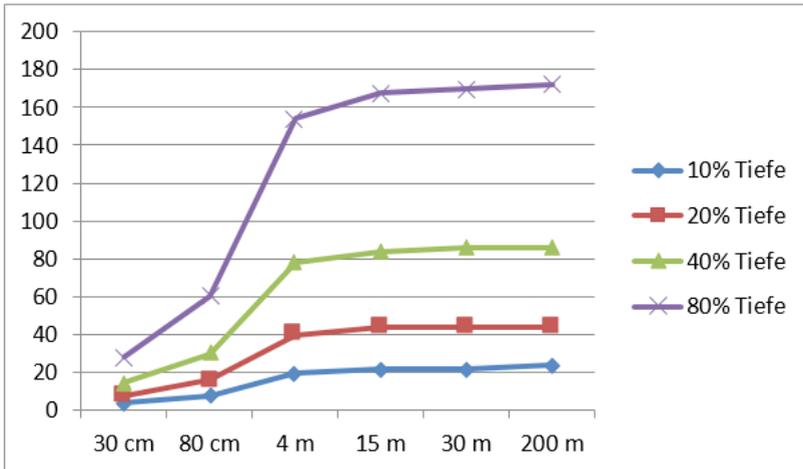


Abbildung 2: Verhältnis von Objektparallaxen[px] zu Entfernung (NVIDIA)

Die absoluten Parallaxen verändern sich im Nahbereich relativ schnell. Ab ungefähr >4 m verändert sie sich bis zur Entfernung »unendlich« nur noch geringfügig, was ungefähr dem natürlichen Sehen ähnelt. Die Plastizität von Objekten nimmt auch hier mit zunehmender Entfernung ab. Das liegt daran, dass sich die Perspektiven eines Objekts (durch beide Augen betrachtet) kaum noch unterscheiden und Volumina folglich flächig wirken.

Je höher der Parameter „Tiefe“ eingestellt wird, desto größer wird auch die relative Parallaxe und somit die absolute Tiefe. Deshalb kann davon ausgegangen werden, dass der NVIDIA-Tiefenregler quasi die Stereobasis repräsentiert. NVIDIA empfiehlt für Games eine Einstellung von 15%.

2.1.4 Workflow-Aspekte

Die Prozess-Vorteile von NVIDIA 3D Vision lauteten:

- Sehr schnelle und einfache Erstellung einer stereoskopischen Aufnahme
- Wiedergabe und Veränderung den Tiefeneffekts in Echtzeit
- Einfache Tiefen- und Konvergenzeinstellung
- Große Auswahl an Spiele-Engines mit Editoren/Mods

Geht man davon aus, dass ein Machinima produziert wird, der nur die integrierten Einstellungen eines bestimmten Spiels nutzt und mit Fraps aufgenommen wird, könnten u.a. folgende mögliche Nachteile festgestellt werden:

- Feste Kameraposition
- Kamerafahrten nur eingeschränkt möglich
- Objekte nicht frei positionierbar, daher keine eigene Tiefenstaffelung möglich
- Kein freies Lichtsetzen möglich
- Schatten, die den Stereo-3D-Effekt stören (z. B. Total War: Rom II)
- Rauch nur in 2D darstellbar (z. B. Total War: Rom II)
- Keine Brennweitenveränderungen möglich
- Keine nachträgliche Veränderung der Kamera nach der Aufnahme

Für alle Spiele gilt, dass der Stereo-Effekt in Echtzeit beurteilt werden kann und eine Veränderung beider Stereo-Parameter jederzeit möglich sind.

2.2 Valve Source Filmmaker

Die zweite analysierte Umgebung stellte der Source Filmmaker von Valve da. Die Basis dieses Editors ist das Spiel Team Fortress 2⁷ und die Source Game Engine. Der Filmmaker wurde speziell für die Erstellung von Machinima entwickelt. Alle Objekte, die im Spiel vorkommen, können auch im Editor verwendet werden. Auch der Source Filmmaker bietet die Möglichkeit, Stereoaufnahmen zu erzeugen. Zentrale Funktionen des Programms sind:

- Aufnehmen
- Editieren
- Freie Kamerapositionierung
- Animierung (Puppeteering und über Keyframes)
- Audiodateien einfügen, wie Dialoge und Effekte
- Manipulierbarkeit von Figuren und Gegenstände

⁷ Team Fortress 2, Valve (2007), eigene Bearbeitung

- Lippensynchronisation
- Rigging
- Lichtsetzen
- Schnitt

2.2.1 Stereo-Parametrisierung

Es gibt im Filmmaker ebenfalls zwei Stereo-Parameter. Diese befinden sich unter den Einstellungen jeder einzelnen Renderkamera. Diese nennen sich „eyeSeparation“ und „zeroParallaxDistance“ (zeroPD). Bei eyeSeparation handelt es sich um die Stereobasis und bei der zeroPD um die Konvergenz. Dies ergab sich durch die Auswertung der unterstehenden Tests der Parallaxenmessung.



Abbildung 3: Parallaxenmessung Team Fortress2

Es wurde über vier verschiedene Konvergenzwerte (jeweils um den doppelten Wert erhöht) die absoluten Parallaxen (mit immer derselben Stereobasis von 1.375 bzw. 2.75) ausgewertet. Je höher die Konvergenz, desto weiter weg von der Kamera befindet sich der Punkt, der bei Wiedergabe genau auf der Leinwand liegt. Auch hier wurde die Entfernung lediglich abgeschätzt, wobei ein Parameter der Kamera (znear bzw. zfar)⁸ zweckentfremdet wurde, um die Entfernung in etwa bestimmen zu können. Da auch hier von Hand gemessen wurde, sind nicht alle Werte Pixel-exakt. Es muss mit einer kleinen Fehlertoleranz gerechnet werden.

⁸ Mit znear kann alles ausgeblendet werden, was unter diesem Wert liegt. Bei zfar alles, was darüber hinaus geht

Die Messungen ergaben folgende Ergebnisse:

Bei gleichbleibender eyeSeparation von 1.375:

Objekte	zeroPD 37.5	zeroPD 75	zeroPD 150	zeroPD 300
Waffenlauf 1 m	27 px	6 px	-6 px	-12 px
Kiste 2.75 m	38 px	17 px	7 px	0 px
Autogrill 5.5 m	41 px	20 px	9 px	4 px
Baum links Nr.2 52 m	43 px	22 px	11 px	6 px
Baum Horizont 80 m	43 px	22 px	11 px	6 px
Rel. Parallaxe	16 px	16 px	17 px	18 px

Tabelle 2: Objektparallaxen vs. Einstellregler zeroPD⁹ (eyeSeparation=1,375)

Bei gleichbleibender eyeSeparation von 2.75:

Objekte	zeroPD 37.5	zeroPD 75	zeroPD 150	zeroPD 300
Waffenlauf 1 m	54 px	11 px	-11 px	-23 px
Kiste 2.75 m	75 px	33 px	11 px	0 px
Autogrill 5.5 m	81 px	38 px	17 px	7 px
Baum links Nr.2 52 m	85 px	42 px	22 px	11 px
Baum Horizont 80 m	86 px	42 px	22 px	11 px
Rel. Parallaxe	32 px	31 px	33 px	34 px

Tabelle 3: Objektparallaxen vs. Einstellregler zeroPD (eyeSeparation=2,75)

Vergleicht man die Tabelleneinträge lässt sich Folgendes ablesen: Bei gleichbleibender Konvergenz und verdoppelter Stereobasis (eyeSeparation), verdoppelten sich die absoluten Parallaxen und ebenfalls die absolute Tiefe (siehe Tabellenzeile „Rel. Parallaxe“). Die Parallaxen steigen im Nahbereich stärker an, als mit zunehmender Entfernung. Die Kurven, die in Abbildung 4 zu sehen sind, zeigen einen sehr ähnlichen Verlauf, wie bei den Tests mit NVIDIA 3D Vision und Battlefield Vietnam.

⁹ Bei den Werten -6 und -12 war ein Teil der Figur somit vor dem Bildschirm.

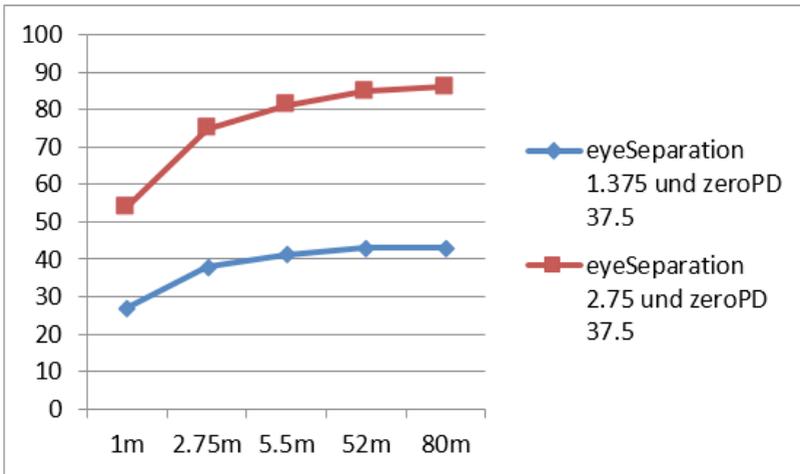


Abbildung 4: Verhältnis von Objektparallaxen[px] zu Entfernung (Filmmaker)

Es wurden weitere Tests durchgeführt, um das Verhalten zwischen veränderter Brennweite und gleichbleibender Stereobasis, Konvergenz und Abstand zum Nahpunkt zu untersuchen. Die zeroPD betrug dabei immer 100. Somit lag sie vor dem Charakter. Es wurden vier Aufnahmen desselben Bildes erstellt. Zwei davon mit 24mm Brennweite, wobei die Stereobasis einmal 1 und einmal 2 betrug. Die letzten beiden Bilder mit 50mm und ebenfalls einer Stereobasis von 1 bzw. 2.

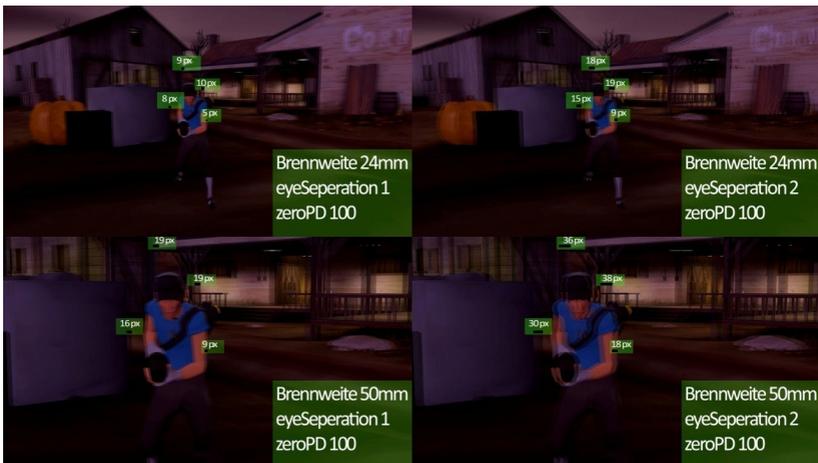


Abbildung 5: Ergebnisse der vier Parallaxenmessungen - Source Filmmaker¹⁰

10 Quelle: Team Fortress 2, Valve (2007), eigene Bearbeitung

Eine gleichbleibende Brennweite und eine Verdoppelung der Stereobasis führten im Schnitt zu einer Verdoppelung der absoluten Parallaxen. Dasselbe geschah auch bei gleichbleibender Stereobasis und einer Verdoppelung der Brennweite. Während doppelte Brennweite und doppelte Stereobasis einer Vervierfachung der absoluten Parallaxen gleichkommen, so führen eine doppelte Brennweite und halbierte Stereobasis zu denselben Parallaxenwerten, wie eine halbierte Brennweite mit der doppelten Stereobasis.

Was der Source Filmmaker besser macht als NVIDIA ist, dass die Konvergenz und die Stereobasis mit Zahlenwerten ausgedrückt werden. Dennoch muss auch hier mit Werten gearbeitet werden, die im Sinne einer Stereokalkulation keine repräsentativen Aussagen haben. Wie sind 10% Tiefe (NVIDIA) und eyeSeparation 1 kalkulatorisch zu verstehen?

2.2.2 Workflow-Analyse

Nach etwas Einarbeitungszeit konnten schnell die Stärken und Schwächen des Editors festgemacht werden.

Vorteile:

- Freie Kamerabewegungen und Einsatz mehrerer Kameras
- Viele Kameraparameter (u.a. Shutterspeed, Blende, Fokus und Depth-of-Field)
- Freie Animation der Charaktere und Objekte
- Freie Objektpositionierung
- Depth-of-Field, Motion Blur und Ambient Occlusion wahlweise schaltbar
- Sehr gute Lippensynchronisation
- Schneller Wechsel zwischen Spiel, Editor, Aufnahme
- Animationen externer Programme importierbar (z.B. Autodesk Motionbuilder)
- Verwendung von Assets
- Videoschnitt im Editor
- Regulierung der Stereobasis und Konvergenz

Nachteile:

- Begrenzte Auswahl an Brennweiten.
- Stufenlose Brennweitenveränderung ist vorhanden, allerdings ohne Anzeige der Brennweite; manuelle Umrechnung nötig
- Comichafter Grafikstil nicht für alle Szenarien geeignet
- Höhere Einarbeitungszeit
- Stereo-3D als Side-by-Side und Over-Under
- Stereo-3D nicht in Echtzeit

2.3 iClone

iClone von Reallusion ist eine Echtzeit-3D-Animations- und Rendering-Software. Mit diesem Programm ist es möglich, einen Machinima-Film von Grund auf zu produzieren. iClone ist flexibel, schnell und leicht zu erlernen.

2.3.1 Stereo-Parametrisierung

Es ist bei iClone lediglich möglich, die Konvergenz einzustellen. Somit bietet iClone dem Nutzer nur die Möglichkeit zu bestimmen, welche Objekte sich vor, auf und hinter der Leinwand befinden. Die Stereobasis kann nicht festgelegt werden. Angesichts der Tatsache, dass stereoskopische 3D-Filme seit einigen Jahren im Trend liegen und NVIDIA mit 3D Vision, wie auch Valve mit ihrem Source Filmmaker das Potenzial erkannt haben, ist es umso überraschender, dass diesbezüglich bei iClone bisher keine Weiterentwicklung stattgefunden hat.

Auch mit iClone wurden Parallaxenmessungen durchgeführt, um zu untersuchen, wie sich der Stereo-3D-Effekt verhält. Bei diesem Test betrug die Kameraentfernung zum Nahpunkt 2 m. Die virtuelle Renderkamera hatte eine Brennweite von 50 mm. Die Konvergenz veränderte sich hierbei jeweils um das Doppelte. Zu beachten sei jedoch, dass eine Verdopplung der Konvergenz, eine Halbierung des Konvergenzwinkels bedeutet. Eine Konvergenz von 50 sagt aus, dass die Zero-Parallaxe-Plane, also der Nullpunkt in 50 cm Entfernung liegt. Der Test wurde mit einer Konvergenz von 50 (50cm) begonnen und mit einer Konvergenz von 1000 (10 m) beendet, da diese die Minimal- bzw. Maximalwerte sind.

Objekte	Konv. 50	Konv. 100	Konv. 200	Konv. 400	Konv. 800	Konv. 1000
Fuchs 2 m	137 px	47 px	2 px	-21 px	-31 px	-35 px
Kiste 4 m	158 px	68 px	24 px	1 px	-11 px	-14 px
Couch 8	169 px	79 px	35 px	13 px	0 px	-3 px
Roboter 16 m	175 px	85 px	40 px	18 px	6 px	6 px
Ruine 32 m	177 px	87 px	43 px	20 px	10 px	7 px
Rel. Paral- laxe	40 px	40 px	41 px	41 px	41 px	42 px

Tabelle 4: Objektparallaxen vs. Einstellregler „Konvergenz“ (iClone)

Wie die Spalte „relative Parallaxe“ zeigt, änderte sich bei einer Konvergenzänderung die absolute Tiefe nicht. Wenn die Konvergenz von 50 auf 100 gestellt wird, verringert sich die absolute Parallaxe jedes einzelnen Objekts um 90 Pixel. Bei weiterer Verdopplung der Konvergenz halbiert sich der genannte Wert nahezu (90 -> 45 -> 23 -> 11 ->3). Bei großen Konvergenzwinkeln steigen die Werte der absoluten Parallaxen sehr stark an. In der Regel wirken sich somit kleine Konvergenzwinkel (hier hohe Konvergenzwerte) angenehmer auf den Stereo-3D-Effekt aus, als große. Hierbei konnte festgestellt werden, dass keine Unterschiede zum Werteverhalten des Source Filmmakers zu erkennen sind. Die Stereobasis konnte selbst nicht verändert werden. Nochmalige Recherchen ergaben, dass tatsächlich nur die Konvergenz variiert werden kann. Da die Brennweite jedoch ebenfalls Einfluss auf die absolute Tiefe hat, wurde ein weiterer Test durchgeführt, in dem analysiert wurde, ob sich in iClone durch eine doppelte Brennweite auch die absoluten Parallaxen bzw. die relative Parallaxe zwischen Nah- und Fernpunkt verdoppeln.

Der Abstand zum Nahpunkt betrug dabei weiterhin 2 m. Die Konvergenz blieb während des gesamten Tests einen Meter vor dem Nahpunkt, somit lag alles hinter dem Bildschirm. Es wurden die Brennweiten 25 mm, 50 mm und 100 mm verwendet.

Objekte	25 mm	50 mm	100 mm
Fuchs 2 m	23 px	48 px	95 px
Kiste 4 m	34 px	69 px	137 px
Couch 8 m	39 px	80 px	158 px
Roboter 16 m	42 px	85 px	171 px
Ruine 32 m	44 px	88 px	175 px
Rel. Parallaxe	21 px	40 px	80 px

Tabelle 5: Objektparallaxen vs. Brennweiten der virtuellen Kameras (iClone)

Unter Berücksichtigung der Messtoleranz kann eindeutig erkannt werden, dass sich bei Verdopplung der Brennweite die absolute sowie die relative Parallaxe verdop-

pehn. Der Unterschied gegenüber einer Stereobasis ist, dass sich auch der Bildausschnitt erheblich verändert, wenn die absolute Tiefe vergrößert werden soll. Wenn eine größere Tiefe mit der Stereobasis realisiert werden könnte, würde der Bildausschnitt weitestgehend unverändert bleiben können.

2.3.2 Workflow-Analyse

Die Workflow-Analyse von iClone (5.0 in der Pro Version) ergab Folgendes:

Personalisierung der Charaktere:

- Erstellung eigener Charakterköpfe mithilfe von Fotos
- Veränderbares Aussehen der Charaktere (Figur, Frisur, Gesichtspartien und Alter)
- Eigener Kleidung einstellbar oder Presets wählbar

Animation:

- Frei erstellbare Animationen für Körper und Gesicht
- Mimik manuell veränderbar oder Rückgriff auf einige Presets
- Lippensynchronisation
- Große Auswahl an vorgefertigten Animationen
- Animationen kombinierbar
- Motion-Capturing mit der Microsoft Kinect möglich

Kamera und Licht:

- Freie Positionierung und Animation der Kamera und des Lichts und all ihrer Parameter
- Anzahl der Kameras frei wählbar
- Kamerafixierung auf ein Objekt während der Fahrt und im Stillstand
- Brennweite stufenlos oder per Preset einstellbar
- Fokus und Schärfebereich einstellbar

Sonstiges:

- Hauseigene Requisiten und Sets
- Animationen von Objekten
- Eine immense Anzahl an Objekten über Trimble 3D Warehouse (3DXchange nötig)
- Import externer Objekte, Figuren und Animationen
- Visual Effects
- Viele Rendermöglichkeiten für Stereo-3D: Anaglyphen, Side-by-Side, Over-Under, Dualstream
- Messungen der Objektentfernung möglich

Der größte Workflow-Nachteil von iClone ist, dass keine Echtzeit-Ansicht in Stereo-3D möglich ist bis auf ein Vorschauenfenster in Anaglyphen für ein Standbild. Das dient zwar zur Orientierung, letztendlich mussten trotzdem Testbilder gerendert werden, um sie final mit einer Shutterbrille beurteilen zu können, da der Tiefeneindruck im Vergleich zur Vorschau deutlich unterschiedlich ausfällt.

3. Prototypische Stereo-3D-Machinima Produktion

Trotz der Tatsache, dass die Stereobasis nicht eingestellt werden kann, wurde sich für iClone als Produktionsmittel für den zu realisierenden Film entschieden. Ausschlaggebend waren die Einfachheit, Flexibilität, Freiheiten sowie der Funktionsumfang, den die Umgebung bietet. Außerdem ähnelte der Grafik-Stil von iClone am ehesten dem Aussehen der realen Welt.

Der Machinima sollte so produziert werden, dass er auf einer Bildschirmgröße von 17,3 Zoll Diagonale ohne Wahrnehmungsprobleme betrachtet werden kann.

3.1 Benötigte Zusatztools

Da es das Ziel war, einen Machinima zu produzieren, der in jeglicher Hinsicht dem Stereo-3D-Realfilm »Die Rolle seines Lebens« nahekommt, musste neben iClone zusätzlich auf weitere Programme zurückgegriffen werden.

Bevor überhaupt mit iClone gearbeitet werden konnte, mussten Requisiten und Sets gebaut werden. Dies geschah mithilfe der Modellierungssoftware Cinema 4D. Um Fragen wie »Welche Requisiten und Texturen beinhalten die Sets?« oder »Wie sind dessen Ausmaße?« beantworten zu können, wurde die erste Schnittfassung des Ste-

reo-3D-Realfilms zu Hilfe gezogen. Somit konnte ein Überblick über Locations und Sets geschaffen werden. Darauf aufbauend wurden Screenshots einzelner Sets erstellt und diese in Cinema 4D nachmodelliert. Alle Maße der Objekte und Sets wurden dabei abgeschätzt. Die fehlende Kompatibilität einiger Softwarefunktionen brachte Veränderung im Aussehen der Requisiten mit sich. So konnte nicht mit dem Extrude-Nurbs von Cinema 4D gearbeitet werden, da dessen erstellte Polygone in iClone falsch dargestellt wurden. Auch jegliche Shader von Cinema 4D konnten nicht in iClone importiert werden. Somit mussten passende Texturen im Internet gesucht werden. Da die Modelle in Cinema 4D als obj-Dateien exportiert werden mussten, wurden alle vorher eingefügten Texturen, nicht mit exportiert. Die Lösung war, die Modelle zusätzlich mit allen Bestandteilen abzuspeichern. Konnten Objekte aus verschiedenen Gründen nicht selbst erstellt werden, gab es die Möglichkeit, diese aus dem Trimble 3D Warehouse herunterzuladen.

Des Weiteren wurde 3DXchange benötigt. Es stellt eine Brücke zwischen Cinema 4D bzw. Trimble 3D Warehouse und iClone her, um Objekte externer Programme für die Nutzung in iClone umzuwandeln. In 3DXchange muss somit jedes extern erstellte Objekt geladen und jegliche Texturen erneut auf die Objekte gelegt werden. Das gilt jedoch nicht für Objekte aus dem Trimble 3D Warehouse. Im letzten Schritt wurden diese nur noch exportiert bzw. in iClone importiert.

3.2 iClone-Workflow

Der erste Schritt ist die Erstellung der Charaktere. Hier wurden aus der relativ kleinen Auswahl an Kleidung die passendsten Stücke ausgewählt. Anschließend wurden die Körperproportionen und Gesichtszüge nach Vorbild der Personen aus dem Realfilm verändert. Nachdem das jeweilige Set mit allen Requisiten positioniert ist, können die Charaktere auf Position gebracht und animiert werden. Kameras werden eingefügt und je nach Einstellung animiert. Anschließend werden Lichter aktiviert, positioniert und die ersten Farbeinstellungen vorgenommen. Durch diese Maßnahmen ist gut zu erkennen, wie eine Animation und die Einstellung im Allgemeinen wirken. Hierbei wurde sich am Realfilm orientiert, um die Einstellungen in etwa gleich lang zu halten. Für die Animationen wurde ein eigener Workflow entwickelt.

3.2.1 Schritt 1 – Grundbewegung

Stand oder saß der Charakter in einer Szene, wurde eine natürliche Ruhebewegung eingefügt. Sollte der Charaktere in einer Szene gehen bzw. laufen, wurde mit eben dieser Ruheanimation begonnen und anschließend eine Laufanimation hinzugefügt. iClone bietet hierfür verschiedene Presets an, die für jede Einstellung verwendet wurden.

3.2.2 Schritt 2 – szenenspezifische Bewegung

Hier wurden geplante Bewegungen der einzelnen Körperteile animiert. Dazu half ein Skelett aus Bones innerhalb jedes Charakters. Durch Anklicken des Animationsbuttons konnte auf die Bones bzw. die Körperteile zugegriffen und diese nach Wunsch bewegt werden.

3.2.3 Schritt 3 – Mimik

Die Mimik wurde meist über Presets eingefügt, wusste aber häufig nicht zu überzeugen. Eigene Versuche diese zu verändern scheiterten an der zu grobstufigen Steuerung.

Während der Produktion traten Limitierungen zutage, die umgangen werden mussten - eine typische Eigenschaft für Machinima. Aus diesem Grund beinhaltet der Film neue Einstellungen und mehr Schnitte als ursprünglich im Drehbuch vorgesehen. Vor dem Rendern wurden letzte Einstellungen an der Animation und dem Licht vorgenommen.

Beim anschließenden Rendern wurde sehr schnell klar, wie sich die fehlende Parametrisierung der Stereobasis bemerkbar macht. Die ersten Stereo-3D-Renderings waren zum Großteil tiefenübertrieben. Die absoluten Parallaxen erstreckten sich teilweise über den halben Bildschirm. Eine justierbare Stereobasis hätte vermutlich viele der aufgetretenen Probleme vermeiden können, vor allem innerhalb kleiner Sets. Des Weiteren könnte dadurch einfacher die absolute Tiefe verringert werden. In diesem Fall war es nur möglich, verschiedene Werte auszuprobieren und das finale Ergebnis zu beurteilen. Eine kleine Hilfe war auch hier die Zweckentfremdung eines Parameters. Über den Fokus, den man an ein Ziel ausrichten kann und so eine Entfernung erhält, konnte der Nahpunkt bestimmt werden. So war es möglich diesen Wert für die Konvergenz zu übernehmen und so zu gewährleisten, dass jegliche Objekte hinter der Leinwand liegen. Aber auch das Gitter auf dem Boden der Umgebung war für eine Entfernungsbestimmung geeignet. Jedes Quadrat hatte die Maße 100 x 100 (1m x 1m), somit war eine grobe Entfernungsmessung möglich. Um die Auswirkungen verschiedener Einstellungen der Parameter auf den Stereo-3D-Effekt besser einschätzen zu können, wurden einige Einstellungen des Films gewählt und drei verschiedene Szenarien durchgespielt:

Wie verändert sich das Stereobild wenn ...

... der Abstand zum Nahpunkt und die Konvergenz gleich bleiben und nur die Brennweite verändert wird? Je größer die Brennweite ist, desto größer die absolute und relative Parallaxe. Schlussfolgerung: Lieber mit kleineren Brennweiten arbeiten.

... sich der Abstand zum Nahpunkt vergrößert, die Konvergenz jedoch immer gleich weit vom Nahpunkt entfernt liegt und die Brennweite nicht verändert wird? Je größer der Abstand, desto kleiner ist die absolute und relative Parallaxe. Schlussfolgerung:

Es sollten eher größere Abstände gewählt werden.

... der Abstand, die Konvergenz und die Brennweite linear vergrößert werden? Je größer alle Werte werden, desto kleiner sind die absoluten und relativen Parallaxen. Schlussfolgerung: Für kleinere Parallaxen, alle drei Parameter gemeinsam um den gleichen Faktor vergrößern.

Mit diesen Erkenntnissen war es nun schneller möglich, ein angenehmes Stereobild zu realisieren. Leider mussten so gut wie alle Kameraparameter und viele Kamera-Abstände neu angepasst werden, da zu schnell zu große Parallaxen entstanden. Alle Brennweiten mussten verkleinert werden. Es kam sogar vor komplette Sets neu positionieren bzw. skalieren zu müssen, da die Entfernung des jeweiligen Fernpunktes zu groß war. Außerdem konnten einige im Storyboard festgelegte Bildausschnitte nicht realisiert werden. Viele Fehler konnten behoben werden, wenn auch nicht alle. So liegen die Parallaxen des Öfteren im Grenzbereich, da das Set nicht mehr erlaubte. Ein weiterer Punkt ist die Dauer des Renderings. iClone konnte mit der Echtzeitfähigkeit punkten, jedoch benötigte das Rendern einiger Einstellungen zwischen 1-2 Stunden. Bei dem Versuch eine Plansequenz zu realisieren, musste gar sechs Stunden gerendert werden, nur um das Ergebnis begutachten zu können. Wenn diese nicht gelungen war, hatte man dadurch mehrere Stunden verloren und musste einen neuen Weg gehen. Somit erstreckte sich die komplette Prozedur des Renderings durch das wiederholte Anpassen der Stereoparameter und das erneute Rausrechnen auf knapp eine Woche. So konnten im Schnitt pro Tag meist nur 3-4 Einstellungen mit akzeptablen Ergebnissen gerendert werden. Der Bedarf einer Stereo-3D-Vorschau in Echtzeit ist daher zwingend notwendig. Für einfache Einstellungen ohne Kamerafahrten oder Brennweitenveränderungen ist die Anaglyphen-Betrachtung im Preview-Fenster dennoch in gewissem Maße ausreichend.

Verbesserungen zum Vorgehen konnten durch diese Erfahrung ebenfalls gemacht werden: Sets sollten nicht als eine Einheit in iClone importiert werden. Dadurch steigt zwar der Zeitaufwand, aber es gibt mehr Spielraum beim Stereo-3D-Rendering, da jedes Objekt einzeln nachträglich verschoben werden kann. Ebenso sollten Sets großzügig aufgebaut werden. In meinem Fall musste die Brennweite deutlich verkleinert werden, da jedoch das Set nach einigen Metern bereits aufgehört hat, war dieses Ende sichtbar. Außerdem sollten keine komplexen Texturen gewählt werden, wenn keine kleinen Parallaxen möglich sind, wobei auch bei kleineren Werten die Textur irritierend wirken kann.

Parallaxen sollten bereits während des Set-Aufbaus ausgewertet werden, so können schnell, und ohne Mehraufwand Objekt- und Kamerapositionen und Brennweiten angepasst werden. Das war ein großes Problem bei der eigenen Produktion, da nicht mit solch enormen Werten gerechnet wurde. Eine Stereobasis ist auch aus diesem Grund notwendig.

4. Fazit

Im Großen und Ganzen ist das Thema Stereo-3D auch bei Machinima angekommen. Die hier analysierten Produktionsumgebungen bieten (mit Ausnahme von iClone) die gleichen Stereo-Parameter zur direkten Manipulation an: Konvergenz und Stereobasis.

Auch wenn damit relativ kostengünstig ein Stereo-3D-Machinima produziert werden kann, ist die wahrnehmungsgerechte Produktion eines solchen immer noch dem Trial & Error-Prinzip unterworfen.

Generell ist die Einarbeitungszeit zunächst gering, somit kann schnell ein Ergebnis geliefert werden. Viele der Einschränkungen bzw. Hürden, die mit Stereo-3D-Realfilm-Produktionen einhergehen, brauchen bei Machinima nicht beachtet zu werden (z. B. das korrekte Alignment der Kameras). Der Unterschied, ob man ein 2D- oder Stereo-3D-Machinima produziert, ist damit letztendlich kleiner als bei 2D vs. Stereo-3D-Realfilmproduktion.

Wenn Spiele verwendet werden, die in der Kompatibilitätsliste von NVIDIA derzeit als gut (oder besser) bewertet wurden, fallen einfache Stereo-3D-Filmsequenzen in diesen Engines gut bis beeindruckend aus. Aber das ist kein Machinima im klassischen Sinne.

Der Source Filmmaker bietet sehr viele Machinima-Funktionen und es gibt einige Assets, die von der Community bereitgestellt werden. Der Editor wurde gut durchdacht und kann - mit kleinen Einschränkungen - als All-In-One Lösung angesehen werden. Es bietet alles von der Animation, über Lighting bis hin zur Verarbeitung von Audiospuren. In Bezug auf die Stereoskopie bietet es dieselben Parameter wie NVIDIA, nur sind noch Kameraparameter gegeben, die verändert werden können, wodurch mehr Möglichkeiten in stereoskopischer Hinsicht vorhanden sind. Durch die fehlende Echtzeitfähigkeit zur Betrachtung des Stereo-3D-Effekts geht allerdings viel Produktionszeit verloren. In diesem Punkt gibt es somit noch Nachholbedarf, um Produktionen effizienter gestalten zu können.

iClone bietet über viele Einstellmöglichkeiten viele Freiheiten und verfügt über ein gutes Animationsrepertoire. Es erhebt den Anspruch einer All-in-One Lösung. Leider kann direkt nur die Stereo-Konvergenz parametrisiert werden. Es muss jedoch gesagt werden, dass es über Brennweite, Abstand, Modellierung, etc. durchaus möglich ist, ein angenehmes stereoskopisches Bild zu erzeugen - auch ohne einstellbare Stereobasis. Dafür sollte das Set maximale Flexibilität bieten. Der Austausch mit anderen Programmen ist gut durchdacht und läuft einwandfrei. Außerdem bietet der Hersteller Reallusion viele Erweiterungen kostengünstig an. Auch zur Prävisualisierung von aufwendigen S3D-Filmsequenzen ist iClone anscheinend gut geeignet.¹¹

¹¹ Ein Beispiel gibt es zum Film „Gravity“.

Teil 8
Web-Based-Training für stereoskopische
Bildgestaltung

Nadine Leonhardt

Abbildungsverzeichnis

Abbildung 1: Struktur der Lernplattform (eigene Darstellung)	312
Abbildung 2: Studioaufbau des Lernvideos (eigene Darstellung)	320
Abbildung 3: Eingabefelder für Diagonale und Seitenverhältnis mit dem errechneten Betrachtungsabstand (eigene Darstellung).....	322
Abbildung 4: Javascript für die Berechnung des Betrachtungsabstandes (eigene Darstellung).....	323

Im Jahr 2011 wurde in deutschen Kinos ein Umsatz von 958,1 Millionen Euro generiert (FFA 2013), davon allein 268 Millionen Euro durch Stereo-3D-Filme (FFA 2012, S. 29).

Viele Produktionsfirmen setzen immer mehr auf Stereo-3D-Filme. Sie sind nicht nur Publikumsmagnet, sondern bringen auch mehr Geld ein, da ein höherer Eintrittspreis verlangt werden kann. Aus ökonomischer Sicht scheint es mehr als verständlich, dass immer mehr Filme in Stereo-3D gedreht werden und das am besten in immer kürzerer Produktionszeiten. Liefen in deutschen Kinos im Jahr 2009 (nur) zehn Stereo-3D-Filme an, so waren es 2011 schon 46 Filme. (Nielsen 2012, S. 20).

Im Moment befindet sich die Filmbranche aber noch am Anfang eines Stereo-3D-Hypes. Logische Konsequenz dessen ist, »[...] dass 90 bis 100 Prozent der Leute, die an einem Stereo-3D-Filmset arbeiten, vor dem jeweiligen Projekt, noch nie etwas mit Stereo-3D zu tun hatten [...]« so eine Schätzung des Stereografs Florian Maier. Es sei schwierig »für einen Filmemacher, der sich seit 20, 30 oder noch mehr Jahren mit einer gewissen Art Filme zu machen beschäftigt hat, komplett loszulassen und fast wieder bei null anzufangen« (Bolliger 2011, S. 45)

Das Betrachten von Stereo-3D-Filmen kann für die Zuschauer schnell unangenehm werden und zu körperlichen Beschwerden führen. Natürlich sind die Filmemacher grundsätzlich bestrebt, dies zu verhindern. Doch das Thema der wahrnehmungsgerechten Stereoskopie ist so komplex, dass vor einem Dreh eines Stereo-3D-Films kaum die Zeit bleibt, sich angemessen mit der Wahrnehmungsphysiologie bzw. -psychologie dahinter zu beschäftigen. Daher hat die Branche hierfür ein neues Berufsbild geschaffen: den Stereografen. »Der Stereograf ist der, der am Set plötzlich am meisten stört, da es oft genug vorkommt, dass das gesamte Set drehfertig ist und er plötzlich noch drei Minuten braucht oder noch eine Probe möchte [...]« weiß der Stereoscopic Supervisor Marc Briede aus eigener Erfahrung (Bolliger 2011, S. 27).

Da kann es schnell zu Konflikten kommen. Das erlebte auch der Kameramann Sebastian Cramer. Er betont, wie wichtig es sei, »[...] dass die Technik und auch die Theorie dahinter transparent (werden).« (Bolliger 2011, S. 31)

Wie jedoch bereits erwähnt, ist die Thematik der Stereoskopie derart komplex, dass dafür ein eigenes Berufsbild geschaffen wurde. »Einer Crew, die in Stereo-3D uner-

fahren ist, in zwei Minuten klarzumachen, warum die Person bitte noch einen Meter weiter von der Kamera wegmüsste, kann zum Problem werden [...] Die Szene wird eingerichtet und geprobt, und kurz bevor es heiß wird, stellt noch jemand eine Blume in den Vordergrund oder stellt einen Darsteller um. Das kann dann wieder bedeuten »ich brauch noch zwei Minuten [...]«, erklärt Marc Briede (Bolliger 2011, S. 27).

Wie aus diesem Zitat herauszulesen ist, ist es oft die Bildgestaltung einer Einstellung, die zum Problem für den stereoskopischen Effekt werden kann. Daher befasst sich dieser Bericht mit der Bildgestaltung in Stereo-3D-Realfilmen und wie diese im Vergleich zum klassischen 2D-Film verändert werden sollte. Da für die Bildgestaltung der Kinematograf der Hauptverantwortliche ist, richtet sich das Ergebnis der Arbeit vor allem an diese Berufsgruppe.

Der Stereograf Florian Maier ist der Meinung, »[...] dass ein 3D-Film keine gestalterischen Einschränkungen besitzt, sondern dass es einfach eine andere Art zu gestalten ist – eine völlig neue, auf die man sich natürlich einlassen muss.« (Bolliger 2011, S. 45)

Genau an dieser Stelle setzt diese Arbeit an. Ziel ist es, dem Kinematografen deutlich zu machen, dass er mit dem Stereografen zusammenarbeiten muss, auch wenn dieser sich in sein gestalterisches Schöpfen einmischt. Weiterhin soll es der Kinematograf zu verstehen wissen, warum er seine Art der Bildgestaltung vielleicht zu Gunsten des Stereo-Effekts verändern muss. Indes soll der Kinematograf aber auch erkennen, dass er nicht all seine Gewohnheiten ablegen muss und gewisse Gestaltungsmerkmale aus 2D-Filmen auch in Stereo-3D-Filmen eingesetzt werden können, bzw. sogar neue Möglichkeiten erhält. Daher ergibt sich für diesen Bericht folgende Fragestellung:

»Wie kann rechtzeitig das Verständnis von Kinematografen für die geänderten Rahmenbedingungen von Stereo-3D-Realfilmproduktionen geschaffen werden?«

1. Berufsbilder

In diesem Beitrag dreht sich vieles um die Zusammenarbeit zwischen dem Kinematografen und dem Stereografen eines Stereo-3D-Films. Während der Beruf Kinematograf schon fast so lange existiert, wie Filme produziert werden, wurde der Beruf des Stereografen erst in den letzten Jahren geschaffen. Beide Berufsbilder werden im Folgenden kurz vorgestellt.

1.1 Der Kinematograf

So unterschiedlich die Wege in diesen Beruf führen, so unterschiedlich sind auch die Sichtweisen auf das Berufsbild des Kinematografen. Um eine einheitliche Vorstellung von diesem Berufsbild zu bekommen, stützt sich die Beschreibung in dieser Arbeit auf den Berufsverband Kinematografie, der in einer Broschüre die verschiedenen Be-

rufsbilder der Abteilung Kamera detailliert aufführt und diese seit 1983 immer aktuell hält¹.

Kurz und knapp beschreibt der Berufsverband die grundsätzliche Aufgabe des Kinematografen derart: »Aufgabe des Kinematografen ist die eigenschöpferische und eigenverantwortliche Bildgestaltung der Filmwerke in Zusammenarbeit mit der Regie und ggf. mit der Ausstattung. Sie umfasst sowohl die künstlerische Gestaltung als auch die technische Mitarbeit bei der Filmherstellung.« (BVK 2013, S. 9)

Dabei ist der Kinematograf aber mit der Technik nur am Rande befasst; darum kümmert sich meist ein dementsprechend qualifiziertes Personal. Das Hauptaugenmerk des Kinematografen liegt auf der kreativen Bildgestaltung als Ergebnis seiner schöpferischen Phantasie. Dabei spielt er eine wesentliche Rolle: »In seiner Verantwortung liegen damit auch die Emotionen der Zuschauer, ihre Aufmerksamkeit, Identifikation, Freude, Trauer oder ihre Angst. Erfolg oder Misserfolg eines Films werden deshalb durch die Entscheidungen des Kinematografen maßgeblich beeinflusst.« (BVK 2013, S. 11)

Der Kinematograf ist in allen Phasen der Produktion tätig. Nicht zuletzt trägt er eine Mitverantwortung bei der Budgetkontrolle, dass die Kosten seines Bereichs den vorgegebenen Rahmen nicht überschreiten. Dazu zählt natürlich auch die Einhaltung des Drehplans. Gerade in diesem Punkt besteht ein besonderes Konfliktpotential während einer Stereo-3D-Produktion, da durch die notwendigen Berechnungen und Anpassungen des Stereografen schnell eine zusätzliche Zeitkomponente entstehen kann.

1.2 Der Stereograf

Bei einer Stereo-3D-Filmproduktion fallen neue Aufgaben an, für die ein spezielles Wissen über die Stereoskopie und die menschliche Wahrnehmung benötigt wird. Somit wurde mit dem Stereographen ein neuer Beruf geschaffen. Eine anerkannte berufliche Ausbildung gibt es allerdings bis dato noch nicht und daher auch keine allgemein gültige Beschreibung des Berufsbildes. Prototypisch sei eine Zusammenfassung der wichtigsten Aufgaben von Kameramann Matthias Bolliger genannt, der sich seit einigen Jahren mit dem Thema befasst und schon Erfahrung mit Stereo-3D-Filmproduktionen hat. Er veröffentlichte dies in einer Ausgabe von »Film & TV Kameramann«: (2011, S. 25).

Der Stereograf hat einen ähnlich großen Einfluss auf das Endprodukt des Films wie der Regisseur und der Kinematograf. Eine enge Zusammenarbeit mit allen Kreativen ist während allen Produktionsphasen notwendig. Er berät den Kinematografen im Vorfeld über die genutzte Kamerakonfiguration, die verwendeten Optiken, wie auch über das eingesetzte Stereo-3D-Rig. Neben der Planung des Arbeitsprozesses und des Workflows berät der Stereograf außerdem Regie und Kinematograf über das stereoskopische Bildkonzept und die technische und gestalterische Durchführbarkeit

¹ Stereo-3D-Filmproduktionen werden darin jedoch nicht berücksichtigt.

beim Dreh und in der Post-Produktion. Sofern möglich testet er dies in der Pre-Production-Phase.

Beim Dreh unterstützt der Stereograf Regie und Kinematograf bei der technischen Umsetzung des gewünschten Bildcharakters und bei der Umsetzung der zuvor besprochenen Tiefendramaturgie. Außerdem ist er verantwortlich für den technischen Aufbau des Stereo-3D-Rigs und dessen Kalibrierung. Weiter muss er die Stereokalibration mit Hilfe von Messtechniken und Berechnungen durchführen; auch dies geschieht in Absprache mit Regie und Kamera, um Verzögerungen durch unvorhergesehene Änderungen der Bildgestaltung zu vermeiden.

Es ist absehbar, dass es zwischen den Aufgabenfeldern des Kinematografen und des Stereografen viele Überschneidungen gibt. Gerade der Kinematograf ist eine solche Einmischung eventuell nicht gewohnt, denn das Berufsbild des Stereografen existiert noch nicht lange und auch ausschließlich in der Sparte des Stereo-3D-Films.

2. Typische Anfängerfehler und deren Vermeidung

Dieser Abschnitt befasst sich nochmals kurz mit den typischen Anfängerfehlern, die einem 2D-Kinematografen bei seinem ersten Stereo-3D-Film unterlaufen könnten. Natürlich können manche Effekte auch als bewußter Stilbruch deklariert werden. Allerdings sollte die Wirkung bestimmter Bildgestaltungselemente bekannt sein, da eine eventuelle Korrektur in der Post-Produktion nur schwerlich und kostenaufwendig umgesetzt werden kann: »If 3D was not properly handled at any given point, there's no way it can be fixed later for a reasonable price. ›We'll fix it in the post‹ is likely the single most expensive phrase ever pronounced in Hollywood (besides ›They won't sue us‹).« (Mendiburu 2009, S.36).

Besonders wichtig ist es für den Kinematografen hierbei aber auch, dass er ganz praktische Hilfestellung zur Vermeidung dieser Fehler bekommt. Daher wird in diesem Bericht für jede Fehlerquelle auch ein Vorschlag zur Best Practice gemacht.

Für diese Arbeit wurden ausschließlich potentielle Fehlerquellen ausgewählt, auf die der Kinematograf auch tatsächlich Einfluss hat. Daher betreffen die Beispiele nur die Bildgestaltung einer Szene. Anm.: Die Effekte, die im Folgenden beschrieben werden, beziehen sich immer auf eine parallele Ausrichtung der beiden Kameras, auch wenn die meisten hiervon auch bei konvergierenden Kameras auftreten – mal mehr oder weniger stark.

2.1 Gestalterische Unschärfe

Es werden drei Arten von Unschärfen unterschieden: Unschärfe im Hintergrund, Unschärfe im Vordergrund und Bewegungsunschärfe. Jede Unschärfe beeinflusst die stereoptische Tiefenwahrnehmung der Zuschauer und verschlechtert diese tendenzi-

ell². Dabei wirkt die Unschärfe im Hintergrund jedoch noch am wenigsten störend, denn auch beim natürlichen Sehen nimmt der Mensch den Hintergrund unscharf wahr, wenn dieser eine höhere Tiefe im Vergleich zu Position des fokussierten Punktes aufweist. In einem Stereo-3D-Film wirkt eine geringe Schärfentiefe so lange nicht störend, bis der Zuschauer versucht, im Hintergrund etwas zu erkennen. Seine Augen können dann nämlich nicht, wie im natürlichen Sehen, auf den Hintergrund akkommodieren. Noch kritischer ist ein unscharfer Hintergrund mit einem eintönigen Muster, da dieses durch die Unschärfe zu einer reinen Farbfläche verschwimmt - es wirkt nur noch zweidimensional (Mendiburu 2009, S. 112).

Jedoch kann man auch diesen Effekt für sich nutzen. Genau wie im 2D-Film kann durch gezielte Unschärfe im Hintergrund ein Objekt in der Fokusebene visuell hervorgehoben werden. So wird der Zuschauer gezwungen, genau dieses Objekt zu betrachten und nicht die Umgebung zu beachten (Kuhn 1999, S. 60).

Unschärfe Objekte im Vordergrund führen beim Zuschauer dagegen immer zu Irritation. Im natürlichen Sehen kommt es nie vor, dass sich im Blickfeld ein unscharfes Objekt im Vordergrund befindet. Der Betrachter passt kurz seine Akkommodation an, bewegt seinen Kopf, sodass er an diesem Objekt vorbeischaun kann. In Einzelfällen werden im realen Sehen unscharfe Vordergrundobjekte durch das Gehirn gelöscht, sodass man es ganz vergisst. Diesen Effekt können wir z.B. an uns selbst beobachten: Die eigene Nase, die eigentlich im Blickfeld beider Augen liegt, wird durch das Gehirn konsequent ausgeblendet.

Ist der Zuschauer erst einmal von einem unscharfen vordergründigen Objekt irritiert, so passiert es schnell, dass er den Bezug zum Film verliert und sich erst mühsam wieder einfinden kann. Kann man nicht vermeiden, dass Objekte im Vordergrund unscharf werden, so müssen diese Objekte meist in der Postproduktion aufwändig entfernt oder durch Ausnutzung von monokularen Tiefen-Effekte retouchiert werden (Letzteres z.B. durch Ablendung oder künstliche Verdeckung).

Durch Bewegungsunschärfe verwischt die horizontale Disparität und damit auch die stereoptische Tiefenwahrnehmung (Mendiburu 2009, S. 112). Bewegungsunschärfe entsteht dabei aber nicht nur durch Bewegung im Bild, sondern auch durch Bewegung der Kamera. Um eine geringere Bewegungsunschärfe zu erreichen, kann eine kürzere Belichtungszeit gewählt werden. Bei sehr kurzen Belichtungszeiten entstehen jedoch stroboskopische Effekte („judder“), die den Bildeindruck des Zuschauers ebenso stören.³

2 Näheres dazu im Beitrag „Unschärfe in Stereo 3D-Produktionen“ in diesem Band

3 Dagegen kann eine höhere Framerate helfen, die wieder zu einer Bewegungsverschmelzung der Bilder führt. Neben dem höheren Datenaufkommen und damit höherem Speicherbedarf liegt das Problem bei dieser Methode, dass Kinos, die bislang in 24fps projiziert haben, nun ihre Projektionstechnik teuer updaten müssten. So bleibt dem Kinematografen dann nur die Wahl zwischen Irritation durch weniger Tiefeneindruck oder Irritation durch stroboskopische Effekte. Bedingt durch Peter Jacksons „The Hobbit – An Unexpected Journey“ haben im Jahr 2012 viele Kinos auf HFR (High Frame Rate), also 48fps aufgerüstet, Untersuchungen der Disney Research zeigen allerdings, dass Frameraten von 120fps von Nöten wären. .

2.2 Nahpunkt und Fernpunkt

Das Prinzip des Nah- und des Fernpunktes ähnelt dem der Schärfentiefe in 2D-Aufnahmen. Es müssen die Entfernungen von der Kamera zum nächsten Objekt im Bild und zum fernsten Objekt gemessen werden. Dies ist auch für die Stereokalkulation ohnehin notwendig (Tauer 2010, S. 44).

Beim Fernpunkt ist darauf zu achten, dass der Tiefenumfang der Einstellung die Augen des Betrachters nicht in eine Divergenzstellung zwingt. Die relative Disparität zwischen Nah- und Fernpunkt sollte zusätzlich maximal 70 Winkelminuten betragen (in Abhängigkeit vom Betrachtungsabstand).

In 2D-Filmen sind Kamerafahrten, bei denen in sehr geringem Abstand zum Beispiel an einem Laternenpfahl vorbei gefahren wird, beliebt. Wird dadurch der zuvor eingerechnete Nahpunkt unterschritten, beeinträchtigt dies den Stereo-3D-Effekt und das Objekt muss in besonders schweren Fällen sogar nachträglich aus dem Bild entfernt werden (Mendiburu 2009, S. 101).

Dennoch kann eine solche Unterschreitung des Nahpunkts auch als Effekt angewendet werden. Für einen sehr kurzen Moment hervorgerufen, verursacht er eine sehr starke Raumwirkung: sogenannten Pop-ups (Tauer 2010, S. 193).

Wird der Nahpunkt durch Teile der Dekoration oder die Bewegungen eines Schauspielers unterschritten, so kann dies einfach durch Neuarrangieren vermieden werden. Schwieriger ist es, wenn weniger kontrollierbare Partikel wie Staub, Regen oder Pollen den Nahpunkt unterschreiten. Diese müssen aufwendig in der Postproduktion entfernt werden. Der schlimmste Fall tritt ein, wenn sich ein Objekt so nah am Stereo-3D-Rig befindet, dass es nur von einer Kamera aufgenommen wird. Das ruft beim Betrachter an dieser Stelle eine binokulare Rivalität hervor, welche unbedingt verhindert werden sollte (Tauer 2010, S. 360f).

2.3 Window-Violation

Ein Objekt, welches „angeschnitten“ in einer Tiefenebene vor der Leinwand positioniert ist, kann eine Window-Violation bzw. eine Rahmenverletzung verursachen. Im Beispiel einer Person, befindet sich diese am linken Rand des Bildes, ist in dem Bild für das linke Auge eventuell noch die Nase der Person zu sehen, während diese im Bild für das rechte Auge schon nicht mehr zu sehen ist. Das Gehirn des Zuschauers entscheidet dann, dass es nicht möglich ist, dass diese Person vor dem Bildschirm schwebt, und drängt das Bild zurück bis auf Leinwandebene (Mendiburu 2009, S. 80). Hier wiegt dann der monokulare Tiefenhinweis der Verdeckung schwerer als die patente Stereopsis (Tauer 2010, S. 264). Kaschieren lässt sich dieser Effekt bei einer Bewegung, wenn sich z.B. eine Person von der Seite in das Bild hinein bewegt. Hier wirkt sich der Effekt nicht negativ auf die Tiefenwahrnehmung aus, denn aus dem natürlichen Sehen ist es das Gehirn gewohnt, dass ein Objekt, das sich sehr nahe am

Gesicht vorbei bewegt, zunächst auch nur mit einem Auge zu sehen ist, bevor es in das Blickfeld des anderen kommt. Wenn die Bewegung schnell genug ist und das Objekt nicht mehr angeschnitten ist, bevor es der Zuschauer überhaupt erst erkennt, dass es sich vor der Leinwand befindet, so stellt eine Window Violation kein Problem dar. Ebenso unproblematisch sind Anschnitte am oberen und unteren Rahmen. In diesem Fall stellt sich das Gehirn vor, die Leinwand wäre nach vorne gebogen. Um ganz auf der sicheren Seite zu sein, kann eine Person auch nur am unteren Bildschirmrand angeschnitten sein, denn der Zuschauer ist es gewohnt, von einer Person direkt vor ihm nur den Oberkörper und nicht die Füße zu sehen (Mendiburu 2009, S. 80).

Der Effekt ist auch bei Großaufnahmen zu bedenken. Ist die Aufnahme so nah, dass der Kopf der Person durch alle vier Ränder des Bildes angeschnitten wird, muss diese Tiefenebene unbedingt hinter die Leinwandebene gesetzt werden (Mendiburu 2009, S. 94).

Für Dialoge werden in 2D-Filmen oft Over-Shoulder-Aufnahmen eingesetzt. Das ist in Stereo-3D eher problematisch, da die angeschnittene Person durch die oben aufgestellten Regeln immer eine Window-Violation verursacht. Soll sich diese Tiefenebene in den Raum vor der Leinwand erstrecken, wirkt es auf den Zuschauer seltsam, einen halben schwebenden Rücken mitten im Raum zu sehen. Will man die Tiefenebene nicht zumindest auf Leinwandebene setzen, so sollte man in diesem Fall besser den ganzen Rücken der Person zeigen (Mendiburu 2009, S. 97).

2.4 Gigantismus

In 2D-Filmen erkennt der Zuschauer die Größe von Objekten aufgrund seiner Seherfahrung durch das Verhältnis zu anderen Objekten im Bild. Der Zuschauer weiß, dass er eine Projektion der Wirklichkeit auf der Leinwand sieht. Wird in einem Close-Up ein Gesicht gezeigt, ist dem Zuschauer klar, dass diese Person nicht 50m groß ist, sondern dass sie einfach von sehr nah gezeigt wird (Mendiburu 2009, S. 100).

Auch wenn ein Close-Up in Stereo-3D gezeigt wird, spielt die Seherfahrung eine wichtige Rolle. Ist die Kamera nah am Gesicht platziert, so muss eine kleine Stereobasis gewählt werden. Dadurch erhält das Gesicht eine eher geringe Tiefe. Da der Mensch aber Dinge, die seinen Augen nahe sind, im natürlichen Sehen mit einer größeren Tiefe wahrnimmt, muss er bei einem solchen stereoskopischen Bild annehmen, dass das Objekt viel größer ist als in der Realität. Denn der Mensch ordnet einer geringen Tiefe eine weite Entfernung zu, aber dennoch erscheint das Gesicht auf der Leinwand sehr groß. So entsteht der Effekt des Gigantismus (Kuhn 1999, S. 47). Doch dieser Effekt wirkt in den seltensten Fällen störend und kann deswegen auch in Kauf genommen werden (Tauer 2010, S. 107). Wird er absichtlich besonders übertrieben, kann er auch zur Inszenierung eingesetzt werden.

2.5 Liliputismus

Der Liliputismus-Effekt tritt vor allem im Vorder- und Mittelgrund eines Bildes auf. Besonders zu beobachten ist der Effekt bei weiten Landschaftsaufnahmen mit Weitwinkelobjektiven. Dabei sind Nah- und Fernpunkt weit entfernt, sodass eine große Stereobasis gewählt wird, um überhaupt Tiefe sichtbar zu machen. Doch da der Zuschauer es gewohnt ist, dass Objekte, bei denen er Tiefe wahrnehmen kann, eine geringe Entfernung zu ihm haben, wirken die Objekte auf der Leinwand viel kleiner als in der Realität. Dabei ist aber noch zwischen dem Liliputismus- und dem Modell-Effekt zu unterscheiden. Bei Letzterem wirkt zwar alles klein, doch das Verhältnis zwischen Tiefe und Größe des Objekts wirkt real. Ein Modell kann dabei kleiner, aber auch größer sein wie in Wirklichkeit. Beim Liliputismus-Effekt dagegen wirkt dieses Verhältnis verzerrt (Kuhn 1999, S. 47). Will man diesen Effekt vermeiden, so kann man einen näheren Vordergrund schaffen und damit die Stereobasis verkleinern. So verschwindet der Liliputismus-Effekt und Tiefe kann trotzdem wahrgenommen werden. Jedoch kann es dadurch zum Kulissen-Effekt kommen, wie folgend erklärt.

2.6 Kulisseneffekt

Der Kulisseneffekt oder Cardboard-Effekt tritt bei gestauchter Tiefenwiedergabe auf. Daß lange Brennweiten den Raum „stauchen“ weiß jeder Photograph und natürlich stauchen sie im gleichen Maße das Volumen von Objekten. Im Extremfall wirken die Objekte in stereoskopischen Aufnahmen plötzlich wie Kulissen, da sie eine viel geringere wahrgenommene Tiefe besitzen als in der Realität (Tauer 2010, S.105).

Der Effekt tritt in der Praxis auf, wenn die Entfernung der Kamera zum Nahpunkt vergrößert wird, um die Stereobasis zu erhöhen und so ein größeres Tiefenbudget zu erreichen. Um den Bildausschnitt zu erhalten, den man sich vorstellt, wird dann eine lange Brennweite gewählt. Dadurch wird der Hintergrund vergrößert, was wiederum Auswirkungen auf die Parallaxen hat, denn der Hintergrund wird weiter nach hinten geschoben. Um aber trotzdem in der Komfortzone zu bleiben, muss die Stereobasis verkleinert werden. Daher lautet eine Faustregel, dass der beste Stereo-3D-Effekt mit einer KB-Brennweite von unter 50 mm erreicht werden kann und einer dementsprechend geringen Entfernung zum Objekt (Mendiburu 2009, S. 109f). Es sei noch darauf hingewiesen, daß dieser Effekt in der Postproduktion durch nachträgliches HIT (Horizontal Image Translation) unbedacht weiter verstärkt werden kann, insbesondere wenn Szenen mittels L/R-Verschiebung nachträglich „nach vorne“ gezogen werden.⁴

2.7 Teilbildkonflikte

Dieser Konflikt entsteht, wenn ein Stereo-Teilbild Informationen oder Anmutungen enthält, die im anderen Teilbild fehlen oder anders geartet sind (Helligkeit, Farbigkeit,

⁴ Besagter Effekt wird in einem weiteren Beitrag in diesem Band ausführlich erläutert.

Schärfe, etc.). Dies kann im visuellen Zentrum des Betrachters zu Problemen führen (Tauer 2010, S. 67). Teilbildkonflikte können bei Wasserreflexen, Spiegelungen bzw. unterschiedlich polarisiertem Licht oder Lens Flares auftreten, denn diese werden von den beiden Kameraperspektiven nie symmetrisch aufgenommen (Mendiburu 2009, S. 107).

Um diese Fehler auszugleichen, reagiert das Gehirn des Betrachters unterschiedlich, je nach Art und Stärke des Teilbildkonflikts. Auch im natürlichen Sehen kommt es zu Teilbildkonflikten, beispielsweise wenn ein sehr nahes Objekt nur von einem Auge gesehen wird. Doch diese bemerkt der Mensch nicht. Ein Auge ist dann dominanter als das andere und so schafft es das Gehirn trotzdem, eine zyklonische Fusion der beiden Teilbilder zu erreichen. Fusioniert werden die beiden Teilbilder durch binokulare Summation. Helligkeits-, Farb-, aber auch geometrische Unterschiede werden dabei gemittelt. Dabei ist die Toleranzgrenze dieses Mechanismus bei jedem Menschen individuell (Tauer 2010, S. 67ff).

Sind die Unterschiede aber zu groß, können die beiden Teilbilder nicht mehr summiert werden. Dann entsteht eine Rivalität zwischen den beiden Teilbildern und es findet ein Auswahlprozess für das bessere Teilbild statt. Das Schlechtere wird dann ausgeblendet. Dieser Auswahlprozess kann mehrere Sekunden andauern und auch ein Wechsel vom einen zum anderen Teilbild ist möglich. Im natürlichen Sehen können durch Suppression auch nur einzelne Stellen eines Teilbilds ausgeschaltet werden. Bewusst nimmt die der Mensch nicht wahr, da seine Augen durch die Sakkadenbewegung immer wieder einen neuen Punkt fixieren und auf ihn akkomodieren. Dann wird immer wieder ein anderer Teil unterdrückt (Tauer 2010, S. 72f). Bei stereoskopischen Abbildungen funktioniert die Suppression nicht, da notwendige Grundinformationen fehlen. Die Augen akkomodieren immer auf die gleiche Ebene, die Leinwand. Daher tritt in diesem Fall binokulare Rivalität schnell auf. Zu unterscheiden sind hierbei die lokale und die globale Rivalität. Die Augen tasten die Umgebung ununterbrochen ab und bei jedem Fixationspunkt wird entschieden, welches Teilbild unterdrückt wird. Sind aber die Unterschiede besonders groß, findet die Rivalität für das ganze Bild statt. Dann wird fast schon monoskopisch gesehen. Die Teilbilder werden aber ständig miteinander verglichen, denn der Wahrnehmungsapparat ist immer darauf aus, eine Fusion der beiden Teilbilder herbeizuführen (Tauer 2010, S. 74).

Da die Toleranzgrenzen des Wahrnehmungsapparats so unterschiedlich sind, ist es bei der Produktion eines Stereo-3D-Films ratsam, solche Teilbildkonflikte ganz zu vermeiden (Mendiburu 2009, S. 107). Gerade Spitzlichter, die ausbrennen und überstrahlen, bergen die Gefahr, nicht von beiden Kamerapositionen gleich aufgenommen zu werden. Auch reflektiertes Licht wird von beiden Kameras unterschiedlich aufgenommen, was insbesondere bei Spiegel-Rigs der Fall ist. Solche Fälle sind schon vor der Aufnahme einfach zu erkennen und sollten vermieden werden.

2.8 Schnitt

Der Stereograf eines Films entwickelt vor dem Dreh auf Grundlage des Drehbuchs ein Tiefenskript. In diesem ist festgelegt, wie viel des Tiefenbudgets zu jedem Zeitpunkt des Films ausgenutzt wird. Auch die Übergänge von einer Einstellung zur nächsten sind darin erkennbar (Mendiburu 2009, S. 88).

Während der Aufnahme müssen Stereograf, Kinematograf und Continuity darauf achten, dass alle Tiefenanschlüsse stimmen und später bei der Betrachtung nicht zu Irritationen führen. Deshalb muss schon vor der Aufnahme klar sein, wo die Aufmerksamkeit des Zuschauers liegen soll und wie die Szene aufzulösen ist. Störend wirkt sich ein Tiefensprung vor allem dann aus, wenn von einer Einstellung mit weit entferntem Vordergrund auf eine Einstellung mit nahem Vordergrund geschnitten wird. Dann müssen die Augen plötzlich – bei gleicher Akkommodation – eine starke Konvergenzstellung einnehmen (Tauer 2010, S. 275).

Aus diesem Grunde sollte auch die Schnittfolge nicht zu schnell sein. Muss sich die Stellung der Augen ständig ändern, wird das anstrengend für den Zuschauer. Daher werden in professionellen Tiefenskripten immer „flache“ Phasen zum Ausruhen der Augen vorgesehen. Um einen angenehmen Übergang von einer bestimmten Tiefe zu einer anderen Tiefe zu schaffen, kann auch während der Einstellung die Tiefe schon an die folgende Tiefe angepasst werden, sodass die Konvergenz der Augen geführt wird. Um dafür Platz zu haben, sollte eine Szene nicht zu hoch aufgelöst sein (Mendiburu 2009, S. 88). Das gilt auch deshalb, weil der Betrachter immer eine gewisse Zeit braucht, um sich im Stereobild zu orientieren.

3. Konzeption eines Web-Based-Trainings

Dieser Abschnitt beschäftigt sich nun damit, wie das benötigte Wissen an die 2D-Kinematografen herangetragen werden kann. Sicherlich könnte ein Kinematograf seine eigenen Erfahrungen mit jedem neuen Stereo-3D-Film sammeln und so irgendwann selbst zu diesen Erkenntnissen kommen. Doch gerade ein Stereo-3D-Film bringt oft ein großes Budget mit und entsprechend hoch ist auch der Verlust für die Produktionsfirma, sollte der Film nicht erfolgreich werden. Gerade weil der Kinematograf eine so hohe Mitverantwortung am Erfolg des Films trägt, sollten ihm keine Anfängerfehler unterlaufen. Deshalb soll eine Lernplattform skizziert werden, über die Kinematografen noch vor ihrer ersten Stereo-3D-Filmproduktion in ersten Kontakt mit dem Thema treten.

3.1 Zielgruppenanalyse

Zielgruppe des »Web-Based-Trainings (WBT)« sind Kinematografen, die schon seit vielen Jahren ihren Beruf ausüben und dabei nur an Filmproduktionen von klassischen 2D-Filmen mitgearbeitet haben. Werden diese dann zum ersten Mal für einen

stereoskopischen Film verpflichtet, sollen sie im Vorfeld diese Lernplattform besuchen. Auch jene Kinematografen, die schon einen oder mehrere Stereo-3D-Filme produziert haben und dabei den schon erwähnten Konflikten mit dem Stereografen begegnet sind, sind Zielgruppe des WBT. Eine genaue Analyse der Zielgruppe soll Anhaltspunkte für die praktische Umsetzung des WBT liefern.

Die didaktische Theorie bietet dabei eine Vielzahl von verschiedenen pädagogischen und psychologischen Diagnostiken an. Um für die Konzeption des WBT eine möglichst praktikable Auswahl an Merkmalen der Zielgruppe zu haben, wird hierzu die überschaubare Auswahl von Holz und Zimmer nach Dr. Dieter Euler (Euler 1992, S. 86) zu Grunde gelegt. Diese beinhaltet acht Merkmale, die nicht unbedingt durch umfangreiche Testverfahren, sondern auch intuitiv oder durch gedankliche Vorstellungskraft bestimmt werden können.

3.1.1 Situative Disposition

Es kann angenommen werden, dass die PC-Vertrautheit der Zielgruppe durchschnittlich hoch ist. Sie besitzen eine gewisse Routine im Umgang mit dem PC, die Hardware ist ihnen vertraut und auch der Dialog mit System- oder Anwendungsprogrammen ist ihnen bekannt. Tiefere Kenntnis über die Software eines PCs besitzen sie aber nicht.

Die Inhaltsvertrautheit der Zielgruppe ist eher gering einzuschätzen. Wird ein Kinematograf für einen Stereo-3D-Film verpflichtet, informiert er sich zwar vielleicht über die Thematik der Stereoskopie nicht nur auf diesem WBT, sondern auch über andere Wege, wie durch persönliche Gespräche mit Experten. Daher kann zwar davon ausgegangen werden, dass der Lernende schon allgemeine Kenntnisse über die Thematik besitzt, doch tiefere Kenntnis, die auch auf individuelle Situationen angewendet werden kann, fehlt ihm noch.

3.1.2 Affektive Disposition

Die Akzeptanz der Zielgruppe, neue Dinge über eine webbasierte Lernplattform zu lernen, wird als eher gering eingeschätzt. Kinematografen, die schon lange in ihrem Beruf arbeiten, sind es gewohnt, neue Dinge eher in der Praxis zu lernen. Daher sind sie es nicht gewohnt, etwas über eine Webseite zu lernen, vor allem, da ihr eigentlicher Beruf so gut wie nichts mit dem Computer oder dem Internet zu tun hat.

Der Lernanspruch der Zielgruppe ist eher emotional begründet. Kinematografen üben ihren Beruf meist aus Leidenschaft aus und lernen neue Dinge daher aus Interesse und Spaß. Teilweise kann der Lernanspruch aber auch sachlich sein, wenn der Besuch der Lernplattform vor einer Stereo-3D-Filmproduktion mit Zeitdruck verbunden ist.

Lernerfolgsdiagnosen werden von der Zielgruppe aus diesem Grund eher persönlich interpretiert. Ihre Art der Bildgestaltung eines Films ist oft sehr individuell und selbst erarbeitet. Daher werden Hinweise auf Fehler leicht als persönliche Kritik aufgefasst.

3.1.3 Kognitive Disposition

Es wird angenommen, dass die Zielgruppe einen aktiven Lernstil besitzt. Probleme werden so aufgenommen, dass eine selbstständige Lösung gesucht wird; die Zielgruppe kennzeichnet also eine explorative Grundhaltung. Neue Lernimpulse werden als Ausgangspunkte für eigene Überlegungen genommen und nicht als fixe Aussagen.

Die bevorzugte Wahrnehmungsform der Lernenden ist nicht abstrakter, sondern gegenständlicher Art. Kinematografen lernen lieber durch Ausprobieren als durch theoretische Anleitung. Da sie in ihrem Beruf selbst Filme produzieren, liegt es nahe, dass sie Inhalte, die in Filmen dargestellt werden, auch einfacher lernen können.

Die Informationsverarbeitungsfähigkeit der Zielgruppe wird als eher hoch eingeschätzt. Es kommt bei einer Filmproduktion oft vor, dass neue Informationen in sehr kurzer Zeit aufgenommen, verarbeitet und behalten werden müssen. Auch das Organisieren von Informationen liegt Kinematografen.

3.2 Lehr-/Lernziele

Um Lernziele richtig definieren zu können, müssen zunächst die beiden Begriffe Lehrziel und Lernziel unterschieden werden. Dabei wird die Unterscheidung von Schübach et al. (2003, S. 17) verwendet:

»Lehrziele bezeichnen das, was die Lehrenden mit ihrer Tätigkeit, dem Lehren, erreichen möchten. Im E-Learning übernimmt das Computerprogramm die Aufgabe des Lehrens. Lernziele bezeichnen das, was die Lernenden mit ihrem Tun, dem Lernen, erreichen möchten«

Aus dieser Unterscheidung ergibt sich auch die Wichtigkeit der aktiven Tätigkeit des Lernenden. Im Idealfall entsprechen sich Lehr- und Lernziele, doch für den Lernenden ist primär das Erreichen der Lernziele von Bedeutung (Schübach et al. 2003, S. 18).

Es gibt mehrere Möglichkeiten, die Lernziele hierarchisch anzuordnen und zu klassifizieren. Zunächst sei hier eine Top-Down-Analyse angewendet, die ein großes Lernziel in schrittweise immer kleinere Lernziele unterteilt (Wendt 2003, S. 102).

Das Richtziel des WBT lässt sich so formulieren: Der Lernende ist in der Lage, die Bildgestaltung eines Stereo-3D-Films in Absprache mit dem Stereografen so zu verändern, dass sie dem neuen Tiefeneffekt angepasst ist.

Daraus lassen sich folgende Grobziele ableiten: Der Lernende erkennt, dass er seine Art der Bildgestaltung ändern muss. Der Lernende kann selbst Argumente formulie-

ren, um mit dem Stereografen eine gemeinsame Diskussionsgrundlage über die Bildgestaltung in Stereo-3D-Filmen zu haben. Der Lernende kann den neuen Tiefeneffekt in sinnvoller Weise nutzen.

Die letztlichen Feinziele werden daraus resultierend so formuliert: Der Lernende ist in der Lage, Problemen, die durch die Bildgestaltung für den Tiefeneffekt auftreten können, vorzubeugen. Dazu weiß der Lernende, wodurch solche Probleme auftreten können. Der Lernende trifft technische Entscheidungen auch im Hinblick darauf, den Tiefeneffekt nicht negativ zu beeinflussen. Da diese Entscheidungen aber immer auch in Absprache mit dem Stereografen erfolgen, versteht der Lernende, was der Stereograf zu ihm sagt und warum er diese Hinweise gibt. Der Lernende weiß, was die Hinweise des Stereografen für die Bildgestaltung bedeuten und kann diese umsetzen. Der Lernende kann aber auch selbst Lösungen anbieten und mit dem Stereografen gleichberechtigt diskutieren. Der Lernende kennt außerdem die Effekte, die bei einer bestimmten Art der Bildgestaltung einer Einstellung entstehen. Diese Effekte kann der Lernende deshalb auch bewusst einsetzen. Die Entscheidungen über Brennweite, Schärfentiefe, Bildausschnitt und die gesamte Bildgestaltung trifft der Lernende unter Berücksichtigung des Tiefeneffekts und in unsicheren Fällen in Absprache mit dem Stereografen.

Weiterhin lassen sich Lernziele auch in die drei Bereiche kognitive, affektive und psychomotorische Lernziele einordnen (Euler 1992, S.90). Als kognitive Lernziele lassen sich für die Lernplattform folgende festhalten: der Lernende weiß, welche Effekte durch eine bestimmte Art der Bildgestaltung in der Tiefendimension auftreten können. Der Lernende versteht den Grund dafür. Der Lernende kann diese Effekte bewusst vermeiden oder anwenden. Der Lernende kann Situationen vor und während den Dreharbeiten daraufhin analysieren, ob Probleme für den Tiefeneffekt auftreten. Er kann das Gelernte auf die jeweilige Situation synthetisieren. Außerdem kann er evaluieren, inwieweit er mit seiner Art der Bildgestaltung dazu beiträgt, welche Qualität der Tiefeneffekt in der Einstellung haben wird.

Auch affektive Lernziele lassen sich für die Lernplattform bestimmen. Der Lernende akzeptiert und nimmt auf, dass sich die Bildgestaltung eines Stereo-3D-Films von der eines 2D-Films unterscheidet. Der Lernende wertet das Gelernte danach, ob er seine Bildgestaltung zu Gunsten des Tiefeneffekts ändert oder, ob ihm ein guter Tiefeneffekt weniger wichtig ist, als seine Art der Bildgestaltung beizubehalten. Entsprechend reagiert er in den verschiedenen Situationen. Er ordnet die Werte danach, welche Probleme für den Tiefeneffekt er eher in Kauf nimmt als andere. Schließlich behält der Lernende immer im Hinterkopf, dass eine bestimmte Bildgestaltung auch bestimmte Auswirkungen auf den Tiefeneffekt hat.

Psychomotorische Lernziele werden aber eher weniger verfolgt. Der Kinematograf kennt seine technischen Möglichkeiten schon aus 2D-Filmen und diese ändern sich bei einem Stereo-3D-Film nicht. Er wird auch nicht versuchen, das Gelernte 1:1 auf eine bestimmte Situation anzuwenden. Das heißt, er wird das Gelernte nicht imitieren.

Das Gelernte ist lediglich der Ausgangspunkt, von dem aus der Lernende für jede Situation eine individuelle Lösung ableiten kann.

3.3 Verbreitungsmedium

Bevor an einem konkreten Konzept für die Lernumgebung gearbeitet werden kann, soll zunächst überprüft werden, inwieweit die Umsetzung der Lernziele für die anvisierte Zielgruppe überhaupt für ein computergestütztes Lernen geeignet ist. Dabei stellt sich zum einen die Frage der Effektivität, also, ob die gesetzten Ziele bezogen auf die Zielgruppe so überhaupt erreichbar sind. Zum anderen muss überlegt werden, ob diese Lernziele durch die Methode des computergestützten Lernens besser zu erreichen sind als durch andere Methoden, ob die Methode also effizient ist. Ein Vergleich mit anderen Methoden würde zu diesen Fragen stichhaltige Antworten liefern. Doch Befunde dieser Art liegen hier nicht vor und daher werden die Fragen durch die Anführung plausibler erscheinender Gründe erörtert (Euler 1992, S. 94).

Es dreht sich bei den angestrebten Lernzielen immer um das gleiche Medium, nämlich Film. Daher ist es naheliegend, den Lerninhalt auch über das Medium Film zu vermitteln. Nicht nur ist die Zielgruppe sehr vertraut mit dieser Art der Darstellung. Auch können in einem Video die aufgezeigten Probleme direkt sichtbar gemacht werden, ohne dabei auf einem zu hohen Abstraktionsniveau zu bleiben, wie das beispielsweise bei einem Text der Fall wäre. Dabei ist es allgemein üblich, Videos am Computer anzusehen. Daher lässt sich sagen, dass die Methode des computergestützten Lernens zur Erreichung der Lernziele bezogen auf die Zielgruppe effektiv ist.

Vorteile gegenüber anderen Methoden bietet das computergestützte Lernen vor allem in der möglichen Darstellungsvariante des Videos. Da der Lerninhalt sich immer auf die Produktion von Videos bezieht, scheint es plausibel, das Video auch als Vermittlungsform zu wählen. Außerdem bietet das Internet die Möglichkeit der kostengünstigen und schnellen Verbreitung der Lernplattform. Diese Vorteile gegenüber anderen Methoden zeigen, dass eine webbasierte Lernumgebung für die Erreichung der Lernziele bezogen auf die Zielgruppe effizient ist.

3.4 Lehrstrategie

Wendt (2003, S. 106) unterscheidet vier verschiedene Lehrstrategien: Faktenwissen, intellektuelle Fähigkeiten, Verhaltensweisen und praktische Fähigkeiten. Für dieses WBT eignet sich die Strategie der Vermittlung von intellektuellen Fähigkeiten. Der Lernende soll kontextübergreifende Problemlösungsstrategien entwickeln und entsprechende Handlungskompetenzen erwerben. Im Lernvideo wird ein bestimmtes Prinzip demonstriert, das der Anwender auf einen anderen Kontext in seiner individuellen Situation bei einer Stereo-3D-Filmproduktion übertragen soll. Das gesamte Lernszenario ist so realitätsnah konstruiert, dass ein schneller und nachhaltiger Wissenstransfer auf wirkliche Problemkonstellationen möglich ist (Wendt 2003, S. 109).

Die Strategie der Vermittlung von Verhaltensweisen eignet sich dagegen kaum für die Lernplattform. Bei dieser Strategie werden ausschließlich Informationen dargestellt, die richtiges oder falsches Verhalten zeigen (Wendt 2003, S. 110). Das WBT soll den Anwender jedoch dafür sensibilisieren, dass es meist nicht nur ein richtiges oder falsches Verhalten gibt, sondern dass für jedes Problem eine individuelle Herangehensweise gefunden werden muss, durch die aber oft auch keine endgültige Lösung gefunden wird, sondern vielmehr Kompromisse eingegangen werden müssen.

Auch die anderen beiden Strategien eignen sich nicht. Eine reine Vermittlung von Faktenwissen reicht nicht aus, um den individuellen Problemen bei einer Stereo-3D-Filmproduktion zu begegnen. Die Vermittlung von praktischen Fähigkeiten ist aus naheliegenden Gründen über ein computergestütztes Lernen in diesem Bereich nicht möglich. Es können immer nur Fallbeispiele vorgestellt werden, doch der Anwender hat keinen Einfluss hierauf, wie er ihn unter realen Bedingungen hätte.

3.5 Lernwegstruktur

Es gibt verschiedene Möglichkeiten, den Lerninhalt darzustellen und unterschiedliche Zugangswege anzubieten. Laut Wendt (2003, S. 112) haben sich zwei Grundformen von Lernwegstrukturen etabliert: die offene und die vorgegebene Lernwegstruktur.

Kinematografen, die diese Lernplattform besuchen, bringen ein sehr unterschiedliches Vorwissen und unterschiedliche Grade der Motivation mit. Daher muss es dem Anwender möglich sein, jederzeit auf alle Inhalte zuzugreifen, egal, an welcher Stelle er sich gerade befindet und ob er eine Lerneinheit fertig absolviert hat oder nicht. Der Anwender kann jeden Lernbaustein so oft wiederholen, wie er will, und kann auch Inhalte überspringen. Dies führt zu einer Netzstruktur. Diese Grundform wird aber mit einer Variation der dualen Lernwegstruktur eingesetzt. Es ist nämlich auch möglich, das Lernmodul linear zu bearbeiten, denn der Anwender wird auch Schritt für Schritt vom Programm geführt. Die Lerninhalte bauen thematisch aufeinander auf, ohne jedoch zwingend voneinander abhängig zu sein. In Abbildung 1 ist die lineare Struktur der Lernplattform dargestellt. Die Linearität kann vom Anwender aber jederzeit unterbrochen werden, sodass eine offene Lernwegstruktur entsteht.

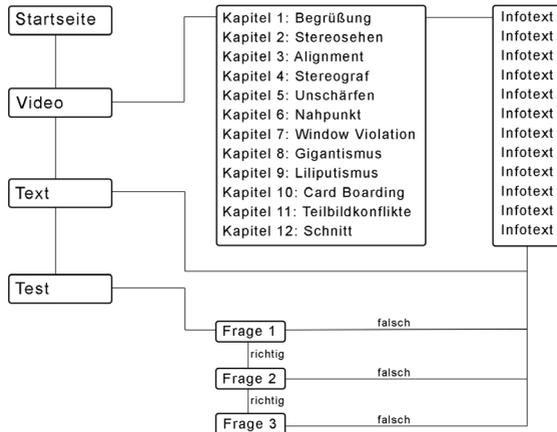


Abbildung 1: Struktur der Lernplattform (eigene Darstellung)

3.6 Darstellung des Lerninhalts

Im Weiteren müssen Überlegungen angestellt werden, wie der Lerninhalt für die Zielgruppe so aufbereitet werden kann, dass er für diese anschaulich und verständlich ist. Wie die Darstellung der Lerninhalte auf den Anwender wirkt, hängt zu einem großen Teil aber auch von den subjektiven Voraussetzungen jedes Einzelnen ab (Euler 1992, S. 102).

Es ist daher notwendig, die Darstellung des Lerninhalts sowohl auf die dem Lernenden bereits bekannten Inhalte als auch die ihm bekannten Darstellungsformen abzustimmen. Bei für den Anwender neuem Inhalt in neuer Form besteht die Gefahr der Überforderung. Bei bekanntem Inhalt und bekannter Form fühlt sich der Anwender schnell unterfordert. Ein bekannter Inhalt in neuer Form fördert das Erinnern. Für dieses WBT wird eine für den Anwender bekannte Darstellungsform gewählt, nämlich das Video, in der ein dem Anwender neuer Inhalt vermittelt wird. Mit dieser Verbindung wird die Wahrnehmung und Verarbeitung des Lernenden gefördert (Euler 1992, S. 103).

Durch die Darstellungsform des Videos ist auch eine hohe Lebensnähe zum Alltag des Lernenden geschaffen. So wird verdeutlicht, welche Bedeutung die dargestellten Informationen für eine sinnvolle Anwendung haben (Euler 1992, S. 121). Die Verbindung von altem und neuem Wissen wird dadurch gefördert, dass die Lerninhalte sich auf einen Erfahrungsbereich beziehen, der dem Lernenden bereits bekannt ist (Euler 1992, S. 122).

Ebenso wichtig ist die Transparenz der Darstellungsform. Es muss zu jeder Zeit für den Anwender kenntlich gemacht werden, wo er sich gerade befindet. Er soll die ganze Zeit über den ganzheitlichen Zusammenhang im Blick behalten (Euler 1992, S. 123).

Um dies zu gewährleisten, ist eine konsistente Ablaufsteuerung von Vorteil. Eine klare Bildschirmaufteilung, die sich auf jeder neuen Bildschirmseite gleich darstellt, verhindert, dass der Anwender sich immer erst wieder neu zurechtfinden muss (Euler 1992, S. 124). So findet sich im Menü eine klar benannte Auflistung aller Seiten, die der Anwender aufrufen kann. Überschriften machen dem Anwender immer deutlich, in welchem Abschnitt er sich gerade befindet. Den Rest des Bildschirms nimmt der Informationsbereich ein, in dem die Lerninhalte dargestellt werden.

Um den Lernenden nicht zu verunsichern, werden die Lernziele zu Beginn nur grob in einem Willkommenstext ausgewiesen. Eine zu detailreiche Erläuterung der Lernziele noch vor Beginn des Lernprozesses könnte den Anwender abschrecken (Euler 1992, S. 127).

3.7 Motivationsdesign

Die Motivation eines Anwenders kann sowohl extrinsisch als auch intrinsisch hervorgerufen sein (Schübach et al. 2003, S. 18). In diesem Fall mag der Anstoß zur Absolvierung des WBT extrinsisch von der Produktionsfirma oder dem Stereografen der Stereo-3D-Filmproduktion gegeben sein. Doch eine intrinsische Motivation wird sich schnell einstellen, denn der Lerninhalt betrifft direkt den Beruf des Kinematografen, der von diesem meist aus Leidenschaft ausgeführt wird.

Das Motivationsdesign der Lernplattform soll den Lernenden weniger motivieren als vielmehr verhindern, dass er demotiviert wird (Niegemann 2006, S. 72). Nach dem Motivationsdesign von Keller (1983, S. 395f.) sind es vier zentrale Merkmale, die für die Motivation des Lernenden verantwortlich sind: Interesse, Relevanz, Erfolgsszuversicht und Zufriedenheit.

Das Interesse und die Neugier des Anwenders werden dadurch geweckt, dass zunächst Situationen aus seinem persönlichen Berufsalltag vorgestellt werden. Dadurch fühlt er sich sogleich individuell angesprochen und identifiziert sich mit der Lernumgebung. Die Relevanz des Lerninhaltes wird für den Lernenden dadurch erkennbar, dass zunächst vertraute Situationen dargestellt werden, die dann in die Lernziele überführt werden. Ihm wird bewusst gemacht, dass es negative Folgen nach sich zieht, wenn er den Lerninhalt nicht umsetzt. Um den Lernenden stets eine Erfolgsszuversicht spüren zu lassen, wird immer wieder betont, dass es nicht das Ziel des WBT ist, dass er die Inhalte auswendig kennt. Stattdessen soll er dafür sensibilisiert werden, dass für eine Stereo-3D-Filmproduktion andere Rahmenbedingungen gelten, wobei der Kinematograf aber immer vom Stereografen unterstützt wird. Die Zufriedenheit des Anwenders stellt sich dadurch ein, dass er weiß, dass er seine alten Gewohnheiten nicht able-

gen muss, sondern lediglich leicht abändern sollte. Positive Folgen stellen sich dann während der Filmproduktion ein, wenn er versteht, was der Stereograf sagt und auch selbst Argumente formulieren kann, weil er die Thematik verstanden hat.

3.8 Interaktionsdesign

Interaktivität macht nur dann Sinn, wenn sie den Lernprozess fördert. Interaktionen müssen plausibel in die Lernanwendung integriert sein, sonst können sie sogar lernhemmend wirken (Niegemann 2006, S. 72). Öffnet der Anwender die Webseite, muss er zunächst einen Button klicken, der das WBT startet und im Vollbildmodus anzeigt. Die Webseite, auf der das WBT stattfindet, ist so gestaltet, dass der Content stets im Mittelpunkt steht. Da eine lineare Struktur des Lernmoduls vorgegeben ist, muss das Menü, in dem der Anwender einen Abschnitt auswählen kann, nicht die ganze Zeit sichtbar sein. Bei einem normalen Programmablauf ist dieses Menü gar nicht nötig. Daher ist das Menü an der Seite versteckt und der Anwender muss den Cursor erst darauf zu bewegen, bevor es erscheint. Dann kann er einen Abschnitt auswählen und seinen eigenen Lernweg wählen.

Am unteren Bildschirmrand befindet sich ebenfalls ein Kontextmenü, in dem zusätzliche Informationen zu den einzelnen Themen zu finden sind. Dieses kann der Anwender ebenfalls sichtbar machen, wenn er seinen Cursor dorthin bewegt. Der Benutzer benötigt eine Anaglyphenbrille, um das WBT in vollem Umfang nutzen zu können. Bevor er sich die einzelnen Kapitel des Lernvideos ansehen kann, muss der Anwender zunächst mit Hilfe eines Testfeldes die Bildschirmhelligkeit seines Monitors an seine Anaglyphenbrille anpassen. Dann muss er in ein Eingabefeld die Diagonale seines Monitors in Zoll eingeben und das Seitenverhältnis seines Bildschirms angeben. Daraus wird ein Betrachtungsabstand errechnet, den er für einen bestmöglichen Tiefeneindruck während des Lernvideos einnehmen sollte. Hat er dies getan, gelangt er mit einem Klick auf einen Button zum ersten Video, das von selbst startet. Der selbst entworfene Player bietet dem Anwender die Möglichkeit, das Video zu stoppen, wieder zu starten, zum nächsten oder zum vorherigen Video zu springen.

Alle Videos starten nacheinander von selbst. Der Anwender kann zu jedem Thema des Videos genauere Informationen erhalten, wenn er mit seinem Cursor den Informationsslider am unteren Bildschirmrand erscheinen lässt. Nach dem letzten Video erscheint ein Button, der den Anwender weiter zum Test führt. Dieser kann geklickt werden, will der Anwender den Test machen. Dies wird im nächsten Kapitel genauer behandelt.

Das WBT, das die ganze Zeit im Vollbildmodus angezeigt wird, kann durch Drücken der Esc-Taste verlassen werden.

3.9 Aufgaben

3.9.1 Aufgabentyp

Um den Lernerfolg des Anwenders überprüfen zu können, wird auf der Lernplattform ein Wissenstest angeboten. Zum Abschluss des WBT ist dieser zwar nicht zwingend notwendig, denn die Zielgruppe ist eine solche Überprüfung ihrer Kenntnisse nicht gewohnt. Dennoch wird dies angeboten, damit der Anwender zumindest die Möglichkeit hat, zu überprüfen, ob er das Gelernte auch richtig verstanden hat.

Zunächst muss ein auf den Lerninhalt und die Zielgruppe zugeschnittener Aufgabentyp gewählt werden (Wendt 2003, S. 145 ff). Ja/Nein-Aufgaben eignen sich eher nicht, da die Wahrscheinlichkeit, dass die richtige Antwort zufällig erraten wird, zu hoch ist. Reihenfolgeaufgaben und Zuordnungsaufgaben sind für diesen Lerninhalt und die Lernziele ungeeignet. Freie Texteingabe-Aufgaben wären für die Lernerfolgskontrolle zwar geeignet, weil sich der Lernende dabei wirklich in seine individuelle Situation hineinendenken kann. Jedoch ist die Auswertung dieses Aufgabentyps durch das Programm nicht möglich.

Es bleibt der Aufgabentyp der Auswahlaufgaben. Zu einer Fragestellung wird dem Anwender eine unterschiedlich große Anzahl an Antwortmöglichkeiten angeboten, von denen er eine oder mehrere auswählen muss. Dieser Aufgabentyp kann dem Anwender ebenso einen Transfer des Gelernten auf andere Situationen abverlangen (Wendt 2003, S. 147).

3.9.2 Feedback

Auswahlaufgaben bieten durch ihre geschlossene Form die Möglichkeit, die Rückmeldung mehrstufig anzubieten. Dabei können diese so geplant werden, dass sie auf die vom Autor als relativ häufig auftretend vermuteten Verständnisschwierigkeiten des Anwenders reagieren. Die hier dargestellten Rückmeldevariationen beruhen auf Euler (Euler 1992, S. 143 f).

Das Feedback auf eine richtige Antwort sollte kurz und aufmunternd auf den Anwender wirken. Dabei sollte die Rückmeldung aber variabel gestaltet sein, sonst verkommt sie schnell zur Floskel und Leerformel, die sich abnutzt, wenn sie zu oft wiederholt wird. Daher wird für den Wissenstest eine Reihe von verschiedenen positiven Feedbacks erstellt.

Bei einer Falschantwort sollte in der Rückmeldung erläutert werden, warum die Antwort als falsch gewertet wird. Außerdem wird auch hier ein mehrstufiges Feedback gegeben. Auf eine falsche Antwort beim ersten Versuch folgt der Hinweis, dass die Antwort falsch war und die Bitte, es noch einmal zu versuchen.

Ist die Antwort auch beim zweiten Versuch noch nicht richtig, erfolgt die Aufforderung, sich noch einmal das Video anzusehen oder sich das entsprechende Kapitel in Textform noch einmal durchzulesen.

Bei einer teilkorrekten Antwort wird angezeigt, welche Antwort richtig gegeben wurde und welche nicht. Auch hier hat der Anwender zunächst die Möglichkeit, einen erneuten Versuch zu unternehmen. Ist die Antwort beim zweiten Mal noch immer nicht ganz richtig, wird er auch hier wieder aufgefordert, sich das Video noch einmal anzusehen oder sich den Text durchzulesen.

3.10 Didaktisches Modell

Didaktische Modelle wurden meist für den Unterricht zwischen Lehrern und Schülern in der Schule entwickelt (Jank/Mayer 1994, S. 38). Auf die Lernplattform wird nun ein Modell übertragen. Dabei muss zunächst herausgefunden werden, welches Modell sich überhaupt auf E-Learning übertragen lässt. Eine Reihe von Modellen fallen gleich im Vorfeld heraus, da sie Dinge voraussetzen, die nur in einem Lehrer/Schüler-Verhältnis möglich sind, nicht aber beim computergestützten Lernen.

Eines der ältesten didaktischen Modelle stammt von Wolfgang Klafki. Er formulierte in den 1960er Jahren seine bildungstheoretische Didaktik: »Bildung und Erziehung haben die Aufgabe, dem unmündigen Menschen zur Mündigkeit zu verhelfen.« (Jank/Mayer 1994, S. 137). Damit ist auch gleich klar, warum Klafkis didaktisches Modell nicht für E-Learning-Anwendungen geeignet ist. Klafki sieht es als Aufgabe des Unterrichts, den Schülern Allgemeinbildung zu vermitteln und nicht unbedingt fachspezifisches Wissen. Das ist in einem E-Learning-Programm nicht überprüfbar und die Inhalte sind viel zu unbestimmt.

Der lehr- bzw. lernzieltheoretische Unterricht nach Heimann, Schulz und Otto dient der reinen Beschreibung von Unterricht. Die Autoren hatten keine Absicht, ein Modell zu entwickeln, welches als Vorlage für Unterrichtsvorbereitung dienen kann. Schulz führt zwar Planungsprinzipien an, füllt diese aber nicht mit Inhalt. So sagen sie dem Lehrer zwar, worüber er entscheiden muss, aber nicht, welche Kriterien dabei gelten. So findet in diesem didaktischen Modell kein Übergang von Unterrichtsanalyse zur Unterrichtsplanung statt (Jank/Mayer 1994, S. 204). So bleibt dieses Modell für die Erstellung eines E-Learning-Programms unbrauchbar.

Die dialektische Didaktik nach Lothar Klingenberg geht davon aus, dass Lehren und Lernen unterscheidbare aber aufeinander bezogene Tätigkeiten sind: »Lehren und Lernen stehen im Unterrichtsprozess [sic!] in einem dialektischen Verhältnis zueinander.« (Jank/Mayer 1994, S. 242). Das bedeutet, dass laut Klingenberg der Lernprozess wechselseitig und dynamisch ist. Es entstehen Konflikte zwischen Lehrer und Schülern, so dass beide Seiten Argumente entwickeln und auf den anderen reagieren müssen. Dieses Modell ist somit nicht auf E-Learning übertragbar, denn das Programm kann keine Argumente entwickeln und kann auch nicht vom Anwender lernen.

Der lernzielorientierte Unterricht nach Christine Möller teilt sich in die drei Teilprozesse Lernplanung, Lernorganisation und Lernkontrolle auf. Während der ersten Phase werden die Lernziele definiert, dann die dazu passenden Methoden gewählt. Zuletzt wird der Lernerfolg des Lernenden überprüft (Jank/Mayer 1994, S. 299). Dieses Konzept funktioniert im Unterricht mit Lehrer/Schüler-Verhältnis nur, wenn der Unterricht komplett planbar wäre, was er aber eigentlich nicht ist. Schüler und Lehrer können nicht mehr spontan reagieren, es werden nur Aktivitäten des Schülers zugelassen, die auf das angestrebte Endverhalten passen (Jank/Mayer 1994, S. 309). Was für einen Unterricht eher störend wirkt, bringt für eine E-Learning-Situation aber Vorteile. Das Programm kann nicht spontan reagieren und wird das daher auch nicht tun. Es akzeptiert von vornherein nur das Verhalten des Anwenders, das zur Erreichung der Lernziele notwendig ist. Daher ist das Modell des lernzielorientierten Unterrichts nach Möller für die Lernplattform geeignet.

Der erfahrungsbezogene Unterricht nach Ingo Scheller bezieht, wie der Name schon sagt, Erfahrungen der Schüler in den Unterricht mit ein. Schellers Modell nimmt Erlebnisse der Schüler, die diese schon vor dem Unterricht gesammelt haben, als Ausgangspunkt, die dann im Unterricht verarbeitet werden. Daraus entstehen dann Erfahrungen, die zu Haltungen gefestigt werden. Schließlich wird so das Handeln der Schüler in sozialen Situationen gesteuert, wodurch wieder neue Erlebnisse gesammelt werden. Der Kreislauf wiederholt sich so immer wieder (Jank/Mayer 1994, S. 314). Für E-Learning eignet sich dieses Modell nicht. Das Programm kann nicht so flexibel sein und für jeden Lernenden mit seinen individuellen Erlebnissen eine angepasste Lernumgebung bieten.

Der handlungsorientierte Unterricht, der von mehreren Theoretikern definiert wurde, weist einige zentrale Merkmale auf. Zunächst ist handlungsorientierter Unterricht ganzheitlich. Das heißt, dass beim Schüler alle Sinne gleichermaßen angesprochen werden sollen und dementsprechend müssen auch die Unterrichtsmethoden ganzheitlich gewählt werden, um alle Sinne anzusprechen. Der Lerninhalt wird nicht aufgrund einer wissenschaftlichen Fachsystematik ausgewählt, sondern aufgrund von Problemen und Fragestellungen. Handlungsorientierter Unterricht ist schüleraktiv, das heißt, die Schüler erarbeiten sich den Lernstoff selbst und der Lehrer gibt wenig vor. Das Ergebnis von handlungsorientiertem Unterricht sind Handlungsprodukte, die beispielsweise aus einer Projektarbeit entstehen. Ausgangspunkt für handlungsorientierten Unterricht sind die Interessen der Schüler, die entwickelt und vertieft werden. Die Schüler werden an der Planung, Durchführung und Auswertung des Unterrichts beteiligt. Für einen handlungsorientierten Unterricht öffnet sich die Schule. Das heißt, Schüler können das Schulgelände verlassen und auch außerhalb lernen, ebenso können Personen von außen in die Schule kommen und dort Rede und Antwort stehen. Zuletzt wird versucht, Kopf- und Handarbeit in ein ausgewogenes Verhältnis zueinander zu bringen (Jank/Mayer 1994, S. 354 ff). Handlungsorientierter Unterricht eignet sich bedingt auch für E-Learning-Maßnahmen. Durch Interaktionen mit dem Programm werden sowohl Kopf als auch Hand des Lernenden angesprochen, am bes-

ten in Simulationen oder Planspielen. Jedoch ist es kaum möglich, auf individuelle Interessen der Lernenden zu reagieren. Für die Lernplattform ist dieses Modell daher eher ungeeignet.

3.11 Realisierung des Web-Based-Trainings

Auf diesen grundlegenden Überlegungen kann nun ein Konzept für die Lernplattform aufgebaut werden. Ausgehend von der Zielgruppe muss zunächst überlegt werden, in welcher Situation ein Besuch der Lernplattform erfolgen wird.

3.11.1 Überlegungen zum Verbreitungsmedium

Kinematografen wollen eher weniger Zeit darin investieren, etwas Neues rein theoretisch zu lernen. Daher ist es wichtig, die wichtigsten Informationen in möglichst kurzer Zeit zu vermitteln. Da es sich beim Lerninhalt um Aspekte der Filmproduktion handelt, liegt es nahe, diesen auch über das Medium Film zu vermitteln. Zudem geht es um stereoskopische 3D-Filmproduktion, daher sollte das Verbreitungsmedium ebenso in der Lage sein, dies darzustellen.

Zu Beginn muss also überlegt werden, welches Medium diese Voraussetzung erfüllt – ein 3D-Fernseher wäre die naheliegendste Lösung. Dazu würde man dann noch einen 3D-fähigen Blu-Ray-Player benötigen und eine der 3D-Technologie angemessenen Brille. Nachteile sind dabei aber, dass diese Technologien vergleichsweise teuer sind und somit nicht vorausgesetzt werden kann, dass die Zielgruppe einen einfachen Zugang zu dieser hat. Außerdem müsste dann jedem Anwender eine Blu-Ray-Disk gepresst und zugeschickt werden. Daher eignet sich diese Verbreitungsart nicht besonders gut.

Eine schnellere und kostengünstigere Art der Verbreitung ist das Internet. Dabei muss zwar auf Polfilter- und Shattersystem verzichtet werden, doch die Anaglyphen-Technik kann hier angewendet werden. Dann kann das Video mit einem roten Bild und einem Bild in Cyan wie ein normales 2D-Video auf einer Webseite eingebunden werden. Der Betrachter benötigt nur eine Anaglyphenbrille, welche es aber sehr kostengünstig zu erwerben gibt. Außerdem kann die Betrachtungssituation somit ziemlich gut vorhergesagt werden, was für die Stereokalkulation des Videos von Vorteil ist.

Eine Webseite als Verbreitungsart bietet außerdem die Möglichkeit, noch anderen Content zusätzlich zum reinen Video einzubinden. So können noch zusätzliche Informationen zu den Themen bereitgestellt und auch ein Wissenstest programmiert werden.

3.11.2 Die Videoproduktion

Als größte Herausforderung bei der Produktion des Videos stellt sich das Konzept heraus. Das komplexe Thema der Stereoskopie soll in so kurzer Zeit wie möglich, aber

gleichzeitig auch für jemanden, der noch nie etwas damit zu tun hatte, verständlich dargestellt werden.

Es stellt sich schnell heraus, dass dem Betrachter zunächst klargemacht werden muss, warum es wichtig ist, sich mit dem Thema zu beschäftigen. Das geht am besten, indem man ihm zeigt, was passiert, wenn man das nicht tut. Die Grundidee des Videos besteht also darin, Fehler zu zeigen, die bei einem Stereo-3D-Film auftreten, wenn man sich zuvor nicht mit der Stereoskopie befasst. Diese Fehler sollten allerdings darauf beschränkt sein, was ein Kinematograf bei der Produktion selbst beeinflussen kann, denn zusätzliche Informationen würden ihn schnell langweilen.

Allein die Auswirkungen von Fehlern in der Bildgestaltung zu zeigen, reicht aber nicht aus. Der Anwender soll auch verstehen, warum dieser Effekt auftritt und was er tun kann, um das zu verhindern. Diese Informationen werden am besten über einen Sprechertext vermittelt.

Der Sprechertext muss es schaffen, in wenigen, einfachen Worten Ursache, Wirkung und Vermeidung der Fehler auszudrücken. Er soll den Anwender persönlich ansprechen, aber den Inhalt dennoch allgemein halten, damit der Lernende versteht, wie er das Gelernte auf seine persönliche Situation übertragen kann.

Neben den Fehlern wird aber auch über die allgemeine Situation an einem Stereo-3D-Filmset gesprochen, über den Stereografen, der neu dazu kommt, und über das Alignment der Kameras, damit der Kinematograf weiß, dass dies zu Zeitverzögerungen führen kann.

Um einen persönlichen Bezug zum Sprecher zu bekommen, tritt dieser als Moderator auf. Durch die Studiosituation wird die Atmosphäre eines Wissensmagazins geschaffen, wie der Anwender sie schon aus dem Fernsehen kennt. So wird eine Akzeptanz dafür geschaffen, dass der Moderator dem Anwender neues Wissen vermitteln will.

Das Studio wird so aufgebaut, dass alle Effekte, die beschrieben werden, kontrolliert hervorgerufen werden können. Im Studio hat man die volle Kontrolle über alle Einflüsse und kann sicher sein, dass immer nur der Effekt auftritt, den man gerade beschreibt. Auch aus organisatorischer Sicht hat die Studiosituation Vorteile. Die Umbauzeiten sind kurz und Dreharbeiten können auch über mehrere Tage dauern, ohne dass der Szenenaufbau verändert wird. Der Aufbau des Studios ist in Abbildung 2 zu sehen.



Abbildung 2: Studioaufbau des Lernvideos (eigene Darstellung)

Für die Aufnahme wurde ein Spiegel-Rig der Firma SwissRIG verwendet. Darin sind zwei Canon EOS C300 mit Super 35mm CMOS-Sensor und FullHD-Auflösung montiert. Die beiden Kameras sind über Genlock synchronisiert. Es werden dazu zwei 24-70mm F2.8 EX DG HSM Objektive von Sigma verwendet. Das Alignment der beiden Kameras erfolgt mit Einsteinium und der Software Uranium, beides an der Hochschule Furtwangen entwickelt.

Die Aufzeichnung erfolgt extern über eine Capture-Karte von Blackmagic Design über HD-SDI. Die dafür verwendete Software Media Express stammt ebenfalls von Blackmagic Design. Eine Datenreduktion wird durch die MJPEG-komprimierte Aufzeichnung mit einer Datenrate von ca. 58000 kb/s erreicht. Die Aufzeichnung erfolgt in einem AVI-Container im Format yuvj422p 1920×1080@25p.

Die Stereokalkulation erfolgt mittels der von Benny Hauser entwickelten Software S3D-Calc in der Version 0.8.5 für den Taschenrechner TI-84 vom Hersteller Texas Instruments. Als Betrachtungssituation wird eine Breite des Videos von 25cm angenommen mit einem idealen PC-Betrachtungsabstand von 50cm.

3.11.3 Post-Produktion

Die Auswahl der besten Takes wird bereits während des Drehs getroffen. So können diese für die weitere Verarbeitung schnell ausgewählt werden. Im Schnittprogramm Adobe Premiere Pro CS6 werden der Grob- und der Feinschnitt vorgenommen. Außerdem wird die absolute Tiefenpositionierung der Ebenen vorgenommen. Die Streams für links und rechts werden dabei immer parallel mitgeführt, sodass keine Vertauschung der beiden Perspektiven passieren kann.

Die AVI-Dateien der ausgewählten Takes werden mittels NukeX 7 und dem Plug-In Ocula 3.0 von TheFoundry feinaligned. Der Farbunterschied, der durch den semitransparenten Spiegel im Rig entsteht, wird durch den Node O_ColourMatcher korrigiert. Außerdem wird das rechte Bild gespiegelt. Aus NukeX werden MOV-Dateien exportiert. In Adobe Premiere werden die AVI-Dateien durch die „feinaligneten“ MOV-Dateien ersetzt. Die beiden Streams für links und rechts werden dann einzeln als TIFF-Sequenzen exportiert. Die beiden TIFF-Sequenzen für links und rechts werden anschließend wieder in Nuke geladen, wo sie als ein Anaglyphenvideo exportiert werden können.

Zusätzlich zum real gefilmten Teil gibt es im Video noch eine Sequenz, die das natürliche Stereosehen erklärt. Die Zeichnungen dafür werden in Photoshop CS6 erstellt und in Premiere Pro CS6 animiert. Diese Sequenz ist nur in 2D zu sehen und wird daher in das fertige Video zwischengeschnitten.

3.11.4 Die Webseite

Das Web-Based-Training wird mit Hilfe von HTML5, CSS und Javascript produziert. Der Aufbau der Webseite enthält keine Unterseiten und Links im klassischen Sinne. Die gesamte Seite wird nur einmal geladen. Inhalte werden dann ein- und ausgeblendet. So kann die Ladezeit beim Öffnen der Seite zwar etwas länger dauern, doch dann kann man alle Inhalte ohne Ladezeit erreichen.

Wird die Webadresse thesis.nadine-leonhardt.de aufgerufen, gelangt man zunächst auf die Startseite. Dabei verschwindet nach dem Laden der Seite eine rote Farbfläche am unteren Bildschirmrand, sodass gleich zu Beginn verdeutlicht wird, dass der Anwender mit diesem Slider interagieren kann. Der Willkommenstext auf der Startseite erklärt Sinn und Zweck des WBT, wobei dort zunächst die Motivation des Lernenden geweckt wird, indem einige rhetorische Fragen gestellt werden. Auch die Bedienung der beiden Slider unten und rechts wird erläutert, auch wenn diese bereits visuell deutlich gemacht wird. Ein Klick auf den Button Starten lässt das WBT in Vollbildmodus wechseln, den der Anwender in den meisten Browsern noch erlauben muss.

Nach diesem Klick wechselt der Anwender weiter in den Abschnitt Video. Dort wird mit einem Text kurz noch einmal verdeutlicht, warum der Anwender das WBT absolvieren sollte, bevor er einen Stereo-3D-Film produziert. Es wird außerdem erklärt, dass gleich ein Film folgt, für dessen Betrachtung eine Anaglyphenbrille notwendig ist. Da es gerade beim Anaglyphenverfahren auf die Farblichkeit sowohl der Brille als auch des Betrachtungsmediums ankommt, wird der Anwender aufgefordert, zu testen, ob die Farbeeinstellungen seines Monitors zu denen seiner Anaglyphenbrille passen. Zu diesem Zweck wurde ein Testfeld eingefügt. Der Anwender wird aufgefordert, seine Bildschirmhelligkeitseinstellungen so lange zu verändern, bis er mit jeweils einem Auge nur noch das entsprechende Feld sieht und das andere verschwindet.

Die Stereokalkulation der Videos wurde auf Grundlage einer vorhergesagten Breite des Videos und eines angenommenen Betrachtungsabstands berechnet. Nun muss aber gewährleistet sein, dass die Tiefenwirkung für jeden Betrachter auf jedem Monitor mit jeder Größe identisch ist, damit die beschriebenen Effekte auch von jedem Anwender gleich wahrgenommen werden. Das heißt, es muss für jeden Anwender ein individueller Betrachtungsabstand errechnet und ihm mitgeteilt werden.

Eine solche Berechnung wäre einfach, wenn die Breite des Monitors abgefragt werden würde. Jedoch wäre es für den Anwender ein zu hoher Aufwand, diese zu messen. Stattdessen kann aber die Diagonale des Monitors in Zoll abgefragt werden, denn es ist wahrscheinlich, dass der Anwender diese Angabe kennt. Zusätzlich muss noch das Seitenverhältnis des Bildschirms angegeben werden. Hierbei kann aus drei Seitenverhältnissen gewählt werden, die am gängigsten sind: 16:9, 4:3 und 16:10. Werden diese beiden Werte vom Anwender angegeben, wird der für ihn optimale Betrachtungsabstand für die Videos ausgegeben (siehe Abbildung 4).

Bitte geben Sie hier die Diagonale Ihres Bildschirms in Zoll an. Wählen Sie anschließend das Seitenverhältnis Ihres Bildschirms aus.

Zoll ▼

Ihr errechneter Betrachtungsabstand beträgt: 30.9 cm

Abbildung 3: Eingabefelder für Diagonale und Seitenverhältnis mit dem errechneten Betrachtungsabstand (eigene Darstellung)

Dahinter steht eine mehrteilige Javascript-Funktion. Zunächst wird die numerische Eingabe der Bildschirmdiagonalen in mit dem Faktor 2,54 multipliziert, um daraus eine Angabe in Zentimeter zu erhalten. Der Winkel Alpha wird mittels trigonometrischer Winkelfunktion aus dem angegebenen Seitenverhältnis errechnet. Mit Hilfe des Winkels Alpha und der Bildschirmdiagonalen wird mit einer weiteren trigonometrischen Winkelfunktion die Breite des Bildschirms berechnet. Die Breite des Bildschirms wird dann mit dem Faktor 0.6 multipliziert, denn das Video nimmt im Vollbildmodus immer 60% der Bildschirmbreite ein. Der entstandene Wert wird mit 2 multipliziert und das Ergebnis als optimaler Betrachtungsabstand mit nur einer Nachkommastelle ausgegeben.

```

104 <script type="text/javascript">
105 function berechnen(faktor) {var breite;
106 breite = document.getElementById("Breite").value;
107
108 var Zentimeter = breite*2.54;
109
110 var Alpha = Math.asin(faktor)*(180/Math.PI);
111
112 Ergebnis = 90-Alpha;
113 Ergebnis = Math.sin(Ergebnis/180*Math.PI);
114 Ergebnis = Ergebnis*Zentimeter*2*0.6;
115 Ergebnis = Math.round(Ergebnis * 10)/10;
116
117 document.getElementById("Ausgabe").innerHTML = unescape("Ihr errechneter
-Betrachtungsabstand betr\u00e4gt: ") + Ergebnis + ' cm';}
118 </script>

```

Abbildung 4: Javascript für die Berechnung des Betrachtungsabstandes (eigene Darstellung)

Klickt der Anwender »Zum Film«, wird der vorherige Abschnitt ausgeblendet und es erscheint das erste Video. Jedes Video wird über ein entsprechendes HTML5 Tag eingebunden. Dieses Tag erlaubt es zwar, durch das Attribut controls auch Steuerelemente eines Videoplayers einzubauen, doch dies eignet sich für dieses WBT nicht. Die Videos werden mit einer Anaglyphenbrille in Stereo-3D angesehen. Will der Anwender Steuerelemente des Browser-eigenen Videoplayers nutzen, so erscheinen diese halbtransparent auf dem abgespielten Video. Dies kann zu großen Irritationen führen, da die Leiste mit den Steuerelementen auf Bildschirmerebene über einem Stereo-3D-Video erscheint. Daher wird dieses Attribut nicht genutzt.

Stattdessen wurde ein eigener Player programmiert, der außerdem Funktionen besitzt, die der Standardplayer nicht enthält. Die Steuerelemente befinden sich alle unterhalb des Videos, sodass hier kein Konflikt entsteht. Mit dem eigenen Player stehen die Funktionen Play, Pause, Vor und Zurück, sowie eine verschiebbare Zeitleiste zur Verfügung. Play und Pause erfüllen dabei die gleichen Funktionen wie im Standardplayer. Jedoch bestand hierbei die Schwierigkeit darin, dass die gesamte Webseite keine Links besitzt und daher auch die Steuerelemente erst einen Wert übergeben bekommen müssen, der ihnen sagt, welches der Videos gerade eingeblendet wird und welches damit gestartet oder pausiert werden soll. Dies ist mit mehreren Funktionen in Javascript gelöst.

Eine Besonderheit bieten die Buttons Vor und Zurück. Denn diese sollen immer zum vorherigen oder zum nächsten Video springen, wobei jedes Video einen eigenen Player besitzt. Diese beiden Buttons müssen daher nicht nur auf vor- oder zurückliegende Videos verweisen, sondern können auch auf ganz andere Abschnitte führen. Klickt man beim ersten Video auf den Zurück-Button, so gelangt man zurück auf die Video-Startseite. Klickt man beim letzten Video auf den Vor-Button, so führt dieser direkt weiter auf die Test-Seite. Die Funktionen für diese beiden Buttons müssen immer

wissen, welches Video gerade abgespielt wird und welches davor oder danach kommen soll. Dies wird mit einer if-else-Anweisung in Javascript ermöglicht.

Springt ein Benutzer entweder durch Klicken auf die Vor-/Zurück-Buttons oder durch eine Auswahl im Menü zu einem anderen Video, so wird das verlassene pausiert und das ausgewählte startet sofort. In Javascript wurde festgelegt, dass alle Videos pausieren, sobald sie nicht mehr angezeigt werden. Kommt der Benutzer noch einmal zum gleichen Video zurück, so wird dies an jener Stelle fortgesetzt, an der es zuvor verlassen wurde.

Das Video auf der nächsten Seite ist in Kapitel unterteilt, die nach den jeweiligen Themen benannt sind und die auch einzeln anwählbar sind.

Der Wissenstest auf der nächsten Seite wurde bereits im Abschnitt Aufgaben angesprochen. Es ist dabei nicht zwingend notwendig, den Test erfolgreich abzuschließen, denn es müssen ja auch im Video nicht alle Themen für den Anwender interessant sein.

In Textform werden dem Lernenden auf einer weiteren Seite noch mehr Informationen über die einzelnen Themen aus dem Video angeboten. Auch diese sind in sinnvolle Kapitel unterteilt.

Unter einem weiteren Menüpunkt werden Extras angeboten. Dort findet sich eine Galerie mit Fotos des Videodrehs aber auch von anderen 3D-Filmen. Außerdem befindet sich dort auch ein Link zu dieser Thesis.

4. Fazit

Stereo-3D-Filme sind beliebt und das nicht nur bei den Zuschauern. Auch die Produktionsfirmen machen derzeit, wenn das Budget es zulässt, gerne aus jedem Stoff einen 3D-Film. Die Produktion eines Stereo-3D-Films ist zwar teurer, aber ebenso sind das auch die Eintrittspreise an den Kinokassen. Das Publikum ist bereit, diesen Preis zu zahlen.

Parallel werden die technischen Bedingungen bei der Produktion immer weiter verbessert, zum Teil effektiv serialisiert und damit auch im Preis gesenkt. Bei großen Produktionen macht es immer weniger Unterschied, ob man in Stereo-3D oder 2D dreht.

Sobald die Technologie und die Kosten kein Hinderungsgrund für den Erfolg von Stereo-3D-Filmen mehr sind, rücken die künstlerische Gestaltung eines Films in den Vordergrund. Immer öfter fragt sich der Zuschauer: »Muss ich diesen Film in 3D sehen? Würde nicht auch eine normale 2D-Fassung genügen?«

Der Zuschauer sollte also immer wieder davon überzeugt werden, dass es sinnvoll ist, dass dieser Film in Stereo-3D zu produziert wurde. Dafür muss der Tiefeneffekt

sinnvoll eingesetzt werden, und zwar um der Geschichte und nicht des Effekts willen. Maßgeblich verantwortlich dafür ist derjenige, der das Bild eines Films gestaltet: der Kinematograf.

Gerade weil jetzt immer mehr Aufmerksamkeit der Zuschauer auf die Bildgestaltung eines Stereo-3D-Films gelenkt wird, sollte der S3D-Kinematograf wahrnehmungsgerecht aber kreativ mit dem Tiefeneffekt umzugehen wissen. Dabei ändern sich einige Rahmenbedingungen der Bildgestaltung im Vergleich zum klassischen 2D-Film. Beachtet er dies nicht, kann das sogar zu Problemen in der Tiefenwahrnehmung der Zuschauer führen und es gibt nichts Irritierenderes für das Publikum, als dass es sich durch einen Fehler im Tiefeneffekt aus der Geschichte herausgeworfen fühlt.

Deshalb ist es wichtig, dass sich der Kinematograf bewusst ist, dass er seine Art der Bildgestaltung dem neuen Tiefeneffekt anpassen muss. Das ist aber keine leichte Aufgabe. Um erst einmal zum Head of Department aufzusteigen, ist jahrelange Berufserfahrung nötig. Zwar muss der Kinematograf in dieser langen Zeit auch viele neue Dinge lernen, doch das meistens durch Learning by doing. Bei einem Stereo-3D-Film wäre das aber fatal, denn je nachdem, was für einen Fehler er in der Bildgestaltung macht, kann der auch nicht mehr in der Postproduktion behoben werden. Er muss es also schon in der Vorbereitung lernen.

Dazu kommt noch der Stereograf, der bei einem Stereo-3D-Film neu dazu kommt und sich sogar in die Bildgestaltung des Kinematografen einmischt. Das ist für diesen absolut ungewohnt und den meisten fällt es schwer, das zu akzeptieren. Deshalb ist es wichtig, dass der Kinematograf schon beim ersten Treffen mit dem Stereografen weiß, warum dieser das tut und warum das auch wichtig ist.

Es ist aber nicht einfach, einem Kinematografen etwas theoretisch beizubringen, was er dann erst später in die Praxis umsetzen kann. Deshalb braucht er Anschauungsmaterial, das er sich gut einprägen kann. Auch ist es wichtig, dass er sich nicht zu lange damit beschäftigen muss, denn er ist es einfach nicht gewohnt, etwas nur theoretisch zu lernen und verliert schnell die Lust daran. Er muss merken, dass derjenige, der ihm das jetzt beibringen will, auch etwas von seinem Berufsalltag versteht und genau weiß, welche Situationen im wirklichen Leben auch auftreten.

Diese Ziele werden mit der Lernplattform erreicht. Das Video schafft es, in weniger als zehn Minuten dem Kinematografen klarzumachen, warum es wichtig ist, sich über die neuen Rahmenbedingungen einer Stereo-3D-Realfilmproduktion im Vorfeld zu informieren. Kurz und vor allem verständlich auch für jeden, der noch nie etwas mit Stereo-3D zu tun hatte, werden die wichtigsten Effekte aufgezeigt, die der Kinematograf mit der Bildgestaltung auslösen oder verhindern kann. Sollte er sich noch weitergehend für das Thema interessieren, zum Beispiel auch, um in einer Diskussion mit dem Stereografen stichhaltige Argumente anbringen zu können, kann er sich in Textform noch intensiver mit den einzelnen Themen befassen. Um sicherzugehen, dass er wirklich die Zusammenhänge verstanden hat, kann er sein Wissen auch in einem Test überprüfen.

N. LEONHARDT

Die entstandene Lernplattform eignet sich daher bestens, bei einem Kinematografen rechtzeitig ein Verständnis für die geänderten Rahmenbedingungen einer Stereo-3D-Realfilmproduktion zu schaffen.

Quellenverzeichnis

Literaturquellen

- BVK - Berufsverband Kinematografie e.V. (Hg.) (2013): Die Berufsbilder. Grafeling: bvkmedien GmbH
- Euler, Dieter (Hg.) (1992): Didaktik des computerunterstützten Lernens. Praktische Gestaltung und theoretische Grundlagen. 1. Aufl. Nürnberg: Bildung-und-Wissen-Verl. (Reihe „Multimediales Lernen in der Berufsbildung“, 3).
- Jank, Werner; Meyer, Hilbert (Hg.) (1994): Didaktische Modelle. 3. Aufl. Frankfurt am Main: Cornelsen Scriptor.
- Keller, J. M. (1983): Motivational design of instruction, In C. M. Reigeluth (Ed.), Instructional-design theories and models: An overview of their current studies. Hillsdale, N.J: Erlbaum
- Kuhn, Gerhard (1999): Stereo-Fotografie und Raumbild-Projektion. Gilching: vfv.
- Mendiburu, Bernard (2009): 3D movie making. Stereoscopic digital cinema from script to screen. Amsterdam, Boston: Focal Press/Elsevier.
- Neuhaus, Stefan (Hg.) (2008): Literatur im Film. Beispiele einer Medienbeziehung. Würzburg: Königshausen & Neumann (Film, Medium, Diskurs, Bd. 22).
- Niegemann, H. M. (2006): Besonderheiten einer Didaktik des E-Learning, In Arnold, Rolf (Hg.): eLearning-Didaktik. Baltmannsweiler: Schneider-Verl. Hohengehren (Grundlagen der Berufs- und Erwachsenenbildung, 48).
- Schupbach, Evi (Hg.) (2003): Didaktischer Leitfaden für E-Learning. Didactic guidelines for e-learning. 1. Aufl. Bern: h.e.p. Verl. (Pädagogik).
- Sczepek, Jorg (2011): PhotoWissen - 1 Bildentstehung, Raumtiefe, Grose. Naturwissenschaften und Psychologie für Fotografen. Bildentstehung, Raumtiefe, Grose. Norderstedt: Books on Demand.
- Spottiswoode, Raymond (Hg.) (1982): The focal encyclopedia of film & television techniques. Reprinted. London [u.a.]: Focal Press.
- Stober, Rudolf (2003): Mediengeschichte. 1. Aufl. Wiesbaden: Westdt. Verl.

- Tauer, Holger (2010): Stereo 3D. Grundlagen, Technik und Bildgestaltung. 1. Aufl. 1 Band. Berlin: Schiele & Schon.
- Wendt, Matthias (2003): Praxisbuch CBT und WBT. Konzipieren, entwickeln, gestalten. Munchen, Wien: Hanser.

Journalistische Quellen

- Bolliger, Matthias: Stereo 3D. In Film & TV Kameramann, Ausgabe 1 & 2 (2011)
- Keuneke, Jurgen: Zwerge und Riesen. Produktionsbericht. In Film & TV Kameramann, Ausgabe 1 (2013), S. 68-73
- Schulz, Thomas: Eine neue Dimension. In Der Spiegel, Ausgabe 29 (2009), S. 78-80

Internetquellen

- 3ality Technica (o.J.): About Us, URL: <http://www.3alitytechnica.com/aboutus.php#3ality> (Zugriff am 09.02.2014)
- Box Office Mojo (o.J.): 2013 Worldwide Grosses, URL: <http://boxofficemojo.com/yearly/chart/?view2=worldwide&yr=2013&p=.htm> (Zugriff am 09.02.2014)
- Browser Statistik (2014): Die Browser-Marktanteile in Deutschland, URL: <http://www.browser-statistik.de/> (Zugriff am 23.02.2014)
- Dirks, Tim - AMC Network Filmsite (o.J.): Timeline of Greatest Film Milestones and Turning Points in Film History - The Year 1922, URL: <http://www.filmsite.org/1922-filmhistory.html> (Zugriff am 09.02.2014)
- FFA - Filmförderungsanstalt (2012): Der Kinobesucher 2011 - Strukturen und Entwicklungen auf Basis des GfK Panels, URL: http://www.ffa.de/downloads/publikationen/kinobesucher_2011.pdf (Zugriff am 06.02.2014)
- FFA – Filmförderungsanstalt (2013): Aktuelle FFA Pressemitteilungen - Detail, URL: http://www.ffa.de/index.php?page=presse_detail&ws=1055 (Zugriff am 06.02.2014)

- Nielsen (2012): Kino Box Office - Insights 2011, URL: http://www.weischer.net/fileadmin/user_upload/Werbeweischer/Dateien_ww/Nielsen_Kino_Box_Office_2011.pdf (Zugriff am 06.02.2014)
- StatCounter (2014): Top 7 Desktop OSs in Germany from Jan 2009 to Jan 2014, URL: <http://gs.statcounter.com/#os-DE-monthly-200901-201401> (Zugriff: 25.02.2014)
- zu Huningen, James (2011): Lexikon der Filmbegriffe, URL: <http://filmlexikon.uni-kiel.de/index.php?action=lexikon&tag=det&id=6184> (Zugriff am 09.02.2014)

