

Peter Wimmer

# **Aufnahme und Wiedergabe stereoskopischer Videos im Anwendungsbereich der Telekooperation**

Diplomarbeit  
zur Erlangung des akademischen Grades Diplom-Ingenieur  
in der Studienrichtung Mechatronik

Angefertigt am Institut für Telekooperation  
der Johannes Kepler Universität Linz

Betreut von Univ.-Prof. Mag. Dr. Gabriele Kotsis

Linz, Mai 2004

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

## Kurzfassung

Die Aufnahme stereoskopischer Videos erfolgt meist mittels zweier Kameras, stellt aber hohe Anforderungen an deren Synchronisierung und den Abgleich. Ebenso erfordert das Abspielen stereoskopischer Videos je nach Wiedergabegerät und Situation unterschiedliche Methoden. Diese Diplomarbeit behandelt zuerst die Grundlagen der Stereoskopie und zeigt anschließend unterschiedliche Wege zur Aufnahme und Wiedergabe auf. Darüber hinaus wird auch die Übertragung stereoskopischer Videos behandelt.

Die Aufgabenstellung umfasst Konzeption, Realisierung und Test eines Prototyps zur Aufnahme, Übertragung und Wiedergabe stereoskopischer Videos. Dazu wurde der *Stereoscopic Player* als Wiedergabeanwendung entwickelt. Er basiert auf DirectShow und unterstützt eine Vielzahl von Dateiformaten, Codecs, stereoskopischen Layouts und Wiedergabemethoden. Neben Dateien kann er Streams und DVDs abspielen und direkt auf Aufnahmegeräte zugreifen. Der *Stereoscopic Multiplexer*, ebenfalls Ergebnis dieser Arbeit, ist ein Treiber, der innerhalb anderer Programme zur Aufzeichnung von stereoskopischen Videos eingesetzt werden kann. Im entwickelten Prototyp auf Basis von zwei Firewire-Industriekameras wurde er zusammen mit dem Windows Media Encoder zur Übertragung stereoskopischer Videos eingesetzt.

## Abstract

Stereoscopic videos are mostly recorded using two separate cameras. However, high demands are made on the synchronisation and adjustment of these cameras. Likewise, playback of stereoscopic videos requires different methods depending on the playback device and other circumstances. This master thesis first presents fundamentals of stereoscopy and proposes different approaches to record and playback stereoscopic videos afterwards. In addition, methods for transmission of stereoscopic videos are presented.

Subject of this thesis is the design, implementation and evaluation of a prototype to record, transmit and playback stereoscopic videos. A video player application, called *Stereoscopic Player*, has been developed to play stereoscopic videos. It is based on DirectShow and is able to handle many file formats, codecs, stereoscopic layouts and viewing methods. Besides video files, it can play streams, DVDs and use capture devices. The *Stereoscopic Multiplexer*, another outcome of this thesis, is a driver that can be used within many other applications to capture stereoscopic content and is also part of this thesis. To implement the prototype based on two industrial Firewire cameras, it is used together with Microsoft's Windows Media Encoder to transmit stereoscopic videos.

---

# Inhalt

---

<b>1. Einleitung</b> .....	<b>1</b>
1.1 Motivation und Ziele.....	1
1.2 Struktur der Diplomarbeit.....	2
1.3 Danksagung.....	3
<b>2. Historisches</b> .....	<b>4</b>
2.1 Geschichte des 3D-Films.....	4
2.1.1 Anfänge.....	4
2.1.2 Der Boom in den 50ern .....	8
2.1.3 Entwicklungen ab 1954 .....	9
2.1.4 IMAX 3D .....	9
2.2 3D-Fernsehen .....	10
2.2.1 Fieldsequential-3D-Verfahren.....	10
2.2.2 Anaglyphen- und Pulfrich-Verfahren .....	11
<b>3. Räumliches Sehen</b> .....	<b>13</b>
3.1 Konvergenz und Akkomodation.....	13
3.2 Parallaxische Verschiebung .....	14
3.3 Monokulare Tiefenhinweise.....	15
3.3.1 Verdeckung .....	15
3.3.2 Größe .....	15
3.3.3 Perspektive .....	16
3.3.4 Dunst .....	16
3.3.5 Bewegungsparallaxe .....	17
3.4 Grenzen der räumlichen Wahrnehmung .....	18
3.4.1 Fusionsbereich.....	18
3.4.2 Bildzerfall.....	19
3.4.3 Divergenz.....	19
3.5 Aufnahme- und Wiedergabegeometrie.....	20
3.5.1 Kameraanordnung und Scheinfenster.....	20
3.5.2 Linsengleichung.....	23
3.5.3 Abbildung im Auge.....	23
3.5.4 Virtueller Abstand .....	25
3.5.5 Virtuelle Größe und Tiefe .....	27

3.6 Liliputismus und Gigantismus .....	28
3.7 Ortho- und Tautostereoskopie.....	29
3.8 Wahl der Stereobasis .....	29
3.9 Raumbildfehler .....	29
<b>4. Wiedergabe .....</b>	<b>29</b>
4.1 Einteilung der Wiedergabemethoden.....	29
4.2 Parallelblick .....	29
4.3 Kreuzblick.....	29
4.4 Anaglyphen.....	29
4.4.1 Echte Anaglyphen .....	29
4.4.2 Graue Anaglyphen .....	29
4.4.3 Farbige Anaglyphen.....	29
4.5 Polarisierungstechnik .....	29
4.6 KMQ.....	29
4.7 Shutterbrillen.....	29
4.7.1 Pageflipping .....	29
4.7.2 Interlaced-Darstellung .....	29
4.7.3 Lineblanking.....	29
4.7.4 Syncdoubling.....	29
4.8 Autostereoskopische Displays.....	29
4.8.1 Linsenrastermethode.....	29
4.8.2 Parallaxenbarriere.....	29
4.9 Head Mounted Displays.....	29
4.10 Stereoskope.....	29
<b>5. Formate .....</b>	<b>29</b>
5.1 Separate Streams.....	29
5.2 Layouts.....	29
5.2.1 Nebeneinander-Layout.....	29
5.2.2 Übereinander-Layout.....	29
5.2.3 Interlaced-Layout.....	29
5.2.4 Sonderlayouts.....	29
5.3 Codecs .....	29
5.4 3D-DVDs.....	29
5.4.1 Fieldsequential DVDs .....	29
5.4.2 TriDVD.....	29
5.4.3 Sensio 3D.....	29
5.4.4 Anaglyphen-DVDs .....	29
5.4.5 ColorCode 3-D .....	29

<b>6. Aufnahme</b> .....	<b>29</b>
6.1 Aufnahme mit zwei Kameras .....	29
6.1.1 Synchronisation auf Frames .....	29
6.1.2 Synchronisation auf Fields .....	29
6.1.3 Hardware-Multiplexing.....	29
6.2 Aufnahme mit einer Kamera .....	29
6.2.1 Passive Bildteiler.....	29
6.2.2 Aktive Strahlenteiler .....	29
6.3 2D-3D Konvertierung.....	29
<b>7. DirectShow</b> .....	<b>29</b>
7.1 Konzept.....	29
7.2 Filtertypen.....	29
7.3 Implementierung von Filtern.....	29
7.4 Alternative Multimedia-Architekuren .....	29
<b>8. Implementierung der Wiedergabe</b> .....	<b>29</b>
8.1 Einführung in den Stereoscopic Player .....	29
8.2 Architektur .....	29
8.2.1 Dateiwiedergabe .....	29
8.2.2 Streaming- und Live-Wiedergabe .....	29
8.2.3 DVD-Wiedergabe .....	29
8.3 Transformationsfilter .....	29
8.4 Renderingfilter .....	29
8.5 Benutzeroberfläche .....	29
8.5.1 Wiedergabeklassen.....	29
8.5.2 Komfortfunktionen.....	29
8.6 Test der Wiedergabefunktionen.....	29
<b>9. Implementierung der Aufnahme</b> .....	<b>29</b>
9.1 Kameraauswahl .....	29
9.2 Stereo Multiplexer .....	29
9.3 Programmintegration.....	29
9.4 Untersuchung der Synchronisation .....	29
9.4.1 Messanordnung .....	29
9.4.2 Ergebnisse .....	29
9.5 Parametersynchronisation.....	29
9.5.1 Synchronisationsmechanismen .....	29
9.5.2 Parameterkorrektur .....	29
9.6 Bildverarbeitungsfunktionen.....	29

<b>10. Streaming</b> .....	<b>29</b>
10.1 Übertragungsvarianten.....	29
10.2 Realisierung.....	29
<b>11. Zusammenfassung und Ausblick</b> .....	<b>29</b>
11.1 Zusammenfassung der Arbeit.....	29
11.2 Geplante Erweiterungen.....	29
<b>A. Bedienung des Stereoscopic Players</b> .....	<b>29</b>
A.1 Einrichtung.....	29
A.2 Wiedergabe.....	29
A.3 Bibliotheksverwaltung.....	29
A.4 Kommandozeilenparameter .....	29
A.4.1 Quellen.....	29
A.4.2 Eingangsformat.....	29
A.4.3 Wiedergabe.....	29
<b>B. Dokumentation der Dateiformate</b> .....	<b>29</b>
B.1 Berechnung der Hashwerte .....	29
B.2 Format von stereoskopischen Informationsdateien .....	29
B.3 Format von Wiedergabelisten .....	29
<b>C. Abbildungsverzeichnis</b> .....	<b>29</b>
<b>D. Tabellenverzeichnis</b> .....	<b>29</b>
<b>E. Literatur- und Linkverzeichnis</b> .....	<b>29</b>

---

# Einleitung

---

*Es werden Thema und Motivation der Diplomarbeit vorgestellt und ein Überblick über den Aufbau der Arbeit gegeben.*

## 1.1 Motivation und Ziele

Menschen besitzen zwei Augen, mit denen sie die Welt wahrnehmen. Zwei Augen schützen den Mensch vor der Erblindung, falls die Funktion eines Auges verloren geht. Diese Redundanz erhöht die Überlebenschance in der Natur wesentlich, insbesondere ermöglicht uns das zweite Auge aber das räumliche Sehen. Durch den Abstand der beiden Augen sehen wir die Welt aus zwei leicht unterschiedlichen Perspektiven. Das Gehirn ist in der Lage, daraus die Entfernung von Gegenständen abzuschätzen.

Die Stereoskopie [78] beschäftigt sich mit der Aufnahme und Wiedergabe räumlicher Bildeindrücke. Für die Aufnahme dreidimensionaler Objekte wird dabei die menschliche Sehweise imitiert. Eine Szene wird aus zwei unterschiedlichen Perspektiven aufgenommen, so wie auch unsere Augen die Szene aus zwei Perspektiven wahrnehmen.

Seit dem 18. Jahrhundert ist die Stereoskopie der Gegenstand umfangreicher Forschungen. Eine ausgezeichnete Zusammenfassung der wissenschaftlichen Arbeiten bietet [41]. Nach anfänglichen Erfolgen in der Pionierzeit der Fotografie hat die Stereoskopie, zumindest in der Wahrnehmung der breiten Massen, an Bedeutung verloren. Ausschlaggebend dafür sind Probleme und hoher Aufwand bei der Aufnahme und Wiedergabe stereoskopischer Bilder, zum Teil verbunden mit Unzulänglichkeiten in der Darstellungsqualität.

Nach dem Übergang von Schwarzweiß zu Farbe in Film und Fernsehen stellt die räumliche Wiedergabe den nächsten evolutionären Schritt dar, genauso wie im Audiobereich der Übergang vom Mono- über den Stereoton zum Raumklang vollzogen wurde. Die genannten Schwierigkeiten haben einen Durchbruch des Raumbildes im Heim- und Unterhaltungsbereich bisher verhindert. Es konnte sich bis dato nur in Nischenmärkten etablieren.

Das Raumbild bietet aber nicht nur ein intensiveres Erlebnis im Unterhaltungsbereich [34, 60]. In Wissenschaft und Technik gibt es unzählige Anwendungsgebiete, bei denen eine räumliche Visualisierung gegenüber einer flachen Abbildung handfeste Vorteile bietet. Beispielsweise ist es bei der Fernsteuerung von Robotern entscheidend, Entfernungen möglichst genau abschätzen zu können [26]. Nur mit einer stereoskopischen Darstellung ist es möglich, den Roboterarm zielgenau zu steuern. In der Biologie erlaubt die Stereoskopie das Erkennen von Details, die sich in einer zweidimensionalen Darstellung nicht erschließen. Manche Objekte besitzen eine so komplexe Struktur, dass sie sich nur in räumlicher Darstellung auf einen Blick erfassen lassen. In der Medizin können räumliche Visualisierungen dem Arzt die Planung von chirurgischen Eingriffen erleichtern [13]. Die Videoübertragung von Operationen erlaubt dem Arzt,

bei schwierigen Eingriffen den Rat von weit entfernt tätigen Experten einzuholen – auch in diesem Szenario erlaubt nur eine räumliche Darstellung, den Operationsbereich präzise wahrzunehmen. Ebenso eignet sie sich für eine realistische Darstellung in der Ausbildung.

Viele der erwähnten Anwendungen erfordern die Übertragung stereoskopischer Videos. Es ist das Hauptziel dieser Diplomarbeit, einen Prototypen zu entwerfen, der sich für unterschiedliche Einsatzzwecke adaptieren lässt. Neben der Übertragung bilden aber auch die Aufnahme und Wiedergabe stereoskopischer Videos Schwerpunkte dieser Arbeit. Der Autor interessiert sich besonders für die Wiedergabe von 3D-Videos und will durch die Implementierung computergestützter, qualitativ besserer Wiedergabemethoden die Akzeptanz von 3D-Videos erhöhen.

Die Computertechnik ist der Schlüssel zu einer hochwertigen und gleichzeitig kostengünstigen Verarbeitung stereoskopischer Inhalte. Diese Arbeit stellt daher die computergestützte Aufnahme und Wiedergabe in den Mittelpunkt, als Übertragungsmedium dient konsequenterweise das Internet. Die Aufgabenstellung der Diplomarbeit umfasst die Konzipierung und Realisierung von Anwendungen zur Aufnahme, Übertragung und Wiedergabe von 3D-Videos.

Beim Umgang mit stereoskopischen Inhalten gibt es zahlreiche Punkte, die für ansprechende Ergebnisse beachtet werden müssen. Die Missachtung eines einzigen Details ist ausreichend, um den Raumeindruck eines Stereobildpaares zu zerstören oder die Betrachtung so zu erschweren, dass nach kurzer Zeit Kopfweg oder Augenschmerzen die Folge sind. Dieser schriftliche Teil der Arbeit geht auf die Probleme ein und zeigt gangbare Lösungswege auf.

## 1.2 Struktur der Diplomarbeit

Der erste große Teil bis einschließlich Kapitel 7 beschäftigt sich mit Grundlagen, die entweder für den erfolgreichen Einsatz stereoskopischer Videos oder das Verständnis der folgenden Kapitel notwendig sind. Kapitel 2 bietet einen Rückblick auf die Geschichte der Stereoskopie, des 3D-Films und 3D-Fernsehens. Danach folgt in Kapitel 3 eine Einführung in die Grundlagen der räumlichen Wahrnehmung und die mathematisch-geometrischen Zusammenhänge der Stereoskopie. Kapitel 4, 5 und 6 behandeln die Aufnahme und Wiedergabe stereoskopischer Videos im Allgemeinen. Der erste Teil wird von einer Einführung in DirectShow, der grundlegenden Architektur der entwickelten Software, abgeschlossen.

Der zweite Teil erläutert den eingeschlagenen Lösungsweg und gibt einen Überblick über die Implementierung. Kapitel 8 beschäftigt sich mit der Realisierung der Wiedergabe. Dazu wurde der *Stereoscopic Player* entwickelt, der neben 3D-Videodateien auch 3D-DVDs abspielen kann und darüber hinaus als Client für die Live-Übertragung fungiert. Die Wiedergabe stellt eine besondere Herausforderung dar, da je nach Situation und vorhandener Hardware unterschiedliche Methoden verwendet werden müssen. Bei der Entwicklung des Stereoscopic Players war es daher oberstes Ziel, für jeden Display- oder Projektortyp die am besten geeignete Wiedergabemethode zur Verfügung zu stellen.

Das Herzstück der Aufnahme und Übertragung ist der *Stereoscopic Multiplexer*, ein Treiber zur Aufnahme von stereoskopischen Videos, der mit fast allen Videoprogrammen kompatibel ist. Bei dessen Implementierung wurde großer Wert darauf gelegt, dass er nicht nur mit jenem Kamerapaar zusammenarbeitet, das speziell für diese Diplomarbeit angeschafft wurde, son-

dern auch mit gänzlich anderen Kameratypen. Der Stereoscopic Multiplexer wird im Kapitel 9 im Detail vorgestellt, darüber hinaus werden die Resultate mit den angeschafften Kameras diskutiert.

Kapitel 10 geht schließlich auf die Übertragung ein, die auf Basis von Microsofts Windows Media realisiert wurde. Der Anhang beschreibt den Stereoscopic Player aus Sicht des Anwenders sowie die eingesetzten Dateiformate.

## 1.3 Danksagung

Ich möchte allen danken, die mich bei der Entstehung dieser Diplomarbeit unterstützt haben. An erster Stelle gilt der Dank meiner Betreuerin Univ.-Prof. Mag. Dr. Gabriele Kotsis.

Ich danke der Johannes Kepler Universität für den Zugang zum ICatcher sowie Christoph Anthes für seine Unterstützung im Umgang mit diesem System.

Mein Dank gebührt weiters den vielen Anwendern, die mit ihren zahlreichen Fehlerberichten und Verbesserungsvorschlägen wesentlich zum Erfolg des Stereoscopic Player beigetragen haben, insbesondere Antonio Scuri, Michal Hušák und Werner Bloos.

Ganz besonderer Dank gilt meinen Eltern und Geschwistern für ihre Unterstützung und Motivation während des gesamten Studiums sowie der kritischen Durchsicht dieser Arbeit.

---

# Historisches

---

*Es wird die Entwicklung der Stereoskopie und insbesondere die Entwicklung des 3D-Films von den Anfängen im 18. Jahrhundert bis in die Gegenwart aufgezeigt. Danach werden das moderne IMAX 3D-Verfahren erklärt sowie Entwicklungen auf dem Gebiet des 3D-Fernsehens behandelt.*

## 2.1 Geschichte des 3D-Films

### 2.1.1 Anfänge

Die Geschichte der Stereoskopie beginnt im Jahr 1833, als Sir Charles Wheatstone das Prinzip der räumlichen Wahrnehmung entdeckte und das Stereoskop erfand. Das Stereoskop besteht in seiner ursprünglichen Ausprägung nach Wheatstone aus zwei V-förmig angeordneten Spiegeln. Diese lenken den Blick des linken und rechten Auges auf zwei Abbildungen um, die das Motiv aus leicht unterschiedlichen Perspektiven zeigen. Die Spiegel sind erforderlich, um den Abstand zwischen Augen und Abbildung ausreichend groß zu halten, denn nur dann können die Augen das Motiv leicht fokussieren.

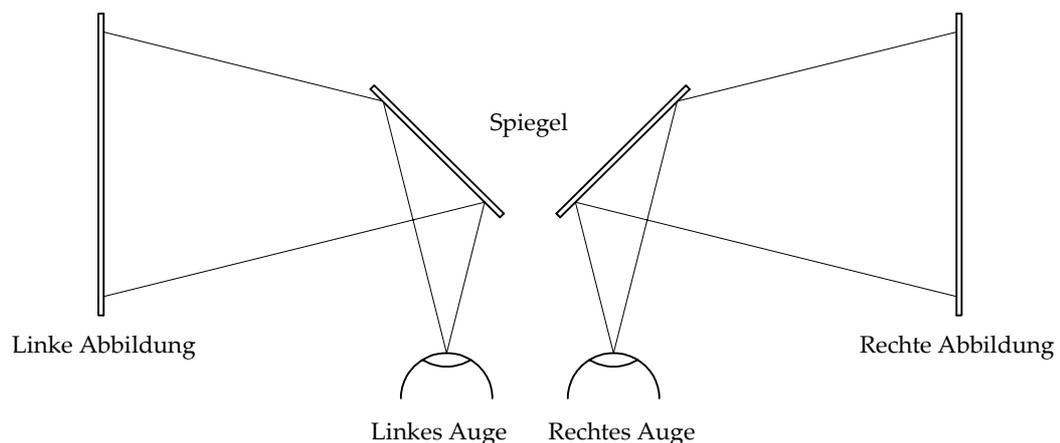


Abbildung 2.1: Wheatstones Stereoskop

Das Stereoskop wurde im Lauf der Zeit weiterentwickelt. Sir David Brewster ersetzte die Spiegel durch Linsen und ermöglichte damit eine kleine, geschlossene Bauform, die gleichzeitig den Augen die Fokussierung erleichterte. Diese Bauform stellt die klassische Ausführung des Stereoskops dar (Abbildung 2.2). Beide Bilder wurden nebeneinander auf einer einzigen

Karte montiert und durften wegen der Verwendung halbiertes Linsen auch breiter als der Augenabstand sein.

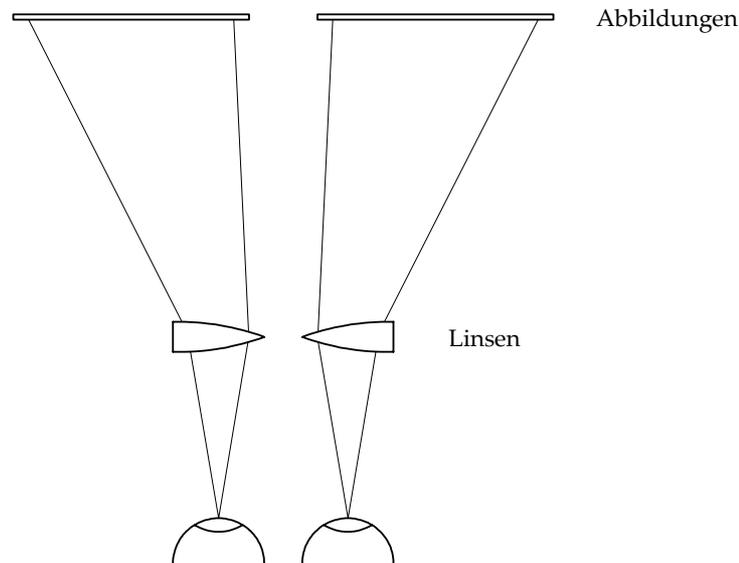


Abbildung 2.2: Brewsters Stereoskop

Die Popularität des Stereoskops erreichte Mitte des 19. Jahrhunderts ihren Höhepunkt, in den folgenden Jahrzehnten nahm das Interesse an stereoskopischen Bildern allmählich ab. Die Begeisterung über die Neuheit legte sich, gleichzeitig wuchs die Konkurrenz durch die konventionelle Fotografie aufgrund dessen einfacher fotomechanischer Reproduktion in Büchern und Zeitschriften. Es erwies sich als Nachteil, dass jeweils nur eine Person ein Stereobild betrachten konnte, während an einem normalen Foto mehrere Leute gleichzeitig teilhaben konnten. Auch die Projektion stereoskopischer Bilder erlangte aufgrund technischer Schwierigkeiten keine große Verbreitung.

Zum Zeitpunkt als Wheatstone das Stereoskop erfand, war die Fotografie noch nicht erfunden, sodass Wheatstone auf Zeichnungen zurückgreifen musste. Erst in der zweiten Hälfte des 19. Jahrhunderts konnten stereoskopische Fotografien erstellt werden. Da die Stereoskopie schon früher verfügbar war, wurde die Stereofotografie praktisch zeitgleich mit der konventionellen Fotografie entwickelt. Erst später schwand die Bedeutung der Stereofotografie und das monoskopische Bild beherrschte den Markt.

1939 wurde auf der New Yorker Weltausstellung der View-Master (siehe Abbildung 2.3) vorgestellt, der statt Papierbildpaaren transparente Dias verwendet, die auf einer Scheibe gegenüber angeordnet werden. Auf diese Weise finden auf einer Scheibe sieben Stereobilder Platz, die durch Drehen der Scheibe nacheinander betrachtet werden können. Der View-Master und die dazupassenden Scheiben werden bis heute hergestellt, seit 1989 von der Firma Tyco Toys.

Obwohl bereits 1727 die Lichtempfindlichkeit von Silbersalzen entdeckt wurde, dauerte es noch mehr als 150 Jahre bis die Gebrüder Lumière 1895 eine brauchbare Filmkamera und einen Projektor konstruierten. 1839 gelang Louis Jacques Mande Daguerres die Verstärkung und Entwicklung eines latenten (belichteten, aber nicht sichtbaren) Bildes. Aber erst die Entdeckung lichtempfindlicherer Substanzen und die Konstruktion lichtstarker Objektive er-

laubte filmtaugliche Belichtungszeiten. Eine weitere entscheidende Vereinfachung ergab sich durch die Verwendung von Nitrozellulose als flexibles Trägermaterial. Nach der Verfügbarkeit des Rollfilms ab 1889 gelang es Thomas Alva Edison 1891, den Kinematographen und das Kinematoskop zur Aufnahme und Wiedergabe von Bildsequenzen zu konstruieren. Das Kinematoskop war kein Projektor, sondern erlaubte nur jeweils einem Benutzer den Film zu sehen. Erst der Cinematograph der Brüder Lumière ermöglichte die Projektion für ein größeres Publikum.

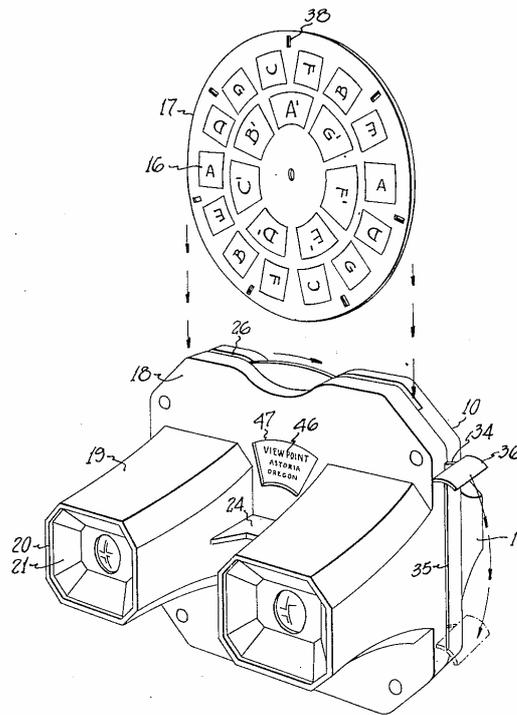


Abbildung 2.3: Prinzip des View-Master  
(aus US-Patent 2.511.334 von Wilhelm B. Gruber)

Nach der Entdeckung der Stereoskopie durch Wheatstone beschäftigten sich zahlreiche Erfinder mit der Darstellung von Stereobildsequenzen: Es wurden Geräte mit klingenden Namen wie Stereophantascope, Bioscope, Stereophotoscope oder Stereotrope konstruiert. Als erste Stereofilmkamera gilt die Konstruktion von William Friese-Green, mit der er Szenen im Hyde Park aufzeichnete. Er baute sie kurz nachdem Edison seinen Kinematographen vorstellte. Im Jahr 1903 drehten die Brüder Lumière ihren berühmten ersten stereoskopischen Kurzfilm mit dem Titel „L'Arrivée du Train“ („Die Ankunft des Zuges“). Er wurde mit zwei Kameras auf 35 mm-Film aufgenommen und dauerte ungefähr eine Minute.

Im Gegensatz zur Projektion von monoskopischen Filmen erwies sich die Stereoprojektion als schwierig. Linkes und rechtes Bild können nicht einfach auf der Leinwand übereinander projiziert werden, da sie nur dem jeweiligen Auge zugeführt werden dürfen. Drei entscheidende Erfindungen prägten die Stereoprojektion: die Anaglyphen-, die Shutter- und die Polarisationsstechnik. Alle drei Verfahren sind bis heute im Einsatz und bilden die Grundlage des 3D-Kinos.

Das Anaglyphenverfahren wurde erstmals von J. Ch. D'Almeida 1858 vorgeschlagen, erste Versuche wurden von Wilhelm Rollmann 1853 in Deutschland unternommen. 1891 patentierte der Franzose Louis Ducos du Hauron das Verfahren und nannte die entsprechenden Bilder Anaglyphen. Das Anaglyphenverfahren basiert auf der Trennung von linkem und rechtem Bild durch Farbfilter. Die Betrachtungsgeräte enthalten z.B. rote und grüne Farbfilter, die dem Auge die jeweilige Ansicht zuführen. Zuerst wurden linker und rechter Film durch Farbfilter getrennt projiziert. Später konnten die beiden Farbkomponenten auf einem einzigen Film kombiniert werden, wodurch auf den zweiten Projektor verzichtet werden konnte und das Verfahren kompatibel zur gewöhnlichen Filmprojektion wurde. Dank der einfachen Wiedergabe sind Anaglyphen bis heute die Grundlage für kostengünstiges 3D-Kino, trotz der schwerwiegenden Nachteile: Es ist bestenfalls eine eingeschränkte Farbwiedergabe möglich, außerdem verursachen Anaglyphen nach einiger Zeit bei vielen Menschen Augenschmerzen.

Der erste Film mit Anaglyphensequenzen, „Jim the Penman“, produziert von Edwin S. Porter mit Unterstützung von William E. Waddell, wurde 1915 im Astor Theater in New York gezeigt. Ein weiteres Anaglyphensystem wurde 1922 von Harry K. Fairall in Los Angeles der Öffentlichkeit vorgestellt. Zwischen 1921 und 1924 produzierten Jacob F. Leventhal und Frederic Ives erfolgreich Anaglyphenfilme, einige weitere 1936 und 1937 für MGM.

Die Shuttertechnik (Verschlusstechnik) zeigt linkes und rechtes Bild zeitsequenziell an. Laurens Hammond stellte 1922 sein Eclipse System vor, das mit synchronisierten Flügelscheiben vor den Projektor-Objektiven abwechselnd linkes und rechtes Bild projizierte. In jedem Sitz war ein Betrachter (Analysator) eingebaut, der wiederum mit einer Flügelscheibe abwechselnd den Blick des linken und rechten Auges auf die Leinwand ermöglichte. Die Motoren von Projektor- und Analysatorshutter wurden von einer gemeinsamen Wechselstromquelle gespeist, sodass die Synchronität sichergestellt war.

Das Shutterssystem hatte den Nachteil, dass es nicht in der Lage war, korrespondierende Bildpaare gleichzeitig anzuzeigen. Dadurch trat in bewegten Szenen ein unerwünschter zeitlicher Versatz auf. Der Vorteil bestand in der uneingeschränkten Tauglichkeit für Farbfilme. Wegen des hohen mechanischen Aufwandes wurde die mechanische Shuttertechnik danach nie wieder kommerziell angewandt. Erst mit der Verfügbarkeit opto-elektronischer Shutter erlangte das System Bedeutung für das 3D-Fernsehen und die computerbasierte 3D-Visualisierung.

Die Verwendung von polarisiertem Licht zur Bildtrennung wurde bereits 1891 von John Anderton vorgeschlagen, scheiterte aber an den unzureichenden Möglichkeiten zur Erzeugung polarisierten Lichts. Erst mit der Erfindung von Flächenpolarisatoren durch Erwin H. Land in Amerika, gefolgt von Zeiss Ikon in Deutschland, konnte die Polarisierungstechnik in der Praxis verwendet werden. Land nannte sein Material Polaroid und begründete damit die gleichnamige Firma.

Im Jahr 1937 erschien in Deutschland der Werbefilm „Zum Greifen nah“ für die Volksfürsorge-Lebensversicherung AG in Polarisierungstechnik. Die Phrase „Zum Greifen nah“ drückt die besondere Faszination des 3D-Films aus und wurde in der Folge für das Marketing vieler weiterer Filme verwendet. In Amerika drehte John A. Norling 1939 im Auftrag der Chrysler Corporation für die Weltausstellung in New York den Werbefilm „Motor Rhythm“.

Einige weitere Filme wurden für das 1951 in London eröffnete Telekinema gedreht, unter anderem „Royal River“ in aufwendigem Technicolor. Der Technicolor-Prozess erforderte drei Filmstreifen bei der Aufnahme, sodass für einen 3D-Film insgesamt sechs Filme gleichzeitig

belichtet werden mussten. Erst Filmmaterialien mit allen farbsensitiven Schichten auf einem Filmstreifen, wie z.B. Eastman Color, erlaubten die einfache und kostengünstigere Herstellung von 3D-Farbfilmen [65].

## 2.1.2 Der Boom in den 50ern

Anfang der 50er Jahre befand sich die Filmindustrie in den USA in der Krise. Das aufkommende Fernsehen hielt die Leute von Kinobesuchen ab und viele kreative Köpfe Hollywoods fielen der Kommunistenhatz der McCarthy-Ära zum Opfer. Hollywood suchte nach etwas Neuem, um die Besucher wieder in die Kinos zu locken.

Am 27. November 1952 hatte der von unabhängigen Filmemachern produzierte 3D-Film „Bwana Devil“ („Bwana, der Teufel“) in Los Angeles Premiere und spielte in einem einzigen Kino innerhalb einer Woche 100.000 Dollar ein. Der Erfolg ließ sich nur durch den Faktor 3D erklären, denn der Film selbst war bestenfalls durchschnittlich.

Die großen Studios erkannten das Potential – das Fernsehen konnte keinen 3D-Effekt bieten – und produzierten nun 3D-Filme im großen Stil. Alleine 1953 wurden 45 3D-Spielfilme herausgebracht, 1954 nochmals 20 [7, 24, 80]. Der 3D-Boom endete so schnell wie er begonnen hatte: Wegen vieler Unzulänglichkeiten hatte das Publikum rasch genug von der 3D-Technik und bevorzugte stattdessen die neuen CinemaScope-Filme mit breitem Bild. Im Jahr 1954 wurden zwar noch einige Filme in 3D gedreht, konnten in den Kinos aber nur mehr konventionell vorgeführt werden. Dies war leicht möglich, da zu Beginn des Boom keine speziellen 3D-Kameras zur Verfügung standen (mit Ausnahme der Kamera von Norling und jener der Brüder Milton und Julian Gunzburg, die für „Bwana Devil“ verwendet wurde). Stattdessen wurden Systeme auf Basis zweier Studiokameras entwickelt, sodass linke und rechte Aufnahme auf separaten Filmen vorlagen und somit einfach mit der althergebrachten Technik gezeigt werden konnten. Kinobesitzer erkannten schnell, dass sich mit einer 3D-Kopie zwei 2D-Vorführungen gleichzeitig durchführen lassen und machten von dieser Möglichkeit Gebrauch, um ihren Gewinn zu erhöhen. Bei den Polarisationsbrillen ließ sich ebenfalls Geld sparen, natürlich auf Kosten der Qualität.

Während die 3D-Projektion mit Polarisierungstechnik in modern ausgestatteten Kinos gut funktionierte, waren viele kleine Kinos nicht darauf vorbereitet: Schlecht eingestellte Projektoren, vertikaler Versatz von linkem und rechtem Bild, Helligkeitsunterschiede und unzureichende Synchronisation verleidetem dem Publikum den Filmgenuss. Kinos ohne geeignete Leinwand zeigten die Filme in der qualitativ unterlegenen Anaglyphentechnik. Viele Zuseher verließen die Kinos mit Kopfschmerzen und hielten die 3D-Technik generell für qualitativ unzureichend. Die genannten Faktoren trugen gemeinsam dazu bei, dass das Interesse an 3D-Filmen rasch abnahm.

Die produzierten Filme waren vom exzessiven Gebrauch von 3D-Effekten geprägt: Ständig wurden Gegenstände in Richtung Publikum geworfen oder gehalten. Zu den wenigen herausragenden Werken gehören „House Of Wax“ („Das Kabinett des Professor Bondi“), „Kiss Me Kate“ („Küss mich Kätschen“), „It Came From Outer Space“ („Gefahr aus dem Weltall“), „Hondo“ („Man nennt mich Hondo“), „Miss Sadie Thompson“ („Fegefeuer“) und Alfred Hitchcocks „Dial M for Murder“ („Bei Anruf Mord“). Letzterer wurde zwar in 3D gedreht, in den Kinos aber nur mehr „flach“ gezeigt.

### 2.1.3 Entwicklungen ab 1954

Nach 1954 wurden in unregelmäßigen Abständen 3D-Filme produziert, ohne jedoch an den großen Erfolg von 1953 anschließen zu können. Oft waren neue Filme mit der Entwicklung neuer Kamerasysteme verbunden, wie z.B. dem United Artists System (vorgestellt 1979), das synchronisierte 70 mm Kameras und Projektoren verwendete. Durch die fehlende Ausrüstung konnten 3D-Filme nur in wenigen Kinos im Polarisationsverfahren gezeigt werden, sodass ihnen ein durchschlagender Erfolg verwehrt blieb. Einer der finanziell erfolgreichsten Produktionen war „The Stuardesses“, ein Softporno von 1969. Anfang der 80er Jahre erlebte der 3D-Film ein kleines Revival. Es wurde erneut eine Reihe 3D Filme produziert, der bekannteste davon ist „Jaws 3-D“ („Der weiße Hai 3-D“). Sowohl die Anzahl der Filme als auch der kommerzielle Erfolg konnten nicht an das Jahr 1953 anknüpfen. In den letzten Jahren wurden einige 3D-Ridefilme für Theme Parks gedreht bzw. animiert, z.B. „Terminator 2: 3D“ in den Universal Studios oder Disney's „Muppet Vision 3D“.

Ab den 40er Jahren wurde auch in der Sowjetunion an der 3D-Technik gearbeitet. Der Schwerpunkt der Entwicklung lag auf der autostereoskopischen Projektion, d.h. die Filme konnten auf einer Drahtgitterleinwand ohne Brille betrachtet werden. Bereits 1947 erschien der abendfüllende Film „Robinson Crusoe“ in Farbe.

### 2.1.4 IMAX 3D

Die IMAX Corporation [34] wurde 1967 in Toronto, Kanada gegründet und entwickelte das so genannte 15/70 Filmformat. Dabei handelt es sich um ein Großformat, basierend auf 70 mm-Film [25] mit 15 Perforationslöchern pro Bild, das damit ca. zehnmals größer als ein gewöhnliches 35 mm-Bild ist. Im Unterschied zu allen anderen Filmformaten wird der Film horizontal durch den IMAX-Projektor transportiert. Die Geschwindigkeit beträgt filmtypische 24 Bilder pro Sekunde, nur beim Ridefilmsystem IMAX-HD kommen 48 Bilder pro Sekunde zur Anwendung. Der Film wird mit dem Rolling Loop System weiterbewegt, das vom Australier Ron Jones erfunden wurde. Jedes Bild wird einzeln an die Linse angepresst und mit Stiften fixiert, um einen optimalen Bildstand und konstante Schärfe zu erreichen.



Abbildung 2.4: Erster IMAX-Projektor [43]

Die Vorstellung des IMAX-Verfahrens erfolgte 1970 auf der EXPO in Osaka, Japan. In den folgenden Jahrzehnten wurden weltweit mehr als 225 IMAX-Kinos in 30 Ländern errichtet, die IMAX-Filme auf riesigen Leinwänden zeigen.

Im Jahr 1986 wurde im Rahmen der EXPO in Vancouver, Kanada, das IMAX 3D System vorgestellt. IMAX-3D arbeitet mit dem Zweibandverfahren, es laufen also zwei Filmstreifen für linkes und rechtes Bild durch die Kamera. Die Wiedergabe erfolgt ebenfalls mit zwei Filmstreifen, die vom Projektor synchron projiziert werden. In den meisten Kinos wird die Polarisationstechnik verwendet, jedoch kommen in einigen auch Shutterbrillen zur Anwendung. Der 3D-Effekt, die riesige Leinwand, die das Blickfeld fast vollständig ausfüllt, und das sechskanalige Tonsystem machen IMAX 3D bis dato zum eindrucksvollsten 3D-System.

## 2.2 3D-Fernsehen

Seit der Einführung des Fernsehens gab es unzählige Versuche, das Fernsehbild um die dritte Dimension zu erweitern [61]. Unzählige Patente dokumentieren die verschiedensten Verfahren, die dazu in Erwägung gezogen wurden. Aufgrund der vorhandenen Infrastruktur ist es einsichtig, dass nur Verfahren, die mit dem 2D-Fernsehen kompatibel sind, Chancen auf Erfolg haben. Es gab Überlegungen, zwei separate Kanäle für die Übertragung von linkem und rechtem Bild zu verwenden, doch ist diese Möglichkeit meist aus ökonomischen Gründen nicht erwünscht. Vor der Einführung des Farbfernsehens wurde vorgeschlagen, die zweite Perspektive in ungenutzten Frequenzbereichen des Fernsehkanals zu übertragen, ähnlich wie heutzutage die Farbinformationen übertragen werden.

Im Prinzip arbeitet ein Fernsehgerät in Europa (PAL-System) mit einer Bildwiederholrate von 25 Bildern pro Sekunde. Das NTSC-System, der Standard in Nordamerika und Japan, verwendet 29,97 Bilder pro Sekunde. Die Bildwiederholraten wurden in Übereinstimmung mit der landesüblichen Frequenz des Stromnetzes festgelegt. Da eine Bildröhre bei 25 Hertz aber unerträglich flimmern würde, hat man bei der Entwicklung der Fernsehsysteme folgenden Trick angewandt: Man überträgt und zeigt zuerst alle ungeraden und danach alle geraden Bildzeilen an. Diese Technik ist unter dem Begriff Zeilensprungverfahren bekannt. Dadurch ergibt sich eine Bildwiederholfrequenz von 50 Halbbildern pro Sekunde, bei gleicher Bandbreite des Übertragungskanals. Die Verwendung des Zeilensprungverfahrens war für die analoge Fernsehtechnik unumgänglich, heutzutage kann durch die Verfügbarkeit digitaler Bildspeicher die übertragene Bildwiederholrate unabhängig von der Bildwiederholfrequenz des Bildschirms gewählt werden, z.B. bei der Videowiedergabe am PC.

### 2.2.1 Fieldsequential-3D-Verfahren

Der zeitliche Versatz zweier Halbbilder ist bei der Anzeige auf modernen Displays mit progressivem Bildaufbau problematisch, da ein aufwendiges Deinterlacing durchgeführt werden muss, um Kamm-Artefakte zu vermeiden. Der Halbbildversatz ist allerdings auch die Grundlage des am weitesten verbreiteten Fieldsequential-3D-Fernsehensystems.

Das Fieldsequential- bzw. Interlaced-3D-System verwendet je ein Halbbild für das rechte und linke Bild, womit von insgesamt 576 (PAL, 480 bei NTSC) sichtbaren Bildzeilen 288 Zeilen pro Halbbild zur Verfügung stehen. Zur Betrachtung werden ein gewöhnliches Fernsehgerät und Shutterbrillen verwendet, die einen opto-elektronischen Shutter je Auge enthalten. Als Shutter

kommen meistens Flüssigkristall-Panels zur Anwendung, die man sich als LC-Displays mit einer Auflösung von nur einem Pixel vorstellen kann.

Während der Elektronenstrahl das Halbbild mit den ungeraden Zeilen schreibt, wird z.B. der rechte Shutter durchlässig geschaltet, der linke Shutter bleibt geschlossen. Während des folgenden Halbbildes mit geraden Zeilen wird der rechte Shutter geschlossen und der linke geöffnet. Es ist offensichtlich, dass jeder Zuseher eine Brille benötigt, die mit dem Videosignal, konkret mit den vertikalen Synchronimpulsen, synchronisiert sein muss.

Am Markt sind sowohl kabelgebundene Shutterbrillen als auch Ausführungen mit Infrarot-Übertragung erhältlich. In beiden Fällen extrahiert ein Controller aus dem Composite-Videosignal die Synchronimpulse und steuert damit die Brillen. Leider existiert keine Norm, die den geraden und ungeraden Zeilen die linke oder rechte Perspektive zuordnet, sodass der Benutzer selbst für die korrekte Einstellung des Controllers sorgen muss. Ein Testbild am Anfang eines Videos kann den Vorgang erleichtern.

Nicht alle Typen von Fernsehgeräten sind für die Fieldsequential-3D-Technik geeignet. Alle Geräte, die die Halbbilder verändern oder verzögert ausgeben, können nicht verwendet werden. Dazu zählen 100 Hz-, Plasma- und LCD-Fernseher sowie viele Projektortypen, insbesondere LCD-Projektoren. Abgesehen vom eingebauten Deinterlacer, der das linke und rechte Bild untrennbar vermischt, sind LCD-Geräte auch wegen der zu geringen Schaltgeschwindigkeit der Pixel ungeeignet.

Obwohl sich die Fieldsequential-Technik senderseitig vollkommen kompatibel zu 2D-Videos verhält, mit jedem Videorekorder aufgezeichnet werden kann und zumindest früher (vor der Entwicklung von 100 Hz-, Plasma und LCD-Geräten) mit allen Fernsehern kompatibel war, konnte sie sich nicht auf breiter Basis durchsetzen und fristet ein Nischendasein. Es sind zwar einige Kameras, Videokassetten und DVDs erhältlich, doch strahlt kein Sender 3D-Programm in diesem Format aus. Folgende Gründe dürften dafür ausschlaggebend sein:

- Die wenigsten Zuseher besitzen Shutterbrillen, aber ein Programm für eine verschwindend geringe Zuschauergruppe ist nicht rentabel. Eine Betrachtung ohne Brillen ist nicht möglich, da sonst linkes und rechtes Bild überlagert wahrgenommen werden. Die Brillen sind teuer im Vergleich zu Anaglyphen- und Pulfrich-Brillen (siehe unten), die meist in Kartonausführung hergestellt werden, und können deshalb nicht an die Zuschauer verschenkt werden. Endkundenpreise liegen je nach Ausführung zum Teil deutlich über 50 €.
- Durch die Shuttertechnik halbiert sich die Bildwiederholfrequenz auf 25 Hz (PAL-Norm) bzw. 30 Hz (NTSC-Norm). Diese niedrige Frequenz wird als extrem starkes Flimmern wahrgenommen und führt dazu, dass die 3D-Darstellung nach kurzer Zeit als unangenehm empfunden wird und zu Kopfschmerzen führen kann. Zwar könnten speziell für 3D optimierte 100 Hz-Fernseher (50 Hz im 3D-Modus) mit integrierter Brillen-Ansteuerung das Problem lösen, werden aber von keinem namhaften Hersteller angeboten.
- Es besteht ein Mangel an geeignetem Sendematerial, mit dem sich über einen längeren Zeitraum ein attraktives Programm zusammenstellen lässt.

## 2.2.2 Anaglyphen- und Pulfrich-Verfahren

Die Anaglyphentechnik kann seit der Einführung des Farbfernsehens nicht nur im Kino, sondern auch für Fernsehsendungen verwendet werden. Schon im Kino wurden Anaglyphen von vielen Zusehern als unangenehm empfunden und waren nur eine Notlösung, um Kosten für

die hochwertige Projektion mit der Polarisierungstechnik zu sparen. Die Qualität von Anaglyphen stellt hohe Anforderungen an die Qualität der Farbwiedergabe, sodass die niedrige Bandbreite der Farbkomponenten eines Videosignals besonders negativ ins Gewicht fällt. Obwohl vereinzelt Sendungen in der Anaglyphentechnik ausgestrahlt wurden, kann sie nicht als brauchbares 3D-Verfahren für das Fernsehen eingestuft werden. Besonders analoge Videoaufnahmen auf VHS-Kassetten sind aufgrund der nochmals reduzierten Bandbreite völlig unbrauchbar.

Fast alle Fernsehsendungen, die mit „3D“ beworben wurden, sind im Pulfrich-Verfahren produziert worden. Es handelt sich dabei um kein echtes 3D-Verfahren, da kein Stereobildpaar vorliegt, sondern der 3D-Effekt durch eine optische Täuschung zustande kommt.

Die Kamera befindet sich beim Pulfrich-Verfahren ständig in Bewegung. Sie bewegt sich entweder von rechts nach links an der Szene vorbei oder im Uhrzeigersinn um ein Objekt herum. Während der Wiedergabe bewegen sich die Objekte daher von links nach rechts über den Bildschirm, der Vordergrund mit einer höheren Geschwindigkeit als der Hintergrund. Anstelle einer Kamerabewegung wäre auch eine entgegengesetzt gerichtete Bewegung der Szene möglich.

Der Benutzer trägt eine Brille mit einem hellen linken und einem abgedunkelten rechten Glas. Das Pulfrich-Verfahren basiert auf einer speziellen Eigenschaft der menschlichen visuellen Wahrnehmung, die von Carl Pulfrich am Anfang des 20. Jahrhunderts in Deutschland entdeckt wurde. Im Vergleich zu einem Bild normaler Helligkeit werden abgedunkelte Bilder verzögert wahrgenommen. Durch die Bewegung nimmt das rechte Auge ein bestimmtes Bild erst wahr, wenn das linke Auge die Szene bereits aus einer leicht anderen Perspektive sieht. Es entsteht ein Versatz zwischen übereinstimmenden Punkten, der umso größer wird je schneller sich das Objekt bewegt, also je näher es sich an der Kamera befindet, und so den 3D-Effekt hervorruft.

Das Pulfrich-Verfahren wird vor allem deshalb gerne für Fernsehsendungen verwendet, weil auch Zuseher ohne 3D-Brille den Film ohne Qualitätsverlust in Farbe genießen können – nur eben ohne 3D-Effekt. Die Brillen lassen sich, genauso wie Anaglyphenbrillen, sehr kostengünstig herstellen und werden meist als Beilage von Fernsehmagazinen vor entsprechenden Sendungen an die Zuseher verteilt. In der Regel wird nicht die ganze Sendung mit Pulfrich-Effekt gesendet, sondern nur Ausschnitte, da die monotone Bewegung auf Dauer als lästig empfunden wird. Besonders Naturdokumentationen und Erotiksendungen eignen sich besonders gut für das Pulfrich-Verfahren.

---

# Räumliches Sehen

---

*Dieses Kapitel erklärt das Phänomen der Tiefenwahrnehmung, beschreibt die Grundlagen des stereoskopischen Sehens (Stereopsis) und beleuchtet auch monokulare Tiefenhinweise. Ferner werden die Grundlagen der stereoskopischen Bildaufnahme und -wiedergabe abgeleitet.*

## 3.1 Konvergenz und Akkomodation

Konvergenz und Akkomodation gehören zur Klasse der okulomotorischen, die Beweglichkeit des Auges betreffenden, Reize. Der Begriff Konvergenz bezeichnet die Fähigkeit der menschlichen Augen, die Sehachsen auf den im Moment betrachteten Raumpunkt auszurichten. Blicken wir auf den Horizont, verlaufen die Sehachsen parallel, bei Betrachtung eines nahen Gegenstands schneiden sich die Achsen an dieser Stelle, verlaufen also konvergierend. Das Gehirn ist in der Lage, aus der Spannung der sechs Muskeln je Auge die Entfernung des fixierten Punktes abzuschätzen. Die Muskulatur ist bei paralleler Ausrichtung der Sehachsen entspannt, mit steigender Konvergenz nimmt die Spannung zu. Den Winkel  $\alpha$  am Schnittpunkt der Sehachsen bezeichnet man als Konvergenzwinkel (siehe Abbildung 3.1).

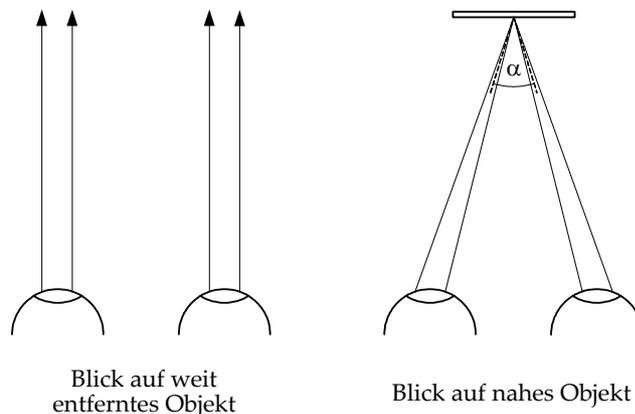


Abbildung 3.1: Konvergenz und Akkomodation

Unter Akkomodation versteht man das Scharfstellen des Auges auf eine bestimmte Entfernung. Dabei wird durch die Anspannung des Ziliarmuskel die Linse verformt und deren Sammelwirkung nimmt zu, um eine scharfe Abbildung naher Objekte zu erreichen.

Konvergenz und Akkomodation arbeiten beim normalen Sehen gekoppelt, der Akkomodationspunkt wird immer dem Konvergenzpunkt angepasst. Bei der Wiedergabe von Stereobil-

dern, egal mit welchem Verfahren, kann das Auge ebenfalls durch die Szene wandern und verschiedene Objekte fixieren. Je nach virtueller Entfernung des Objekts ist dabei eine unterschiedliche Konvergenz nötig. Im Gegensatz dazu muss die Akkomodation jedoch immer auf dieselbe Entfernung eingestellt werden – die Entfernung des Bildschirms oder der Leinwand (Abbildungs- bzw. Wiedergabeebene) vom Auge. Man spricht dabei von der Trennung von Konvergenz und Akkomodation. Je mehr sich Konvergenz- und Akkomodationspunkt unterscheiden, also je weiter ein Objekt vor oder hinter der Abbildungsebene erscheint, desto unnatürlicher und anstrengender wird dessen Betrachtung.

Abbildung 3.2 zeigt die Situation bei hilfsmittelfreier Betrachtung eines Bildpaares mit dem Parallelblick (siehe Kapitel 4.2). Während die Augen auf einen unendlich weit entfernten Punkt ausgerichtet sind und nicht konvergieren, fokussieren sie auf die nahe Abbildungsebene.

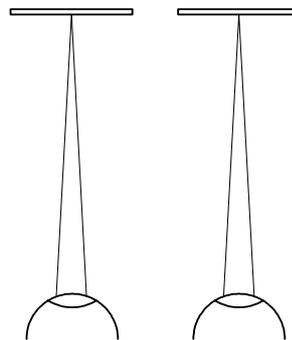


Abbildung 3.2: Konvergenz und Akkomodation bei Wiedergabe eines Stereobildes

## 3.2 Parallaktische Verschiebung

Der Mensch nimmt den größten Teil des Gesichtsfeldes binokular, also mit zwei Augen, wahr. Die frontale Anordnung der Augen in einem Abstand von rund 6,5 cm führt zu leicht unterschiedlichen Abbildungen auf der Netzhaut, die uns die Abschätzung von Entfernungsdifferenzen ermöglichen. Diese Anordnung der Augen ist keinesfalls typisch für alle Lebewesen: Bei vielen Tieren sind sie außenseitig angeordnet, um durch ein Gesichtsfeld von 360° Feinde von allen Seiten wahrnehmen zu können.

Der fixierte Punkt, z.B. Punkt 1 der Abbildung 3.3, wird auf der Fovea abgebildet, jener Stelle der Netzhaut mit höchster Sehschärfe. Die Abbildung des zweiten Punktes befindet sich im linken Auge links, im rechten Auge rechts der Fovea. Punkt 3 wird zwar in beiden Augen links der Fovea abgebildet, jedoch unterscheiden sich die Abstände deutlich, der Punkt wird ebenfalls an nicht-korrespondierenden Netzhautorten abgebildet. Diese Abstandsunterschiede werden in der Physiologie Querdissipation oder binokulare Disparität genannt, im Zusammenhang der Stereoskopie spricht man von parallaktischer Verschiebung oder Parallaxe. Die Differenz von maximaler und minimaler Querdissipation einer stereoskopischen Abbildung oder eines einzelnen Objekts wird Deviation genannt.

Die Querdissipation ist eine essentielle Voraussetzung für die räumliche Wahrnehmung, da das menschliche Gehirn in der Lage ist, die Verschiebung als Tiefeninformation zu interpretieren. Die genaue Funktionsweise der Verarbeitung der beiden Bilder im Gehirn ist nicht restlos

geklärt und wird an dieser Stelle nicht weiter erläutert. Es sei diesbezüglich auf die Fachliteratur verwiesen [8, 20].

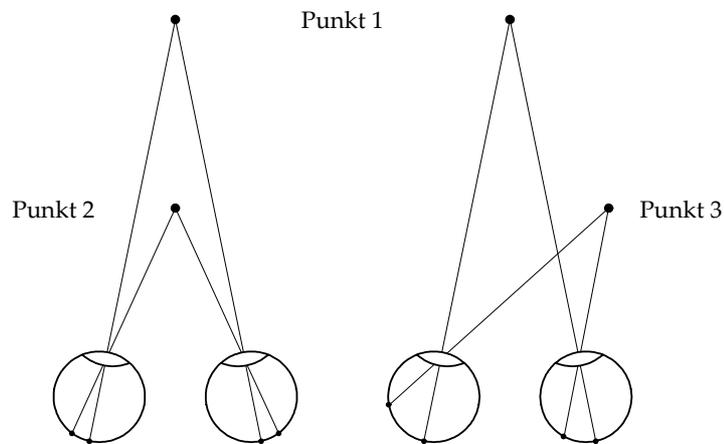


Abbildung 3.3: Querdisparation

### 3.3 Monokulare Tiefenhinweise

Neben Konvergenz, Akkomodation und Querdisparation gibt es viele weitere visuelle Hinweise, die eine Abschätzung der Entfernung erlauben und auch bei monokularem (einäugigem) Sehen zur Verfügung stehen [9].

#### 3.3.1 Verdeckung

Wird ein Objekt von einem anderen teilweise verdeckt, dann lässt sich daraus schließen, dass das verdeckte hinter dem verdeckenden Objekt liegt. Die Verdeckung erlaubt nur die Ermittlung der Tiefenreihenfolge, die tatsächlichen Abstände lassen sich nicht bestimmen.

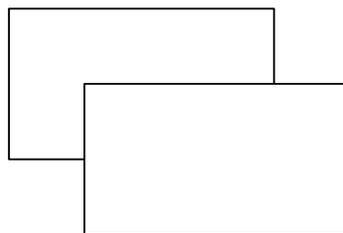


Abbildung 3.4: Verdeckung als Tiefenhinweis

#### 3.3.2 Größe

Aus der wahrgenommenen Größe eines Objekts lässt sich auf dessen Entfernung rückschließen, sofern die tatsächliche Größe bekannt ist. Je kleiner es erscheint, desto weiter ist es entfernt, je größer es erscheint, desto näher liegt es beim Betrachter.

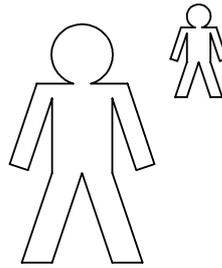


Abbildung 3.5: Größe als Tiefenhinweis

### 3.3.3 Perspektive

Parallele Linien werden auf der Netzhaut nicht parallel abgebildet, sondern sie konvergieren zu einem Punkt (Fluchtpunkt). Man denke zum Beispiel an eine Straße, deren Ränder sich mit zunehmender Entfernung immer weiter zu nähern scheinen, bis sie sich am Horizont treffen. Da bekannt ist, dass die Straße ihre Breite beibehält, interpretiert man die abnehmende Breite der Netzhautabbildung als Tiefeninformation.

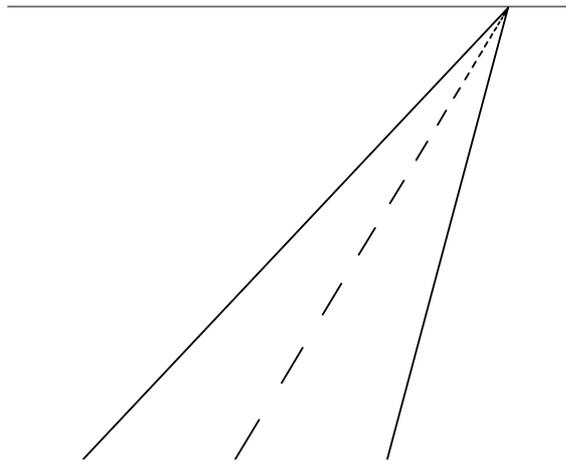


Abbildung 3.6: Perspektive als Tiefenhinweis

### 3.3.4 Dunst

Die Tiefenwahrnehmung aufgrund von Konvergenz und Querdisparation ist auf die nähere Umgebung beschränkt. Ab einer Entfernung von ungefähr 2,5 Kilometern ist der Augenabstand des Menschen nicht mehr ausreichend, um durch das binokulare Sehen die Tiefe von Objekten unterscheiden zu können. Für große Entfernungen liefert der atmosphärische Dunst (Wasserdampf, Staubpartikel) entscheidende Hinweise, um die Entfernungsunterschiede von Objekten im Raum wahrzunehmen. Mit zunehmender Entfernung wird die Gegend von einem bläulichen Schleier überlagert und verliert an Kontrast und Farbsättigung.

Abbildung 3.7 zeigt zweimal die gleiche Gegend, das linke Bild mit und das rechte Bild ohne atmosphärischen Dunst. Man sieht, dass sich die Bergrücken ohne Dunst kaum von dahinter liegenden Bergen abheben, während mit Dunst die Berge räumlich gestaffelt erscheinen.

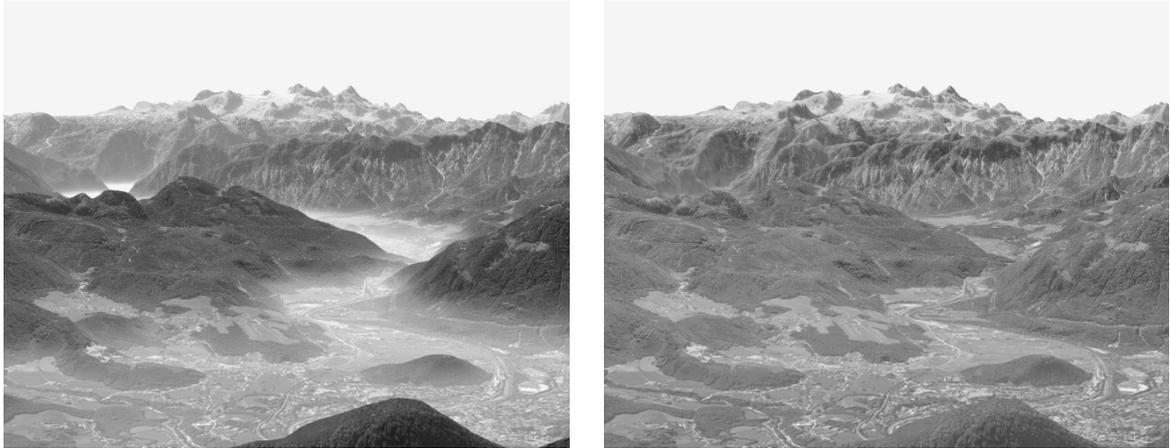


Abbildung 3.7: Dunst als Tiefenhinweis

Während bei einer stereoskopischen Kameraaufnahme alle Tiefenhinweise automatisch im Bild vorhanden sind, müssen sie bei computergenerierten Inhalten explizit berücksichtigt werden. Ein gutes, realistisch wirkendes Stereobildpaar erfordert nicht nur eine geeignete Wahl des Kameraabstandes (Stereobasis), sondern auch das Vorhandensein der monokularen Tiefenhinweise.

### 3.3.5 Bewegungsparallaxe

Die Bewegungsparallaxe gehört zur Kategorie der bewegungsinduzierten, monokularen Tiefenhinweise und entsteht, wenn sich entweder der Betrachter oder das Umfeld bewegen. Je näher sich ein Objekt beim Betrachter befindet, desto schneller erscheint dessen Bewegung, während der Horizont immer stillzustehen scheint.

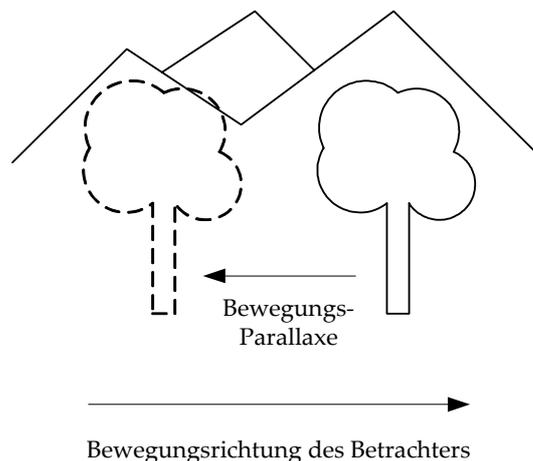


Abbildung 3.8: Bewegungsparallaxe als Tiefenhinweis

Abbildung 3.8 zeigt eine Situation wie sie z.B. beim Blick aus einem Zug oder Auto auftritt. Der Baum bewegt sich verhältnismäßig rasch durch das Blickfeld und muss sich deshalb nahe beim Betrachter befinden, während die stillstehende Bergkette weit entfernt sein muss.

## 3.4 Grenzen der räumlichen Wahrnehmung

Der geringe Abstand der Augen führt dazu, dass sich aus der Querdissipation ab einer Grenzentfernung von mehr als ungefähr 2,5 km keine Tiefeninformationen gewinnen lassen. Dieser Wert folgt daraus, dass sich Differenzen des Konvergenzwinkels von bis zu 5 Bogensekunden unterscheiden lassen. Im Vergleich dazu beträgt das Auflösungsvermögen des Auges nur 1 Bogenminute.

Die Grenzauflösung von 5 Bogensekunden ist allerdings nur für die Betrachtung einer realen Szene von Bedeutung, da die erforderliche Auflösung von stereoskopischen Abbildungen nicht erreicht wird und somit die Grenzentfernung auf ca. 100 Meter sinkt (bei Verwendung des Augenabstandes als Aufnahme-Stereobasis).

Der menschliche Sehapparat unterliegt weiteren Einschränkungen, die das räumliche Sehen betreffen und bei der Aufnahme und Wiedergabe stereoskopischer Bilder berücksichtigt werden müssen. Die folgenden Kapitel 3.4.1 bis 3.4.3 gehen näher darauf ein.

### 3.4.1 Fusionsbereich

Unter dem Panum'schen Fusionsraum versteht man jenen Tiefenbereich vor und hinter der Fixationsentfernung, innerhalb dessen Objekte zu einem Bild verschmelzen und einfach gesehen werden. Die Fusion funktioniert nicht nur im Bereich des Sehentrums, sondern immer dann, wenn ein Punkt auf korrespondierende Stellen auf der Netzhaut abgebildet wird. Dies ist genau für jene Punkte der Fall, die von beiden Augachsen im gleichen Winkel abweichen. Die Menge dieser Punkte ergibt einen Kreis, der theoretischer Horopter genannt wird. Empirisch ermittelte Horoptoren weichen von der idealen Kreisform ab.

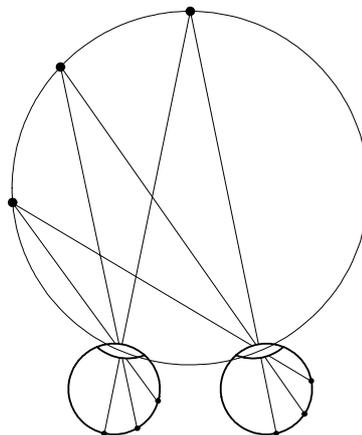


Abbildung 3.9: Theoretischer Horopter

In der Literatur [29, 37] wird der zulässige Konvergenzwinkelunterschied des Panum'schen Fusionsraums im Bereich des Sehentrums mit ungefähr 6 Bogenminuten angegeben, mit steigendem Abstand vom Sehzentrum erhöht sich der Wert. Außerhalb des Panum'schen Fusionsraums kommt es zu keiner vollständigen Fusion der Bilder, stattdessen werden Doppelbilder wahrgenommen. Alle genannten Werte zur Tiefenwahrnehmung sind empirisch ermittelte Durchschnittswerte, die sich von Person zu Person unterscheiden und auch in der

Literatur nicht immer einheitlich angegeben werden. In Experimenten ist es sowohl gelungen, den Fusionsbereich weit über die üblichen Grenzen auszudehnen, als auch die Fusion innerhalb des Panum'schen Fusionsraums zu stören.

### 3.4.2 Bildzerfall

In der Stereoskopie ist weniger der Fusionsbereich als jener Bereich von Bedeutung, in dem die räumlichen Bezüge in der Szene wahrgenommen werden können. Umfasst die Abbildung einen zu großen Tiefenbereich, dann können die Elemente eines Stereobildes nicht mehr als Ganzes, sondern Vorder- und Hintergrund nur mehr isoliert voneinander gesehen werden. Konvergieren die Augen z.B. auf den Vordergrund, dann ist die Querdisparation des Hintergrunds so groß, dass er als weit horizontal versetztes Doppelbild erscheint und seine Entfernung nicht mehr in der Szene eingeordnet werden kann. Können Bildebenen nur mehr einzeln erfasst werden, dann spricht man von Bildzerfall.

Für den zulässigen Tiefenbereich lässt sich kein fester Wert angeben, weil eine starke Abhängigkeit vom Bildinhalt besteht. Zum Beispiel ist ein freistehendes Objekt im Vordergrund problematischer als ein Objekt, das sich vom Hintergrund gleich weit in den Vordergrund erstreckt, aber den Augen die Möglichkeit bietet, am Objekt entlang zu wandern und dadurch die Tiefe leichter zu erfassen.

Hermann Lüscher [42] gab die zulässige Differenz des Konvergenzwinkels zwischen Nah- und Fernpunkt mit 70 Bogenminuten an. Dieser Wert wird von vielen Stereofotografen als Richtwert verwendet und ist in der Fachliteratur häufig als 70-Minuten-Bedingung zu finden, obwohl er gemäß jüngeren Erkenntnissen zu niedrig sein dürfte. Größere Werte sind besonders dann vertretbar, wenn sich der Tiefenbereich lokal auf kleinere Werte beschränkt.

### 3.4.3 Divergenz

Von divergenter Augenstellung spricht man, wenn die Sehachsen nicht auf einen Punkt konvergieren oder parallel verlaufen, sondern sich mit zunehmendem Abstand voneinander entfernen. Während eine divergente Augenstellung beim natürlichen Sehen niemals auftritt, ist eine horizontale Divergenz bei einer großformatigen, stereoskopischen Projektion möglich. Für die Stereowiedergabe gilt eine horizontale Divergenz bis  $1^\circ$  als akzeptabel, darüber hinaus wird das Auge stark angestrengt oder eine Fusion ist überhaupt nicht mehr möglich. Manche Quellen sehen eine Divergenz bis maximal  $1,5^\circ$  als zulässig an. Idealerweise vermeidet man eine unnatürliche, divergente Augenstellung gänzlich [14, 41].

Bei vertikaler Divergenz ist eine Sehachse stärker nach unten oder oben gerichtet als die andere. Die Ursachen für eine vertikale Divergenz liegen in mangelhaft ausgerichteten Kameras oder Projektoren (bei Wiedergabe mit zwei Projektoren). Vertikaler Versatz aufgrund von Aufnahme Fehlern kann unter Auflösungsverlust im Zuge der Nachbearbeitung am PC korrigiert werden, bei einer Echtzeitübertragung und der Justierung zweier Projektoren muss jedoch unbedingt auf eine exakte Ausrichtung geachtet werden. Um eine angenehme Betrachtung zu gewährleisten, darf der vertikale Versatz  $0,5^\circ$  nicht überschreiten.

## 3.5 Aufnahme- und Wiedergabegeometrie

Das Ziel der Stereoskopie ist eine ansprechende Wiedergabe der räumlichen Tiefe. Daraus folgt unmittelbar die Frage nach der korrekten Platzierung der beiden Kameras für die Aufnahme und der korrekten Aufbereitung der Bilder für Medien mit unterschiedlicher Wiedergabetechnik und -größe [28]. Kann dasselbe Stereofoto sowohl für eine großformatige Projektion als auch für einen Druck mit nur wenigen Zentimetern Größe verwendet werden?

### 3.5.1 Kameraanordnung und Scheinfenster

Im Prinzip folgt die Platzierung der Kameras der menschlichen Augenanordnung. Linke und rechte Kamera werden auf gleicher Höhe nebeneinander angeordnet und nehmen linkes und rechtes Bild in gleicher Weise auf, wie unsere Augen die Szene wahrnehmen. Linkes und rechtes Einzelbild ergeben zusammen ein Stereobild, dass bei stereoskopischer Betrachtung zu einem virtuellen Raumbild verschmilzt. Durch Verändern des Kameraabstandes lässt sich der Tiefeneffekt steuern, vorausgesetzt linke und rechte Aufnahmeeinheit befinden sich in getrennten Gehäusen. Die genauen Zusammenhänge werden in den folgenden Kapiteln hergeleitet.

Beim natürlichen Sehen konvergieren die Augachsen und schneiden sich im fixierten Punkt, dessen parallaktische Verschiebung auf der Netzhaut verschwindet. Durch Zusammenschwenken der Kameraachsen lässt sich diese Situation nachbilden, allerdings entstehen dadurch im Gegensatz zur parallelen Ausrichtung trapezförmige Verzerrungen.

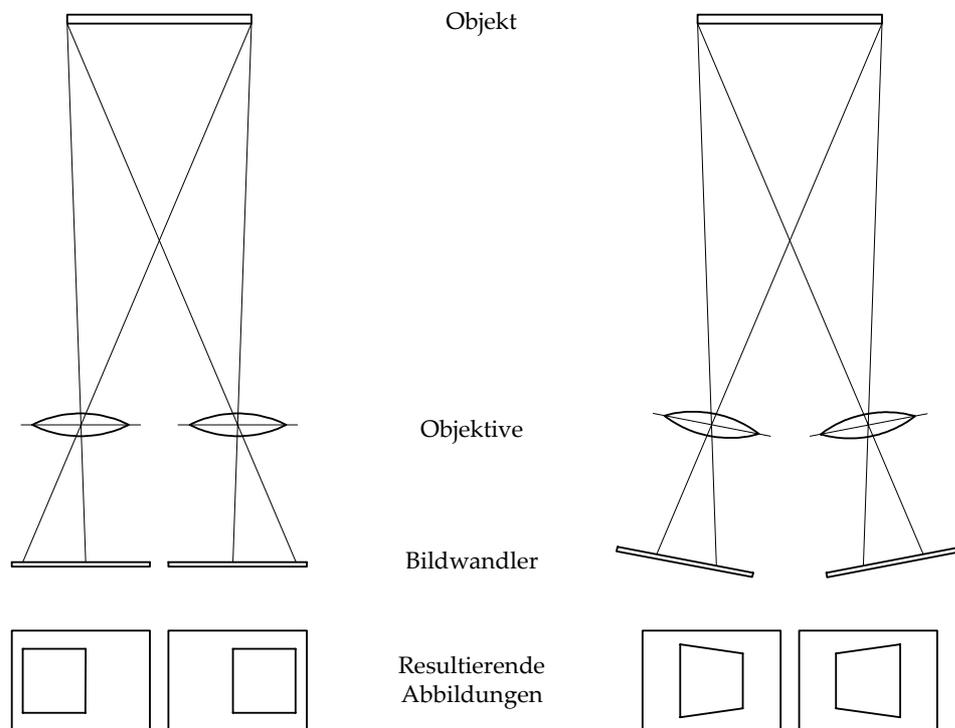


Abbildung 3.10: Parallele und konvergierende Kameraanordnung

Dem Wunsch nach konvergierenden Kameraachsen liegt die Scheinfensterregel zugrunde. Unter dem Scheinfenster versteht man die Abbildung der Bildbegrenzung von linkem und rechtem Halbbild im virtuellen Raum, die Szene wirkt wie durch ein Fenster betrachtet. Dieses Scheinfenster liegt genau in jener Entfernung, in der sich die Kameraachsen schneiden und die parallaktische Verschiebung verschwindet. Die Scheinfensterweite gibt den Normalabstand von der Kamera zu jener gedachten Ebene in der realen Szene an, in der die Abbildung der Halbbilderbegrenzungen erscheinen würde. Im Falle paralleler Kameraachsen kommt das Scheinfenster in unendlicher Entfernung zu liegen, mit zunehmendem Konvergenzwinkel rückt es in den Vordergrund. Die parallaktische Verschiebung wird für Objekte hinter dem Scheinfenster positiv, für Objekte vor dem Scheinfenster negativ gezählt.

Die Scheinfensterregel besagt, dass aus ästhetischen Gründen kein Objekt vom Rahmen des Scheinfensters angeschnitten werden darf. Daraus folgt, dass alle Bestandteile der Szene hinter dem Scheinfenster liegen müssen, also eine positive parallaktische Verschiebung aufweisen müssen. Ausgenommen sind Objekte, die aus dem Scheinfenster ragen, ohne dieses zu berühren. Es ist offensichtlich, dass die Scheinfensterregel bei paralleler Kameraausrichtung verletzt wird. Um sowohl die Scheinfensterregel zu erfüllen als auch Trapezverzerrungen zu vermeiden, kann das Scheinfenster durch linksseitiges Beschneiden des linken Halbbildes und rechtsseitiges Beschneiden des rechten Halbbildes in den Vordergrund geschoben werden.

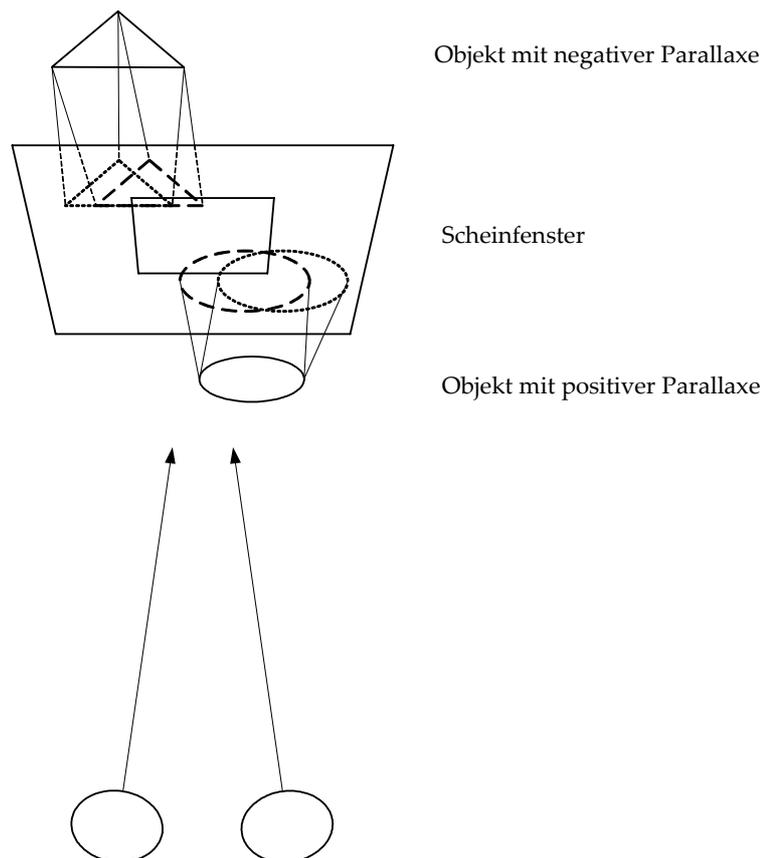


Abbildung 3.11: Scheinfenster und parallaktische Verschiebung

Bei der Wiedergabe liegt das Scheinfenster in der Monitor- bzw. Leinwandebene, außer man führt eine Verschiebung der beiden Halbbilder durch, sodass sich deren Ränder nicht mehr

decken. Demzufolge erscheint die gesamte Szene weiter im Vordergrund oder Hintergrund, die Lage des Scheinfensters in Bezug auf die restlichen Elemente der Szene bleibt indessen unverändert. Wird das linke Halbbild nach rechts und das rechte Halbbild nach links versetzt, verschiebt sich die Szene nach vorne, bei umgekehrter Richtung nach hinten.

Stereofotografen bevorzugen die parallele Kameraausrichtung, da die Scheinfensterkorrektur durch Verschieben der Dias im Rahmen einfach erledigt werden kann. Im Videobereich wird dagegen manchmal eine konvergierende Anordnung verwendet, um einen Auflösungsverlust aufgrund der Beschneidung zu vermeiden. Solange der verzerrungsbedingte, vertikale Versatz unter der Auflösungsgrenze des Bildwandlers bleibt, kann dieser toleriert werden, jedoch ist diese Voraussetzung bereits bei Verwendung zweier Videokameras mit einer Auflösung von 480 (NTSC, VGA) oder 576 Zeilen (PAL) in den meisten Aufnahmesituationen nicht mehr erfüllt.

Wenn eine Beschneidung durchgeführt wird und das Originalseitenverhältnis erhalten bleiben soll, muss diese an allen vier Rändern durchgeführt werden. Meist wird danach das verkleinerte Halbbild auf die ursprüngliche Auflösung skaliert, da alle gebräuchlichen Bandformate eine feste Auflösung verlangen und auch die meisten Videocodecs Auflösungsrestriktionen unterliegen.

Durch außermittige Platzierung der Bildwandler in der Kamera lässt sich eine Scheinfensterkorrektur ohne Auflösungsverlust und Trapezverzerrungen realisieren. Zur Anpassung an unterschiedliche Aufnahmesituationen sollte sich die Position der Bildwandler verändern lassen. Bei der Auslegung der Objektive ist darauf zu achten, dass die Abbildung am Bildwandler unabhängig von dessen Position ausreichend scharf und nicht vignettiert ist.

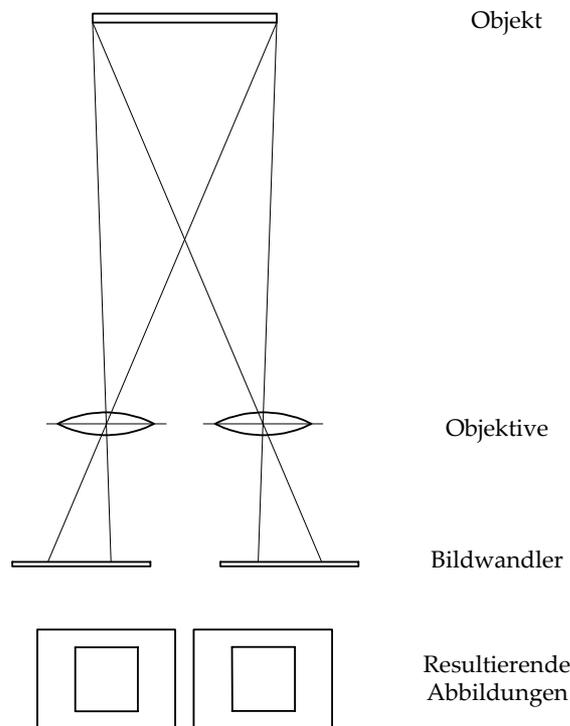


Abbildung 3.12: Scheinfensterkorrektur durch außermittige Bildwandler

### 3.5.2 Linsengleichung

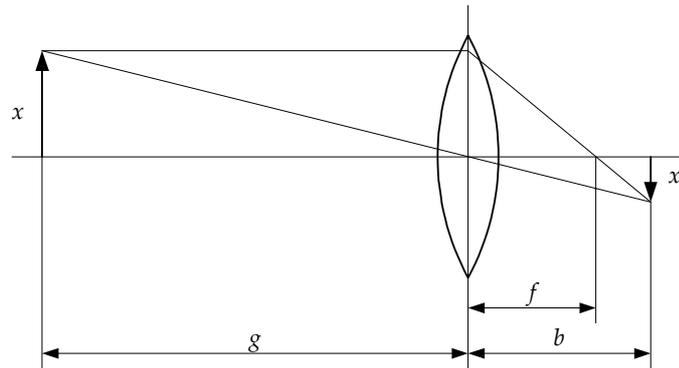


Abbildung 3.13: Zur Linsengleichung

In den folgenden Ableitungen wird die Linsengleichung für dünne Linsen verwendet. Sie lautet

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{b} \quad (3.1)$$

wobei  $f$  die Brennweite,  $g$  die Gegenstandsweite und  $b$  die Bildweite bezeichnet. Im Falle der Abbildung auf der Netzhaut oder am Bildwandler der Kamera ist die Gegenstandsweite sehr viel größer als die Bildweite und kann vernachlässigt werden.

$$\frac{1}{f} \approx \frac{1}{b} \Rightarrow f \approx b \quad (3.2)$$

Die Bildweite kann daher durch die Brennweite approximiert werden. Für den Vergrößerungsfaktor  $x'/x$  folgt:

$$\frac{x'}{x} = \frac{b}{g} \approx \frac{f}{g} \quad (3.3)$$

### 3.5.3 Abbildung im Auge

Auf den folgenden Seiten werden formelmäßige Zusammenhänge für die Aufnahme und Wiedergabe stereoskopischer Abbildung hergeleitet, gültig sowohl für Videos als auch Fotos. Die Herleitung und Nomenklatur der Variablen orientiert sich an [29] und [31].

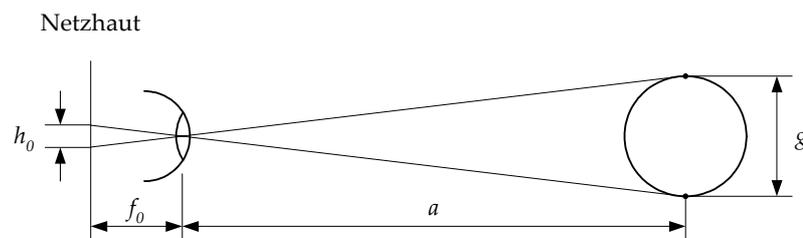


Abbildung 3.14: Größenabbildung im Auge

Unter Zuhilfenahme der zuvor getroffenen Vereinfachungen wird die Größe (Breite bzw. Höhe) eines Objekts, wie in Abbildung 3.14 illustriert, gemäß

$$h_0 = \frac{f_0}{a} \cdot g \quad (3.4)$$

auf die Netzhaut transformiert. Reale Größen werden ohne Index, Größen des Auges werden mit dem Index 0 bezeichnet.  $a$  bezeichnet die Gegenstandsweite, also den Abstand des Objekts vom Auge,  $g$  dessen Größe bzw. Höhe,  $f_0$  die Brennweite des Auges und  $h_0$  die Höhe der Netzhautabbildung.

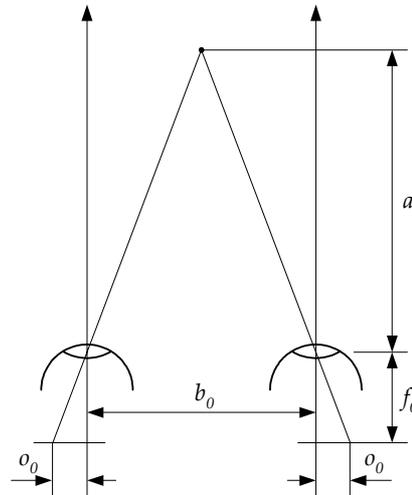


Abbildung 3.15: Parallaxenabbildung im Auge

Das Sehzentrum gewinnt Informationen über den Abstand eines Objekts aus dessen parallaktischer Verschiebung der Netzhautabbildung. Abbildung 3.15 zeigt die Parallaxe  $o_0$  für einen Punkt mit dem Abstand  $a$ . Wir legen fest, dass die Parallaxe des Unendlichkeitspunktes 0 beträgt und mit abnehmender Entfernung des Objekts zunimmt. Aus diesem Grund wurden die Augachsen in Abbildung 3.15 parallel eingezeichnet. Beziehung (3.5) kann mit dem Strahlensatz hergeleitet werden.  $b_0$  ist der Augenabstand von durchschnittlich 6,5 cm.

$$o_0 = \frac{b_0 \cdot f_0}{2 \cdot a} \quad (3.5)$$

Zur Berechnung der Tiefenabbildung im Auge wird gemäß Abbildung 3.16 von einem Objekt mit der Tiefe  $t$  ausgegangen. Der Punkt mit dem geringsten Abstand zu den Augen wird Raum-Nahpunkt genannt, die zugehörige Entfernung Nahpunktweite  $a_N$ . Der am weitesten entfernte Punkt heißt Raum-Fernpunkt, dessen Abstand von den Augen Fernpunktweite  $a_F$ . Die Abbildungen dieser beiden Punkte werden Bild-Nahpunkt und Bild-Fernpunkt genannt. Die Deviation  $d_0$  berechnet sich durch zweimaliges Anwenden des Strahlensatzes zu:

$$d_0 = \frac{b_0 \cdot f_0}{a_N} - \frac{b_0 \cdot f_0}{a_F} = b_0 \cdot f_0 \cdot \frac{a_F - a_N}{a_F \cdot a_N} \quad (3.6)$$

Die Differenz  $a_F - a_N$  entspricht der Objekttiefe  $t$ , das Produkt aus Fern- und Nahpunktweite kann durch den geometrischen Mittelwert  $a$  ausgedrückt werden.

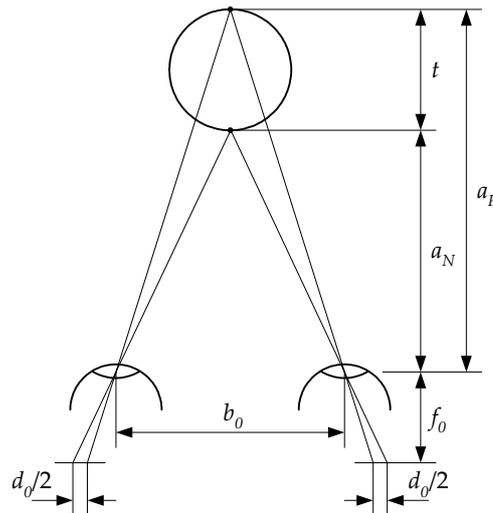


Abbildung 3.16: Deviations- bzw. Tiefenabbildung im Auge

$$a = \sqrt{a_F \cdot a_N} \quad (3.7)$$

$$d_0 = \frac{b_0 \cdot f_0}{a^2} \cdot t \quad (3.8)$$

### 3.5.4 Virtueller Abstand

Um die Wirkung einer stereoskopischen Abbildung beurteilen zu können, definieren wir den virtuellen Abstand. Dieser gibt an, in welcher Entfernung sich das virtuelle Objekt zu befinden scheint. Die virtuelle Entfernung entspricht in der Regel nicht dem tatsächlichen Abstand des Objekts von der Kamera zum Zeitpunkt der Aufnahme, sondern hängt, wie wir sehen werden, von verschiedenen Faktoren ab. Zum Beispiel wird ein Objekt bei Betrachtung auf einem Computermonitor näher (und auch kleiner) erscheinen als bei der Projektion auf einer riesigen Leinwand.

Die Berechnung der virtuellen Entfernung erfolgt über die Herleitung der parallaktischen Verschiebung des virtuellen Bildes auf der Netzhaut. Dazu berechnen wir die parallaktische Verschiebung, die sich bei Betrachtung des Bildes ergibt und nehmen an, dass ein Objekt in jener Entfernung erscheint, an der ein reales Objekt dieselbe parallaktische Verschiebung hervorrufen würde.

Im ersten Schritt ermitteln wir die parallaktische Verschiebung am Bildwandler der Kamera,  $o_k$ . Die Herleitung erfolgt gemäß der bereits bekannten Abbildung 3.15 und Formel (3.5), da die Abbildung in der Kamera äquivalent zu jener im Auge verläuft. Kameragrößen werden mit dem Index  $k$  gekennzeichnet. Der Abstand der beiden Objektivachsen, Stereobasis genannt, wird mit  $b_k$  bezeichnet und hat im Gegensatz zum Augenabstand  $b_0$  keinen festen Wert, sondern ist jene Größe, durch deren Veränderung die gewünschte Tiefenwirkung eingestellt werden kann. Formeln zur Wahl einer geeigneten Stereobasis werden später hergeleitet.

$$o_k = \frac{b_k \cdot f_k}{2 \cdot a} \quad (3.9)$$

Die Halbbilder werden vergrößert auf dem Monitor, der Leinwand oder am Papier wiedergegeben. Wir definieren hierzu einen Vergrößerungsfaktor  $V$ , der das Größenverhältnis zwi-

schen der Wiedergabeabbildung und der Abbildung am Bildwandler angibt. Der Index  $p$  (projiziert) bezeichnet Größen des Wiedergabemediums.

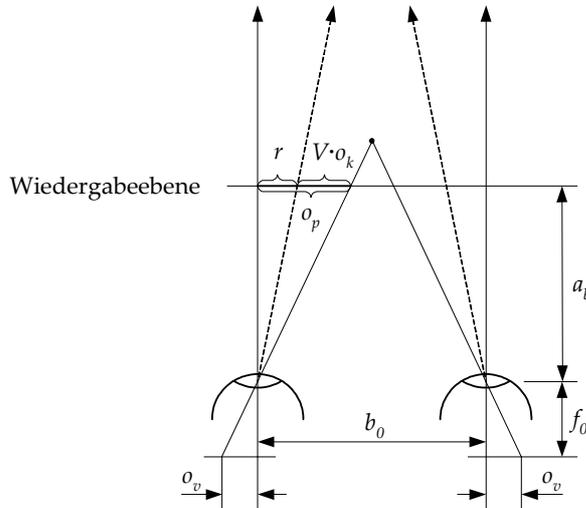


Abbildung 3.17: Parallaxische Verschiebung

Die parallaxische Verschiebung wird mit dem Vergrößerungsfaktor  $V$  am Wiedergabemedium abgebildet. Es ist dabei aber nicht garantiert, dass die Unendlichkeitspunkte im Augenabstand zu liegen kommen, sodass eine Verschiebung um  $r$  berücksichtigt werden muss, welche die virtuelle Entfernung stark beeinflusst. Die strichlierten Linien in Abbildung 3.17 konvergieren zur Abbildung des realen Unendlichkeitspunkts, dessen Entfernung in der virtuellen Abbildung endlich ist.

$$o_p = V \cdot o_k + r = V \cdot \frac{b_k \cdot f_k}{2 \cdot a} + r \tag{3.10}$$

Eine Abbildung mit Unendlichkeitspunkten im Abstand von ca. 6,5 cm ( $r = 0$ ) lässt sich nur auf einer Leinwand oder mit dem Parallelblick erzielen, ansonsten liegen die Unendlichkeitspunkte näher beisammen und werden mit einem positiven Wert für  $r$  berücksichtigt. Negative Werte für  $r$  führen zu einer divergenten Augenstellung, falls der Unendlichkeitspunkt (Horizont) in der Abbildung auftritt, und müssen deshalb vermieden werden. Ansonsten sind negative Werte erlaubt, solange die parallaxische Verschiebung 6,5 cm, bezogen auf die Wiedergabeebene, nicht überschreitet (wie im Kapitel 3.4.3 beschrieben, kann auch eine etwas größere Verschiebung noch toleriert werden). Die Gefahr, negative Werte für  $r$  zu erhalten, besteht besonders bei großformatigen Projektionen.

Das Wiedergabemedium wird aus der Entfernung  $a_b$  betrachtet. Unter Berücksichtigung der Brennweite des Auges lässt sich die Parallaxe auf der Netzhaut berechnen. Für Größen der Abbildung des virtuellen Objekts auf der Netzhaut verwenden wir den Index  $v$ .

$$o_v = \frac{f_0}{a_b} \cdot o_p = \frac{f_0}{a_b} \cdot \left( V \cdot \frac{b_k \cdot f_k}{2 \cdot a} + r \right) \tag{3.11}$$

Der Zusammenhang für die Netzhaut-Parallaxe eines realen Objekts und dessen Entfernung ist durch Gleichung (3.5) gegeben. Anstelle des realen Abstandes  $a$  setzen wir den virtuellen Abstand  $a_v$  ein.

$$o_v = \frac{b_0 \cdot f_0}{2 \cdot a_v} \quad (3.12)$$

Durch Gleichsetzen von (3.11) und (3.12) ergibt sich die virtuelle Entfernung  $a_v$ .

$$a_v = \frac{a_b \cdot b_0}{2} \cdot \frac{1}{o_p} = \frac{1}{V} \cdot \frac{a_b \cdot b_0}{b_k \cdot f_k} \cdot \frac{b_k \cdot f_k}{b_k \cdot f_k + 2 \cdot r \cdot a} \cdot a \quad (3.13)$$

An dieser Stelle führen wir die Abkürzung  $R(a)$  ein, in der Stereofotografie Rahmungsfaktor genannt. Weiters definieren wir den Tiefenfaktor  $T$  und schreiben unter Berücksichtigung dieser Größen die Beziehung für die virtuelle Entfernung  $a_v$  neu an.

$$R(a) = \frac{b_k \cdot f_k}{b_k \cdot f_k + 2 \cdot r \cdot a} \quad (3.14)$$

$$T = \frac{a_b}{V \cdot f_k} \quad (3.15)$$

$$a_v = \frac{b_0}{b_k} \cdot T \cdot R(a) \cdot a \quad (3.16)$$

### 3.5.5 Virtuelle Größe und Tiefe

Die Abbildung der Objektgröße und –tiefe in der Kamera erfolgt gleich wie deren Abbildung im Auge (siehe Kapitel 3.5.3). Der Augenabstand  $b_0$  und die Brennweite  $f_0$  der Formeln (3.4) und (3.8) werden durch die Stereobasis  $b_k$  und die Kamerabrennweite  $f_k$  ersetzt.

$$h_k = \frac{f_k}{a} \cdot g \quad (3.17)$$

$$d_k = \frac{b_k \cdot f_0}{a^2} \cdot t \quad (3.18)$$

Sowohl Objekthöhe als auch die Deviation werden auf dem Wiedergabemedium mit dem Faktor  $V$  vergrößert dargestellt.

$$V = \frac{h_p}{h_k} = \frac{d_p}{d_k} \quad (3.19)$$

$$h_p = V \cdot \frac{f_k}{a} \cdot g \quad (3.20)$$

$$d_p = V \cdot \frac{b_k \cdot f_k}{a^2} \cdot t \quad (3.21)$$

Als nächster Schritt erfolgt die Abbildung des Stereobildes vom Wiedergabemedium auf die Netzhaut des Auges. Wiederum nehmen wir an, dass das Wiedergabemedium aus einer Entfernung  $a_b$  betrachtet wird und berechnen die Größe und Deviation des virtuellen Objekts auf der Netzhaut. Die Transformation erfolgt äquivalent zu Formel (3.4), sowohl für die Größe  $h_v$  als auch für die Deviation  $d_v$ .

$$h_v = \frac{f_0}{a_b} \cdot h_p = V \cdot \frac{f_0 \cdot f_k}{a_b \cdot a} \cdot g \quad (3.22)$$

$$d_v = \frac{f_0}{a_b} \cdot d_p = V \cdot \frac{f_0 \cdot b_k \cdot f_k}{a_b \cdot a^2} \cdot t \quad (3.23)$$

Im letzten Schritt ermitteln wir, wie groß und tief das virtuelle Objekt erscheint. Die virtuelle Größe  $g_v$  bestimmen wir aus der virtuellen Bildhöhe auf der Netzhaut  $h_v$ , die virtuelle Objektiefe  $t_v$  aus der virtuellen Deviation  $d_v$ .

$$h_v = \frac{f_0}{a_v} \cdot g_v \quad (3.24)$$

$$d_v = \frac{b_0 \cdot f_0}{a_v^2} \cdot t_v \quad (3.25)$$

Gleichsetzen von (3.22) mit (3.24) und (3.23) mit (3.25) liefert die gewünschten Ergebnisse, der Vollständigkeit halber ist auch Gleichung (3.16) nochmals aufgeführt.

$$g_v = V \cdot \frac{a_v \cdot f_k}{a \cdot a_b} \cdot g = \frac{b_0}{b_k} \cdot R(a) \cdot g \quad (3.26)$$

$$t_v = V \cdot \frac{a_v^2 \cdot b_k \cdot f_k}{a^2 \cdot a_b \cdot b_0} \cdot t = \frac{b_0}{b_k} \cdot T \cdot R^2(a) \cdot t \quad (3.27)$$

$$a_v = \frac{b_0}{b_k} \cdot T \cdot R(a) \cdot a \quad (3.16)$$

Eine Korrektur der parallaktischen Verschiebung mit  $r > 0$  hat einen Rahmungsfaktor  $R(a) < 1$  und eine lineare Abnahme der Entfernung und Größe, aber eine quadratische Abnahme der Tiefe zur Folge. Das virtuelle Objekt erscheint verglichen mit den ursprünglichen Proportionen flacher.

## 3.6 Liliputismus und Gigantismus

Die Ergebnisse von Gleichung (3.26), (3.27) und (3.16) überraschen, denn man hätte erwartet, dass die virtuelle Tiefe  $t_v$  mit zunehmender Stereobasis  $b_k$  ebenfalls zunimmt. Gemäß Gleichung (3.27) ist aber genau das Gegenteil der Fall. Die Ursache liegt in der virtuellen Entfernung begründet, die mit zunehmender Stereobasis abnimmt. Das Objekt wandert in den Vordergrund und erscheint dadurch kleiner, denn die Abbildungsgröße auf der Netzhaut bleibt unabhängig von der Stereobasis konstant. Ein nahes Objekt kann aber nur dann eine gleich große Abbildung wie ein entferntes Objekt erzeugen, wenn es kleiner als das entfernte ist [27].

Die Verkleinerung der virtuellen Abbildung aufgrund einer Aufnahmestereobasis größer als der Augenabstand wird als Liliputismus oder Modellwirkung bezeichnet. Das Raumbild wirkt plastischer, erscheint aber verkleinert. Eine große Stereobasis wird für Motive ohne Vordergrund benötigt, z.B. wenn als Aufnahmestandort ein steiler Gipfel gewählt wird und eine durch ein Tal getrennte Gebirgskette aufgenommen wird. Eine Stereobasis in der Größenord-

nung des Augenabstandes würde ein Bild ohne jede erkennbare Tiefenwirkung zur Folge haben, erst durch eine Stereobasis im Meter- bis 100-Meter-Bereich gewinnt das Bild an Tiefe. Der Preis dafür ist der Liliputismus: Beim Betrachten des Raumbildes entsteht der Eindruck als betrachte man eine Modelllandschaft. Leider wirken Bilder mit großer Stereobasis oft wenig ästhetisch.

Aufnahmen mit vergrößerter Basis sind nur dann möglich, wenn linker und rechter Bildwandler nicht fest in ein Gehäuse montiert sind, sondern in Form zweier separater Kameras in beliebigem Abstand positioniert werden können.

Der Gigantismus, manchmal auch Mammutwirkung genannt, ist das Gegenteil des Liliputismus. Er geht mit Stereobasen unter dem Augenabstand einher. Das Raumbild wird in den Hintergrund geschoben, die Größe und Tiefe werden verstärkt. Kleine Stereobasen werden für Makroaufnahmen benötigt. Die Ergebnisse sind im Gegensatz zum Liliputismus meist ansprechend, z.B. die Abbildung eines Insekts, dessen Flügelspannweite plötzlich im Meterbereich zu liegen scheint. Makroaufnahmen erfordern immer einen Rahmungsfaktor  $R(a) > 0$ , um das wenige Zentimeter entfernte Objekt in einer ausreichenden virtuellen Entfernung zu platzieren.

In der Regel sind mit der Standardausführung von Stereokameras keine Makroaufnahmen möglich, da die Kameras aufgrund Ihrer Abmessungen nicht nahe genug zusammengeschoben werden können, sondern es werden weitere optische Vorrichtungen benötigt.

### 3.7 Ortho- und Tautostereoskopie

Ein Raumbild heißt orthostereoskopisch, wenn es gegenüber dem natürlichen Sehen beim Betrachter einen formrichtigen Eindruck erzeugt. Ist eine orthostereoskopische Abbildung auch noch maßstabsgetreu, dann heißt sie tautostereoskopisch.

Das orthostereoskopische Raumbild bildet also alle Objekte mit der gleichen Größen- und Tiefenskalierung ab, eine Kugel erscheint als Kugel und nicht als Ei. Eine tautostereoskopische Abbildung erfüllt die zusätzliche Bedingung, dass alle Objekte in Originalgröße dargestellt werden.

Wann ist ein Raumbild ortho- bzw. tautostereoskopisch? Dazu betrachten wir die folgenden Beziehungen mit den Abbildungskoeffizienten  $s_g$  (Größe),  $s_t$  (Tiefe) und  $s_a$  (Entfernung):

$$\begin{aligned} g_v &= s_g \cdot g \\ t_v &= s_t \cdot t \\ a_v &= s_a \cdot a \end{aligned} \tag{3.28}$$

Eine orthostereoskopische Abbildung muss offenbar die Bedingung

$$s_g = s_t \tag{3.29}$$

erfüllen, während für Tautostereoskopie

$$s_g = s_t = s_a = 1 \tag{3.30}$$

gelten muss. Die Abbildungskoeffizienten lassen sich aus einem Koeffizientenvergleich von Gleichung (3.28) mit den Gleichungen (3.16), (3.26) und (3.27) ermitteln.

$$\begin{aligned} s_g &= \frac{b_0}{b_k} \cdot R(a) \\ s_t &= \frac{b_0}{b_k} \cdot T \cdot R^2(a) \\ s_a &= \frac{b_0}{b_k} \cdot T \cdot R(a) \end{aligned} \quad (3.31)$$

Um die Bedingung für Tautostereoskopie zu erfüllen, muss  $R(a) = 1$  sein und damit  $r = 0$ . Folglich müssen die Unendlichkeitspunkte am Wiedergabemedium im Augenabstand zu liegen kommen. Weiters muss die Aufnahmestereobasis  $b_k$  dem Augenabstand  $b_0$  entsprechen und der Tiefenfaktor  $T$  muss 1 ergeben.

$$T = \frac{a_b}{V \cdot f_k} = 1 \quad (3.32)$$

Die Abhängigkeit des Tiefenfaktors vom Betrachtungsabstand hat zur Folge, dass sich nur bei der Betrachtung vom orthostereoskopischen Punkt ein ortho- bzw. tautostereoskopisches Raumbild ergibt.

Unter der Annahme, dass  $R(a) = 1$  ist, ist die einzige Voraussetzung für eine orthostereoskopische Wiedergabe dass  $T = 1$  gilt. Die Stereobasis kann theoretisch frei gewählt werden, praktisch muss sie der Aufnahmesituation angepasst werden, um einen ausreichenden Tiefeneindruck zu erhalten oder um Bildzerfall zu vermeiden.

Im Falle  $R(a) \neq 1$  folgt aus Bedingung (3.29):

$$s_g = s_t \quad (3.29)$$

$$\frac{b_0}{b_k} \cdot R(a) = \frac{b_0}{b_k} \cdot T \cdot R^2(a) \quad (3.33)$$

$$T \cdot R(a) = \frac{a_b}{V \cdot f_k} \cdot R(a) = 1 \quad (3.34)$$

Bedingung (3.34) ist abstandsabhängig und lässt sich nur für eine einzige (mittlere) Entfernung  $a = a_m$  im Raumbild erfüllen, nähere oder weiter entfernte Objekte erscheinen nicht mehr exakt formtreu. Da Objekte nahe der gewählten Entfernung  $a_m$  vom ermittelten Betrachtungspunkt  $a_b$  aus gesehen halbwegs unverzerrt erscheinen, wird er quasi-orthostereoskopischer Punkt genannt.

### 3.8 Wahl der Stereobasis

Die Bedingungen für Ortho- und Tautostereoskopie liefern zwar Regeln für ein mathematisch korrektes Raumbild, die Praxis hat aber gezeigt, dass sie nicht die ästhetisch ansprechendsten Ergebnisse liefern. Der Betrachter bevorzugt meist eine Abbildung mit leicht übertriebener Tiefe, und findet andererseits einen zu großen Tiefenumfang als störend.

Es hat sich bewährt, ausgehend von der Annahme einer maximal zulässigen Differenz des Konvergenzwinkels zwischen Nahpunkt und Fernpunkt, die Aufnahmeparameter festzulegen. Wir nennen den Konvergenzwinkel des Nahpunktes  $\alpha_1$  und jenen des Fernpunktes  $\alpha_2$ , somit folgt für den Differenzwinkel:

$$\alpha = \alpha_1 - \alpha_2 \quad (3.35)$$

Der Betrachtungsabstand  $a_b$  ist wesentlich größer als die Deviation  $d_p$  auf dem Wiedergabemedium, sodass mit

$$\tan\left(\frac{\alpha}{2}\right) = \frac{d_p}{2 \cdot a_b} \quad (3.36)$$

der Winkel  $\alpha$  gut angenähert wird (siehe Abbildung 3.18). Da der Winkel  $\alpha$  klein ist, kann der Tangens vernachlässigt werden und die Beziehung einfach nach  $d_p$  aufgelöst werden. Der Winkel  $\alpha$  muss in Radianten eingesetzt werden!

$$\tan\left(\frac{\alpha}{2}\right) \approx \frac{\alpha}{2} \quad (3.37)$$

$$d_p = \alpha \cdot a_b \quad (3.38)$$

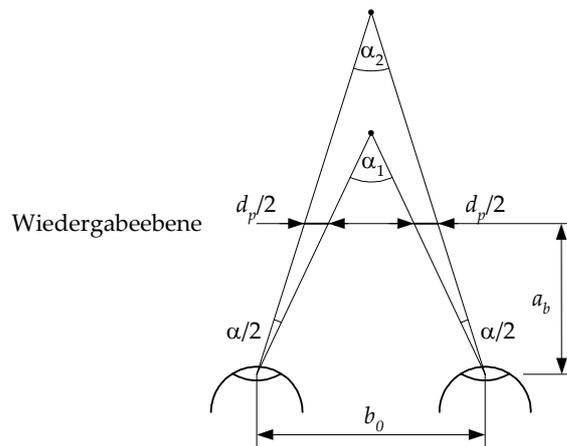


Abbildung 3.18: Zur Ermittlung der zulässigen Deviation

Es ist zweckmäßig, die zulässige Deviation auf die Gesamtbreite des Bildes  $w_p$  zu beziehen. Für das Verhältnis von Bildbreite  $w_p$  zu Betrachtungsabstand  $a_b$  führen wir die Abkürzung  $P$ , die Präsenz, ein.

$$P = \frac{w_p}{a_b} \quad (3.39)$$

$$\frac{d_p}{w_p} = \frac{\alpha \cdot a_b}{w_p} = \frac{\alpha}{P} \quad (3.40)$$

Zur Ermittlung der maximal zulässigen Deviation in einem Stereobild setzen wir in (3.40) die maximal zulässige Differenz der Konvergenzwinkel und die maximal gewünschte Präsenz ein. Obwohl die 70-Minuten-Bedingung heute als zu restriktiv gilt, hat sie sich in der Praxis

bewährt, sodass wir an dieser Stelle für  $\alpha$  einen Wert von 70 Bogenminuten (0,02 rad) wählen. Für die Präsenz eignet sich ein Wert von 0,6. Er wird in den meisten Wiedergabeszenarien nicht überschritten. Tabelle gibt Präsenzwerte für die Wiedergabe auf einem Monitor und auf einer Leinwand an.

Szenario	$w_p$	$a_b$	$P$
19"-Röhrenmonitor	36 cm	65 cm	0,55
Projektion, kleine Leinwand	1,5 m	2 – 6 m	
Projektion, mittlere Leinwand	3 m	4 – 12 m	0,75 – 0,25
Projektion, große Leinwand	6 m	8 - 24 m	

Tabelle 3.1: Typische Präsenzwerte

$$\left. \frac{d_p}{w_p} \right|_{\max} = \left. \frac{d}{w} \right|_{\max} = \frac{0,02}{0,6} = \frac{1}{30} \quad (3.41)$$

Formel (3.41) ist ein zentrales Ergebnis und besagt, dass die Deviation zwischen Nah- und Fernpunkt maximal 1/30 der Raumbildbreite betragen darf.

Bei gegebener Nah- und Fernpunktweite kann durch Adaption von Beziehung (3.6) die erforderliche Aufnahmebasis berechnet werden. Erwartungsgemäß geht in die Berechnung sowohl die Kamerabrennweite  $f_k$  als auch die Breite des Bildwandlers  $w_k$  ein.

$$d_k = b_k \cdot f_k \cdot \frac{a_f - a_n}{a_f \cdot a_n} \quad (3.42)$$

$$b_k = \frac{d_k}{f_k} \cdot \frac{a_f \cdot a_n}{a_f - a_n} = \frac{d_k}{w_k} \cdot \frac{w_k}{f_k} \cdot \frac{a_n}{1 - a_n / a_f} \quad (3.43)$$

Einsetzen von (3.41) liefert unser Endergebnis. Trotz zahlreicher Näherungen und Annahmen liefert Formel (3.44) in den meisten Aufnahmesituationen eine Stereobasis, die zu ansprechenden Ergebnissen führt.

$$b_k = \frac{1}{30} \cdot \frac{w_k}{f_k} \cdot \frac{a_n}{1 - a_n / a_f} \quad (3.44)$$

Ist der Horizont in der Aufnahme sichtbar, gilt für die Fernpunktweite  $a_f \rightarrow \infty$  und (3.44) kann nochmals vereinfacht werden:

$$b_k = \frac{1}{30} \cdot \frac{w_k \cdot a_n}{f_k} \quad (3.45)$$

Im Gegensatz zu zweidimensionalen Aufnahmen ist im Raumbild Tiefenunschärfe unerwünscht, da das Auge beim Umherwandern im virtuellen Raum alle Tiefenbereich scharf sehen sollte. Um eine große Tiefenschärfe zu erhalten, sollte die Blende der Kamera möglichst weit geschlossen werden.

Es sei abschließend darauf hingewiesen, dass die Faustformel Stereobasis = Nahpunktweite / Brennweite, die häufig in der Literatur zur Stereofotografie auftaucht, nur für Aufnahmen auf Kleinbildfilm gilt [30, 32].

### 3.9 Raumbildfehler

Raumbildfehler bezeichnen Störungen im reproduzierten Raumbild, die einer ermüdungsfreien Betrachtung im Weg stehen oder die Verschmelzung der Halbbilder gänzlich verhindern. Oft sind die Fehler während der Wiedergabe nicht sofort zu erkennen und äußern sich nur durch ein unangenehmes Gefühl, es entsteht der Eindruck, dass „etwas nicht stimmt“. Raumbildfehler können sowohl durch Aufnahme-, Bearbeitungs- oder Wiedergabefehler entstehen.

Die häufigsten Fehler sind im Folgenden zusammengefasst, zum Teil wurden sie auch schon in vorhergehenden Kapiteln detailliert behandelt.

- *Störbilder*, auch Geisterbilder oder Ghosting genannt, entstehen durch ungenügende Trennung der Halbbilder. Das linke Auge nimmt Teile des rechten Halbbildes wahr und umgekehrt. Störbilder sind *das* Problem vieler Wiedergabemethoden.
- *Pseudostereoskopische Abbildungen* entstehen durch vertauschen des linken und rechten Bildes. Ursprünglich nahe Objekte erscheinen im Hintergrund, ferne Objekte im Vordergrund. Mit Ausnahme einfacher geometrischer Objekte ergibt die Vertauschung der Halbbilder kein sinnvolles Raumbild. Besonders Anfänger neigen dazu, linke und rechte Perspektive zu verwechseln, sofern sie nicht gekennzeichnet sind.
- *Retinale Rivalität* entsteht, wenn die beiden Augen unterschiedliche Bildinhalte wahrnehmen. Eine der Abbildungen wird dabei zugunsten der dominanten Abbildung vom Sehzentrum unterdrückt. In der Praxis tritt retinale Rivalität in Stereofotos auf, die mit dem Sukzessivverfahren erzeugt wurden (linke und rechte Perspektive werden nacheinander mit einer Kamera aufgenommen), wenn sich die Szene in der Zeit zwischen den Aufnahmen verändert hat (sich bewegende Personen, Fahrzeuge, sich im Wind bewegende Äste, Blätter und vieles mehr). Im Falle synchroner Bildpaare kann retinale Rivalität durch Reflexionen verursacht werden, die nur in einer der Aufnahmen sichtbar sind.
- Von *Größenfehler* spricht man, wenn die Halbbilder verschieden groß wiedergegeben werden. Ursachen können unterschiedliche Aufnahme- oder Projektionsobjektive sein, besonders bei Verwendung von Zoomobjektiven ist größte Vorsicht geboten.
- *Höhenfehler* bezeichnen einen vertikalen Versatz von linken und rechten Bildpunkten und sind die Folge mangelhafter Sorgfalt bei der Kamera- oder Projektorjustierung. Partielle Höhenfehler entstehen außerdem in Folge der Trapezverzerrung bei konvergenten Aufnahmeachsen (Kapitel 3.5.1).
- *Rotationsfehler* werden ebenfalls durch mangelnde Genauigkeit bei Aufnahme oder Wiedergabe hervorgerufen.
- *Bildzerfall* oder *übertriebene Tiefenwirkung* sind eine Folge der fehlerhaften Wahl von Stereobasis und Kamerabrennweite (Kapitel 3.4.2 und 3.8).
- *Divergente Augachsen* ermüden die Augen oder verhindern die Verschmelzung der Halbbilder gänzlich (Kapitel 3.4.3)
- *Helligkeits-, Farbton- oder Farbsättigungsunterschiede* der Halbbilder können sowohl durch mangelnden Abgleich von Aufnahme- als auch Wiedergabegeräten auftreten.

---

# Wiedergabe

---

*Zur Wiedergabe stereoskopischer Videos stehen zahlreiche Methoden für unterschiedliche Bildschirm- und Projektortypen zur Verfügung, deren Qualität, Aufwand, Kosten und Einsatzgebiete stark voneinander abweichen. Dieses Kapitel stellt alle bedeutenden Methoden vor.*

## 4.1 Einteilung der Wiedergabemethoden

Die stereoskopischen Wiedergabemethoden können nach der Art der eingesetzten Hilfsmittel in vier Kategorien eingeteilt werden.

- Die *hilfsmittelfreie Stereowiedergabe* erlaubt das Betrachten von stereoskopischen Inhalten ohne spezielle Brillen, Displays oder andere Hilfsmittel. Es gibt zwei sehr ähnliche Verfahren, den Parallelblick und den Kreuzblick. Beide erfordern die bewusste Trennung von Konvergenz und Akkomodation des Auges und müssen deshalb trainiert werden.
- *Passive Systeme*: Ein Wiedergabesystem wird als passiv bezeichnet, wenn die Wiedergabe auf einem konventionellem, zweidimensionalen Medium erfolgt und mit Brillen ohne elektronische Ansteuerung betrachtet wird. Im Gegensatz zu den hilfsmittelfreien Betrachtungsmethoden erfordern passive Systeme keine Übung – der räumliche Eindruck stellt sich Dank der Brillen sofort ein. Zu den passiven Systemen zählen Anaglyphen, die Polarisationstechnik und das KMQ-Verfahren.
- *Aktive Systeme* verwenden ebenfalls ein zweidimensionales Wiedergabemedium, allerdings verfügen sie über eine aktive, elektronische Ansteuerung der Brille. In diese Kategorie fallen alle Varianten von Shutterbrillensystemen.
- Head Mounted Displays, autostereoskopische Displays und andere Konstruktionen, die speziell für die 3D-Wiedergabe optimiert oder entwickelt wurden, werden zur Gruppe der *3D-Displays* zusammengefasst.

## 4.2 Parallelblick

Zum Betrachten einer stereoskopischen Abbildung mit dem Parallelblick (Straight Eye Viewing) ordnet man linkes und rechtes Halbbild nebeneinander an. Anschließend versucht man, den Konvergenzwinkel der Augen zu verringern, um die beiden Halbbilder zu fusionieren. Obwohl ein weit entfernter, virtueller Punkt fixiert wird, muss anschließend auf die Abbildungsebene akkomodiert werden.

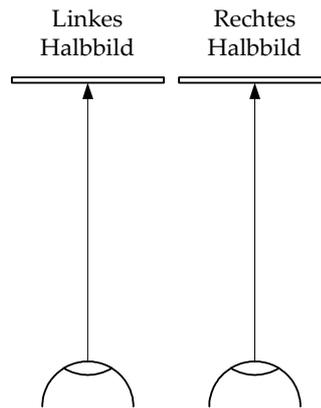


Abbildung 4.1: Parallelblick

Verlässt der Konvergenzpunkt die Abbildungsebene, wird jedes Halbbild doppelt wahrgenommen, sodass man insgesamt vier Halbbilder sieht. Ziel ist es, die zwei inneren Halbbilder zur Deckung zu bringen und anschließend die Schärfe in die Darstellungsebene zurückzuverlagern. Ist der Parallelblick hergestellt, wird das mittige Raumbild von den zwei verbliebenen, zweidimensionalen Halbbildern links und rechts flankiert.

Der Parallelblick gelingt am leichtesten, wenn man die Augen entspannt und durch die Papirebene oder den Monitor „hindurchsieht“. Eine helle Umgebung verkleinert die Iris, erhöht die Tiefenschärfe und erleichtert dadurch die Fokussierung.



Abbildung 4.2: Parallelblick-Beispiel

Es ist offensichtlich, dass die Halbbildbreite auf 6,5 cm beschränkt werden muss, um eine divergente Augenstellung zu vermeiden. In der Praxis ist es vorteilhaft, eine etwas geringere Breite zu verwenden, da einerseits die Verschmelzung erleichtert wird, andererseits durch die Scheinfensterkorrektur Hintergrundelemente eine größere Verschiebung aufweisen, aber trotzdem unter 6,5 cm bleiben müssen. Die Halbbildbreite von Abbildung 4.2 beträgt 6 cm, der Versatz des Hintergrunds 6,25 cm.

Durch eine entsprechende Einstellung der Druckgröße kann die Breitenbeschränkung problemlos erfüllt werden. Die Aufbereitung von Stereobildern für die Monitorarstellung birgt aber Probleme, denn die endgültige Abbildungsgröße ist sowohl von der Monitorgröße als auch von der gewählten Bildschirmauflösung abhängig. Kann der Anwender die Darstel-

lungsgröße nicht selbst verändern, zum Beispiel bei Stereobildern auf einer Internetseite, dann sollte die zulässige Größe zumindest auf einem 19"-Monitor (sichtbare Breite: 36 cm) bei einer Auflösung von 1024 × 786 Pixel nicht überschritten werden. Daraus folgt, dass das gesamte Stereobild nicht breiter als 380 Pixel sein sollte. Nimmt man eine Divergenz von 1,5° in Kauf und legt einen Betrachtungsabstand von 50 cm zugrunde, erhöht sich der zulässige Wert auf 410 Pixel.

Optische Betrachtungshilfen mit Sammellinsen können die Betrachtung erleichtern, allerdings ist dann die Hilfsmittelfreiheit nicht mehr erfüllt. Stark kurzsichtige Personen erzielen denselben Effekt durch Abnehmen der Brille. Keillinsenbetrachter nach dem Prinzip des Brewster-Stereoskops (siehe Kapitel 2.1.1) erlauben auch die Betrachtung von Bildern mit einer Breite über 6,5 cm.

### 4.3 Kreuzblick

Der Kreuzblick (Schieltechnik, Crossed Eye Viewing) basiert ebenfalls auf nebeneinander angeordneten Halbbildern, nur werden diese vertauscht angeordnet: das linke Halbbild rechts, das rechte Halbbild links. Der Konvergenzpunkt liegt vor der Abbildungsebene. Der Kreuzblick kann am einfachsten hergestellt werden, indem man den Daumen in ungefähr 10 cm Entfernung vor die Augen hält und anblickt. Im Hintergrund erkennt man unscharf vier Halbbilder. Anschließend verschiebt man den Daumen nach vorne oder hinten bis sich die zwei inneren Halbbilder decken und verlagert den Fokus danach in den Hintergrund, ohne die Konvergenz zu ändern.

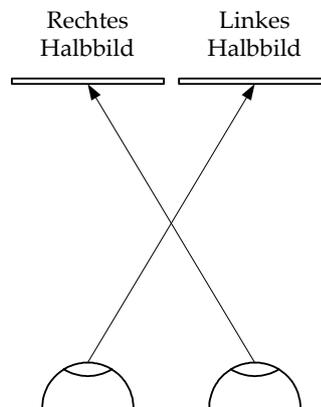


Abbildung 4.3: Kreuzblick

Das Raumbild erscheint nah (vor dem Schnittpunkt der beiden Sehachsen) und klein. Gemäß der Definitionen von Kapitel 3 erhöht sich die parallaktische Verschiebung um

$$r = \frac{w_p + 6,5 \text{ cm}}{2}. \quad (4.1)$$

$w_p$  bezeichnet die Halbbildbreite, wobei vorausgesetzt werden muss, dass kein Abstand zwischen den Halbbildern liegt. Im Gegensatz dazu erhöht sich die parallaktische Verschiebung beim Parallelblick nicht ( $r = 0$ ). In beiden Fällen muss zusätzlich der Einfluss der Scheinfensterkorrektur auf  $r$  berücksichtigt werden.

Vielen Menschen bereiten beide hilfsmittelfreie Verfahren auch nach zahlreichen Trainingsversuchen Schwierigkeiten, manche sind überhaupt nicht in der Lage, Konvergenz und Akkomodation zu trennen. Es ist individuell verschieden, ob der Parallelblick oder der Kreuzblick leichter erlernt werden kann. Beide führen aber auf jeden Fall zu einer raschen Ermüdung der Augen und eignen sich deshalb nicht für lange Präsentationen. Besonders für die Betrachtung von Videosequenzen sind beide Verfahren schlecht geeignet, da das Beibehalten der Augenstellung nach Szenenwechseln oft nicht gelingt. Die Halbbildtrennung ist prinzipiell immer perfekt, Störbilder sind deshalb nie vorhanden.



Abbildung 4.4: Kreuzblick-Beispiel

## 4.4 Anaglyphen

Das Anaglyphenverfahren überlagert die Halbbilder und realisiert die Bildtrennung mit Farbfilterfolien. Linkes bzw. rechtes Bild werden unterschiedlich eingefärbt und der Betrachter trägt eine Brille mit Farbfilter derselben Farben, wobei es zweckmäßig ist, Kombinationen der Primärfarben Rot, Grün und Blau zu verwenden. Je nach Art der Farbkodierung unterscheidet man verschiedene Anaglyphenvarianten [17, 62].

Anaglyphen sind mit fast allen Medien kompatibel, sie funktionieren auf sämtlichen Bildschirm- und Projektortypen, im Druck ebenso wie auf Film und Fotopapier. Einzig für die Fernsehübertragung sind Anaglyphen kaum geeignet, da die geringe Farbbandbreite heutiger Fernsehsysteme die Qualität drastisch mindert. Besonders analoge VHS-Videoaufnahmen leiden unter der nochmals reduzierten Bandbreite. Aber auch digitale Videosysteme sind aufgrund der meist verwendeten Farbunterabtastung nur bedingt für die Aufzeichnung von Anaglyphen geeignet.

Die Qualität des Anaglyphenverfahrens wird individuell unterschiedlich beurteilt. Manche Menschen empfinden die farblich unterschiedliche Wahrnehmung der beiden Perspektiven als unangenehm. Es ist sogar möglich, dass das Bild des schwächeren Auges zu Gunsten des dominanten Auges ausgeblendet wird und der 3D-Effekt verschwindet, obwohl die betroffene Person reale Szenen problemlos stereoskopisch sehen kann.

### 4.4.1 Echte Anaglyphen

Echte Anaglyphen (True Anaglyphs, Pure Anaglyphs) verwenden sowohl für linke als auch rechte Filterfolie eine der Primärfarben. Übliche Kombinationen sind Rot-Blau und Rot-Grün, wobei erstere bei weitem überwiegt. Es hat sich als De-facto-Standard etabliert, den roten Filter für das linke Auge zu verwenden. Aus Sicht der Kanaltrennung ist der Variante Rot-Blau gegenüber Rot-Grün der Vorzug zu geben, da die Wellenlängendifferenz von Rot und Blau größer ist als jene von Rot und Grün. Ein größerer Abstand der beiden Hauptwellenlängen im Spektrum erfordert eine weniger steilflankige Filtercharakteristik bzw. hat bei gleicher Charakteristik einen geringeren Störbildanteil zur Folge [83].

Zur Erzeugung eines Anaglyphenbildes gehen wir davon aus, dass linkes und rechtes Halbbild als gleich große, digitale Rasterbilder vorliegen. Die Farbe jedes Bildpunktes (Pixel) wird durch Angabe seines Rot-, Blau- und Grünanteils beschrieben, die üblicherweise mit je 8 Bit quantisiert werden. Für die folgenden Formeln ist es unerheblich, ob ein Wertebereich von 0 bis 255 oder ein normierter Wertebereich im Intervall 0 bis 1 zugrunde gelegt wird.

Echte Anaglyphenbilder können keine Farben reproduzieren, darum werden beide Halbbilder zuerst in ein Graustufenbild transformiert. Da das Auge für die drei Primärfarben unterschiedlich empfindlich ist, berechnet sich der Helligkeitswert nicht als Mittelwert der Primärfarbanteile, sondern wird gewichtet. Üblicherweise kommt folgende Formel zur Anwendung:

$$y = 0,299 \cdot r + 0,587 \cdot g + 0,114 \cdot b \quad (4.2)$$

Im Falle eines Rot-Blau-Anaglyphenbildes wird für die Rot-Komponente jedes Pixels der Helligkeitswert des linken Halbbildes (Index 1) und für die Blau-Komponente der Helligkeitswert des rechten Halbbildes (Index 2) gewählt. Die Farbkomponenten des Anaglyphenbildes erhalten den Index a.

$$\begin{pmatrix} r_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} y_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ y_2 \end{pmatrix} \quad (4.3)$$

Unter Verwendung von Beziehung (4.2) können wir die Berechnungsvorschrift für ein echtes Rot-Blau-Anaglyphenbild in Matrizenform anschreiben.

$$\begin{pmatrix} r_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0,299 & 0,587 & 0,114 \end{pmatrix} \cdot \begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} \quad (4.4)$$

### 4.4.2 Graue Anaglyphen

Graue Anaglyphen (Gray Anaglyphs) unterscheiden sich von echten Anaglyphen durch die Verwendung aller drei Primärfarben. Für das rechte Halbbild werden grüner und blauer Kanal gemeinsam verwendet und mit einem Filter der Farbe Cyan betrachtet. Damit erreicht man eine hellere, farbneutralere Wiedergabe, muss aber stärkere Störbilder in Kauf nehmen, besonders wenn die verwendeten Farbfilter der Farbcharakteristik des Wiedergabemediums nicht optimal angepasst sind [83]. Die Berechnung erfolgt gemäß Formel (4.5).

Abbildung 4.5 auf Seite 29 zeigt ein graues Anaglyphenbild, dass mit der beiliegenden Brille betrachtet werden kann. Alternativ zur Farbkombination Rot-Cyan sind unter dem Marken-

namen „ColorCode 3-D“ Gelb-Blau-Brillen erhältlich. In diesem Fall bilden roter und grüner Kanal das linke und der blaue Kanal das rechte Halbbild.

$$\begin{pmatrix} r_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0,299 & 0,587 & 0,114 \\ 0,299 & 0,587 & 0,114 \end{pmatrix} \cdot \begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} \quad (4.5)$$

### 4.4.3 Farbige Anaglyphen

Farbige Anaglyphen (Color Anaglyphs) erlauben im Gegensatz zu den zuvor beschriebenen Varianten eine eingeschränkte Farbwiedergabe. Statt der Helligkeitskomponente werden die Farbkomponenten der Halbbilder direkt für die Anaglyphendarstellung übernommen.

$$\begin{pmatrix} r_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} \quad (4.6)$$

Während grüne und blaue Farbtöne gut reproduziert werden, da beide Kanäle dem gleichen Auge zugeführt werden, erscheinen rote Farben in der Regel unbunt. Farbige Anaglyphen verursachen bei der Darstellung bunter Objekte retinale Rivalität. Beispielsweise führt ein roter Gegenstand zu einer hohen Intensität im roten linken Kanal, während die grüne und die blaue Komponente des rechten Kanals keine Intensität aufweisen und schwarz erscheinen. Die retinale Rivalität lässt sich abschwächen, indem die Farbsättigung beider Halbbilder vor dem Erstellen des Anaglyphenbildes abgeschwächt wird. Da rote Farbtöne kaum wahrgenommen werden können, bietet es sich außerdem an, für den Rotkanal wie bei echten oder grauen Anaglyphenbildern die Helligkeitskomponente zu verwenden, ohne dass sich die Farbwiedergabe dabei signifikant verschlechtert. Diese Variante, vom Autor halbfarbiges Anaglyphenbild (Half Color Anaglyph) genannt, berechnet sich nach folgender Formel:

$$\begin{pmatrix} r_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} \quad (4.7)$$

Die fast vollständige Eliminierung der retinalen Rivalität gelingt, indem die Rot-Komponente des Anaglyphenbildes aus den Grün- und Blau-Anteilen des Halbbildpaares abgeleitet wird. Damit wird eine übereinstimmende Helligkeit im linken und rechten Kanal erzielt. Die Berechnungsvorschrift für die rote Komponente berücksichtigt die unterschiedliche Empfindlichkeit der Augen für Grün und Blau. Nach Anwendung von Formel (4.8) kann eine Aufhellung des Rotkanals durchgeführt werden, um den Helligkeitsverlust durch den Rot-Filter sowie die geringere Sensitivität des Auges für Rot (im Vergleich zu Grün) zu kompensieren. Eine Gammakorrektur mit  $\gamma \approx 1,5$  hat sich bewährt.

$$\begin{pmatrix} r'_a \\ g_a \\ b_a \end{pmatrix} = \begin{pmatrix} 0 & 0,837 & 0,163 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} \quad (4.8)$$

$$r_a = r'_a{}^{\frac{1}{\gamma}} \quad (4.9)$$



Abbildung 4.5: Graues Anaglyphenbild



Abbildung 4.6: Farbiges Anaglyphenbild

## 4.5 Polarisationsstechnik

Die Polarisationsstechnik nutzt die Polarisations Eigenschaften des Lichtes zur Kanaltrennung aus. Es wird zwischen linearer und zirkularer Polarisation unterschieden. Beide Varianten werden für die Stereoprojektion verwendet, wobei die lineare Polarisation bei weitem am häufigsten eingesetzt wird.

Die Wiedergabe erfolgt mit zwei Projektoren, bei denen im Falle der linearen Polarisation um  $90^\circ$  versetzte Polarisationsfilter, auch Polfilter genannt, vor den Objektiven platziert werden. In den allermeisten Fällen wird der V-Anordnung der Vorzug gegeben, d.h., es wird eine Polarisationsrichtung von  $45^\circ$  für die rechte und  $135^\circ$  für die linke Darstellung verwendet. Die beiden Bilder werden auf einer polarisationserhaltenden Leinwand übereinander projiziert und durch eine Brille mit Polfiltern in derselben Anordnung betrachtet.

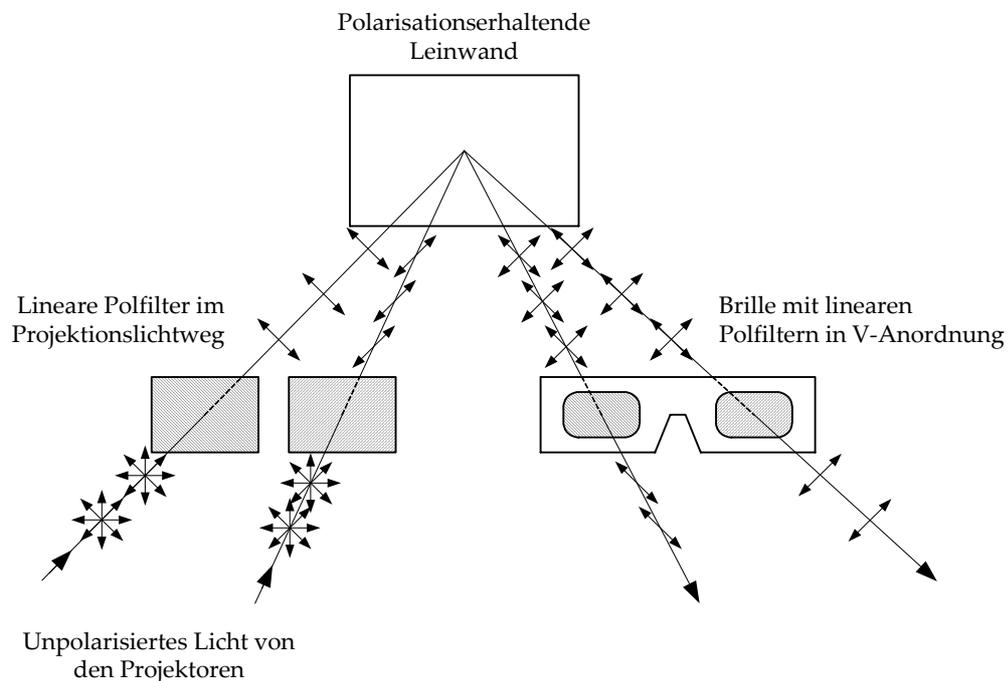


Abbildung 4.7: Prinzip der Polarisationsstechnik

Eine optimale Unterdrückung von Störbildern wird nur erreicht, wenn die Filter vor den Projektoren (Polarisatoren) und in den Brillen (Analysatoren) im exakt gleichen Winkel ausgerichtet sind. Damit ist auch schon die größte Schwäche der linearen Polfiltertechnik angesprochen: Neigt der Zuseher den Kopf, entsteht ein Differenzwinkel zwischen Polarisator und Analysator und der Störbildanteil nimmt stark zu. Dieses Problem wird bei Verwendung von zirkularen Polfiltern (Abbildung 4.8) umgangen. Allerdings zeigen diese nur für eine bestimmte Wellenlänge, die meist in der Mitte des sichtbaren Spektrums gewählt wird, ein optimales Verhalten. An den Rändern des sichtbaren Spektrums muss man wiederum eine schlechtere Kanaltrennung in Kauf nehmen. Da die lineare Technik (zumindest bei exakter Ausrichtung) gegenüber der zirkularen Technik einen geringeren Störbildanteil aufweist, wird sie häufiger verwendet.

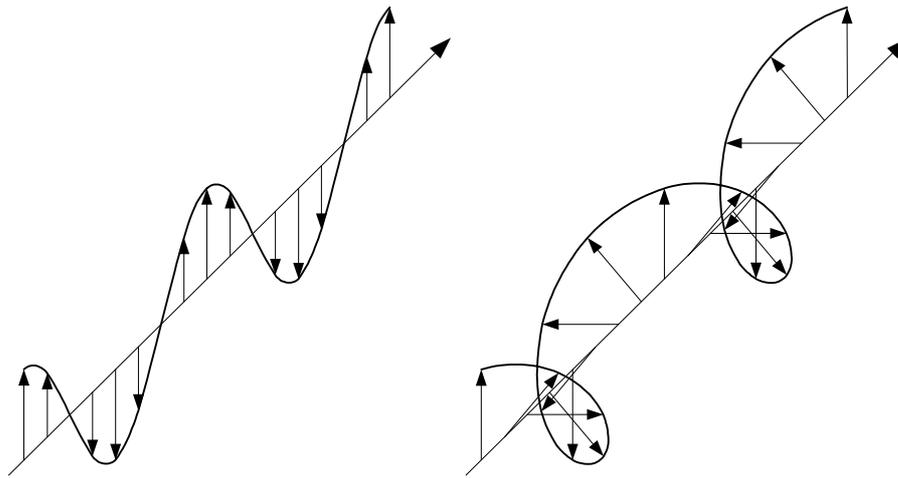


Abbildung 4.8: Linear und zirkular polarisierte Welle

Die tatsächlich erreichbare Qualität hängt stark von den verwendeten Filtern und der Leinwand ab. Meist wird als Projektionsfläche eine metallisierte Oberfläche verwendet. Eine solche polarisationserhaltende Leinwand wird Silberleinwand genannt, obwohl sie in der Regel nicht mit Silber, sondern mit Aluminiumpartikel beschichtet wird. Die Polarisatoren absorbieren mindestens die Hälfte des Lichtes, sodass im Vergleich zu einer zweidimensionalen Präsentation lichtstärkere Projektoren benötigt werden. Polarisationsbrillen können preiswert (unter 0,5 € pro Stück) in Kartonausführung erworben werden. Die Polarisationstechnik ist deshalb das Standardverfahren für hochqualitative Projektionen für ein großes Publikum.

Nicht jeder Projektortyp ist für die Polarisationstechnik gleichermaßen gut geeignet. LCD-Projektoren emittieren aufgrund der verwendeten LCD-Panels vorpolarisiertes Licht, DLP-Projektoren mit drei Spiegelchips verwenden unter Umständen ebenfalls polarisierende Strahlenteiler. Die externen Polfilter führen in der Folge zu einem starken Helligkeitsverlust oder zu einer einseitigen, fast vollständigen Auslöschung des Lichts. Es gibt allerdings speziell angepasste Projektoren, die bereits die benötigten Filter für das jeweilige Halbbild enthalten [5].

Die Grün-Komponente von LCD-Projektoren ist in der Regel vertikal vorpolarisiert, Rot- und Blaukomponente horizontal. Mit Hilfe von optischen Filterfolien kann die Polarisation um  $45^\circ$  in beide Richtungen gedreht werden und die Grünkomponenten von linkem und rechtem Halbbild können in Software ausgetauscht werden. Dank der Vorpolarisierung des Lichts sind keine Polarisatoren notwendig, sodass keine Helligkeit verloren geht. Allerdings schwankt der Grad der Vorpolarisierung zwischen verschiedenen Projektormodellen erheblich und lässt sich durch Filter nicht mehr erhöhen. Deshalb können nicht mit allen Projektoren gute Ergebnisse erzielt werden. Diese Methode für LCD-Projektoren ist unter dem Namen StereoBright bekannt.

Die Polarisationstechnik ist auch mit Röhrenbildschirmen kompatibel, wenn vor der Bildschirmoberfläche ein Flüssigkristallpanel befestigt wird, das synchron zur Bildwiederholrate zwischen zwei Polarisationszuständen wechselt. Am Bildschirm werden linke und rechte Perspektive mit derselben Frequenz gewechselt, genauso wie dies bei der Shuttertechnik der Fall ist. Erfolgen die Wechsel ausreichend schnell (mindestens 100 Hz Bildwiederholrate) ergibt sich eine flimmerfreie Darstellung. Aufgrund der erforderlichen synchronen Ansteuerung fällt dieses Verfahren in die Kategorie der aktiven Systeme. Da ein Bildschirm großen-

bedingt nur von wenigen Leuten gleichzeitig betrachtet werden kann, kompensiert der Kostenvorteil von Polarisationsbrillen den Preis des notwendigen LC-Panels nicht, sodass Shutterbrillen die billigere Lösung darstellen. Vorteile ergeben sich allerdings bei öffentlich zugänglichen Installationen, wo mit häufigen Beschädigungen von Shutterbrillen zu rechnen wäre.

## 4.6 KMQ

KMQ steht für die Initialen von Dr. Christoph Koschnitzke, Reiner Mehnert und Dr. Peter Quick, die das Verfahren Anfang der 80er Jahre an der Universität Hohenheim entwickelten. Linkes und rechtes Bild werden übereinander angeordnet (rechtes Halbbild oben, linkes Halbbild unten) und mit einer Prismenbrille betrachtet. Das KMQ-Verfahren eignet sich für alle gängigen Medien. Allerdings wird die Darstellungsfläche von Bildschirmen nur mit Panoramaaufnahmen gut ausgenutzt, bei Aufnahmen im videotypischen 4:3- oder 16:9-Format bleiben links und rechts Teile der Bildfläche ungenutzt. Dieses Verfahren hat bis jetzt keine nennenswerte Bedeutung erlangt, wohl auch deshalb, weil der Betrachtungsabstand der Brille angepasst werden muss.

## 4.7 Shutterbrillen

Bildröhren arbeiten mit einem Elektronenstrahl (genau genommen sind es drei Elektronenstrahlen, je einer pro Grundfarbe), der die Phosphorbeschichtung der Bildröhre zum Leuchten bringt. Der Elektronenstrahl wird von der magnetischen Ablenkeinheit zeilenweise von links oben nach rechts unten geführt. Horizontale Synchronimpulse beenden eine Bildzeile und bewirken den Rücklauf des Elektronenstrahls nach links zum Anfang der nächsten Zeile. Am Ende der letzten Zeile folgt ein vertikaler Synchronimpuls, der den Rücklauf in die linke, obere Ecke einleitet [39].

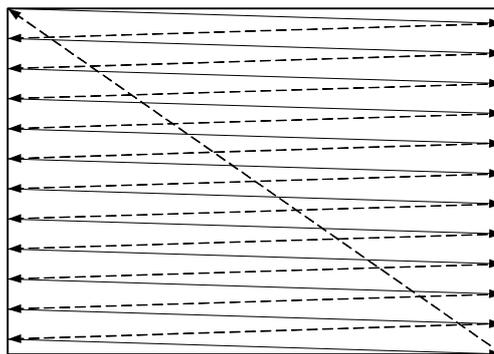


Abbildung 4.9: Progressiver Bildaufbau einer Bildröhre

Shutterbrillen enthalten je Auge ein opto-elektronisches Verschlusselement, das zwischen durchsichtigem und undurchsichtigem Zustand umgeschaltet werden kann. In der Praxis kommen Flüssigkristall-Shutter zum Einsatz, die im Ruhezustand durchsichtig sind und

durch Anlegen einer Spannung undurchsichtig werden. Am Bildschirm werden abwechselnd das linke und rechte Bild dargestellt und der Shutter des jeweiligen Auges für die Dauer des Einzelbildes geöffnet. Die Brillen werden von einem Controller gesteuert, der in das Videosignal eingeschleift wird und die Shutter im Takt der vertikalen Synchronimpulse gegengleich öffnet und schließt.

Shutterbrillen stellen die kostengünstigste Möglichkeit zur hochqualitativen Wiedergabe von stereoskopischen Inhalten auf Röhrenbildschirmen dar. Sie sind ab 50 € erhältlich, die günstigsten kabellosen Brillen mit Infrarotsteuerung kosten knapp unter 100 €. Die Shuttertechnik eignet sich auch für die Projektion, sofern ein Röhrenprojektor verwendet wird.



Abbildung 4.10: Shutterbrille

Für eine flimmerfreie Darstellung ist eine Bildwiederholfrequenz von mindestens 100 Hz notwendig, damit 50 Hz pro Auge erreicht werden. Für eine ruckelfreie Videowiedergabe sollte die Bildwiederholfrequenz des Monitors ein Vielfaches der Bildwiederholrate des Videos betragen. Videos gemäß der PAL-Norm werden am besten mit 100 Hz ( $4 \times 25$  Hz) betrachtet, Videos gemäß der NTSC-Norm mit 120 Hz ( $4 \times 30$  Hz). Höhere Frequenzen verschlechtern die Qualität, weil das Nachleuchten des Bildschirmphosphors zu einem erhöhten Störbildanteil führt. Besonders der grüne Phosphor zeichnet sich durch eine lange Nachleuchtdauer aus, weshalb bei Röhrenprojektoren für die Stereowiedergabe eine grüne Bildröhre mit Short-Decay-Phosphor zum Einsatz kommen sollte. Die Schaltgeschwindigkeit von LCD-Displays ist für die Verwendung von Shutterbrillen nicht ausreichend. Selbst wenn das Display ein Eingangssignal mit 100 Hz akzeptiert, reagieren die Flüssigkristalle nicht rasch genug auf den Wechsel zwischen linkem und rechtem Halbbild, starke Störbilder sind die Folge.

Je nach Prinzip der abwechselnden Bildgenerierung unterscheidet man vier Verfahren: Pageflipping, Interlaced-Darstellung, Lineblanking und Syncdoubling.

### 4.7.1 Pageflipping

Pageflipping ist die naheliegendste Methode: Die 3D-Software stellt abwechselnd linkes und rechtes Bild dar. Die Schwierigkeit besteht darin, dass der Bildinhalt mit jedem vertikalen Synchronimpuls gewechselt werden muss, auch wenn der Prozessor zu diesem Zeitpunkt mit anderen Berechnungen ausgelastet ist. Daher ist es für zuverlässige Lösungen unumgänglich, dass der Bildwechsel vom Treiber oder der Grafikhardware durchgeführt wird.

Professionelle OpenGL-Treiber bieten Unterstützung für stereoskopische Vierfachpufferung (Quadbuffered OpenGL). Zwei der Puffer (Frontbuffer) werden für das aktuelle Bildpaar benötigt und abwechselnd dargestellt, während das neue Bildpaar in zwei weiteren Puffern (Backbuffer) aufgebaut wird. Erst wenn die Backbuffer vollständig gefüllt sind, werden Front- und Backbuffer ausgetauscht. Leider wird die stereoskopische Darstellung nur von teuren Profikarten unterstützt, günstige Grafikkarten für den Spielbereich werden ohne stereotaugliche Treiber geliefert.

Die Firma NVIDIA bietet für ihre Grafikkarten eine Treibererweiterung an, die eine stereoskopische Darstellung von allen ursprünglich monoskopischen OpenGL- und Direct3D-Programmen ermöglicht. Dazu modifiziert der Treiber den Betrachterstandpunkt der Szenenbeschreibung, erzeugt zwei Halbbilder aus leicht unterschiedlichen Perspektiven und stellt sie im Pageflipping- oder Anaglyphenverfahren dar. Der Treiber bietet außerdem eine Programmierschnittstelle, um vorgefertigte Halbbilder unter Direct3D zur Anzeige übergeben zu können [53].

Wird die Brille ausschließlich durch die vertikalen Synchronimpulse gesteuert, kann der Controller nicht zwischen linkem und rechtem Bild unterscheiden. Dadurch erhält man mit einer Wahrscheinlichkeit von 50% ein pseudostereoskopisches Bild. Zwar bieten die meisten Controller einen Schalter zum Wechsel von links und rechts, trotzdem ist es natürlich vorteilhaft, die Brille von vornherein richtig anzusteuern. Professionelle Grafikkarten bieten dazu einen eigenen Anschluss für die Shutterbrille, der den Controller überflüssig macht. Ansonsten wird am DDC-Pin des Monitorausgangs signalisiert, ob gerade das linke oder rechte Bild übertragen wird.

Sind für eine Grafikkarte keine Stereotreiber verfügbar, muss auf eine der folgenden Methoden ausgewichen werden, um stereoskopische Bilder oder Videos anzuzeigen.

## 4.7.2 Interlaced-Darstellung

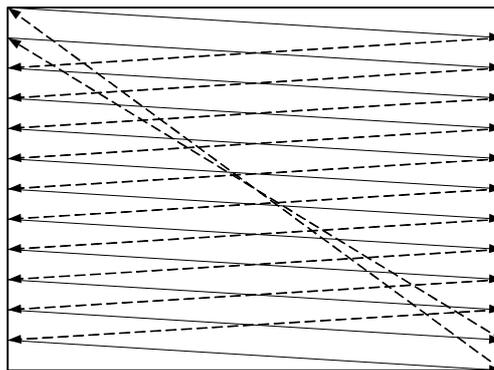


Abbildung 4.11: Bildaufbau mit Zeilensprung

Man spricht vom Interlaced-Modus, wenn ein Monitor im Zeilensprungverfahren betrieben wird. Das Bild wird in zwei Durchgängen übertragen und dargestellt, zuerst alle ungeraden und im Anschluss daran alle geraden Bildzeilen. Jeder Durchgang wird von einem vertikalen Synchronimpuls abgeschlossen. Es bietet sich an, die geraden Bildzeilen für das linke Halbbild und die ungeraden Bildzeilen für das rechte Halbbild (oder umgekehrt) zu verwenden. Durch das Zeilensprungverfahren werden die beiden Halbbilder automatisch sequenziell übertra-

gen, ohne dass ein Programm oder Treiber für das korrekte Timing des Bildwechsels sorgen muss. Die 3D-Wiedergabesoftware verschachtelt die Halbbilder zeilenweise und kopiert sie anschließend mit gewöhnlichen Grafikfunktionen in den Bildspeicher.

Gegenüber dem Pageflipping-Verfahren reduziert sich die vertikale Auflösung um die Hälfte. Zusätzlich entstehen schwarze Abstände zwischen den Bildzeilen, die nur bei ausreichend hohen Auflösungen nicht störend auffallen. Wird ein Monitor im Interlaced-Modus betrieben, wirkt die Darstellung unruhig – es kommt zu Zeilenflimmern. Aus diesem Grund wird dieser Modus im normalen Computerbetrieb heutzutage nicht mehr verwendet. Da im Stereobetrieb jede zweite Bildzeile nicht gesehen werden kann, wird die Betrachtung des zweidimensionalen Desktops erschwert. Folglich sollte die 3D-Wiedergabe nur im Vollbildmodus erfolgen. Moderne Grafikkarten lassen sich entweder nur über spezielle Tools oder erst gar nicht in die Interlaced-Betriebsart versetzen. Selbst wenn sich der Modus aktivieren lässt, sind manche Hardwarebeschleunigungsfunktionen nicht funktionsfähig, sodass die Interlaced-Darstellung auch im Bereich der Stereoskopie immer mehr an Bedeutung verliert.

### 4.7.3 Lineblanking

Das Lineblanking-Verfahren bietet sich an, wenn die Grafikkarte die Interlaced-Betriebsart nicht unterstützt. Die Halbbilder werden auch hier zeilenweise verschachtelt dargestellt, aber als Vollbilder übertragen. Erst der Shutterbrillen-Controller erzeugt die benötigte Abfolge von linkem und rechtem Halbbild, indem er abwechselnd die ungeraden und geraden Zeilen ausblendet, zum Beispiel im ersten Vollbild alle ungeraden Zeilen, im nächsten alle geraden Zeilen und so weiter. Die Brillen werden wie bei allen anderen Varianten synchron zum Bildwechsel angesteuert.

Als Nachteile sind zu nennen, dass der Controller zur Modifikation der Videosignale einen wesentlich höheren Hardwareaufwand erfordert. Das Ausblenden jeder zweiten Zeile reduziert außerdem die Helligkeit um die Hälfte. Da die Shutter ebenfalls 50% der Zeit geschlossen sind und selbst im offenen Zustand ungefähr 65% des Lichts absorbieren bleiben insgesamt nur mehr ca. 9% der ursprünglichen Bildschirmhelligkeit übrig. In den anderen Shutterbrillen-Betriebsmodi werden immerhin 18% erreicht [5].

### 4.7.4 Syncdoubling

Syncdoubling erfordert ebenso wie Lineblanking eine Modifikation des Monitorsignals durch den Controller. Allerdings müssen nicht die hochfrequenten, störungsempfindlichen Bildsignale, sondern nur das vertikale Synchronsignal modifiziert werden. Nach der Übertragung der Hälfte der Bildzeilen wird vom Controller ein vertikaler Synchronimpuls eingefügt, der bewirkt, dass der Monitor die verbleibende Hälfte der Zeilen als eigenes Bild darstellt. Softwareseitig werden die Halbbilder, getrennt durch einen schwarzen Bereich, übereinander mit etwas weniger als halber Höhe dargestellt. Der schwarze Trennbereich ist notwendig, da der Elektronenstrahl eine gewisse Zeit für den Rücklauf in die linke, obere Ecke benötigt. Der Einfügepunkt des Synchronimpulses entspricht übrigens nicht der vertikalen Mitte des sichtbaren Bildes, da auch im konventionellen Betrieb der Zeilentakt während dem Strahlrücklauf weiterläuft und unsichtbare Zeilen zur Folge hat. Die erforderliche Höhe des Trennbalkens unterscheidet sich je nach Auflösung und Bildwiederholrate. Sie wurde richtig gewählt, wenn die Halbbilder ohne vertikalen Versatz dargestellt werden.

## 4.8 Autostereoskopische Displays

Die Grundidee autostereoskopischer Displays kommt den Anforderungen an ein optimales 3D-Display sehr nahe: Autostereoskopische Displays können ohne Brille betrachtet werden und bieten eine uneingeschränkte Farbwiedergabe. Allerdings schränken sie die zulässige Betrachtungsposition auf einen relativ kleinen Bereich ein – oft sind nur wenige Zentimeter Abweichung von der Idealposition erlaubt – und reduzieren die horizontale Auflösung gegenüber der vertikalen um mindestens die Hälfte. Fortgeschrittene Entwicklungen bieten durch Headtracking einen erweiterten Betrachtungsbereich, erlauben aber trotzdem nur für eine einzige Person eine optimale Wiedergabe. Andere Ansätze bieten durch die Verwendung von mehr als zwei Perspektiven eine größere Bewegungsfreiheit, haben aber aufgrund des hohen Aufzeichnungsaufwandes für Video-Anwendungen nur geringe Bedeutung [16].

Die Funktionsweise autostereoskopischer Displays soll anhand zweier unterschiedlicher Konstruktionen vorgestellt werden, der Linsenrastermethode und der Parallaxenbarriere.

### 4.8.1 Linsenrastermethode

Linsenrasterdisplays (Lenticular Screens) basieren auf LCD-Displays, die eine exakt konstante Position aller Pixel garantieren. Davor wird eine Linsenrasterfolie montiert, die manchem Leser von den „Wackel- oder Kippbildern“ bekannt sein dürfte, welche je nach Betrachtungsrichtung ein unterschiedliches Motiv zeigen. Anstelle unterschiedlicher Motive werden für eine 3D-Darstellung stereoskopische Halbbilder verwendet und spaltenweise verschachtelt. Bei richtiger Betrachtungsposition sieht das linke Auge nur die vertikalen Pixelreihen des linken Halbbildes und das rechte Auge nur die Reihen des rechten Halbbildes. Abbildung 4.12 illustriert den Sachverhalt.

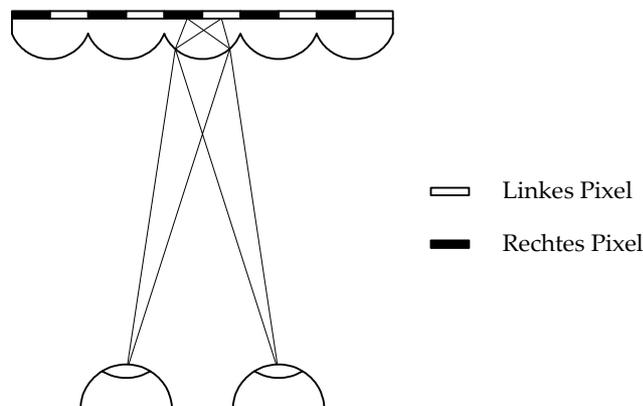


Abbildung 4.12: Linsenrasterbildschirm

### 4.8.2 Parallaxenbarriere

Ein vertikaler Raster mit einem gewissen Abstand von der Displayoberfläche blockiert die Sicht des linken Auges auf Pixelspalten des rechten Halbbildes und umgekehrt (siehe Abbildung 4.13). Im Jahr 2003 stellte Sharp eine andere Variante mit einer Parallaxenbarriere zwischen Hintergrundbeleuchtung und dem eigentlichen LCD-Panel vor, die die Beleuchtung je nach Blickwinkel blockiert oder durchlässt. Die Parallaxenbarriere ist wiederum als mono-

chromes LCD-Panel realisiert. Wird dieses Panel vollständig transparent geschaltet, stellt das Display ein gewöhnliches, zweidimensionales Bild dar.

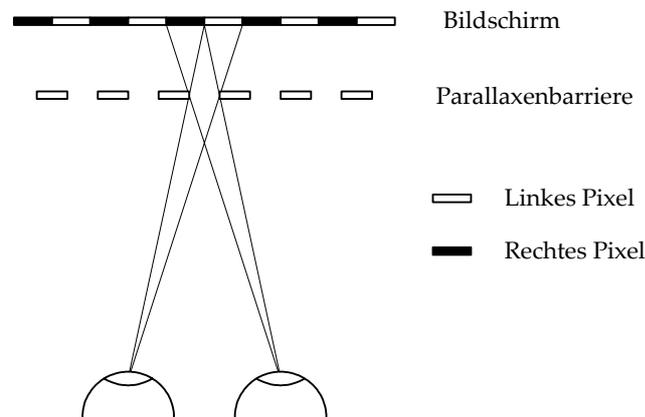


Abbildung 4.13: Display mit Parallaxenbarriere

## 4.9 Head Mounted Displays

Head Mounted Displays (HMDs) werden vom Benutzer aufgesetzt und enthalten pro Auge ein Mikrodisplay, das durch die integrierte Optik vergrößert gesehen wird. Damit das HMD für die 3D-Wiedergabe geeignet ist, müssen den beiden Displays unterschiedliche Bilder zugeführt werden können. Dies geschieht im Idealfall über getrennte Eingänge, kann aber mittels eines Multiplexverfahrens (Pageflipping- oder Interlaced-Format) auch mit nur einem Eingang erfolgen. 3D-taugliche HMDs mit einer Auflösung von 800 x 600 Pixel und einer diagonalen Bildöffnung von ca. 30° sind ab 1000 € erhältlich (Stand Februar 2004), Preise für professionelle Geräte mit besseren technischen Daten liegen um ein Vielfaches darüber.



Abbildung 4.14: Head Mounted Display (5DT HMD 800) [19]

## 4.10 Stereoskope

An dieser Stelle sei die Möglichkeit erwähnt, linkes und rechtes Halbbild auf getrennten Bildschirmen wiederzugeben und eine angenehme Betrachtung über eine Spiegelkonstruktion (ähnlich dem Wheatstone-Stereoskop, siehe Abbildung 2.1, Seite 4) oder eine andere optische Vorrichtung zu ermöglichen. Die Nachteile liegen im aufwendigen, sperrigen Aufbau und der eingeschränkten Bewegungsfreiheit des Betrachters. Die meisten derartigen Konstruktionen sind für einen einzigen Benutzer ausgelegt.

Abbildung 4.15 zeigt einen 3D-Display-Prototypen von Kodak, der die Abbildungen der beiden externen LCD-Displays nach dem Stereoskop-Prinzip zu einem Raumbild vereint. Die stereoskopische Darstellung füllt ein Blickfeld von  $45 \times 36^\circ$  [56].

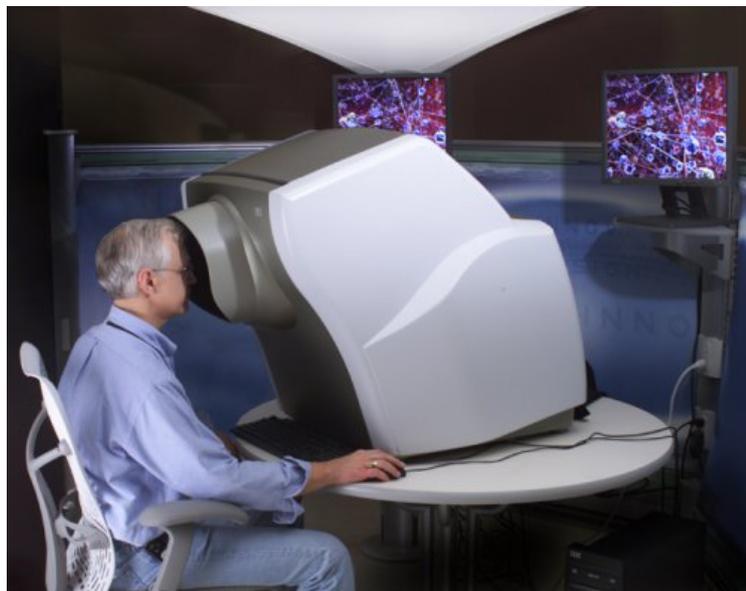


Abbildung 4.15: 3D-Display von Kodak

---

# Formate

---

*Zur Speicherung und Übertragung von stereoskopischen Videos stehen verschiedene Kodierungsformate zur Verfügung. Einerseits können die beiden Streams auf verschiedene Art und Weise zu einem gemeinsamen kombiniert werden, andererseits muss ein geeigneter Codec ausgewählt werden. Kapitel 5 stellt die gebräuchlichsten Varianten vor.*

## 5.1 Separate Streams

Stereoskopische Videos erfordern Arbeitsweisen, die sich von monoskopischen Videos unterscheiden, da zwei Perspektiven verarbeitet werden müssen. Im einfachsten Fall werden linke und rechte Aufnahme in getrennten Dateien gespeichert bzw. separat übertragen.

Im Folgenden werden einige Argumente gegen diese Vorgangsweise geliefert. Prinzipiell können alle aufgeführten Probleme gelöst werden, entweder durch erhöhten Hardwareaufwand oder durch Softwarelösungen. Trotzdem ist es in den meisten Fällen einfacher, eines der stereoskopischen Layouts von Kapitel 5.2 zu verwenden, um beide Videodatenströme (Streams) zu kombinieren.

- Werden gleichzeitig zwei Videodateien auf einen Datenträger geschrieben oder von diesem gelesen, müssen die Schreib- bzw. Leseköpfe abwechselnd auf Sektoren an unterschiedlichen Stellen des Datenträgers zugreifen. Die Transferrate und somit auch die maximal mögliche Datenrate für die Videos sinken.
- Hardwarebeschleunigte Videocodecs arbeiten bei zweifacher Instanzierung möglicherweise weniger effizient oder gar nicht.
- Ist die Prozessorleistung nicht ausreichend, um alle Einzelbilder eines Videos in Echtzeit zu dekodieren, überspringen Decoder Bilder (Frames), um die vorgegebene Abspielgeschwindigkeit einzuhalten. Die Decoder von linkem und rechtem Datenstrom arbeiten unabhängig voneinander und überspringen in der Regel unterschiedliche Bilder, selbst die Anzahl der ausgelassenen Bilder muss nicht übereinstimmen. Die Folge ist eine mangelhafte Synchronisation der Videos, die von nachfolgenden Komponenten in der Verarbeitungskette aufwendig wiederhergestellt werden muss.
- Bei der Nachbearbeitung muss größte Sorgfalt darauf verwendet werden, dass gleiche Einstiegs- und Ausstiegsschnittpunkte sowie gleiche Effekteinstellungen für linkes und rechtes Video verwendet werden. Außerdem erhöht sich der Arbeitsaufwand durch das zweite Schnittprojekt. Jeder Fehler führt zu einer fehlerhaften Synchronisation oder zu anderen unerwünschten Unterschieden zwischen den beiden Perspektiven.

- Zwei Dateien sind umständlicher zu handhaben als eine. Beim Betrachten des Videos müssen zwei Dateien geöffnet werden, wobei die Gefahr besteht, dass versehentlich anstelle eines Stereopaars zwei Dateien unterschiedlichen Inhalts geladen werden. Dem Anwender stellt sich ferner die Frage, ob die Tonspur in der linken, der rechten oder in beiden Dateien enthalten ist.

Eine Abwandlung dieses Verfahrens verwendet eine einzige Datei, die beide Videostreams in verschachtelter Form enthält. Dazu wird ein Containerformat benötigt, das neben dem üblichen Video- und Audio-Stream einen zusätzlichen Videostream akzeptiert. Dies ist zum Beispiel mit dem Audio-/Video-Interleaved- (\*.avi) und dem Quicktime-Format (\*.mov) möglich, wurde aber bisher nur von der Firma Dynamic Digital Depth [12] basierend auf Quicktime realisiert. Ihr Plugin OpticBOOM konnte sich aber nicht am Markt behaupten. DVDs unterstützen ebenfalls parallele Videospuren (Multiangle), allerdings ist kein Produkt bekannt, in dem diese Funktion zur Kodierung stereoskopischer Videos benutzt worden wäre.

## 5.2 Layouts

In den allermeisten Fällen werden linkes und rechtes Halbbild zu einem einzigen Bild kombiniert und anschließend mit denselben Methoden weiterverarbeitet, die auch für zweidimensionale Inhalte verwendet werden [68]. Die Halbbilder lassen sich auf viele verschiedene Arten zusammenfügen, die gebräuchlichsten werden auf den folgenden Seiten erläutert. Zwar erzeugen einige 3D-Aufnahmesysteme systembedingt Aufnahmen mit anderen Layouts, für den fertigen Film sollte aber ausschließlich das Übereinander-, Nebeneinander- oder Interlaced-Layout gewählt werden. Nur diese gewährleisten eine hohe Kompatibilität mit stereoskopischer Wiedergabesoftware.

### 5.2.1 Nebeneinander-Layout

Beim Nebeneinander-Layout (Side-By-Side Layout) werden linkes und rechtes Halbbild nebeneinander platziert, entweder das linke Bild auf der linken und das rechte Bild auf der rechten Seite oder umgekehrt. Die erste Variante kann ohne Umkodierung mit dem Parallelblick, die zweite Variante mit dem Kreuzblick betrachtet werden. Das Nebeneinander-Layout ist die beste Wahl, wenn beide Halbbilder in voller oder halber horizontaler Auflösung vorliegen. Im zweiten Fall weist das resultierende Bild ein übliches Seitenverhältnis auf.

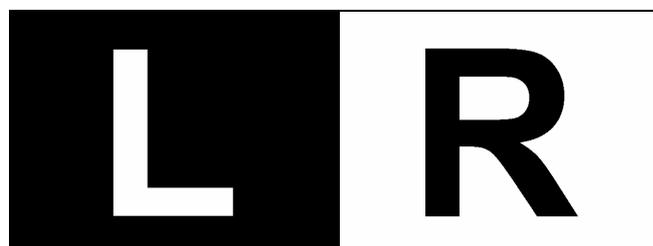


Abbildung 5.1: Nebeneinander-Layout

## 5.2.2 Übereinander-Layout

Das Übereinander-Layout (Over/Under Layout, Top/Bottom Layout) wird weniger oft als das Nebeneinander-Layout verwendet. Es wird hauptsächlich eingesetzt, wenn das Ausgangsmaterial nur mit halber vertikaler Auflösung vorliegt. In diesem Fall ergibt sich, wie in Abbildung 5.2 gezeigt, ein übliches Seitenverhältnis. Bilddaten werden im Speicher meistens zeilenweise abgelegt. Für die Weiterverarbeitung der Daten kann es von Vorteil sein, wenn die Daten des linken und rechten Bildes, wie im Falle des Übereinander-Layouts, in zusammenhängenden Speicherbereichen zu liegen kommen.

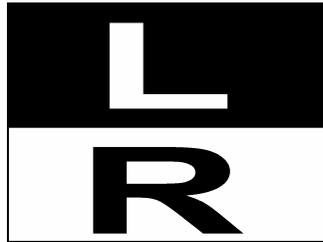


Abbildung 5.2: Übereinander-Layout

## 5.2.3 Interlaced-Layout

Die 3D-Wiedergabe auf Fernsehgeräten basiert auf dem Zeilensprungverfahren. Die Halbbilder der geraden und ungeraden Bildzeilen werden nacheinander dargestellt und die Shutterbrille synchron dazu angesteuert. 3D-Videokassetten und 3D-DVDs enthalten das rechte Halbbild in ungerade Bildzeilen und das linke Halbbild in geraden Bildzeilen (oder umgekehrt), weil mit dieser Kodierung eine direkte Wiedergabe am Fernsehgerät möglich ist. Man spricht dabei vom Fieldsequential-3D-Verfahren.



Abbildung 5.3: Vertikales Interlaced-Layout

Die entsprechende Kodierung kann auch unabhängig vom Fieldsequential-3D-Verfahren verwendet werden und wird (vertikales) Interlaced-Layout genannt. Aufgrund der Zeilenverschachtelung steht pro Halbbild nur die halbe vertikale Auflösung zur Verfügung. Eine Variante mit doppelter Bildhöhe und voller Auflösung ist möglich, wird aber in der Praxis nie verwendet. Das Interlaced-Layout kommt typischerweise dann zum Einsatz, wenn 3D-Videos auf Fernsehgeräten betrachtet oder ohne spezielle 3D-Wiedergabesoftware auf Monitoren in Interlaced-Betriebsart wiedergegeben werden sollen.

Für PC-basierte 3D-Lösungen sind das Nebeneinander- oder Übereinander-Layout dem Interlaced-Layout vorzuziehen, da sie unproblematischer zu verarbeiten sind. Unbedingte

Voraussetzung zur Speicherung von Interlaced-Videos ist ein verlustfreier Codec oder ein Codec mit Halbbildunterstützung, der gerade und ungerade Bildzeilen strikt getrennt komprimiert. Ansonsten beeinflussen sich die Halbbilder aufgrund der Qualitätsverluste gegenseitig und Störbilder sind die Folge.

Soll ein Video mit voller vertikaler Auflösung in das Interlaced-Format konvertiert werden, darf die Reduzierung der Zeilenzahl nicht durch Elimination jeder zweiten Zeile erfolgen. Stattdessen muss die Auflösung durch Mittelwertbildung zweier aufeinander folgender Zeilen oder mittels eines anderen hochwertigen Skalierungsalgorithmus reduziert werden. Wird diese Vorschrift missachtet, sind Aliasing-Artefakte die Folge.

### 5.2.4 Sonderlayouts

Neben den beschriebenen Layouts sind viele weitere Varianten denkbar. Viele optische Strahlenteiler erzeugen Aufnahmen in speziellen Layouts, oft sind die Halbbilder um 90° gedreht nebeneinander angeordnet. Es muss Teil der Nachbearbeitung sein, derartiges Material in ein gebräuchliches Format zu konvertieren.

Auf keinen Fall sollte ein Video direkt im Anaglyphenverfahren kodiert werden. Erstens verursacht die Farbunterabtastung aller verlustbehafteten Codecs einen starken Qualitätsverlust. Zweitens gibt es keine Möglichkeit, aus Anaglyphen die farbigen Halbbilder wieder zu extrahieren, falls später doch ein anderes Wiedergabeverfahren zum Einsatz kommen soll.

Ebenso sollte darauf verzichtet werden, Videos für den Shutterbrilleneinsatz oder für autostereoskopische Displays im vertikalen bzw. horizontalen Interlaced-Format in der Zielauflösung des jeweiligen Displays zu kodieren. Stattdessen sollte das Video in der Quellauflösung in einem Standardlayout gespeichert werden und durch eine geeignete Wiedergabesoftware in Echtzeit umgewandelt werden. Dadurch vermeidet man Performanceprobleme, die unweigerlich auftreten, wenn das Video in Displayauflösung dekomprimiert werden muss. Außerdem wird die Qualität des Interlaced-Layouts genauso wie beim Anaglyphenformat durch Farbunterabtastung stark vermindert, und die hochfrequenten Bildanteile aufgrund der zeilen- bzw. spaltenweise Verschachtelung werden von sämtlichen verlustbehafteten Codecs nicht effizient komprimiert. Codecs mit Zeilensprungunterstützung sind immerhin für vertikales Interlacing geeignet, wie schon in Kapitel 5.2.3 erwähnt.

## 5.3 Codecs

Stereoskopische Videos können mit jedem beliebigen Videocodec komprimiert werden, sofern das Nebeneinander- oder Untereinanderlayout verwendet wird. In der Praxis scheiden jene Codecs aus, die von den verfügbaren Playern bis dato nicht wiedergegeben werden können. Dazu zählen alle Quicktime-Codecs, für die kein Video für Windows-Pendant verfügbar ist, sowie das Format des Streaming-Spezialisten Real Networks [58]. Am häufigsten werden MPEG-1, MPEG-2 und mit MPEG-4 [49] verwandte Codecs wie DivX [15] oder Xvid [85] zur Kodierung von stereoskopischen Videos verwendet. Auch Windows Media [47] kommt gelegentlich zum Einsatz. Kommerzielle 3D-Videos erscheinen ausschließlich auf DVD (von einigen VHS-Kassetten abgesehen) und sind damit in MPEG-2 kodiert.

Linkes und rechtes Halbbild sind einander sehr ähnlich und unterscheiden sich im optimalen Fall nur durch die parallaktische Verschiebung. Ein gewöhnlicher Codec nutzt diese Redundanz nicht aus, weder für eine effizientere Kodierung der Schlüsselbilder (Key- oder Intraframes) noch für die Bewegungskompensation zur Vorhersage des nächsten Bildes. Ein hinsichtlich stereoskopischer Inhalte optimierter Videocodec könnte die Kodiereffizienz durch die folgenden Ansätze erhöhen:

- Aus der linken Perspektive und der parallaktischen Verschiebung für jedes Pixel (Disparity Map) kann die rechte Perspektive fast vollständig vorhergesagt werden. Anstelle des rechten Bildes müssen nur die Parallaxen-Daten und der Vorhersagefehler kodiert werden.
- Die Bewegungsvektoren der linken Perspektive können zur Vorhersage der Parallaxendaten des nächsten Frames verwendet, sodass ebenfalls nur der Vorhersagefehler kodiert werden muss.

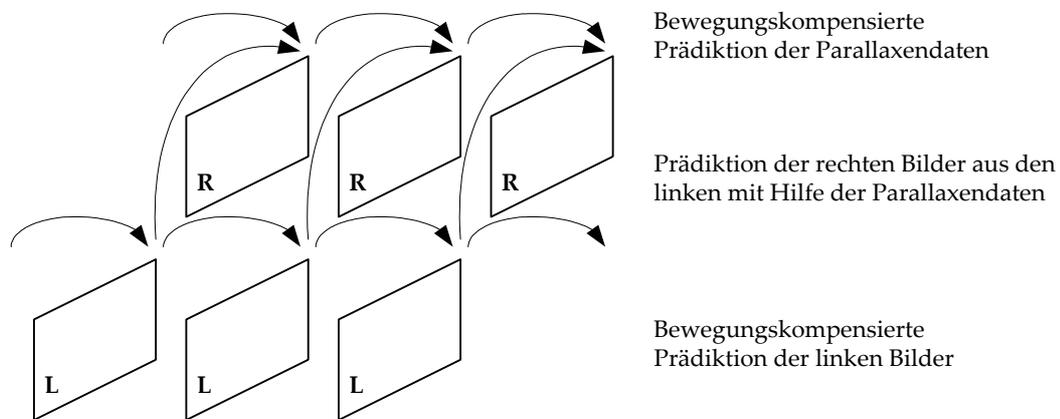


Abbildung 5.4: Prädiktion der zweiten Perspektive in einem stereoskopischen Videocodec

Eine Reihe wissenschaftlicher Arbeiten [54, 72] beschreibt die Kodierung stereoskopischer Videos auf Basis der verschiedenen MPEG-Varianten, jedoch ist aus diesen bisher keine frei verfügbare Implementation entstanden. Meist wird vorgeschlagen, die stereoskopischen Zusatzinformationen unter Ausnutzung der *Temporal Scalability* kompatibel zu monoskopischen Decodern in den Datenstrom einzubetten.

## 5.4 3D-DVDs

Obwohl das Angebot an 3D-Filmen derzeit sehr gering ist, sind DVDs in verschiedenen stereoskopischen Layouts erschienen, die jeweils einer speziellen Wiedergabemethode angepasst sind. Die meisten Filme sind außerdem nur in der amerikanischen NTSC-Norm erhältlich und weisen zum Teil erhebliche technische Mängel auf, die auf Unkenntnis im Umgang mit stereoskopischen Inhalten zurückgeführt werden können.

### 5.4.1 Fieldsequential DVDs

Die meisten 3D-DVDs werden im Interlaced-Layout angeboten, weil damit die direkte Wiedergabe am Fernseher möglich wird. Als einzige zusätzliche Hardware ist ein Controller mit passender Shutterbrille erforderlich, der in das Videosignal eingeschleift wird. Das Field-sequentialverfahren bzw. das Interlaced-Layout wurden in den Kapiteln 2.2.1 und 5.2.3 detailliert beschrieben.

### 5.4.2 TriDVD

Das TriDVD-Format wurde 1997 von den Firmen Neotek [50] und 3DTV Corp. [3] mit dem Ziel entwickelt, durch eine computerbasierte Wiedergabe eine flimmerfreie Darstellung zu ermöglichen. Das TriDVD-Format war nicht nur für den Einsatz auf DVDs bestimmt, sondern konnte auch für die dateigestützte Wiedergabe verwendet werden. Die Videos wurden im Übereinander-Layout kodiert, wobei zwischen den Halbbildern ein schwarzer Balken eingefügt wurde, um eine direkte Wiedergabe mit dem Syncdoubling-Verfahren zu ermöglichen.

Das TriDVD-System erlangte keinerlei Bedeutung und wurde mittlerweile eingestellt. Die Ursachen für den Misserfolg dürften einerseits im Mangel an attraktiven Filmen zu suchen sein, andererseits war eine Wiedergabe nur am Computer mit Shutterbrillen im Syncdoubling-Modus möglich. Eine Software zur Wiedergabe mit anderen Verfahren wurde nicht angeboten.

### 5.4.3 Sensio 3D

Sensio 3D von TEG Sensorial Technologies [73] ist eine 3D-Wiedergabelösung, in deren Mittelpunkt der Sensio-Prozessor steht. Er nimmt analoge Videosignale von einem DVD-Player entgegen und übergibt sie in aufbereiteter Form per VGA-Ausgang an einen Monitor oder Projektor.

Sensio-DVDs sind im Nebeneinander-Layout codiert, wobei das Bild horizontal gestaucht wird, um der DVD-Auflösung von 720 x 480 Pixel zu entsprechen. Eine PAL-Variante mit 720 x 576 Pixel ist nicht vorgesehen. Am VGA-Ausgang des Sensio-Prozessors liegt ein Signal mit einer Auflösung von 640 x 480 Pixel im Pageflipping-Modus an. Es sind Ausgangsfrequenzen von 59,94 Hz und 71,928 Hz möglich. Da sich diese Werte pro Auge allerdings halbieren, ist kein flimmerfreier Betrieb möglich.



Abbildung 5.5: Sensio 3D-Prozessor

Manche Projektoren puffern die Bilddaten intern und verursachen damit eine Verzögerung zwischen der Übertragung und Darstellung der Bilddaten, sodass die Shutter der 3D-Brillen zu falschen Zeitpunkten geöffnet bzw. geschlossen werden. Der Sensio-Prozessor bietet die Möglichkeit zur entsprechenden Kompensation des Shutterbrillen-Timings.

Der Sensio-Prozessor kam im September 2003 zu einem Preis von fast 3000 Dollar auf den Markt. Es wird sich in den nächsten Monaten herausstellen, ob dem Sensio-Format Erfolg beschieden ist, der nicht zuletzt vom Angebot an passenden DVDs abhängen wird. Diese Medien wären jedenfalls auch eine gute Ausgangsbasis für die computergestützte Darstellung mit anderen 3D-Wiedergabemethoden.

#### **5.4.4 Anaglyphen-DVDs**

Anaglyphen-DVDs enthalten den 3D-Film aufbereitet in einem der Anaglyphenverfahren. Da DVDs mit einer 4:2:0-Farbunterabtastung kodiert sind, reduziert sich die Farbauflösung auf ein Viertel, sodass die Bildschärfe stark abnimmt. Überdies besteht keine Möglichkeit, die ursprünglichen Halbbilder zu rekonstruieren. Dadurch ist es unmöglich, Anaglyphen-DVDs als Grundlage für andere Wiedergabemethoden zu verwenden.

Aus den genannten Gründen stellen Anaglyphen-DVDs keine ernsthafte Alternative zu den zuvor beschriebenen Formaten dar.

#### **5.4.5 ColorCode 3-D**

Die dänische Firma ColorCode 3-D Center ApS [10] vertreibt 3D-DVDs im ColorCode 3-D-Verfahren. Bei ColorCode 3-D handelt es sich um eine patentierte Variante (US-Patent 6.687.003) des Anaglyphenverfahrens, die mit Gelb-Blau-Brillen betrachtet wird. Durch eine proprietäre Aufbereitung der Halbbilder soll eine bessere Farbwiedergabe gegenüber den konventionellen Anaglyphenvarianten erzielt werden. Ansonsten weisen ColorCode 3-D DVDs dieselben Schwächen wie konventionelle Anaglyphen-DVDs auf. Allerdings enthalten alle zuletzt erschienenen Scheiben zumindest eine zusätzliche Fassung im Fieldsequential-Verfahren.

---

# Aufnahme

---

*Dieses Kapitel diskutiert unterschiedliche Varianten zur Aufnahme stereoskopischer Videos. Neben der Beschreibung des klassischen Aufnahmeverfahrens mit zwei Kameras werden alternative Lösungen mit nur einer Kamera aufgezeigt.*

## 6.1 Aufnahme mit zwei Kameras

Seit den Anfängen des 3D-Films Anfang des 20. Jahrhunderts wurde von der Möglichkeit Gebrauch gemacht, linke und rechte Perspektive mit zwei getrennten Kameras aufzunehmen. Seit damals wurde auch versucht, das systemimmanente Problem dieser Technik zu lösen: den Gleichlauf beider Kameras, Synchronisation genannt. Der Begriff Synchronisation bezieht sich dabei nicht nur auf das Timing – Stereobildpaare müssen zum exakt gleichen Zeitpunkt belichtet werden – sondern auch auf alle anderen Aufnahmeparameter wie Zoom, Fokus, Belichtung, Farbton (Weißabgleich), Kontrast und viele mehr [64]. Wird ein Film als Aufnahmeemedium verwendet, muss bei der Erstellung der Filmkopien ebenfalls darauf geachtet werden, dass keine Farb- oder Helligkeitsabweichungen entstehen.



Abbildung 6.1: Regisseur James Cameron mit digitaler 3D-Kamera  
Dieses Aufnahmesystem wurde für den IMAX-Film „Ghosts of the Abyss“ verwendet.

Im Folgenden wird davon ausgegangen, dass elektronische Bildwandler zum Einsatz kommen und alle Aufnahmen am Computer digital nachbearbeitet werden, oder im Falle einer Echtzeitübertragung zumindest digitale Verarbeitungsschritte durchlaufen. Die Synchronisation gestaltet sich je nach verwendeten Kameras unterschiedlich schwierig. Grundsätzlich ist es empfehlenswert, zwei baugleiche Kameras einzusetzen, wenn möglich sogar aus derselben Herstellungsladung. Je mehr manuelle Einstellmöglichkeiten vorhanden sind, desto besser, denn dann können diese identisch gewählt werden. Da dieses Verfahren bei manueller Durchführung mit hohem Aufwand verbunden ist und Fehler provoziert, ist es von Vorteil, wenn sich die Einstellungen einer Kamera einfach auf die zweite übertragen lassen, beispielsweise indem sie auf einer Speicherkarte abgelegt werden. Durch Differenzen im werksseitigen Abgleich ist es dennoch möglich, dass Unterschiede im Bild auftreten. In diesem Fall muss eine geringfügige Korrektur der Parameterwerte durchgeführt werden.



Abbildung 6.2: Consumer-Kameras und ste-fra CAM-Halterung [6]

Consumer-Kameras bieten in der Regel keine vernünftigen Möglichkeiten, um die Parameter mit vertretbarem Aufwand abzugleichen, sodass sie meist im Automatikmodus betrieben werden. Je nach Aufnahmesituation und Kamerapaarung schwanken die Ergebnisse zwischen perfekt und unbrauchbar. Viele Bildabweichungen lassen sich später gut korrigieren, einschließlich leichter Abweichungen des Zooms, ziehen aber einen hohen Nachbearbeitungsaufwand nach sich. Im Falle einer Echtzeitübertragung sind einige Bildkorrekturen allerdings nur mit sehr hoher Rechenleistung oder zum Teil gar nicht möglich. Das gilt vor allem dann, wenn sich Kameraeinstellung oder Motiv laufend ändern und unerwünschte Abweichungen von der Software selbstständig erkannt werden sollen.

Die zeitliche Synchronisation kann mittels Genlock sichergestellt werden, sofern die Kamera einen entsprechenden Eingang bietet. Genlocking zwingt die Kameras synchron zu einem externen Taktgeber zu arbeiten. Ansonsten kann durch gleichzeitiges Einschalten der Kameras versucht werden, einen möglichst geringen zeitlichen Offset zwischen den Aufnahmezeit-

punkten herzustellen. Ein Messgerät, das die Synchronisationsdifferenz zwischen den beiden analogen Videoausgängen misst, kann zur Kontrolle dienen. Je nach der gewünschten Toleranz muss der Einschaltvorgang mehr oder weniger oft wiederholt werden. Eine bessere Lösung ist das gleichzeitige Einschalten der Kameras über Steuereingänge wie z.B. Sonys LANC-Buchse. Dazu ist eine spezielle Fernbedienung nötig, die an beide Kameras gleichzeitig angeschlossen werden kann. Zwar driften die internen Taktgeber auseinander, aufgrund der hohen Genauigkeit von Quarzoszillatoren allerdings so langsam, dass der Versatz zumindest in den ersten Minuten nicht ins Gewicht fällt.



Abbildung 6.3: ste-fra LANC-Steuerung

Abbildung 6.3 zeigt die ste-fra LANC-Steuerung [6], die neben dem Einschalten der Camcorder auch noch die gleichzeitige Steuerung anderer Funktionen ermöglicht, beispielsweise Zoom und Fokus. Da die Zoomsteuerung nur durch Auf-/Abwärtskommandos und nicht durch Vorgabe eines exakten Werts erfolgt, driften die Zoomeinstellungen mit jedem Kommando weiter auseinander. Sie können aber wieder synchronisiert werden, indem die maximale Weitwinkel- oder Teleposition bis zum Anschlag angefahren wird.

### 6.1.1 Synchronisation auf Frames

Der erste Schritt der Nachbearbeitung ist das zeitrichtige Zusammenfügen der linken und rechten Aufnahmen. Wurden synchronisierte Kameras verwendet, werden jeweils Bildpaare mit gleichem Aufnahmezeitpunkt kombiniert. Im Falle unsynchronisierter Kameras werden Bildpaare mit der geringsten Differenz des Aufnahmezeitpunkts gewählt. Der Synchronisationsfehler kann höchstens die Hälfte der Framedauer betragen, bei 25 Bildern pro Sekunde (PAL) beispielsweise 1/50 Sekunde, bei 30 Bildern pro Sekunde (NTSC) 1/60 Sekunde.

In bewegungsarmen Sequenzen ist es schwierig, die richtige Zuordnung zu finden. Um die Synchronisation zu erleichtern, sollte am Anfang jeder Szene ein markantes Ereignis mit aufgenommen werden, am besten ein Schlag mit einer Klappe, ansonsten eine schnelle Handbewegung. Alternativ kann man sich an den Waveform-Darstellungen der Tonspuren orientieren, sofern die verwendete Software diese unterstützt.

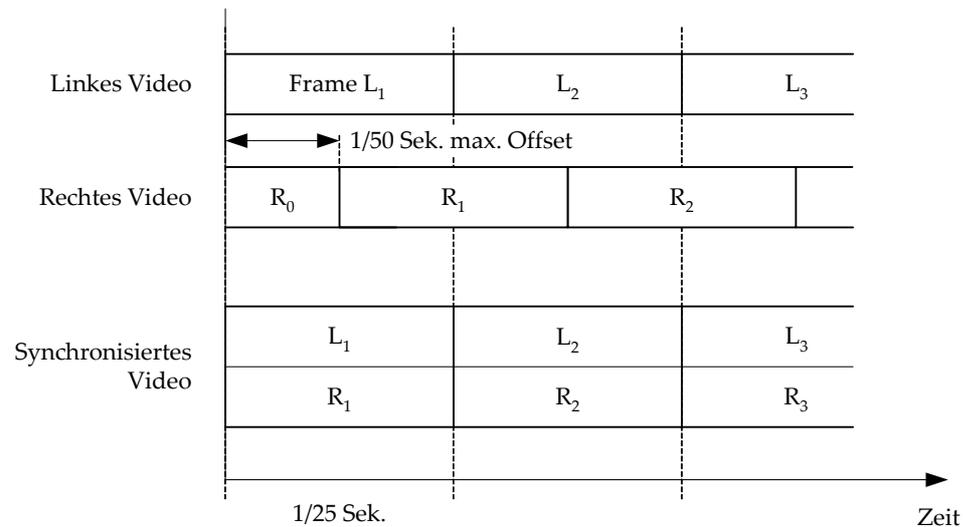


Abbildung 6.4: Synchronisation auf Frame-Ebene

### 6.1.2 Synchronisation auf Fields

Gängige Videokameras arbeiten im Zeilensprungverfahren, sodass vor der Wiedergabe auf einem progressiven Bildschirm ein Deinterlacing durchgeführt werden muss, um Kammartefakte zu vermeiden. Bei zweidimensionalen Aufnahmen ist das Zeilensprungverfahren für die computergestützte Wiedergabe grundsätzlich unerwünscht, weil das Deinterlacing die Bildschärfe im Vergleich zu einer progressiven Aufnahme mit derselben Auflösung reduziert. Im Falle stereoskopischer Aufnahmen von unsynchronisierten Kameras kann aber Nutzen aus dem Zeilensprungverfahren gezogen werden, indem die zwei Halbbilder eines Frames als eigenständige Bilder behandelt werden. Durch die Synchronisation auf Field-Ebene halbiert sich der maximale Offset im Vergleich zur Synchronisation auf Frame-Ebene. Im Deutschen werden Fields üblicherweise als Halbbilder bezeichnet, in dieser Arbeit wird der Begriff in diesem Kontext wegen der Verwechslungsgefahr mit stereoskopischen Halbbildern (linkes und rechtes Halbbild bilden ein Stereobildpaar) aber vermieden. Durch Beibehaltung der synchronisierten Fields als eigenständige Frames erhält man ein stereoskopisches Video mit doppelter Bildwiederholrate (aber unruhigen horizontalen Kanten). Alternativ kann man je zwei Fields zusammenmischen und erhält wieder die ursprüngliche Bildwiederholrate.

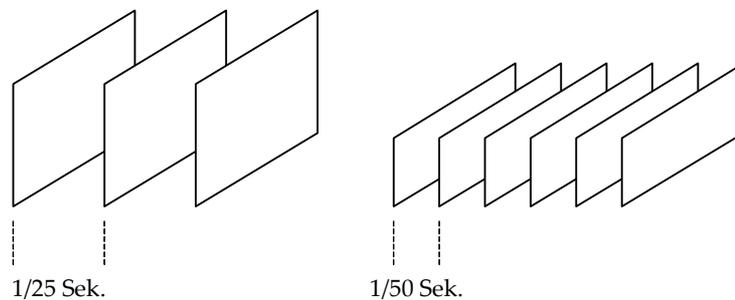


Abbildung 6.5: Progressives versus fieldsequenzielles Video

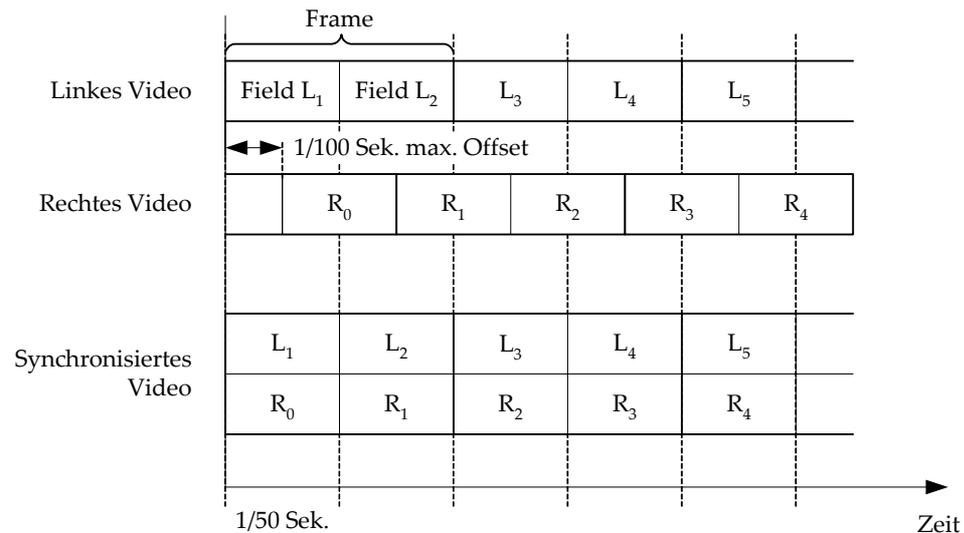


Abbildung 6.6: Synchronisation auf Field-Ebene

### 6.1.3 Hardware-Multiplexing

Mit Hardware-Multiplexing wird jener spezielle Prozess bezeichnet, der zwei Videosignale in Echtzeit zu einem Fieldsequential-3D-Signal zusammenmischt, das mit Shutterbrillen direkt am Fernseher betrachtet werden kann. Hardware-Multiplexing war vor dem Durchbruch computerbasierter Synchronisationsmethoden das Standardverfahren zur 3D-Videoaufnahme.

Voraussetzung sind zwei Kameras, die unbedingt per Genlock synchron betrieben werden müssen. Die analogen Ausgänge beider Kameras werden am Multiplexer angeschlossen, der das stereoskopische Ausgangssignal abwechselnd aus Fields der linken und rechten Kamera bildet.

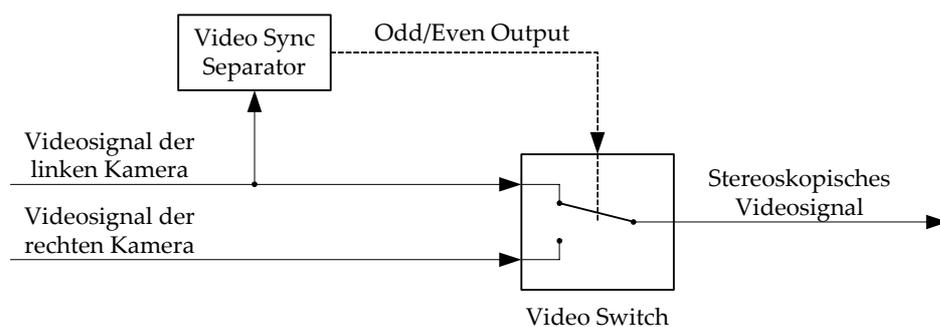
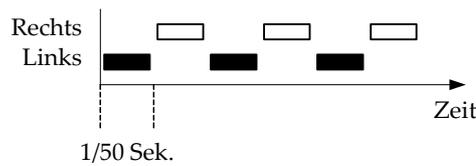


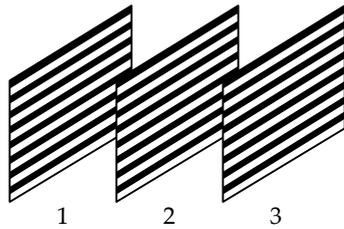
Abbildung 6.7: Blockschaubild eines analogen Videomultiplexers

Die Multiplexerschaltung kann einfach und kostengünstig mit wenigen elektronischen Bauteilen realisiert werden: Ein Video Sync Separator extrahiert aus dem linken oder rechten Videosignal (Composite-Signal oder Y-Komponente bei Y/C-Verbindung) ein digitales Signal mit dem Wert logisch null solange ungerade Zeilen übertragen werden und dem Wert logisch eins wenn gerade Zeilen übertragen werden. Dieses so genannte Odd/Even-Signal steuert einen elektronischen Schalter (Video Switch), der nach jedem Halbbild die jeweils andere Kamera an den Ausgang durchschaltet [63].

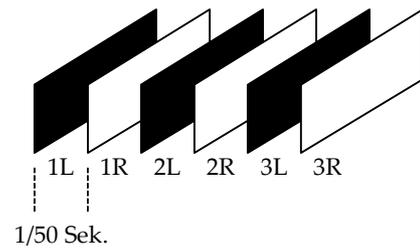
Aufnahmetiming des Hardware Multiplexers



Frames, bestehend aus zeitversetzten Fields

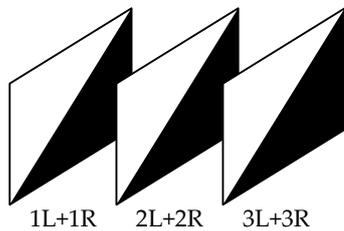


Separierte Fields



Aufbereitete Frames für progressive Wiedergabe

ohne zusätzliche Frames  
1/50 Sek. Fehlsynchronisation



mit zusätzlichen Frames zur Synchronisationskorrektur

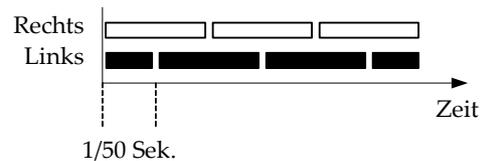
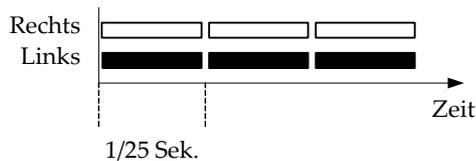
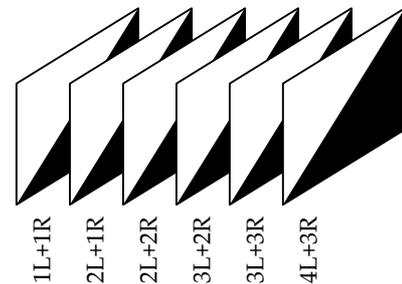


Abbildung 6.8: Aufbereitung fieldsequentieller Aufnahmen für progressive Wiedergabe

Systembedingt weisen aufeinander folgende Halbbilder eine Differenz des Aufnahmezeitpunkts von 1/50 Sekunde (PAL) bzw. 1/60 Sekunde (NTSC) auf. Erfolgt die Wiedergabe auf einem Fernseher, entsprechen die zeitlichen Verhältnisse jenen der Aufnahme. Aufgrund des starken Flimmerns dieser Methode geht der Trend jedoch zu anderen Wiedergabemethoden. Während am Fernseher ein optimales Timing erreicht wird, ergibt sich mit anderen progressiven Methoden die schlechtestmögliche Synchronisation. Fügt man Zwischenbilder ein, kann der zeitliche Versatz aber auch bei progressiver Wiedergabe kompensiert werden (siehe Abbildung 6.8).

## 6.2 Aufnahme mit einer Kamera

Die 3D-Aufnahme mit einer Kamera wird durch Verwendung eines optischen Bild- bzw. Strahlenteilers möglich. Dieser wird vor dem Objektiv befestigt und erzeugt die beiden benötigten Perspektiven. Die Stereobasis der meisten Kameravorsätze kann nicht verändert werden. Anhand des Funktionsprinzips lassen sich zwei grundlegende Typen unterscheiden: passive Bild- und aktive Strahlenteiler. Kameravorsätze sind einfach in der Handhabung, weisen allerdings auch zahlreiche Nachteile auf:

- Pro Halbbild steht bestenfalls nur die halbe Auflösung der Kamera zur Verfügung, während ein Zweikamerasystem die volle Auflösung für beide Halbbilder bietet (von Verlusten durch das Deinterlacing abgesehen).
- Die zusätzlichen Bauteile vermindern die Abbildungsqualität der Optik. Besonders bei Verwendung hoch auflösender Kameras fällt die reduzierte Schärfe ins Gewicht.
- Reflexionen im Kameravorsatz können Bildstörungen verursachen.
- Bei unsymmetrischem Aufbau muss mit Abweichungen zwischen den Halbbildern (z.B. Größenunterschiede) gerechnet werden.
- Bei ungenügender Anpassung des Vorsatzes an die Kameraoptik kann es an den Rändern zu Abschattungen kommen. Besonders häufig tritt das Problem bei Weitwinkelobjektiven auf.

Trotz der genannten Nachteile dominieren Kameravorsätze aus ökonomischen Gründen den Amateurmarkt, während professionelle Aufnahmen bis dato ausschließlich mit zwei Kameras produziert werden.

### 6.2.1 Passive Bildteiler

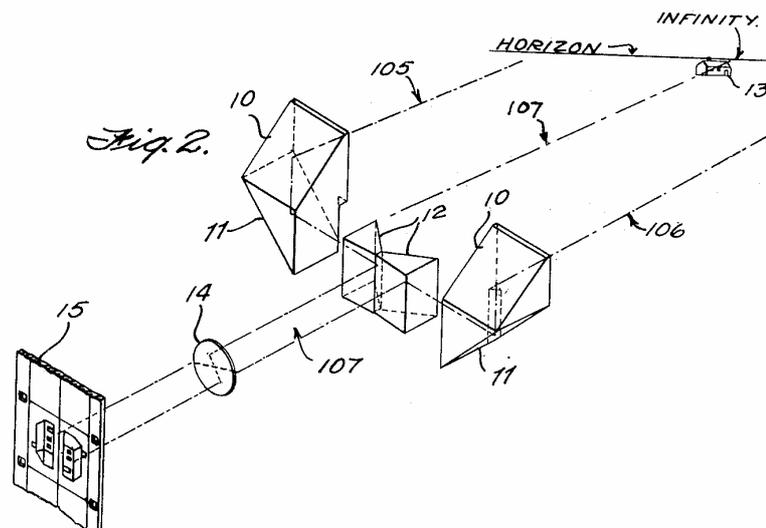


Abbildung 6.9: Passiver Bildteiler  
(aus US-Patent 2.282.947 von Ingvar E. de Cherbinin)

Passive Strahlenteiler vereinen linkes und rechtes Bild kontinuierlich auf optischem Weg. Die Halbbilder werden über- oder nebeneinander angeordnet. Je nach Konstruktion werden sie um  $90^\circ$  gedreht, um das Bildformat der Kamera besser auszunutzen. Viele Strahlenteiler ordnen zwei Halbbilder im Hochformat nebeneinander an, für die Videoaufnahme sind aber nur jene Systeme gut geeignet, die Halbbilder mit einem Seitenverhältnis 4:3 oder breiter erlauben.

Während passive Strahlenteiler für die Stereofotografie gut geeignet sind, muss bei Videoaufnahmen auf Kosten der Auflösung ein Deinterlacing durchgeführt werden (ausgenommen bei progressiven Kameras).

## 6.2.2 Aktive Strahlenteiler

Aktive Strahlenteiler sind ausschließlich für die Verwendung mit Videokameras geeignet. Die beiden Halbbilder werden nicht örtlich angeordnet, sondern der Kamera zeitsequenziell zugeführt. Die Kamera nimmt also abwechselnd die linke und rechte Perspektive auf. Dabei macht man sich den Umstand zunutze, dass ein Videoframe aus zwei zeitlich separierten Fields besteht. Der Strahlenteiler wird auf das vertikale Synchronsignal der Kamera synchronisiert, das aus dem Ausgangssignal der Kamera extrahiert wird, sodass eine Aufnahme im Field-sequential-Format (siehe Kapitel 2.2.1 und 5.4.1) entsteht.



Abbildung 6.10: Nu-View-Adapter

Der Nu-View-Adapter ist der am weitesten verbreitete aktive Strahlenteiler. Er enthält einen prismatischen Strahlenteiler und einen Spiegel, um die zweite Perspektive zu erzeugen (siehe Abbildung 6.11). In beiden Strahlengängen sind opto-elektronische Shutter platziert, die gegengleich geöffnet und geschlossen werden. Der Spiegel kann bewegt werden, um die Konvergenz zu verändern.

Durch die unterschiedliche Länge von linkem und rechtem Strahlengang ergeben sich Größenunterschiede zwischen den beiden Perspektiven, die bei nahen Gegenständen den Raumeindruck stören. Ein weiteres Problem ergibt sich durch den Strahlenteiler, der aufgrund seines Funktionsprinzips als Polfilter wirkt – mit unterschiedlicher Polarisationsrichtung für linke und rechte Perspektive. Dadurch werden reflektierende Oberflächen sowie alle anderen

Objekte, die vorpolarisiertes Licht emittieren, unterschiedlich hell abgebildet. Die Frontscheibe zum Schutz der Optik führt zu weiteren unsymmetrischen Reflexionen.

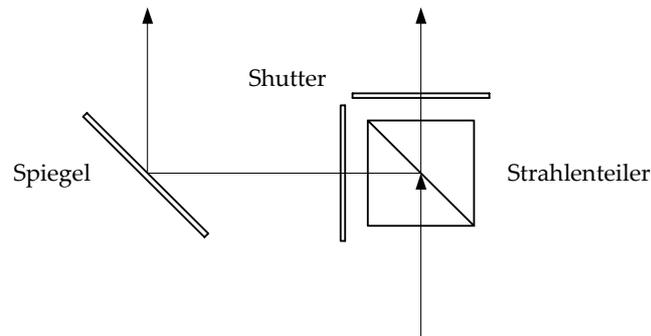


Abbildung 6.11: Nu-View Funktionsprinzip

## 6.3 2D-3D Konvertierung

Die Aufnahme von 3D-Filmen ist im Vergleich zu 2D-Aufnahmen aufwendig. Das Verlangen nach einem Verfahren zur Konvertierung von 2D nach 3D ist deshalb groß. Damit würde mit einem Schlag die riesige Menge an 2D-Material in den Archiven auch für die 3D-Wiedergabe zur Verfügung stehen.

Mit heutiger Technik ist es allerdings nicht möglich, zweidimensionales Material automatisch zu konvertieren. Zwar lassen sich bei manueller Bearbeitung gute Ergebnisse erzielen, doch übersteigt der Aufwand jenen einer echten 3D-Aufnahme. Zahlreiche Firmen bieten Programme für die automatische Umwandlung an, jedoch ist keines dieser Programme in der Lage, einen brauchbaren 3D-Effekt zu erzeugen (abgesehen von einigen speziell ausgewählten Beispieldateien).

Während das menschliche Gehirn in der Lage ist, auch bei einäugigem Sehen Objekte zu segmentieren und anhand monokularer Tiefenhinweise deren räumliche Tiefe abzuschätzen, stellt diese Aufgabe für Computer immer noch ein unlösbares Problem dar. Heutige Algorithmen sind auf Bewegung in der Szene angewiesen, um aus der Geschwindigkeit eines Objekts auf dessen Tiefe schließen zu können. Weiters wird Bewegung benötigt, um jene Bildbereiche für die zweite Perspektive rekonstruieren zu können, die in der zweidimensionalen Ansicht durch Vordergrundelemente verdeckt sind. Nur wenn die verdeckten Bereiche in vorhergehenden oder folgenden Bildern sichtbar sind, kann die zweite Perspektive vollständig erzeugt werden. Ansonsten müssen die betroffenen Stellen retuschiert werden.

Echtzeit-2D-3D-Konverter beschränken sich darauf, die zweite Perspektive durch Verzögerung des zweidimensionalen Videos um ein oder mehrere Bilder zu erzeugen. Ein dreidimensionales Bild entsteht nur bei Kamerafahrten, wobei die Blickrichtung normal zur Bewegungsrichtung stehen muss. Je nach Bewegungsrichtung müssen entweder dem linken oder dem rechten Auge die verzögerten Bilder zugeführt werden. Bewegungen innerhalb der Szene selbst müssen vermieden werden, da ein Objekt bei horizontaler Bewegung durch die Bewegungsparallaxe sonst in falscher Tiefe wahrgenommen wird und bei vertikalen Bewegungen eine unerwünschte vertikale parallaktische Verschiebung entsteht.

---

# DirectShow

---

*Dieses Kapitel gibt eine kurze Einführung in DirectShow. Es behandelt die Grundlagen dieser Multimedia-Architektur soweit sie für das Verständnis von Kapitel 8 und 9 notwendig sind. Am Ende erfolgt ein Blick über den Tellerrand auf Multimedia-Architekturen in anderen Betriebssystemen.*

## 7.1 Konzept

DirectShow ist eine Architektur zur Verarbeitung von Multimediadaten unter Microsoft Windows. Obwohl es hauptsächlich für Video- und Audiodaten zur Anwendung kommt, ist DirectShow auch für alle anderen Arten von Daten offen. DirectShow ist eine Weiterentwicklung von ActiveMovie und seit Version 8 Bestandteil von DirectX [44]. Es ist für alle Windows-Versionen ab Windows 95 bzw. Windows NT 4 verfügbar.

DirectShow verwendet Filtergraphen, um die Multimediadaten zu verarbeiten. Dabei handelt es sich um gerichtete Graphen, an deren Knoten sich DirectShow-Filter befinden. Die Kanten verbinden je einen Ausgangs- mit einem Eingangspin eines Filters. Es ist Voraussetzung für eine Verbindung zweier Filter, dass sich diese auf einen gemeinsamen Medientyp einigen können. Wird der Graph gestartet, beginnen die Daten in Form einzelner Datenpakete, Media-samples genannt, durch den Graphen zu fließen. Filter empfangen Datenpakete an ihren Eingangspins und geben sie in veränderter Form an ihren Ausgängen weiter. Jedes Datenpaket ist mit einem Zeitstempel (Timestamp) versehen, der dessen Präsentationszeitraum festlegt (beispielsweise Beginn- und Endzeitpunkt eines Videoframes). Die Anwendung steuert den Filtergraphen über den Filtergraph-Manager.

Im Standardumfang unterstützt DirectShow eine Vielzahl an Videoformaten sowie Ein- und Ausgabegeräten. Mittels Filtern von Drittherstellern lässt sich der Funktionsumfang beliebig erweitern. So bieten zum Beispiel zahlreiche Codec-Hersteller DirectShow-Filter für ihre Formate an, sodass sämtliche auf DirectShow basierende Wiedergabeprogramme ohne jede Anpassung mit neuen Formaten zurechtkommen. Zum Testen von Filtergraphen steht ein Tool namens GraphEdit zur Verfügung. Es erlaubt das interaktive Erstellen und Ausführen von Graphen.

Abbildung 7.1 zeigt einen Filtergraphen zur Wiedergabe einer MPEG-1 Datei. DirectShow-Filter werden durch Blöcke symbolisiert, die durch Pfeile zu einem gerichteten Graphen verbunden sind. Der *File Writer* liest die MPEG-Datei und übergibt die Daten an den *MPEG-I Stream Splitter*, der sie in einen Video- und Audiodatenstrom aufteilt. Die Videodaten werden vom *MPEG Video Decoder* dekodiert und anschließend vom *Video Renderer* am Bildschirm ausgegeben. Der *MPEG Audio Decoder* übernimmt die Dekodierung des Audiodatenstroms und

ein DirectShow-Filter mit dem Namen *Default DirectSound Device* übergibt sie an die Soundkarte. Alle in diesem Beispiel verwendeten Filter sind im Standardumfang von DirectShow enthalten.

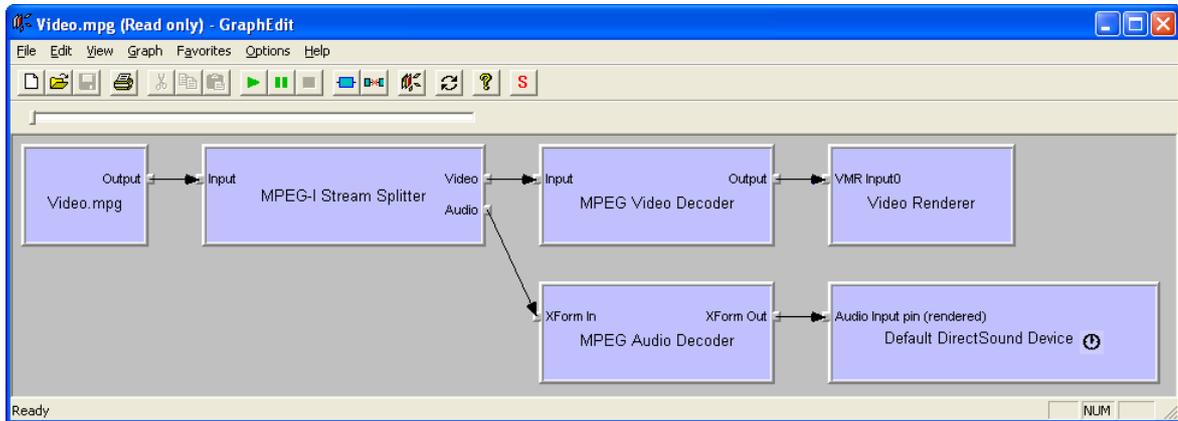


Abbildung 7.1: Typischer Filtergraph zur Wiedergabe einer MPEG-1 Videodatei

DirectShow ist in der Lage, ausgehend von einer Quelldatei einen Filtergraphen zu deren Wiedergabe automatisch zu erzeugen. Dazu sind in der Systemregistrierung zu jedem Filter Informationen zu dessen unterstützten Medientypen sowie eine Prioritätsangabe (Merit) abgelegt. Diese Informationen werden zum Auffinden geeigneter Zwischenfilter genutzt, um den Quellfilter mit den Ausgabefiltern verbinden zu können.

## 7.2 Filtertypen

Nach Anzahl und Art der Pins lassen sich Filter in unterschiedliche Kategorien aufteilen:

- *Quellfilter* besitzen keinen Eingangspins, aber einen oder mehrere Ausgangspins. In diese Kategorie fallen Filter, die Dateien lesen wie z.B. *File Source (Async)* und *File Source (URL)*, sowie alle Arten von Capture-Filtern.
- *Splitter, Parser bzw. Demultiplexer* teilen den Datenstrom von ihrem einzigen Eingangspins auf zwei oder mehrere Ausgänge auf. Die meisten Filter dieser Kategorie dienen zum Aufteilen verschachtelter Video- und Audiodaten und enthalten auch die Implementierung zum Lesen des jeweiligen Containerformats. Typische Vertreter sind der *AVI Splitter*, der *DV Splitter*, der *QuickTime Movie Parser*, der *MPEG-1 Stream Splitter*, der *MPEG-2 Splitter* sowie der *MPEG-2 Demultiplexer*.
- *Multiplexer* vereinen zwei oder mehr Eingangsdatenströme zu einem gemeinsamen Ausgangsdatenstrom. Der *AVI Mux* und der *DV Muxer* fallen in diese Kategorie.
- *Transformationsfilter* besitzen in der Regel einen Eingangspins und einen Ausgangspins. Stimmen Medientyp von Ein- und Ausgang überein, handelt es sich um einen Effektfiter, ansonsten um einen Decoder oder Encoder.

In DirectShow werden Medientypen durch die Struktur `AM_MEDIA_TYPE` beschrieben. Sie setzt sich aus dem `majorType`, dem `subtype` und weiteren Detailinformationen zusammen.

Das Feld `majorType` beschreibt den Hauptmedientyp, den das Feld `subtype` genauer spezifiziert. Unter anderem sind in DirectShow folgende Hauptmedientypen vordefiniert, viele weitere können dem DirectX SDK entnommen werden:

- `MEDIATYPE_Video`
- `MEDIATYPE_Audio`
- `MEDIATYPE_Interleaved`

`MEDIATYPE_Video` wird für alle Arten von Videoformaten, einschließlich komprimierter, verwendet. `MEDIATYPE_Interleaved` beschreibt verschachtelte Video- und Audiodaten, die beispielsweise bei DV-Datenströmen vorkommen.

Exemplarisch seien hier einige Untertypen des Haupttyps `MEDIATYPE_Video` aufgeführt:

- `MEDIASUBTYPE_RGB32`: RGB Farbraum, 32 Bits pro Pixel, 8 Bit ungenutzt
- `MEDIASUBTYPE_RGB24`: RGB Farbraum, 24 Bits pro Pixel
- `MEDIASUBTYPE_YUY2`: YUV 4:2:2 Farbraum
- `MEDIASUBTYPE_UYVY`: YUV 4:2:2 Farbraum
- `MEDIASUBTYPE_YV12`: YUV 4:2:0 Farbraum

Die detaillierte Formatbeschreibung in einer `AM_MEDIA_TYPE` Struktur erfolgt durch einen Verweis auf eine weitere Datenstruktur variablen Typs. Für unkomprimierte Videodaten wird entweder eine Struktur von Typ `VIDEOINFOHEADER` oder `VIDEOINFOHEADER2` benutzt.

## 7.3 Implementierung von Filtern

DirectShow-Filter sind Objekte nach dem *Component Object Model* (COM) und müssen das Interface `IBaseFilter` sowie für jeden Pin das Interface `IPin` implementieren. Abhängig vom Filtertyp müssen weitere Interfaces implementiert werden.

Mediasamples sind ebenfalls COM-Objekte, die das Interface `IMediaSample` implementieren. Der Zugriff auf den eigentlichen Datenbereich erfolgt mit der Methode `GetPointer`. Zum Weiterreichen der Datenpakete gibt es zwei Modi: Push oder Pull. Im ersten Fall ruft der Ausgangspin die Methode `IMemInputPin::Receive` des verbundenen Eingangspins auf, im zweiten Fall verwendet der Eingangspin das `IAsyncReader`-Interface um Daten vom Ausgangspin abzurufen. Der Pull-Modus wird zwischen Quell- und Parser-Filtern bevorzugt, ansonsten dominiert der Push-Modus.

Es besteht die Möglichkeit, dass eine Pinverbindung anstelle der tatsächlichen Mediendaten nur Kontrollinformationen austauscht und der tatsächliche Datenfluss in Form einer Hardwareverbindung stattfindet. In diesem Fall implementieren die Pins statt eines `IMemAllocator`-Interfaces (stellt die Mediasamples zur Verfügung) andere Schnittstellen.

Das DirectX SDK [45] enthält eine Reihe von Basisklassen, die die grundlegende Funktionalität eines Filters implementieren, zum Beispiel das Aushandeln eines Medientyps beim Verbinden zweier Filter. Ein eigener Filter wird durch Ableiten von einer geeigneten Basisklasse und Überschreiben aller relevanten Methoden erstellt.

Die allgemeine Basisklasse `CBaseFilter` unterliegt keinerlei Einschränkungen in Bezug auf Anzahl und Art der Pins. Die gewünschten Pins müssen selbst implementiert und erzeugt

werden, wobei wiederum die Klassen `CBasePin`, `CBaseInputPin` und `CBaseOutputPin` die Programmierarbeit erleichtern.

Für Transformationsfilter mit einem Eingangs- und Ausgangspin verwendet man die Klasse `CTransformFilter`. Im einfachsten Fall beschränkt sich die Implementierung auf die Methoden `CheckInputType`, `CheckTransform`, `DecideBufferSize` und `GetMediaType` um die unterstützten Formate zu spezifizieren, sowie die Methode `Transform`, um die eigentlichen Berechnungen auf den Mediasamples durchzuführen. Die Klasse `CTransformInPlaceFilter` stellt eine Optimierung von `CTransformFilter` dar, falls Eingangs- und Ausgangstyp exakt übereinstimmen und die Berechnungen ohne Kopieren der Daten durchgeführt werden kann. In diesem Fall wird für die ein- und ausgehenden Daten derselbe Speicherblock verwendet. Zur Entwicklung von Rendering-Filtern sind die Klassen `CBaseRenderer` und `CBaseVideoRenderer` vorgesehen.

Damit ein Filter anderen Anwendungen zur Verfügung steht, muss er in die Systemregistrierung aufgenommen werden. Ist der Filter darauf vorbereitet, sich selbst zu registrieren, genügt ein Aufruf von `regsvr32.exe` auf der Kommandozeile mit der Filter-DLL als Parameter. Ein Filter kann mit `regsvr32.exe` installiert werden, falls die DLL-Funktionen `DllRegisterServer` und `DllUnregisterServer` vorhanden sind. Am einfachsten werden sie durch den Aufruf der Funktion `AMovieDllRegisterServer2` implementiert.

## 7.4 Alternative Multimedia-Architekturen

Unter Windows ist DirectShow (abgesehen von den Vorgängern Video für Windows und ActiveMovie) die einzige bedeutende Multimedia-Architektur zur Verarbeitung von Streamingdaten. Auf Apple-Systemen nimmt Quicktime [4] diese Stellung ein. SGI stellt für ihre Systeme die Digital Media Library zur Verfügung, die als offener Standard OpenMP künftig auch für andere Plattformen verfügbar sein wird. Unter Linux bzw. Unix haben die Network-Integrated Multimedia Middleware (NMM) [51] und das GStreamer-Projekt [22] Bedeutung erlangt [79].

Die Network-Integrated Multimedia Middleware wird vom Lehrstuhl für Computergrafik an der Universität des Saarlandes entwickelt und verfolgt im Gegensatz zu DirectShow einen netzwerkzentrierten Ansatz. Ein Node, das Äquivalent zu einem DirectShow-Filter, kann auf jedem beliebigen Rechner eines Netzes laufen. Das NMM-Projekt hat bereits einen Funktionsumfang erreicht, der zum Teil an DirectShow heranreicht und dieses in manchen Bereichen sogar übertrifft.

NMM bietet die Möglichkeit, komplexe Graphen automatisch vervollständigen zu lassen und unterstützt bei Formatverhandlungen zwischen Filtern die Optimierungskriterien Qualität und Kosten. Beispielsweise führt eine höhere Auflösung zu einer besseren Bildqualität und damit zu einer höheren Bewertung der Qualität, erfordert aber auch mehr Rechenleistung und Speicher, was durch das Kosten-Kriterium berücksichtigt wird.

GStreamer ist eine Multimedia-Architektur, die sich stark an DirectShow anlehnt. Gst-Elemente werden zu einer Pipeline zusammengefügt, indem Ausgangs- mit Eingangspads verbunden werden. Das Format einer Verbindung wird dabei durch die Angabe des MIME-Typs (Multipurpose Internet Mail Extension) und weiterer Detailinformationen festgelegt.

---

# Implementierung der Wiedergabe

---

*Zur Darstellung stereoskopischer Videoinhalte wurde der Stereoscopic Player entwickelt. Dieses Kapitel behandelt aufbauend auf dem Vorigen die Grundlagen dieser Wiedergabesoftware und skizziert Funktionsweise und Zusammenspiel der einzelnen Komponenten.*

## 8.1 Einführung in den Stereoscopic Player

In der Vergangenheit mussten stereoskopische Videos für eine bestimmte Wiedergabemethode kodiert werden. So kommt für die Shutterbrillenwiedergabe das Interlaced-Format (für Interlaced-Darstellung oder Lineblanking) oder das Übereinander-Format mit schwarzem Trennbalken (für Syncdoubling) zur Anwendung. Ein Wechsel der Wiedergabemethode erfordert eine Umkodierung und verlustbehaftete Neukomprimierung des Videos. Einige Formate können gar nicht oder nur unter starken Qualitätsverlusten in andere Formate gewandelt werden (beispielsweise Anaglyphen) und schließen Anwender mit anderen Wiedergabesystem aus.

Natürlich könnte man mehrere Versionen in unterschiedlichen Formaten anbieten, jedoch erhöhen sich bei dieser Vorgehensweise mit jeder zusätzlichen Variante der Arbeitsaufwand und der benötigte Speicherplatz. Dies fällt sowohl bei der Verteilung über CDs, DVDs als auch über das Internet negativ ins Gewicht. Verwendet der Anwender je nach Situation unterschiedliche Wiedergabemethoden, z.B. Shutterbrillen am Röhrenbildschirm seines Desktoprechners und Anaglyphen am LCD-Display des Notebooks, dann wird er gezwungen, mehrere Versionen des gleichen Videos herunterzuladen. Damit entsteht ein hohes Übertragungsvolumen für den Inhaltenanbieter und hohe Downloadkosten für den Anwender.

Besser ist es, eine einzige Version des Videos anzubieten, die unabhängig von der Wiedergabemethode kodiert ist. Erst eine spezielle 3D-Wiedergabesoftware rechnet die Daten während der Wiedergabe in das gewünschte Ausgabeformat um. Damit spart man einerseits Kosten, und erhält als positiven Nebeneffekt bei einigen Wiedergabemethoden eine signifikant bessere Bildqualität, z.B. bei der Anaglyphendarstellung. Am besten eignen sich das Nebeneinander- und Übereinanderlayout zur wiedergabeunabhängigen Kodierung. Ersteres hat den Vorteil, dass es mit dem Parallel- oder Kreuzblick auch ohne jede 3D-Software und andere Hilfsmittel betrachtet werden kann. Zweiteres hat implementierungstechnische Vorteile, da beide Halbbilder in je einem zusammenhängenden Speicherbereich liegen.

Das zuvor beschriebene Konzept der Echtzeitumwandlung wird von der Wiedergabesoftware *Stereoscopic Player* [82] umgesetzt, die zu diesem Zweck im Rahmen der Diplomarbeit entstanden ist. Alle folgenden Erläuterungen spiegeln den Entwicklungsstand der Version 0.8.4 wieder. Ähnliche Ansätze verfolgen auch der kostenlose StereoMovie Player [71] sowie die

kommerziellen Produkte Depth Charge von VRex [77], DepthQ [33] von Lightspeed Design [40], der nicht mehr erhältliche Parallax Player von StereoGraphics [69] und der Pop Media Player von Third Dimension Technologies [74].

Der Stereoscopic Player ist in der Lage, Videodateien, Internet-Streams, DVDs und Aufnahmegeräte zu öffnen und für alle bedeutenden Wiedergabemethoden aufzubereiten. Er unterstützt die Wiedergabe von Videos im Nebeneinander-, Übereinander- und Interlaced-Layout ebenso wie die Wiedergabe monoskopischer Videos. Letztgenannte werden zwar der gewählten Wiedergabemethode entsprechend aufbereitet, es wird aber beiden Augen dasselbe monoskopische Bild zugeführt – eine 2D-3D-Konvertierung wird nicht durchgeführt. Abstände zwischen den Halbbildern im Nebeneinander- oder Untereinander-Format können entfernt werden, ebenso kann das Seitenverhältnis unabhängig von der Auflösung angegeben werden.

Im Folgenden werden die unterstützten Wiedergabemethoden und ihre unterschiedlichen Anwendungsbereiche erläutert. Sie decken alle bedeutenden stereoskopischen Verfahren ab und erfüllen den Anspruch des Stereoscopic Players, universell einsetzbar zu sein.

- Die *unveränderte Ausgabe des Quellvideos* dient hauptsächlich als Unterstützung bei der korrekten Konfiguration des Quellformats.
- Oftmals besteht der Wunsch, ein 3D-Video zweidimensional zu betrachten. Damit dazu keine eigenständige 2D-Fassung des Videos erstellt werden muss, unterstützt der Player die *monoskopische Ausgabe*.
- Die *Zweibildschirmausgabe* wird zur Ausgabe des Videos auf getrennten Projektoren (Projektion in Polarisierungstechnik) oder Monitoren (Betrachtung über Stereoskop-ähnliche Spiegelvorrichtung) verwendet.
- Eine Variante der Zweibildschirmausgabe stellt die *StereoBright-Technik* dar, die für die Projektion in Polarisierungstechnik mit LCD-Projektoren benötigt wird.
- Für eine optimale Betrachtung mit Shutterbrillen müssen die Halbbilder im Pageflippingverfahren ausgegeben werden. Dies lässt sich mit dem *NVIDIA Stereotreiber* oder mit vierfachgepufferter *OpenGL-Ausgabe* erreichen. Der NVIDIA Stereotreiber unterstützt neben Pageflipping weitere Ausgabeformate, die damit im NVIDIA Stereotreiber-Modus auch mit dem Stereoscopic Player genutzt werden können.
- Das *Nebeneinander-Format* wird für die hilfsmittelfreie 3D-Wiedergabe und für die Betrachtung mit (Brewster-)Stereoskopen benötigt. Zwischen den Halbbildern kann ein schwarzer Trennbalken variabler Größe eingefügt werden.
- Die Syncdoubling-Technik erfordert die Ausgabe im *Untereinander-Format* mit schwarzem Trennbalken. Das Untereinander-Format wird außerdem für die KMQ-Technik benötigt.
- Die *zeilenweise Verschachtelung* ist die Grundlage der Shutterbrillenwiedergabe in der Interlaced- und Lineblanking-Betriebsart.
- Die *spaltenweise Verschachtelung* wird für einige autostereoskopische Displays benötigt.
- *Echte Anaglyphen* lassen sich sowohl in der Variante für Rot-Blau- als auch Rot-Grün-Brillen ausgeben. *Graue, farbige und halbfarbige Anaglyphen* kann der Stereoscopic Player für Rot-Cyan- und Gelb-Blau-Brillen erzeugen.

Wo sinnvoll, steht jeweils auch die komplementäre Wiedergabemethode zur Verfügung, z.B. sowohl echte Anaglyphen für Rot-Blau- als auch Blau-Rot-Brillen.

## 8.2 Architektur

Der Stereoscopic Player basiert auf DirectShow. Die Entscheidung zugunsten von DirectShow als grundlegende Multimedia-Architektur fiel aufgrund dessen umfassender Unterstützung fast aller bedeutender Videoformate. Dadurch entfällt die Arbeit, die Unterstützung jedes Videoformats separat implementieren zu müssen. Sofern für ein Videoformat ein DirectShow-kompatibler Decoder zur Verfügung steht, dann kann dieses Format auch im Stereoscopic Player verwendet werden.

### 8.2.1 Dateiwiedergabe

Öffnet der Anwender eine Datei, wird zunächst mit den DirectShow-Funktionen zum automatischen Aufbau eines Filtergraphen ein Graph für die zweidimensionale Wiedergabe erstellt. Anschließend wird dieser für die 3D-Wiedergabe modifiziert, indem an geeigneter Stelle der *Stereo Transformation Filter* eingefügt wird, der die Aufbereitung des Videos für die unterschiedlichen Wiedergabeformate übernimmt. Für den NVIDIA Stereotreiber- und den OpenGL-Modus muss der Standard-Video-renderer durch den *NVIDIA Stereo Renderer* bzw. den *OpenGL Stereo Renderer* ersetzt werden. Beide Renderer sind nicht Bestandteil von DirectShow und mussten selbst entwickelt werden.

Der Stereo Transformation Filter führt Berechnungen auf den Bilddaten durch und kann deshalb nur an jenen Stellen in den Graphen eingefügt werden, wo unkomprimierte Videoframes vorliegen. Im Graphen von Abbildung 8.1 ist diese Voraussetzung nur zwischen dem Videodecoder und dem Videorenderer erfüllt.

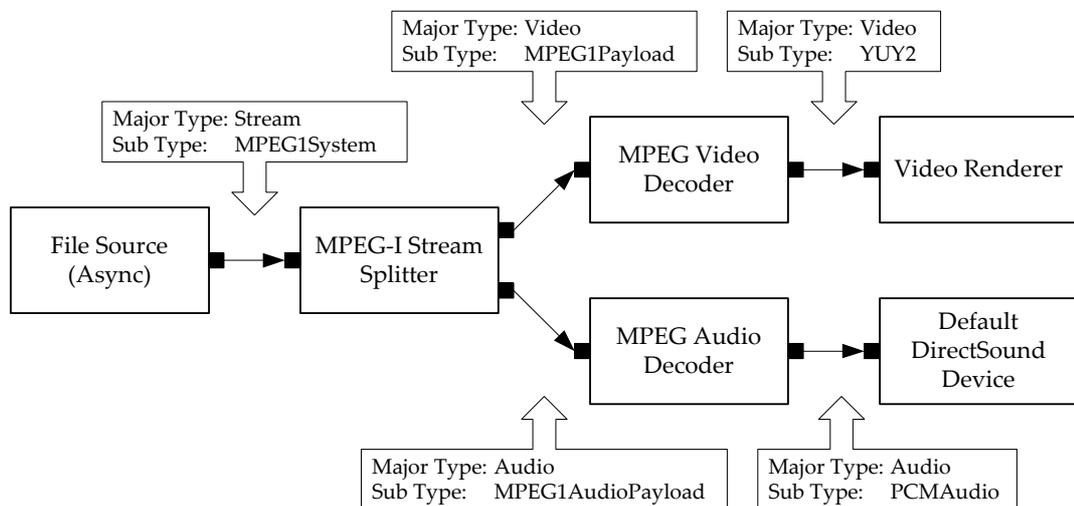


Abbildung 8.1: Medientypen in einem MPEG-1 Wiedergabegraph

Der Algorithmus des Stereoscopic Player sucht zunächst den Videorenderer und löst die Verbindung mit dessen Upstream-Filter. Darüberhinaus werden weitere Zwischenfilter entfernt, deren Funktion entweder der Stereo Transformation Filter selbst übernehmen kann (z.B. Farbraumkonverter) oder überflüssig sind. In der aktuellen Implementierung wird die Entscheidung über den Verbleib eines Filters anhand einer Blacklist unerwünschter Filter getroffen. Anschließend wird der Stereo Transformation Filter in den Graphen eingefügt, der Eingang

mit dem verbliebenen Teilgraphen und der erste Ausgang mit dem Video Renderer verbunden. Abbildung 8.2 zeigt das Ergebnis.

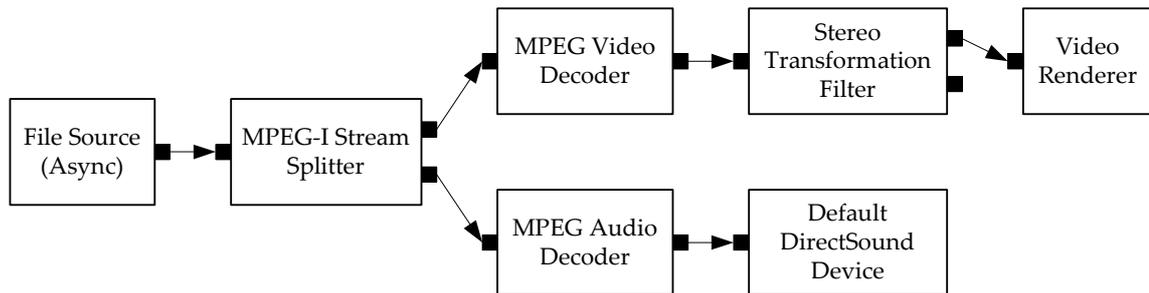


Abbildung 8.2: 3D-Wiedergabegraph einer MPEG-I-Datei

Für die Zweibildschirmausgabe und die StereoBright-Methode wird ein weiterer Renderer eingefügt und mit dem zweiten Ausgang des Stereo Transformation Filter verbunden. Der erste Renderer wird mit linken und der zweite Renderer mit rechten Halbbildern versorgt. Die Ausgabefenster der Videorenderer werden auf dem jeweiligen Bildschirm platziert.

Läuft ein PC im Zweibildschirmbetrieb, erstreckt sich der Windows-Desktop über beide Bildschirme, meist sind diese nebeneinander angeordnet. Oft wird vorgeschlagen, mit Hilfe des Nebeneinander-Format die Zweibildschirmausgabe eines 3D-Videos zu realisieren. Anstelle der aufwendigen Realisierung mit zwei Videorenderern wird das Video desktopfüllend wiedergegeben, sodass die linke Perspektive am linken Bildschirm und die rechte Perspektive am rechten Bildschirm zu liegen kommt. Versuche mit verschiedenen Wiedergabeprogrammen zeigen sofort, dass dieses Verfahren nicht funktioniert – einer der beiden Bildschirme bleibt dunkel. Die Ursache liegt darin, dass moderne Grafikkarten Hardware-Overlays für die Videowiedergabe verwenden, die bis dato – unabhängig vom Hersteller – auf einen Bildschirm beschränkt sind. Ein Verzicht auf die Hardwarebeschleunigung umgeht dieses Problem, führt aber zu signifikanten Performanceeinbußen.

Es sei an dieser Stelle angemerkt, dass fast alle Grafikkarten nur ein einziges Overlay unterstützen. Folglich muss für die Darstellung der zweiten Perspektive auf gewöhnliche DirectX-Surfaces zurückgegriffen werden. Viele Grafiktreiber bieten getrennte Bildeinstellungsmöglichkeiten für Overlays und den Desktop, sodass bei Verwendung der Zweibildschirmausgabe im Stereoscopic Player unbedingt auf gleiche Einstellungen geachtet werden muss.

## 8.2.2 Streaming- und Live-Wiedergabe

Die Wiedergabe von Streams über ein Netzwerk funktioniert äquivalent zur Dateiwiedergabe: Der Filtergraph-Manager erzeugt einen Wiedergabegraphen, der anschließend für die stereoskopische Wiedergabe modifiziert wird.

Bei der Live-Wiedergabe von einem Capture-Gerät sind einige Besonderheiten zu beachten. So besitzen einige Capture-Filter Eingangspins, die ebenfalls mit zusätzlichen Filtern verbunden werden müssen. Diese Hilfsfilter repräsentieren und steuern die interne Signalführung des Capture-Geräts. Eine TV-Karte erfordert beispielsweise einen Video-Crossbar-Filter, der die Steuerung der Eingangsquelle (Video, S-Video, Tuner) übernimmt. DirectShow erleichtert dem Programmierer mit dem *Capture Graph Builder*-Objekt, das auch vom Stereoscopic Player

verwendet wird, die Arbeit. Es fügt alle Hilfsfilter ein und unterstützt auch den automatischen Aufbau des restlichen Graphen.

Ein Capture-Geräte kann eine Reihe von Konfigurationsdialogen zur Verfügung stellen, die in der Benutzeroberfläche zugänglich gemacht werden müssen. Eine Besonderheit stellt die Eigenschaftsseite zur Konfiguration des gewünschten Ausgangsformats dar, die vom Ausgangspin zur Verfügung gestellt wird. Das Ausgangsformat kann nur dann erfolgreich geändert werden, wenn der Ausgangs-Pin nicht verbunden ist. Bevor die Eigenschaftsseite angezeigt wird, müssen alle Downstream-Filter entfernt und nach dem Schließen des Dialogs der Graph neu aufgebaut werden.

### 8.2.3 DVD-Wiedergabe

Die DVD-Wiedergabe wird vom Standardumfang von DirectShow nur unvollständig unterstützt. Zwar steht zum Lesen der DVD-Strukturen und zur Navigation der *DVD Navigator* zur Verfügung, jedoch fehlen ein MPEG-2- und ein Dolby Digital-Decoder. Abbildung 8.3 zeigt einen 3D-DVD-Wiedergabegraphen, die Decoder sind dabei Platzhalter für entsprechende Filter von Drittanbietern.

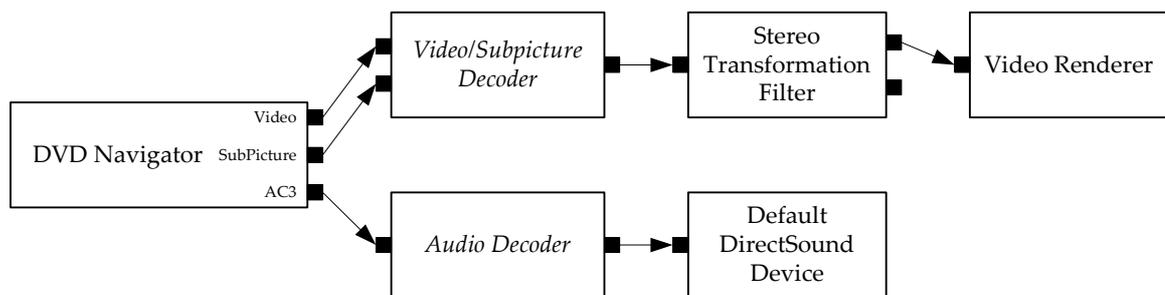


Abbildung 8.3: 3D-DVD-Wiedergabegraph

Bedauerlicherweise traten mit allen getesteten kommerziellen Videodecodern Schwierigkeiten auf, sodass ein eigener Decoder, der *GPL MPEG-1/2 Decoder* [81] entwickelt werden musste. Aus lizenzrechtlichen Gründen unterstützt er keine Wiedergabe verschlüsselter DVDs. Der GPL MPEG-1/2 Decoder basiert auf dem Decoder des *Guliverkli-Projekts* [23], der wiederum auf die freie Bibliothek *libmpeg2* [38] zurückgreift.

Ein Videodecoder muss folgende Eigenschaften erfüllen, um für die Wiedergabe von 3D-DVDs im Stereoscopic Player geeignet zu sein:

- Da die meisten 3D-DVDs im Interlaced-Format vorliegen und ein Deinterlacing die Halbbilder untrennbar vermischen würde, muss die Deinterlacing-Funktion abschaltbar sein, z.B. über die Eigenschaftsseite des Decoders. Im Idealfall bietet der Decoder ein Interface, mit dem der Stereoscopic Player das Deinterlacing-Verhalten in Abhängigkeit des Stereoformats automatisch konfigurieren kann. In der aktuellen Implementierung funktioniert dies nur mit dem GPL MPEG-1/2 Decoder.
- Bei abgeschaltetem Deinterlacing sollte der Decoder die Farbinterpolation vom 4:2:0- in den 4:2:2- oder RGB-Farbraum gemäß dem MPEG-2 Standard für jedes Field getrennt vornehmen. Anderenfalls vermischen sich die Farbkomponenten der stereoskopischen Halbbilder, und Störbilder sind die Folge.

- Der Videodecoder muss eine Verbindung mit dem Stereo Transformation Filter erlauben, d.h., er darf nicht voraussetzen, dass die Verbindung mit einem der Videorenderer erfolgt.
- Der Wechsel vom NTSC in den PAL-Modus oder umgekehrt muss über dynamische Formatwechsel realisiert werden und nicht über einen proprietären Mechanismus.

## 8.3 Transformationsfilter

Der Stereo Transformation Filter, enthalten in der DLL (Dynamic Link Library) *Stf.ax*, ist das Herzstück des Stereoscopic Player. Der Filter kann auch außerhalb des Stereoscopic Players verwendet werden, z.B. in GraphEdit, um stereoskopische Videos in ein anderes Format zu wandeln.

Der Filter besteht aus der Klasse `CSTF`, abgeleitet von `CBaseFilter`. Der Eingangspins wird durch die Klasse `CSTFInput`, abgeleitet von `CBaseInputPin`, die beiden Ausgangspins durch die Klasse `CSTFOutput`, abgeleitet von `CBaseOutputPin`, repräsentiert.

Der *Video Renderer* Filter erlaubt ausschließlich Verbindungen im RGB-Format, wobei die Farbtiefe jener des Desktops entsprechen muss. Erst während der Wiedergabe initiiert der Video Renderer einen dynamischen Formatwechsel zu einem YUV-Format, sofern ein Hardware-Overlay zur Verfügung steht. Der neuere *Video Mixing Renderer 7* (ab Windows XP verfügbar) sowie der *Video Mixing Renderer 9* (ab DirectX 9 verfügbar) verwenden ausschließlich DirectX-Funktionen zur Wiedergabe und greifen nie auf das Windows GDI (Graphics Device Interface) zurück. Sie erlauben deshalb von Anfang an Verbindungen in YUV-Formaten.

Da der Stereo Transformation Filter sowohl mit allen Videorenderern zusammenarbeiten als auch eine bestmögliche Performance bieten soll, muss er in der Lage sein, am Ausgang sowohl RGB-Formate als auch zumindest ein gebräuchliches YUV-Format zu unterstützen. Konkret sind es RGB 24, RGB 32 und YUY2.

Da das RGB 8-Format derzeit nicht unterstützt wird, funktioniert der Stereoscopic Player bei einer Desktop-Farbtiefe von 8 Bit (256 Farben) nicht. 256 Farben sind für die Videowiedergabe ohnehin ungeeignet, deshalb stellt dies keine funktionale Einschränkung dar. Eine Desktop-Farbtiefe von 16-Bit Farbtiefe erlaubt der Stereoscopic Player nur unter Windows XP oder höher. In diesem Fall wird der *Color Space Converter* zwischen dem Stereo Transformation Filter und dem Video-Renderer eingefügt. Der Stereoscopic Player überprüft die Desktopfarbtiefe während dem Programmstart und gibt bei nicht geeigneten Einstellungen entsprechende Fehlermeldungen aus. Obwohl der 16-Bit Modus unter Windows XP prinzipiell verwendet werden kann, wird dies nicht empfohlen, da der zusätzliche Konvertierungsfilter die Wiedergabepformance verschlechtert.

YUY2 ist neben UYVY der gebräuchlichste YUV 4:2:2 Farbraum. YUY2 und UYVY unterscheiden sich lediglich in der Anordnung der Bytes im Speicher. Im PC98 Standard wurde die hardwaremäßige Unterstützung beider Farbräume für Grafikkarten vorgeschrieben, dementsprechend umfassend werden sie von aktuellen Modellen unterstützt.

Es ist nicht immer sinnvoll, eine Ausgabe im YUY2-Format zuzulassen. Sowohl Anaglyphen, die StereoBright-Methode als auch das vertikal verschachtelte Interlaced-Format verlieren durch die halbierte horizontale Farbauflösung deutlich an Qualität. Im Stereo Transformation Filter wurde deshalb die Möglichkeit vorgesehen, Verbindungen im YUY2-Format zu verhin-

dern. Der Stereoscopic Player macht von dieser Möglichkeit Gebrauch, bietet aber auch die Möglichkeit, auf leistungsschwächeren PCs die YUY2-Ausgabe zuzulassen. Zu Testzwecken lassen sich im Stereo Transformation Filter auch RGB-Verbindungen unterdrücken.

Der Stereo Transformation Filter verarbeitet die eingehenden Videodaten standardmäßig in Originalauflösung. Erst der Video-Renderer führt mit Unterstützung der Grafikhardware die Skalierung auf die endgültige Größe durch. Der Stereo Transformation Filter ist allerdings auch in der Lage, am Ausgang Bilder beliebiger Größe zu liefern. In diesem Fall werden die Daten intern bilinear interpoliert. Der Stereoscopic Player nutzt die Skalierung nur für zeilen- oder spaltenweise Interlaced-Darstellung, denn diese beiden Wiedergabemethoden müssen gemäß ihrem Funktionsprinzip in der Zielauflösung aufbereitet werden.

Während RGB 24, RGB 32 und YUY2 als Ausgangsformate unterstützt werden, versteht sich der Eingang auch auf UYVY. Das zweite YUY 4:2:2 Format ist notwendig, weil einige Aufnahmegeräte nur jeweils eines der beiden Formate unterstützen. Dies trifft auch für jene Kameras zu, die für diese Diplomarbeit verwendet wurden – sie unterstützen lediglich UYVY. Bezüglich der Wahl unterschiedlicher Eingangs- und Ausgangsformate unterliegt der Stereo Transformation Filter keinerlei Einschränkungen. Es sind alle denkbaren Kombinationen möglich. Wenn erforderlich, wird intern eine Umwandlung des Farbraumes durchgeführt. Abgesehen von der Skalierung – diese wird im Eingangsformat mit 8 Bit Genauigkeit durchgeführt – erfolgen Farbraumkonvertierung und Aufbereitung für die gewählte stereoskopische Wiedergabemethode in einem Schritt mit mehr als 8 Bit Genauigkeit, um einen Qualitätsverlust durch Rundungsfehler zu vermeiden.

Der Stereo Transformation Filter wurde in C++ entwickelt. Als Entwicklungsumgebung kam Visual Studio .NET 2003 zu Einsatz.

## 8.4 Renderingfilter

Spezielle Renderer sind notwendig, wenn eine Wiedergabemethode den Zugriff auf spezielle Funktionen des Grafiktreibers erfordert. Ansonsten werden entsprechend der Wiedergabemethode aufbereitete Bilder mit einem der Videorenderer aus dem Standardumfang von DirectShow ausgegeben.

Konkret werden für die Ausgabe über den NVIDIA Stereotreiber und für vierfachgepufferte OpenGL-Ausgabe spezielle Renderer benötigt. Der NVIDIA Stereotreiber arbeitet grundsätzlich nur im Vollbildmodus und wurde ursprünglich entwickelt, um Direct3D- und OpenGL-Spiele stereoskopisch betrachten zu können. Der Treiber erzeugt transparent für die Anwendung die linke und rechte Perspektive, indem er den Betrachtungsstandpunkt modifiziert. Zur Ausgabe von vorberechneten Bildpaaren bietet NVIDIA das StereoBlt SDK an (in mehreren Varianten für unterschiedliche DirectX-Versionen). Zur Verwendung des StereoBlt SDK muss man zunächst ein Direct3D-Objekt erzeugen – ein DirectDraw-Objekt ist nicht ausreichend. Anschließend kopiert man linkes und rechtes Halbbild nebeneinander in eine DirectX-Surface und fügt eine Signatur in der linken unteren Ecke ein. Die Signatur ist später am Bildschirm in Form bunter Pixel sichtbar, stört in der Praxis aber kaum. In welcher Form die endgültige Ausgabe erfolgt, hängt von den Treibereinstellungen ab. Der aktuelle Referenztreiber mit der Versionsnummer 56.74, verfügbar auf der NVIDIA Homepage, unterstützt offiziell Pageflipping und Farbanaglyphendarstellung. Durch Setzen eines Wertes in der Systemregistrie-

rung lassen sich aber noch einige weitere Modi freischalten. Einige davon lassen sich als Alternative zu den nativen Modi des Stereoscopic Player verwenden (Zweibildschirmwiedergabe, Interlaced-Darstellungen).

Der selbstentwickelte *NVIDIA Stereo Renderer*, enthalten in der DLL *Nvr.ax*, wurde in C++ unter Visual Studio .NET 2003 entwickelt und basiert auf dem StereoBlt SDK für DirectX 7. Es wurde Version 7 der Vorzug gegeben, da diese alle benötigten Funktionen bietet und seit Windows 2000 auf allen System vorinstalliert ist. Der NVIDIA Stereo Renderer akzeptiert am Eingang nur YUY2-Daten (andere Formate ließen sich leicht ergänzen) in Kreuzblickanordnung. Um ein Umkopieren der Bilddaten in eine DirectX-Surface zu vermeiden, implementiert der Renderer das *IMemAllocator*-Interface selbst und übergibt *Mediasamples*, die einen Zeiger auf die DirectX-Surface enthalten, an den Upstream-Filter (in der Regel den Stereo Transformation Filter). Dieser verwendet damit bereits die DirectX-Surface als Ausgangspuffer, ein Umkopieren der Daten im NVIDIA Stereo Renderer entfällt. Die YUY2-Surface wird anschließend in eine RGB-Surface umkopiert, dabei wird gleichzeitig auf die gewünschte Zielauflösung skaliert. Sowohl die Farbraumkonvertierung als auch die Skalierung werden von der Grafikkhardware durchgeführt und belasten die Wiedergabepformance kaum. Abschließend wird die Stereosignatur in die RGB-Surface eingefügt. Die weiteren Schritte unterscheiden sich nicht von anderen DirectX-Anwendungen: Die Surface wird mit einer Blit-Operation in den Backbuffer kopiert und dieser mit einer Flip-Operation präsentiert. Da der NVIDIA Stereotreiber nur im exklusiven Vollbildmodus arbeitet, kann über das *INVRRenderer*-Interface die Monitоруауflösung und Farbtiefe (16 oder 32 Bit) gewählt werden. Die tatsächliche Ausgabegröße wird mit dem *IVideoWindow*-Interface gesetzt. Darüber hinaus kann im Falle mehrerer Monitore jener festgelegt werden, auf dem die Ausgabe erfolgt. Die Bildwiederholrate wird in den Einstellungen des Stereotreibers festgelegt.

Der *OpenGL Stereo Renderer*, enthalten in der DLL *Glr.ax*, wurde ebenso wie die anderen Filter in C++ entwickelt [46]. Im Gegensatz zu DirectX ist der Stereomodus in der OpenGL-Spezifikation [55] fest verankert und funktioniert herstellerübergreifend. Weiters wird die Stereowiedergabe in einem Fenster unterstützt. Leider stehen Treiber mit OpenGL-Unterstützung nur für professionelle Grafikkarten zur Verfügung, Consumerkarten bleiben außen vor. Der OpenGL Stereo Renderer unterstützt am Eingang das Untereinander-Format im RGB Farbraum. Beträgt die Desktopfarbtiefe 24 Bit, wird nur das RGB 24 Format akzeptiert, sonst nur das 32 Bit Format. Der Eingangspuffer wird in eine Textur umkopiert und im Anschluss als ein texturiertes Rechteck gerendert. Der Weg über eine Textur ist notwendig, um eine gefilterte Darstellung anstelle von Pixelreplikation zu erzielen. Im Gegensatz zum NVIDIA Stereo Renderer muss der OpenGL Stereo Renderer selbst einen Eingangspuffer bereitstellen, da OpenGL keinen direkten Zugriff auf die Texturdaten erlaubt und damit kein Zeiger an den Upstream-Filter übergeben werden kann. Die Umwandlung vom YUV in den RGB-Farbraum muss ebenfalls vom Upstream-Filter in Software durchgeführt werden. Eine Verbesserung der Performance ließe sich durch die Verwendung von herstellerepezifischen OpenGL-Erweiterungen zum Zugriff auf die Texturdaten und zur Farbraumkonvertierung erzielen.

## 8.5 Benutzeroberfläche

Die Benutzeroberfläche des Stereoscopic Player wurde im Gegensatz zu den DirectShow-Filtern in Borland Delphi 7 programmiert. Delphi ist ein RAD-Werkzeug und erlaubt die rasche

Entwicklung ansprechender Benutzeroberflächen. Die Wahl unterschiedlicher Programmiersprachen fördert außerdem die Trennung der datenverarbeitenden Komponenten von der Benutzerschnittstelle. Es sei an dieser Stelle nochmals darauf hingewiesen, dass die entwickelten DirectShow-Filter auch außerhalb des Stereoscopic Player einsetzbar sind.

Zu den Aufgaben der in Delphi geschriebenen Teile des Stereoscopic Player gehört die Interaktion mit dem Benutzer genauso wie die Erstellung und Steuerung der Wiedergabegraphen sowie die Verwaltung der Videoinformationen und der Bibliothek. Die folgenden Ausführungen beziehen sich ausschließlich auf die Kernfunktionen des Players. Eine vollständige Beschreibung aller Klassen und Module des Players würde den Rahmen dieser Arbeit sprengen.

### 8.5.1 Wiedergabeklassen

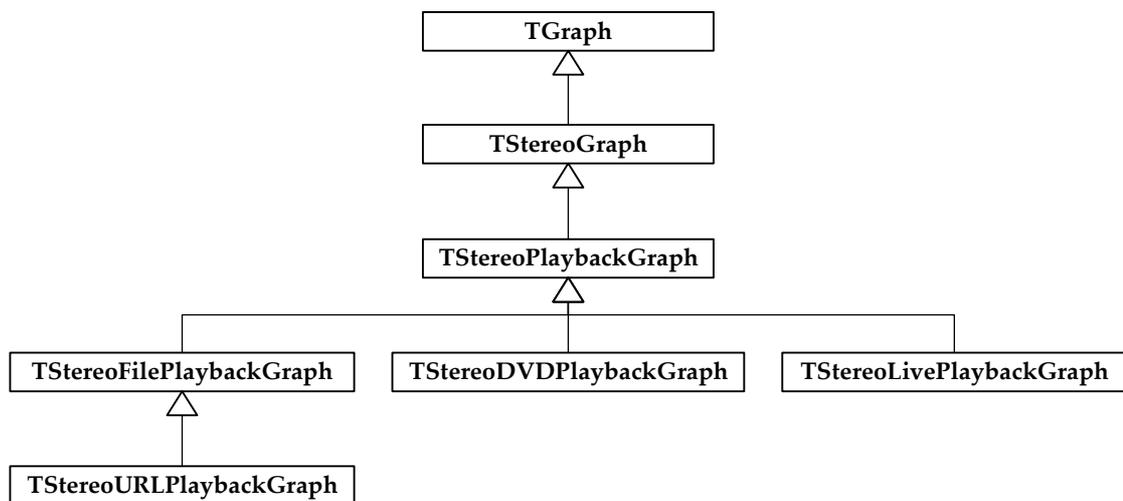


Abbildung 8.4: Klassenhierarchie der Wiedergabegraphen

Die Basisklasse für einen Wiedergabegraphen ist TGraph. Sie kapselt einen DirectShow-Filtergraphen und stellt die benötigten Methoden zur Verfügung, um einen Graphen aufzubauen (z.B. Filter suchen, einfügen und verbinden) und zu steuern (z.B. Wiedergabe starten, stoppen, aktuelle Position abfragen und setzen). Der Aufruf von Steuerungsmethoden nimmt unter Umständen erhebliche Zeit in Anspruch, z.B. wenn der Zugriff auf die Quelldatei über das Internet erfolgt. Um während dem Aufruf dieser Methoden ein Blockieren der Benutzeroberfläche zu verhindern, werden diese von einem eigenen Thread aufgerufen, der vom Konstruktor von TGraph erstellt wird.

TStereoGraph erweitert TGraph um grundlegende Methoden zum Einfügen und Konfigurieren der stereospezifischen DirectShow-Filter. TStereoPlaybackGraph erweitert wiederum TStereoGraph und implementiert den Umbau eines gewöhnlichen Wiedergabegraphen für die stereoskopische Wiedergabe. Neben dem Einfügen des Stereo Transformation Filters und bei Bedarf der beiden stereoskopischen Renderer verwaltet TStereoPlaybackGraph die Eigenschaften des Videofensters und des Audio-Renderers und sorgt für eine korrekte Abwicklung des Wechsels von Eingangsformat und Wiedergabemethode. Den Aufbau des gewöhnlichen Graphen übernehmen je nach Medium TStereoFilePlaybackGraph, TStereoURLPlaybackGraph, TStereoDVDPlaybackGraph oder TStereoLivePlaybackGraph.

## 8.5.2 Komfortfunktionen

Im Vergleich zu einem 2D-Video gestaltet sich die Wiedergabe eines stereoskopischen Videos für den Anwender aufwendig. Zum einen muss er aus den zahlreichen angebotenen Wiedergabemethoden die für sein System passende auswählen und eventuell sogar zusätzliche Hardware anschließen bzw. konfigurieren. Zum anderen muss er für jede Datei und jede DVD dessen Stereoformat korrekt festlegen.

Einmal ausgewählt, muss die Wiedergabemethode in der Regel nicht mehr gewechselt werden. Der Stereoscopic Player merkt sich die zuletzt verwendete Methode. Außerdem lassen sich eine bevorzugte monoskopische und stereoskopische Wiedergabemethode festlegen und per Knopfdruck aufrufen. Alle anderen Methoden können über das Menü trotzdem schnell angewählt werden.

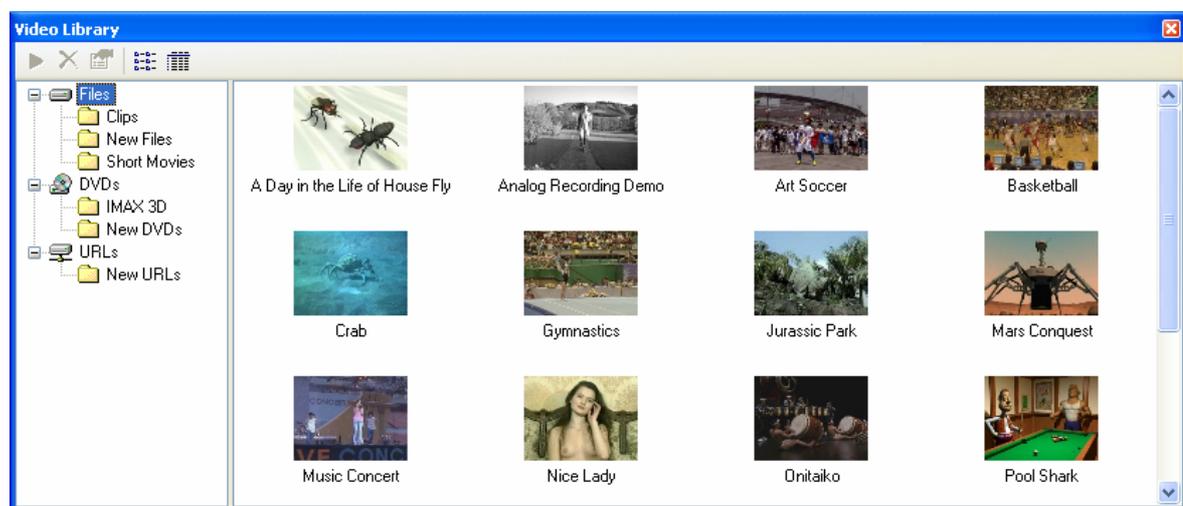


Abbildung 8.5: Bibliothek des Stereoscopic Players

Während es sinnvoll ist, die Wiedergabemethode global zu speichern, unterscheidet sich das Stereoformat je Datei, URL oder DVD. Aus diesem Grund bietet der Stereoscopic Player eine Bibliothek, welche die Einstellungen für bereits geöffnete Dateien, URLs und DVDs speichert. Der Anwender muss damit nur beim erstmaligen Öffnen eines Videos die Einstellungen vornehmen. Neben den grundlegenden Einstellungen können zu jedem Eintrag weitere Informationen, einschließlich eines Vorschaubildes, angegeben werden.

Trotzdem stellt die richtige Konfiguration eines Videos viele Anwender vor Schwierigkeiten: Während das grundlegende Layout (nebeneinander, untereinander, interlaced) bei Betrachtung in unveränderter Darstellung und vorhandenem Wissen über Stereoformate sofort ersichtlich ist, können linkes und rechtes Halbbild nur durch stereoskopische Betrachtung oder ein Links-Rechts-Testbild am Anfang des Films (in der Art von Abbildung 5.1 bis Abbildung 5.3) unterschieden werden. Ein Testbild ist in den meisten Fällen nicht vorhanden, sodass nur durch Betrachten des Videos die richtige Halbbildanordnung ermittelt werden kann. Besonders unerfahrenen Anwendern fällt es oft schwer, die pseudostereoskopische Darstellung von der richtigen zu unterscheiden.

Eine nahe liegende Lösung des Problems besteht darin, die Aufgabe der richtigen Konfiguration vom Anwender zum Autor des Videos zu verlagern. Theoretisch stellt diese Vorgangs-

weise kein Problem oder einen Mehraufwand dar, schließlich weiß der Autor über das verwendete Format Bescheid – allerdings gibt es bis dato keinen Ansatz, geschweige denn einen Standard, zur Übermittlung dieser Informationen.

Der Stereoscopic Player definiert ein eigenes Dateiformat für *stereoskopische Informationsdateien* zur Speicherung dieser Einstellungen. Öffnet der Anwender eine Videodatei, sucht der Stereoscopic Player im selben Verzeichnis eine gleichnamige stereoskopische Informationsdatei mit der Erweiterung *svi* (Stereoscopic Video Information File). Als Alternative dazu kann die stereoskopische Informationsdatei auch auf einem Webserver abgelegt sein. In diesem Fall wird ein 64-Bit-Hashwert des Dateiinhalts als Dateiname verwendet. Der Stereoscopic Player versucht per http-Request die Informationsdatei herunterzuladen. Schlägt der Download fehl, weil die Datei nicht am Server vorhanden ist, dann bedeutet das, dass keine Informationsdatei verfügbar ist. In diesem Fall wird die Suche auf anderen Servern fortgesetzt, sofern weitere Server definiert wurden.

Ein Server muss eine Textdatei mit dem Namen *UpdateLibrary.dat* zur Verfügung stellen, die den String `Stereovideo-Updateserver[V1.0]` enthält, andernfalls wird der Server vom Stereoscopic Player ignoriert. Das Vorhandensein dieser Datei wird am Anfang jeder Verbindungsaufnahme überprüft. Der 64-Bit-Hashwert wird nicht nur zum Download von stereoskopischen Informationsdateien verwendet, sondern dient auch zur Identifikation von Dateien und DVDs in der Bibliothek. Einträge in der Bibliothek sind damit nicht an den Pfad und Namen einer Datei gebunden (ausgenommen URLs, da der Datenstrom nicht vorab zur Bildung des Hashwerts zur Verfügung steht) und lassen sich darüber hinaus ohne Zugriff auf die Videodateien über Webserver aktualisieren.

Der Stereoscopic Player sucht in folgender Reihenfolge nach den korrekten Einstellungen, wobei die Suche abgebrochen wird, sobald Informationen gefunden wurden:

1. Eintrag in der internen Bibliothek
2. Stereoskopische Informationsdatei im selben Verzeichnis der Videodatei
3. Stereoskopische Informationsdatei auf einem Webserver

Falls der Player an keinem der drei Orte fündig wurde, wird das zuletzt gewählte Stereolayout vorgeschlagen, wobei Dateien, URLs, DVDs und Aufnahmegeräte getrennt behandelt werden. Alle anderen Parameter erhalten ihre Standardwerte.

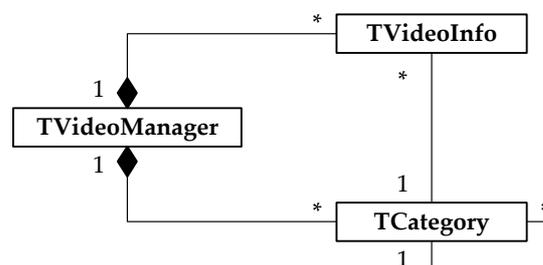


Abbildung 8.6: UML-Diagramm der Bibliotheksklassen

Die Bibliothek ordnet Videos in eine hierarchische Struktur von Kategorien ein. Derzeit sind drei Wurzelkategorien definiert: Dateien, DVDs und URLs. Die Bibliothek wird durch die Klasse `TVideoManager` implementiert, die eine Liste der Kategorien und Videos verwaltet.

Kategorien werden durch `TCategory` repräsentiert, Videos durch `TVideoInfo`. Das Feld `Mediatype` von `TVideoInfo` legt fest, ob es sich um eine Datei, DVD oder URL handelt. Jedes Video verweist auf die Kategorie, der es zugeordnet ist. Kategorien verweisen ebenfalls auf ihre Eltern-Kategorien, ausgenommen Wurzelkategorien.

## 8.6 Test der Wiedergabefunktionen

Zum Testen der Anaglyphenmethoden standen Rot-Blau, Rot-Grün und Rot-Cyan-Brillen in Kartonausführung sowie Rot-Cyan-Brillen in Plastikausführung von [57] und Gelb-Blau-Brillen von ColorCode 3-D Center ApS [10] zur Verfügung.

Für die Shutterbrillenwiedergabe wurden Brillen von [35] sowie Asus verwendet, die an einem selbstentwickelten Controller betrieben wurden. Der Controller unterstützt sowohl Pageflipping, Lineblanking als auch Syncdoubling. Dazu muss das VGA-Signal durch den Controller durchgeschleift werden. Darüber hinaus akzeptiert er Composite- sowie S-Video-Eingangssignale und lässt sich auch über die parallele und serielle Schnittstelle steuern. Es können sowohl Shutterbrillen mit Klinckenstecker (Spannung regelbar im Bereich von 0 bis 14 Volt) als auch Brillen mit dreipoligem VESA Mini-DIN-Stecker betrieben werden. Die Ansteuerung von Infrarotbrillen ist hardwareseitig vorgesehen und wird in zukünftige Firmwareversionen integriert werden.



Abbildung 8.7: Shutterbrillen-Controller *3D Control*

Die Wiedergabe in Polarisationstechnik konnte im Virtual Reality-Labor der Johannes Kepler Universität Linz getestet werden. Dazu stand ein EON ICatcher-System [18], bestehend aus einem Cyviz-Demultiplexer [11], zwei DLP-Projektoren und einer Steward-Silberleinwand zur Verfügung.

Die Überprüfung der Funktionsfähigkeit der Zweibildschirm Ausgabe wurde mit zwei Computerbildschirmen durchgeführt, die Ausgabe aller übrigen Wiedergabemodi wurde ebenfalls anhand deren Bildschirmausgabe beurteilt.

---

# Implementierung der Aufnahme

---

*Kapitel 9 widmet sich den Herausforderungen der Aufnahme stereoskopischer Videos. Es werden die gewählten Lösungen vorgestellt und Messergebnisse mit einem ausgewählten Kamerapaar präsentiert. Im Mittelpunkt dieses Kapitels steht aber der Stereoscopic Multiplexer, ein Treiber der innerhalb anderer Programme zur Aufnahme stereoskopischer Videos verwendet werden kann.*

## 9.1 Kameraauswahl

Die Probleme der 3D-Aufnahme wurden bereits in Kapitel 6 ausführlich erläutert. Da als Ziel der Diplomarbeit eine Live-Übertragung definiert wurde, waren folgende Kriterien für die Auswahl der Kameras entscheidend:

1. Möglichkeit der Synchronisierung möglichst vieler Kameraparameter. Eine laufende manuelle Kontrolle der Halbbildpaare auf Abweichungen und ein händlicher Abgleich kommen für die Echtzeitverarbeitung nicht in Frage.
2. Digitaler Ausgang mit weit verbreiteter Schnittstelle, damit die Lösung auf möglichst vielen Rechnern ohne spezielle Framegrabber-Karten lauffähig ist, und später für Video-Konferenzen adaptiert werden kann.
3. Es ist von Vorteil, wenn die Kameras unkomprimierte Bilddaten liefern. In diesem Fall entfällt die Dekomprimierung vor der Weiterverarbeitung der Bilddaten.
4. Erfassung von progressiven Bildern um den Rechenaufwand und Qualitätsverlust durch Deinterlacing zu vermeiden.

Gewöhnliche DV-Videokameras erfüllen Punkt 1, 3 und 4 nicht und wurden deshalb nicht in Betracht gezogen. Es sei aber darauf hingewiesen, dass DV-Kameras gut für 3D-Aufnahmen geeignet sind, wenn keine Echtzeitbearbeitung erforderlich ist und Abweichungen der Halbbilder im Rahmen der Nachbearbeitung korrigiert werden können.

Die Wahl fiel schlussendlich auf zwei DFW-VL500 Firewire-Kameras von Sony [66]. Es handelt sich dabei um Industriekameras gemäß der *IIDC 1394-based Digital Camera Specification* [2, 75]. Die Sony DFW-VL500 unterstützt eine Reihe von Auflösungen und Farbräumen. Tabelle 9.1 gibt Aufschluss über alle möglichen Modi. Unabhängig von Auflösung und Farbformat können Bildwiederholraten von  $3\frac{3}{4}$ ,  $7\frac{1}{2}$ , 15 und 30 Bildern pro Sekunde (Frames per Second, fps) verwendet werden. Die angegebenen Formate gelten nur für den verwendeten Originaltreiber von Sony. Alternative Treiber bilden die Kamerafarb Räume zum Teil auf andere Direct-Show-Formate ab.



Abbildung 9.1: Sony DFW-VL500 Kameras

Auflösung	Farbraum	Format	Datenrate (30 fps)
640 x 480	YUV 4:2:2, 8 Bit pro Komponente	UYVY	141 Mbps
640 x 480	YUV 4:1:1, 8 Bit pro Komponente	IYU1	105 Mbps
640 x 480	8 Bit Graustufendarstellung	RGB8	70 Mbps
320 x 240	YUV 4:2:2, 8 Bit pro Komponente	UYVY	35 Mbps
160 x 120	YUV 4:4:4, 8 Bit pro Komponente	IYU2	13 Mbps

Tabelle 9.1: Betriebsmodi der Sony DFW-VL500 Kamera

Obwohl in allen Fällen die Datenrate zweier Kameras unter der nominellen Bandbreite des Firewire-Busses von 400 Mbps bleibt, ist es bei der höchsten Qualitätsstufe (640 x 480, YUY 4:2:2, 30 fps) nicht möglich, beide Kameras an einem Controller zu betreiben. Das liegt einerseits darin, dass der verwendete isochrone Übertragungsmodus nicht die gesamte Bandbreite belegen kann. Andererseits steht durch den Protokoll-Overhead grundsätzlich nicht die volle Bandbreite zur Verfügung.

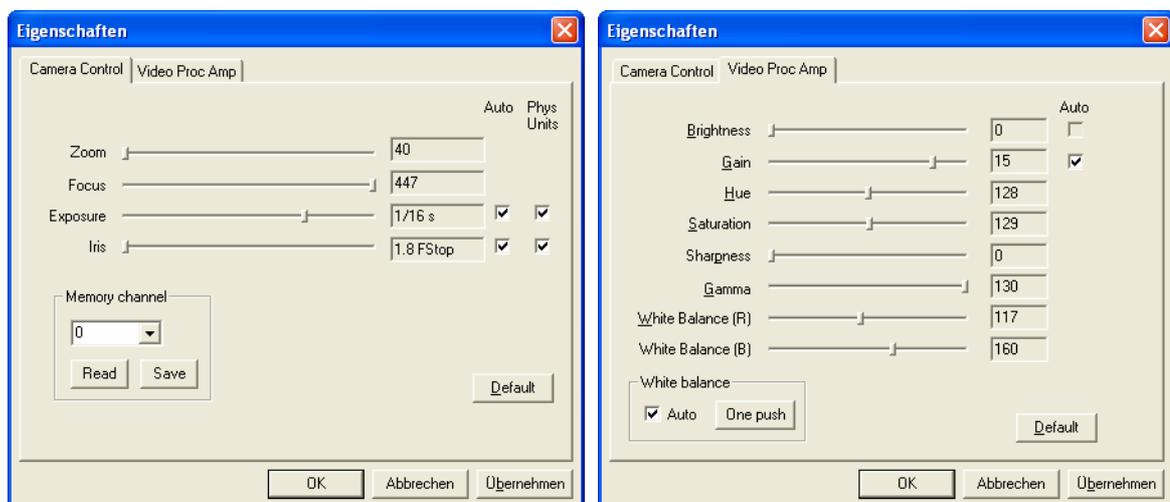


Abbildung 9.2: Eigenschaften des DFW-VL500 Treibers

Der Kameratreiber bietet einen Konfigurationsdialog zum Einstellen aller Kameraparameter (Abbildung 9.2), wobei es möglich ist, alle Parameter numerisch exakt einzustellen. Über den *WDM Capture Filter* kann die Kamera in DirectShow angesprochen werden. Dieser Filter stellt die Interfaces *IAMCameraControl* und *IAMVideoProcAmp* zur Verfügung, um die Parameter programmgesteuert verändern zu können. Da diese Interfaces nicht den vollen Funktionsumfang der Kamera unterstützen, wurden von Sony zusätzliche Konstanten definiert. Über das *IKsPropertySet*-Interface lassen sich weitere Funktionen ansprechen, die von den standardmäßigen DirectShow-Interfaces nicht abgedeckt werden. Unter anderem ist damit ein direkter Zugriff auf die Kameraregister möglich.

## 9.2 Stereo Multiplexer

Wie bereits in Kapitel 5.1 dargelegt, ist es sinnvoll, den linken und rechten Datenstrom zum frühest möglichen Zeitpunkt zu kombinieren und gemeinsam weiterzuverarbeiten. Die beiden Kameras stehen unter DirectShow als zwei Instanzen des *WDM Capture Filters* zur Verfügung, wobei der WDM Capture Filter unter dem Namen des jeweiligen Aufnahmegeräts aufscheint.

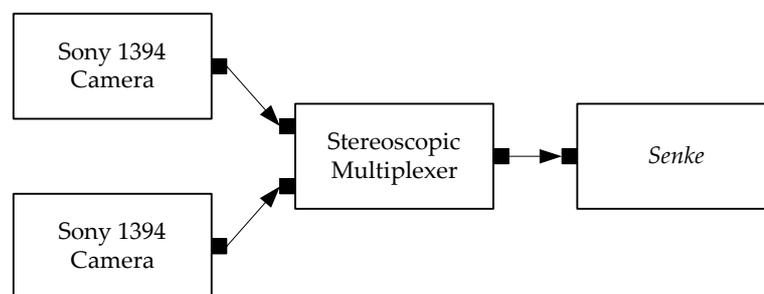


Abbildung 9.3: Aufnahmegraph

Das Ziel ist ein DirectShow-Graph wie in Abbildung 9.3 gezeigt: Ein Multiplexer nimmt den Datenstrom von linker und rechter Kamera entgegen und gibt ihn in kombinierter Form an den Downstreamfilter, symbolisiert durch die Senke, weiter. Die Senke steht stellvertretend für eine Reihe von anderen Filtern, z.B. einen Video-Renderer oder eine Kombination aus *AVI Mux* und *File Writer*, um das stereoskopische Video auf Festplatte aufzunehmen. Mit dem *Smart Tee* Filter ist die gleichzeitige Aufnahme und Vorschau möglich.

Ein derartiger Multiplexer ist im Standardumfang von DirectShow nicht enthalten und musste selbst entwickelt werden. Daher entstand der *Stereoscopic Multiplexer*, enthalten in der DLL *Sm.ax*, der Videodaten in den Formaten YUY2, UYVY, RGB 24 und RGB 32 entgegen nimmt und in Nebeneinanderanordnung (Kreuzblick) am Ausgang zur Verfügung stellt. In der derzeitigen Implementierung muss das Format der beiden Eingangspins mit dem Format des Ausgangspins übereinstimmen. Selbstverständlich muss auch die Auflösung an den Eingangspins identisch sein.

Der Stereoskopische Multiplexer wird durch die Klasse *CSM* realisiert, abgeleitet von *CBaseFilter*. Die Pins werden durch *CSMInputPin* und *CSMOutputPin* realisiert, abgeleitet von *CBaseInputPin* und *CBaseOutputPin*.

Der Stereoscopic Multiplexer wertet die Zeitstempel der empfangenen Frames aus, um den Zeitversatz zwischen linkem und rechtem Halbbild zu minimieren. Es werden jeweils die Halbbilder mit dem geringsten Versatz kombiniert. Kann zu einem Halbbild kein zeitlich passendes Halbbild des anderen Streams gefunden werden (die Startzeitpunkte der Bilder dürfen nicht mehr als die Hälfte der Präsentationsdauer auseinanderliegen), wird es verworfen. Dadurch wird vermieden, dass die Synchronität verloren geht, falls eine der Quellen nicht zuverlässig alle Frames liefert, z.B. wenn die Hardware an ihrem Leistungslimit arbeitet. Der verwendete Algorithmus setzt voraus, dass beide Eingangstreams dieselbe Framerate und damit gleiche Präsentationsdauer pro Frame aufweisen. Andernfalls kommt es zu unvorhersehbarem Verhalten – im schlimmsten Fall werden alle Bilder verworfen.

Trotz Berücksichtigung der Zeitstempel ist nicht garantiert, dass die resultierenden Halbbildpaare tatsächlich zum gleichen Zeitpunkt aufgenommen wurden. Zum einen werden auch Frames mit abweichenden Startzeitpunkten kombiniert, sofern die Abweichung unter der Hälfte der Präsentationsdauer liegt. Zum anderen werden die Zeitstempel vom Aufnahmetreiber vergeben und müssen nicht unbedingt in direktem Zusammenhang mit dem tatsächlichen Aufnahmezeitpunkt stehen.

### 9.3 Programmintegration

Der Stereo Multiplexer kann mit Hilfe von GraphEdit flexibel für unterschiedliche Aufgaben eingesetzt werden. Allerdings ist es von Vorteil, wenn für typische Aufgaben wie die Aufnahme und Übertragung von 3D-Videos eine spezialisierte Anwendung zur Verfügung steht. Es gibt zwei Möglichkeiten, diese Anforderung zu erfüllen.

1. Entwicklung von Anwendungen, die auf die Aufnahme und Verarbeitung von 3D-Videos zugeschnitten sind. Vorteil dieser Variante ist die bestmögliche Anpassung der Benutzeroberfläche an die Bedürfnisse der Stereoskopie. Als Nachteile sind der hohe Entwicklungsaufwand und die geringe Flexibilität bei neuen Anforderungen zu nennen.
2. Integration des Stereoscopic Multiplexers in bestehende Anwendungen. Sobald linke und rechte Halbbilder zu einem einzigen Videostrom wurden, unterscheidet sich die Kompression, Übertragung und Speicherung von 3D-Videos in keiner Weise von 2D-Videos. Es liegt also nahe, ursprünglich für gewöhnliche Videos entwickelte Software für stereoskopische Videos einzusetzen.

Die zweite Variante stellt zweifellos die elegantere Methode dar, da sie die große Anzahl bestehender Anwendungen für den 3D-Einsatz öffnet. Allerdings sind nur die wenigsten Programme im Quellcode verfügbar, um den Stereoscopic Multiplexer und die Unterstützung der zweiten Kamera zu integrieren.

Die Anpassung bestehender Programme wird allerdings überflüssig, wenn der Stereoscopic Multiplexer direkt als Aufnahmeggerät verwendet werden kann. Die Realisierung dieser Idee erfordert einerseits, dass der Stereoscopic Multiplexer als Aufnahmeggerät erkannt wird, andererseits muss er in der Lage sein, ohne Mithilfe der Anwendung die echten Aufnahmeggeräte nachzuladen.

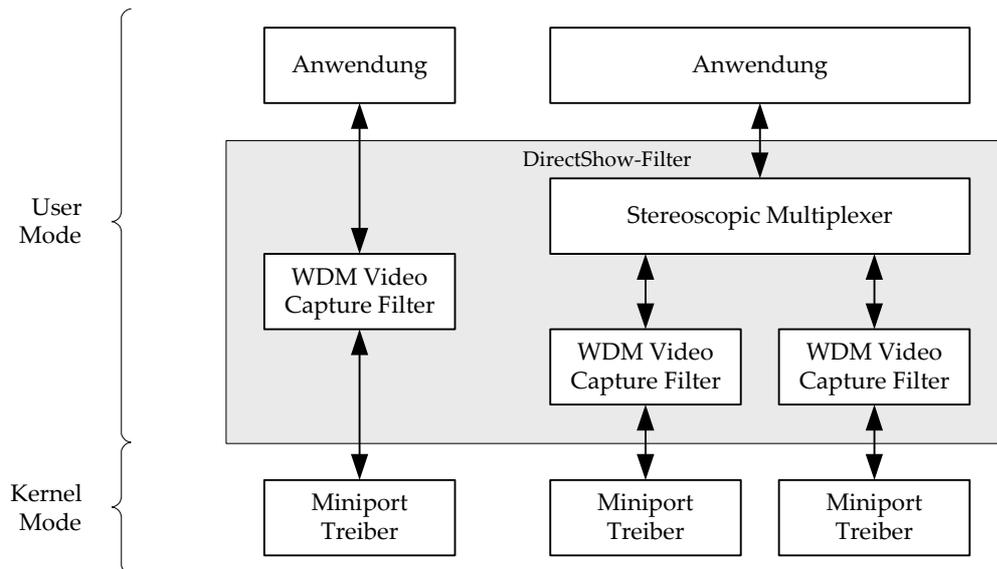


Abbildung 9.4: Zugriff auf Aufnahmegeräte

Die linke Hälfte von Abbildung 9.4 illustriert den Zugriff einer Anwendung auf ein Aufnahmegerät. Sofern die Anwendung auf DirectShow basiert, lädt sie eine Instanz des WDM Video Capture Filters. Die Anwendung kommuniziert ausschließlich mit dem DirectShow-Filter, der darunter liegende Treiber bleibt ihr verborgen. Aufgrund dieser Zugriffshierarchie akzeptiert die Anwendung anstelle des WDM Video Capture Filter den Stereoscopic Multiplexer, obwohl dieser kein Hardwaregerät kapselt. Damit der Stereoscopic Multiplexer als Aufnahmegerät erkannt wird, muss er in der Systemregistrierung entsprechend eingetragen werden und alle wichtigen Schnittstellen des WDM Video Capture Filter anbieten, konkret das `IKsPropertySet`- und `IAMStreamConfig`-Interface für den Ausgangspin. Der Stereoscopic Multiplexer wird von allen Anwendungen erkannt, die Aufnahmegeräte mittels DirectShow ansprechen. Alte Anwendungen, die noch auf Video für Windows basieren, können nicht verwendet werden, ebenso solche, die nur DV-Geräte unterstützen.

Der Stereoscopic Multiplexer wird von der Anwendung in den Graphen eingefügt, die beiden echten Aufnahmegeräte jedoch nicht. Damit bleiben die Eingangspins des Multiplexers vorerst unverbunden. Es ist dessen Aufgabe, beide Geräte bei Bedarf nachzuladen und in den Graphen einzufügen. In der derzeitigen Implementierung geschieht dies, sobald die Anwendung den Ausgangspin verbindet. Wird die Verbindung des Ausgangspins getrennt, werden die Filter an den Eingängen ebenfalls wieder entfernt. Anhand Abbildung 9.4 sieht man, dass der Stereoscopic Multiplexer im User-Mode des Betriebssystems arbeitet und dadurch die Systemstabilität nicht gefährden kann. Systemabstürze können nur von den Miniport-Treibern im Kernel-Mode verursacht werden.

Das Videoformat der Eingangs- und des Ausgangspins sowie das linke und rechte Aufnahmegeräte müssen vom Benutzer im Konfigurationsdialog des Multiplexers festgelegt werden (siehe Abbildung 9.5). Wird das gewählte Format von den Aufnahmegeräten nicht direkt unterstützt, wird mit zusätzlichen Transformationsfiltern versucht, trotzdem eine Verbindung herzustellen. Aus Performancegründen ist es empfehlenswert, ein von den Treibern direkt unterstütztes Format auszuwählen. In manchen Fällen sind Transformationsfilter aber unvermeidlich, z.B. bei Verwendung von DV-Kameras: Die komprimierten Datenströme können vom Stereoscopic Multiplexer nicht verarbeitet werden, sondern müssen zuerst durch je einen

DV Video Decoder in eines der unterstützten Formate gewandelt werden. Das Einfügen der zusätzlichen Filter geschieht automatisch und unsichtbar für den Anwender. Im Allgemeinen ist den UYVY- und YUY2-Formaten gegenüber RGB 24 und RGB 32 aus Performancegründen der Vorzug zu geben, außer die Kameras unterstützen nur RGB.

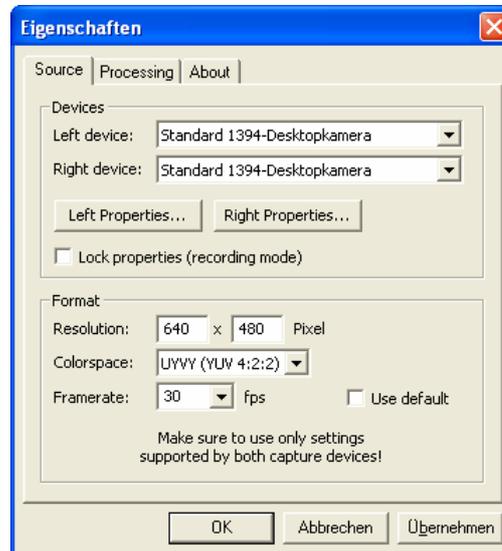


Abbildung 9.5: Konfigurationsdialog des Stereoscopic Multiplexers

## 9.4 Untersuchung der Synchronisation

### 9.4.1 Messanordnung

Zur Messung der Synchronisation müssen beide Kameras eine zeitlich genau definierte Bewegung aufnehmen, sodass in der Folge aus dem Versatz der aufgezeichneten Bewegungsphasen auf die Differenz der Auslösezeitpunkte rückgerechnet werden kann.

In der Praxis hat es sich bewährt, mit kurzer Belichtungsdauer ein Testbild von einem Röhrenmonitor abzufilmen. Das Zeitverhalten des Elektronenstrahls ist mit ausreichender Genauigkeit bekannt. Wenn die Belichtungsdauer unter der Dauer eines Bildaufbauzykluses liegt, ist auf der Aufnahme nur ein Teil des Testbildes sichtbar, der restliche Bereich erscheint schwarz. Sind beide Kameras synchronisiert, ist auf beiden Aufnahmen jeweils derselbe Teilbereich sichtbar, ansonsten kann aus dem vertikalen Versatz auf den zeitlichen Versatz der Aufnahmezeitpunkte rückgerechnet werden.

Dazu verwendet man ein Testbild wie in Abbildung 9.6 gezeigt, das ein einfaches Ablesen der Zeilen ermöglicht. Allerdings lässt sich anhand eines einzelnen Testbildes nicht festlegen, welchem Bildaufbauzyklus eine Aufnahme entstammt. Eine zusätzliche Kennzeichnung der Testbilder behebt das Problem, ist aber schwierig zu realisieren, da die Kennung mit jedem Bildaufbauzyklus geändert werden muss. Zur Durchführung der Messungen wurde der *Camera Sync Tester* entwickelt, der fünf Testbilder, die sich durch die Anzahl der vertikalen Balken unterscheiden, zyklisch wiederholt. Der *Camera Sync Tester* verwendet den exklusiven Vollbildmodus von DirectX 7, um den Wechsel nach jedem Bildaufbauzyklus zu ermöglichen. Es gibt zwar keine Garantie, dass die Bilder tatsächlich störungsfrei in der geforderten

Frequenz angezeigt werden, auf allen getesteten Systemen funktionierte das Programm aber einwandfrei. Man kann davon ausgehen, dass der Camera Sync Tester störungsfrei läuft, sofern die Grafikkarte DirectX 7 hardwareseitig unterstützt.

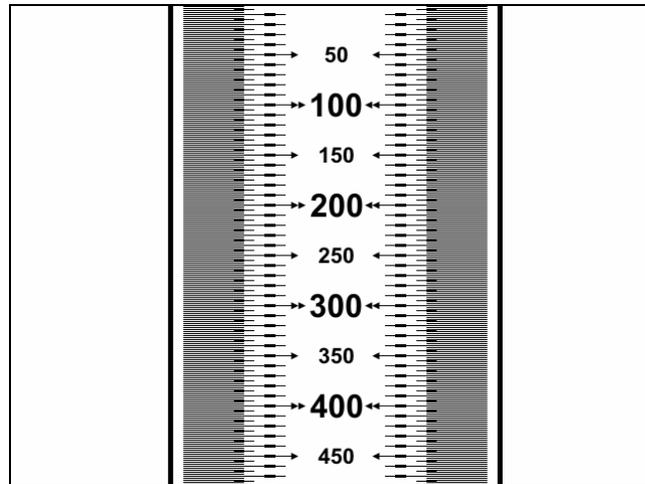


Abbildung 9.6: Testbild zur Messung der Kamerasynchronisation

Abbildung 9.7 zeigt das Testbild, mit den beiden Firewire-Kameras und dem Stereoscopic Multiplexer vom Monitor abgefilmt. Beide Aufnahmen weisen zwei vertikale Balken auf, entstammen also demselben Bildaufbauzyklus.

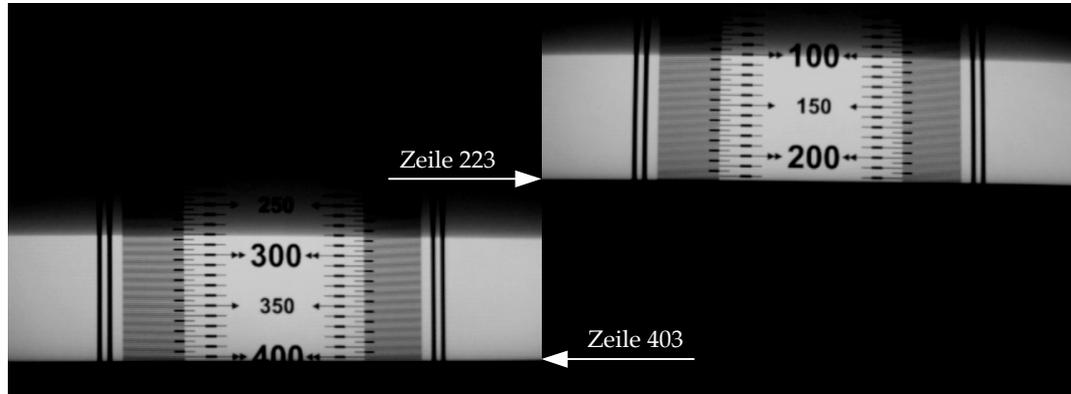


Abbildung 9.7: Testbild, mit zwei Kameras vom Monitor abgefilmt

Zur Berechnung der Zeilendifferenz müssen die jeweils letzten sichtbaren Zeilen  $z_1$  und  $z_2$  aus den Aufnahmen abgelesen werden. Es ist wichtig, die untere Begrenzung der Ausschnitte zu vergleichen, da der obere Rand durch das Nachleuchten des Monitorphosphors nicht ausreichend scharf abgegrenzt wird. Die Dauer, die der Elektronenstrahl zur Darstellung einer Zeile benötigt, ergibt sich direkt aus der Horizontalfrequenz des Monitorsignals. Die meisten Bildschirme geben in ihrem Bildschirmmenü Auskunft über die aktuelle Horizontal- und Vertikalfrequenz. Die Zeilendauer multipliziert mit der Zeilendifferenz ergibt den gesuchten Synchronisationsoffset  $t_{Offset}$  der Kameras.

$$t_{\text{Zeile}} = \frac{1}{f_{\text{Horiz}}} \quad (9.1)$$

$$t_{\text{Offset}} = \frac{z_2 - z_1}{f_{\text{Horiz}}} \quad (9.2)$$

Zeigen die beiden Aufnahmen unterschiedliche Bildwiederholzyklen, erkennbar an einer unterschiedlichen Anzahl vertikaler Balken, muss  $t_{\text{Offset}}$  um ein entsprechendes Vielfaches der Bilddauer korrigiert werden. Die Bilddauer  $t_{\text{Bild}}$  folgt aus der Vertikalfrequenz,  $d_{\text{Bild}}$  bezeichnet die Differenz der Bildwiederholzyklen.

$$t_{\text{Bild}} = \frac{1}{f_{\text{Vert}}} \quad (9.3)$$

$$t_{\text{Offset}} = \frac{z_2 - z_1}{f_{\text{Horiz}}} + \frac{d_{\text{Bild}}}{f_{\text{Vert}}} \quad (9.4)$$

## 9.4.2 Ergebnisse

Alle Messungen wurden mit den Sony DFW-VL500 Kameras bei einer Auflösung von 640 x 480 Pixel bei 30 Bildern pro Sekunde und einer Belichtungsdauer von 1/256 Sekunde durchgeführt. Zur Erfassung der Halbbildpaare wurde der Stereoscopic Multiplexer im Stereoscopic Player geladen. Zum Ablesen der Werte wurde die Wiedergabe pausiert.

Im ersten Versuch wurde der Synchronisationsoffset nach gleichzeitigem Einschalten der Kameras gemessen. Dazu wurden die Kameras an der internen Schnittstelle eines Notebooks und einer zusätzlichen PC Card betrieben. Da die 4-polige interne Schnittstelle nicht in der Lage ist, die Kamera mit Strom zu versorgen, wurde ein Hub mit eigenem Netzteil zwischengeschaltet. Die PC Card wurde ebenfalls von einem eigenen Netzteil versorgt. Vor jeder Messung wurden beide Netzteile über eine gemeinsame Steckdosenleiste gleichzeitig aus- und eingeschaltet. Es ergab sich ein durchschnittlicher Offset von 10,55 ms mit einer mittleren Abweichung von nur 0,25 ms. Daraus folgt, dass sich durch gleichzeitiges Einschalten eine feste Beziehung der Aufnahmezeitpunkte innerhalb enger Grenzen herstellen lässt, allerdings mit einem Offset von 10 ms und nicht dem Idealwert von 0 ms. Der erzielte Wert ist trotzdem für die meisten Aufnahmesituationen ausreichend, insbesondere in Anwendungsszenarien mit hauptsächlich bewegungsarmen Szenen wie z.B. Videokonferenzen.

In der folgenden Messung wurde versucht, der Ursache für die Verzögerung der zweiten Kamera von 10 ms auf den Grund zu gehen. Dazu wurden die Kameras an der jeweils anderen Schnittstelle betrieben. Das Ergebnis (Mittelwert 9,61 ms, mittlere Abweichung 0,44 ms) ist ähnlich zur vorhergehenden Messung. Es erlaubt den Rückschluss, dass die Abweichung durch die Kameras selbst und nicht durch die verwendete Schnittstellenkarte oder den Hub verursacht wird. Eine weitere Messung an einem anderen PC mit zwei integrierten OHCI-Controllern am PCI-Bus bestätigt diese Annahme.

Senkt man die Bildwiederholrate von 30 auf 15 Bilder pro Sekunde ergibt sich ein anderes Bild: Zwar stellt sich manchmal wiederum ein Offset von rund 10 ms ein, in den meisten Fällen liegt er aber bei durchschnittlich -23,39 ms mit einer mittleren Abweichung von 0,26 ms. Die Differenz dieser beiden Werte beträgt ziemlich genau 33,33 ms und entspricht damit der Periodendauer des internen Kamerataktes. Bei einer Bildwiederholrate von 15 Bildern pro Sekunde entscheidet der Zufall darüber, ob die Bildsensoren während korrespondierender

Zyklen ausgelesen werden. Nach einer weiteren Reduzierung der Bildrate auf 7,5 oder 3,75 Bilder pro Sekunde ist der Versatz bei Betrachtung einer raschen Bewegung bereits bei normaler Betrachtung der Halbbildpaare offensichtlich, sodass sich diese Modi nicht für die Stereoaufnahme eignen (ausgenommen spezielle Aufnahmesituationen mit extrem langer Belichtungsdauer).

Neben dem Synchronisationsoffset nach dem Einschalten ist auch die Drift nach längerer Betriebsdauer von Interesse. Dazu wurde über einen Zeitraum von einer Stunde alle 15 Minuten der Offset ermittelt. Innerhalb dieser Stunde änderte sich der Offset um 16,3 ms, pro Minute also um 0,27 ms.

Sowohl der Synchronisationsoffset bei gleichzeitigem Einschalten als auch die Drift dürften durch Bauteiltoleranzen in der Kameraelektronik verursacht werden und gelten damit nur für das vermessene Kamerapaar. Werden die Kameras zu unterschiedlichen Zeitpunkten eingeschaltet, ergibt sich eine zufällige Verteilung des Synchronisationsoffsets, alle Messwerte lagen aber innerhalb des Intervalls -16,67 ms bis 16,67 ms, sodass man davon ausgehen kann, dass die Zeitstempel der Mediasamples ausreichend gut mit den Aufnahmezeitpunkten korrelieren und der Stereoscopic Multiplexer somit in der Lage ist, den maximalen zeitlichen Versatz unter 1/60 Sekunde zu halten.

## 9.5 Parametersynchronisation

### 9.5.1 Synchronisationsmechanismen

Ein fehlerfreies Raumbild erfordert eine exakte Einstellung der Kameras, um Abweichungen in Helligkeit, Farbe und Kontrast zwischen den Halbbildern zu vermeiden. Der Konfigurationsdialog der Sony DFW-VL500-Kameras bietet die Möglichkeit, alle relevanten Bildparameter zu verändern. Der Dialog der linken bzw. rechten Kamera kann über den Konfigurationsdialog des Stereoscopic Multiplexer aufgerufen werden. Somit ergibt sich folgender Arbeitsablauf:

1. Einstellung der ersten Kamera um ein optimales Bild zu erzielen.
2. Wählen der gleichen Einstellungen für die zweite Kamera.

Die große Anzahl der Parameter macht den zweiten Schritt aufwendig sowie fehleranfällig und weckt den Wunsch nach einer automatischen Übertragung der Parameter. Es bietet sich an, die Parameter der ersten Kamera über das IAMCameraControl- bzw. IAMVideoProcAmp-Interface auszulesen und auf der zweiten Kamera zu setzen. Eine weitere Vereinfachung ergibt sich durch die Möglichkeit, die Parameter direkt im Stereoscopic Multiplexer einzustellen und damit den Umweg über den Dialog des Kameratreibers zu vermeiden.

Die soeben beschriebene, unidirektionale Übertragung der Parameter von Kamera 1 nach Kamera 2 erweist sich in der Praxis als zu eingeschränkt. Da sich die Parameter Zoom und Fokus durch Tasten an der Kamera direkt verändern lassen und auch eine Änderung der restlichen Parameter der zweiten Kamera über den Konfigurationsdialog des Treibers nicht ausgeschlossen werden kann, ist eine bidirektionale Synchronisierung wünschenswert: Egal ob der Parameter der ersten oder zweiten Kamera variiert wird, die Änderung sollte auf die jeweils andere Kamera übertragen werden. Das Verhalten muss allerdings noch von der Art des Parameters abhängig gemacht werden.

- *Manuelle Parameter* ändern ihren Wert nur durch einen Eingriff des Anwenders, sei es durch Betätigen einer Taste an der Kamera oder über den Konfigurationsdialog. Manuelle Parameter werden vom Stereoscopic Multiplexer uneingeschränkt bidirektional synchronisiert, sofern die entsprechende *Sync*-Option für den jeweiligen Parameter aktiviert wurde. Eine Kopie des Parameterwerts wird in einer Variablen gespeichert und periodisch mit den Werten der Kameras verglichen. Ändert sich der Wert einer Kamera, wird er sowohl in die Variable als auch die andere Kamera übernommen.
- *Automatische Parameter* werden von der Kamera selbstständig gewählt. Typischerweise kann die Belichtungssteuerung, der Weißabgleich und der Fokus in den automatischen Modus versetzt werden. Die verwendeten Sony DFW-VL500-Kameras bieten zwar eine automatische Belichtungssteuerung und einen automatischen Weißabgleich, allerdings keinen Autofokus.

Wird der Automatikmodus auf beiden Kameras aktiviert, ergeben sich häufig leichte Unterschiede zwischen den Halbbildern, da die Algorithmen aufgrund unterschiedlicher Eingangsdaten und Störeinflüssen zu unterschiedlichen Ergebnissen gelangen. Der Stereoscopic Multiplexer umgeht diese Problematik, indem nur die linke Kamera automatische Parameter verwendet. Die Parameter der rechten Kamera werden im manuellen Modus belassen und unidirektional mit der linken Kamera synchronisiert. Die unidirektionale Synchronisierung automatischer Parameter setzt voraus, dass die Kamera und der Treiber auch im automatischen Modus in der Lage sind, den aktuellen Wert des Parameters zu ermitteln. Die verwendeten Kameras erfüllen diese Voraussetzung.

- *One Push Parameter* sind eine Besonderheit des Sony-Treibers [67]. Sie sind weder im `IAMCameraControl`- noch im `IAMVideoProcAmp`-Interface vorgesehen. Über eine zusätzlich definierte Konstante kann der One Push Modus trotzdem über diese Interfaces aktiviert werden. Das Setzen des One Push Modus bewirkt eine einmalige automatische Wahl des jeweiligen Parameters, danach wird in den manuellen Modus zurück gewechselt. Der Stereoscopic Multiplexer aktiviert den One Push Modus nur für Parameter der linken Kamera. Da das One Push Flag beim nächsten Abruf des Parameterwertes vom Treiber entfernt wird, wechselt der Stereoscopic Multiplexer daraufhin in den manuellen Modus. Die verwendeten Kameras unterstützen den One Push Modus nur für den Weißabgleich.

Die Implementierung der Parametersynchronisation ist allgemein gehalten und sollte mit allen Kameratypen funktionieren, die einen Zugriff auf die Parameter über die standardisierten `DirectShow`-Schnittstellen anbieten. Abschließend sei erwähnt, dass der Abgleich von DV Kameras mangels Zugriff auf die internen Kamerazustände nicht möglich ist. Zwar definiert der AV/C-Standard [1] einen Befehlssatz zur Steuerung von Kameras, allerdings wurde die Möglichkeit, Parameter numerisch exakt auszulesen und zu setzen, nicht vorgesehen.

## 9.5.2 Parameterkorrektur

Erste Tests des Stereoscopic Multiplexers zeigten, dass trotz exakter Übereinstimmung der Parameterwerte Farb- und Helligkeitsabweichungen der Halbbilder auftraten. Die Ursache hierfür ist in Toleranzen bei der Herstellung und dem werksseitigen Abgleich der Kameras zu suchen. Die größten Abweichungen traten bei der Helligkeit zu Tage und es zeigte sich, dass sie durch die Wahl unterschiedlicher Iriswerte korrigiert werden können. In der Folge wurden bei gleicher Wahl aller anderen Parameter eine Reihe von Irispaarungen ermittelt, die zu

identischer Helligkeit führen und diese in ein Diagramm eingetragen. Für jede Messung wurden die Iriswerte so lange manuell verändert, bis keine Helligkeitsabweichung mehr wahrnehmbar war. Während die Beurteilung von Helligkeitsunterschieden in strukturierten Bildteilen sehr schwer fällt, gelingt sich am Übergang zu überbelichteten Bildteilen leicht, indem man die Größe der gesättigten Fläche vergleicht. Zur Ermittlung der einzelnen Messpunkte wurde daher hauptsächlich auf überlichtete Bildteile geachtet, die durch entsprechende Wahl der anderen Kameraparameter erzwungen wurden.

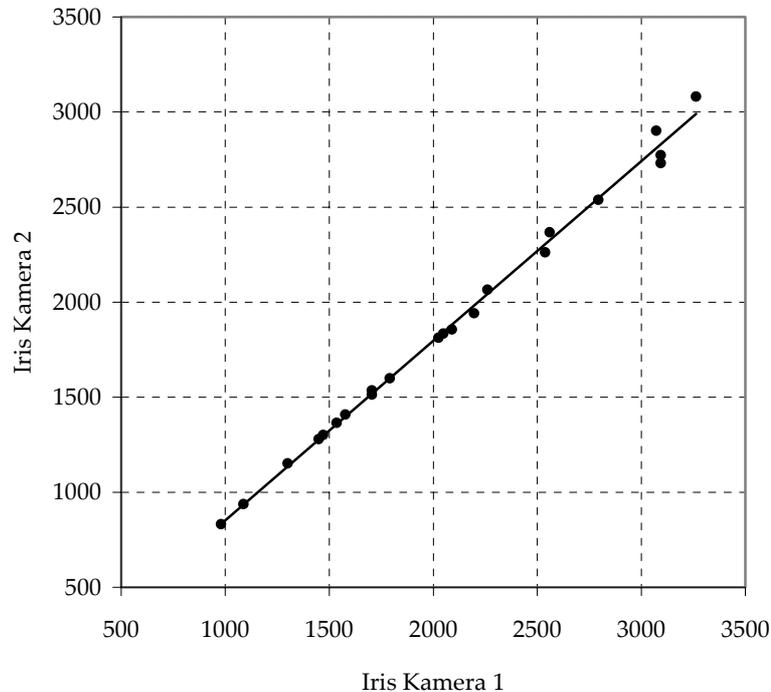


Abbildung 9.8: Korrekturkennlinie des Iris-Parameters

Anhand Abbildung 9.8 erkennt man, dass Iris-Werte von Kamera 2 etwas niedriger als jene von Kamera 1 gewählt werden müssen. Außerdem ist ersichtlich, dass offenbar ein linearer Zusammenhang vorliegt, der sich durch die Gerade

$$f(x) = 0,9452 \cdot x - 93,5 \quad (9.5)$$

gut approximieren lässt. Damit ist eine Korrekturfunktion für den Iris-Parameter gefunden.

Zur effizienten Handhabung der Parameterkorrektur ist es unumgänglich, dass diese vom Stereoscopic Multiplexer durchgeführt wird. Da eine feste Kodierung der Korrekturfunktion  $f(x)$  die Verwendbarkeit des Multiplexers mit anderen Kameras einschränken würde, wird die Korrekturfunktion in der Systemregistrierung hinterlegt. Da eine Gerade zur Beschreibung einer Korrekturfunktion im Allgemeinen nicht ausreicht, können komplexere Zusammenhänge in Form von stückweise linearen Funktionen beschrieben werden. Es besteht die Möglichkeit, für jeden Parameter eine Wertetabelle zu definieren, zwischen den vorgegebenen Punkten wird linear interpoliert.

Die Einstellungen des Stereoscopic Multiplexers werden unter `HKEY_LOCAL_MACHINE\Software\WimmerSoft\Stereoscopic Capture Driver` gespeichert. Die Schlüssel

CameraControl und ProcAmp enthalten weitere Unterschlüssel, die der Parameternummer gemäß den Definitionen des IAMCameraControl und des IAMVideoProcAmp-Interfaces entsprechen. In diesen Unterschlüsseln wird die Wertetabelle in Form von DWORD-Einträgen festgelegt.

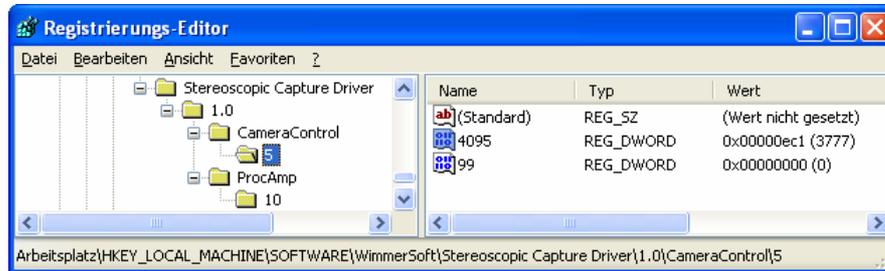


Abbildung 9.9: Tabelle zur Parameterkorrektur

Abbildung 9.9 zeigt die notwendigen Einträge für die zuvor erläuterte Korrekturfunktion des Iris-Parameters. Der rechte Randwerte des Regelbereich, 4095, wird auf 3777 umgesetzt, der linke Randwert 0 ergibt sich als Funktionswert von 99. Werte außerhalb des Regelbereichs werden, falls notwendig, begrenzt und sind daher zulässig. Wurde die Funktion nicht auf dem gesamten Regelbereich definiert, wird der erste oder letzte Funktionsabschnitt linear fortgesetzt.

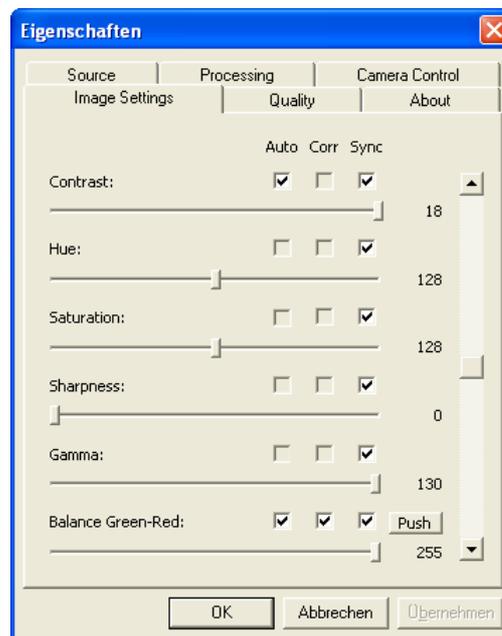


Abbildung 9.10: Eigenschaftsseite zum Einstellen der Bildparameter

Die bidirektionale Parametersynchronisation (siehe Abbildung 9.11) erfordert die inverse Korrekturfunktion  $f^{-1}(x)$ . Der Stereoscopic Multiplexer kann die eindeutige, inverse Korrekturfunktion nur dann bilden, wenn die Korrekturfunktion bijektiv und somit streng monoton steigend oder fallend ist. Nicht bijektive Funktionen führen zu Programmfehlern oder unvorhersehbarem Verhalten und dürfen nicht verwendet werden.

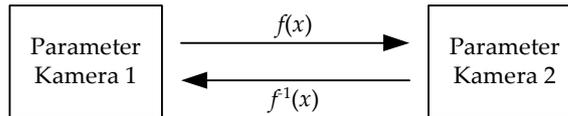


Abbildung 9.11: Bidirektionale Parametersynchronisation

Die Parametersynchronisation lässt sich über die *Sync*-Option für jeden Parameter einzeln steuern (siehe Abbildung 9.10, Seite 29). Bei deaktiviertem *Sync* unterbleibt die periodische Überprüfung auf Änderungen des Parameterwerts, sodass über die Konfigurationsdialoge der beiden Treiber unterschiedliche Werte gewählt werden können. Das Abwählen der *Corr*-Option unterbindet nur die Parameterkorrektur. Die Option *Lock properties (recording mode)* unterbindet ebenfalls die Parametersynchronisation, versetzt zuvor aber alle Parameter in den manuellen Modus, um Einstellungsänderungen während der Aufnahme zu verhindern.

## 9.6 Bildverarbeitungsfunktionen

Leider erwies sich die interne Bildverarbeitung der Sony DFW-VL500-Kameras teilweise als unzulänglich. Zum einen verursacht die Gammakorrektur Farbverfälschungen – helle Objekte erscheinen deutlich zu bunt und weisen sprunghafte Farbänderungen auf, weiters führt die Schärfungsfunktion zu einem Bild mit vertikalen Streifen und verschlechtert den Bildeindruck, anstatt ihn zu verbessern. Sowohl Gammakorrektur als auch Schärfungsfunktion mussten durch entsprechende Routinen im Stereoscopic Multiplexer ersetzt werden. Die Gammakorrektur hellt die Bilddaten gemäß der Funktion

$$f(x) = x^{\frac{1}{\gamma}} \quad \text{für } 0 \leq x \leq 1 \quad (9.6)$$

für die Darstellung auf Monitoren mit nichtlinearer Kennlinie auf. Windows-Systeme arbeiten mit einem Gammawert von 2,2, Apple-Systeme mit 1,8 – entsprechend sollte der Gammawert  $\gamma$  in dieser Größenordnung gewählt werden. Aus Performancegründen werden die Funktionswerte der Gammafunktion in einer Lookuptable abgespeichert, sodass sich die Transformation eines Pixels auf drei Lookuptabellenzugriffe beschränkt – je einen Zugriff für die Rot-, Grün- und Blaukomponente.

$w(-1, -1)$	$w(-1, 0)$	$w(-1, 1)$
$w(0, -1)$	$w(0, 0)$	$w(0, 1)$
$w(1, -1)$	$w(1, 0)$	$w(1, 1)$

0	-1	0
-1	5	-1
0	-1	0

Abbildung 9.12: Allgemeine 3x3 Filtermaske und spezielle Maske zur Unschärfmaskierung

Die Bildschärfung wird mit Hilfe der Unschärfmaskierung im Ortsbereich vorgenommen [21]. Sie wird in der derzeitigen Implementierung nur für UYVY- und YUY2-Formate unterstützt. Dazu wird die 3x3-Filtermaske  $w(x, y)$  auf die Helligkeitskomponente angewandt. Aus den Eingangswerten  $f(x, y)$  berechnen sich die Ausgangswerte  $g(x, y)$  gemäß Beziehung (9.7).

$$g(x, y) = \sum_{s=-1}^1 \sum_{t=-1}^1 w(s, t) \cdot f(x + s, y + t) \quad (9.7)$$

Alle DirectShow-YUV-Formate basieren auf dem Standard ITU-R BT.601. Er definiert einen Farbraum mit Helligkeitswerten im Bereich von 16 bis 235 und Farbwerten im Bereich von 16 bis 240. Die Umrechnung von RGB nach YUV erfolgt über Beziehung (9.8) [70].

$$\begin{aligned} K_r &= 0,299 \\ K_b &= 0,114 \\ L &= K_r \cdot R + K_b \cdot B + (1 - K_r - K_b) \cdot G \\ Y &= \frac{219 \cdot L}{255} + 16 \\ U &= \frac{224 \cdot 0,5 \cdot (B - L)}{(1 - K_b) \cdot 255} + 128 \\ V &= \frac{224 \cdot 0,5 \cdot (R - L)}{(1 - K_r) \cdot 255} + 128 \end{aligned} \quad (9.8)$$

In Matrizenschreibweise ergibt sich (gerundet):

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,257 & 0,504 & 0,098 \\ -0,148 & -0,291 & 0,439 \\ 0,439 & -0,368 & -0,071 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} \quad (9.9)$$

Die Sony DFW-VL500-Kameras nutzen sowohl für die Helligkeits- als auch die Farbkomponenten den vollen 8-Bit-Wertebereich von 0 bis 255 aus. In diesem Fall dient Beziehung (9.10) zur Umrechnung zwischen YUV und RGB-Komponenten.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ -0,1687 & -0,3313 & 0,5 \\ 0,5 & -0,4187 & -0,0813 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \quad (9.10)$$

Die Option *Convert to ITU-R BT.601* veranlasst den Stereoscopic Multiplexer, die Eingangsdaten mit vollem Wertebereich zu interpretieren und im ITU-R BT.601-Farbraum auszugeben. Ohne diese Konvertierung würden durch den erhöhten Kontrast Details in sehr dunklen und hellen Bereichen verloren gehen.

Eine weitere Bildverarbeitungsfunktion des Stereoscopic Multiplexers ist die Einstellung des Schwarz- und Weißlevels zur Anpassung des Kontrasts. Beide Funktionen werden durch eine Modifikation der Lookuptabelle realisiert und erfordern keinen zusätzlichen Rechenaufwand.

Darüberhinaus können sowohl das linke als auch das rechte Halbbild horizontal und vertikal gespiegelt werden. Diese Funktionen werden benötigt, wenn aus Platzgründen eine Kamera kopfüber montiert wird oder Konstruktionen mit (halbdurchlässigen) Spiegeln zur Reduzierung der Stereobasis eingesetzt werden.

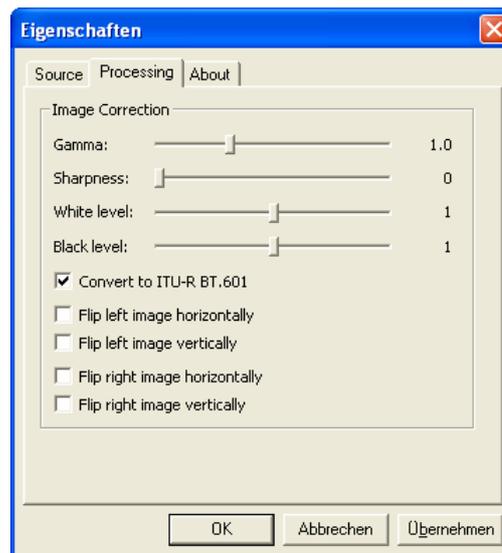


Abbildung 9.13: Bildverarbeitungsoptionen des Stereoscopic Multiplexers

---

# Streaming

---

*Dieses Kapitel beschreibt die Integration der Komponenten aus den vorigen Kapiteln zu einem Übertragungssystem für stereoskopische Videos, dem erklärten Ziel dieser Diplomarbeit. Zuvor werden aber die Möglichkeiten zur Übertragung stereoskopischer Videos im Allgemeinen beleuchtet.*

## 10.1 Übertragungsvarianten

Stereoskopische Videos können (oder müssen manchmal sogar) mit den Verfahren für zweidimensionale Inhalte übertragen werden. Prinzipiell gibt es zwei grundlegende Varianten zur Übertragung stereoskopischer Videos:

- Man verwendet einen einzelnen Übertragungskanal, der die Halbbilder, in einem Stereoformat kodiert, gemeinsam überträgt. Synchronisierung und Multiplexing werden senderseitig durchgeführt.
- Man verwendet je einen Übertragungskanal pro Perspektive. Synchronisierung und Multiplexing werden empfängerseitig durchgeführt.

Der Autor vertritt die Meinung, dass der ersten Variante sofern möglich der Vorzug gegeben werden sollte (siehe auch Kapitel 5.1). Die Synchronisierung zum frühestmöglichen Zeitpunkt und anschließende kombinierte Übertragung schließt Synchronisationsprobleme aufgrund unterschiedlicher Übertragungsverzögerung, fehlerhaft übertragender Einzelbilder und anderer Unregelmäßigkeiten aus, die sonst mit großem Aufwand kompensiert werden müssten. Die Übermittlung des Timecodes, je nach Medium unterschiedlich schwierig zu realisieren, kann ebenfalls entfallen.

Auf die zweite Variante muss zurückgegriffen werden, wenn der Übertragungskanal Einschränkungen unterliegt und die Verwendung des gewünschten Stereoformats nicht zulässt. Beispielsweise erfordert eine fernsehkompatible Übertragung eine feste Auflösung und Bildwiederholrate, sowohl bei analoger als auch digitaler Ausstrahlung. Entweder muss man sich mit der halben Auflösung pro Halbbild begnügen (in diesem Fall wird das Fieldsequential-Verfahren bevorzugt), oder man verwendet getrennte Kanäle für die linke und rechte Perspektive. Der entscheidende Vorteil der zweiten Variante ist, dass mit einem herkömmlichen Fernsehgerät ein zweidimensionales Bild empfangen werden kann, während für die 3D-Wiedergabe zusätzlich der zweite Kanal empfangen werden muss. Digitales Fernsehen [59] bietet weitere Möglichkeiten, um die zweite Perspektive kompatibel zum 2D-Empfang im Datenstrom unterzubringen. Auf diese wird an dieser Stelle allerdings nicht mehr näher eingegangen.

## 10.2 Realisierung

Ziel dieser Diplomarbeit ist die Übertragung von stereoskopischen Videos über das Internet. Im Gegensatz zu bisherigen Arbeiten auf diesem Gebiet [36] wurde auf die Entwicklung eines eigenen Streamingsservers oder auf Modifikation einer bestehenden 2D-Streaminganwendung verzichtet.

Stattdessen wurde ein wesentlich eleganterer Weg beschritten: Existierende Software wird unverändert zur Übertragung benutzt. Dazu muss vorausgesetzt werden, dass beide Perspektiven, wie schon in Kapitel 5 propagiert, als gemeinsamer Datenstrom zur Verfügung stehen. Am besten eignen sich das Übereinander- und Untereinander-Format, die beide mit allen gängigen Streamingprogrammen kompatibel sind. Dank des Stereoscopic Multiplexers kann ein passender Datenstrom live zur Verfügung gestellt werden, Synchronisation und Multiplexing erfolgen transparent für die Serveranwendung.

Clientseitig gestaltet sich die Realisierung der stereoskopischen Wiedergabe komplizierter. Während Serverprogramme die Auswahl der Aufnahmequelle und damit eine einfache Integration des Stereoscopic Multiplexers erlauben, ist eine Auswahl der Wiedergabekomponente im Client in der Regel nicht vorgesehen. Die große Vielfalt stereoskopischer Wiedergabemethoden und deren individuelle Anforderungen lassen darüber hinaus eine spezialisierte Anwendung sinnvoll erscheinen. Es bietet sich an, den Stereoscopic Player als Client-Anwendung einzusetzen. Sofern das eingesetzte Streamingsystem einen DirectShow-Filter zum Empfang des Datenstroms mitbringt, kann der Stereoscopic Player ohne Änderungen als Client eingesetzt werden.

Das Testsystem wurde mit den Windows Media-Komponenten von Microsoft realisiert. Die Entscheidung zu Gunsten von Windows Media fiel aufgrund der Verfügbarkeit der benötigten DirectShow-Unterstützung. Darüber hinaus ist es kostenlos erhältlich. Serverseitig dient der *Windows Media Encoder 9 Series* [47] zur Aufnahme und Kodierung des Videos. Als Aufnahmequelle wird der Stereoscopic Multiplexer ausgewählt. Es können alle angebotenen Codecs eingesetzt werden. Für höchste Kompressionseffizienz eignet sich der *Windows Media Video 9 Codec* [48] am besten. Steht eine hohe Bandbreite zur Verfügung, eignen sich ältere Codecs wie *Windows Media Video V7* aufgrund der geringeren Anforderungen an die Rechnerleistung besser. Eine Echtzeitcodierung ist nur im CBR- (konstante Bitrate) oder *Quality VBR*-Modus (variable Bitrate, gleich bleibende Qualität) möglich. Tests haben gezeigt, dass im Quality VBR-Modus Bildfehler auftreten (kontinuierlich abnehmende Bildqualität bis zum nächsten Keyframe), sodass nur der CBR-Modus verwendet werden sollte.

Die Wiedergabe von Windows Media Streams erfolgt über den *Windows Media Source Filter* oder den *WM ASF Reader*. Der Windows Media Source Filter wurde vom Windows Media Player 6.4 zur Wiedergabe von Windows Media-Inhalten verwendet, wurde aber in den folgenden Versionen durch den *WM ASF Reader* ersetzt. Der Stereoscopic Player setzt bevorzugt den WM ASF Reader ein, ist aber auch mit dem Windows Media Source Filter kompatibel. Beide Filter puffern den empfangenen Datenstrom und verursachen damit eine Verzögerung von zirka 10 Sekunden. Um ein Blockieren der Benutzeroberfläche des Stereoscopic Players während des Puffervorgangs zu verhindern, wird der Zugriff auf den Filtergraphen über einen eigenen Thread realisiert. Zur Anzeige des Pufferstatus enthält der Player spezielle Anpassungen für beide Quellfilter.

Der Stereoscopic Multiplexer sowie der Stereoscopic Player sind die zentralen Ergebnisse dieser Diplomarbeit und erfüllen mit Unterstützung des Windows Media Encoders deren Zielsetzung. Der eingeschlagene Lösungsweg zeichnet sich durch hohe Flexibilität aus und ist eine hervorragende Ausgangsbasis für zukünftige Erweiterungen. Insbesondere sei darauf hingewiesen, dass das entwickelte System leicht für andere Aufgaben adaptiert werden kann. So bietet es sich an, für stereoskopische Videotelefonie-Anwendungen den Windows Media Codec durch einen Codec mit geringerer Latenz zu ersetzen. Im Idealfall kann diese Erweiterung ohne Änderungen an der bereits existierenden Software durchgeführt werden.

---

# Zusammenfassung und Ausblick

---

*Dieses abschließende Kapitel gibt einen Überblick über die Ergebnisse dieser Diplomarbeit und streicht die Bedeutung des Stereoscopic Players und des Stereoscopic Multiplexers für die Realisierung von 3D-Videosystemen heraus. Darüber hinaus wird ein Ausblick auf mögliche weiterführende Arbeiten und geplante Erweiterungen des Stereoscopic Players und Stereoscopic Multiplexers gegeben.*

## 11.1 Zusammenfassung der Arbeit

Es ist nicht absehbar, ob dem 3D-Kino und dem 3D-Fernsehen in den nächsten Jahren der Durchbruch gelingen wird. Neben den etablierten IMAX-3D-Filmen lief im Jahr 2004 „Spy Kids 3D: Game Over“ („Mission 3D“) von Regisseur Robert Rodriguez erfolgreich in den Kinos. Aufgrund des wachsenden Interesses von namhaften Hollywood-Regisseuren wie Robert Rodriguez und James Cameron an 3D-Produktionen könnten in den nächsten Jahren durchaus weitere Filme folgen. Ein dauerhafter Erfolg wird aber nur dann möglich sein, wenn die vorherrschende Anaglyphentechnik durch die Polarisationstechnik ersetzt wird.

Mangels massentauglicher, hochqualitativer 3D-Wiedergabetechniken ist die Einführung des 3D-Fernsehens auf breiter Front innerhalb dieses Jahrzehnts nicht mehr zu erwarten. 3D-Produkte für den Heimbereich werden sich auf 3D-DVDs – zum Großteil im Anaglyphenverfahren – beschränken. Gerade dadurch gewinnt aber der Stereoscopic Player, Resultat dieser Diplomarbeit, als einfach zu handhabende Alternative besondere Bedeutung, denn er ermöglicht die kostengünstige und hochqualitative Wiedergabe von 3D-Videos auf PC-Systemen. Durch die umfassende Unterstützung aller bedeutenden stereoskopischen Wiedergabeverfahren und der meisten Dateiformate eignet sich der Player sowohl für ambitionierte Amateure als auch für den professionellen Einsatz. Die intuitiv zu bedienende Benutzeroberfläche hat besonders zur hohen Akzeptanz im Amateurbereich beigetragen. Die große Anzahl positiver Rückmeldungen von allen Anwendergruppen bestätigen diese Aussagen.

Der Stereoscopic Multiplexer zeichnet sich durch seine vielseitigen Einsatzmöglichkeiten aus. Da er sich gegenüber Anwendungen als Treiber präsentiert, kann er in einer Vielzahl von Programmen zu unterschiedlichsten Zwecken verwendet werden. Er nimmt die Bilddaten von zwei Aufnahmegeräten entgegen und gibt einen synchronisierten Datenstrom im Nebeneinander-Layout an die Anwendung weiter. Zusammen mit dem Stereoscopic Player bietet er eine stereoskopische Echtzeitvorschau bei 3D-Aufnahmen (z.B. mit DV-Camcordern). Die Aufnahme von 3D-Videos auf Festplatte gelingt mit einer Vielzahl an Capture-Programmen und mit Hilfe von GraphEdit ist sogar die Aufnahme mit gleichzeitiger Vorschau möglich. Die Übertragung von 3D-Videos über das Internet wurde durch den gemeinsamen Einsatz von

Stereoscopic Multiplexer, Stereoscopic Player und Windows Media Encoder realisiert. Dank der offenen Architektur ist ein Wechsel zu anderen Streaming-Komponenten ohne große Veränderungen an Player oder Multiplexer möglich.

Neben der entwickelten Software wurden die Grundlagen der 3D-Technik dokumentiert, sodass der Leser nach der Lektüre dieser Arbeit die mathematischen Zusammenhänge zur Wahl der Stereobasis sowie die unterschiedlichen Verfahren zur Aufnahme, Kodierung und Wiedergabe kennt und in die Lage versetzt wird, den Stereoscopic Multiplexer und Stereoscopic Player erfolgreich einzusetzen.

## 11.2 Geplante Erweiterungen

Für Videokonferenz-Anwendung ist der Windows Media Codec aufgrund seiner hohen Latenz ungeeignet. Ziel zukünftiger Arbeiten muss es sein, eine geeignete Basis für ein Videokonferenzmodul zu finden und in den Stereoscopic Player zu integrieren. Die Auswahl eines anderen Codecs ist nicht ausreichend, zusätzlich muss auch das Senden von Video- und Audiodatenströmen sowie ein Protokoll für das Sitzungsmanagement in den Player integriert werden. Im Hinblick auf eine einfache Handhabung von Videokonferenzen scheint es für die letzte Ausbaustufe sogar sinnvoll, eine eigenständige Anwendung für 3D-Videokonferenzen zu entwickeln.

Während die Erweiterungen hinsichtlich Videokonferenzfunktionen mittelfristig geplant sind, werden einige andere Ergänzungen des Stereoscopic Players schon in den nächsten Monaten abgeschlossen sein. Derzeit befindet sich ein Modul zur Lokalisierung in Entwicklung. Es ist vorgesehen, dass sich andere Sprachversionen mittels sprachspezifischer Dateien ohne Neukompilierung des Quellcodes hinzufügen lassen. Sie werden sich ohne Programmierkenntnisse erstellen lassen, sodass durch die Mithilfe von Anwendern die Lokalisierung für zahlreiche Sprachen in kurzer Zeit möglich wird. An der Benutzeroberfläche ist ebenfalls eine große Änderung geplant: Durch die Unterstützung von austauschbaren Skins soll der Stereoscopic Player ein moderneres Aussehen erhalten.

Auf technischer Seite sind zahlreiche neue Funktionen zur Verbesserung des Raumbildes und der Performance geplant. Dazu zählt beispielsweise die Möglichkeit, die parallaktische Verschiebung sowie die vertikale Ausrichtung der Halbbilder justieren zu können. Beide Funktionen werden benötigt, wenn die entsprechenden Vorgänge während der Nachbearbeitung nicht korrekt durchgeführt wurden oder eine Anpassung für besondere Wiedergabeverhältnisse (z.B. große Projektionsfläche) notwendig ist. Weiters soll es zukünftig möglich sein, das Bild an den Rändern zu beschneiden, sodass Breitbildfilme, die oben und unten mit schwarzen Balken kodiert wurden, auch auf Breitbilddisplays formatfüllend wiedergegeben werden können. Prozessorspezifische Optimierungen und die Unterstützung eines YUV 4:2:0-Formats am Eingang des Stereo Transformation Filters werden helfen, die Performance des Players weiter zu steigern. Die Verlagerung einiger Berechnungsschritte vom Haupt- zum Grafikprozessor (z.B. Anaglyphenberechnung) zielt ebenfalls in diese Richtung.

Für den Stereoscopic Multiplexer ist sowohl die Unterstützung zusätzlicher Farbformate als auch weiterer Ausgangslayouts geplant. Darüber hinaus soll die Möglichkeit geschaffen werden, den Konfigurationsdialog aufzurufen, auch wenn die Anwendung keine entsprechende Möglichkeit bietet.

---

# Bedienung des Stereoscopic Players

---

Dieses Anhang-Kapitel widmet sich der Bedienung des Stereoscopic Player aus der Sicht des Anwenders. Es beschreibt alle Schritte, die notwendig sind, um mit diesem Player erfolgreich stereoskopische Videos wiedergeben zu können.

## A.1 Einrichtung

Vor dem ersten Einsatz muss der Stereoscopic Player auf dem System eingerichtet werden. Die Installation des Stereoscopic Players erfolgt durch Aufruf des *Microsoft Installer*-Skripts *StereoscopicPlayer084.msi*, anschließend ist den Anweisungen des Installationsprogramms Folge zu leisten. Nach dem Abschluss der Installation ist sowohl am Desktop als auch im Startmenü unter *Stereoscopic Applications* ein Symbol zum Starten des Stereoscopic Players vorhanden.

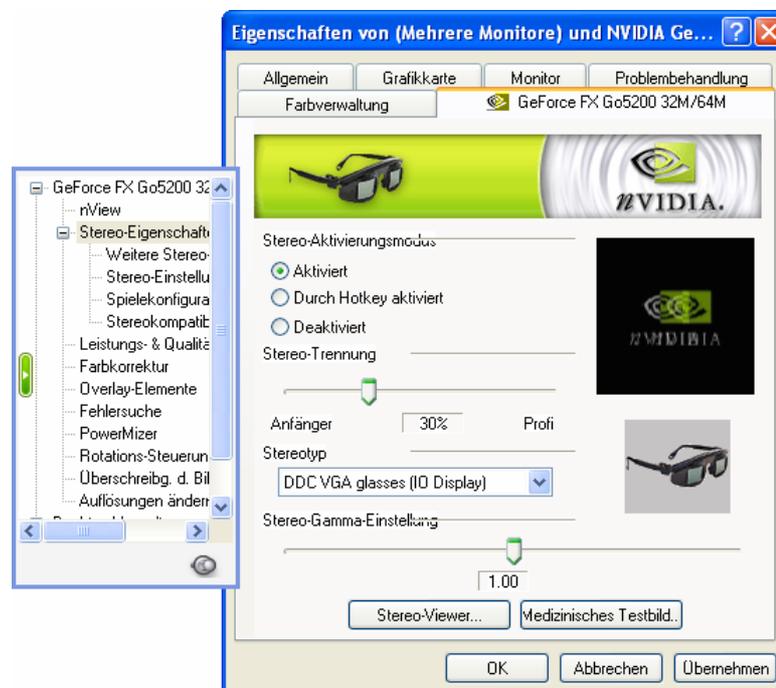


Abbildung A.1: Konfigurationsdialog des NVIDIA Stereo Treibers

Um die volle Funktionsfähigkeit des Players zu gewährleisten, sollten sowohl DirectX als auch die Windows Media-Komponenten auf den neuesten Stand gebracht werden [47]. Für die Zweibildschirmausgabe und die StereoBright-Methode ist entweder Windows XP oder DirectX 9 [44] nötig. Zum Empfang von Windows Media Streams (siehe Kapitel 10.2) ist die Installation des Windows Media Players 9 empfehlenswert, aber nicht zwingend notwendig, falls ältere Versionen der Windows Media Codecs zum Einsatz kommen.

Die Ausgabe über den NVIDIA Stereo Treiber benötigt mindestens DirectX 7 und setzt natürlich den NVIDIA Consumer Stereotreiber voraus. Es ist zu beachten, dass nur Kombinationen von Grafik- und Stereotreiber mit gleicher Versionsnummer funktionsfähig sind. Der Stereo-Modus muss im Konfigurationsdialog des Treibers aktiviert werden (siehe Abbildung A.1, Seite 29), andernfalls werden statt der stereoskopischen Darstellung linkes und rechtes Bild nebeneinander wiedergegeben. Gleiches passiert, wenn der Stereotreiber nicht ordnungsgemäß installiert wurde. Während die Bildwiederholrate in den Treibereinstellungen festgelegt wird, werden Auflösung und Farbtiefe erst im Stereoscopic Player ausgewählt (*File|Settings|General*, siehe Abbildung A.2). Im selben Dialog werden auch die bevorzugte monoskopische und stereoskopische Wiedergabemethode ausgewählt.

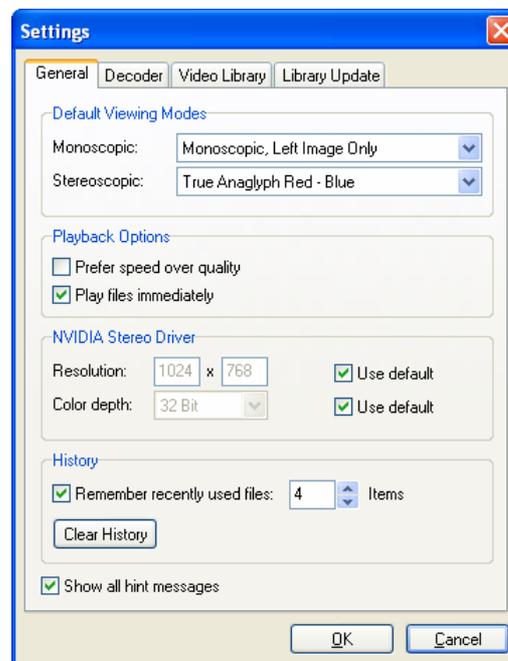


Abbildung A.2: Konfigurationsdialog des Stereoscopic Players

Der Stereoscopic Player basiert auf DirectShow und kann deshalb all jene Formate wiedergeben, für die geeignete DirectShow-Filter oder Video für Windows-Codecs installiert sind. Soll eine Datei geöffnet werden, die Video- oder Audiostreams in nicht unterstützten Formaten enthält, müssen zuvor passende Codecs von Drittanbietern installiert werden. MPEG-2, MPEG-4 [49] und die MPEG-4-Verwandten DivX [15] und Xvid [85] gehören zu den populären Formaten, die DirectShow nicht von Haus aus unterstützt. Abbildung A.3 zeigt eine Fehlermeldung die typischerweise auftritt, wenn der benötigte Video- oder Audiocodec nicht installiert ist. Die DVD-Wiedergabe erfordert ebenfalls zusätzliche Decoder. Der Autor dieser Diplomarbeit hat einen Videodecoder entwickelt, der speziell auf die Anforderungen der 3D-DVD-Wiedergabe zugeschnitten ist. Unter anderem wird der *GPL MPEG-1/2 Decoder* vom

Player automatisch konfiguriert, sodass das Deinterlacing bei der Wiedergabe von DVDs im Interlaced-Format nicht vom Anwender deaktiviert werden muss. Aus rechtlichen Gründen bleibt die Wiedergabe von kopiergeschützten, verschlüsselten DVDs mit dem GPL MPEG-1/2 außen vor. Die Verwendung anderer Decoder ist zum Teil möglich, aktuelle Informationen können der Homepage des Stereoscopic Players entnommen werden. Als Audiodecoder wird der ebenfalls freie *AC3Filter* [76] empfohlen. Der Einsatz anderer Audiodecoder ist unproblematisch, weil die Audiowiedergabe vom Stereoscopic Players genauso wie von gewöhnlichen Playern gehandhabt wird.



Abbildung A.3: Fehlermeldung bei fehlendem Codec

## A.2 Wiedergabe

Nach dem Aufruf des Stereoscopic Players erscheint dessen Hauptfenster (siehe Abbildung A.4). Es umfasst die Menüleiste, den vorerst leeren Wiedergabebereich, eine Zeitleiste, eine Symbolleiste und ein Fenster mit Statusinformation. Die Zeitleiste dient zum schnellen, direkten Anspringen beliebiger Positionen des Videos. Sie ist nur anwählbar, wenn die Dauer des Videos bekannt ist und wahlfreier Zugriff auf beliebige Stellen unterstützt wird. Diese Voraussetzungen sind beispielsweise bei Live-Streams über das Internet oder der Live-Wiedergabe von Aufnahmegeräten nicht erfüllt. Die Symbolleiste umfasst Schaltflächen zur Wiedergabesteuerung sowie einen Lautstärkeregler. Die letzten beiden Schaltflächen im linken Bereich dienen zur schnellen Anwahl der bevorzugten monoskopischen oder stereoskopischen Wiedergabemethode, die zuvor in den Einstellungen festgelegt wurden. Im Statusfenster werden Titel und Quelle (Datei, Verzeichnis, URL oder Gerät) des geladenen Videos angezeigt.

Je nach Quelle werden Videos mit vier unterschiedlichen Menübefehlen geöffnet:

- *File|Open File* dient zum Öffnen einzelner Videodateien oder von Wiedergabelisten. Im Falle von Wiedergabelisten kann danach mit den Befehlen *Playback|Next Chapter* und *Playback|Previous Chapter* zwischen einzelnen Videos gewechselt werden.
- *File|Open DVD* öffnet eine DVD, wobei das Verzeichnis, das die DVD-Daten enthält, vom Anwender ausgewählt werden muss. Damit wird ermöglicht, DVD-Kopien von der Festplatte wiedergegeben zu können.

- *File|Open URL* öffnet Videos, die auf Servern abgelegt sind. Im Normalfall wird das Video vollständig heruntergeladen, bevor die Wiedergabe startet. Verweist die URL (Uniform Resource Locator) auf einen Streaming-Server wie den Windows Media Encoder, dann startet die Wiedergabe nach einer kurzen Pufferphase sofort.
- *File|Live Video* erlaubt es, ein Aufnahmegerät als Quelle zu verwenden. Es werden sowohl ältere Geräte mit Video für Windows-Treiber als auch neue Geräte mit WDM-Treiber unterstützt. Ist der Stereoscopic Multiplexer installiert, steht er ebenfalls in diesem Menü zur Verfügung. Er wird benötigt, wenn linke und rechte Perspektive von unterschiedlichen Geräten geliefert werden.



Abbildung A.4: Hauptfenster des Stereoscopic Players

Im zweiten Fall, der DVD-Wiedergabe, erscheinen im File-Menü die zusätzlichen Einträge *Video Decoder Properties* und *Audio Decoder Properties*. Mit diesen lassen sich, sofern vorhanden, die Konfigurationsdialoge der Decoder aufrufen. Änderungen an der Audiodecoderkonfiguration sind nur notwendig, wenn die Art der Soundausgabe geändert werden soll (Stereo, analoge 5.1 Ausgabe, digitale SPDIF-Ausgabe). Im Gegensatz dazu muss bei allen Videodecodern mit Ausnahme des GPL MPEG-1/2 Decoders (dieser wird automatisch konfiguriert) für die Wiedergabe von Fieldsequential-DVDs (siehe Kapitel 5.4.1) die Deinterlacing-Funktion deaktiviert werden. Bietet der Decoder keine entsprechende Option, kann er für die Wiedergabe von Fieldsequential-DVDs nicht verwendet werden.

Im Falle der Live-Wiedergabe stehen zur Konfiguration des Aufnahmegerätes im *File*-Menü ebenfalls weitere Menüpunkte zur Verfügung, sofern diese Funktionen vom Gerät unterstützt werden. Mittels *Format Properties* wird festgelegt, in welcher Auflösung und welchem Farbraum das Gerät die Bilddaten anliefert. Üblicherweise sind keine Änderungen notwendig, da der Player automatisch ein passendes Format anfordert. Interlaced-3D-Quellen erfordern möglicherweise manuelle Anpassungen, denn es muss unbedingt darauf geachtet werden, dass die volle Zeilenzahl des Videos erhalten bleibt (576 Zeilen bei PAL, 480 bei NTSC). Ansonsten wird eines der Halbbilder entweder gar nicht erfasst, oder die Halbbilder werden

überblendet und können nicht mehr getrennt werden. Ist die Auswahl eines Farbformats möglich, sollte YUY2, UYVY, RGB32 oder RGB24 der Vorzug gegeben werden, mit Priorität gemäß dieser Reihenfolge. Bei Wahl eines anderen Formats müssen auf Kosten der Performance zusätzliche DirectShow-Filter zur Farbraumkonvertierung eingesetzt werden, oder die Wiedergabe scheitert gänzlich. Der Menüpunkt *Device Properties* dient zur Konfiguration gerätespezifischer Parameter. Unter *Input Properties* wird der aktive Videoeingang (z.B. TV-Tuner, Video, S-Video) festgelegt. Handelt es sich um eine TV-Karte, kann mit *Tuner Properties* die Sendereinstellung vorgenommen werden.

Wurde die Quelle erfolgreich geöffnet, muss im nächsten Schritt deren Stereoformat spezifiziert werden. Dazu wird in *File|Layout* zunächst das Layout ausgewählt (siehe Kapitel 5.2). Wurde ein Übereinander- oder Nebeneinander-Layout gewählt, muss im Menü *File|Separation* der Abstand der Halbbilder in Pixel gesetzt werden, sofern diese einen Abstand voneinander aufweisen. Ist das Format des Videos nicht bekannt, kann es hilfreich sein, das unveränderte Video zu betrachten. Dazu muss im Menü *View|Viewing Method* die Wiedergabemethode *Source* ausgewählt werden. Abschließend ist in *File|Aspect Ratio* das Seitenverhältnis festzulegen. Entweder wird es direkt festgelegt, z.B. auf 4:3 oder 16:9, oder es wird aus der Auflösung des Videos abgeleitet: *Default* bedeutet, dass das Seitenverhältnis eines Halbbildes unverändert beibehalten wird, *Default, Half Height*, dass jedes Halbbild nur mit der halben vertikalen Auflösung vorliegt, also zur Wiedergabe auf die doppelte Höhe gestreckt werden muss. Äquivalent dazu verhält sich der Menüpunkt *Default, Half Width*, er streckt das Bild auf die doppelte Breite.

Die gewählten Einstellungen für eine Datei, DVD oder URL werden in der integrierten Bibliothek des Players gespeichert. Wird das Video erneut geöffnet, werden die zuletzt verwendeten Einstellungen wiederhergestellt. In folgenden Fällen ist der Player in der Lage, die korrekten Einstellungen auch bei der erstmaligen Wiedergabe selbst zu ermitteln:

- Es ist eine stereoskopische Informationsdatei für die zu öffnende Videodatei vorhanden. Sie trägt den gleichen Dateinamen wie die Videodatei und befindet sich im selben Verzeichnis, unterscheidet sich von dieser aber durch die Erweiterung *svi* (Stereoscopic Video Information File).
- Für die Videodatei oder die DVD sind Informationen auf einem Internetserver verfügbar. Voraussetzung ist, dass während dem Öffnen des Videos eine Verbindung ins Internet besteht, ein Server in den Einstellungen (*File|Settings|Library Update*) eingetragen ist und *Contact update service when opening a new file* aktiviert ist.

Die gewünschte Wiedergabemethode wird im Menü *View|Viewing Method* festgelegt. Mit den Schaltflächen *View Monoscopic* und *View Stereoscopic* in der Symbolleiste kann schnell zu den bevorzugten Wiedergabemethoden gewechselt werden. Wird eine Methode gewählt, die beide Halbbilder über- oder nebeneinander ausgibt, kann mit *View|Separation* ein schwarzer Trennbalken zwischen den Halbbildern eingefügt werden. Beispielsweise wird dieser Balken für Shutterbrillen in Syncdoubbling-Betriebsart benötigt (siehe Kapitel 4.7.4). Darüber hinaus kann mit *View|Aspect Ratio* eine Ausgabe mit halber Höhe oder Breite erzwungen werden. Es ist wichtig, den Unterschied zwischen *File|Aspect Ratio* und *View|Aspect Ratio* zu verstehen – ersteres definiert das Seitenverhältnis des Videos unabhängig von der Wiedergabemethode, zweiteres beeinflusst nur die gewählte Wiedergabemethode und sollte nicht zur Kompensation von Video-bezogenen Fehlkonfigurationen verwendet werden.

Der Menüpunkt *View|Full Screen Scaled* aktiviert den Vollbildmodus. Steht mehr als ein Bildschirm zur Verfügung, kann der Bildschirm für die Vollbilddarstellung unter *View|Full Screen Monitor* ausgewählt werden. Wurde eine Zweibildschirmwiedergabemethode ausgewählt, können sowohl der Bildschirm für die linke als auch die rechte Perspektive ausgewählt werden, womit der Player auch bei drei oder mehr Bildschirmen uneingeschränkt einsatzfähig ist. Befindet sich der Player im Vollbildmodus, bleibt das Hauptfenster in verkleinerter Form trotzdem verfügbar. Es kann zur Steuerung verwendet werden, sofern es auf einem Bildschirm platziert wird, der nicht von der Vollbildwiedergabe belegt wird (mit Ausnahme der NVIDIA Stereotreiber-Betrachtungsmethode). *View|Full Screen Unscaled* verwendet ebenfalls einen Bildschirm exklusiv für die Videodarstellung, allerdings wird die zuvor benutzte Darstellungsgröße beibehalten, umgeben von einem schwarzen Rahmen.

Die Steuerung der Wiedergabe erfolgt über die Menüpunkte *Play* (Wiedergabe starten), *Pause* (Wiedergabe pausieren) und *Stop* (Wiedergabe stoppen) im *Playback*-Menü oder die entsprechenden Schaltflächen in der Symbolleiste. *Step Forward* und *Step Backward* springen einzelbildweise vor- und rückwärts. Das Zurückspringen wird nicht von allen Codecs unterstützt und steht bei der DVD-Wiedergabe generell nicht zur Verfügung. Dafür bleibt der DVD-Wiedergabe *Fast Forward* (schneller Vorlauf) und *Rewind* (schneller Rücklauf) vorbehalten. Ebenfalls auf die DVD-Wiedergabe sind die Menüpunkte *Root Menu* und *Title Menu* zum Aufruf der entsprechenden DVD-Menüs beschränkt.

Während der Live-Wiedergabe von einem Aufnahmegerät kann die Wiedergabe ausschließlich angehalten und wieder fortgesetzt werden, alle anderen Funktionen sind deaktiviert. Eine Ausnahme stellen DV-Camcorder und DV-Videorecorder dar – hier dienen sie zur Steuerung des Bandlaufwerks.

## A.3 Bibliotheksverwaltung

Die Video-Bibliothek (siehe Abbildung 8.5, Seite 29) wird über *File|Video Library* aufgerufen. In der linken Fensterhälfte wird die hierarchische Struktur der Bibliothek dargestellt, die der Verzeichnisstruktur im Windows-Explorer nachempfunden ist. Es gibt drei unveränderliche Wurzeleinträge: *Files*, *DVDs* und *URLs*, die selbst keine Videos, sondern nur Untereinträge, Kategorien genannt, enthalten können. Wird eine Kategorie durch anklicken ausgewählt, werden in der rechten Fensterhälfte deren Videos angezeigt. Wurzeleinträge können ebenfalls ausgewählt werden, in diesem Fall werden die Videos aller untergeordneten Kategorien aufgelistet. Das Anlegen und Löschen neuer Kategorien erfolgt über die Menüpunkte *New Category* und *Delete* des Kontextmenüs der linken Fensterhälfte. Im Kontextmenü der rechten Fensterhälfte können Videos mit *Open* gestartet, mit *Delete* gelöscht und mit *Move To* in eine andere Kategorie verschoben werden. Das Verschieben von Videos und Umstrukturieren der Kategorien ist auch mittels Drag&Drop möglich. Prinzipiell sind Verschiebeoperationen nur innerhalb eines Wurzeleintrags möglich.

*Properties* öffnet die Eigenschaftsseiten eines Videos. Die Seite *General* erlaubt das Ändern des Titels und zeigt einige Informationen über das Video an. Unter *Settings* können dieselben Einstellungen wie in den Menüs *File|Layout*, *File|Separation* und *File|Aspect Ratio* des Hauptfensters getätigt werden. Auf der Seite *Information* können ergänzende Kommentare zum Video sowie dessen Hersteller oder Herkunft angegeben werden. Die Seite *Preview* dient zum Laden des Vorschaubildes, das in der Bibliothek angezeigt wird. Die jeweilige Grafikdatei muss eine

Auflösung von 160 x 120 Pixel aufweisen, im JPEG-Format gespeichert und kleiner als 10 kByte sein, um als Vorschaubild geeignet zu sein.

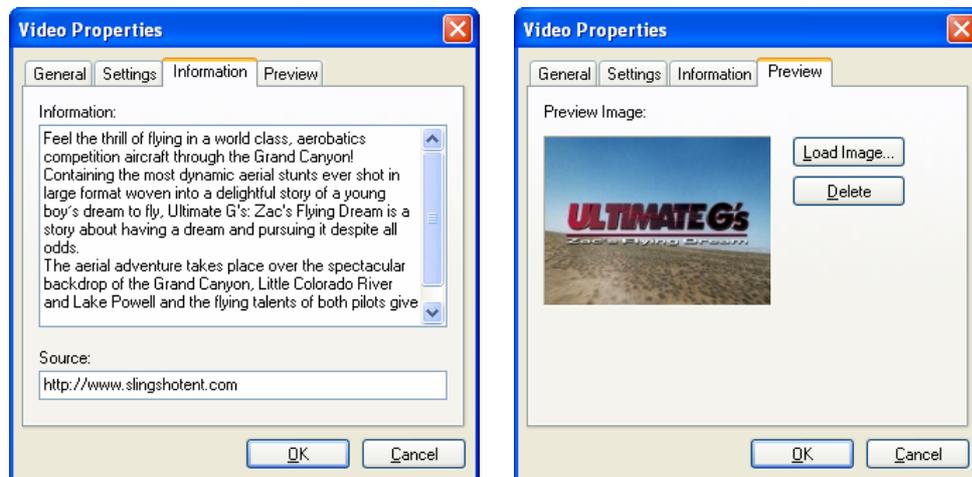


Abbildung A.5: Eigenschaftsseiten eines Videos

Bibliothekseinträge (mit Ausnahme von Videos in der Wurzelkategorie *URLs*) lassen sich über das Internet auf den neuesten Stand bringen. Dazu wählt man in einem der Kontextmenüs *Internet Update* und anschließend einen der drei Untermenüpunkte:

- *Selected Videos Only* aktualisiert nur die markierten Videos.
- *New Videos Only* berücksichtigt nur bis dato noch nie aktualisierte Videos.
- *All Videos* aktualisiert alle Videos.

Eine weitere Funktion im Kontextmenü eines Videos ist der Export einer stereoskopischen Informationsdatei oder einer Datei für die Bibliotheksaktualisierung. Eine stereoskopische Informationsdatei wird mit *Export | Stereoscopic Information File* erzeugt. Sie wird automatisch im selben Verzeichnis der Videodatei angelegt und erhält den gleichen Dateinamen, abgesehen von der Erweiterung *svi*. Die Videodatei sollte nur zusammen mit der *svi*-Datei an andere Personen weitergegeben werden, um diesen den Aufwand der Stereoformat-Einstellung abzunehmen.

Dateien für die Bibliotheksaktualisierung werden in einem Verzeichnis erstellt, das zuvor vom Anwender ausgewählt werden muss. Als Dateiname wird eine Nummer verwendet, welche die Datei oder DVD eindeutig identifiziert, die Erweiterung lautet ebenfalls *svi*. Die *svi*-Datei muss anschließend auf einen Webserver für Bibliotheksaktualisierungen kopiert werden. Soll die Datei am offiziellen Aktualisierungsserver verfügbar gemacht werden, ist Kontakt mit dem Autor dieser Diplomarbeit aufzunehmen. Das Funktionsprinzip eines solchen Servers wurde bereits in Kapitel 8.5.2 erläutert. Stereoskopische Informationsdateien und Dateien für die Bibliotheksaktualisierung unterscheiden sich nur durch den Dateinamen, nicht aber im Inhalt.

Von Zeit zu Zeit sollte die Bibliothek mit dem Menüpunkt *Backup Library* gesichert werden. Die Backups werden im Unterverzeichnis *Library Backups* des Stereoscopic Player-Installationsverzeichnisses abgelegt. Jede Sicherung erfolgt in eine eigene Datei mit Sicherungsdatum und -uhrzeit im Dateinamen. Mit dem Menüpunkt *Import* können Backups ebenso wie *svi*-Dateien in die Bibliothek eingespielt werden. Ist in den Einstellungen *Preserve Categories* akti-

viert, verbleiben die Videos in ihren bisherigen Kategorien, ansonsten werden sie gemäß den Vorgaben verschoben und nicht vorhandene Kategorien angelegt. Dieses Verhalten betrifft alle Varianten von Aktualisierungen oder Importen in die Bibliothek.

## A.4 Kommandozeilenparameter

Der Stereoscopic Player unterstützt eine Reihe von Kommandozeilenparameter. Sie können verwendet werden, um nach dem Start automatisch ein vorgegebenes Video zu öffnen und dessen Format festzulegen. Weiters kann die Wiedergabemethode festgelegt werden, in den Vollbildmodus gewechselt und die Wiedergabe gestartet werden. Die vollständige Syntax der Kommandozeilenparameter lautet:

```
StereoPlayer.exe [{Filename | -file:Filename | -dvd:Directory | -url:URL | -svi:Filename | -spp:Filename} [-il:Input Layout] [-is:Input Separation] [{-ihw] [-ihh] | [-ax:Aspect Ratio X] [-ay:Aspect Ratio Y]} [-inf:Information File] [-src:Source] [-img:JPEG Filename] [-fp] [-play] [-fss] [-ol:Output Layout] [-os:Output Separation] [-ohw] [-ohh]
```

### A.4.1 Quellen

Parameter	Beschreibung
<i>Filename</i>	Dieser Parameter kann verwendet werden, um Videodateien, stereoskopische Informationsdateien oder Wiedergabelisten zu öffnen. Für DVDs oder URLs muss allerdings der <code>-dvd</code> bzw. <code>-url</code> Parameter verwendet werden.
<b>-file:</b> <i>Filename</i>	Öffnet eine Videodatei. Als Dateiname muss eine Videodatei angegeben werden, die vom Stereoscopic Player geöffnet werden kann.
<b>-dvd:</b> <i>Directory</i>	Öffnet eine DVD. Der Pfad kann mit oder ohne VIDEO_TS-Verzeichnis angegeben werden.
<b>-url:</b> <i>URL</i>	Öffnet eine URL.
<b>-svi:</b> <i>Filename</i>	Importiert eine stereoskopische Informationsdatei in die Bibliothek und versucht das Video zu öffnen. Die Videodatei oder DVD-Daten müssen sich im selben Verzeichnis befinden, damit sie gefunden werden können.
<b>-spp:</b> <i>Filename</i>	Öffnet eine Wiedergabeliste.

Tabelle A.1: Kommandozeilenparameter – Quellen

### A.4.2 Eingangsformat

Die Verwendung der folgenden Parameter ist nur zulässig, wenn eine Quelle gemäß Tabelle A.1 angegeben wurde.

Parameter	Beschreibung
<b>-il:</b> <i>Input Layout</i>	Legt das stereoskopische Layout des Videos fest. <i>Mono:</i> Monoskopisch <i>InterlacedRF:</i> Interlaced, rechtes Halbbild zuerst <i>InterlacedLF:</i> Interlaced, linkes Halbbild zuerst

	<i>SideBySideRF</i> : Nebeneinander, rechtes Halbbild links
	<i>SideBySideLF</i> : Nebeneinander, linkes Halbbild rechts
	<i>OverUnderRT</i> : Übereinander, rechtes Halbbild oben
	<i>OverUnderLT</i> : Übereinander, linkes Halbbild oben
<b>-is</b> : <i>Input Separation</i>	Gibt den Abstand zwischen linkem und rechtem Halbbild an. Dieser Parameter darf nur für Neben- und Untereinanderlayouts einen Wert größer als 0 aufweisen.
<b>-ihw</b>	Das Video hat nur die halbe horizontale Auflösung. Dieser Parameter kann nicht zusammen mit den Parametern <i>-ihh</i> , <i>-iax</i> und <i>-ay</i> verwendet werden.
<b>-ihh</b>	Das Video hat nur die halbe vertikale Auflösung. Dieser Parameter kann nicht zusammen mit den Parametern <i>-ihw</i> , <i>-ax</i> und <i>-ay</i> verwendet werden.
<b>-ax</b> : <i>Aspect Ratio X</i> <b>-ay</b> : <i>Aspect Ratio Y</i>	Die Parameter <i>-ax</i> und <i>-ay</i> werden verwendet, um das Seitenverhältnis des Videos anzugeben und müssen immer gemeinsam eingesetzt werden. Sie sind nützlich wenn das Video anamorph kodiert ist. Bei quadratischen Pixeln können stattdessen auch die Parameter <i>-ihw</i> und <i>-ihh</i> verwendet werden.
<b>-inf</b> : <i>Information File</i>	Dateiname (einschließlich Pfad) einer Textdatei, die Informationen über das Video enthält. Die Dateierweiterung muss <i>txt</i> lauten.
<b>-src</b> : <i>Source</i>	Herkunft des Videos. Hierbei kann es sich um eine Internetadresse, einen Namen oder um eine andere Art von Quelle handeln.
<b>-img</b> : <i>JPEG Filename</i>	Dateiname einer Grafik im JPEG-Format, die als Vorschaubild des Videos in der Bibliothek angezeigt werden soll. Die Auflösung muss 160 x 120 Pixel betragen und die Dateigröße darf 10240 Bytes nicht überschreiten.
<b>-fp</b>	Dieser Parameter zwingt den Stereoscopic Player, die aufgeführten Parameter auch dann zu verwenden, wenn für das Video bereits ein Eintrag in der Bibliothek existiert. Der existierende Eintrag wird aktualisiert.

Tabelle A.2: Kommandozeilenparameter – Eingangsformat

### A.4.3 Wiedergabe

Parameter	Beschreibung
<b>-ol</b> : <i>Output Layout</i>	Legt die stereoskopische Wiedergabemethode fest.
	<i>Source</i> : Unveränderte Wiedergabe
	<i>MonoL</i> : Monoskopisch, nur linkes Halbbild
	<i>MonoR</i> : Monoskopisch, nur rechtes Halbbild
	<i>Dual</i> : Zweibildschirmausgabe
	<i>StereoBright</i> : Zweibildschirmausgabe, StereoBright
	<i>OpenGL</i> : Vierfachgepuffertes OpenGL
	<i>SideBySideRF</i> : Nebeneinander, rechtes Halbbild links
	<i>SideBySideLF</i> : Nebeneinander, linkes Halbbild links
	<i>OverUnderRT</i> : Übereinander, rechtes Halbbild oben

---

	<i>OverUnderLT</i> : Übereinander, linkes Halbbild oben
	<i>InterlacedRT</i> : Zeilenweise verschachtelt, rechtes Halbbild zuerst
	<i>InterlacedLT</i> : Zeilenweise verschachtelt, linkes Halbbild zuerst
	<i>InterlacedRF</i> : Spaltenweise verschachtelt, rechtes Halbbild zuerst
	<i>InterlacedLF</i> : Spaltenweise verschachtelt, linkes Halbbild zuerst
	<i>TrueAnaRB</i> : Echte Anaglyphen, rot-blau
	<i>TrueAnaBR</i> : Echte Anaglyphen, blau-rot
	<i>TrueAnaRG</i> : Echte Anaglyphen, rot-grün
	<i>TrueAnaGR</i> : Echte Anaglyphen, grün-rot
	<i>GrayAnaRC</i> : Graue Anaglyphen, rot-cyan
	<i>GrayAnaCR</i> : Graue Anaglyphen, cyan-rot
	<i>GrayAnaYB</i> : Graue Anaglyphen, gelb-blau
	<i>GrayAnaBY</i> : Graue Anaglyphen, blau-gelb
	<i>HColAnaRC</i> : Halbfarbige Anaglyphen, rot-cyan
	<i>HColAnaCR</i> : Halbfarbige Anaglyphen, cyan-rot
	<i>HColAnaYB</i> : Halbfarbige Anaglyphen, gelb-blau
	<i>HColAnaBY</i> : Halbfarbige Anaglyphen, blau-gelb
	<i>ColAnaRC</i> : Farbige Anaglyphen, rot-cyan
	<i>ColAnaCR</i> : Farbige Anaglyphen, cyan-rot
	<i>ColAnaYB</i> : Farbige Anaglyphen, gelb-blau
	<i>ColAnaBY</i> : Farbige Anaglyphen, blau-gelb
<b>-os:Output Separation</b>	Abstand von linkem und rechtem Halbbild in 1/10 Prozent, ein Wert von 10 bedeutet also, dass ein schwarzer Trennbalken mit 1% Breite bezogen auf die Halbbildbreite eingefügt wird. Dieser Parameter darf nur für Neben- und Untereinander-Wiedergabemethoden einen Wert größer als 0 aufweisen.
<b>-ohw</b>	Ausgabe erfolgt mit halber Breite. Wird weder dieser Parameter noch <b>-ohh</b> angegeben, erfolgt die Ausgabe im ursprünglichen Seitenverhältnis.
<b>-ohh</b>	Ausgabe erfolgt mit halber Höhe. Dieser Parameter wird benötigt, wenn die Wiedergabe mit Shutterbrillen im Syncdoubling-Verfahren erfolgt. In diesem Fall muss eine der Übereinander-Wiedergabemethoden gewählt werden.
<b>-play</b>	Startet die Wiedergabe. Dieser Parameter kann nur verwendet werden wenn eine Quelle gemäß Tabelle A.1 angegeben wurde.
<b>-fss</b>	Wechselt in den Vollbildmodus. Dieser Parameter ist ebenfalls nur funktionsfähig, wenn ein Video geladen wird.

---

Tabelle A.3: Kommandozeilenparameter – Wiedergabe

Da der Stereoscopic Player den Abstand und die Ausgabegröße für jede Wiedergabemethode getrennt speichert, wirken sich die Parameter **-os**, **-ohw** und **-ohh** nur auf die in **-ol** angegebene oder die zuletzt verwendete Wiedergabemethode aus.

---

# Dokumentation der Dateiformate

---

*Dieses Kapitel beschreibt den Aufbau von stereoskopischen Informationsdateien, damit andere Softwareentwickler in der Lage sind, dieses Dateiformat zu verwenden. Die Beschreibung des Wiedergabelisten-Formats richtet sich auch an Anwender, da Wiedergabelisten mit dem Stereoscopic Player derzeit noch nicht erzeugt werden können.*

## B.1 Berechnung der Hashwerte

Videodateien werden anhand eines eindeutigen, 64-Bit-Hashwertes identifiziert. Der Hashwert wird aus 114 Bytes des Dateiinhalts mittels exklusiver Oder-Verknüpfung berechnet. Nach jeweils zwei Bytes wird eine Schiebeoperation um ein Bit nach Links durchgeführt. Die Schiebeoperation im letzten Schleifendurchgang ist überflüssig und führt zum Verlust von einem Bit. Der Algorithmus lässt sich zum jetzigen Zeitpunkt aufgrund der existierenden stereoskopischen Informationsdateien allerdings nicht mehr ändern. Beim Erzeugen von Dateien für die Bibliotheksaktualisierung wird der Absolutbetrag des Hashwertes als Dateiname verwendet.

```
function GetFileID(Filename: String): Int64;
var
  i: Integer;
  FileStream: TFileStream;
  NextByte: Byte;
begin
  Result := 0;

  try
    FileStream := TFileStream.Create(Filename, fmOpenRead);
  except
    Exit;
  end;

  for i := 1 to 57 do begin
    FileStream.Position := (FileStream.Size-1) * (2*i-1) div (2*57);
    FileStream.ReadBuffer(NextByte, SizeOf(NextByte));
    Result := Result xor NextByte;

    FileStream.Position := (FileStream.Size-1) * i div 57;
    FileStream.ReadBuffer(NextByte, SizeOf(NextByte));
    Result := Result xor NextByte;

    Result := Result shl 1;
  end;
```

```
    FileStream.Free;  
end;
```

Zur Identifikation von DVDs wird der Hashwert ausschließlich aus dem Inhalt der Datei *VIDEO\_TS.IFO* berechnet. Gegen eine Berücksichtigung aller Dateien spricht einerseits, dass die Berechnung aufgrund der langsamen Zugriffszeiten eines DVD-Laufwerks zu lange dauern würde. Andererseits können die Videodaten auf einer DVD in verschlüsselter Form vorliegen, sodass sich der Hashwert von Original-DVD und entschlüsselter Kopie ändern würde. Der Algorithmus liest die Datei in 64-Bit-Blöcken und verknüpft diese mit einer exklusiven Oder-Operation. Ebenso wie bei Dateien wird der Absolutbetrag des Hashwertes für Dateinamen herangezogen.

```
function GetDVDID(Directory: String): Int64;  
var  
    i: Integer;  
    FileStream: TFileStream;  
    NextValue: Int64;  
begin  
    Result := 0;  
  
    try  
        FileStream := TFileStream.Create(Directory + '\VIDEO_TS.IFO',  
                                         fmOpenRead);  
    except  
        Exit;  
    end;  
  
    for i := 1 to FileStream.Size div SizeOf(NextValue) do begin  
        FileStream.ReadBuffer(NextValue, SizeOf(NextValue));  
        Result := Result xor NextValue;  
    end;  
  
    FileStream.Free;  
end;
```

## B.2 Format von stereoskopischen Informationsdateien

Die Bibliothek, gespeichert in der Datei *Library.dat*, und stereoskopische Informationsdateien verwenden dasselbe Dateiformat. Es handelt sich um Binärdateien, die aus drei Abschnitten bestehen, dem Header, einer Liste aller Kategorien und einer Liste aller Videoinformationen. Eine stereoskopische Informationsdatei enthält genau eine Videoinformation, während eine Bibliotheksdatei eine beliebige Anzahl enthalten darf. Die hierarchische Struktur wird durch das Feld Eltern-ID abgebildet. Jede Kategorie mit Ausnahme der Wurzelkategorien besitzen eine übergeordnete Kategorie, die auch in der Datei enthalten sein muss. Ebenso muss jedes Video einer Kategorie zugeordnet sein. Die Wurzelkategorien *Files*, *DVDs* und *URLs* müssen ebenfalls in der Datei enthalten sein. Ihre IDs müssen den Werten von Tabelle B.1 entsprechen und die Eltern-ID-Felder den Wert 0 enthalten.

Bezeichnung	ID
Files	2810800629978329
DVDs	2810833035958617
URLs	657670585994094975

Tabelle B.1: Vordefinierte Wurzelkategorien

Es ist zu beachten, dass die ID eines Videos nicht mit dem Hashwert übereinstimmt. Die ID ist eine Zufallszahl, die beim erstmaligen Anlegen der Videoinformation erzeugt wird, während sich der Hashwert aus dem Dateinhalt ableitet (siehe Kapitel B.1). Die Videoinformationen enthalten Angaben über Auflösung, Dauer und Dateigröße des Videos damit ein rascher Aufbau der Bibliothek ohne Zugriff auf die jeweiligen Medien möglich ist. Die Bibliothek erlaubt aus Gründen der Übersichtlichkeit das Ausblenden von Videos, die noch nie wiedergegeben wurden, dazu gehören unter Umständen die Beispieleinträge oder manuell importierter Einträge. Um solche Videos zu erkennen, existiert ein Flag, das erst beim Öffnen des Videos in den Videoinformationen und rekursiv in allen übergeordneten Kategorien gesetzt wird. Ein weiteres Flag kennzeichnet Videoinformationen, die bereits über das Internet aktualisiert wurden. Die Flags *Öffentliches Video* und *Öffentliche Kategorie* können nur im Entwicklermodus des Players gesetzt werden, der mit der Tastenkombination *Ctrl+Alt+D* aktiviert wird. Sie dienen dem Autor zum Erstellen der Bibliotheksdatei, die mit dem Player ausgeliefert wird, indem alle nicht-öffentlichen Videos gelöscht werden. Die entsprechenden Befehle (*Private Video*, *Public Video*, *Delete Private Items*) sind im Kontextmenü der Bibliothek zu finden. Ein weiterer Befehl des Entwicklermodus ist *Reset All Items*, mit dem der Status aller Videos und Kategorien auf „noch nicht wiedergegeben“ zurückgesetzt werden kann.

Feldname	Länge	Bemerkung
Signatur	21 bzw. 25 Bytes	StereoVideoInfo[V1.0] oder StereoVideo-Library[V1.0]
Kategorien	4+X Bytes	Anzahl der Kategorien, gefolgt von den eigentlichen Kategorie-Blöcken (siehe Tabelle B.3).
Videos	4+X Bytes	Anzahl der Videos, gefolgt von den eigentlichen Videoinformations-Blöcken (siehe Tabelle B.4).

Tabelle B.2: Aufbau von stereoskopischen Informationsdateien

Feldname	Länge	Bemerkung
ID	8 Bytes	Zufallszahl, die die Kategorie eindeutig identifiziert.
Eltern-ID	8 Byte	ID der übergeordneten Kategorie
Letzte Änderung	8 Byte	Datum und Zeit im Delphi-TDateTime-Format
Flags	1 Byte	Bit 1: Kategorie enthält bereits wiedergegebenes Video Bit 2-7: Nicht genutzt Bit 8: Öffentliche Kategorie
Titel	2+X Bytes	Länge des String, gefolgt vom String selbst

Tabelle B.3: Aufbau eines Kategorie-Blocks

Feldname	Länge	Bemerkung
Medientyp	1 Byte	0 = Datei, 1 = DVD, 2 = URL
ID	8 Bytes	Eindeutige Zufallszahl
Hashwert	8 Bytes	Hashwert der Videodatei oder DVD bzw. 0 bei einer URL
Kategorie-ID	8 Bytes	ID der Kategorie, welcher das Video zugeordnet ist
Letzte Änderung	8 Bytes	Datum und Zeit im Delphi-TDateime-Format
Titel	2+X Bytes	Länge des String, gefolgt vom String selbst
Dateiname	2+X Bytes	Dateiname, DVD-Pfad (ohne VIDEO_TS) oder URL
Information	2+X Bytes	Länge des String, gefolgt vom String selbst
Quelle	2+X Bytes	Länge des String, gefolgt vom String selbst
Layout	1 Byte	0 = Monoskopisch 1 = Zeilenweise verschachtelt, rechtes Halbbild zuerst 2 = Zeilenweise verschachtelt, linkes Halbbild zuerst 3 = Nebeneinander, rechtes Halbbild links 4 = Nebeneinander, linkes Halbbild links 5 = Übereinander, rechtes Halbbild oben 6 = Übereinander, linkes Halbbild oben
Abstand	2 Bytes	Abstand zwischen linkem und rechtem Halbbild (nur bei Neben- und Untereinander-Layout, sonst 0)
Seitenverhältnis X	2 Bytes	0 bei Standardseitenverhältnis
Seitenverhältnis Y	2 Bytes	0 bei Standardseitenverhältnis
Breite	2 Bytes	Breite des Videos in Pixel
Höhe	2 Bytes	Höhe des Videos in Pixel
Dateigröße	8 Bytes	Größe der Videodatei oder der DVD-Daten in Bytes
Dauer	8 Bytes	Dauer des Videos im DirectX-REFTIME-Format
Flags	1 Byte	Bit 1: Halbe horizontale Auflösung Bit 2: Halbe vertikale Auflösung Bit 3: Video wurde bereits wiedergegeben Bit 4: Videoinformation wurde aktualisiert Bit 5-7: Nicht genutzt Bit 8: Öffentliches Video Bit 1 und 2 schließen sich gegenseitig aus und werden ignoriert, wenn ein Seitenverhältnis angegeben ist.
JPEG-Bild	4+X Bytes	Größe der eingebettete JPEG-Datei, gefolgt von den Bilddaten selbst (160 x 120 Pixel, max. 10240 Bytes).

Tabelle B.4: Aufbau eines Videoinformation-Blocks

## B.3 Format von Wiedergabelisten

Wiedergabelisten werden sich in zukünftigen Versionen des Stereoscopic Player in der Video-bibliothek bilden lassen, derzeit müssen sie aber noch manuell eingegeben werden. Wieder-

gabelisten sind Textdateien mit der Erweiterung *spp* (Stereoscopic Player Playlist). Sie beginnen mit einem Header, der in der ersten Zeile den Text `Playlist[V1.0]` enthält, gefolgt vom Titel und einer eindeutigen Nummer in der zweiten Zeile. Die eindeutige Nummer ist optional und wurde im Hinblick auf die zukünftige Integration der Wiedergabelisten in die Bibliothek vorgesehen. Als eindeutige Nummer eignen sich beispielsweise aktuelles Datum und Uhrzeit. In den folgenden Zeilen sind die Videodateien aufgelistet, die sequentiell wiedergegeben werden sollen. Auf die Angabe des vollständigen Pfades kann verzichtet werden, wenn sich die Videodatei im gleichen Verzeichnis wie die Wiedergabeliste befindet oder die Videodatei in einem Unterverzeichnis abgelegt ist. Kommentarzeilen müssen mit einem Strichpunkt eingeleitet werden.

Nach jedem Videodateinamen können weitere Parameter, getrennt durch einen senkrechten Strich, angegeben werden. Der erste Parameter legt fest, ob die Datei wiederholt werden soll, bis der Anwender manuell zum nächsten Video springt. Dazu muss `loop` angegeben werden, ansonsten `noloop` (oder der Parameter bleibt leer). Zweiter und dritter Parameter legen Einstiegs und Ausstiegspunkt fest. Entweder werden die Zeitpunkte in Nanosekunden oder die Framenummern mit vorangestelltem *f* angegeben. Die zweite Variante ist nicht möglich, wenn ein Videocodec mit variabler Framerate benutzt wird. Das folgende Beispiel illustriert die Möglichkeiten von Wiedergabelisten.

```
Playlist[V1.0]
Demo Playlist|110520041930

; Jetzt beginnt die Liste der Videodateien.
Video 1.avi
Unterverzeichnis\Video 2.avi

; Die nächste Datei wird wiederholt.
Video 3.avi|loop

; Wiedergabe ab Sekunde 2.
Video 4.avi|noloop|20000000

; Wiedergabe von Sekunde 2 bis 4.
Video 5.avi||20000000|40000000

; Wiedergabe von Frame 50 bis 100.
Video 6.avi||f50|f100
```

---

# Abbildungsverzeichnis

---

Abbildung 2.1: Wheatstones Stereoskop .....	4
Abbildung 2.2: Brewsters Stereoskop .....	5
Abbildung 2.3: Prinzip des View-Master (aus US-Patent 2.511.334 von Wilhelm B. Gruber).....	6
Abbildung 2.4: Erster IMAX-Projektor [43] .....	9
Abbildung 3.1: Konvergenz und Akkomodation .....	13
Abbildung 3.2: Konvergenz und Akkomodation bei Wiedergabe eines Stereobildes.....	14
Abbildung 3.3: Querdisparation.....	15
Abbildung 3.4: Verdeckung als Tiefenhinweis.....	15
Abbildung 3.5: Größe als Tiefenhinweis .....	16
Abbildung 3.6: Perspektive als Tiefenhinweis .....	16
Abbildung 3.7: Dunst als Tiefenhinweis .....	17
Abbildung 3.8: Bewegungsparallaxe als Tiefenhinweis.....	17
Abbildung 3.9: Theoretischer Horopter .....	18
Abbildung 3.10: Parallele und konvergierende Kameraanordnung .....	20
Abbildung 3.11: Scheinfenster und parallaktische Verschiebung .....	21
Abbildung 3.12: Scheinfensterkorrektur durch außermittige Bildwandler .....	22
Abbildung 3.13: Zur Linsengleichung.....	23
Abbildung 3.14: Größenabbildung im Auge .....	23
Abbildung 3.15: Parallaxenabbildung im Auge .....	24
Abbildung 3.16: Deviations- bzw. Tiefenabbildung im Auge.....	25
Abbildung 3.17: Parallaktische Verschiebung.....	26
Abbildung 3.18: Zur Ermittlung der zulässigen Deviation .....	29
Abbildung 4.1: Parallelblick .....	29
Abbildung 4.2: Parallelblick-Beispiel.....	29
Abbildung 4.3: Kreuzblick.....	29
Abbildung 4.4: Kreuzblick-Beispiel .....	29
Abbildung 4.5: Graues Anaglyphenbild .....	29
Abbildung 4.6: Farbiges Anaglyphenbild .....	29
Abbildung 4.7: Prinzip der Polarisierungstechnik .....	29
Abbildung 4.8: Linear und zirkular polarisierte Welle .....	29
Abbildung 4.9: Progressiver Bildaufbau einer Bildröhre.....	29
Abbildung 4.10: Shutterbrille.....	29

Abbildung 4.11: Bildaufbau mit Zeilensprung.....	29
Abbildung 4.12: Linsenrasterbildschirm .....	29
Abbildung 4.13: Display mit Parallaxenbarriere.....	29
Abbildung 4.14: Head Mounted Display (5DT HMD 800) [19] .....	29
Abbildung 4.15: 3D-Display von Kodak .....	29
Abbildung 5.1: Nebeneinander-Layout.....	29
Abbildung 5.2: Übereinander-Layout.....	29
Abbildung 5.3: Vertikales Interlaced-Layout.....	29
Abbildung 5.4: Prädiktion der zweiten Perspektive in einem stereoskopischen Videocodec...	29
Abbildung 5.5: Sensio 3D-Prozessor .....	29
Abbildung 6.1: Regisseur James Cameron mit digitaler 3D-Kamera .....	29
Abbildung 6.2: Consumer-Kameras und ste-fra CAM-Halterung [6].....	29
Abbildung 6.3: ste-fra LANC-Steuerung.....	29
Abbildung 6.4: Synchronisation auf Frame-Ebene .....	29
Abbildung 6.5: Progressives versus fieldsequentielles Video .....	29
Abbildung 6.6: Synchronisation auf Field-Ebene.....	29
Abbildung 6.7: Blockschaltbild eines analogen Videomultiplexers .....	29
Abbildung 6.8: Aufbereitung fieldsequentieller Aufnahmen für progressive Wiedergabe .....	29
Abbildung 6.9: Passiver Bildteiler (aus US-Patent 2.282.947 von Ingvar E. de Cherbinin).....	29
Abbildung 6.10: Nu-View-Adapter.....	29
Abbildung 6.11: Nu-View Funktionsprinzip.....	29
Abbildung 7.1: Typischer Filtergraph zur Wiedergabe einer MPEG-1 Videodatei .....	29
Abbildung 8.1: Medientypen in einem MPEG-1 Wiedergabegraph .....	29
Abbildung 8.2: 3D-Wiedergabegraph einer MPEG-I-Datei.....	29
Abbildung 8.3: 3D-DVD-Wiedergabegraph .....	29
Abbildung 8.4: Klassenhierarchie der Wiedergabegraphen.....	29
Abbildung 8.5: Bibliothek des Stereoscopic Players .....	29
Abbildung 8.6: UML-Diagramm der Bibliotheksklassen.....	29
Abbildung 8.7: Shutterbrillen-Controller <i>3D Control</i> .....	29
Abbildung 9.1: Sony DFW-VL500 Kameras.....	29
Abbildung 9.2: Eigenschaften des DFW-VL500 Treibers .....	29
Abbildung 9.3: Aufnahmegraph .....	29
Abbildung 9.4: Zugriff auf Aufnahmegeräte.....	29
Abbildung 9.5: Konfigurationsdialog des Stereoscopic Multiplexers.....	29
Abbildung 9.6: Testbild zur Messung der Kamerasynchronisation.....	29
Abbildung 9.7: Testbild, mit zwei Kameras vom Monitor abgefilmt.....	29
Abbildung 9.8: Korrekturkennlinie des Iris-Parameters .....	29
Abbildung 9.9: Tabelle zur Parameterkorrektur .....	29
Abbildung 9.10: Eigenschaftsseite zum Einstellen der Bildparameter .....	29
Abbildung 9.11: Bidirektionale Parametersynchronisation.....	29

Abbildung 9.12: Allgemeine 3x3 Filtermaske und spezielle Maske zur Unschärfmaskierung	29
Abbildung 9.13: Bildverarbeitungsoptionen des Stereoscopic Multiplexers .....	29
Abbildung A.1: Konfigurationsdialog des NVIDIA Stereo Treibers.....	29
Abbildung A.2: Konfigurationsdialog des Stereoscopic Players .....	29
Abbildung A.3: Fehlermeldung bei fehlendem Codec.....	29
Abbildung A.4: Hauptfenster des Stereoscopic Players.....	29
Abbildung A.5: Eigenschaftsseiten eines Videos .....	29

---

# Tabellenverzeichnis

---

Tabelle 3.1: Typische Präsenzwerte.....	29
Tabelle 9.1: Betriebsmodi der Sony DFW-VL500 Kamera .....	29
Tabelle A.1: Kommandozeilenparameter – Quellen.....	29
Tabelle A.2: Kommandozeilenparameter – Eingangsformat .....	29
Tabelle A.3: Kommandozeilenparameter – Wiedergabe .....	29
Tabelle B.1: Vordefinierte Wurzelkategorien .....	29
Tabelle B.2: Aufbau von stereoskopischen Informationsdateien.....	29
Tabelle B.3: Aufbau eines Kategorie-Blocks.....	29
Tabelle B.4: Aufbau eines Videoinformation-Blocks .....	29

---

## Literatur- und Linkverzeichnis

---

- [1] 1394 Trade Association: AV/C Camera Subunit Specification, 1999
- [2] 1394 Trade Association: IIDC 1394-based Digital Camera Specification (Ver. 1.3), 1999
- [3] 3DTV Corp.  
*<http://www.3dmagic.com>*
- [4] Apple Computer, Inc.: Quicktime  
*<http://www.quicktime.com>*
- [5] Barco N.V.: Stereoscopic Projection – 3D Projection Technology  
*[http://www.barco.com/projection\\_systems/downloads/Barco\\_stereoscopic\\_proj.pdf](http://www.barco.com/projection_systems/downloads/Barco_stereoscopic_proj.pdf)*
- [6] Bloos, Werner: ste-fra LANC-Steuerung und ste-fra CAM-Halterung  
*<http://www.digi-dat.de>*
- [7] Bungert, Christoph: stereo3d.com  
*<http://www.stereo3d.com>*
- [8] Campenhausen, Christoph von: Die Sinne des Menschen (2. Aufl.). Stuttgart; New York: Georg Thieme Verlag, 1993
- [9] Ciccarello, Liborio; Berger, Dirk: Tiefenwahrnehmung durch binokulare Stereopsis. Seminararbeit, Universität Mannheim  
*<http://www.uni-mannheim.de/fakul/psycho/irtel/lehre/seminararbeiten/w96/Tiefe/binoc.html>*
- [10] ColorCode 3-D Center ApS  
*<http://www.colorcode3d.com>*
- [11] CYVIZ AS  
*<http://www.cyviz.com>*
- [12] Dynamic Digital Depth  
*<http://www.ddd.com>*
- [13] del Río, Angel; Bartz, Dirk; Jäger, R.; Gürvit, Özlem; Freudenstein, Dirk: Efficient Stereoscopic Rendering in Virtual Endoscopy Applications. In: Proc. of WSCG, pp. 95-101, 2003
- [14] Deutsche Gesellschaft für Stereoskopie e.V.: DGS-Qualitätsstandards für die Kleinbild-Stereodiaprojektion  
*[http://home.t-online.de/home/agp.herbig/download/dgs\\_standard.pdf](http://home.t-online.de/home/agp.herbig/download/dgs_standard.pdf)*
- [15] DivXNetworks, Inc.: DivX Codec  
*<http://www.divx.com>*

- [16] Dodgson, Neil A.: Autostereo Displays: 3D without Glasses. In: EID '97 (Electronic Information Displays), Esher, Surrey, 18<sup>th</sup>-20<sup>th</sup> November 1997
- [17] Dubois, Eric: A Projection Method to Generate Anaglyph Stereo Images. In: Proc. IEEE Int. Conf. Acoustics Speech Signal Processing, Salt Lake City, Vol. 3, pp. 1661-1664, May 2001
- [18] EON Reality, Inc.  
*<http://www.eonreality.com>*
- [19] Fifth Dimension Technologies, Inc.  
*<http://www.5dt.com>*
- [20] Goldstein, E. Bruce: Sensation and Perception (6<sup>th</sup> Ed.). Belmont: Wadsworth Publishing, 2001
- [21] Gonzalez, Rafael C.; Woods, Richard E.: Digital Image Processing. Second Edition. Upper Saddle River: Prentice Hall, 2002
- [22] GStreamer Open Source Multimedia Framework  
*<http://www.gstreamer.net>*
- [23] Guliverkli Project  
*<http://sourceforge.net/projects/guliverkli>*
- [24] Hagemann, Peter A.: Der 3D Film. München: Verlag Monika Nüchtern, 1980
- [25] Hauerslev, Thomas (Hrsg.): in70mm.com  
*<http://www.in70mm.com>*
- [26] Hawkins, Rhys: Digital Stereo Video: Display, Compression and Transmission. Master Thesis, Australian National University, February 2002  
*[http://escience.anu.edu.au/research/papers/02\\_Rhys\\_Hawkins/thesis-small.pdf](http://escience.anu.edu.au/research/papers/02_Rhys_Hawkins/thesis-small.pdf)*
- [27] Herbig, Gerhard P.: Anmerkungen zur Stereofotografie mit verkleinerter oder vergrößerter Basis  
*[http://home.t-online.de/home/agp.herbig/download/variable\\_stereobasis.pdf](http://home.t-online.de/home/agp.herbig/download/variable_stereobasis.pdf)*
- [28] Herbig, Gerhard P.: Die 3 goldenen Regeln der Stereofotografie  
*[http://home.t-online.de/home/agp.herbig/download/goldene\\_regeln.pdf](http://home.t-online.de/home/agp.herbig/download/goldene_regeln.pdf)*
- [29] Herbig, Gerhard P.: Die 70 Minuten-Bedingung und ihre Folgen  
*<http://home.t-online.de/home/agp.herbig/download/70-minuten.pdf>*
- [30] Herbig, Gerhard P.: Stereofotografie heute  
*[http://home.t-online.de/home/agp.herbig/download/stereofotografie\\_heute.pdf](http://home.t-online.de/home/agp.herbig/download/stereofotografie_heute.pdf)*
- [31] Herbig, Gerhard P.: Tiefenwahrnehmung bei der Stereoprojektion  
*<http://home.t-online.de/home/agp.herbig/download/tiefenwahrnehmung.pdf>*
- [32] Herbig, Gerhard P.: Zur Wahl der richtigen Stereobasis  
*<http://home.t-online.de/home/agp.herbig/download/stereobasis.pdf>*
- [33] Hušák, Michal; Chris Ward: DepthQ: Universal System for Stereoscopic Video Visualization on Win32 Platform. In: Proceedings of SPIE Vol. 5291A Stereoscopic Displays and Virtual Reality Systems XI

- [34] IMAX Corp.  
<http://www.imax.com>
- [35] i-O Display Systems, LLC:  
<http://www.i-glassesstore.com>
- [36] Johanson, Mathias: Stereoscopic Video Transmission over the Internet. Framkom Research Corporation, presented at WIAPP'01 in San Jose, July 2001  
<http://confero.alkit.se/stereovideo/wiapp-paper.pdf>
- [37] Kuhn, Gerhard: Stereofotografie und Raumbildprojektion. Gilching: vfv Verlag für Foto, Film und Video, 1999
- [38] Lespinasse, Michel; Holtzman, Aaron: libmpeg2  
<http://libmpeg2.sourceforge.net>
- [39] Limann, Otto; Pelka, Horst: Fernsehtechnik ohne Ballast – Einführung in die Schaltungstechnik der Fernsehempfänger. 18. Aufl. Feldkirchen: Franzis, 1996
- [40] Lightspeed Design, Inc.  
<http://www.lightspeeddesign.com>
- [41] Lipton, Lenny: Foundations of the Stereoscopic Cinema. New York: Van Nostrand Reinhold Company, 1982  
<http://info.curtin.edu.au:8080/~iwoods/stereoscopic/library/foundation.cgi>
- [42] Lüscher, Hermann: Stereophotographie. Berlin: Union Deutsche Verlagsgesellschaft, 1931
- [43] McCullough, Hugh: Cinephoto.co.uk – IMAX  
<http://www.cinephoto.co.uk/imax.htm>
- [44] Microsoft Corp.: DirectX 9  
<http://www.microsoft.com/directx>
- [45] Microsoft Corp.: DirectX 9 Software Development Kit  
[http://msdn.microsoft.com/library/en-us/directx9\\_c/directx/directx9cpp.asp](http://msdn.microsoft.com/library/en-us/directx9_c/directx/directx9cpp.asp)
- [46] Microsoft Corp.: OpenGL  
<http://msdn.microsoft.com/library/en-us/dnanchor/html/opengl.asp>
- [47] Microsoft Corp.: Windows Media 9 Series  
<http://www.microsoft.com/windowsmedia>
- [48] Microsoft Corp.: Windows Media Format 9 Series Software Development Kit  
<http://msdn.microsoft.com/library/en-us/wmform/html/introducingwindowsmediaformat.asp>
- [49] Moving Picture Experts Group  
<http://www.chiariglione.org/mpeg>
- [50] Neotek  
<http://www.neotek.com>
- [51] Network-Integrated Multimedia Middleware  
<http://graphics.cs.uni-sb.de/NMM>

- [52] NVIDIA Corp.  
*<http://www.nvidia.com>*
- [53] NVIDIA Corp.: 3D Stereoscopic Development Guide, August 2001  
*[http://developer.nvidia.com/object/3D\\_Stereoscopic\\_Dev.html](http://developer.nvidia.com/object/3D_Stereoscopic_Dev.html)*
- [54] Ohm, Jens-Rainer: Stereo/Multiview Video Encoding Using the MPEG Family of Standards. In: Proc. of SPIE Vol. 3639, Stereoscopic Displays and Virtual Reality Systems VI, May 1999
- [55] OpenGL Spezifikation  
*<http://www.opengl.org>*
- [56] optics.org: Kodak unveils 3D immersive display. March 2004  
*<http://optics.org/articles/news/10/3/15/1>*
- [57] Philip Kurzenberger e.K.  
*<http://www.3d-brillen.de>*
- [58] RealNetworks, Inc.  
*<http://www.realnetworks.com>*
- [59] Reimers, Ulrich (Hrsg.): Digitale Fernsehtechnik: Datenkompression und Übertragung für DVB. Berlin; Heidelberg: Springer, 1995
- [60] Rittermann, Marco; Schuldt, Michael: 3D Television Production Based on MPEG-4 Principles. In: 11<sup>th</sup> International Conference in Central Europe on Computer Graphics, Visualization, and Computer Vision 2003; Plzen; 3<sup>rd</sup>-7<sup>th</sup> February, WCSG 2003 Posters
- [61] Sand, Rüdiger: 3-DTV Research and Development in Europe. In: Proc. of SPIE Vol. 1457, Stereoscopic Displays and Applications II, August 1991
- [62] Sanders, William; McAllister, David F.: Producing Anaglyphs from Synthetic Images. In: Proceedings Electronic Imaging '03, Santa Clara, January 2003
- [63] Scheiwiller, Peter M.; Dumbreck, Andrew A.; Chapman A. D.: Videotape Recording of 3D Television Pictures. In: Proc. of SPIE Vol. 1256, Stereoscopic Displays and Applications, September 1990
- [64] Scheiwiller, Peter M.; Murphy S. P.; Dumbreck, Andrew A.: A Compact Zoom Lens for Stereoscopic Television. In: Proc. of SPIE Vol. 1457, Stereoscopic Displays and Applications II, 1991
- [65] Schmidt, Ulrich: Digitale Film- und Videotechnik. München; Wien: Fachbuchverlag Leipzig im Carl Hanser Verlag, Juli 2002
- [66] Sony Corp.: DFW-VL500 Technical Manual (Ver. 1.0)
- [67] Sony Corp.: Sony 1394 Camera WDM Driver Readme File, January 2002
- [68] Steinmetz, Ralf; Nahrstedt, Klara: Multimedia Fundamentals Volume 1 – Media Coding and Content Processing. Upper Saddle River: Prentice Hall, 2002
- [69] StereoGraphics Corp.  
*<http://www.stereographics.com>*

- [70] Sullivan, Gary; Estrop, Stephen: Video Rendering with 8-Bit YUV Formats, August 2003  
<http://msdn.microsoft.com/library/en-us/dnwm/html/YUVFormats.asp>
- [71] Suto, Masuji: StereoMovie Player  
<http://www3.zero.ad.jp/esuto/stvply/indexe.htm>
- [72] Tseng, Belle L.; Anastassiou, Dimitris: Perceptual Adaptive Quantization of Stereoscopic Video Coding Using MPEG-2's Temporal Scalability Structure. In: International Workshop on Stereoscopic and Three Dimensional Imaging (IWS3DI'95), Santorini, September 1995
- [73] TEG Sensorial Technologies, Inc.: Sensio 3D  
<http://www.sensio3d.com>
- [74] Third Dimension Technologies, Inc.  
<http://www.realtime-3d.com>
- [75] Ulrich, Iwan; Nourbakhsh, Illah: Firewire Untethered: High-Quality Images for Notebook Computer. In: Advanced Imaging Magazine, January 2000, pp. 69-70
- [76] Vigovsky Alexander: AC3Filter  
<http://ac3filter.sourceforge.net>
- [77] VRex, Inc.  
<http://www.vrex.com>
- [78] Waack, Fritz G.; Herbig, Gerhard P. (Überarb.): Allgemeine technische Begriffe der Stereoskopie (Raumbildwesen)  
<http://home.t-online.de/home/agp.herbig/download/stereoglossar.pdf>
- [79] Wambach, Patrick: Formatdefinition und Formatverhandlung von Multimedia-Geräten. Diplomarbeit, Universität des Saarlandes  
<http://graphics.cs.uni-sb.de/NMM/Publications/thesis/wambach2001.pdf>
- [80] Wittkowsky, Wilfried: Der 3D-Film  
<http://www.wittkowsky.net/3d-film>
- [81] Wimmer, Peter: GPL MPEG-1/2 Decoder  
<http://sourceforge.net/projects/gplmpgdec>
- [82] Wimmer, Peter: Stereoscopic Player  
<http://mitglied.lycos.de/stereo3d>
- [83] Woods, Andrew J.; Rourke, Tegan: Ghosting in Anaglyphic Stereoscopic Images. In: Stereoscopic Displays and Virtual Reality Systems XI, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 5291, San Jose, 2004
- [84] Wright, Richard S.; Sweet, Michael R.: OpenGL SuperBible. 2<sup>nd</sup> Edition. Corte Madera: Waite Group Press, 1999
- [85] Xvid Team: Xvid Codec  
<http://www.xvid.org>

Alle Verweise auf Seiten des Internets gültig am 23. Mai 2004.

## Lebenslauf

### *Persönliche Angaben:*

Name: Peter Wimmer  
Geburtsdatum: 23. November 1979  
Familienstand: Ledig  
Nationalität: Österreich  
Eltern: Ing. Maximilian Wimmer, Landesbediensteter  
Anna Wimmer, kaufmännische Angestellte  
Geschwister: Christian Wimmer, Informatikstudent  
Isabella Wimmer, Schülerin

### *Ausbildung:*

1986 - 1990 Öffentliche Volksschule 9 in 4020 Linz  
1990 - 1994 Bundesgymnasium  
4020 Linz, Khevenhüllerstraße 1  
1994 - 1998 Bundesrealgymnasium  
4020 Linz, Fadingerstraße 4  
Juni 1998 Matura mit ausgezeichnetem Erfolg  
seit 1998 Studienrichtung Mechatronik  
Johannes Kepler Universität Linz  
April 2001 Erste Diplomprüfung mit ausgezeichnetem Erfolg

### **Eidesstattliche Erklärung**

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Linz, Mai 2004

Peter Wimmer