

Ein Erweitertes Kameramodell Methoden und Werkzeuge für die dynamische Kamerasteuerung in interaktiven Systemen

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur (Dr.-Ing.)

angenommen durch die Fakultät für Informatik
der Otto-von-Guericke-Universität Magdeburg

von
geboren am

Dipl.-Inf. Ralf Helbing
26. 11. 1968 in Burg

Gutachterinnen/Gutachter:

Prof. Dr. Thomas Strothotte
Prof. Dr. Elisabeth André
Prof. Dr. Bernhard Preim

Promotionskolloquium:

Magdeburg, den 23. Januar 2004

Danksagung

Zu Beginn möchte ich allen danken, die mir auf die eine oder andere Weise beim Schreiben dieser Arbeit geholfen haben. Dazu gehört besonders mein Betreuer Prof. Dr. Thomas Strothotte, von dem ich viel lernen konnte und mein Mitstreiter Dr. Knut Hartmann. Ohne ihre Hilfe wäre diese Arbeit wahrscheinlich nie entstanden.

Dank sagen möchte ich auch dem Kollektiv des Instituts für Simulation und Grafik, besonders Petra Specht, die immer dafür gesorgt hat, daß man sich am ISG wohlfühlen konnte. Vielen Dank auch an die Techniker.

Mein Dank gilt auch den Gutachtern Prof. Dr. Thomas Strothotte, Prof. Dr. Elisabeth André und Prof. Dr. Bernhard Preim für ihre Bereitschaft zur Begutachtung und ihre kritischen Hinweise.

Während der Zusammenarbeit und bei zahlreichen Diskussionen mit meinen Kollegen habe ich viele wertvolle Anregungen, Hinweise und Kritiken erhalten. Besonders erwähnen möchte ich hier Nick Halper, Felix Ritter, Dr. Stefan Schlechtweg, Dr. Maic Masuch, Dr. Michael Rüger, Dr. Axel Hoppe, Dr. Alf Ritter, Dr. Steffi Beckhaus und Dr. Martin Scholz.

Mein Dank geht auch an alle Studenten und Kollegen, deren Programmiergeschick zum VISDOX-System beigetragen hat, besonders: Dr. Knut Hartmann, Dr. Andreas Raab, Jens Fänger, Roland Jesse und Thomas Witzel.

Nicht zuletzt möchte ich auch meinen Eltern, Freunden und Verwandten danken, die mir immer wieder Kraft und Ausdauer bei der Arbeit gaben.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Anwendungsfelder von Kameramodellen	4
1.1.1	Technische Dokumentation	4
1.1.2	Unterhaltungsbereich	5
1.1.3	Nicht-Photorealistische Anwendungen	6
1.1.4	Kameraplaner	7
1.1.5	Simulation optischer Eigenschaften	8
1.1.6	Zusammenfassung	10
1.2	Ziele der Arbeit	11
1.3	Ergebnisse der Arbeit	12
1.4	Gliederung der Arbeit	14
2	Ein Konzept für ein erweitertes Kameramodell	17
2.1	Modelle in der Computergrafik	18
2.2	Intelligente Interaktive Systeme	20
2.3	Die intelligente Kamera als Ausdrucksmittel	21
2.4	Basisdaten	24
2.5	Darstellungsstile	26
2.6	Zeit	27
2.7	Interaktivität und Steuerung	29
2.8	Personalität	32
2.9	Zusammenfassung	35
3	Hervorhebungstechniken in interaktiven Anwendungen	37
3.1	Begriffe	38
3.1.1	Hervorhebungen	38

3.1.2	Akzentuierungen	39
3.1.3	Direkte und indirekte Hervorhebungen	40
3.1.4	Lokalität von Hervorhebungen	41
3.1.5	Dynamische Hervorhebungen	41
3.1.6	Situationen	43
3.2	Sichtbarkeit	44
3.2.1	Ausblenden verdeckender Objekte	45
3.2.2	Transparenz	46
3.2.3	Drahtgitter	48
3.2.4	Zusammenfassung	49
3.3	Zeichenattribute zur Hervorhebung	50
3.3.1	Farbe	51
3.3.2	Zeichenstil	52
3.3.3	Licht und Schatten	52
3.4	Abstraktion	53
3.5	Unschärfe	54
3.6	Zusammenfassung	55
4	Filmische Gestaltungsmittel	59
4.1	Filmsprache	60
4.2	Filmtechniken außerhalb des Films	62
4.2.1	Interaktive Technische Dokumentationen	62
4.2.2	Interaktive Entertainmentsysteme	64
4.3	Hervorhebungsmittel im Film	66
4.4	Zusammenfassung	72
5	Planung von Kamerafahrten	75
5.1	Betrachterposition	76
5.1.1	Konflikte	80
5.1.2	Sonderfall Sichtbarkeit	82
5.1.3	Genetische Algorithmen	83
5.1.4	Bildbasierte Methoden	84
5.2	Reaktive Kameraplanung	85
5.3	Kamerafahrten	88
5.3.1	Roadmap-Methoden	89
5.3.2	Potentialfelder	91

5.3.3	Gemeinsamkeiten und Unterschiede von Robotik und Kameraplanung	92
5.3.4	Kameraplanung in interaktiven Systemen	94
5.4	Umweg: Ein neues Verfahren zur Planung von Kamerafahrten	95
5.4.1	Anforderungen an Kamerafahrten in interaktiven technischen Dokumentationen	95
5.4.2	Umfahren von Hindernissen durch Ablenkung	96
5.4.3	Nachbearbeitung des Kamerapfades	101
5.4.4	Wahl der Blickrichtung	104
5.4.5	Fazit	106
5.5	Zusammenfassung	106
6	Fallstudien	109
6.1	VISDOK	110
6.1.1	Technische Dokumentation und Computergrafik	111
6.1.2	VISDOK als multimediales Präsentationssystem	112
6.1.2.1	Verbindung von 3D-Grafik und Wissensbasis	113
6.1.2.2	Medienpräsentation und Interaktionsmanagement	116
6.1.2.3	Medienrealisierung	117
6.1.3	Interaktionsplanung	119
6.1.4	Beispiel: Eine VISDOK-Sitzung	122
6.1.5	Hervorhebungen	124
6.1.5.1	Schattierte Polygone	124
6.1.5.2	Liniengrafik	125
6.1.5.3	Kamerafahrten	126
6.1.5.4	Schatten	127
6.1.5.5	Sichtbarkeit	127
6.1.6	Situationen in VISDOK	128
6.1.7	Das Erweiterte Kameramodell in VISDOK	130
6.2	AGI ³ LE	134
6.2.1	Wissensbasierte Textanalyse und Illustration	135
6.2.2	Die Intelligente Kamera in AGI ³ LE	137

7	Ausblick: Kamerasteuerung in Unterhaltungssoftware	143
7.1	Trends in Computerspielen	144
7.2	Filmtechniken im Computerspiel	146
7.2.1	Abkehr von der Ich-Form	146
7.2.2	Einbettung der Kamerasteuerung	147
7.3	Anforderungen an eine Kamerasteuerung für Computerspiele	149
7.4	Architektur einer Kamerasteuerung für Computerspiele . . .	151
7.4.1	Beispiel Dialog	153
7.4.2	Beispiel Verfolgung	155
7.5	Zusammenfassung	156
8	Zusammenfassung	159
A	Keramamodelle in 3D-Programmierbibliotheken	163
B	Potential Visibility Regions	171

Seit in Computern grafische Daten verarbeitet werden, muß auch eine Sicht auf diese Daten beschrieben werden, die festlegt, wohin und worauf man schaut, also: was sichtbar sein soll und was nicht. Dies trifft in besonderem Maße auf dreidimensionale Daten zu, die sich von verschiedenen Seiten betrachten lassen. So liegt es nahe, die Sicht auf die Welt durch eine Kamera-Metapher zu beschreiben.

Kameras sind den meisten Menschen vertraute, gut etablierte technische Geräte und technologisch wesentlich älter als Computer.¹ Damit drängen sich die Kamera als Modell auf, um Daten zu beschreiben, die auch für die 3D-Darstellung am Computer benötigt werden. Wie bei allen Metaphern aber sind auch hier Gemeinsamkeiten und Unterschiede zwischen dem der Metapher zugrundeliegenden Begriff *Kamera* und seiner Anwendung in der Computergrafik zu beachten.

Traditionell wird eine Kamera in der Computergrafik in enger Analogie zur echten Kamera beschrieben und aufgefaßt. Obgleich diese in ihrer über viele Jahre andauernden technischen Entwicklung zahlreiche Veränderungen erfahren hat und noch erfährt, bezieht man sich dabei auf sehr grundlegende Eigenschaften, die sich zudem leicht auf die Bedürfnisse der 3D-Grafik übertragen lassen. Das betrifft in erster Linie die Punkte: „Wo stehe ich, wohin

¹ ohne die rasante Weiterentwicklung besonders bei den Digitalkameras in Abrede zu stellen

blicke ich?“ [BLINN 1988]. Die Antwort auf diese Fragen, gegeben in Form einiger Zahlen (Koordinaten), ist für die Computergrafik schlicht eine mathematische Notwendigkeit, gehen diese Koordinaten doch in den der Bilddarstellung zugrundeliegenden Rechenprozeß ein.

Diese Grundmotivation der Kamerametapher aus den Anfangstagen der Computergrafik hat nach wie vor ihre Gültigkeit. Jedoch eignen sich traditionelle geometrische Kameramodelle aus Position und Blickrichtung nur für die einfachsten Anwendungen und werden den vielfältigen Möglichkeiten moderner Computergrafik nicht gerecht, da hierfür mehr Wissen über die Kamera erforderlich ist als nur ihre Position und Blickrichtung. So könnte – um in der engen Analogie zur Fotokamera zu bleiben – in einer Anwendung zeitweise ein etwas über- oder unterbelichtetes Bild wünschenswert sein, ohne daß der Benutzer deswegen die Helligkeit des Monitors ändern soll. Softwaretechnisch wäre eine Helligkeitsänderung machbar. Ein rein geometrisches Kameramodell allerdings kann diese Einstellmöglichkeit nicht bieten, müßte also erweitert werden.

Computergrafik wird heute in verschiedensten Softwaresystemen mit teilweise sehr unterschiedlicher Zielstellung eingesetzt. Daraus ergeben sich ebenso unterschiedliche Bedürfnisse und Ansprüche an das verwendete Kameramodell, so daß nicht immer alle Eigenschaften einer realen Kamera eine sinnvolle Entsprechung in der Software finden. So kann es sein, daß etwa immer der gleiche Kamerastandpunkt eingenommen wird und die Software nur andere Blickwinkel zuläßt. Darstellungsalgorithmen könnten in solchen Fällen spezialisiert, vereinfacht und optimiert, nicht benötigte Aspekte einer Kamera könnten vernachlässigt werden. Die Verwendung der Kamera-Metapher geht immer nur so weit, wie es für die Anwendung erforderlich ist.

Auf der anderen Seite stellt die (Foto-)Kamera in Wirklichkeit ein komplexes technisches System dar, bestehend aus Gehäuse mit Kameramechanik (z. B. Verschlußmechanismus und Meßelektronik) sowie den ebenfalls komplexen Objektiven. Auch die Eigenschaften dieser Kamerakomponenten wirken sich auf die Bilderzeugung aus. Da die Kamera letztlich einen Film belichtet, gehen auch dessen Eigenschaften mit in das Bild ein. Hinzu kommen scheinbar nebensächliche Eigenschaften wie Gewicht, was seine Auswirkungen nicht so sehr beim Fotografieren, sondern mehr beim Filmen zeigt, wo die Kamera nicht selten bewegt werden muß.

In der noch heute andauernden Entwicklung immer realistischerer Verfahren in der Computergrafik wird unter anderem das Ziel verfolgt, den in der Realität ablaufenden Bilddarstellungsprozeß mit all seinen zugrundeliegenden physikalischen Gesetzmäßigkeiten zu simulieren. Dabei spielen auch die obengenannten, subtileren Eigenschaften der Kamera eine wichtige Rolle, sei es bei der Simulation von Unschärfe [HEIDRICH et al. 1997] oder den Verzerrungseffekten unterschiedlicher Linsensysteme [MUSGRAVE 1987, POTMESIL und CHAKRAVARTY 1982, KOLB et al. 1995].

Nicht alle dieser Eigenschaften lassen sich in der Computergrafik tatsächlich oder mit vertretbarem Aufwand umsetzen; einige sind sogar – wie im Verlauf der Arbeit gezeigt wird – oft unerwünscht. Jedoch bestimmen sie die Möglichkeiten und Grenzen der traditionellen Fotografie und Kinematografie. Vielfach bilden diese technischen Unzulänglichkeiten sogar die Grundlage für besondere Ausdrucksmittel, wie im Falle der Unschärfe, die gezielt eingesetzt werden kann, um den Fokus eines Bildes zu setzen. Die Situation, daß hierbei aus der Not eine Tugend gemacht wird, ist auch in vielen anderen Gebieten der Kunst anzutreffen: so hat z. B. der Umstand, daß bei Strichzeichnungen bzw. Radierungen die Intensität der Linien nur schwer zu variieren ist, zur Entstehung von Schraffurtechniken geführt. Letztlich entsteht der besondere Reiz einer Kunstform oft aus „technischen Beschränkungen“ ihrer Ausdrucksmittel.

In Analogie dazu werden die Darstellungsmöglichkeiten einer Grafiksoftware in starkem Maße von ihrem Kameramodell und dessen Umsetzung bestimmt. Dieses Repertoire ist weit umfangreicher als nur die Positionsbestimmung der Kamera. Darüber hinaus löst sich die Computergrafik sowohl in den Möglichkeiten der virtuellen Kamera als auch im Kameramodell von den Beschränkungen der realen Kamera. So erlaubt die beliebige Positionierbarkeit einer im Raum frei schwebenden virtuellen Kamera Perspektiven und Kamerabewegungen, die in der Realität nicht oder nur mit hohem Aufwand möglich sind.

1.1 Anwendungsfelder von Kameramodellen

Bereits bei der Herleitung mathematischer Formeln aus der analytischen Geometrie, die z. B. der Projektion von 3D-Bildern auf eine zweidimensionale Fläche zugrundeliegen, werden bewußt oder unbewußt Vorstellungen einer einfachen Kamera oder der Funktion des menschlichen Auges Pate gestanden haben. Im Umfeld der Computergrafik finden Kameramodelle erstmals Erwähnung im Zusammenhang mit dem CORE-System [DUNN und HERZOG 1977], wobei hier mehr die Metapher der Kamera selbst eine Rolle spielt und nicht die zu simulierenden Eigenschaften einer Fotokamera. Die Benutzung der Kamerametapher in der Literatur (z. B. in Nachschlagewerken oder Lehrbüchern zur Computergrafik) hat aber sicher dazu beigetragen, auch noch andere Kameraeigenschaften zu betrachten als nur die, die man zur Erklärung von in der Computergrafik benutzten Algorithmen braucht.

In den folgenden Abschnitten wird daher eine Reihe von Anwendungsfeldern der Computergrafik vorgestellt, die jeweils besondere Anforderungen an unterstützte Darstellungsmerkmale und Kameraeigenschaften stellen.

1.1.1 Technische Dokumentation

In der technischen Dokumentation werden *informierende Bilder* [WEIDENMANN 1993] zur Illustration oder zur Abbildung technischer Geräte genutzt. Diese Bilder kommunizieren Erscheinungsmerkmale von Gegenständen und verdeutlichen Zusammenhänge, häufig in enger Abstimmung mit dazugehörigen geschriebenen oder generierten Texten [FEINER 1985]. Konventionelle technische Dokumentationen setzen oft schematische oder anderweitig vereinfachte Darstellungen ein [SCHOLZ 1999]. Abbildungen werden auf wesentliche Merkmale beschränkt, um die Aufmerksamkeit des Betrachters zu lenken. In interaktiven technischen Dokumentationen besteht aufgrund der Geschichte von Technischen Dokumenten (die ja in ihrer Bildsprache auf traditionellen Illustrationen aufbauen) eine enge Beziehung zur nicht-photorealistischen Grafik, deren im Vergleich zum Photorealismus stärkere Ausdruckskraft für die Vermittlung kommunikativer Ziele eingesetzt wird.

Die Arbeit eines Nutzers mit einer interaktiven technischen Dokumentation wird oft in Anlehnung an ein Gespräch zwischen Personen als *Diskurs* be-

zeichnet. Dementsprechend versucht auf Seiten des Computers eine *Diskursplanung*, bestimmte Regeln der zwischenmenschlichen Kommunikation zu simulieren. Im Verlauf eines Diskurses ergeben sich ständig neue Situationen der Wissensvermittlung (z. B. der Wissensstand des Nutzers erweitert sich, die Erklärung schreitet fort, neue Aspekte rücken in den Blickpunkt). In interaktiven Systemen kann die Abbildung laufend den sich ändernden Diskursituationen angepaßt werden, die im Verlauf der Inhaltsvermittlung auftreten. Da die Diskursplanung selbst auf einer hohen Abstraktionsebene und nicht in computergrafischen Kategorien wie Geometrie und Zeichenstil arbeitet, wird eine *intelligente Kamera* benötigt, die in der Realisierung der Diskursstrategie die Vorgaben der Planung interpretiert und umsetzt. In Verbindung mit nicht-photorealistischen Techniken ist es hier besonders hilfreich, wenn Eigenschaften wie Linienstile und andere Hervorhebungsmittel nicht im geometrischen Modell festgelegt sind, sondern dynamisch von der Kamera verwaltet und an die aktuelle Situation angepaßt werden.

Ein besonderes Merkmal interaktiver Systeme im Vergleich zu bedrucktem Papier ist die Möglichkeit einer dynamischen Präsentation, etwa mit Hilfe von Animationen. Das Hinzukommen einer zeitlichen Dimension wirft zugleich neue Aspekte der Kameraführung auf, wie etwa ihr Verhalten bei einer Kamerafahrt oder die vielfältigen Gestaltungsfreiheiten eines zeitlichen Ablaufs.

1.1.2 Unterhaltungsbereich

Bereits im Kinofilm wird die Art der Kameraführung benutzt, um Stimmungen zu vermitteln. Wie bereits im vorigen Abschnitt sind auch hier die Kameraeigenschaften dynamischer Natur und beschreiben das Verhalten der Kamera über mehrere Bilder bzw. Bildsequenzen hinweg.

Besonders 3D-Computerspiele werden derzeit meist in der Ich-Form gespielt, d. h. sie benutzen sogenannte *first-person*-Kamerasteuerungen (siehe auch Abschnitt 2.8), bei denen der Spieler mit der Spielfigur verschmilzt und die Kamera vom Spieler direkt gesteuert wird. Aber auch in Spielen (oder ganz allgemein, in interaktiven Systemen) kann die Kamera benutzt werden, um der aktuellen Situation entsprechend verschiedene filmische Gestaltungsmittel anzuwenden. Dazu zählt die Wahl der Perspektive, die Abfolge und Montage

verschiedener Einstellungen, oder auch die Art der Kamerabewegung. Voraussetzung dafür ist allerdings eine Abkehr vom Prinzip der Ich-Form, so daß sich die Kamera wieder losgelöst von der Spielfigur bewegen kann.

Anwendungen, die dem Nutzer die Kamerasteuerung aus der Hand nehmen, um so die Art der Kameraführung als zusätzliches Gestaltungselement zu erhalten, stehen grundsätzlich vor dem Problem, einen für die jeweilige Anwendung spezifischen Situationsbegriff zu entwickeln. Aus diesem läßt sich dann ableiten, wodurch Situationen charakterisiert sind und wie sie sich auf die grafische Präsentation etwa mit Mitteln der Kameraführung auswirken sollen. Dies kann auf verschiedene Arten geschehen: TOMLINSON et. al. beschreiben ein interaktives System mit verschiedenen Spielfiguren, bei dem sich ähnlich wie im Film die Stimmung einer Person über die Kamera mitteilt [TOMLINSON et al. 2000]. Emotionen wie Freude, Angst oder Ärger werden durch die durchschnittliche Dauer der Einstellungen, harte oder weichere Kamerabewegungen, etc. vermittelt. Das Kameramodell bleibt auf die traditionellen Koordinatenangaben beschränkt, und die Kamerasteuerung obliegt einer speziell dazu geschaffenen unsichtbaren Kamerafigur. Diese empfängt von anderen Spielfiguren Signale, die deren aktuellen Zustand charakterisieren und leitet daraus eine Vorstellung über die aktuelle Situation ab, die sich dann in der Kameraführung widerspiegelt.

1.1.3 Nicht-Photorealistische Anwendungen

Als Negation des Begriffes Photorealismus läßt sich der Nicht-Photorealismus (NPR, siehe auch [STROTHOTTE und SCHLECHTWEIG 2002]) nicht ohne Rückgriff auf das erklären, wovon er sich abgrenzt. NPR ist ein Sammelbegriff für verschiedene Verfahren der Computergrafik, die vielleicht auf den ersten Blick wenig Gemeinsamkeiten aufweisen. Ihr besonderes Kennzeichen ist die Abkehr vom in der traditionellen Computergrafik verfolgten Ziel der möglichst *photo-realistischen*, d. h. vorbildnahen Simulation der natürlichen Bild-erzeugung. Obwohl – oder auch gerade weil – dieses Ziel sicher nicht in vollem Umfang erreichbar ist, bleibt es doch Gradmesser für den bereits erreichten Fortschritt des Photorealismus. Photorealismus ist also als das hin-

ter den Grafikalgorithmen stehende *Ziel* aufzufassen; photorealistische *Bilder* aus dem Computer gibt es dagegen nicht.²

Bei der vom Photorealismus angestrebten vorbildgetreuen Simulation natürlicher Prozesse müssen sämtliche das Bild „verfälschende“ Effekte vermieden werden. Fragen zum Sinn und Zweck des Bildes können keinen Einfluß auf die Bildentstehung haben. Sie können in der photorealistischen Computergrafik immer nur soweit zum Zuge kommen, wie sie sich bereits im geometrischen Modell, der Beleuchtung, der Kameraposition, etc. niederschlagen.

In der nicht-photorealistischen Computergrafik spielt dagegen, vereinfacht ausgedrückt, das geometrische Modell nur insofern eine Rolle, wie es zur Erlangung des kommunikativen Zieles oder eines gewünschten Ausdrucksstils notwendig ist; Abweichungen vom physikalisch korrekten Realismus etwa durch gesonderte Betonung bestimmter Bildbereiche, Weglassen von Detail an unwichtigen Stellen, Verzerrungen, klare Abgrenzungen der Objekte durch Liniengrafik bzw. Strichzeichnungen sind ausdrücklich erlaubt und gewollt. Dabei kann die gleiche Szenenbeschreibung benutzt werden wie im photorealistischen Fall;³ der Unterschied liegt nicht im *was*, sondern im *wie* der Darstellung.

Insofern die nicht-photorealistische Computergrafik auf der photorealistischen aufbaut, benutzt auch sie (bewußt oder unbewußt) eine Kameraposition, um zu beschreiben, wo der Betrachter steht und wohin er sieht.

1.1.4 Kameraplaner

Bei der rechnergestützten Simulation realer Kamerafahrten bzw. der Einsatzvorbereitung für robotergesteuerte Kameras [DRUCKER 1994] werden zusätzlich

² Derzeit ist auch die photorealistische Computergrafik nicht in der Lage, dieselben Bilder zu generieren wie eine reale Kamera, sondern macht zahlreiche Kompromisse bezüglich der benötigten Rechenzeit, des Speicherbedarfs und der Detailliertheit der Modelle. Nicht zuletzt können auch die verwendeten Darstellungsalgorithmen selbst immer nur auf bereits bekannten Naturgesetzen beruhen und sind daher bestenfalls eine Annäherung an die Realität.

³ Für die jeweiligen Spezialeffekte eines NPR-Verfahrens sind zusätzliche Informationen nötig, wie z. B. Beschaffenheit des Papiers, Stiftstärke, Linienstil, Art des Pinsels, etc. Andere Informationen sind dagegen für viele NPR-Anwendungen entbehrlich, so etwa der Farbton des Materials im Falle einfarbiger Strichzeichnungen.

zu den optischen Eigenschaften der Kamera auch deren physische Eigenschaften wie Größe und Masse benutzt, damit beim Filmen benutztes zusätzliches Gerät (Krane, Schienen, etc.) darauf abgestimmt werden können. Außerdem muß der geplante Kamerapfad kollisionsfrei sein; Vereinfachungen, die bei einer rein virtuellen Kamera noch möglich waren, wie die Vernachlässigung der Kameragröße, sind hier nicht mehr zulässig. Die zusätzlichen Daten gehen in eine Reihe von Nebenbedingungen (*constraints*) ein, die z. B. DRUCKER in seinem Optimierungsverfahren zur Kamerapositionierung benutzt.

Auf die Bilddarstellung haben diese Kameraeigenschaften insofern Einfluß, als sie einerseits die Orte einschränken, an denen sich die Kamera befinden kann und andererseits Bewegungsbahnen der Kamera durch die oben erwähnten Hilfsmittel vorgegeben sind. Im Film als besonders ungewöhnlich empfundene Perspektiven kommen häufig auch durch den Einsatz neuartiger Kamertechnik zustande, etwa im Zuge der Miniaturisierung. Andererseits kann ein Verzicht auf aufwendige Technik z. B. bei Flucht- oder Verfolgungssequenzen für eine besondere Spannung und Echtheit sorgen, wenn die Kamera im Lauf nicht durch teure Mechanik abgedämpft wird (Steadycam), sondern sich die Erschütterungen im Bild zeigen.

1.1.5 Simulation optischer Eigenschaften

Kameras stellen trotz aller moderner Feinessen technisch unvollkommene Systeme dar, die die Welt zwar sehr gut, aber nicht vollständig exakt abbilden. Viele der dabei entstehenden Artefakte sind auf die optischen Eigenschaften der verwendeten Linsensysteme zurückzuführen.

Mit der möglichst exakten Simulation der optischen Eigenschaften von Kameraobjektiven wird unter anderem das Ziel verfolgt, die Abbildungseigenschaften – mit allen Unzulänglichkeiten – möglichst früh zu erkennen. Das wiederum kann dazu genutzt werden, Bildfehler vor einer etwaigen Bildverarbeitung (maschinelles Sehen) herauszurechnen.

Das in den meisten Systemen verwendete Prinzip der *camera obscura*⁴ (siehe Abbildung 1.1) ist stark idealisiert und stellt gemessen an heute in Kameras

⁴ Urform der Kamera, die ohne eine gläserne Linse auskommt. Die Kamera war tatsächlich eine Kammer bzw. Kiste, an deren Rückseite sich eine lichtempfindliche Fotoplatte befand.

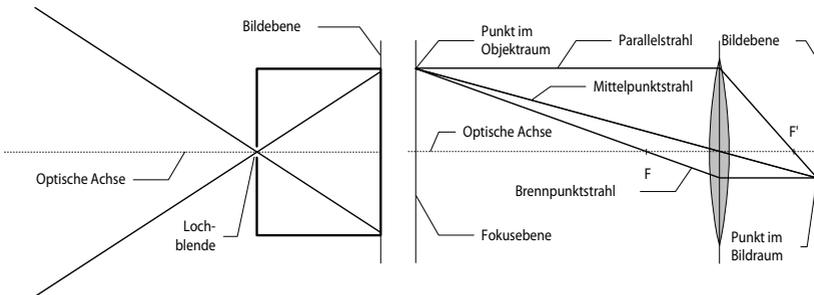


Abbildung 1.1: Vereinfachte Linsenmodelle und ihre Strahlverläufe.

Links: Camera Obscura. Licht kann nur durch die (als unendlich klein angenommene) Öffnung an der Frontseite der Kamera eindringen. Dadurch gibt es für jeden Punkt auf der Bildebene genau einen Strahl in den Objektraum; das Bild ist immer scharf.

Rechts: Linsenmodell mit der vereinfachenden Annahme einer unendlich dünnen Linse mit nur einer Brechungsebene. Da das Licht die Linse nicht mehr nur an einer Stelle durchquert, gibt es für jede Entfernung der Bildebene von der Linse nur noch eine Ebene im Objektraum, die scharf dargestellt wird. Mit der Möglichkeit, die Größe der Blendenöffnung zu variieren, wird allerdings die Schärfentiefe erhöht.

benutzten optischen Systemen eine grobe Vereinfachung dar, die allerdings verschiedene Vorteile mit sich bringt. In der Praxis kann durch eine unendlich kleine Öffnung auch nur unendlich wenig Licht gelangen. In der Computergrafik wird aber kein lichtempfindliches Material benötigt und so ist es egal, welche Belichtungszeit gebraucht wird, um ein Bild zu erzeugen.⁵ Die Schärfentiefe ist unbegrenzt, da das Licht nur an einer einzigen Stelle durch das Objektiv tritt. Da keine Linsen vorhanden sind, gibt es keine Brechungs- und Beugungseffekte mit den daraus resultierenden Farbverfälschungen.⁶ Außerdem wird die gesamte Bildfläche gleichmäßig belichtet, wogegen in der Realität Vignettierungen (Abdunklungen des Bildes zum Rand hin) auftreten.

⁵ In den Basisalgorithmen der Computergrafik, die sich auf Einzelbilder beziehen, taucht generell keine zeitliche Komponente auf. Erst wenn Effekte wie Bewegungsunschärfe simuliert werden sollen, spielt Zeit wieder eine Rolle.

⁶ Hierzu zählen unter anderem chromatische Aberrationen des Objektivs, die daraus resultieren, daß die optischen Eigenschaften einer Linse von der Lichtwellenlänge (Farbe) abhängen, sphärische Aberration (Strahlen durch die Objektivmitte vereinigen sich nicht am gleichen Punkt wie Strahlen am Rand des Objektivs), Ungenauigkeiten der Linsenoberfläche, etc.

ten. Diese für Konstrukteure von Objektiven geradezu traumhaften Zustände sind in der Computergrafik selbstverständlich und vereinfachen den Bildberechnungsprozeß dramatisch.

Wesentlich komplizierter wird es erst, wenn auch die eigentlich ungewollten Artefakte eines optischen Linsensystems mit in die Bildberechnung eingehen sollen. Bereits 1981 stellen POTMESIL et. al. ein auf der *thin lens approximation* (siehe Abbildung 1.1) basierendes Modell vor, welches die optische Blende simuliert und damit Effekte wie Schärfentiefe und Aufhellungen bzw. Abdunklungen durch die unterschiedlich stark einfallende Lichtmenge ermöglicht [POTMESIL und CHAKRAVARTY 1982]. Spätere Arbeiten setzen noch wesentlich aufwendigere Modelle ein, die jede einzelne Linse eines Linsensystems geometrisch abbilden, um daraus ein angenähertes Linsenmodell zu erstellen. KOLB et. al. konnten in einem Raytracer zeigen, daß die Berechnungen im Linsensystem die Rechenzeit nicht wesentlich erhöhen (wohl aber die Notwendigkeit, zur Nachbildung der Unschärfe für jeden Bildpunkt ein Vielfaches an Strahlen zu berechnen) [KOLB et al. 1995]. HEIDRICH et. al. benutzen ein ähnliches Modell für ein bildbasiertes⁷ Verfahren, welches – mit entsprechend leistungsfähiger Grafikhardware – echtzeitfähig ist [HEIDRICH et al. 1997].

1.1.6 Zusammenfassung

Die Berücksichtigung verschiedener Kameraeigenschaften in grafischen Systemen ist natürlich kein Selbstzweck. Vielmehr bilden sie eine Grundlage für verschiedenste Effekte, die – einzeln oder kombiniert – die Vermittlung anwendungsspezifischer Inhalte unterstützen können.

So kann zum Beispiel gezielt eingesetzte Unschärfe einen Hinweis auf den derzeitigen Fokus geben (siehe Abschnitt 4.3). Transparenz („Röntgenauge“) kann Störendes durchsichtig machen und so den Blick auf wichtigere Objekte freigeben. Nicht-photorealistische Zeichenstile können die grafische Präsentation abwechslungsreicher und interessanter machen.

⁷ Bildbasierte Verfahren operieren auf den erzeugten 2D-Bildern, nicht auf den 3D-Modellen. Sie greifen daher nicht in den eigentlichen Zeichenprozeß ein, sondern führen eine Nachbearbeitung an seinem Ergebnis durch. Dadurch lassen sie sich leichter mit vorhandenen Grafikengines kombinieren.

Für die Wirkung dieser Effekte ist es aus Anwendungssicht belanglos, ob sie ihr Vorbild im Verhalten einer Film-, Foto- oder Videokamera und den damit verbundenen physikalisch-chemischen Bilderzeugungsprozessen haben oder ob andere Darstellungsverfahren, etwa aus der Bildenden Kunst Pate standen. Die gemeinsame Eigenschaft der oben beschriebenen Effekte ist vielmehr, daß sie auf gezielten Variationen bestimmter Parameter des jeweiligen Darstellungsprozesses oder – im Falle der Computergrafik, des Zeichenalgorithmus – basieren.

1.2 Ziele der Arbeit

Die heute noch innerhalb der Anwendungen verteilten Funktionalitäten, Eigenschaften, Darstellungsmodi und Bewegungsabläufe der Kamera dienen meist⁸ den gleichen Zielen: der Lenkung des Betrachterfokus und der Vermittlung von Informationen, die über die rein geometrisch korrekte Darstellung der Szene hinausgehen. Mit der Umsetzung solcher Darstellungsziele wird die Kamera zu einer Schnittstelle zwischen den Möglichkeiten der benutzten grafischen Zeichenalgorithmen und den situationsabhängigen Kommunikationsabsichten der Anwendung.

Das bisher dargestellte Kameramodell ist damit allerdings überfordert. Es stellt sich daher die Frage, ob sich zusätzliche Aspekte der Bilddarstellung unter einem erweiterten (und erweiterbaren) Kameramodell zusammenfassen lassen, mit dessen Hilfe eine *intelligente Kamera* die visuellen Ausdrucksmittel koordinieren kann. Die virtuelle Kamera würde zu einer Schaltstelle für die intelligente, situationsabhängige Präsentation des zugrundeliegenden Daten- bzw. Szenenmodells. Hiermit würden bisher separate behandelte Kameratechniken für neue Effekte kombinierbar.

In dieser Arbeit wird daher die Hypothese aufgestellt, daß für bestimmte Klassen von Anwendungen (besonders die Intelligenten Interaktiven Systeme) eine derart erweiterte Sicht auf die Kamera neue Möglichkeiten der Inhaltsvermittlung eröffnet.

Dazu ist es zunächst notwendig, sich ein wenig vom Gedanken an eine Kamera als realem Vorbild zu lösen und sie abstrakt als ein Gerät zu erkennen, das

⁸ mit Ausnahme der Kamera-Simulation vor Filmaufnahmen

uns die Welt als „technisches Auge“ sehen läßt. Da dieses „Sehen“ in der Computergrafik etwas anders abläuft als in der Realität, gibt es bei dieser erweiterten Kamera Gemeinsamkeiten und Unterschiede zum realen Vorbild. Wichtig ist jedoch, daß die Eigenschaften der Kamera bestimmen, welche Teile der Realität (der Szene) abgebildet werden und wie sie im Bild erscheinen. Für die Art der Darstellung bietet die verwendete Grafikkbibliothek (der Renderer, bzw. die Engine) zusammen mit den vorhandenen Grafikmodellen (Szenen, etc) eine Reihe von Möglichkeiten und Freiheitsgraden, die analog zu denen realer Kameras ihren Niederschlag im Kameramodell finden müssen. Dabei gibt es statische Aspekte, die sich bereits auf Einzelbilder auswirken und dynamische, die erst in einer Bildfolge zum Tragen kommen.

Ziel der Arbeit ist es somit, die verschiedenen Aspekte eines Kameramodells zusammenzufassen und um spezifisch computergrafische Eigenschaften zu erweitern. In einem zweiten Schritt wird aufgezeigt, wie die Freiheitsgrade dieses erweiterten Kameramodells für verschiedenste Hervorhebungstechniken und die Erzeugung ausdrucksstarker Computergrafiken (sowohl statisch als auch dynamisch) genutzt werden können.

1.3 Ergebnisse der Arbeit

Die vorliegende Arbeit stellt ein neuartiges, erweitertes Kameramodell vor, und demonstriert seine Einbettung in Intelligente Interaktive Systeme. Dabei wurden folgende Ergebnisse erzielt:

- Untersuchungen zu grafischen Hervorhebungsmitteln in interaktiven Systemen

Dazu zählen sowohl illustrative Mittel, die in Einzelbildern wirken, als auch dynamische Mittel, die sich erst während einer kurzen Animation entfalten. Aus den gewünschten Mitteln lassen sich Anforderungen an die Grafikkomponente des Systems und die benötigten 3D-Modelle ableiten.

- Vergleichende Betrachtungen zu technischen und unterhaltenden interaktiven Systemen sowie Filmen

Ziel dieser Untersuchungen war es, grafische Hervorhebungsmittel herauszuarbeiten, die der Spezifik des jeweiligen Mediums gerecht werden. Hierbei werden insbesondere Unterschiede der beiden großen Gruppen interaktiver 3D-Anwendungen (der technischen und der unterhaltenen Systeme) deutlich.

- Entwurf eines erweiterten Kameramodells zur Erfassung verschiedener darstellungs- und interaktionsrelevanter Eigenschaften intelligenter Kameras

Dieses Modell ermöglicht die Weiterentwicklung der in der Computergrafik bekannten virtuellen Kameras zu *intelligenten* Kameras, deren Freiheitsgrade weit über geometrische Angaben hinausgehen und den gesamten Prozeß der computergrafischen Bilddarstellung umfassen. Im Zusammenhang mit den gewählten Anwendungen aus dem Bereich der interaktiven Systeme werden dabei insbesondere die oben erwähnten Hervorhebungseffekte berücksichtigt, die so zu einem Ausdrucksmittel der intelligenten Kamera werden.

- Entwurf und Implementation eines Verfahrens zur dynamischen Planung von Kamerafahrten

Im Vordergrund stand hierbei die schnelle Erzeugung einer sanften und kollisionsfreien Bewegung durch die Szene. Dies ermöglicht interaktiven Anwendungen, den Benutzer zu bestimmten Szenenobjekten zu führen. Positionswahl und Parameter der Wegplanung werden so zu weiteren Eigenschaften der intelligenten Kamera.

- Entwicklung eines Systems zur interaktiven technischen Dokumentation (VISDOC) sowie zur medizinischen Illustration (AGI³LE)

Diese Systeme demonstrieren prototypisch die vorgestellten neuen Konzepte. Sie implementieren verschiedene Hervorhebungstechniken, etwa Variationen der Schattierung, der Linienstile sowie Schatten und Transparenz. Hinzu kommen dynamische Effekte wie die erwähnten Kamerafahrten.

1.4 Gliederung der Arbeit

In **Kapitel 1** wurden bereits Anwendungsfelder von Kameramodellen diskutiert. Diese umfassen jeweils bestimmte Einzelaspekte und finden ihre Anwendung in Systemen, die ihren Schwerpunkt in der Demonstration bzw. Ausnutzung eben dieser Aspekte (z. B. Bildschärfe, Steuerung der Kamera) haben. Außerdem wurde gezeigt, daß die Kamerametapher im Sinne einer Kamera als „betrachtendes Medium“ erweitert und damit von der engen Analogie zum Vorbild der Film- bzw. Fotokamera losgelöst werden kann.

Basierend auf den zuvor aufgezählten statischen und dynamischen Eigenschaften einer virtuellen Kamera wird in **Kapitel 2** das *Erweiterte Kameramodell* vorgestellt und in den Prozeß der computergestützten Bilddarstellung eingeordnet. Dieses Modell beschreibt nicht nur Eigenschaften, die die Darstellung von Einzelbildern betreffen, sondern auch dynamische Aspekte des Zeitverhaltens und die für interaktive Anwendungen wichtige Steuerung der Kamera. Weiterhin werden exemplarisch Anwendungsfelder für die verschiedenen Teilaspekte (Achsen) des Modells angeführt.

Die so erfaßten Kameraparameter legen die Grundlage für diverse Effekte und Ausdrucksmittel in interaktiven Anwendungen. Die dort verwendeten dynamischen Hervorhebungstechniken bilden eine der Hauptmotivationen für diese Arbeit. Ein erweitertes Kameramodell dient dabei als Werkzeugkasten, dessen Instrumente für die Unterstützung kommunikativer Ziele genutzt werden.

Kapitel 3 stellt Hervorhebungsmittel für den Einsatz in interaktiven Systemen vor. Der Begriff der Hervorhebung wird definiert und vom verwandten Begriff der Akzentuierung abgegrenzt. In interaktiven 3D-Systemen spielt außerdem die – nicht immer gegebene – Sichtbarkeit der hervorgehobenen Objekte eine große Rolle. Es werden daher Wege diskutiert, wie Sichtbarkeit garantiert und/oder verbessert werden kann. Weitere Themen dieses Kapitels sind Zeichenattribute zur Hervorhebung und Abstraktionen als Mittel der Deakzentuierung.

All diese Techniken basieren jeweils auf den in einem anwendungsspezifischen Kameramodell erfaßten Freiheitsgraden der Grafikengine, die für als „auffällig“ empfundene grafische Effekte genutzt werden können. Dem erweiterten

Kameramodell kommt daher die Aufgabe zu, diese Effekte zu koordinieren und situationsgerecht einzusetzen.

Wirken die grafischen Mittel von Kapitel 3 noch auf Einzelbilder, so widmet sich **Kapitel 4** Techniken, die sich die Dynamik einer bewegten Kamera zunutze machen. Filmische Gestaltungsmittel bieten einen großen Fundus an dynamischen Hervorhebungsmitteln, die in interaktiven Systemen derzeit noch wenig genutzt werden.

Kapitel 4 untersucht daher Filmtechniken auf ihre Eignung in interaktiven Systemen, insbesondere Techniken zur Führung des Betrachterfokus. Dabei kristallisieren sich automatisch generierte Kamerafahrten als nützliches Mittel der gesteuerten Erkundung einer 3D-Welt heraus.

In **Kapitel 5** werden verschiedene Ansätze für die dynamische Planung von Kamerapositionen und -bewegungen diskutiert. Dabei werden die Kameraperspektiven nicht mehr auf der Ebene der 3D-Koordinaten beschrieben, sondern *inhaltlich*, als System von Anforderungen an das zu zeigende Bild. Die möglichen Typen von Anforderungen stellen eine weitere Eigenschaft des erweiterten Kameramodells dar.

Für die Planung von Kamerabewegungen werden neben dem Verfahren der reaktiven Kameraplanung verschiedene Ansätze aus der Robotik vorgestellt und zu den Anforderungen an eine Kameraplanung in interaktiven Systemen in Beziehung gesetzt. Diese Betrachtungen münden in die Vorstellung eines neuen Verfahrens für die Planung von Kamerafahrten.

Wie das erweiterte Kameramodell in interaktiven Anwendungen zum Tragen kommt, wird in Fallstudien im **Kapitel 6** demonstriert. Dabei handelt es sich zunächst um „aus Kamerasicht“ relativ einfache, technisch orientierte Systeme wie VISDOK und AG³LE.

Im **Ausblick** wird, basierend auf einer reaktiven Kameraplanung, eine Systemarchitektur für unterhaltende Grafiksysteme vorgestellt, die sich an handelnden Spielfiguren orientiert und bedeutend mehr Freiheiten der Kameraführung ermöglicht, aber auch zahlreiche neue Fragenstellungen beinhaltet. Schließlich gibt **Kapitel 8** eine Zusammenfassung der Arbeit.

2 Ein Konzept für ein erweitertes Kameramodell

Nachdem bereits im vorangehenden Kapitel der Begriff des Kameramodells benutzt wurde, wird hier zunächst eine genauere Bestimmung des Modellbegriffs nachgeholt. Diese geht insbesondere auf die Bedeutung von Modellen in der Computergrafik ein. Hierbei wird die Funktion und der Stellenwert des Kameramodells unterstrichen.

Für die Motivation dieser Arbeit spielten besonders die Anwendungsfelder der interaktiven technischen Dokumentation und der Computerspiele eine große Rolle. Bei näherer Betrachtung fällt auf, daß diese beiden Genres der interaktiven Computergrafik eine Reihe von Gemeinsamkeiten aufweisen. Dieser Abschnitt schlägt deshalb eine einheitliche Betrachtung der beiden Kategorien vor.

Im Weiteren wird untersucht, wie die Kamera in solchen Systemen ihren Platz als Ausdrucksmittel finden kann. Das danach vorgestellte erweiterte Kameramodell bildet die Grundlage einer solchen Einbindung. Die verschiedenen Ausdrucksebenen (Einzelbild, Animation, etc) der Kameraführung finden ihren Niederschlag in den verschiedenen Dimensionen des erweiterten Kameramodells. Die weiteren Abschnitte erläutern die Bedeutung und möglichen Einsatzfelder der verschiedenen Dimensionen des erweiterten Kameramodells.

2.1 Modelle in der Computergrafik

In allen Computerprogrammen müssen veränderliche Faktoren, die das Verhalten der Software beeinflussen können oder sollen, in einem Datenmodell gespeichert und während des Programmablaufs ausgewertet werden. Modelle treten in der Computergrafik in verschiedener Form auf, die sich – wie auch in anderen Sparten der Informatik – grob in Datenmodelle und Ablaufmodelle unterteilen lassen.

Datenmodelle beschreiben die von der Anwendung benötigten Daten und ihre Struktur; speziell in der Computergrafik also geometrische und andere physikalische Eigenschaften.

Geometrische Modelle dienen der Ablage, dem Austausch und der Verwaltung von zu zeichnenden Szenen bzw. Szenenobjekten. Diese liegen meist als Dateien vor und enthalten die geometrischen Koordinaten, die bei geeigneter Interpretation die gewünschte Darstellung ergeben. Hinzu kommen meist noch Informationen über das Erscheinungsbild, also Farben oder andere Aspekte der Oberflächenbeschaffenheit. Bei Animationen wird zusätzlich noch ein *Animationsmodell* benötigt, was aber ebenso der Datenablage dient und die zeitliche Veränderung der Szene beschreibt.¹

Ablaufmodelle stellen Vereinfachungen physikalischer Prozesse (z. B. der Lichtreflexion) dar und beschreiben die grundlegenden Annahmen und Kompromisse, die dem Darstellungsprozeß zugrundeliegen.

Der Darstellungsprozeß selbst folgt mehr oder weniger exakten *mathematisch-physikalischen Modellen*, aus denen sich bestimmte Algorithmen bzw. Berechnungen ableiten lassen. Hierzu zählen etwa die Beleuchtungsmodelle² von PHONG oder COOK und TORRANCE, die Farbmodelle sowie das weiter unten beschriebene einfache Kameramodell der Computergrafik.

Darüber hinaus definieren die Modelle bzw. ihre konkrete Implementation auch den Satz veränderlicher Parameter, die als Einflußgrößen in die Berech-

¹ Das Animationsmodell beschreibt nicht den Verlauf der Datenverarbeitung und ist aus diesem Grund kein Ablaufmodell im hier beschriebenen Sinn.

² Siehe auch [COOK und TORRANCE 1981, PHONG 1975].

nung eingehen können. Diese können frei wählbar sein oder sich während der Berechnung ergeben. Im Falle des PHONG'schen Beleuchtungsmodells sind das z. B. die diffuse Farbkomponente einer Lichtquelle bzw. die Oberflächennormale eines beleuchteten Flächenpunktes. Selbst für den einfachen Vorgang des Zeichnens einer Linie gibt der benutzte Algorithmus vor, mit welchen Daten (z. B. Anfangs- und Endpunkt oder Anfangspunkt, Richtung und Länge) die Linie beschrieben werden muß. Gleiches gilt für Dreiecke oder komplexere, nicht-polygonale Objekte wie Freiformflächen, etc. Diese frei wählbaren Daten sind dann Bestandteile der Szene, wenn sie wie im Beispiel die zur Szene gehörenden Lichtquellen oder eine einzelne Linie beschreiben. Sie können aber auch als Konstanten in die Berechnung eingehen, wenn ihre Veränderbarkeit im Anwendungskontext als nicht sinnvoll erscheint.

Da die Modelle nur eine Vereinfachung bilden, wird es immer Effekte geben, die sich mit ihnen nicht darstellen lassen und für die auch keine Daten vorhanden sind, die sie beschreiben könnten. Wie eingangs erwähnt, hat bereits eine reale Kamera eine Vielzahl von Eigenschaften, die nicht mit vertretbarem Aufwand implementiert werden können. Diejenigen Kameraeigenschaften aber, die entweder im Kamera-Datenmodell der Software nicht gespeichert werden oder die im Verlauf der Bildgenerierung nicht abgefragt werden, gehen zwar eventuell als invariante Parameter in das Bild ein, können aber nicht dynamisch variiert werden. Wenn dagegen eine dynamische Veränderung gewünscht wird, muß diese auch mit dem verwendeten Modell darstellbar sein. Welche der vielen möglichen Eigenschaften nun tatsächlich modelliert und verändert werden sollen, hängt also von den Bedürfnissen der Anwendung ab.

Die in der Einführung genannten Anwendungsfelder stellen unterschiedliche Anforderungen an die virtuelle Kamera, ihre zu variierenden Eigenschaften und deren dynamische Steuerung. Tabelle 2.1 faßt diese Aspekte nochmals zusammen.

In keinem der Anwendungsgebiete wird derzeit für die Kameraeigenschaften ein über das einfache Modell der 3D-Grafik hinausgehendes Kameramodell verwendet. Jedoch bietet sich diese Zusammenfassung gerade im Hinblick auf eine gemeinsame und abgestimmte Steuerung an, um die verschiedenen visuellen Effekte, die die Kamera erlaubt, besser zu beherrschen und damit

Tabelle 2.1: Grafische Anwendungen und ihre Kameraeigenschaften

3D-Grafik allgemein	Position, Richtung, Öffnungswinkel, ...
Intelligente Interaktive Systeme	verschiedene Darstellungsmodi zur Hervorhebung, Interaktionseigenschaften zur Navigation
Nicht-Photorealistische Darstellungen	Schattierungsmodi, Beleuchtungsmodelle, Zeichenstile
Kameraplanung	physikalische Eigenschaften der zu simulierenden Kamera
Simulation optischer Eigenschaften	Beschreibung der einzelnen Linsen (Brechkraft, Wölbung, Dicke, Abstände, Blenden)

die Kamera besser als Ausdrucksmittel in interaktiven Systemen zu erschließen.

2.2 Intelligente Interaktive Systeme

Interaktive technische Dokumentationen und Unterhaltungssoftware weisen trotz unterschiedlicher Ausrichtung und Zielgruppen eine Reihe von Gemeinsamkeiten auf:

1. Sie präsentieren dem Benutzer eine virtuelle Welt, in der mit mehr oder weniger starken Einschränkungen interaktiv navigiert werden kann.
2. Es sind Interaktionen mit einzelnen Objekten der Welt möglich, die zu Veränderungen innerhalb der virtuellen Welt führen können.
3. Der Benutzer hat ein Ziel bzw. eine Aufgabe innerhalb der Welt zu lösen (und sei es nur, sie kennenzulernen).
4. Zusätzlich zur grafischen Darstellung werden erläuternde Texte präsentiert.
5. Textuelle und grafische Information müssen kohärent sein.
6. Der präsentierte Zustand der Welt kommuniziert eine bestimmte Situation, die sich u. a. aus vorangegangenen Interaktionen ergibt.

7. Die Wahrnehmung dieser Situation kann durch verschiedenste grafische Ausdrucksmittel und textuelle Hervorhebungen unterstützt werden.
8. Diese Mittel sind nicht in der Beschreibung der virtuellen Welt enthalten, sondern werden von der Anwendung bereitgestellt.

In der Tat ist es zuweilen schwierig, eine klare Grenze zwischen instruierenden und unterhaltenden Anwendungen zu ziehen. Außerdem ist fraglich, ob diese Unterscheidung im Hinblick auf die in dieser Arbeit betrachteten Methoden der dynamischen Kamerasteuerung überhaupt bedeutsam ist (mehr dazu im Abschnitt 4.2). Der Begriff *Intelligente Interaktive Systeme* faßt daher diese beiden Anwendungsklassen überall dort zusammen, wo die Gemeinsamkeiten überwiegen.

Bei näherer Betrachtung der obigen Punkte fällt auf, daß die oberen vier Eigenschaften von vielen einfachen Systemen bis hin zu VRML-Betrachtern verwirklicht werden und – abgesehen von der Interaktivität – keine besonderen Anforderungen an die Kamera oder ihre Steuerung stellen. Dies gilt nicht mehr für die unteren vier Punkte. Das Kommunizieren einer bestimmten Situation durch angepaßte Darstellungsmittel verlangt nach einer *intelligenten Kamera*, die die vorhandenen zeichnerischen und dynamischen Möglichkeiten geeignet kombiniert und nutzt. Damit wird das Vorhandensein einer solchen intelligenten Kamera zu einem wichtigen Merkmal Intelligenter Interaktiver Systeme.

2.3 Die intelligente Kamera als Ausdrucksmittel

In den in der Einleitung erwähnten Anwendungsgebieten – speziell in den Intelligenen Interaktiven Systemen – werden zahlreiche Darstellungstechniken benutzt, die kommunikativen Zielen Ausdruck verleihen oder einfach nur Ausdruck eines bestimmten visuellen Stils sind. Unter *Kameratechniken* sollen speziell diejenigen Darstellungsmittel verstanden werden, die durch die zielgerichtete Veränderung von Darstellungsparametern entstehen, die einer virtuellen Kamera zugeordnet werden können.

Diese Darstellungsparameter aber lassen sich nicht mehr nur durch wenige Koordinaten wie Kameraposition, etc. beschreiben, sondern umfassen prinzipiell alle Freiheitsgrade,³ die die benutzte *Grafikengine* (als Gesamtheit der zur Bilddarstellung benutzten Grafikmodule) zuläßt und die nicht schon in der Szenenbeschreibung benutzt werden. Wegen der Vielfalt der Grafikengines, ihrer Möglichkeiten und Einschränkungen, ist es an dieser Stelle nicht möglich, alle derartigen Ausdrucksmittel anzugeben; stattdessen konzentriert sich die Arbeit auf Mittel, die im hauptsächlich betrachteten Anwendungsfeld der Intelligenten Interaktiven Systeme eine Rolle spielen (siehe Kapitel 3 und Abbildung 2.1).

Das erweiterte Kameramodell⁴ spannt mehrere Achsen (Dimensionen) auf, die jeweils eine neue darstellungsrelevante Eigenschaft widerspiegeln. Diese Eigenschaften umfassen all diejenigen Variablen der Bilddarstellung, die nicht zur Szenendefinition gehören, sondern unabhängig von der dargestellten Szene veränderbar sind. Damit sind sowohl die Achsen selbst als auch ihre Ausprägungen von der Anwendung abhängig, so daß für jede zu betrachtende Anwendung immer eine Untermenge der vorgestellten Eigenschaften wichtig ist.

Einige der Faktoren (z. B. Darstellungsstile) werden in existierenden Systemen direkt der Grafikengine zugeordnet. Diese wertet die Daten aber nur aus. Für einen Nutzer, der die von der Grafikengine zur Verfügung gestellten Gestaltungsmittel ausnutzen und kombinieren möchte, ist es aber intuitiver, sie als Präsentationsvariablen unter einem erweiterten Kameramodell (siehe Abbildung 2.2 auf Seite 24) zusammenzufassen.

Neben bildrelevanten Aspekten besitzen Kameras auch noch interaktionsrelevante Eigenschaften (in Abbildung 2.1 unten dargestellt). So beeinflussen Anordnung und Bedeutung der verschiedenen Bedienelemente immer auch die Art des Umgangs mit der Kamera. Aus diesem Grunde kommt der ergonomischen Gestaltung einer Kamera – gerade im professionellen Bereich – große Bedeutung zu. Überträgt man diese Beobachtung auf virtuelle Kame-

³ Diejenigen Freiheitsgrade, die von der Anwendung explizit als Ausdrucksmittel genutzt werden, werden auch als *Präsentationsvariablen* bezeichnet.

⁴ Die Bezeichnung *Erweitertes Kameramodell* betont, daß reale Kameras für viele der vorgestellten Eigenschaften als Vorbild dienen. Erweiterte Kameramodelle enthalten aber darüber hinaus auch Eigenschaften, denen kein reales Pendant gegenübersteht.

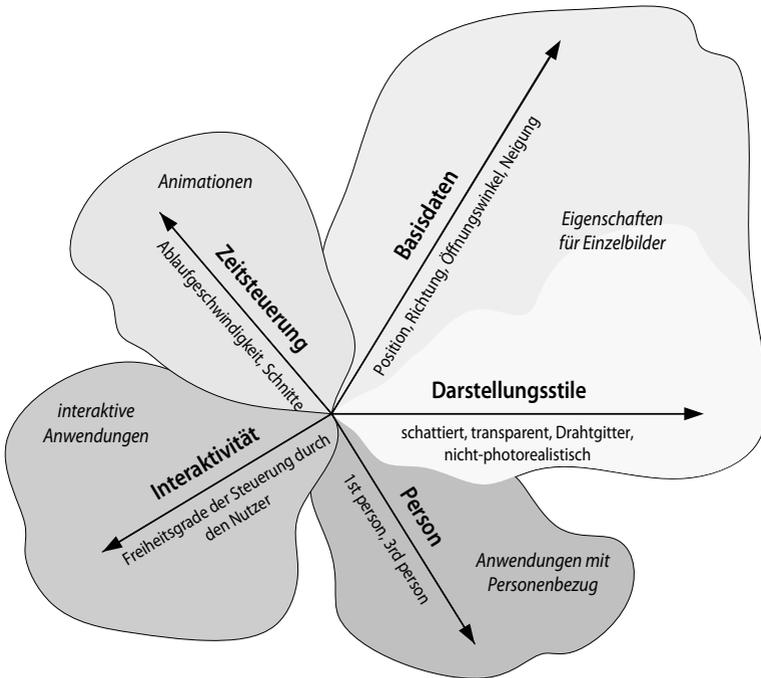


Abbildung 2.1: Erweitertes Kameramodell: Die verschiedenen Achsen repräsentieren verschiedene Aspekte der Steuerung einer virtuellen Kamera.

ras, wird die Art der Kameramanipulation – speziell der interaktiven Navigation in der 3D-Welt – durch den Nutzer zu einem ebenso wichtigen Aspekt des Kameramodells wie die Darstellungsparameter. Dazu kommt, daß sich solche Steuermodi auch situationsbedingt während des Betriebs ändern können. Aus diesem Grunde erfaßt das erweiterte Kameramodell auch interaktive Aspekte.

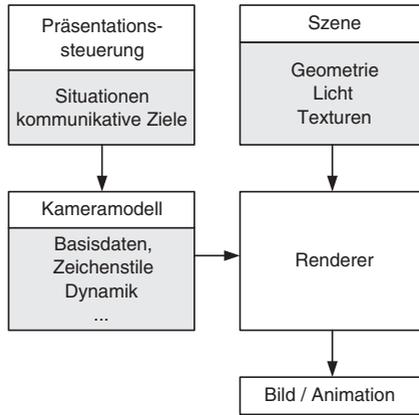


Abbildung 2.2: Abgrenzung von Kameramodell und Renderer. Ersteres enthält eine Reihe von der Szene unabhängige, kamerabezogene Präsentationsvariablen, die von den Algorithmen zur Bildberechnung im Renderer benutzt werden. Die Algorithmen sind statisch, die Szene und das Kameramodell jedoch dynamisch. Die Szene stellt zwar den abzubildenden Inhalt dar, sie bestimmt allerdings nicht das endgültige Erscheinungsbild als Ergebnis des Zeichnens: hierfür werden die Angaben aus dem Kameramodell benötigt, die in der einen oder anderen Form Einfluß auf die Bilddarstellung haben.

2.4 Basisdaten

Grundlage jedes Kameramodells sind einige Daten, die sich an die gängige Vorstellung einer Film- bzw. Fotokamera anlehnen (siehe Abbildung 2.3). Zu den Grundaufgaben des Kameramodells gehört die Beantwortung der in der Einführung erwähnten Fragen „Wo stehe ich, wohin blicke ich?“ also die Angabe von Position und Blickrichtung. Mathematisch gesehen, ergeben diese beiden Angaben die Transformation von Welt- in Kamerakoordinaten mittels einer Drehung und einer Verschiebung. Anschließend wird die 3D-Welt auf ein 2D-Bild projiziert (für eine genauere Beschreibung der mathematischen Hintergründe siehe Anhang A auf Seite 163).

Ob das am Ende entstehende Bild photorealistisch ist, eine einfache Liniengrafik oder ein nicht-photorealistentes Bild, ist für die Transformationen, die die 3D-Modelle nach 2D überführen, unerheblich. Daher bildet das zugrundelie-

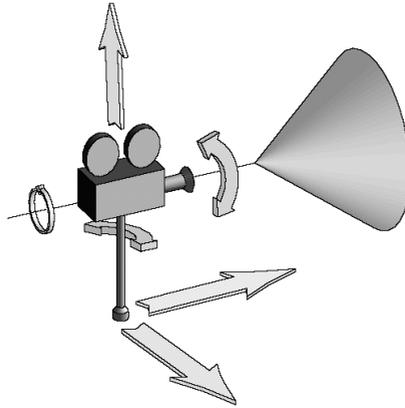


Abbildung 2.3: Sieben Freiheitsgrade der Kamera aus [Drucker 1994]. Die geraden Pfeile symbolisieren die drei Freiheitsgrade der Kameraposition; die gekrümmten Pfeile stehen für die Orientierung. Die Wahl der Oben-Richtung geschieht durch den hinteren kreisförmigen Pfeil und der siebente Freiheitsgrad besteht im Öffnungswinkel des Sichtkegels.

gende Kameramodell, welches nur die in Abbildung 2.3 dargestellten Basisdaten enthält, den kleinsten gemeinsamen Nenner der Kameramodelle in der 3D-Computergrafik.

Wie die Weltkoordinaten sind auch die Kamerakoordinaten in ihrem Wertebereich zunächst nicht begrenzt. Da die zur Verfügung stehende Darstellungsfläche nur eine bestimmte Größe hat, muß festgelegt werden, welcher Ausschnitt der Projektionsebene im Bild dargestellt werden soll. Diese Überführung von 3D-Daten in 2D-Daten ist Bestandteil jedes Darstellungsalgorithmus der 3D-Computergrafik, der zweidimensionale Bilder erzeugt.

Dazu dient eine weitere wesentliche Größe: der Öffnungswinkel des Objektivs. Dieser steht in Beziehung zur bei realen Kameraobjektiven angegebenen Brennweite. Der Öffnungswinkel beschreibt die Größe der Sichtfeldes. Dieses ist in der Breite eventuell etwas anders als in der Höhe, je nach Seitenverhältnis des Bildschirmfensters. Auch hier gehen Grafikbibliotheken verschiedene Wege; so kann das Sichtfeld als rechteckiger Ausschnitt der Projektionsebene angegeben werden oder als Öffnungswinkel in der Vertikalen, gefolgt vom Seitenverhältnis des Rechtecks.

Zumeist weniger offensichtlich ist die Notwendigkeit, auch die Richtung anzugeben, die im fertigen Bild nach *oben* zeigen soll; eine Veränderung dieser Richtung käme einer Drehung der Kamera um die Blickrichtung gleich. Im Normalfall wird diese Richtung mit der natürlichen Vertikalen zusammenfallen, d. h. die Kamera steht aufrecht.

Für die 3D-Darstellung ebenfalls notwendig, aber in ihrer Wirkung auf das Bild weniger auffällig sind Angaben über die Art der Projektion (perspektivisch, orthogonal) und die vordere und hintere Begrenzung des Sichtkörpers (siehe Anhang A).

Die bis hierher beschriebenen Parameter dienen als Grundlage vieler anderer Facetten des Kameramodells. In der Tat beschränken sich viele implementierte Kameramodelle darauf. Um die Kamera jedoch von einer höheren Ebene als der der 3D-Koordinaten zu steuern, sind einige weitere Eigenschaften erforderlich, wie sie im Folgenden beschrieben werden.

2.5 Darstellungsstile

Genauso wie ein Film unterschiedliche Eigenschaften haben kann, die sich auf die Reproduktion der fotografierten Szene auswirken, so kann auch die virtuelle Kamera die Szene in verschiedener Weise abbilden.

Dem Film am nächsten kommt sicherlich die schattierte Darstellung, in der Flächen gefüllt dargestellt werden. In der Computergrafik (besonders in Computerspielen) kommen heute meist Texturen⁵ zur Anwendung; im einfachsten Fall geht man aber von einer Grundfarbe je Fläche aus. Das Beleuchtungsmodell wandelt diese Farbe dann entsprechend des Lichteinfalls ab. Diesen Vorgang bezeichnet man in der Computergrafik als *Schattieren*

⁵ Texturen sind ein Kunstgriff der Computergrafik, bei dem Flächen nicht mit bestimmten Farben, sondern mit Mustern gefüllt werden. Viele komplexe Objekte lassen sich dadurch ausreichend realistisch zeichnen. So können z. B. Bäume mit tausenden Blättern durch wenige ineinander verschränkte Flächen mit speziell vordefinierten, transparenten Bildern von Zweigen simuliert werden. Trotz der zusätzlich zu erstellenden Texturbilder sind texturbasierte Szenen wesentlich leichter zu modellieren, weil die Modelle weniger Flächen besitzen. Zudem werden sie auf moderner Grafikhardware, die die Darstellung von Texturen beschleunigt, wesentlich schneller gezeichnet.

(*shading*). Schattierte Darstellungen wirken im Vergleich zu anderen Methoden wie Drahtgittern oder Comic-Zeichenstilen zwar am realistischsten und lassen sehr gut Formen erkennen,⁶ sind aber trotzdem nicht immer zweckmäßig. So sind z. B. gleichfarbige Kanten nicht unterscheidbar, was Objektkonturen vor einem ähnlich gefärbten Hintergrund schwer erkennbar macht. Gefüllte Flächen verdecken sich gegenseitig, so daß immer erst eine Navigation nötig ist, wenn man hinter eine Fläche sehen will. Wenn der Zweck der Darstellung nicht im bloßen, möglichst realistischen Betrachten der Szenerie steht, sondern darin interagiert und etwas verändert werden soll, sind gefüllte Flächen oft hinderlich.

Besonders in der 3D-Modellierung (CAD) sind daher auch heute noch Drahtgitterdarstellungen gefragt, die oft einen besseren Überblick ermöglichen, lassen sie doch auch das Erkennen von bzw. Interagieren mit verdeckten Flächen zu, die bei einer schattierten Darstellung nicht mehr zugänglich sind.

2.6 Zeit

Die ersten beiden Gruppen von Kameraeigenschaften waren noch auf Einzelbilder bezogen. In interaktiven Anwendungen spielt aber auch die Darstellung bewegter Bilder eine große Rolle, sei es nun, daß sich die Szene selbst verändert oder daß durch Nutzerinteraktionen die Kamera direkt oder indirekt bewegt wird (oder beides). Dies bringt neben statischen Eigenschaften der Kamera auch dynamische Aspekte ins Spiel, die sowohl die Veränderung gewisser Kameraparameter über die Zeit hinweg bzw. den Umgang mit einer über die Zeit veränderlichen Szene beschreiben.

Animationen haben meist zum Ziel, einen bestimmten Inhalt zu vermitteln oder zu untermauern. Dies kann wie z. B. bei Computersimulationen dem Erkenntnisgewinn dienen oder etwa bei Spielen zur Unterhaltung beitragen. In beiden Fällen aber ist es wichtig, die Dauer auf die Aufmerksamkeitsspanne

⁶ Die Orientierung der Flächenelemente zu den Lichtquellen und damit ihre Reflektion variiert mit der Form. So kann der menschliche Wahrnehmungsapparat (und auch der künstliche, siehe [STEWART und LANGER 1997, WINSTON 1975]) von Helligkeitsverläufen auf die Form rückschließen. Ohne diesen Effekt wären Formen nur an Silhouetten erkennbar.

der Betrachter abzustimmen, wenn der damit transportierte Inhalt verstanden bzw. aufgenommen werden soll.

An dieser Stelle ist es sinnvoll, zwischen der dem Benutzer gezeigten *Präsentation* und der dieser zugrundeliegenden *Animation* zu unterscheiden. Erstere ist das, was der Betrachter zu sehen bekommt; letztere bildet gleichermaßen die Handlung, das Rohmaterial, aus dem dann für die Präsentation ausgewählt wird. Hierbei findet eine unumkehrbare Transformation statt: nicht nur kann die Animation wesentlich mehr Rohmaterial enthalten als die präsentierte Endfassung,⁷ auch gehen in den Auswahlprozeß zahlreiche Überlegungen ein, die erst nach der Animation entstanden sind.

Zu kurze Präsentationen führen dazu, daß Details nur unvollständig wahrgenommen werden und die Animation gegebenenfalls noch einmal betrachtet werden muß. Zu lange Präsentationen dagegen provozieren entweder ein mentales „Abschalten“ des Betrachters, bevor die beabsichtigte Aussage erkennbar wird oder sie erzeugen Langeweile, weil die Aussage bereits lange vor Ende der Animation ausreichend klar gemacht ist. Beides wirkt kontraproduktiv.

Um die zeitlichen Gegebenheiten einer Animation an die Bedürfnisse der Zielgruppe anzupassen, gibt es im Wesentlichen drei Möglichkeiten:

- Zeitraffer
- Zeitlupe
- Zeitsprünge

Die ersten beiden setzen einen konstanten Faktor zwischen den Verlauf der Animationszeit und der realen Betrachterzeit, so daß die Animation in einen bestimmten Zeitrahmen paßt. So wird man das Wachstum einer Pflanze durch einen Zeitraffer darstellen, dessen Einzelbilder jeweils an einem anderen Tag entstanden sind. Eine Animation eines FEM-simulierten Crashtests dagegen müßte eventuell um mindestens Faktor zehn verlangsamt sein, damit der genaue Ablauf der Verformungen studiert werden kann. In bestimmten Situationen kann es auch sinnvoll sein, das Vorzeichen des Zeitfaktors zu ändern, also rückwärts zu gehen.

⁷ Als Beispiel außerhalb der Computergrafik mag hier wieder der Kinofilm dienen; ein Verhältnis von 10 : 1 von Rohmaterial zu Filmlänge ist hier nicht unüblich.

Besteht allerdings eine Animation in ihrer Urform aus interessanten Momenten, die durch längere Abschnitte von geringem Interesse getrennt sind, so hilft die alleinige Verwendung von Zeitlupe oder Zeitraffer nicht weiter, zumal es ja auch nicht immer angebracht ist, von der natürlichen Ablaufgeschwindigkeit abzuweichen. Vielmehr werden dem Betrachter wesentliche Ausschnitte gezeigt, und die eventuell dazwischenliegenden unwesentlichen Teile weggelassen. Zeitsprünge sind also noch wesentlich flexibler als Änderungen des Zeitfaktors (die natürlich auch damit kombinierbar sind). Im Film etwa sind Zeitsprünge das nahezu ausschließliche Mittel, eine ablaufende Handlung auf Spielfilmlänge zu bringen, da hier das Ändern des Zeitfaktors zum alleinigen Zweck der Zeitbegrenzung als besonders störend empfunden würde. Zeitsprünge können auch in die Vergangenheit führen, entweder als Wiederholungen oder als Rückblenden.

Ein weiterer Anwendungsfall, in dem das Zeitverhalten eine Rolle spielt, sind Kamerafahrten. Diese können durch eine ansonsten unbewegte Szene führen oder parallel zu einer Animation ablaufen. Auch hier ist zu beachten, daß die Kamerafahrt einen kommunikativen Zweck erfüllen soll, der meist darin besteht, dem Betrachter ein besseres dreidimensionales Gefühl der Szene zu vermitteln, als das mit feststehender Kamera möglich wäre. Bei der Fahrt trifft daher prinzipiell die schon für die eigentliche Animation gemachte Überlegung zu, daß die Dauer in Beziehung zum gezeigten Inhalt stehen muß. Zusätzlich ist hier allerdings noch auf eine möglichst störungsfreie Fahrt zu achten, was nicht nur Konsequenzen für die Geometrie des Bewegungspfades (keine Ecken, nur Kurven) hat, sondern auch für ein sanftes Anfahren und Abbremsen spricht.

2.7 Interaktivität und Steuerung

Unter dieser Kategorie werden alle Möglichkeiten des Nutzers oder der Software zusammengefaßt, die Kamera in ihren Basisdaten oder Darstellungsmodi zu verändern. Dazu zählen besonders die zur Verfügung gestellten Navigationsmöglichkeiten, also die Bewegung der Kamera an einen neuen Ort oder die Ausrichtung auf einen neuen Blickpunkt. Diese Bewegungen können sowohl direktmanipulativ als auch indirekt gesteuert werden.

Direkte und indirekte Steuerung Bei einer direkten Steuerung wird der vom Eingabegerät (z. B. der Maus) gelesene Wert in eine neue 3D-Koordinate umgerechnet und die Kamera dorthin versetzt oder gedreht. Ebenfalls der direkten Steuerung zuzurechnen sind Systeme, bei denen die Kamera mit Eigenschaften wie Masse und Trägheit ausgestattet ist und daher einem Bewegungswunsch einen gewissen Widerstand entgegensetzt sowie Systeme, die vor einer Bewegung der Kamera den Bewegungspfad auf Kollisionen mit Szenenobjekten testen und so das Durchdringen von Wänden, etc. verhindern.

Für direkte Steuerungen gibt es zwei unterschiedliche Kontrollparadigmen, die sich sehr gut z. B. an VRML-Betrachtern ausprobieren lassen: Für große Szenen (etwa Räume) eignet sich am besten der „Flug“-Modus oder als eine Variante der „Lauf“-Modus, der eine konstante Augenhöhe einhält (*eye-in-hand*-Metapher [WARE und OSBORN 1990]). Mausebewegungen in diesen Modi werden üblicherweise in eine Vor-Zurück- bzw. Rechts-Links-Bewegung der Kamera umgesetzt. Drehungen der Kamera (und damit Änderungen der Laufrichtung) werden durch gleichzeitiges Drücken von Sondertasten (Maus oder Tastatur) bewirkt. Dieser Bewegungsmodus entspricht also einer Veränderung der Kameraposition bzw. -blickrichtung, d. h. der Betrachter „bewegt sich selbst“ durch die Welt.

Für einzelne, kleinere Objekte ist eine andere Vorgehensweise besser geeignet: der „Untersuchungs“-Modus (*world-in-hand*-Metapher). Hierbei wird nicht die Kamera bewegt, sondern die Welt, bzw. das untersuchte Objekt. Mausebewegungen werden als eine Drehung der Welt um die Bildschirmwaagerechte bzw. -senkrechte interpretiert,⁸ wogegen Bewegungen der Kamera in Blickrichtung oder zu den Seiten über Sondertasten gesteuert werden. Dies entspricht dem in-die-Hand-nehmen des Objektes.

Mathematisch sind beide Möglichkeiten äquivalent. Welche der Varianten eingesetzt wird, hängt damit von der Art des Betrachtungsgegenstandes ab und welche Erkundungstechniken man damit verbindet. Im Falle eines großen Raumes oder auch schon eines Autos wird man den Flug- bzw. Laufmodus bevorzugen, da er eher als das in-die-Hand-nehmen dem natürlichen

⁸ Alternativ kommt auch der *Archeball*-Algorithmus [SHOEMAKE 1992] zum Einsatz, der die Mausebewegungen auf eine virtuelle Kugel überträgt, deren Orientierung für die Kamera benutzt wird.

Umgang mit dem Gegenstand entspricht. Einen kleinen Kugelschreiber wird man dagegen kaum „durchschreiten“ wollen.

Bei einer indirekten Steuerung wird vom Benutzer oder der Software eine Zieiauswahl (z. B. ein zu besichtigendes Szenenobjekt) getroffen und die Kamera bewegt sich automatisch dorthin. Dem Nutzer steht für diese Bewegung keine zu manipulierende Kamera zur Verfügung, sondern die Bewegung richtet sich an dafür angebotenen Szenenmerkmalen aus.

Beide Steuerungsmodi haben ihre unterschiedlichen Einsatzfelder und können im gleichen System zu verschiedenen Zeitpunkten aktiv sein. Eine direkte Steuerung gibt dem Nutzer die volle Kontrolle über die Kamera, bietet aber keine Navigationshilfen. Diese Steuerungen sind daher anfällig für das oft zitierte *lost in cyberspace* Syndrom⁹ [ELM und WOODS 1985, EDWARDS und HARDMAN 1989, RAAB und RÜGER 1996]. Demgegenüber ist eine indirekte Steuerung zielorientierter und bietet eine Auswahl von „Sehenswürdigkeiten“, die aber auf Kosten der Kontrolle durch den Anwender geht. So kann in Situationen, in denen ein bestimmter Ort schnell aufgesucht werden soll, die indirekte Steuerung wesentlich effektiver sein als langwieriges Suchen (mit der oft sehr realen Möglichkeit der Verirrens). Ist man dann allerdings am Ziel angekommen, ist eine direktere Steuerung wünschenswert, damit die Umgebung „auf eigene Faust“ erkundet werden kann.

Gängige Eingabegeräte für die Kamerasteuerung sind Maus und Tastatur, wobei letztere oft benutzt wird, um die Interpretation der Mauseingaben abzuwandeln, also wahlweise die Kamera zu bewegen oder zu drehen. Problematisch ist insbesondere die Tatsache, daß die Maus ein zweidimensionales Gerät ist, die Kamerakoordinaten aber durch drei Dimensionen beschrieben werden. Im VR-Bereich werden daher häufig Eingabegeräte mit einer höheren Anzahl von Freiheitsgraden benutzt, wie etwa Spaceball (3D-Maus), Datenhandschuhe oder Phantom. Letztere eignen sich gut für Steuerungen nach der *eye-in-hand*-Metapher, erstere passen besser zur genaueren Untersuchung nach der *world-in-hand*-Metapher [RITTER et al. 2000].

⁹ Dieses Phänomen wurde zuerst in Hypertext-Systemen wie dem WWW untersucht, tritt aber auch in graphischen Umgebungen auf. Wegen der potentiellen Verdeckung wichtiger Bildelemente sind 3D-Systeme in besonderer Weise betroffen.

2.8 Personalität

Die Eigenschaft der Personalität beschreibt einen wichtigen Aspekt der Herangehensweise an die Kamerasteuerung und wird deswegen gesondert behandelt. Personalität tritt hauptsächlich in Anwendungen zutage, die einen irgendwie gearteten Personenbezug aufweisen, darunter besonders in Computerspielen, aber auch in technischen Dokumentationen, in denen etwa Bedienpersonal vorkommt. Generell muß es sich aber nicht um eine tatsächliche Person handeln. Den Fokus bzw. das *Subjekt* kann auch ein unbelebter Gegenstand bilden. Man unterscheidet zwischen einer Betrachtungsweise in der ersten bzw. der dritten Person.

Erste Person In der Ich-Form zeigt die Kamera die Szene aus der Perspektive einer virtuellen Spielfigur, die der Spieler selbst steuert. Die Kamera ist also an den Spieler gebunden.¹⁰ Diese Metapher bedingt eine Kamerasteuerung, die durch die Bewegungsmöglichkeiten der Spielfigur bestimmt ist, d. h. die Kamera kann nur an Orte, die der Spieler entsprechend den Spielregeln auch erreichen kann. Damit wird zwar die eigentlich in der 3D-Grafik vorhandene Freiheit der Kamerasteuerung erheblich eingeschränkt, aber gerade diese Einschränkung macht den Reiz des Spieles aus: Räume können nur dadurch in Augenschein genommen werden, daß sich die Spielfigur selbst dorthin begibt und sich dabei auch den damit verbundenen Gefahren aussetzt.

Diese Metapher korrespondiert mit der direkten Steuerung (siehe Abschnitt 2.7), da der Nutzer seine Spielfigur direkt kontrolliert und dabei die Kamera automatisch mitsteuert.

Dritte Person Steht die Kamera außerhalb des Subjekts, spricht man von *third person view*. Diese Metapher räumt der Kamera deutlich mehr Flexibilität ein. Sie kann sich frei im Raum bewegen und unterliegt nicht mehr der direkten Kontrolle des Betrachters.

In einfachen Fällen kann die Kamera einfach die Spielfigur begleiten und dabei wie in der ersten Person an die Figur gebunden bleiben (allerdings nicht

¹⁰ Im Film wird diese Art der Abbildung *subjektive Kamera* genannt.

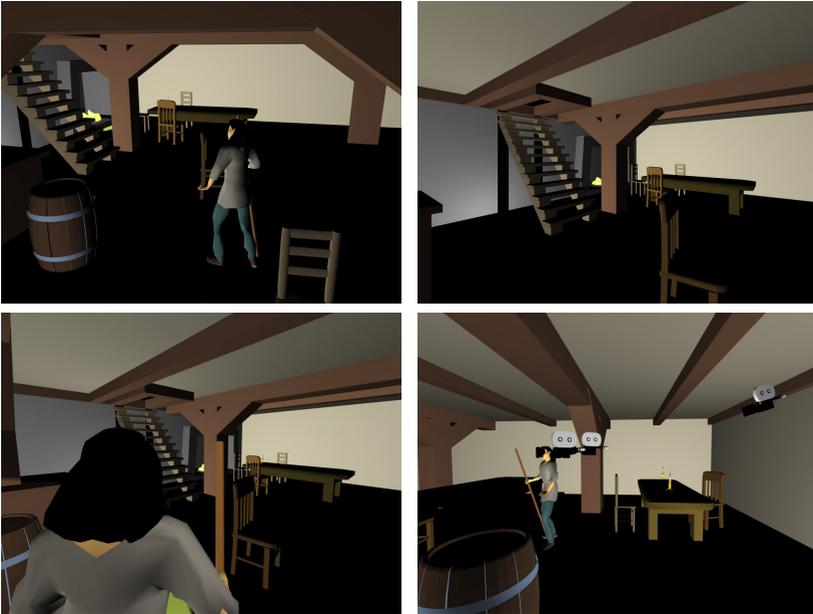


Abbildung 2.4: Links oben: Totale mit Blick auf die Spielfigur. Rechts oben: Subjektive Sicht in einer Szene; der Spieler selbst ist nicht sichtbar. Links unten: Sicht aus der dritten Person in Form einer Ansicht von schräg hinter der Spielfigur. Rechts unten, zur besseren Orientierung: Blick auf alle benutzten Kameras: first-person, third-person und Totale (von links nach rechts).

so starr). Meist wird eine Position schräg hinter dem Akteur eingenommen, so daß sich dem Betrachter fast die gleiche Ansicht wie der Spielfigur bietet.

Diese vereinfachte Umsetzung der Metapher entspricht jedoch noch der direkten Steuerung, da die Kamera weiterhin an die Spielfigur gebunden ist. Obwohl die Kamera von der Spielfigur physisch getrennt ist, also quasi außerhalb steht, wird damit unter Umständen eine stärkere Identifikation des Spielers mit der Spielfigur bewirkt, da diese im Unterschied zur ersten Variante die ganze Zeit im Bild sichtbar ist.

Kameras in der dritten Person sind mit deutlich mehr Aufwand für das System verbunden: Bereits bei der oben beschriebenen einfachen Variante ergibt sich

für die Kamerasteuerung die zusätzliche Aufgabe, sicherzustellen, daß sich hinter der Spielfigur (dort, wo die Kamera plaziert werden soll) keine Hindernisse befinden. Solche Probleme sind bei der starren Kamerasteuerung in der ersten Person prinzipiell nicht vorhanden. Hinzu kommen noch weitere, wie etwa die Berechnung von Übergängen zwischen Kamerapositionen und die Auswahl der konkreten Position bezüglich der Spielfigur.

Für diese Auswahl kann es eine Reihe inhaltlicher Beschreibungsformalisten geben, die definieren, was im Bild zu sehen sein soll (siehe dazu auch Abschnitt 5.1). Für jede dieser Zielvorgaben muß die intelligente Kamera ein Verfahren (im einfachsten Fall eine Formel) beherrschen, das letztlich die gesuchte Kameraperspektive liefert. Der hier mit *Personalität* überschriebene Aspekt der virtuellen Kamera umfaßt also auch die der Kamera verständlichen inhaltlichen Kriterien einer Kameraperspektive.

Filmische Gestaltungsmittel Noch wesentlich aufwendiger wird die Steuerung aber beim Versuch, die Spielsituation zu *filmen*.¹¹ Filme wechseln häufig zwischen erster und dritter Person; der Einsatz von Filmtechniken bedeutet daher eine Abkehr von der klassischen Trennung zwischen erster und dritter Person. Dabei können bestimmte aus dem Film bekannte Techniken auf im Spielverlauf vorkommende Situationen angewandt werden und die Illusion eines vorher durchgeplanten und gefilmten Ablaufes erzeugen, obwohl eine solche Planung in Anwendungen mit einem so hohen Grad an Interaktivität wie Spielen nicht möglich ist. Neben Anleihen aus dem Film kann auch eine Reihe von Techniken z. B. zur Hervorhebung bestimmter im Spiel vorkommender Objekte (wie etwa verborgener Waffen oder Schätze) genutzt werden, die Hinweise an den Spieler geben.

Filmtechniken sind, wie bereits der Name sagt, für Filme konzipiert und funktionieren am besten, wenn die Anwendung möglichst viele Gemeinsamkeiten

¹¹ Mit dem Versuch, die Arbeit von Filmemachern zu automatisieren, begibt sich die Computergrafik wieder einmal in künstlerische Bereiche. Das Ziel ist aber wie auch bei der Nachbildung nicht-photorealistischer Illustrationstechniken in der Computergrafik nicht die Ablösung des Künstlers und der Ersatz von menschlicher Kreativität durch schablonenhafte Automatismen. Vielmehr sollen vorhandene Computeranwendungen von der gesteigerten Ausdruckskraft profitieren und attraktiver werden. Nicht zuletzt fällt durch die Beschäftigung mit traditionellen Arbeitsmethoden ein Lerneffekt für die Entwickler ab, der auch dann vorteilhaft sein kann, wenn die anvisierten Methoden am Computer keinen Erfolg haben sollten.

ten mit einem Film hat.¹² Dazu zählen insbesondere die Existenz einer irgendwie gearteten Handlung und das Auftreten von handelnden Charakteren, wobei deren Physiognomie Menschen oder anderen Lebewesen vergleichbar sein sollte. Bisherige Systeme aus dem Bereich der technischen Dokumentation erfüllen diese Voraussetzung in weit geringerem Maße als Computerspiele und profitieren daher in ihrer Präsentation längst nicht so stark von der Ausdruckskraft filmischer Gestaltungsmittel.¹³ Nichtsdestotrotz können aber auch hier z. B. Kamerafahrten zur Fokusverlagerung von einem Gerät auf ein anderes, oder zur geführten Erkundung eines Raumes benutzt werden. Auch Schnittsequenzen aus einführenden Übersichtseinstellungen, die dann in Detailaufnahmen übergehen, sind durchaus sinnvoll. Weniger erfolgversprechend sind Versuche, ganze filmische Erzählstrukturen (z. B. Parallelhandlungen) oder dramatisierende Mittel (z. B. Vertigoeffekt) zu übertragen.

Mit filmischen Gestaltungsmitteln wird der intelligenten Kamera eine weitere Klasse von Freiheitsgraden hinzugefügt: Die verschiedenen Möglichkeiten der Verknüpfung und Montage von Einstellungen (Kameraperspektiven). Während die einzelnen Einstellungen noch auf Basis der im vorigen Abschnitt erwähnten Beschreibungsformalisten in Echtzeit geplant werden können, erfordert spätestens die Verknüpfung der Einstellungen einen gewissen zeitlichen Vorlauf, der fast zwangsläufig zum Konflikt mit der Interaktivität führt.

2.9 Zusammenfassung

Das hier aufgestellte Modell umfaßt die wesentlichen Aspekte einer Kamerasteuerung, erhebt aber nicht den Anspruch der Abgeschlossenheit. Vielmehr dienen die vorgestellten Bereiche als Beispiele für Kameraeigenschaften, die zwar in Anwendungen benutzt werden, aber üblicherweise an verschiedenen

¹² Das betrifft nicht nur die Frage, ob sie ihre beabsichtigte Wirkung entfalten, sondern auch, ob sie überhaupt anwendbar sind. Mehr dazu in Abschnitt 4.2.

¹³ Emotionale Beteiligung bzw. das Einbringen menschlicher Charaktere sind in technischen Dokumentationen und Animationen nicht immer zwingend erforderlich, bedeuten aber in jedem Fall hohen zusätzlichen Aufwand bei der Modellierung. Allerdings ist zu erwarten, daß sich dieser Extraaufwand im Zuge des technischen Fortschritts verringert.

Stellen modelliert und kontrolliert werden. Es ist nicht auszuschließen, sondern sogar zu erwarten, daß zukünftige neue Anwendungen der Computergrafik weitere Freiheitsgrade der Bilddarstellung und Kamerasteuerung schaffen und ausnutzen. Wesentlich am erweiterten Kameramodell ist daher nicht so sehr die Menge der erfaßten Kamera- bzw. Bilddarstellungsparameter, sondern deren Bündelung in *einer* Kontrollinstanz.

Die Zusammenfassung der verschiedenen Kameraaspekte unter dem Oberbegriff des Erweiterten Kameramodells dient außerdem dem Zweck, das Verständnis für die Zusammengehörigkeit und den gemeinsamen koordinierten Einsatz verschiedener Kameraeffekte zu verbessern.

Im nächsten Schritt kann dieses Modell zur Grundlage einer intelligenten Kamera als Kontrollinstanz zur Koordination dieser Effekte ausgebaut werden, so daß Anwendungen die Dienste der intelligenten Kamera und ihres Kameramodells in Anspruch nehmen können, um bestimmte Situationen bzw. kommunikative Ziele umzusetzen. Im dieser Arbeit zugrundeliegenden Anwendungsfeld sind das vor allem Hervorhebungen und, damit verbunden, Fragen der Sichtbarkeit, sowie die Möglichkeiten der interaktiven und automatischen Steuerung der Kameraperspektive.

3

Hervorhebungstechniken in interaktiven Anwendungen

Intelligente Interaktive Systeme benötigen eine Benutzerführung, die (neben anderen Aufgaben) die Inhaltsvermittlung unterstützt, die Navigation ermöglicht, Aktionen oder Selektionen einzelner Objekte zulässt.

Die dabei durchgeführten Interaktionen können einerseits zu Änderungen in der virtuellen Welt führen und andererseits die Darstellung der Szene oder die Kameraansicht abwandeln. Ein häufiger Grund dafür ist das Bestreben,

- bestimmte Teile des Bildes hervorzuheben
- frühere Hervorhebungen rückgängig zu machen
- die Sichtbarkeit bestimmter Objekte zu gewährleisten.

Hierbei bedient sich die Anwendung jeweils der Möglichkeiten, die in ihrem Kameramodell bzw. der Grafikbibliothek angelegt sind.

Dieses Kapitel analysiert und vergleicht verschiedene Hervorhebungstechniken für den Einsatz in Intelligenten Interaktiven Systemen. Nach einigen Begriffsklärungen wird zunächst auf verschiedene Hervorhebungstechniken eingegangen, wobei deren Einsatzfälle und Konsequenzen beleuchtet werden. Weiterhin werden Kriterien erarbeitet, mit denen im Vorfeld die Eignung bestimmter Hervorhebungstechniken überprüft werden kann.

Eine wesentliche Voraussetzung für Hervorhebungen ist die Sichtbarkeit der betroffenen Objekte. Diese ist in interaktiven 3D-Welten nicht selbstverständlich. Aus diesem Grunde wird dem Zusammenwirken von indirekten Hervorhebungen und Sichtbarkeit besondere Beachtung geschenkt. Im Anschluß werden mit der grafischen Abstraktion und der Unschärfe zwei weitere Techniken vorgestellt, die als Hervorhebungsmittel in Frage kommen.

3.1 Begriffe

3.1.1 Hervorhebungen

Visuelle Hervorhebungen basieren auf den aus der Kognitionspsychologie bekannten *Pop-Out-Effekten*, die *präattentiv*, d. h. in weniger als 200 ms und ohne zielgerichtete Aufmerksamkeit erfaßt werden. Pop-Out-Effekte sind Unterschiede zwischen dargestellten Objekten hinsichtlich bestimmter Darstellungseigenschaften (in der Kognitionspsychologie *Primitive* genannt). Zu diesen Primitiven gehören laut [TREISMAN 1985] Kurven, Orientierung, Farben, Linienstile, Bewegung, Geschlossenheit von Flächen, Kontrast und Helligkeit.

Unter *Hervorhebungen* sollen also vom „Normalfall“¹ abweichende Zeichentechniken verstanden werden, die den so dargestellten Teil besser kenntlich machen.

Dabei ist die Hervorhebung vor allem dazu gedacht, einen offensichtlichen und klar erkennbaren Effekt zu erzeugen und mit dieser Offensichtlichkeit eine kommunikative Funktion zu erfüllen. Diese kann z. B. darin bestehen, auf ein Objekt hinzuweisen oder Ausdruck einer gesonderten Stellung dieses Objektes sein.

In Bild 3.1 könnten z. B. Pfeile angebracht sein, die auf die „nach einem Feuer 1567 aufwendig restaurierte“ Turmspitze verweisen, oder der bereits erkennbar mit anderen Materialien gebaute untere Teil der Vorderseite könnte nochmals weiß umrahmt werden, weil ihn der Betrachter gerade mit der Maus

¹ Selbstverständlich gibt es keinen für alle Anwendungen gültigen Normalfall. Dieser liegt vielmehr im Auge des Betrachters: Alles, was nicht in irgendeiner Weise hervorsteht, wird als „normal“ aufgefaßt.



Abbildung 3.1: Beispiel einer Akzentuierung aus [Hoppe 1998]. Die linke Seite zeigt das Originalbild eines Klosters, wie es beim photorealistischen Rendern entstanden ist. Aufgrund der Lichtverhältnisse sind die im Schatten liegenden Bereiche des Klosters nicht zu erkennen. In der Version rechts wurden die Stützpfeiler und das rechte Tor akzentuiert.

angedlickt hat. In beiden Fällen wäre klar erkennbar, daß die so dargestellten Veränderungen im Bild nicht zum normalen Zeichenstil gehören, sondern eine spezielle Bedeutung haben.

3.1.2 Akzentuierungen

Eng verwandt mit dem Begriff der Hervorhebung ist die *Akzentuierung*. Auch hier wird ein Objekt in seiner Darstellung verändert, aber sehr viel subtiler [HOPPE 1998], d. h. ohne den erwähnten Pop-Out-Effekt. Es geht nicht um einen offensichtlichen Hinweis oder eine Sonderstellung einzelner Bildbestandteile, sondern vielmehr um eine Nachbearbeitung mit dem Ziel, Unzulänglichkeiten des Originalbildes oder des zugrundeliegenden Darstellungsverfahrens auszugleichen.

Auch die oben erwähnte weiße Umrahmung würde den unteren Vorderteil des Klosters im Bild in gewisser Weise akzentuieren; die „echten“ Akzentuierungen im Bild jedoch versuchen, den photorealistischen Stil des Bildes intakt zu lassen und keine erkennbare Abweichung vom Normalfall erzeugen. Man kann daher sagen, Hervorhebungen akzentuieren und Akzentuierungen heben hervor, allerdings verschieden stark und mit unterschiedlichen Zielen und Mitteln.

3.1.3 Direkte und indirekte Hervorhebungen

Hervorhebungen müssen nicht unbedingt nur auf das hervorzuhebende Objekte wirken. Wir unterscheiden daher prinzipiell zwei Arten von Hervorhebungen: *direkte* und *indirekte* [HELBING und PREIM 1998].

Direkte Hervorhebungen variieren den Zeichenstil für das Hervorgehobene so, daß er für eine klare Unterscheidung zum Rest der Szene sorgt. Nicht fokussierte Objekte bleiben unverändert.

Indirekte Hervorhebungen zeichnen alle oder einen Teil der restlichen Szenenbestandteile derart anders, daß sie in ihrer Bedeutung gegenüber dem Hervorzuhebenden zurücktreten (sie können auch ganz weggelassen). Häufig werden Details weggelassen oder zumindest in ihrer Erkennbarkeit erschwert, so daß sich die Aufmerksamkeit unwillkürlich auf die verbleibenden gut erkennbaren Bildbereiche richtet [STROTHOTTE et al. 1994]. Das fokussierte Objekt wird so gezeichnet wie immer und nur durch die Abschwächung anderer Objekte als hervorgehoben erkennbar.

Beide Varianten sind kombinierbar. Dabei entstehen dann drei Hervorhebungsgrade: unwichtige Objekte werden durch indirekte Hervorhebung defokussiert, fokussierte Objekte werden direkt hervorgehoben und damit verbundene, nicht ganz so wichtige Objekte normal dargestellt.

Metagraphische Objekte sind abstrakte, zusätzlich ins Bild gebrachte Elemente wie Pfeile oder Kreise. Sie sind klar als nicht zum Originalbild gehörig erkennbar. Obwohl hier weder die fokussierten Objekte noch der Rest des Bildes anders dargestellt werden, sind metagraphische Objekte den direkten

Hervorhebungen zuzurechnen, weil sie die Aufmerksamkeit explizit auf die Fokusobjekte dirigieren, anstatt wie im indirekten Falle, von etwas abzulenken.

3.1.4 Lokalität von Hervorhebungen

PREIM und RITTER führen mit der *Lokalität* ein weiteres Klassifizierungsmerkmal von Hervorhebungen ein. Sie unterscheiden in [PREIM und RITTER 2002] drei Kategorien:

lokal Die Wirkung beschränkt sich auf das hervorgehobene Objekt selbst.

Dies entspricht einer *direkten* Hervorhebung im Sinne von Abschnitt 3.1.3. Als grafische Mittel werden jegliche Veränderungen der betroffenen Objekte selbst sowie deren Schatten angeführt; zusätzlich sind metagrafische Symbole möglich.

regional Das Objekt und seine unmittelbare Umgebung werden verändert.

Dies kann durch *Cutaway*-Ansichten erreicht werden, die manchmal bei medizinischen Volumendaten Verwendung finden sowie durch Modifikationen verdeckender Objekte mit dem Ziel, den Blick auf die Fokusobjekte freizugeben. Ebenfalls in diese Kategorie fallen 3D-*Fisheye-Zoom*-Darstellungen (siehe Abbildungen 3.2 bis 3.3 auf Seiten 42–42).

global Das gesamte Bild ist betroffen.

Ursache kann z. B. eine veränderte Kameraperspektive sein, oder eine *indirekte* Hervorhebung, bei der *alle* unwichtigen Objekte deakzentuiert sind.

3.1.5 Dynamische Hervorhebungen

Dynamische Hervorhebungen umfassen solche Effekte, die nicht innerhalb eines einzelnen Bildes wirken, sondern erst durch eine schrittweise Veränderung von Darstellungsparametern über eine ganze Bildserie hinweg entstehen. Im Gegensatz dazu stehen statische Effekte wie im vorherigen Abschnitt beschrieben. Sie wirken bereits im Einzelbild.

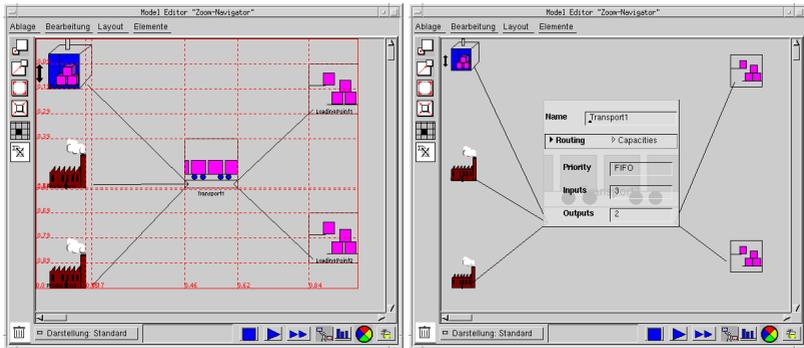


Abbildung 3.2: Kontinuierlicher Zoom in Benutzungsschnittstellen (aus [Rüger 1998]). Der begrenzte Platz auf dem Bildschirm wird optimal den wechselnden Anforderungen entsprechend verwaltet. Dabei macht das System Annahmen über den Fokus des Nutzers, z. B. basierend auf seiner Interaktion mit den Bedienelementen. Das Layout der Symbole im Fenster wird dabei entsprechend angepaßt (verzerrt). Der Hervorhebungseffekt, der hauptsächlich durch die kurze Animation beim Zoomen entsteht, ist gegenüber der intelligenteren Platzaufteilung sekundär.

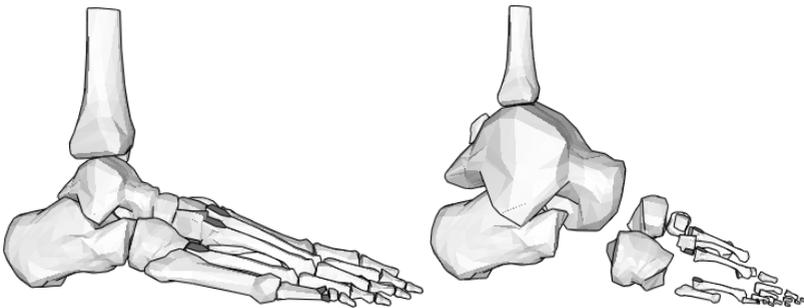


Abbildung 3.3: Die Anwendung des auf 3D erweiterten kontinuierlichen Zooms [Raab und Rüger 1996] auf Abbildungen realer Objekte kann zu Bildern führen, die wegen der zahlreichen nichtlinearen Verzerrungen nicht mehr ausreichend mit dem Gegenstand übereinstimmen. Als Hervorhebungsmittel in medizinischen Illustrationen wie auch für technische Dokumentationen ist der 3D-Fisheye-Zoom weniger geeignet, da sich die Größenverhältnisse der einzelnen Teilobjekte in unnatürlicher Form verändern.

Beispiele für einfache dynamische Hervorhebungen sind gleitende Farbübergänge (Pulsieren, Blinken). Größen- oder Ortsveränderungen von Objekten erfordern etwas mehr Aufwand zur Planung und Implementation. Kamerabewegungen schließlich erfordern sowohl die Planung einer neuen Kameraposition als auch eines Pfades dorthin.

Dynamische Hervorhebungen ergänzen statische Techniken. Gedruckte Bilder können aufgrund ihres statischen Charakters ausschließlich diese statischen Hervorhebungen benutzen. Rechnergestützte interaktive Systeme besitzen diese Einschränkung nicht und sind daher potentiell ausdrucksstärker, da sie über eine reichere Palette von Techniken verfügen.

3.1.6 Situationen

Eine wichtige Voraussetzung für die automatisierte Anwendung grafischer wie auch filmischer Mittel in interaktiven Systemen ist die Definition eines wie auch immer gearteten Situationsbegriffs. Dieser dient dazu, die einsetzbaren Darstellungsmittel und Interaktionstechniken zu den möglichen Zuständen des Systems in Beziehung zu setzen und ihren zeitlichen Ablauf zu koordinieren. Die Menge und Art der Darstellungsmittel wie auch der Situationen hängt vom betrachteten System bzw. Medium ab.

- In vielen interaktiven Systemen, besonders aber in Technischen Dokumentationen wird die Situation von *kommunikativen Zielen* bestimmt. Der Begriff der kommunikativen Ziele ist allerdings eher für Illustrationen und andere statische Inhalte gedacht und berücksichtigt keine interaktiven Aspekte.
- Für Computerspiele wird zuweilen der Begriff *activity* benutzt, der die aktuell vom Spieler oder vom Computer ausgeführten Vorgänge bzw. Spielzüge beschreibt [CHRISTIANSON et al. 1996].
- Die Filmkunst hingegen unterscheidet mittlerweile unzählige Situationen mit zahlreichen Untersituationen und Varianten und bietet jeweils angepaßte Einstellungssequenzen an. Werke wie [ARIJON 1976] bestehen im Prinzip aus einer Sammlung von Situationen und zugehörigen Filmtechniken.

Insbesondere muß es keine 1:1-Beziehung zwischen Situation und Darstellung geben. So kann die gleiche Situation mit durchaus unterschiedlichen oder zumindest variierten Mitteln dargestellt werden, was besonders bei einem eher geringen Situationsvorrat Langeweile vermeiden hilft. Umgekehrt können auch verschiedene Situationen mit derselben Darstellung gezeigt werden. Trotzdem ist die Kenntnis der aktuellen Situation die Entscheidungsgrundlage für den koordinierten Mitteleinsatz. *Mittel* sind in diesem Zusammenhang nicht nur auf die Bilddarstellung bezogen, sondern auch auf die interaktiven Elemente, z. B. ob und wie der Nutzer im Moment navigieren darf.

3.2 Sichtbarkeit

Eine wesentliche Voraussetzung für jede Anwendung von Hervorhebungstechniken ist die Sichtbarkeit der hervorzuhebenden Szenenelemente [SELIGMANN UND FEINER 1991]. Dieser Umstand ist aus der Perspektive handgezeichneter Darstellungen leicht zu übersehen: Ein Zeichner wird immer eine Perspektive wählen, die eine optimale Sicht auf die wesentlichen Aspekte einer Darstellung gestattet. Innerhalb dieser im Hinblick auf die Hervorhebungen ausgewählten Perspektive werden dann gewisse Zeichenattribute oder Darstellungsweisen variiert.

In interaktiven Systemen ist dieser Zustand nicht unbedingt gegeben. Der Nutzer kann frei navigieren und dabei eine beliebige Betrachterposition einnehmen oder sogar in die „falsche Richtung“ sehen. Nutzerinteraktionen, die das Zeigen eines bestimmten Gegenstandes erfordern („zeige mir Objekt A“) oder auch eine vom System geplante Erklärung, bei der die beteiligten Gegenstände erkennbar sein müssen („Motor M treibt Keilriemen K an“), werden durch Hervorhebungstechniken unterstützt.

Der Einsatz direkter und indirekter Hervorhebungsmittel versagt aber, wenn das Hervorgehobene nicht im Blickfeld liegt oder verdeckt wird. Zudem muß eine minimale Größe der hervorgehobenen Objekte auf dem Bildschirm gewährleistet werden, so daß auch die Größe des Sichtfeldes bzw. die Entfernung der Kamera vom fokussierten Objekt zu beachten ist. Das System muß daher

zunächst die Sichtbarkeit garantieren, was auf zwei Arten realisiert werden kann:

- Veränderungen der Szene: Weglassen oder Transparenz verdeckender Objekte
- Veränderungen der Kameraposition

Beide Möglichkeiten haben unterschiedliche Anwendungsfälle und Konsequenzen. Im weiteren Verlauf dieses Abschnitts werden verschiedene Varianten der Szenenveränderung untersucht. Für die Veränderung der Kameraposition sind einige zusätzliche Betrachtungen wichtig, die in Abschnitt 4 folgen und in das im Abschnitt 5 vorgestellte Verfahren münden.

Veränderungen der Szene umfassen veränderte Darstellungsattribute einzelner Objekte bis hin zum Weglassen. Nicht betrachtet werden dagegen Bewegungen von Objekten, um sie etwa „aus dem Weg zu räumen“. Hierfür müßten zusätzlich Durchdringungen der weggeräumten Objekte mit weiteren Objekten vermieden werden, was weiteren Aufwand bedeutet. Während die Kollisionserkennung zwar aufwendig, aber eventuell noch handhabbar ist (siehe [GOTTSCHALK et al. 1996]), sind wichtige Gütekriterien wie die Auswirkungen auf das Arrangement der restlichen Szenerie und die Plausibilität der Szene, wenn überhaupt, nur schwierig abzuschätzen.

3.2.1 Ausblenden verdeckender Objekte

Das Weglassen störender Objekte ist der algorithmisch einfachste Weg, die Sichtbarkeit dahinterliegender Objekte sicherzustellen. Die Objekte werden nicht aus der Szene entfernt, sondern nur vorübergehend nicht mitgezeichnet. Ändert sich der Fokus, tauchen sie wieder auf. Da sich für die fokussierten Objekte nichts ändert, handelt es sich hierbei um eine indirekte Hervorhebung. Problematisch an diesem Verfahren ist, daß dadurch der Zusammenhang zum Rest der Szene verlorengehen kann, besonders wenn es sich bei den weggelassenen Objekten um Dinge aus dem Umfeld des Fokus handelt.

Diese Methode erscheint daher nur sinnvoll, wenn die Anzahl der entfernten Objekte im Vergleich zur Gesamtszene relativ klein ist. Allerdings ist dann

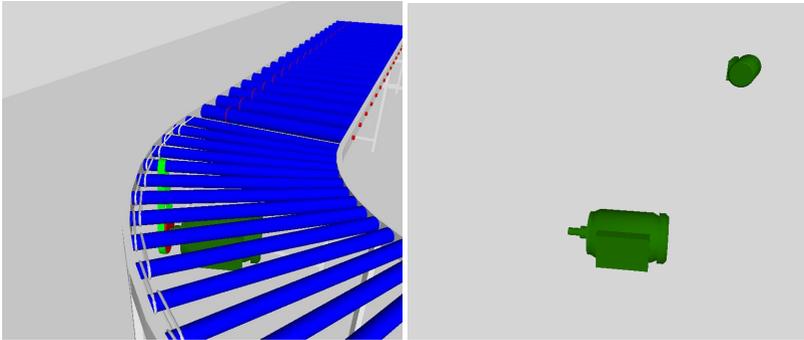


Abbildung 3.4: Eine Szene (a) ohne erkennbare Hervorhebung und (b) mit Weglassen der defokussierten Objekte. Im rechten Bild fehlt klar der Kontext, der eine sinnstiftende räumliche Einordnung der Motoren möglich macht.

auch der Effekt am geringsten. Es ist deswegen sinnvoller, zumindest eine Andeutung (z. B. einige Konturlinien) der entfernten Szenenbestandteile beizubehalten. Für eine interaktive Software ist das Weglassen auch deshalb nicht zu empfehlen, weil der Programmierer im Unterschied zu einem Zeichner den Effekt der Weglassungen auf die Verständlichkeit der Darstellung nicht vorab beurteilen kann. Prinzipiell trifft diese Überlegung zwar auch auf alle anderen Techniken zu, jedoch sind die allesamt weniger drastisch.

3.2.2 Transparenz

Eine Alternative zum simplen Weglassen stellt Transparenz dar. Der Betrachter kann durch unwichtige Objekte hindurch sehen und das verdeckte Fokusobjekt erkennen. Er erhält damit gleichzeitig einen Hinweis auf die geringere Bedeutung der durchsichtigen Objekte. Da diese aber immer noch sichtbar sind, bleibt, anders als beim einfachen Weglassen, der Kontext erhalten. Grundsätzlich werden bei dieser Methode zuerst alle undurchsichtigen Objekte gezeichnet. Anschließend werden ihnen die durchsichtigen Objekte überlagert, wobei sich deren Farbe entsprechend dem Transparenzwert mit

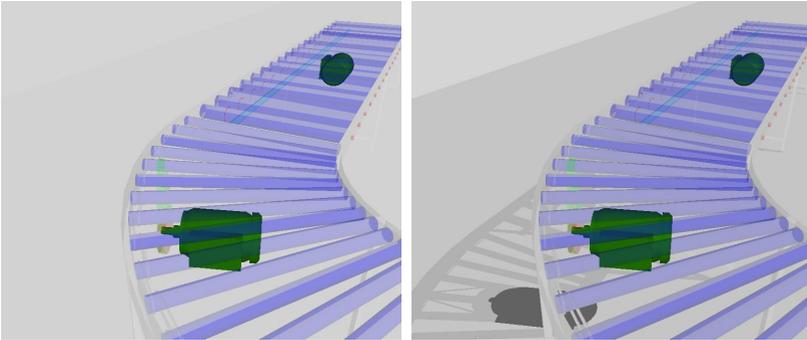


Abbildung 3.5: Defokussierte Objekte wurden transparent dargestellt. Die Verwendung von unterschiedlich starken Schatten (ebenfalls mit Berücksichtigung des Fokus) verbessert den 3D-Eindruck.

der darunterliegenden Farbe mischt.² Jedem verdeckenden Objekt muß nun ein solcher Transparenzwert zugewiesen werden.

Es ist allerdings nicht ganz einfach, die richtigen Transparenzwerte zu finden, mit denen die durchsichtigen Objekte gezeichnet werden sollen. Ein niedriger Wert bedeutet hierbei hohe Transparenz, also geringen Einfluß des transparenten Objektes auf die endgültige Farbe. Der Wert muß hoch genug sein, damit das verdeckende Objekt trotz allem noch erkennbar ist (sonst könnte man es ja auch weglassen und damit den Kontext verlieren). Er muß aber auch niedrig genug sein, so daß dahinterliegende Objekte gut sichtbar bleiben. Aber egal, wie klein der Wert gewählt wird, jedes verdeckende Objekt trägt ein wenig Farbe auf das verdeckte Objekt auf. Die Farbbeiträge aller verdeckenden Objekte akkumulieren sich und wenn ein Fokusobjekt von zu vielen anderen Objekten verdeckt wird, wird es schlecht oder gar nicht mehr erkennbar sein, selbst wenn jedes einzelne verdeckende Objekt transparent ist. Damit beschränkt sich die Anwendbarkeit dieser Technik auf Fälle, in denen

² Bis in die jüngste Zeit war die korrekte Darstellung transparenter Flächen relativ aufwendig, weil dafür eine Zeichenreihenfolge eingehalten werden muß, die der Tiefenanordnung der Objekte entlang der aktuellen Blickrichtung entspricht. Modernere Grafikhardware eröffnet hier neue Möglichkeiten [EVERITT 2001], für die die folgenden Betrachtungen allerdings weiterhin gelten.

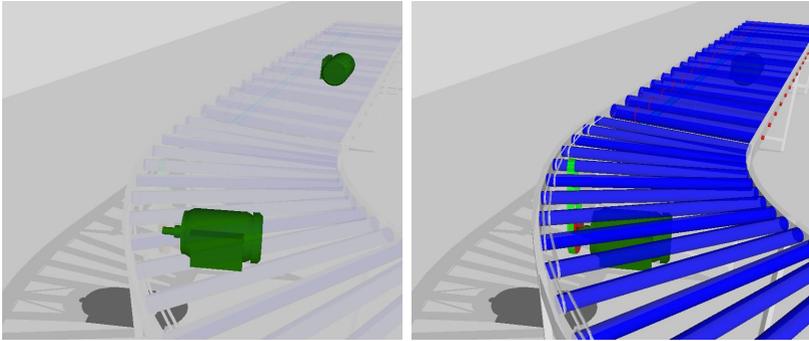


Abbildung 3.6: Zu starke (a) und zu schwache (b) Transparenz führen zu unnötig stark verschlechterter Darstellung der Umgebung des Fokus bzw. zu übermäßig starker Verdeckung.

die Anzahl der verdeckenden Objekte relativ gering ist bzw. eine gute Abschätzung für einen optimalen Transparenzwert getroffen werden kann.

3.2.3 Drahtgitter

Einen der Transparenz ähnlichen Effekt erzielt man mit Drahtgitterdarstellungen. Waren diese zu Beginn der CAD-Ära aufgrund des damaligen Entwicklungsstandes der 3D-Hardware noch der meistgenutzte Darstellungsmodus,³ werden sie heute hauptsächlich wegen ihrer „Durchsichtigkeit“ genutzt, weil so auch mit weiter hinten liegenden Eckpunkten und Linien interagiert werden kann, die bei Verwendung gefüllter Flächen nicht sichtbar wären. Dadurch entfallen zeitaufwendige Operationen, die das Objekt erst in Position drehen müssten, bevor die gewünschten Punkte zugänglich würden.

Wie bei der Transparenz sollten also auch hier hintenliegende Objekte klar sichtbar sein. Dies ist aber nur dann zu erwarten, wenn das Drahtgitter nicht übermäßig dicht ist (siehe Abbildung 3.8 auf Seite 50). Heutige Modelle dagegen sind oftmals sehr detailliert und bestehen aus einer großen Anzahl recht

³ Für gefüllte Flächen müssen normalerweise wesentlich mehr Pixel gezeichnet werden, was entsprechend hohe Anforderungen an die Speicherbandbreite der Grafikkhardware stellt. Heutige 3D-Systeme sind hard- und softwaremäßig auf gefüllte (texturierte) Flächen optimiert und eher langsamer, wenn es um Linien geht.

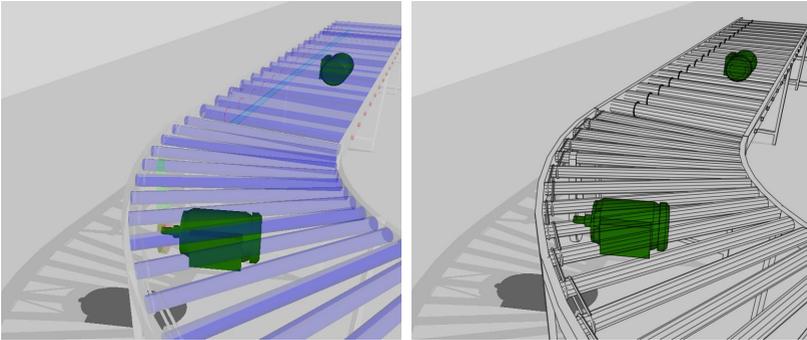


Abbildung 3.7: Anstelle der Transparenz (a) wird eine Drahtgitterdarstellung benutzt, um das fokussierte, aber verdeckte Objekt sichtbar zu machen.

kleiner Polygone, die in der Projektion auf dem Bildschirm teilweise nur noch wenige Pixel groß sind. Dies führt zu sehr vielen Polygonkanten und macht dahinterliegende Objekte auch dann noch fast unsichtbar, wenn nur die sichtbaren Kanten gezeichnet werden (was ja im Durchschnitt nur eine Halbierung der Kantenanzahl bedeutet). Die Anzahl der Kanten kann außerdem zwischen verschiedenen Objekten stark variieren. Einen Ausweg stellt die Filterung der Kanten nach einem Schärfekriterium dar, bei der nur noch Kanten dargestellt werden, deren angrenzende Polygone einen gewissen Mindestwinkel zueinander einnehmen.

Linien haben außerdem die angenehme Eigenschaft, daß sie zusätzlich zu der normalen gefüllten Darstellung genutzt werden können, um besondere kommunikative Ziele (z. B. Hervorhebungen, siehe Abschnitt 3.3) zu vermitteln. Daher ist es immer von Vorteil, wenn die benutzte 3D-Software außer der Flächendarstellung auch noch Drahtgitter unterstützt.

3.2.4 Zusammenfassung

Die sichtbarkeitsrelevanten Veränderungen, die die beschriebenen Techniken hervorrufen, basieren jeweils auf einer mehr oder weniger ausgeprägten Deakzentuierung der unwichtigen Objekte. Diese Deakzentuierung ist selbst ein in-

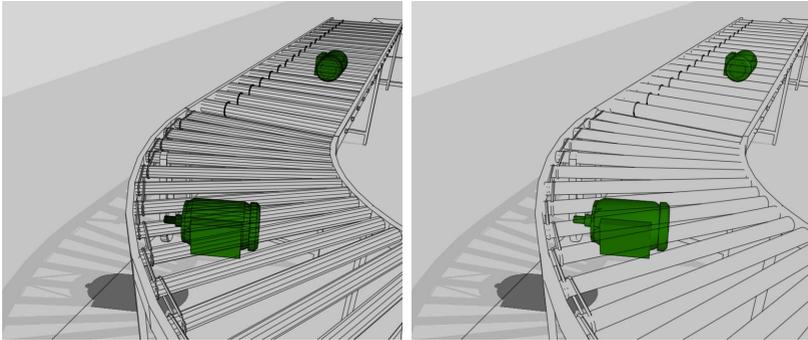


Abbildung 3.8: Variationen des Mindestwinkels zwischen zwei benachbarten Flächen im Modell führen zu unterschiedlich dichten Drahtgitterdarstellungen. Im linken Bild werden alle Modellkanten angezeigt, wogegen rechts nur die schärfsten Kanten (Winkel über 90°) und die Silhouetten dargestellt werden.

direktes Hervorhebungsmittel (siehe Abschnitt 3.1.3), das bereits unabhängig von einer eventuell vorgenommenen direkten Hervorhebung der wichtigen Objekte wirkt. Beschränkt man die Veränderungen auf die tatsächlich störenden Objekte (deren Bestimmung je nach Blickwinkel zusätzlichen Aufwand beim Zeichnen erfordert), so verliert die Deakzentuierung ihren Hervorhebungscharakter.

3.3 Zeichenattribute zur Hervorhebung

Hervorhebungen in interaktiven Dokumentationen erfüllen immer eine Hinweisfunktion. Anders als Akzentuierungen⁴ sind sie nicht nur dazu da, um bestimmte Aspekte besser sichtbar zu machen. Damit ein Hinweis als solcher wahrgenommen wird, muß er eindeutig erkennbar sein und sich klar von Rest der Zeichnung unterscheiden. Anderenfalls besteht die Gefahr, daß er übersehen oder in seiner Bedeutung nicht erkannt wird.

Variationen der Zeichenstile erfüllen die Funktion, Teile einer Szene vom Rest unterscheidbar zu machen. Die Attributsätze, die beim Zeichnen eine Rolle

⁴ siehe Abschnitt 3.1.2

spielen, sind von Anwendung zu Anwendung unterschiedlich. Sie hängen u. a. von den benutzten Modellen und der Zielplattform sowie den Darstellungsalgorithmen ab. Aus diesem Grund muß jede Anwendung Hervorhebungsparameter heranziehen, die nicht ohnehin schon in den normalen Darstellungen benutzt werden.

3.3.1 Farbe

Ein beliebter Hervorhebungsparameter ist die Farbe. Einem Objekt wird eine Sonderfarbe zugewiesen, die sowohl einen hohen Kontrast zur Umgebung wie auch zur ursprünglichen Farbe hat. Eine solche Farbe zu finden, kann in einer 3D-Anwendung schwierig sein, da hier die Umgebung im Allgemeinen eine unbekannte und uneinheitliche Farbe haben wird. Außerdem unterliegen die Farben eventuell noch Beleuchtungseffekten. Alles in allem macht das eine Analyse der generierten Bilder erforderlich, um eine zur Umgebung möglichst kontrastreiche Farbe zu finden. Da diese Farbe sich in jeder neuen Hervorhebungssituation ändern kann, wird sich auch keine Signalwirkung (*Pop-Out-Effekt*, siehe Abschnitt 3.1.1) etablieren lassen, was die Wirksamkeit der Hervorhebung einschränkt.

Wird nun beispielsweise aufgrund ihrer allgemein akzeptierten Signalwirkung die rote Farbe durchgängig als Hervorhebungsmittel benutzt, so wird das nur dann funktionieren, wenn im Modell keine roten Objekte enthalten sind. Nicht nur könnten diese roten Objekte nicht noch roter werden, wären also durch Rotfärbung nicht mehr hervorhebbar; auch für die restlichen Objekte müßte sichergestellt sein, daß der Nutzer weiß, daß sie im Normalfall nicht rot gezeichnet werden. Obendrein verliert das Rot seine Signalwirkung, wenn sich das hervorgehobene Objekt inmitten roter Objekte befindet. All dies ist ohne Restriktionen der Modelle (z. B. Verbot roter Objekte) schwer sicherzustellen. Für Systeme, in denen bekannt ist oder sichergestellt werden kann, daß es keine roten Objekte gibt, ist dieses einfache Vorgehen sinnvoll. Aber genau in diesem Fall gilt die oben gemachte Aussage, daß eine Hervorhebung sich am besten eines Darstellungsparameters bedient, der in der normalen Anwendung sonst nicht auftritt.

3.3.2 Zeichenstil

Modelle für 3D-Szenen bestehen oft hauptsächlich aus Polygondaten mit Farbinformation, allerdings ohne daß dabei ein bestimmter Darstellungsstil festgelegt wird. Die Polygone werden üblicherweise gefüllt dargestellt, wobei aber nicht einfach die zugeordnete Farbe benutzt wird, sondern eine Schattierung, wie sie an beleuchteten Körpern zu beobachten ist. Die schattierte Darstellung dient einem durch die Helligkeitsverläufe verbesserten 3D-Eindruck (*shape from shading*) und wird auch benutzt, ohne daß eine Lichtquelle in der Szenenbeschreibung enthalten ist. In diesen Fällen wird meist eine Lichtquelle unmittelbar oder in einem gewissen Abstand an die Kamera gebunden.

Eine Anwendung hat also in vielen Fällen gewisse Freiheiten, die Modelldaten zu interpretieren. Diese Freiheiten können zu Hervorhebungszwecken genutzt werden.

Beispielsweise kann die Anwendung die Polygone als Drahtgitter interpretieren und erhält damit eine Möglichkeit zur Hervorhebung, die nicht die oben beschriebenen Nachteile der Verwendung von Sonderfarben hat. Werden Objekte mit Konturlinien oder als Silhouette nachgezeichnet, werden sie von anderen Objekten klar unterscheidbar.

3.3.3 Licht und Schatten

Eine andere Möglichkeit ist eine Variation der Beleuchtung, so daß weniger wichtige Objekte nur abgeschwächt beleuchtet werden. Der visuelle Effekt dieser Methode und damit die Erkennbarkeit der Hervorhebung hängt allerdings nicht unwesentlich von der Grundfarbe und den Materialeigenschaften der Objekte ab. So können helle Objekte auch ohne starke Beleuchtung noch klar und kontrastreich bleiben. Die Hervorhebung manifestiert sich in solchen Fällen nicht so sehr im eigentlichen Erscheinungsbild eines Objektes, sondern in seiner Änderung, die als eine Abweichung von Normalzustand erkennbar wird.

In ähnlicher Weise kann Schatten als zusätzlicher Effekt dienen, wobei es gleichgültig ist, ob die Schatten tatsächlich durch physikalisch exakte Berechnungen aufgrund eines globalen Beleuchtungsmodells entstehen. Als flexibel

steuerbares Hervorhebungsmittel läßt sich Schatten viel besser dann einsetzen, wenn seine Stärke nicht von physikalischen Eigenschaften der schattenwerfenden Körper abhängt, sondern kommunikativen Zielen wie etwa der Wichtigkeit eines Objektes unterliegt (siehe Abbildung 3.5 auf Seite 47).

3.4 Abstraktion

Nicht so sehr zur Hervorhebung als vielmehr für den gegenteiligen Effekt sind abstrahierte graphische Darstellungen geeignet [SCHLECHTWEG 1999, STROT-HOTTE 1998, BUTZ und KRÜGER 1996].

Der Abstraktionsbegriff der Computergrafik ist sicher nicht der gleiche wie in der abstrakten Malerei, oder gar in der Philosophie. Seine Grundbedeutung allerdings – die Reduktion auf das Wesentliche – kann im Kontext interaktiver Computergrafik als *Erkennen und Eliminieren von für die Bildaussage entbehrlichem Detail* verstanden werden. Das läßt offen, ob das derzeit gebräuchlichste grafische Abstraktionsmittel – die Reduktion der geometrischen Komplexität [DEUSSEN et al. 1999], also Anzahl der Flächen, Linien bzw. Punkte eines Modells – der einzige Weg zur Abstraktion sein muß. Weiterhin läßt sich zeigen, daß sich Abstraktion und Realismus nicht unbedingt ausschließen [SCHIRRA und SCHOLZ 1998].

Zur Abstraktion werden zeichnerische Ressourcen (etwa die Anzahl oder Gesamtstrecke der gezeichneten Linien oder Flächen) unter den zu zeichnenden Objekten entsprechend ihrer Wichtigkeit verteilt (siehe Abbildung 3.9). Das kann bei den weniger wichtigen Objekten zu einer Ausdünnung der gezeichneten Linien bzw. einer groberen Darstellung führen. Daraus ergibt sich ein deakzentuierender Effekt, der jedoch weniger offensichtlich ist als die bisher aufgeführten Hervorhebungsmittel.

In jedem Fall setzt Abstraktion durch kontrollierte Zuweisung von Zeichenressourcen einen augenfälligen Zusammenhang zwischen der Wichtigkeit eines gezeichneten Objektes und der Menge der von ihm verbrauchten zeichnerischen Ressourcen voraus. Diese Annahme trifft in vielen Fällen zu, nämlich immer dann, wenn die geometrischen Modelle gerade so detailliert (oder ressourcenintensiv) sind, wie es für die Anwendung gebraucht wird. Zuweilen sind die Modelle aber deutlich aufwendiger oder enthalten ungewollte Details

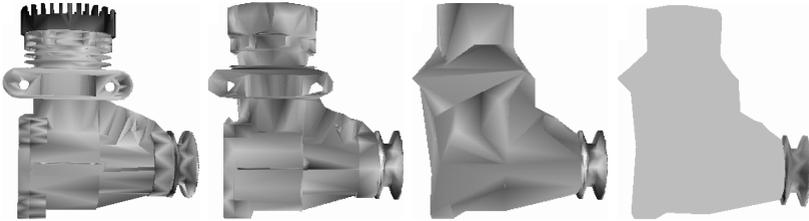


Abbildung 3.9: Beispiel für eine abstrahierte technische Darstellung eines Motors aus [Krüger und Rist 1995]. Links die komplette Darstellung mit allen Polygonen. Daneben eine weniger detaillierte Ansicht, die besonders im oberen Teil gut sichtbare Unterschiede aufweist. Erst im dritten Bild fällt auf, daß die Drehscheibe rechts von der Abstraktion ausgenommen ist: vom restlichen Motor ist nur noch die grobe Form zu erkennen. Ganz offensichtlich wird es im letzten Bild. Hier wird sogar auf die Schattierung der Polygone verzichtet.

und Artefakte (z. B. Modelle aus 3D-Scannern). In solchen Fällen muß nicht nur zur Deakzentuierung, sondern auch schon für die normale Darstellung eine geometrische Abstraktion erfolgen.

Abstrahierte Darstellungen werden also häufiger aus einem anderen Grund eingesetzt: die verringerte Anzahl von Linien bzw. Polygonen senkt den Zeichenaufwand zum Teil beträchtlich, was sich besonders dann auszahlt, wenn die geometrischen Modelle ungewollt detailliert sind. Wenn etwa der Motor aus Abbildung 3.9 auf dem Bildschirm nur noch wenige Pixel bedeckt, dafür aber in zahlreichen Instanzen auftaucht, wird selbst bei schneller Grafikhardware die Interaktivität beeinträchtigt. Ein etwaiger Hervorhebungs- bzw. Abschwächungseffekt ist hier bestenfalls ein Nebenprodukt. Die Verfahren zur Polygonreduktion sind im Gegenteil gerade darauf bedacht, trotz der drastisch verringerten Polygonanzahl eine vom Originalmodell nicht unterscheidbare Darstellungsqualität zu liefern [ECK et al. 1995, HOPPE 1996].

3.5 Unschärfe

Den exakten photorealistischen Abbildungen aus dem Computer liegt fast immer das vereinfachte Abbildungsmodell der Lochkamera zugrunde (siehe Abschnitt 1.1.5). Dieses Modell hat die Unschärfe aus computergenerierten Bil-

dern verbannt. Um sie zurückzubekommen, müßte wesentlich höherer Rechenaufwand betrieben werden als für „korrekte scharfe“ Bilder. Warum sollte man nun viel Mühe auf etwas verwenden, was man eigentlich nicht will?

Unschärfe tritt nicht nur in realen Kameras auf, sondern auch im menschlichen Wahrnehmungsapparat. KOSARA et. al. konnten in Untersuchungen zeigen, daß Unschärfe ebenfalls zu den in Abschnitt 3.1.1 erwähnten Pop-Out-Effekten gehört [KOSARA et al. 2002].

Schärfentiefe wird seit langem als indirektes Hervorhebungsmittel⁵ in der Fotografie benutzt. Schärfenverlagerungen zur Verschiebung des Fokus zwischen Bildebenen treten auch in Filmen auf. In fotografischen Abbildungen wirkt Unschärfe noch auf das gesamte Bild. Objekte werden nur aufgrund ihrer Lage zur Fokusebene (die durch ihre Entfernung von der Kamera definiert ist) scharf oder unscharf abgebildet. Der Einsatz von Schärfentiefe zur Hervorhebung ist also an eine entsprechende Bildkomposition gebunden, die unterschiedlich wichtige Objekte auch in unterschiedlichen Bildebenen anordnet. Diese Einschränkung besteht in der Computergrafik nicht mehr. Hier können Objekte mit unterschiedlicher Schärfe dargestellt werden, auch wenn sie die gleiche Entfernung zur Kamera haben.

Die genaue Entstehung der Unschärfe im Linsensystem einer Kamera oder im menschlichen Auge ist für ihren präattentiven Effekt letztlich bedeutungslos. Daher kann sie in interaktiven Systemen auch ohne Simulation einer Linse durch andere Mittel wie Texturen oder nachträgliches Verwischen (*blurring*) des Bildes erzeugt werden. Sofern derartige Techniken von der Grafikengine unterstützt werden, steht mit der Unschärfe ein vielseitiges indirektes Hervorhebungsmittel zur Verfügung.

3.6 Zusammenfassung

Dieses Kapitel stellte einige wichtige Begriffe vor, mit denen Hervorhebungen klassifiziert werden können. In Abgrenzung zum weniger starken Mittel

⁵ Sieht man die scharfe Darstellung als den Normalfall an, wird die Hervorhebung durch Deakzentuierung der unwichtigen Objekte erzeugt.

der Akzentuierung werden Hervorhebungen durch ihre Signalwirkung charakterisiert. Dies hat Einfluß auf die zur Hervorhebung in frage kommenden Darstellungsmittel.

Hervorhebungen sind letztlich immer Abwandlungen der verfügbaren Darstellungsmittel. Je nach Blickwinkel (Modellentwurf oder Systementwurf) richtet sich die Wahl der Darstellungsmittel nach den Möglichkeiten der Grafikengine bzw. das Design der Grafikengine nach den gewünschten Darstellungsmitteln.

Entscheidend für die Möglichkeit, überhaupt unmittelbar erkennbare Hervorhebungen zu erzeugen, ist das Vorhandensein von Darstellungsmitteln, die von den Szenenmodellen normalerweise nicht ausgenutzt werden. Damit ist die Grafikengine in der Lage, Modelldaten je nach Situation verschieden zu interpretieren und durch diese Uminterpretation eine Signalwirkung zu erzielen. Die oben aufgezählten Methoden sind Beispiele dafür.⁶

Der Interpretationsspielraum kann dann auf seiten der Anwendung ausgeschöpft werden, wobei auch mögliche Nebenwirkungen oder Defizite beachtet werden müssen. Im Beispiel der abstrahierten Darstellung aus Abbildung 3.9 auf Seite 54 wären zum Hinweis auf die Drehscheibe durchaus auch Alternativen denkbar. So könnte anstelle der Abstraktion des Motors auch Transparenz benutzt werden, um den Motor zu deakzentuieren. Beide Techniken wären sogar kombinierbar. Das muß nicht immer gelten: Bei einer Kombination mit Drahtgitterdarstellung anstelle von Transparenz würde durch die Liniendarstellung die im Zuge der Abstraktion veränderte Struktur des geometrischen Modells offensichtlich. Es ist zu vermuten, daß dieser Effekt eher ablenkend wirkt. Beide Alternativtechniken haben zudem den Nebeneffekt, daß hinter dem Motor befindliche Objekte sichtbar werden, was ebenfalls ablenken könnte.

Eine Auswahl der beschriebenen Techniken wurde in den interaktiven Systemen, die im Abschnitt 6 vorgestellten werden, implementiert. Dort finden

⁶ Besonders Nicht-Photorealistische Anwendungen bieten noch zahlreiche weitergehende Möglichkeiten. Dies erklärt sich einerseits aus den andersartigen zeichnerischen Mitteln, aber mehr noch aus ihrer Losgelöstheit vom einschränkenden Diktat des Realismus. Eine veränderte Art der Schattierung etwa bedeutet hier nicht automatisch, daß das betroffene Objekt sich irgendwie verändert hätte, da bereits die ursprüngliche Schattierung nicht realistisch war.

sich auch Erläuterungen, die die Auswahl der Techniken und ihre Kombination begründen und motivieren.

Intelligente Interaktive Systeme müssen bei der Wahl von Hervorhebungsmitteln die verfügbaren Möglichkeiten ihrer Grafikengine ausnutzen und geeignet kombinieren. Hierfür wird eine Kontrollinstanz benötigt, die die Darstellungsparameter der Bilderzeugung (Zeichenstile, deren lokale, regionale oder globale Variation, Annotationen, Abstraktionen, etc) in Abhängigkeit von Situationen und kommunikativen Zielen steuert. Das im letzten Abschnitt vorgestellte erweiterte Kameramodell ist für diese Aufgabe geeignet, wenn es als Vermittlungsstelle zwischen den Eigenschaften der Grafikengine und den Bedürfnissen der Präsentationssteuerung über die jeweils passenden Umsetzungsregeln ergänzt wird (siehe Abbildung 2.2 auf Seite 24).

4

Filmische Gestaltungsmittel

Die im letzten Kapitel besprochenen Hervorhebungsmittel waren meist Adaptionen traditioneller Illustrationstechniken an die Mittel der Computergrafik (Variationen des Zeichenstils, Hinzufügen bzw. Weglassen von Details, Ändern von Proportionen). Sofern die zugrundeliegenden Parameter stetig sind, lassen sich diese Verfahren auch animieren. Ihre Wirkung kommt dann gegenüber dem Einzelbild zur vollen Entfaltung.

In dem Maße jedoch, wie sich grafische Anwendungen nicht mehr nur Einzelbildern, sondern der Animation bedienen, eröffnet sich auch die Möglichkeit, Gestaltungsmittel aus weiteren Kunstgattungen zu adaptieren. Gemeint ist hier vor allem die Filmkunst, die lange vor der Computergrafik „die Bilder laufen“ ließ und dabei zahlreiche Ausdrucksmittel etablierte, die sich auch außerhalb des Filmes anwenden lassen.

In ihrer ursprünglichen, rein auf Spielfilme bezogenen Bedeutung umfassen filmische Gestaltungsmittel (auch als Filmtechniken bezeichnet) zunächst die Bereiche Kamera, Ton und Licht, aber auch den Aufbau der Szene, das Aussehen und Verhalten der Darsteller, also alle Aspekte eines Films, die neben der Filmstory zum Gesamtwerk Film beitragen (siehe Abbildung 4.1). Die Filmstory, also der Handlungsablauf, kann in schriftlicher Form vorliegen, bei-

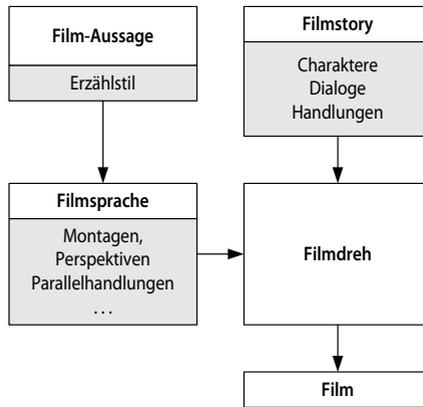


Abbildung 4.1: Das Verhältnis zwischen filmischen Erzähltechniken und Filminhalt in Analogie zur Beziehung zwischen 3D-Modell und Kameramodell aus Abbildung 2.2 auf Seite 24. Eine Neuverfilmung eines bekannten Stoffes ist demnach eine filmsprachliche Neuformulierung des bekannten Inhalts, mit gegebenenfalls einer neuen Aussage, einem anderen Filmstil, anderen filmischen Mitteln (Tricks, etc) oder einer neuen Besetzung.

spielsweise als Roman oder Dokumentation.¹ Die spezifisch filmische Erzählweise aber wird erst durch den Einsatz filmischer Gestaltungsmittel erreicht. Der folgende Abschnitt gibt daher zunächst einen kurzen Überblick über die Benutzung dieser Mittel im Film, bevor dann auf Anwendungen außerhalb des Films eingegangen wird.

4.1 Filmsprache

Die Filmsprache² ist es, die den Film als Medium ausmacht, ihn vom Buch oder dem Theater unterscheidet. Lesen und ganz allgemein Sprachen müssen normalerweise erlernt werden; die Filmsprache bildet hier keine Ausnahme.

¹ Die Vorlage wird nicht direkt benutzt, sondern als *Filmskript* mit Blick auf die Erfordernisse einer filmischen Umsetzung aufbereitet.

² Genaugenommen gibt es mehr als eine Filmsprache, da verschiedene Filmgenres (oder sogar einzelne Filme) ihre eigene Sprache haben. Gemeint ist hier die Sprache des Films in Abgrenzung zu der anderer Medien.

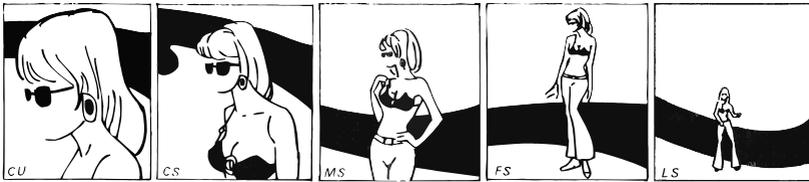


Abbildung 4.2: Arijon unterscheidet fünf Einstellungsgrößen, beginnend mit *closeup*, über *close shot*, *medium shot*, *full shot* bis zum *long shot* [Arijon 1976]. Neben der hier gezeigten Einteilung in fünf Größen gibt es weitere, mit mehr oder weniger Stufen. Ihnen gemeinsam ist die Orientierung am Kopf (bei den Nahaufnahmen), dann am Oberkörper, bis sie in der Totalen die ganze Person mit ihrem Umfeld zeigen.

Da das Medium Film aber mittlerweile ein Jahrhundert alt ist und besonders seit der Verbreitung des Fernsehens fast alle Menschen (nicht nur der westlichen Welt) mit Filmen aufwachsen, kann von der Verständlichkeit der Filmsprache ausgegangen werden.

Die Zeichen der Filmsprache, die Filmtechniken, sind stark ikonisch geprägt, und – anders als natürliche Sprachen – kaum formalisiert und konventionalisiert [WOLLEN 1969]. Das macht sie leichter verständlich als Texte, deren Verständnis ja das Erlernen einer Sprache voraussetzt. Bedeutungen werden nicht nur durch die Handlung selbst transportiert, sondern auch durch die Art ihrer Darstellung im Film [HICKETHIER 1993]. Diese beiden Ebenen sollten aber nicht getrennt betrachtet werden, sondern spielen im Film immer zusammen und bilden ein Bedeutungskontinuum [MONACO 1980]. Für den Filmemacher heißt das: „Die Darstellung muß zum Inhalt passen!“

Filmtechniken sind immer auf die handelnden Charaktere ausgerichtet. In Filmen geht es meist um Menschen und ihre Erlebnisse und Schicksale. Charaktere teilen sich durch Sprache, aber auch Körperhaltung mit sowie durch Mimik und Gestik. Wir lesen in den Gesichtern – besonders den Augen – Emotionen, und dies nicht nur im täglichen Leben, sondern auch im Film. Somit haben Filmtechniken einen starken Menschbezug, was man z. B. gut an der oft gebrauchten Erklärung von Einstellungsgrößen (in Abbildung 4.2) sehen kann. Aber nicht nur Einstellungsgrößen, sondern auch der Bildaufbau (wer ist wo im Bild zu sehen), das Aneinanderfügen verschiedener Einstellun-

gen, die Schnittfolge, die Bewegungen der Kamera und viele andere Elemente der *Grammatik* der Filmsprache richten sich an handelnden Charakteren aus. Beim Versuch, Filmtechniken als zusätzliches Darstellungselement in anderen Medien als Film zu nutzen, stellt sich daher zunächst die Frage nach der Übertragbarkeit der Gestaltungsmittel.

4.2 Filmtechniken außerhalb des Films

Die im Film verwendeten Gestaltungsmittel sind das Produkt eines viele Jahrzehnte andauernden ästhetischen Diskurses, in dessen Verlauf immer wieder neue Ausdrucksmittel erprobt und zum Teil auch wieder verworfen wurden. Es ist nicht zu erwarten, daß sich Gestaltungsmittel aus einem Medium einfach auf ein anderes übertragen lassen. Daher ist zunächst eine Betrachtung der Gemeinsamkeiten und Unterschiede von Filmen und den verschiedenen Spielarten interaktiver Systeme angebracht (siehe Tabelle 4.1 für eine Zusammenfassung).

4.2.1 Interaktive Technische Dokumentationen

In der Anwendungsdomäne, um die es in dieser Arbeit hauptsächlich geht, der technischen Dokumentation, fallen wesentliche Motivationen, die Filmtechniken begründen, weg. Dokumentationen sollen vorrangig Informationen vermitteln. Sie sollten zwar ansprechend aussehen, aber hier hat die Korrektheit immer Vorrang vor der Ästhetik.

Stimmungen und Menschbezug Das Primat der Korrektheit gilt für unterhaltende Elemente. In jedem Fall sind wichtige Filmelemente (Humor, Emotionen, Tragik, Spannung, das komplexe Geflecht zwischenmenschlicher Beziehungen) in einer technischen Dokumentation nicht zu erwarten.

Auch die äußerst vielseitige Ausdruckskraft eines menschlichen Gesichtes und die durch sie transportierten Inhalte und Stimmungen spielen in technischen Dokumentationen keine Rolle. Soweit Menschen überhaupt auftauchen, dann zur Demonstration von Handlungen, Größen-

verhältnissen, usw., aber nicht zur schauspielerischen Darstellung irgendeiner Dramatik.

Das mit vielen Filmtechniken verfolgte Ziel der emotionalen Beteiligung des Zuschauers am Geschehen wird bei technischen Dokumentationen nicht verfolgt.

Erfafbarkeit Betrachter technischer Dokumentationen wollen die präsentierte Information möglichst leicht und ohne lange Suche erfassen können. Kameraperspektiven müssen daher so gewählt werden, daß eine schnelle Orientierung und gute Erkennbarkeit der wichtigen Bildelemente gewährleistet sind. Das trifft ebenso in den meisten filmischen Kameraeinstellungen zu.

Im Gegensatz dazu ist das bewußte Nicht-Sichtbarmachen wichtiger Information – im Film ein beliebtes Mittel, um Spannung zu erzeugen und Interesse zu wecken – in technischen Dokumentationen unüblich.

Narration Viele filmische Erzähltechniken entfalten sich erst über längere Zeit. Technische Animationen dagegen sind im Vergleich zu Spielfilmen recht kurz und erzählen keine Geschichte, so daß im Film benutzte narrative Elemente wie Parallelhandlungen (die durch geschickte Montage noch zusätzlich Spannung erzeugen) oder Rückblenden nicht anwendbar sind.

Gestaltungsfreiraum Die für viele Fragen der Bildkomposition nötige Gestaltungsfreiheit bezüglich der Szene ist in Dokumentationen nicht immer gegeben, da die Szene ja so gezeigt werden soll, wie sie ist, ohne daß wie beim Filmdreh noch Änderungen vorgenommen werden können, wenn etwas „nicht gut aussieht“.

Der einzig verbleibende Freiheitsgrad für die Bildgestaltung ist dann die Positionierung der Kamera.

Ausprobieren Bei automatisch erzeugten technischen Dokumentationen etwa in einem interaktiven System ist außerdem das beim Film übliche Probieren und Experimentieren („Nochmal!“, „Alles auf Anfang!“) unmöglich. Alle Einstellungen werden sofort gezeigt und vom Zuschauer konsumiert; nichts kann zurückgenommen werden. Die Auswahl der Filmtechniken und ihre Ausführung muß nach vorher festgelegten Re-

geln erfolgen und kann zwar während der Entwicklung an Testfällen, aber nicht am konkreten Beispiel geprobt werden.

Ton Bei all den genannten optischen Mitteln soll auch nicht übersehen werden, daß eine weitere wesentliche Komponente von Filmen, die akustische Untermalung der Handlung mittels Filmmusik oder anderen Zusatzeffekten, bei technischen Dokumentationen (abgesehen von Präsentationen) unüblich ist.

Insgesamt wird sich also nur ein Teil des filmischen Erzählrepertoires auf technische Dokumentationen übertragen lassen. Dazu zählt vor allem die Positionierung der Kamera, der eine Hinweisfunktion im Sinne von Hervorhebungen zukommt.

4.2.2 Interaktive Entertainmentsysteme

Unter Interaktiven Entertainmentsystemen (siehe Abschnitt 1.1.2) sollen Grafikanwendungen verstanden werden, die Computergrafiken (meist in Echtzeit) als Reaktion auf Nutzerinteraktionen erzeugen, aber im Unterschied zu den zuvor besprochenen technischen Dokumentationen einen zusätzlichen Unterhaltungsanspruch und – damit verbunden – andere Zielgruppen haben. Neben Computerspielen zählen hierzu auch interaktive Installationen (*location-based entertainment* [FUNGE et al. 1999, TOMLINSON et al. 2000]).

Im Hinblick auf die stilistischen Mittel sind Entertainmentsysteme zwischen der technischen Dokumentation und dem Film einzuordnen. Vom Film sind sie als Unterart der Intelligenten Interaktiven Systeme klar durch ihre Bindung an den Computer und dessen Möglichkeiten als Medium – insbesondere der Interaktivität – abgegrenzt. Die Grenzen zu technischen Dokumentationen jedoch sind fließend: Etwa dort, wo neben dem reinen Unterhaltungsanspruch auch Lehrreiches vermittelt werden soll (*Edutainment*), zeichnen sich z. B. Computerspiele durch einen wesentlich stärkeren Menschbezug aus und geben Raum für viele filmische Gestaltungsmittel, die in eher nüchternen Anwendungen fehl am Platze wären. Viele im Film erzeugte Stimmungen wie Spannung, Humor, etc. kommen auch in Spielen vor. Es gibt eine Handlung, und es tauchen Menschen oder menschenähnliche Figuren auf, die dem Spieler oder Betrachter als Bezugspunkte dienen und bei ihren Aktionen „gefilmt“ werden können.

Weitere Unterschiede zu technischen Systemen ergeben sich aus der veränderten Erwartungshaltung an Spiele: die Benutzer wollen unterhalten werden, die Toleranz gegenüber Schwachstellen und Unzulänglichkeiten in der Präsentation ist geringer. Und nicht zuletzt „arbeiten“ sie länger mit dem System als bei technischen Dokumentationen.

Auch die Handlungsabläufe sind unter Umständen wesentlich schneller (etwa in Kampfszenen), komplexer (bei Parallelhandlungen), und die Schauplätze wechseln unter Umständen öfter. Fokusverlagerungen zwischen verschiedenen Orten treten häufiger auf; die Vertrautheit mit den Schauplätzen muß nicht jedesmal neu hergestellt werden. Fahrten brauchen aber immer etwas Zeit und würden den Spielfluß bei übermäßigem Gebrauch hemmen. Daher sind hier Kamerafahrten als alleiniges Mittel der Kamerapositionierung ungeeignet. Vielmehr müssen desöfteren harte Schnitte eingebaut werden. Die Kamerasteuerung muß daher über geeignete Mechanismen verfügen, um die Entscheidung *Fahrt oder Schnitt* zu treffen.

Der für Filmtechniken entscheidende Unterschied zwischen Dokumentations- bzw. Lernsoftware und Computerspielen ist daher nicht der unterhaltende Charakter, sondern die Existenz bzw. Abwesenheit von (Spiel-)Figuren, mit denen sich der Spieler/Nutzer identifizieren kann. Dies sind bei Spielen einzelne Charaktere oder ganze Familien bis hin zu Völkern. In eher abstrakten Spielen (z.B. Tetris) bieten sich dagegen kaum Ansatzpunkte für filmische Gestaltungsmittel.

Darüberhinaus läßt sich argumentieren, daß die allzustarke Fixierung auf filmische Mittel in interaktiven Systemen letztlich die Entwicklung eigener Erzählmuster hemmt, denn die hervorstechendste Eigenschaft der Spiele, ihre Interaktivität, ist in Filmen gänzlich unbekannt. WILLMANN bemerkt hierzu:

Wenn die Videogames ihre Ambitionen verwirklichen wollen, von ihrer Wirkung, von ihrer Aussagekraft her endlich mit dem Film gleichzuziehen, dann müssen sie sich, trotz aller Parallelen, wieder mehr von dessen Vorbild lösen. So wenig, wie das Kino seine eigene Sprache im doch auf den ersten Blick so eng verwandt scheinenden Theater gefunden hat, so wenig wird das den Bildschirmspielen beim ihnen eben nur scheinbar nahestehenden Film gelingen – der höchstens als thematischer Ideenlieferant geeignet ist. Viel eher

müßten sie sich orientieren an Ausdrucksformen, die mit den selben Parametern arbeiten, die auch einem interaktiven Medium voll zur Verfügung stehen – mit Raum, Zeit, Rhythmus, Spannung und Entspannung, Dichte, Farben, Formen. [WILLMANN 2002]

Diese Haltung stützt die These, daß die Ausdrucksmittel, derer sich Werke eines bestimmten Mediums bedienen, in starkem Maße von den vorhandenen technischen Möglichkeiten dieses Mediums bestimmt werden. Diese aber unterscheiden sich zwischen Film und Computer. Weitere Unterschiede finden sich in inhaltlichen Fragen (siehe Tabelle 4.1) und im künstlerischen Anspruch.

Filmische Mittel können daher nicht unreflektiert in Intelligente Interaktive Systeme übernommen werden, sondern nur Grundlage einer sorgfältigen Adaption sein. Dennoch ist die Beschäftigung mit filmischen Mitteln für Intelligente Interaktive Systeme keine Zeitverschwendung, denn sie hilft, den Adaptionprozeß voranzutreiben. Es ist zu erwarten, daß im Zuge dieser Adaption längerfristig neue, dem neuen Medium Computer angemessene Ausdrucksmittel entstehen, die im Ursprungsmedium nicht in dieser Form realisierbar wären.

4.3 Hervorhebungsmittel im Film

Eine der wesentlichen Neuerungen des Filmes im Vergleich zum Theater ist die Einführung einer Kamera, die sich unabhängig von Zuschauer bewegen kann und wechselnde Perspektiven zuläßt. Gleichzeitig hat der Zuschauer nicht mehr die Möglichkeit, sich seine eigene Perspektive zu suchen. Vielmehr zwingt ihm die Kamera ihre „Sicht der Dinge“ auf. Es ist daher nicht verwunderlich, daß der Gebrauch der Kamera maßgeblich zu neuen Ausdrucksmitteln beigetragen hat. Zahlreiche Filme bieten Beispiele dafür, wie mit Hilfe der Kamera schwer zu erkennende bzw. gemessen an ihrer Bedeutung zu unauffällige Gegenstände oder Menschen hervorgehoben werden.

Als visuelles Medium stehen dem Film zunächst einmal die gleichen Hervorhebungsmittel zur Verfügung, die auch für einzelne Bilder anwendbar sind. Allerdings gilt die Einschränkung, daß – abgesehen von speziellen Effektsequenzen – die Bilder nicht computergeneriert oder gezeichnet, sondern ge-

Tabelle 4.1: Unterschiede zwischen den Genres Film, Interaktiver Technischer Dokumentation und Computerspiel

	Film	Interaktive Dokumentation	Computerspiele
Menschbezug	stark, Menschen sind handelnde Charaktere, sollen den Zuschauer berühren	objektzentriert, Akteure nur als Platzhalter innerhalb von objektbezogenen Aktionen	Menschen bzw. menschenähnliche Charaktere handeln, Bezugsperson ist oft Spielerfigur
Handlung	<i>vorherbestimmt</i> , lang, komplex, fiktional, evtl. mit zahlreichen Nebenhandlungen	<i>interaktiver</i> Diskurs, kommunikative Ziele der Wissensvermittlung	<i>interaktiv</i> zu lösende Spielaufgabe mit Handlung (<i>Story</i>) als motivierendem Hintergrund
Gestaltungsraum	Regisseur mit zahlreichen Einflußmöglichkeiten auf Aktionen von Charakteren und Arrangements von Gegenständen	Handlungen und räumliche Anordnungen vorgegeben	Gestaltung von Räumen vorab durch Leveldesign, Spielerhandlungen interaktiv, nicht vorhersehbar
Zusatzeffekte	Filmmusik zur Verstärkung von Stimmungen	Textannotation, verschiedenste Techniken z.B. aus Abschnitt 3	dito, zusätzlich Musik, meist nicht speziell (dynamisch) auf die Handlung abgestimmt
Probieren	u.U. zahlreiche Iterationen (<i>Proben</i> bzw. <i>Takes</i>) bis zur Zufriedenheit der Crew	kein menschlicher Kameramann „am Set“, daher Entwurf und Test der Planungsmethoden nur vorab an Testscenarien, „Filmen“ der Inhalte in Echtzeit ohne Proben	

filmt werden müssen. Damit fallen selektive Variationen des Zeichenstils oder der Farben, wie sie bei Computergrafiken oder Handzeichnungen möglich sind, weg. Stattdessen lassen sich hauptsächlich folgende Mittel beobachten:

Beleuchtung einzelne Bildteile (z. B. Gesichter) werden angeleuchtet; andere Bildteile sind wegen der Dunkelheit wenig bis gar nicht erkennbar.

Diese stärkste Form der Hervorhebung mit Licht kann niemals willkürlich eingesetzt werden, sondern muß zur Situation passen; die Beleuchtung sollte im Allgemeinen realistisch wirken. Beleuchtungseffekte wirken im allgemeinen eher subtil und sind besser zur Akzentuierung geeignet als zur Hervorhebung.

Schärfentiefe nur der gerade wichtige Bildteil ist scharf erkennbar; davor oder dahinter liegende Ebenen verschwimmen.

Dies verlangt eine Separation der potentiellen Fokuspunkte entlang der Bildtiefe und beeinflußt daher auch die Bildkomposition. Unschärfe ist im Film kaum zu vermeiden und eigentlich nichts Besonderes. Mit der Verlagerung der Bildschärfe zwischen den Fokuspunkten jedoch wird sie als bewußtes Mittel der Hervorhebung erkennbar (siehe Abschnitt 1.1.5 auf Seite 8 und 3.5 auf Seite 54).

Zoom und Kamerabewegungen entfalten ihre Wirkung erst durch den zeitlichen Verlauf. Der näherkommende neue Fokuspunkt wird nach und nach größer und dominanter im Bild und kommuniziert dadurch die Verlagerung des Fokus.

Zooms und Kamerabewegungen unterliegen nicht den Einschränkungen der beiden anderen Verfahren und sind daher leichter in den Film zu integrieren.

Das bekannteste Mittel dürfte der Zoom sein: Zur Vergrößerung eines Objektes wird die Brennweite der Linse verlängert, so daß eine Ausschnittsvergrößerung entsteht und der bislang nur schwer auszumachende Gegenstand sichtbar wird. Das Detail wird auf Kosten der Umgebung, die dabei aus dem Bild wandert, hervorgehoben. Das funktioniert in beide Richtungen: Eine Verkürzung der Brennweite läßt das Detail wieder in den Kontext zurücktreten. Beobachten läßt sich diese Technik z. B. in Kriminalfilmen, wo etwa der Attentäter in seinem Versteck (meist nur für einen kurzen Moment) gezeigt wird, bevor die Handlung fortsetzt.

Das Zoomen kann stufenlos erfolgen, seltener auch als Montage von Einstellungen mit konstanter, aber verschiedener Brennweite. Letzteres verlangt dem Betrachter eine höhere Konzentration ab, um die Detailaufnahme als solche zu erkennen und eignet sich daher besser für eher moderate Brennweitesprünge, oder wie in *DIE VÖGEL* von *HITCHCOCK*, für dramatische Effekte.



Abbildung 4.3: Unterschied zwischen Kamerafahrt (links) und Zoom (rechts). Der gemeinsame Ausgangspunkt der Bild zeigt die Person kaum erkennbar hinter einem Balken. Man erkennt die Stuhllehne, die das auf dem Boden stehende Faß verdeckt. Bereits im zweiten Bild ist die Kamera auf ihrer Fahrt so weit an die Person herangerückt, daß nunmehr das Faß im Vordergrund steht.

Der stufenlose Zoom dagegen läßt auch bei sehr kurzer Dauer den graduellen Übergang miterleben, so daß der Zuschauer immer Totale und Großaufnahme in Relation setzen kann.

Fast die gleiche Wirkung wie beim Zoom erzielt man mit dem Bewegen der Kamera hin zum fraglichen Gegenstand, ohne daß die Brennweite verändert wird. Auch hier wird das Objekt im Ergebnis größer abgebildet, allerdings mit einem kleinen, aber entscheidenden Unterschied: die Position der Kamera hat sich verändert und damit auch die Größenverhältnisse im Bild sowie der Blickwinkel auf Objekte außerhalb der Bildmitte.

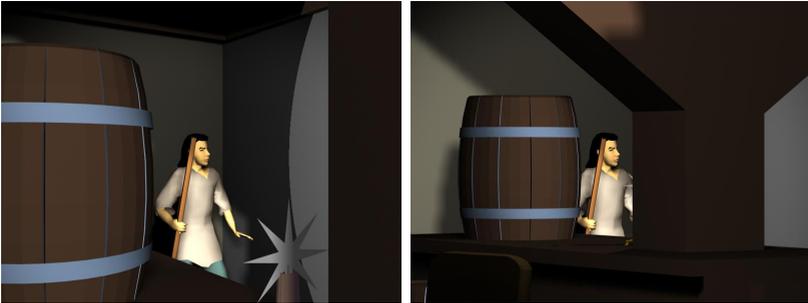


Abbildung 4.4: Im dritten Bild steht die Kamera über dem Tisch, und man erkennt Gegenstände, die hinter dem Balken sind. Trotz des gleichen Abbildungsmaßstabs wie in der Zoom-Version rechts wirkt die Person größer, da sie wegen der veränderten Kameraposition nunmehr fast bis zu den Knien sichtbar ist.



Abbildung 4.5: Die Nahaufnahme zeigt außer dem Kopf nur noch den Knüppel, den die Person hält. Dabei wird deutlich die unveränderte Weitwinkelcharakteristik der linken, zur Figure hin bewegten Kamera sichtbar.

Diese auf den ersten Blick unscheinbare Änderung bewirkt aber während einer Animation eine wesentliche Verbesserung bei der räumlichen Wahrnehmung der Szene (siehe Abbildungen 4.3 bis 4.5 auf Seiten 69–70). Dadurch, daß Objekte in unterschiedlichen Entfernungen verschieden stark von der Kamerabewegung betroffen sind, kann das Gehirn auf ihre Anordnung im Raum schließen (*Bewegungsparallaxe*). Das gelingt sonst nur beim stereoskopischen Sehen, was ja zwei Kameraperspektiven *gleichzeitig* auswertet.

Kamerafahrt und Zoom als Hervorhebungsmittel Kamerafahrten und Zoom haben aber neben der primären Funktion der Detailbetonung oder auch -abschwächung noch einen Nebeneffekt, der sich aus ihrer Interpretation als Element der Filmsprache ergibt: Die so ins Blickfeld gerückten Bildteile werden als wichtig erkannt, auch wenn das anderweitig noch nicht so offensichtlich war.

Damit entspricht das Hervorheben sowohl im Film als auch in einem interaktiven System nicht nur dem Wunsch des Betrachters, irgendetwas besser sehen zu können, sondern *postuliert* ihn geradezu dadurch, daß es ihn gleichzeitig erfüllt. Der Betrachter kann sich, da er ja keine Kontrolle über die Kamera hat, diesem Bild nicht entziehen und schließt intuitiv aus der Bildveränderung auf die darunterliegende Information.

Die Entscheidung darüber, was wann wichtig ist, obliegt im Film dem Regisseur. In interaktiven Systemen ist die Software dafür verantwortlich. Wie im Film wird sich die Wichtigkeit verschiedener Objekte entsprechend der Situation dynamisch (siehe Abschnitt 1.1.1) ändern. Die Software benötigt also eine Möglichkeit, um diese Veränderungen – so sie kamera-relevant sind – zu erkennen und daraus geeignete Aktivitäten zur Kamerasteuerung abzuleiten.

Zoom oder Kamerafahrt In interaktiven Systemen wird zur Hervorhebung von Detail durch Vergrößerung sowohl der Zoom verwendet als auch die Kamera bewegt [BUTZ 1997, KARP und FEINER 1993]. Während aber in Filmproduktionen die Zoomtechnik zu beachtlichen Kosteneinsparungen führt, weil sie den Aufwand für Kamerafahrten (Schienen, Krane, Planung, etc.) verringert und damit trotz der in der Abbildung 4.4 gezeigten Mängel³ eine gewisse Berechtigung hat, ist die Anwendung in interaktiven Systemen weniger zwingend.

Eine virtuelle Kamera kann zwar auch nicht überall, aber auf jeden Fall leichter positioniert werden, kommt sie doch ohne Stative oder andere Vorrichtungen aus. Damit fällt in der Computergrafik das ökonomische Moment

³ Obwohl von Puristen eher abgelehnt [BITOMSKY 1972] ist Zoom bei maßvoller Anwendung ein legitimes gestalterisches Mittel im Film, das zudem wie etwa in HITCHCOCK's Vertigo zur Bereicherung der Filmsprache beigetragen hat.

des Zooms weitestgehend weg. Kamerafahrten müssen aber auch im Computer geplant und implementiert werden, was entsprechende Softwaremodule voraussetzt. Geht es allerdings nur um eine Detailvergrößerung bzw. -verkleinerung (wie sie der Zoom noch leisten könnte), hat man es auch nur mit geraden Fahrten ohne Hindernisse zu tun, so daß sich der Aufwand in Grenzen hält. Der Implementationsaufwand – auch bei komplizierteren Fahrten – ist in jedem Fall einmalig und unabhängig von der Anzahl der durchgeführten Kamerafahrten, während im Film jede Fahrt neuen Aufwand verursacht. Zoomen ist darüber hinaus nur dann möglich, wenn das fragliche Zielobjekt überhaupt von der aktuellen Kameraposition aus sichtbar ist. Bestehen Verdeckungen durch andere Szenenobjekte, ist eine Umgehung der Hindernisse durch eine veränderte Kameraposition unausweichlich (siehe Abschnitt 5).

Ein weiterer Punkt, der für Kamerafahrten spricht, ist die oben erwähnte bessere 3D-Wahrnehmung aufgrund der Bewegungsparallaxe, die dem Betrachter ein Gefühl für die virtuelle Szene gibt [WARE 2000, ROGERS und GRAHAM 1979, COREN et al. 1994]. Da dies neben anderen Anliegen der Sinn einer 3D-Grafik ist, bietet die Kamerafahrt eine willkommene Gelegenheit, die Erkundung der Szene zu unterstützen.

Damit haben zwar sowohl Zoom als auf Kamerafahrt die gleiche Bedeutung zur (De-)Fokussierung von Objekten. Beide arbeiten zudem kontinuierlich, so daß der Betrachter die Übersicht behält. Der Zoom jedoch schränkt die Art der Annäherung bzw. Entfernung auf eine gerade Linie ein, während echte Kamerafahrten auch beliebige andere Bahnen beschreiben können. Zusätzlich wirkt die Bewegungsparallaxe als unterstützendes Element der 3D-Wahrnehmung. Aus diesen Gründen sind interaktiven 3D-Anwendungen Kamerafahrten besser geeignet als Zooms, um Fokusverlagerungen zu untermauern.

4.4 Zusammenfassung

Das letzte Kapitel beschäftigt sich mit grafischen Mitteln zur Hervorhebung von Bildelementen, die bereits bei Einzelbildern wirken. Viele Intelligente Interaktive Systeme zeichnen sich aber durch Echtzeitgrafik aus, können also auch Effekte nutzen, die erst in einer Folge von Einzelbildern wirken. Dies

und die Handhabung einer virtuellen Kamera, die sich frei positionieren lässt, werfen die Frage nach möglichen Gemeinsamkeiten mit dem Medium Film auf.

Neben den vermuteten Gemeinsamkeiten gibt es jedoch auch klare Unterschiede: Die Filmsprache hat ihre Wurzeln im Theater. Ihre Ausformung fand zu großen Teilen in einer anderen kulturgeschichtlichen Periode statt als die des Mediums Computergrafik. Diese ist von völlig anderen technischen Abläufen und Möglichkeiten geprägt. Nicht zuletzt ist die Computerbranche wesentlich stärker technisch dominiert als die eher künstlerisch orientierte Filmwelt.

Im Mittelpunkt dieses Kapitels stand daher die Frage, welche Mittel des Films geeignet sind, Intelligente Interaktive Systeme bei der Inhaltsvermittlung zu unterstützen. Dabei treten neben den in Abschnitt 2.2 beobachteten Gemeinsamkeiten dieser Systeme auch wieder Unterschiede zutage, die eine differenzierte Behandlung technischer und unterhaltender Systeme nahelegen.

Als eines der prägnantesten Mittel des Films kann die Bewegung der Kamera angesehen werden. Änderungen der Perspektive betreffen das gesamte Bild und sind daher als globales Hervorhebungsmittel (siehe Abschnitt 3.1.4) anzusehen. Besonders Bewegungen hin zu fokussierten oder weg von defokussierten Objekten sind unabhängig von der Anwendungsdomäne und bieten sich als dynamisches Hervorhebungsmittel an. Die einfachere zu implementierende Variante, der Zoom, vermittelt weniger Information über die virtuelle Szene, hat aber ihre Berechtigung, wenn der Implementations- oder Planungsaufwand für Kamerafahrten zu hoch ist.

5

Planung von Kamerafahrten

In interaktiven 3D-Systemen ist häufig – etwa beim Einsatz von Hervorhebungsmitteln oder objektbezogenen Interaktionen – die Sichtbarkeit bestimmter Objekte gefordert. Der Einsatz von Transparenz und anderen zuvor beschriebenen Veränderungen der grafischen Darstellung bis hin zum Weglassen störender Objekte löst das Problem aber nicht vollständig. Er versagt sogar vollends, wenn das Objekt zu klein dargestellt wird oder gar nicht im Blickfeld liegt. In diesen Fällen wird eine neue Kameraperspektive benötigt.

In Abschnitt 4.3 wurden Gründe aufgeführt, die für eine graduelle Veränderung der Kameraposition, also eine Fahrt sprechen. Das könnte einfach durch eine lineare Überführung von Ausgangs- in Zielposition (sowie der dazugehörigen Blickrichtungen) geschehen. In einer 3D-Umgebung würde dies jedoch leicht zu „Störeffekten“ führen, die durch ihre ungeschickte Wirkung vom eigentlichen Sinn der Kamerafahrt ablenken würden. Dazu gehören Ruckeln (durch plötzliches Anfahren und Abbremsen), zu schnelle Richtungswechsel und nicht zuletzt Kollisionen der Kamera mit Szenenobjekten. Die Vermeidung dieser Störungen macht die eigentliche Herausforderung bei der Planung von Kamerafahrten aus.

In den hier behandelten Anwendungen kommt hinzu, daß der genaue Zielort der Kamera nicht immer exakt vorgegeben ist; vielmehr dient die Kamerafahrt einem kommunikativen Ziel. So wird oft nur ein zu zeigendes Objekt

spezifiziert und die Kamera muß sich ihren neuen Standpunkt erst suchen, bevor sie die Fahrt planen und antreten kann. Daher zerfällt die Planung einer Kamerafahrt prinzipiell in zwei Teilaufgaben: das Finden einer inhaltlich (also nicht in Form von konkreten Koordinaten) beschriebenen neuen Kameraposition und die anschließende Fahrt dorthin.

5.1 Betrachterposition

Verschiedene Arbeiten widmen sich dem Finden einer *geeigneten* Kameraperspektive [OLIVIER et al. 1999, HALPER und OLIVIER 2000, OLIVIER et al. 1999, BARES und LESTER 1999, BARES et al. 2000, DRUCKER 1994]. Welche Eigenschaften diese haben soll, wird in Form von Anforderungen beschrieben, die das zu erzeugende Bild beschreiben. Dabei gibt es jeweils anwendungsabhängig unterschiedliche, eventuell sogar einander widersprechende Anforderungen an die gesuchte Position bzw. Blickrichtung.

Auf der Anwendungsseite stehen Wünsche wie „Zeige Objekt A aus einer möglichst günstigen Blickrichtung“ oder „Zeige den Kopf von Figur A, während die Kamera hinter dem Kopf von Figur B ist“. Diese visuellen Gestaltungswünsche ergeben sich aus der aktuellen Situation bzw. den Darstellungszielen. Für kurze Animationen in technischen Systemen genügen recht einfache Regeln, wogegen für charakterbasierte Entertainment-Anwendungen komplexere filmsprachliche Idiome sinnvoll sind.

Für die Umsetzung im Rechner müssen diese Wünsche zunächst in eine geeignete mathematische Form gebracht werden, d. h. sie werden in eine Reihe verschiedener *Zielvorgaben* übersetzt. Zielvorgaben (oder auch Nebenbedingungen) konkretisieren die unscharf formulierten Gestaltungsregeln in geometrischer Hinsicht und ermöglichen so die Berechnung einer Kameraposition.

In den Vorgaben wird jeweils eine bestimmte Art der Sicht auf die Objekte gekapselt, die unabhängig von der konkreten Position der Objekte ist. Die letztlich gesuchte Kameraeinstellung muß erst aus den aktuellen Objektdaten (Größe, Position, Orientierung) berechnet werden. Wenn sich die Objekte bewegen, kann die einmal definierte Sichtweise durch wiederholte Neuberechnung immer wieder hergestellt werden. Weil die Zielvorgaben die mögliche

Kameraposition bzw. -orientierung beschränken, nennt man sie auch *constraints*.

Zu den möglichen Zielvorgaben gehören:

Zielobjekte An sich handelt es sich hierbei noch um keine Zielvorgabe, aber das Zielobjekt mit seinen geometrischen Daten (Schwer- oder Mittelpunkt, Größe, Ausrichtung, etc) bildet die Grundlage für die anderen Nebenbedingungen.

Normalerweise sollen Zielobjekte im Bild zu sehen sein, d. h. sie müssen im Sichtfeld der Kamera liegen. Im Falle eines einzelnen Zielobjektes wird die Kamera darauf ausgerichtet.

Blickwinkel Horizontaler bzw. vertikaler Winkel, aus dem das Zielobjekt betrachtet wird.

Diese Vorgabe ermöglicht Konstrukte wie: „Zeige den Kopf von vorn“ oder „Folge der Figur von hinten, halb rechts“, wobei die ungefähren Angaben noch in konkrete Winkel oder Winkelbereiche übersetzt werden müssen. Diese Winkel sind relativ zur Orientierung des Objektes.¹ Änderungen dieser Orientierung, also z. B. Drehungen eines Kopfes, würden die Kamera auf einem Orbit um den Kopf mitschwenken. Diese Vorgabe ist daher wenig geeignet, wenn häufige und schnelle Drehungen des Zielobjektes zu erwarten sind.

Objektgröße im Bild Diese ergibt sich aus realer Objektgröße, Abstand zur Kamera und Kameraöffnungswinkel (Brennweite). Wenn letzterer nicht verändert werden soll, also nicht gezoomt wird, und die Objektgröße konstant ist, legt diese Bedingung den Abstand zur Kamera fest.

Damit entspricht diese Vorgabe der *Einstellungsgröße* (siehe auch Abbildung 4.2 auf Seite 61). Soll die Objektgröße über mehrere Bilder hinweg eingehalten werden, ergibt sich ein neues Problem: Die Objektgröße im Bild zu bestimmen, erfordert ein geeignetes Maß, etwa relative Höhe oder Breite des Objektes im Bild oder auch Größe der bedeckten Bildfläche. Solche Maße sind aber außer vom Abstand der Kamera auch noch

¹ Für die Beispiele sei davon ausgegangen, daß die Objekte so modelliert sind, daß z. B. ein Kopf mit der Nasenspitze in Richtung der positiven Z-Achse zeigt, die Richtung *vorn* also einen Winkel von 90° mit der X-Achse bildet.

von der Form des Objektes in der Projektion, also der Blickrichtung auf das Objekt abhängig. Da letztlich ja der Abstand gesucht ist, soll dieser selbst bei verändertem Betrachtungswinkel eingehalten werden, auch wenn das reine Objektgrößenkriterium einen veränderten Abstand fordern würde. Veränderte Blickwinkel würden sonst zu aus Benutzersicht unmotivierten Vor- und Zurückbewegungen der Kamera führen.

Abstand Ein einzuhaltender Abstand vom Objekt kann entweder euklidisch oder in XYZ angegeben werden. Die Koordinaten wiederum können Weltkoordinaten oder aber an das Objekt gebunden sein (lokale Koordinaten). Letztere drehen sich mit dem Objekt mit.

Diese Vorgabe ist nicht zu verwechseln mit der Objektgröße. Wird diese gleichzeitig mit einem Abstand vorgegeben, ergibt sich ein überbestimmtes System mit einem potentiellen Zielkonflikt.

Der Abstand muß nicht in allen drei Dimensionen gefordert sein, so daß z. B. ein gewünschter vertikaler Abstand noch die Wahl der anderen beiden Dimensionen offen läßt.

Abstandsvorgaben sind sehr direkt, aber dadurch auch weniger ausdrucksstark als die vorhergehenden, da mit absoluten Zahlenwerten gearbeitet wird. Letztlich handelt es sich um eine Konkretisierung von Forderungen, die sich auch mit anderen Vorgaben formulieren ließen. So wäre eine frontale Großaufnahme eines Kopfes (Angabe von Blickwinkel, Einstellungsgröße und Zielobjekt) auch über das Abstandskriterium formulierbar.²

Winkel zur Blickrichtung Der von der Kamera ausgehende Blickrichtungsstrahl wird auf die Bildmitte abgebildet. Objekte, die genau in der Blickrichtung liegen, erscheinen also in der Mitte des Bildes. Für jeden Punkt der Szene kann man einen Winkel zur Blickrichtung (horizontal und vertikal) bestimmen. Ein positiver vertikaler Winkel zur Blickrichtung bedeutet dann, daß das Objekt in der oberen Bildhälfte liegt.

Diese Vorgabe richtet also die Kamera so aus, daß ein Objekt nicht in der Bildmitte sondern, an einer *gewünschten anderen Stelle im Bild* erscheint. Das erlaubt eine Komposition des Bildes im Hinblick auf die

² Wenn man von gegebener realer Kopfgröße und Brennweite ausgeht, ließe sich dieses Beispiel auch direkter über die Forderung: „Abstand zum Objekt in $m = (0 \ 0 \ 1)$ “ ausdrücken.

Anordnung der Zielobjekte. Da die Komposition mit nur einem Objekt wenig Sinn hat, wird man meist mindestens zwei Objekte gleichzeitig anordnen wollen. Hierdurch entstehen schnell Zielkonflikte (siehe Abschnitt 5.1.1).

Winkel zur Verbindungslinie (*line-of-interest*) zweier Objekte Besonders in Filmdialogen wird oft eine gedachte Linie zwischen den beiden Objekten (Köpfen) gezogen. Diese Linie teilt den umliegenden Raum in zwei Hälften. Steht ein Beobachter in der einen Hälfte, erscheinen die Objekte in der entgegengesetzten Reihenfolge (von links nach rechts), wie aus der anderen Hälfte betrachtet.

Etabliertes Prinzip im Film ist, daß sich die Kamera immer auf der gleichen Seite dieser Verbindungslinie aufhält, um in allen Einstellungen die gleiche Reihenfolge der Personen zu zeigen und den Betrachter nicht zu verwirren. Falls unbedingt nötig, sollte ein Überschreiten der Linie gestalterisch wohlbegründet sein und nicht in einer Montage, sondern in einer für den Betrachter sichtbaren Kamerabewegung geschehen.

Sichtbarkeit Zwischen der Kameraposition und den Zielobjekten sollen keine verdeckenden Objekte liegen.

Dieser Bedingung kommt eine Sonderstellung zu. Sind für die Berechnungen zu den anderen Zielvorgaben nur Daten der beteiligten Objekte erforderlich, so müssen hier alle Szenenobjekte einbezogen werden. Gleichzeitig schränkt das Sichtbarkeitskriterium die Wahl der Kameraposition nicht soweit ein, daß es allein schon ausreichend wäre. Es wird daher immer zusätzlich zu anderen Zielvorgaben herangezogen.

Jedes dieser Kriterien führt über entsprechende Berechnungen zu Wunschkoordinaten für die Kamera. Die aufgeführten Beispiele gehen zunächst von *einem* Zielobjekt aus. Jedoch kann es in vielen Situationen wünschenswert sein, mehrere Objekte gleichzeitig zu beobachten. Damit wären weitere, komplexere Vorgabetypen möglich, wie etwa: „Zeige Person B zwischen A und C“. Solche umgangssprachlich leicht formulierten Wünsche stecken allerdings voller unbewußter Zusatzforderungen und werden dadurch schnell zu komplexen Vorgabe-Systemen. Das Beispiel könnte bedeuten: „Zeige A links von B“ und „Zeige B links von C“ und „halte die Kamera auf der gleichen Höhe wie die

A, B und C“. Diese Verknüpfung läßt sich nur dann umsetzen, wenn B tatsächlich zwischen A und C steht und alle (etwa) die gleiche Höhe haben. Die Anwendung sollte solche Vorgaben also nicht blind treffen, sondern nur dann, wenn sie auch eingehalten werden können.

5.1.1 Konflikte

Zum Zeitpunkt, in dem die Anwendung die aus ihrer Sicht günstigste Kameraposition in Form von Wünschen beschreibt, ist nicht sicher, ob die daraus abgeleiteten Zielvorgaben überhaupt miteinander vereinbar sind. Das System muß also eventuell miteinander in Konflikt stehende Vorgaben handhaben.

In Anlehnung an den Sprachgebrauch in der Robotik wird der Lösungsraum auch als *Konfigurationsraum* und die Kombination aus Kameraposition, -blickrichtung und Öffnungswinkel als *Kamerakonfiguration* bezeichnet [LATOMBE 1991]. Der Konfigurationsraum wird also von je drei Dimensionen für Position und Orientierung sowie einer Dimension für den Kameraöffnungswinkel aufgespannt (siehe Abbildung 2.1 auf Seite 23). Generell reduziert sich der Lösungsraum (als Teilgebiet des siebendimensionalen Suchraums) mit jeder Vorgabe.

Werden diese Vorgaben nun für zwei oder mehrere Objekte gewünscht, muß sichergestellt sein, daß die Schnittmenge dieser Teilräume nicht leer ist. Diese Feststellung durch eine geometrische Betrachtung der Teilräume zu treffen, ist aufgrund ihrer hohen Dimensionalität nur eingeschränkt praktikabel. Daher muß diese Frage vorab mittels der Vorgaben und ihrer Parameter beantwortet werden. Für bestimmte Kombinationen der obengenannten Forderungen läßt sich das feststellen: So ist es z. B. nicht möglich, gleichzeitig einen bestimmten Abstand in x-, y- und z-Richtung einzuhalten und zusätzlich noch einen beliebigen Blickwinkel zum Objekt zu gewährleisten.³

Sind Nebenbedingungen für mehrere Objekte einzuhalten, so ergeben sich noch wesentlich mehr Konfliktmöglichkeiten: so ist es z. B. nicht möglich, gleichzeitig zwei Objekte aus vorgegebenen horizontalen Blickwinkeln zu betrachten und dabei auch noch die Objektgröße beider Objekte unabhängig

³ Dieser Fall wird in der Praxis kaum auftreten und soll hier nur den Umstand illustrieren, daß bei der Festlegung von Vorgaben an die Kamera zahlreiche unerwartete Schwierigkeiten auftreten können.

voneinander zu wählen (siehe Abbildung 5.1). In dieser Konstellation ist der vertikale Blickwinkel zunächst nicht festgelegt, was bei entsprechend ungünstigen Konstellationen dazu führen kann, daß sich die Kamera einen weit oben oder unten liegenden Beobachtungspunkt sucht, weil sie anders den rechnerisch geforderten Abstand (siehe Objektgrößenkriterium) zu den Objekten nicht einhalten kann.⁴ Gibt man für beide Objekte auch noch den vertikalen Blickwinkel vor, so muß die Kamera auf dem Schnittpunkt zweier von den jeweiligen Zielobjekten ausgehenden Strahlen sein; existiert dieser Schnittpunkt nicht, liegt ein Konflikt vor.

Theoretisch könnte man Konflikte vermeiden, indem man nur eingeschränkte Zielvorgaben bzw. deren Kombinationen zuläßt, von denen bekannt ist, daß sie sich immer eindeutig auflösen lassen. Das würde jedoch den ursprünglich auf Flexibilität und Ausdrucksstärke ausgerichteten Vorrat der Zielvorgaben wieder stark einschränken und funktioniert zudem allenfalls für die analytisch berechenbaren Vorgaben, nicht jedoch für die Sichtbarkeit. Daher bleibt nur der Weg, einmal erkannte Konflikte weitestmöglich abzuschwächen oder aufzulösen. Dazu gibt es im Wesentlichen drei Möglichkeiten:

Modifikationen in der Szene Durch Verändern, Umordnen bzw. Weglassen von Objekten werden Problemfälle entschärft.

Sofern wie beim Film ausreichende Kontrolle und Gestaltungsfreiheit über die Szene besteht, ist dies die einfachste Variante. Da solche Umordnungen aber nie frei von „Nebenwirkungen“ sind, muß die neue Szene anschließend auf eventuelle Fehler und auf Plausibilität geprüft werden. In interaktiven Systemen müßte dafür eine automatische Evaluierungskomponente geschaffen werden, die je nach Art der Nebenwirkungen recht komplex werden kann.

Anpassen der Vorgaben Die Parameter oder die Bedingung selbst werden so angepaßt, daß sich eine Lösung ergeben kann.

⁴ Die inhaltliche Spezifikation von Bildern über gewünschte Eigenschaften der Kameraperspektive ist nicht perfekt: Dieser Fall demonstriert sehr schön den Einfluß impliziter Vorgaben, die die Vorstellung des Betrachters von der gewünschten Kameraperspektive unbewußt mitbestimmen. Werden solche Anforderungen vergessen, ergeben sich häufig ungewöhnliche und unbeabsichtigte Perspektiven. Dies gilt insbesondere dann, wenn die Vorgaben in Form von Regeln festgelegt werden, die später automatisch umgesetzt werden, ohne daß der derjenige, der sie aufgestellt hat, ihr Versagen beobachten oder verhindern kann.

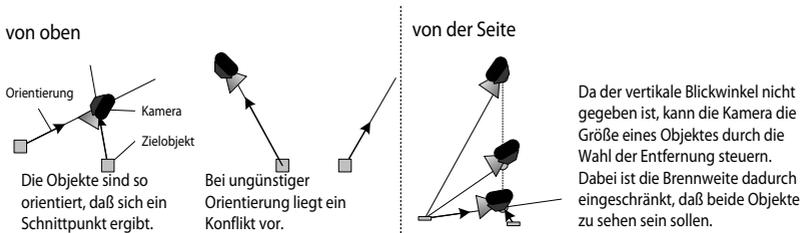


Abbildung 5.1: Kamerapositionierung bei gegebenen horizontalen Blickwinkeln auf die Zielobjekte. Die Nebenbedingung schränkt den möglichen Aufenthaltsort der Kamera auf eine senkrechte Linie ein; in den meisten Fällen wird zusätzlich noch der vertikale Blickwinkel eingeschränkt sein.

Letztlich kommt es nicht auf die genauen Parameterwerte an, sondern das sichtbare Endergebnis. Im Beispiel von Abbildung 5.1 (links unten) wäre dies etwa eine Drehung eines der Objekte, so daß die Strahlen wieder zusammenfallen.

Zulassen von Toleranzbereichen Quasi die Umkehrung der vorherigen Punktes: Die Parameter werden zwar nicht direkt angepaßt, aber sich ergebende Abweichungen werden toleriert, indem erlaubte Schwankungsbreiten vorgegeben werden.

Dies erlaubt zusätzlich eine Bestimmung der *Güte* einer Kamerakonfiguration (z. B. Objektgröße darf zwischen 14 und 20 Prozent des Bildes betragen; 17 Prozent bedeuten eine Güte von 100 Prozent). Existieren mehrere tolerante Nebenbedingungen, müssen die einzelnen Güterwerte – eventuell gewichtet – zu einem Endwert zusammengefaßt werden.

5.1.2 Sonderfall Sichtbarkeit

Wie bereits erwähnt, kommt dem Sichtbarkeitskriterium eine Sonderstellung zu. Falls die Szene nicht – abgesehen von den Zielobjekten – völlig leer ist, wird der häufigste Konflikt im Zusammenhang mit Verdeckungen auftreten. Diese können nur unter Berücksichtigung aller restlichen Szenenobjekte erkannt und behandelt werden. Obwohl die potentiellen Hindernisse bekannt

sind, läßt sich aus ihnen nicht wie bei den anderen Nebenbedingungen durch geschicktes Umstellen entsprechender Formeln eine Lösung berechnen.

Vielmehr muß unter Einbeziehung aller verdeckenden Flächen die Sicht von einer gegebenen Position auf das Zielobjekt geprüft werden. Der damit verbundene potentiell hohe Aufwand unterscheidet das Sichtbarkeitskriterium von allen anderen Nebenbedingungen. Es ist auch nicht möglich, die Sichtbarkeitsanforderungen in die aus den Nebenbedingungen abgeleiteten Formeln einzubeziehen.

Abgesehen von den oben beschriebenen Konflikten wird daher spätestens die Sichtbarkeit Kompromisse bezüglich der Nebenbedingungen fordern, so daß also beim Berechnen der bestmöglichen Kamerakonfiguration nicht davon ausgegangen werden kann, daß diese alle Anforderungen tatsächlich erfüllt. Daher muß ein Maß für die teilweise Erfüllung der Anforderungen gefunden werden, das eine *Güteabschätzung* von Kamerakonfigurationen erlaubt.

Insgesamt ergibt damit sich die Situation eines komplexen Systems aus sich möglicherweise ausschließenden Anforderungen, aus deren mathematischer Formulierung sich kein direkter Lösungsweg (etwa in Form eines Gleichungssystems) ableiten läßt. In solchen Szenarien werden meist Optimierungsverfahren eingesetzt. Auch die Suche nach der besten Perspektive kann als Optimierungsproblem [DRUCKER 1994, OLIVIER et al. 1999] mit einem mehrdimensionalen Lösungsraum aufgefaßt werden, in das die Anforderungen als Nebenbedingungen (*constraints*) eingehen.

Die Suche in einem so hoch dimensionierten Suchraum gestaltet sich naturgemäß sehr rechenintensiv, zumal in diesem Fall auch die Zielfunktion relativ komplex ausfällt: Es muß für die zu testende Kameraperspektive die Einhaltung der Nebenbedingungen geprüft werden, was letztlich heißt, ein Bild aus der fraglichen Perspektive zu erzeugen und entsprechend zu analysieren.

5.1.3 Genetische Algorithmen

OLIVIER et. al. benutzen genetische Programmierung, wobei aus einer initialen, zufälligen ausgewählten Population Kamerakonfigurationen solange gekreuzt und ausgewählt werden, bis sich eine findet, die die gewünschten Nebenbedingungen *gut genug* erfüllt [OLIVIER et al. 1999]. Da bei genetischen

Algorithmen keine analytischen Verfahren zum Einsatz kommen, die den Lösungsraum basierend auf den geforderten Nebenbedingungen vorab eingrenzen, kann eine große Auswahl von Zielvorgabentypen unterstützt werden, ohne daß für jede Kombination von Vorgaben auch ein Rechenweg existieren muß, der zu einer Menge von Kamerakonfigurationen führt.⁵

Die Vielfalt der unterstützten Typen von Nebenbedingungen erlaubt große Flexibilität bei der Beschreibung der gesuchten Kameraperspektive, verlangt aber bei der Auswertung der generierten Bilder zur Bestimmung des Gütemaßes höheren Aufwand. Die Zeit, die zur Planung benötigt wird, hängt zwar von der Szenenkomplexität sowie Art und Anzahl der Zielvorgaben ab, läßt sich aber – der Natur genetischer Algorithmen entsprechend – nicht vorhersagen oder auch nur abschätzen. Daher muß bei der Anwendung eine zeitliche Obergrenze gesetzt werden, nach deren Ablauf die Suche abgebrochen werden kann.

Durch das große Reservoir spezifizierbarer Typen von Zielvorgaben sind die auf Optimierungsmethoden basierten Verfahren sehr flexibel und ausdrucksstark. Durch den Verzicht auf schnelle, analytische Lösungsverfahren ist allerdings derzeit kein Echtzeiteinsatz möglich.

5.1.4 Bildbasierte Methoden

Wird das aktuelle oder das von einer potentiellen Kamerakonfiguration aus erzeugte Bild untersucht, um daraus Hinweise auf die gesuchte Kamerakonfiguration abzuleiten, spricht man von *bildbasierten* Methoden⁶ [MARCHAND und COURTNEY 2000, HALPER et al. 2001]. Bei der Verfolgung von Zielen mittels an Robotern angebrachter Kameras steht immer nur das aktuelle Bild (bzw. ältere Bilder) zur Verfügung. In der Computergrafik dagegen lassen sich auch

⁵ Eine Eingrenzung des nach genetischen Verfahren abgesuchten Lösungsraumes ist zwar in einfachen Fällen möglich und sinnvoll, widerspricht aber eigentlich dem Sinn der genetischen Suche, zumal dann die Kompromisse hinsichtlich der Gütefunktion anders ausfallen.

⁶ Genaugenommen ist natürlich auch der oben angeführte genetische Ansatz bildbasiert, werden doch die zu testenden Individuen danach bewertet, wie gut sie die Kriterien einhalten, indem ein Bild erzeugt und analysiert wird. Diese Analyse wird jedoch nicht dazu benutzt, die Suche gezielter zu steuern oder gar eine Lösung der Problems abzuleiten; es wird lediglich geprüft, ob eine Lösung vorliegt.

„Hypothesen“ testen, d. h. es können Kamerakonfigurationen ausprobiert werden, indem das entstehende Bild auf die Erfüllung der Vorgaben getestet wird (diese Methode wird auch bei den genetischen Verfahren benutzt). Für Echtzeitanwendungen ist das aber nur praktikabel, wenn es möglich ist, die Szene schnell in einen von der Grafikhardware bereitgestellten unsichtbaren Puffer zu zeichnen und diesen ebenso schnell auszulesen und zu analysieren.

MARCHAND UND COURTY beschreiben ein Verfahren, das aus einer Definition von Sichtzielen eine Transformationsmatrix für die Kamera ableitet. Die Sichtziele geben die gewünschte Lage im Bild von bestimmten Raumpunkten an. Die Einhaltung dieser Ziele bindet dabei Freiheitsgrade der Kamera. Sofern die Sichtziele die Kamerakonfiguration nicht eindeutig festlegen, können die noch offenen Freiheitsgrade für die Vermeidung eventueller Verdeckungen ausgenutzt werden [MARCHAND und COURTY 2000].

HALPER UND HELBING stellten eine bildbasierte Methode (siehe Anhang B) vor, bei der unter Ausnutzung moderner Grafikhardware eine schnelle Bestimmung von unverdeckten Kamerapositionen (*potential visibility regions*) im Raum möglich ist [HALPER et al. 2001]. Dieses Verfahren benötigt als Eingabe die zu überprüfende Kameraposition und liefert die nächstliegende Position, die keine Verdeckungen des Ziels mehr aufweist. Die Methode läßt sich auf mehrere Ziele erweitern, wobei die Szene je Ziel einmal gezeichnet werden muß.

Die zu Beginn dieses Abschnitts aufgeführte Auswahl inhaltlicher Kriterien für eine Kameraperspektive ermöglicht interaktiven Systemen eine intelligente, adaptive Perspektivwahl, die der aktuellen Situation gerecht wird. Damit bereichert sie die Palette der Ausdrucksmittel einer virtuellen Kamera in ähnlicher Weise wie die Unterstützung und Kombination verschiedener Zeichenstile.

5.2 Reaktive Kameraplanung

Häufig tritt das Problem, eine neue Kameraposition zu bestimmen, in einer abgewandelten Form auf: Die Kamera ist an einer Position, die einen „guten“ Blick auf das Zielobjekt gestattet, d. h. sie erfüllt die jeweils gewünschten

Nebenbedingungen. Nun bewegt sich aber das Zielobjekt, so daß eine Neu-
auswertung der Nebenbedingungen nötig wird. Diese liefert im Allgemeinen
eine andere Kamerakonfiguration als die aktuelle. Wenn, wie üblich, Sichtbar-
keit eine der Nebenbedingungen ist, also die Kamera in der alten und der neu-
en Konfiguration im freien Raum ist, und unter der Annahme, daß sich das
Zielobjekt nicht viel bewegt hat, und deswegen auch die Kamerakonfigurati-
on nur wenig geändert werden muß, könnte man auch *ohne eine Kamerafahrt*
zu planen, einfach auf die neue Konfiguration wechseln.

Diesen Fall bezeichnen wir in Anlehnung an die reaktive Planung von Mul-
timediapräsentationen als *Reaktive Kameraplanung* [HALPER et al. 2001]. Re-
aktive Verfahren versuchen ausdrücklich, die Kamerakonfiguration der sich
verändernden Szene anzupassen und zielen daher besonders auf dynamische
Szenarien, bei denen die Kamerasteuerung getrennt von der Steuerung der
Szene erfolgt.⁷ Diese getrennte Steuerung basiert allein auf der Szenengeome-
trie und muß ohne gezielte Hinweise auskommen, die etwa eine Interaktions-
steuerung oder ein spezielles Szenendesign liefern könnte.⁸

Kohärenz Eine angenehme, sanfte Kamerabewegung setzt die zeitliche Ko-
härenz der Kamerakonfigurationen voraus. Im Film ist diese Kohärenz ein
wesentlicher Beitrag zum übergeordneten Ziel der Kontinuität. Letztere er-
streckt sich noch auf viele andere Punkte, so etwa die Übergänge zwischen Ein-
stellungen durch diverse Blendeneffekte oder die Gestaltung von zusammen-
gehörenden Einstellungen (Positionierung von Akteuren, ihre Bewegungs-
richtung, ihr gleichbleibendes Aussehen, speziell an unterschiedlichen Dreh-
tagen, etc).⁹ Aus Sicht des Betrachters, der aus dem Film einen hohen Grad
an Kontinuität gewohnt ist, ist Kohärenz wichtiger als eine um jeden Preis
erzwungene Einhaltung von Zielvorgaben an die Kamera.

⁷ Dies ist ein gravierender Unterschied insbesondere zum Kinofilm, bei dem die Handlung
(Schauspieler) und die Kameraführung (Kameramann) von einem Regisseur koordiniert wer-
den.

⁸ Die Szenendaten könnten mit bevorzugten Kamerapositionen angereichert sein; die mög-
lichen Veränderungen der Szene könnten vorab bekannt sein und damit auch die entsprechen-
den Reaktionen der Kamerasteuerung, etc. Diese Maßnahmen erhöhen natürlich den Auf-
wand beim Erstellen der Szene bzw. der gesamten Anwendung beträchtlich, ohne jedoch die
Flexibilität einer echten intelligenten Kamerasteuerung zu bieten.

⁹ Besonders in größeren Filmproduktionen wird die für *Continuity* verantwortliche Person ei-
gens im Abspann erwähnt.

Reaktive Verfahren machen sich die Kohärenz der Bewegung des Zielobjekts zunutze und berechnen, basierend auf Zielvorgaben wie den oben genannten, für jedes Bild eine neue Kamerakonfiguration. Dabei tritt das Hauptproblem der reaktiven Kamerasteuerung zutage: die Annahme, daß eine kohärente Bewegung des Ziels auch automatisch zu einer kohärenten Bewegung der Kamera führt. Diese Voraussetzung trifft zwar in den allermeisten Fällen zu, aber leider nicht immer.¹⁰ Der wichtigste Grund dafür sind Verdeckungen: wenn das Zielobjekt bei seiner Bewegung hinter einem anderen Objekt verschwindet, so muß die Kamera, um den Sichtbarkeitsanforderungen zu genügen, von ihrer bisherigen Bewegungsbahn abweichen, was sich visuell in einem Sprung äußert (siehe Bild 5.2). Hier tritt also trotz einer kohärenten Bewegung des Ziels eine Diskontinuität der Kamerabewegung auf. Da sich der Betrachter nach jedem Sprung der Kamera neu im Bild orientieren muß, steigt die kognitive Last beim Nutzer. Da die Sprünge außerdem unmotiviert erscheinen (das Ziel bewegt sich ja kohärent), werden sie schnell als störend empfunden [BARES und LESTER 1999].

Ein weiteres Problem besteht darin, daß man oft gar nicht davon ausgehen kann, daß sich das Zielobjekt kohärent bewegt. Beginnt dieses nämlich, sich in inkohärenter Weise zu bewegen, z. B. durch schnelle Hin- und Herbewegungen, dann wird die Kamera starr nachgeführt, was zu einem unruhigen Bild führt.

Für eine auch unter diesen Umständen gewünschte Kohärenz der Kamera müssen daher besondere Vorkehrungen getroffen werden, die letztlich dazu führen, daß die Zielvorgaben zumindest vorübergehend verletzt werden können (siehe Bild 5.3 auf Seite 89).

Reaktive Verfahren eignen sich am besten für Anwendungen, in denen die Kamera die ganze Zeit unter Kontrolle der Software ist und keine unmittelbare Steuerung durch den Benutzer möglich ist. Sobald die Kamera phasenweise interaktiv bewegt werden kann, ist die Kohärenz nicht mehr gegeben: Die Annahme eines nur geringen Unterschieds zwischen alter und neuer Kamerakonfiguration wird ungültig, d. h. es muß ein geeigneter Übergang zwischen inter-

¹⁰ Angenommen, sie trifft in 99 Prozent der Fälle zu, dann bedeutet das, daß sie bei schon 20 Bildern je Sekunde alle fünf Sekunden versagt. Unter diesen Umständen ist es auch bei einer 99,9-prozentigen Erfolgsquote noch sinnvoll, nach Verbesserungen zu streben.

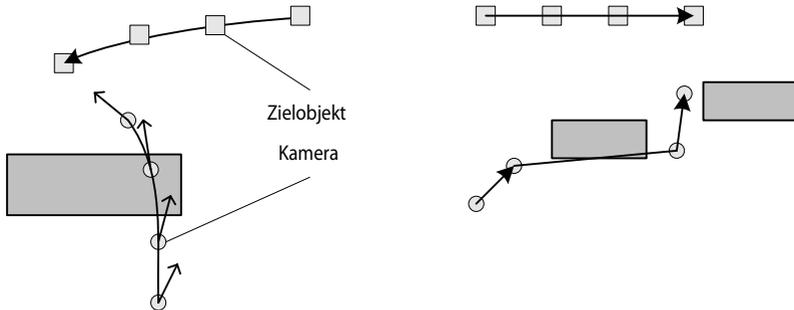


Abbildung 5.2: Links: Ohne Kollisionstests läuft eine reaktive Kameraplanung Gefahr, die Kamera in ein Hindernis zu schicken. Anstelle eines Kollisionstests würde auch eine Sichtbarkeitsprüfung verhindern, daß die Kamera in Phase 2 – also noch vor der eigentlichen Kollision – den Sichtkontakt zum Zielobjekt verliert.

Rechts: Die Sichtbarkeit ist gewährleistet, aber fehlende zeitliche Kohärenz führt in Phase 3 zu einem plötzlichen Sprung an eine Stelle hinter das Hindernis. Bereits hier kann ein eventuell als Zielvorgabe geforderter Kamerablickwinkel nicht eingehalten werden.

aktiver und automatischer Steuerung erfolgen, entweder durch einen Sprung, eine Überblendtechnik oder eine Kamerafahrt.

Da bei reaktiver Kameraplanung die Kamerakonfiguration laufend neu berechnet und evaluiert wird, entsteht der Eindruck einer Kamerafahrt, auch wenn keine Fahrt im eigentlichen Sinne geplant wird. Die angesprochenen Probleme beim Einsatz einer rein reaktiven Steuerung – Inkohärenz, Verdeckungen und Kollisionen – lassen sich durch geschickte Verbindung mit anderen Verfahren (Trägheitssteuerung, Vorausschau, etc) zumindest lindern.

5.3 Kamerafahrten

Die Planung der eigentlichen Kamerafahrt geht von einer gegebenen Start- und Zielposition aus. Als Startposition kann man in interaktiven Systemen meist die derzeitige Kameraposition annehmen. Wege zur Ermittlung einer geeigneten Zielposition wurden im vorhergehenden Abschnitt besprochen.

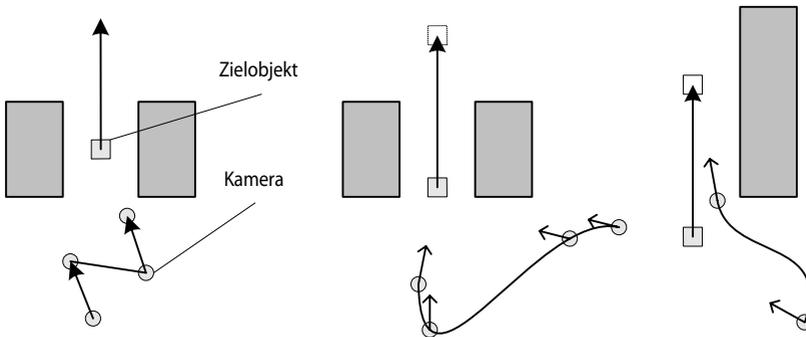


Abbildung 5.3: Links: Ohne Vorausschau kann die reaktive Planung nicht erkennen, daß der aktuelle Weg der Kamera in ein Hindernis führt. Unter Beibehaltung von Kohärenz und Sichtbarkeit wären daher Korrekturen erforderlich, bevor die Sicht blockiert wird oder plötzliche Sprünge erzwungen werden.

Mitte und Rechts: Bereits in den ersten Phasen ist durch die Vorausschau (weißes Quadrat) bekannt, daß der aktuelle Pfad der Kamera zukünftig korrigiert werden muß. Deshalb wird die Kamera unter Vorwegnahme der Hindernisse rechtzeitig abgelenkt. Dabei muß wie schon in Bild 5.2 gegebenenfalls eine Verletzung der ursprünglichen Zielvorgaben (etwa des Blickwinkels) toleriert werden.

Das hauptsächlich zu lösende Problem¹¹ bei der Bewegung vom Start zum Ziel ist dann die Berechnung eines kollisionsfreien Pfades durch die Szene.

In der Robotik wird dieses Problem in seinen vielen Facetten etwa seit den 80er Jahren behandelt. So gibt LATOMBE in seinem Standardwerk zur Planung von Roboterbewegungen [LATOMBE 1991] verschiedene Klassen von Problemstellungen und mögliche Lösungsansätze an.

5.3.1 Roadmap-Methoden

In Analogie zu einem Straßennetz wird der gesamte befahrbare Raum in einem Vorverarbeitungsschritt als Netz von Wegen organisiert. Dabei wird zu-

¹¹ Je nach Art des ursprünglich berechneten Pfades und seiner Verwendung bei der Kamerafahrt kommen dann noch verschiedene Zusatzprobleme hinzu, wie etwa die Berechnung von Blickrichtungen.

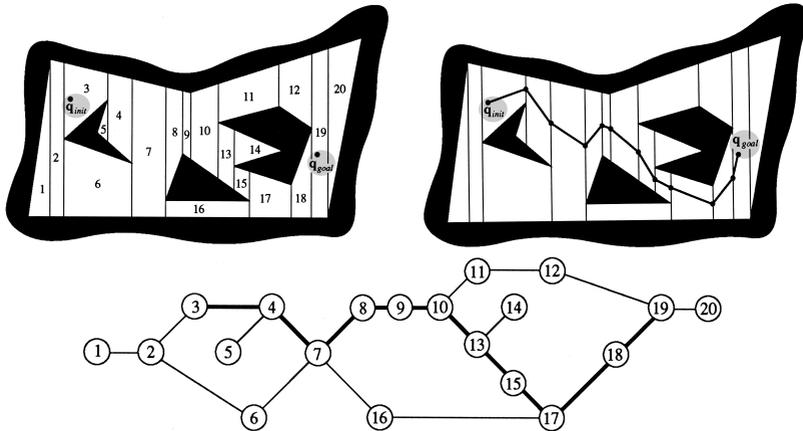


Abbildung 5.4: Roadmap-Methode und Zellzerlegung nach [Latombe 1991].

Oben links: Die derzeitige und die Zielposition sind mit q_{init} und q_{goal} angegeben. Der polygonale freie Raum (begrenzt durch die Umrandung und drei Hindernisse) wird zunächst an jeder Ecke in vertikale Streifen zerlegt (Zellzerlegung). Dann können die Start- und die Zielzelle bestimmt werden.

Unten: Die Roadmap repräsentiert die Verbindungen zwischen den Zellen. In diesem Graph wird nach einem Weg von der Start- zur Zielzelle gesucht (dicke Linie).

Oben rechts: Eine Möglichkeit, den Weg zu befahren, ist, von Zelle zu Zelle zu gehen und die Wegpunkte auf die Mitte der Grenze zweier Zellen zu legen.

erst der freie Raum in Zellen unterteilt. Anschließend werden benachbarte Zellen durch Wege verbunden. Abbildung 5.4 veranschaulicht dies mit einer *visibility roadmap*; alternativ dazu werden auch Voronoi-Diagramme benutzt, die auf eine „Skelettierung des freien Raumes“ hinauslaufen. Im dreidimensionalen Raum und bei komplexer Geometrie der Hindernisse stoßen Roadmap-Verfahren allerdings an Grenzen, da die Anzahl der Eckpunkte und damit die Größe und der Aufwand zum Aufbau und Durchsuchen der Roadmap rasant ansteigt.

5.3.2 Potentialfelder

Eine andere, ebenfalls der Robotik entlehnte Planungsmethode bedient sich sogenannter Potentialfelder. Dabei wird jedes Hindernis mit einem den Roboter (bzw. die Kamera) abstoßenden Feld umgeben, wobei die abstoßende Kraft ihr Maximum am Hindernis selbst hat (und dadurch seine Durchdringung verhindert) und sich dann in den umgebenden Raum hin bis auf 0 abschwächt. Gleichzeitig wirkt auf die Kamera eine anziehende Kraft, die vom gewünschten Zielort ausgeht. Damit steht zu jedem Zeitpunkt eine Bewegungsrichtung der Kamera durch Aufaddieren der Kräfte fest.

Theoretisch müßte, um die an einem bestimmten Ort wirkenden Kräfte zu bestimmen, der genaue Abstand zu jedem möglichen Hindernis bestimmt werden, was bei vielen polygonalen Objekten recht aufwendig werden kann. In der Praxis wird dieser Abstand daher näherungsweise basierend auf einer vorangehenden Diskretisierung des Raumes bestimmt, d. h. für jeden Punkt in einem dreidimensionalen Raster wird festgestellt, ob er im freien Raum liegt oder nicht und gegebenenfalls ein entsprechender Kraftvektor ermittelt. Die Genauigkeit dieser Diskretisierung bestimmt die Präzision der Kameraplanung. Ändert ein Objekt seine Position, Orientierung, etc., so müssen die betreffenden Stellen im Raster aktualisiert werden.

Problematische Aspekte bei Potentialfeldern liegen neben dem erwähnten Kompromiß zwischen genauer Diskretisierung und möglichst wenigen Rasterzellen in der Existenz lokaler Minima.¹² In Anwendungen, in denen dieses Problem entweder ganz vermieden wird oder nur geringe Bedeutung hat, können Potentialfelder eine effiziente Planungsmethode bieten. Beispielsweise führen BECKHAUS et al. in ihrem CUBICALPATH-System zu diesem Zweck spezielle „Navigationsobjekte“ ein (siehe Abbildung 5.5), die die Kamera aus solchen „schwierigen“ Regionen hinausleiten [BECKHAUS et al. 2001, BECKHAUS et al. 2000, BECKHAUS 2002].

¹² Der Begriff stammt aus der Optimierung und bezeichnet Orte, an denen zwar die Zielfunktion (hier die Summe der Kraftvektoren) minimiert wird, die aber nicht die gesuchte Lösung – also das globale Minimum – sind. Bei Potentialfeldern z. B. treten besonders in verwinkelten Szenen leicht Orte auf, an denen sich die Kraftvektoren aufheben, wie etwa in Abbildung 5.5 (links).

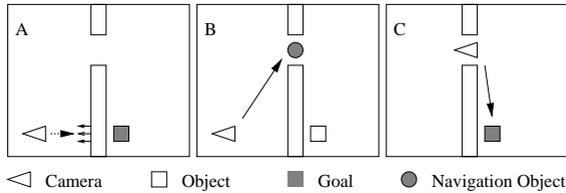


Abbildung 5.5: Navigationsobjekte zur Unterstützung der Wegberechnung bei Potentialfeldplanern (aus [Beckhaus et al. 2001]). In der Szene nicht sichtbare Hilfsobjekte leiten die Kamera aus „Sackgassen“ heraus bzw. vermeiden, daß sie dort hingerät. Im Falle von Türen oder anderen engen Durchgängen kann so die Suche zumindest stark beschleunigt werden. Die Objekte lassen sich gegebenenfalls automatisch erzeugen.

5.3.3 Gemeinsamkeiten und Unterschiede von Robotik und Kameraplanung

Neben vielen Gemeinsamkeiten zur Planung von Roboterbewegungen gibt es natürlich auch Unterschiede zwischen beiden Anwendungen, die sich auf die Auswahl und Eignung der möglichen Planungsverfahren auswirken. Diese werden im Folgenden skizziert:

Konfigurationsraum Hierunter versteht man in der Robotik die Menge aller Positionen und Orientierungen des Roboters. Diese rein geometrischen Daten sind vergleichbar mit den Basisdaten des erweiterten Kameramodells. Teile des Konfigurationsraumes werden von Hindernissen belegt. Den übrigen, zugänglichen Teil des Raumes bezeichnet man als freien Raum (*free space*). Im einfachsten Fall eines kreisförmigen Roboters in der Ebene hätte dieser Raum die Dimensionalität zwei, d. h. jeder Ort kann durch Angabe eines Koordinatenpaares beschrieben werden. Die weitaus meisten Kamerafahrten im dreidimensionalen Raum dürften mit einer Dimensionalität von fünf auskommen, d. h. drei im Raum und zwei für den Blickwinkel der Kamera. In diesem Falle würden Brennweite und eine eventuelle Verdrehung der Kamera um die Sichtachse vernachlässigt bleiben. Für das eigentliche Hauptproblem, die Kollisionsvermeidung, genügt es sogar, nur die Position zu betrachten.

Es kann für alle bekannten Klassen von Suchalgorithmen gezeigt werden, daß deren Komplexität sowohl für Rechenzeit als auch für Speicherbedarf exponentiell zur Dimensionalität des betrachteten Konfigurationsraumes ist [LATOMBE 1991]. Die meisten beschriebenen Robotikverfahren beschränken sich auf die Ebene, was zwar legitim ist, solange die Roboter nicht „fliegen“, aber auch die Komplexität der Bewegungsplanung relativ gering hält. Insbesondere der Übergang von der Planung in der zweidimensionalen Ebene zum dreidimensionalen Raum würde bedeuten, daß die Verfahren nicht mehr in Echtzeit arbeiten können.

Form Modelliert man einen Roboter entsprechend seiner tatsächlichen Form und Ausdehnung, muß man berücksichtigen, daß er im Unterschied zur punktförmigen Kamera nicht an jeder Stelle im Raum positioniert werden kann und seine Positionierbarkeit an einem bestimmten Ort von seiner augenblicklichen Orientierung abhängt. Das betrifft nicht nur die Frage, *wie* sondern auch, *ob* ein bestimmter Punkt überhaupt erreichbar ist. Zudem bedeutet eine Kollision in der Robotik nicht nur eine kurzzeitige visuelle Ungeschicklichkeit, sondern schlimmstenfalls die Zerstörung der kollidierenden Gegenstände. Dieser Umstand erfordert in der Roboterplanung aufwendige Vorkehrungen, die bei der Kameraplanung entfallen können.

Ästhetik Dieser Punkt spielt in der Robotik kaum eine Rolle, sollte aber bei Kamerafahrten beachtet werden. Hervorzuheben sind hier unschöne Diskontinuitäten (Ruckeln, Springen) der Kamerabewegung. In der Robotik werden diese, sofern sie überhaupt von Bedeutung sind, bereits durch die Charakteristik der Antriebsmechanik (Trägheitseffekte, minimale Lenkradien, etc) verringert. Für Kamerafahrten sind dagegen besondere Vorkehrungen nötig.

Unsicherheit und Ungenauigkeit Bewegt sich ein Roboter in einer dynamischen Umgebung, so kann nicht garantiert werden, daß einmal geplante Fahrtstrecken auch tatsächlich frei bleiben. In solchen Fällen werden aufmontierte Kameras bzw. andere Sensoren benutzt, die die Umgebung überwachen und deren Beobachtungen in die Fahrtp lanung einfließen. Weiterhin hängt die Genauigkeit der Messungen von der Entfernung vom Roboter ab.

Kamerafahrten können zwar auch in dynamischen Szenen stattfinden, jedoch unterliegt selbst diese Dynamik der Kontrolle des Systems und ist vorher bekannt. Trotzdem birgt dieser Punkt gewisse Gemeinsamkeiten zwischen Kameraplanung und Robotik und führt in beiden Fällen zu reaktiven Systemen.

Zusammengefaßt vereinfacht die Annahme einer punktförmigen Kamera die Kameraplanung im Vergleich zur Roboterplanung; der Übergang von 2D nach 3D aber verkompliziert sie. Insgesamt führen die genannten Unterschiede zu anderen Prioritäten und Kompromissen und somit zu anderen Herangehensweisen an die Planung von Kamerafahrten, als sie in der Robotik üblich sind.

5.3.4 Kameraplanung in interaktiven Systemen

Einige intelligente interaktive Systeme planen neben anderen Präsentationstechniken auch Kamerabewegungen. Meist wird dabei jedoch der wichtige Aspekt der kollisionsfreien Planung von Kamerafahrten ausgeklammert. So beschreibt TOMLINSON ein System, bei dem eventuelle Sichtbarkeitsprobleme einfach dadurch gelöst werden, daß die Kamera etwas nach oben rückt [TOMLINSON et al. 2000]. Das vereinfacht die Kameraplanung erheblich, wobei deren Schwerpunkt in diesem System auch nicht auf der größtmöglicher Robustheit im Hinblick auf Hindernisse liegt. Der Ansatz ist sogar völlig ausreichend, wenn – wie in diesem Fall – die Szenen auf größtenteils freien Ebenen spielen und kaum Verdeckungen auftauchen.

BARES dokumentiert mehrere Systeme, die sich der 3D-Kameraplanung inklusive Kamerabewegungen widmen [BARES et al. 2000, BARES und LESTER 1999]. Auch hier werden keine Strategien zur Vermeidung von Kollisionen bei der Kamerafahrt beschrieben. Die Sichtbarkeitstests der Zielobjekte beziehen sich ausschließlich auf die Endzustände der Kamerafahrt. Zudem finden Fahrten überhaupt nur bei sehr kurzen Distanzen zum Ziel statt, was den Verzicht auf Kollisionsprüfungen erlaubt. Längere Fahrten werden durch Schnitte ersetzt. Diese sind ebenfalls aus dem Film bekannt und akzeptiert. Zudem sparen sie im Vergleich zu Kamerabewegungen viel Zeit, besonders, wenn die Szene als bekannt vorausgesetzt werden kann und keine Orientierungsproble-

me auftreten. Für komplexere Welten sind automatische Fahrten jedoch eine sinnvolle Orientierungshilfe.

5.4 Umweg: Ein neues Verfahren zur Planung von Kamerafahrten

Im Vordergrund des neu entwickelten Verfahrens stand das Ziel, möglichst schnell *ansprechende* Kamerafahrten zu erzeugen, die einerseits möglichst geringe Wartezeiten im interaktiven Betrieb erfordern, aber trotzdem frei von ungewollten Artefakten wie Kollisionen oder Ruckeln sind [HELBING et al. 1998].

Das Erfordernis der Interaktivität mit angestrebten Reaktionszeiten von unter einer Sekunde zieht selbstverständlich Kompromisse nach sich; so ist in keiner Weise garantiert, daß die geplanten Kamerafahrten ein wie auch immer definiertes Gütemaß erfüllen oder maximieren.

Entwurf und Implementierung des Verfahrens erfolgten im Umfeld eines interaktiven technischen Dokumentationssystems im Sinne von Abschnitt 4.2.1 (für eine Vorstellung des Systems siehe Abschnitt 6.1).

5.4.1 Anforderungen an Kamerafahrten in interaktiven technischen Dokumentationen

In interaktiven Systemen ist nicht nur die für die Planung zur Verfügung stehende Zeit sehr kurz;¹³ auch die Kamerafahrt selbst liegt im Bereich von wenigen Sekunden; meist soll nur ein Übergang von einem betrachteten Objekt zu einem anderen signalisiert werden. Diese Zeit reicht im allgemeinen aus, um dem Betrachter die Möglichkeit zu geben, die veränderte Situation zu erkennen, sich daran anzupassen und insgesamt den Überblick zu behalten.

¹³ Die genaue Definition von *kurz* liegt hier wie so oft im Auge des Betrachters. Wichtig ist, daß die Kameraplanung den Ablauf der Interaktion nicht stört.

Im Unterschied zu Spielen oder Installationen wie in [TOMLINSON et al. 2000]¹⁴ beschrieben kann bei Dokumentationen aber nicht davon ausgegangen werden, daß die Gestaltung der Szene oder ihre datentechnische Aufbereitung im Hinblick auf etwaige Bedürfnisse des Dokumentationssystems erfolgt oder daß Schwächen oder Fehler des Systems (wie etwa des Planungsverfahrens für Kamerafahrten) bekannt sind oder berücksichtigt werden könnten. Vielmehr liegen der Szenengestaltung exakt die Daten des dokumentierten Gegenstandes zugrunde (etwa aus dem Planungsprozeß einer technischen Anlage oder Konstruktionsunterlagen); ebendiese sollen ja durch das System auch dokumentiert werden.

Insgesamt bedeutet das, daß für dieses Verfahren keinerlei Hilfestellung etwa durch speziell angepaßte Szenendaten mit Zusatzinformationen für die Kameraplanung¹⁵ oder besonders planungsfreundliche Szenen¹⁶ vorausgesetzt werden kann. Es müssen daher Vorkehrungen getroffen werden, die auch für den Fall, daß sich keine „schöne“ Fahrt planen läßt, oder diese Planung unvertretbar lange dauern würde, noch ein annehmbares Ergebnis liefern.

5.4.2 Umfahren von Hindernissen durch Ablenkung

Seine Geschwindigkeit erreicht das in der vorliegenden Arbeit entwickelte Verfahren dadurch, daß es ganz allgemein vermeidet, Berechnungen durchzuführen, deren Ergebnis nicht unbedingt gebraucht wird. So wird zunächst versucht, auf direktem Wege vom Start zum Ziel zu gelangen; erst wenn das wegen auf dem Weg befindlicher Hindernisse nicht möglich ist, werden Ausweichmanöver geplant. Dieser Ansatz folgt also dem Prinzip der *lazy evalua-*

¹⁴ Das dort beschriebene System macht sich den Umstand zunutze, daß seine Szenen „im Freien“ spielen und sich mehr oder weniger auf die Ebene beschränken.

¹⁵ Im Falle unseres Verfahrens wären das vorberechnete *bounding boxes* bzw. geeignete hierarchische Cluster von Polygonen mit ihren entsprechenden *bounding boxes*, die erhebliche Zeiteinsparungen zur Laufzeit ermöglichen würden.

¹⁶ Dies könnte z. B. in Spielszenen erreicht werden, indem etwaige Ziele für Kamerafahrten, also im Spiel wichtige Objekte, auch entsprechend plaziert werden, so daß sie garantiert gut zu sehen sind, sich also leicht eine geeignete Kameraposition finden läßt. Die Szenenbeschreibung könnte sogar diese geeigneten Positionen vorab zur Verfügung stellen, was den Planungsbedarf stark verringert. Gleichermaßen könnten, wenn Fahrten nur zwischen vorher bekannten und statischen Punkten erfolgen, diese Kamerapfade komplett vorberechnet werden, so daß die Planung insgesamt nicht mehr zur Laufzeit stattfinden muß.

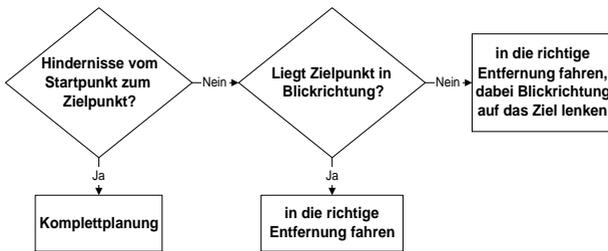


Abbildung 5.6: Planungsmodi: Eine wirkliche Planung ist natürlich nur dann erforderlich, wenn das Zielobjekt von der momentanen Kameraposition aus nicht *gut* sichtbar ist. Selbst dann kann sich die Planung des Bewegungspfad als trivial erweisen, wenn nämlich die angestrebte *bessere* Kameraposition ohne Hindernisse erreichbar ist. Sind aber Hindernisse vorhanden, so wird eine Planung nötig, die diese Hindernisse umgeht.

tion, wenn auch die gesamte Planung im implementierten System *vor* Antritt der Kamerafahrt stattfindet. Wie in Abbildung 5.6 gezeigt, wird außerdem noch der Fall berücksichtigt, daß die Kamera bereits auf das Zielobjekt bzw. auf dem gewünschten Sichtstrahl (aber mit der falschen Entfernung zum Ziel) liegt.

Kollisionstest Die Planung selbst basiert auf einer sukzessiven Unterteilung des Bewegungspfad an Hindernissen. Der zugrundeliegende Kollisionstest besteht aus drei Teilen:

triviale Ablehnung Ist der Abstand zwischen Linienmitte und *bounding box* größer als Linienlänge plus Box-Diagonale, kann die Linie die Box nicht erreichen. Dieser Test (siehe Abbildung 5.7) ist der schnellste und wird daher zuerst ausgeführt. Sind die Linien sehr kurz, besteht eine besonders hohe Wahrscheinlichkeit, daß die nachfolgenden Tests vermieden werden können.

Test gegen die *bounding box* Testet die Linie gegen die sechs Flächen der *bounding box*. Es werden, wie in [GOTTSCHALK et al. 1996], orientierte *bounding boxes* benutzt. Gegenüber achsparallelen *bounding boxes* haben frei orientierte den Vorteil, daß sie (etwa im Falle von schräg im Raum liegenden Objekten) eine wesentlich engere Umschließung

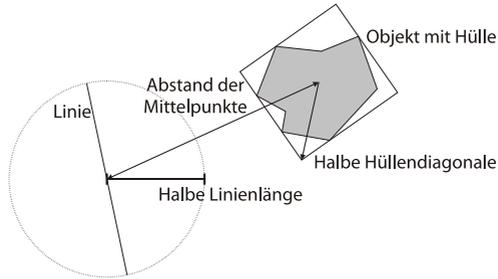


Abbildung 5.7: Kollisionstest Linie-Polygon. Zunächst wird geprüft, ob die halbe Summe der Längen von Diagonale und Linie kleiner als der Abstand der Linienmitte zur Objekthülle ist. Ist dies nicht der Fall, müssen Tests gegen die Flächen der Objekthülle oder sogar gegen die einzelnen Polygonflächen folgen.

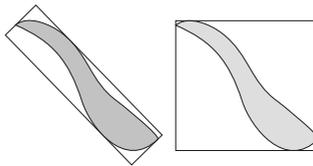


Abbildung 5.8: Vergleich achsparalleler und am Objekt orientierter Hüllkörper.

des Körpers ermöglichen und daher weniger falsche Treffer produzieren (siehe Abbildung 5.8). Demgegenüber steht eine recht moderate Erhöhung der Berechnungskomplexität für den Test.¹⁷

Test gegen die einzelnen Flächen Speziell bei vielen Flächen pro Objekt der aufwendigste Test, ist er doch sinnvoll, besonders bei *hohlen* Körpern, z. B. ein Tisch mit seinen Beinen, um falsche Treffer zu vermeiden. Die Benutzung dieses Tests kann je nach Geometrie der Hindernisse (wie im Beispiel des Tisches) zu recht ungewöhnlichen Kamerafahrten führen; wenn diese nicht gewünscht sind, kann der Test auch entfallen und das Objekt zählt als getroffen, auch wenn keine Fläche berührt wurde.

¹⁷ Die Berechnung minimal umschließender *bounding boxes* ist allerdings recht anspruchsvoll und wird vom beschriebenen Algorithmus weder geleistet noch benötigt.

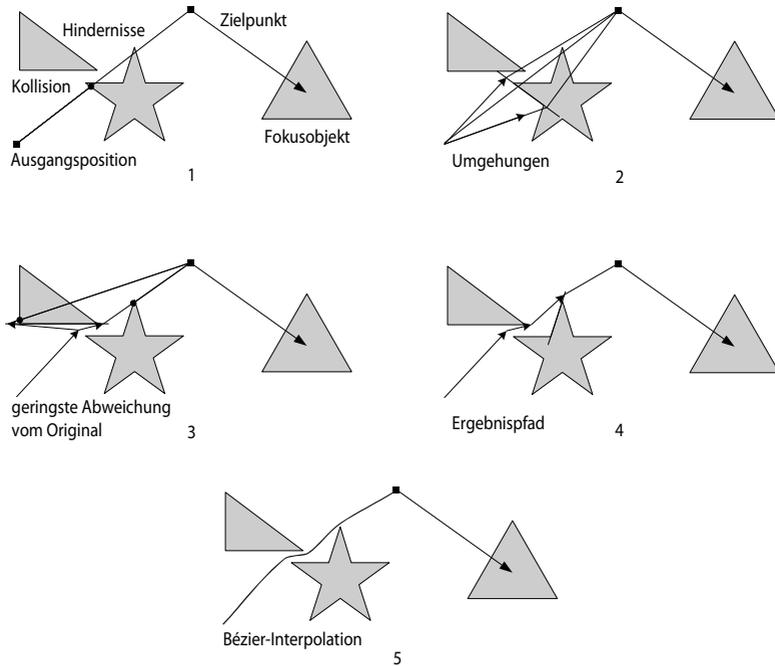


Abbildung 5.9: Ausweichmanöver beim Umgehen von Hindernissen, hier in 2D veranschaulicht.

Ausweichpunkte Trifft der Pfad auf ein Hindernis, so wird ein Ausweichpunkt erzeugt, der außerhalb dieses Hindernisses liegt (siehe Abbildung 5.9). Der ursprüngliche Pfadabschnitt wird dadurch in zwei Segmente zerlegt, eines von Start des Segments zum Ausweichpunkt und eines von dort zum Endpunkt. Diese beiden neuen Pfadabschnitte müssen nun in gleicher Weise auf Hindernisse getestet (und gegebenenfalls unterteilt) werden wie das ursprüngliche Segment. Die Wahl des Ausweichpunktes ist dabei zunächst offen; wesentlich ist nur, daß er einerseits außerhalb des erkannten Hindernisses liegen muß, andererseits noch nicht in einem neuen Hindernis liegen darf.

Erzeugt man anstatt einem mehrere Ausweichpunkte, so wird das Verfahren

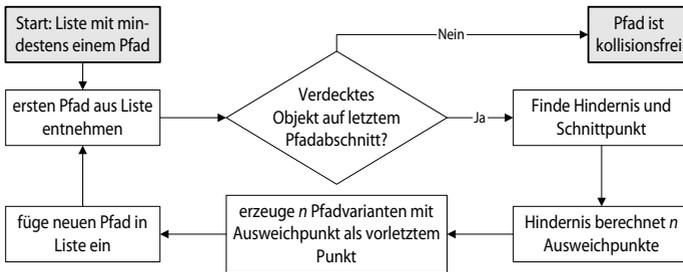


Abbildung 5.10: Planungsalgorithmus. Die Breitensuche wird über eine Liste von Pfadkandidaten realisiert, deren erstes Element jeweils getestet wird. Ist der Pfad nicht geeignet, so werden in jedem Fall Nachkommen erzeugt und ans Ende der Liste gehängt. Dabei können die neuen Pfadkandidaten noch umsortiert werden, so daß bessere Pfade auch bessere Chancen haben, ausgewählt zu werden.

wesentlich robuster, da nun je Unterteilungsschritt mehrere Varianten getestet werden können. So können „Sackgassen“ bei der Planung vermieden werden. Außerdem wird das Verfahren beschleunigt, da die Variante gewählt werden kann, die am schnellsten (mit den wenigsten Unterteilungsschritten) einen kollisionsfreien Pfad liefert.

Dies wird realisiert durch eine hierarchische Organisation der Pfade, bei der, ausgehend vom Ur-Pfad (der direkten Verbindung von Start und Ziel), die Kinder-Pfade jeweils durch Unterteilung des Vaterpfades erzeugt werden. Diese Baumstruktur wird während der Planung aufgebaut und in einer Breitensuche abgearbeitet (siehe Abbildung 5.10).

Beim Erzeugen der Ausweichpunkte wird zunächst in keiner Weise getestet, inwieweit die dadurch erzeugten neuen Pfade *geeignet* sind. Mögliche Problemfälle sind:

- zu große Abweichungen der Ausweichpunkte von gewünschten Streckenverlauf (was bei der Kamerafahrt in übermäßig starken Kurven resultiert),
- Ausweichpunkte, die aus Sicht des Originalpfades nach hinten führen,
- Ausweichpunkte, die wieder innerhalb eines Objektes liegen.

Die zu großen Abweichungen sind zwar unschön, aber manchmal nicht zu vermeiden und es ist auch klar, daß es Richtungsänderungen geben muß, schließlich will man ja ausweichen. Der Algorithmus versucht daher, die unvermeidbaren Kurven insoweit zu begrenzen, daß es Pfade mit geringer Abweichung vom Ursprung zuerst prüft.

Im Falle einer Richtungsumkehr wird ähnlich verfahren. Pfade, die die bisherige Bewegungsrichtung umkehren würden, werden erst dann geprüft, wenn alle anderen Varianten versagen.

In Fall der verdeckten Ausweichpunkte (siehe Bild 2, Abbildung 5.9 auf Seite 99, der rechte Ausweichpunkt) wird ein alternativer Punkt erzeugt, der noch außerhalb des verdeckenden Objektes liegt.

5.4.3 Nachbearbeitung des Kamerapfades

Die nun vorliegende Rohversion des Pfades liegt zwar garantiert außerhalb von Szenenobjekten, eignet sich aus ästhetischen Gründen aber trotzdem nicht für Kamerafahrten. Da der Pfad aus aneinandergereihten geraden Abschnitten besteht, würde die Fahrt bei jeder Richtungsänderung unweigerlich ruckeln. Um das zu vermeiden, wird der Pfad grundsätzlich als Bézierkurve (siehe Abbildung 5.11) interpretiert. Das hat allerdings zur Folge, daß der tatsächlich abgefahrene Pfad Objekte schneiden kann, die der getestete Pfad nicht berührte. Daher muß nun zusätzlich die Bézierkurve getestet werden.

Außerdem tritt hin und wieder ein weiterer ungewollter Effekt auf: Da der Abstand der Ausweichpunkte nur basierend auf der Größe des getroffenen Polygons und des Abstandes der Kollisionspunktes zum Polygonrand getroffen wird, kann es vorkommen, daß der Pfad bei seinen Umleitungen scheinbar in der Szene „herumirrt“. Geometrisch bedeutet dies, daß auch zwischen zwei nicht direkt aufeinanderfolgenden Pfadpunkten kein Hindernis liegt, mithin dieser Teil des Pfades übersprungen werden kann. Da sich dieser Umstand im Nachhinein sehr leicht feststellen und schnell beseitigen läßt, kann er im Interesse einer schnellen Planung, die das Herumirren nicht durch aufwendige Zusatztests zu vermeiden sucht, in Kauf genommen werden. Beim Entfernen der überflüssigen Umleitungen muß allerdings überprüft werden, ob die dabei entstehende Bézierkurve kollisionsfrei bleibt.

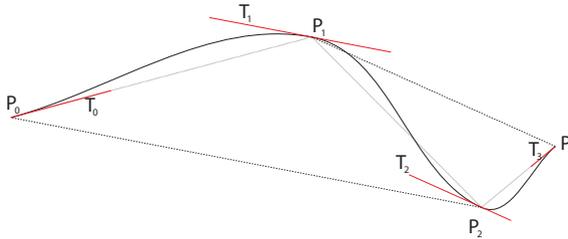


Abbildung 5.11: Konstruktion einer dreiteiligen Bézierkurve aus den Liniensegmenten des Pfades. Die Pfadplanung liefert zunächst nur die Punkte P_0 bis P_3 . Um Kurven zu erzeugen, müssen zusätzlich Tangenten angelegt werden, die den Kurvenverlauf zwischen den Endpunkten bestimmen. Indem die Tangenten parallel zur Verbindungslinie zwischen den beiden Nachbarpunkten gelegt werden, erreicht man den Eindruck einer glatten Kurve. Der Eindruck der Glätte wird durch die gleiche Richtung der Tangenten an den angrenzenden Kurvensegmenten erreicht; die Tangentenlänge jedoch darf auf beiden Seiten der Punkte unterschiedlich sein (proportional zu den beiden Liniensegmenten), damit ein Überschießen auf Seiten des kürzeren Segmentes vermieden wird.

Dies geschieht, indem die Kurve in *ausreichend* kurze Linien unterteilt wird. Die Länge der Liniensegmente, bzw. deren Anzahl kann man überschlagsweise so wählen, daß je Bild ein Liniensegment getestet wird, also bei einer zehnsекündigen Fahrt mit etwa 30 Bildern je Sekunde etwa 300 Segmente. Da die vorherige Planung Umleitungen mit geringer Abweichung von der Originalstrecke bevorzugt, wird die Kurve generell sehr nah an der geplanten kollisionsfreien Linie liegen und die Tests werden fast immer fehlschlagen. Zudem verwirft der benutzte Kollisionstest kurze Linien besonders schnell.

Schneidet die Bézierkurve aber doch ein Objekt, so wird versucht, die Kurve enger an die Linie zu bringen (siehe Abbildung 5.12). Dabei werden naturgemäß die Kurven etwas enger, aber eine aufwendige Neuplanung des Pfades entfällt. Um sicherzugehen, werden anschließend die betroffenen Kurvensegmente neu getestet (In den Abbildungen die von P_0 bis P_4). Falls die neue Kurve immer noch Kollisionen verursacht, kann der Schritt wiederholt werden.

Parametrierung Darunter versteht man bei Bézier- und anderen Kurven den Werteverlauf des Kurvenparameters t über die Kurve hinweg, also das

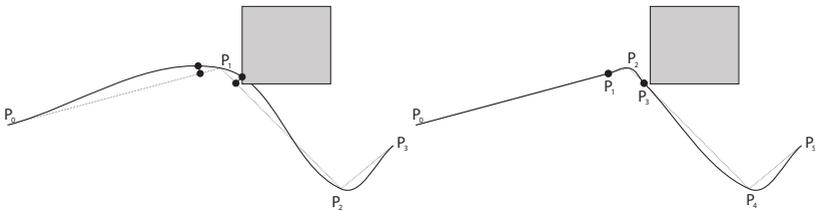


Abbildung 5.12: Links: Die Kurve – aber nicht die geplante Linie – schneidet ein Hindernis. Dabei kommt das Hindernis der geplanten Linie im weiteren Verlauf sogar noch wesentlich näher als am Ort der Kollision mit der Kurve. Es wird die Stelle ermittelt, an der die Kollision auftritt, sowie der nächste Knotenpunkt auf der Kurve (hier P_1). Der dazu symmetrische Punkt auf der anderen Seite der Kurve wird ebenfalls berechnet. Rechts: Die Projektionen beider Punkte auf die geplante Linie werden zusätzlich in die Kurve eingebaut. Da die Tangentenberechnung an einem Knotenpunkt immer die angrenzenden Punkte einbezieht, liegen die Tangenten der neuen Punkte exakt auf der ursprünglich geplanten Linie. Obwohl das Hindernis etwas später auch der Linie recht nah kommt, reicht dieser Schritt fast immer aus. Es ist also nicht nötig, die Stelle zu suchen, an der das Hindernis der Linie am nächsten kommt.

Verhältnis des Zuwachses von t zur dementsprechenden Wegstrecke auf der Kurve. In einer Animation, bei der t über die Zeit hinweg linear interpoliert wird, entspricht die Parametrierung der Geschwindigkeit auf der Kurve. Bei der Verwendung von Bézierkurven wie rechts im Bild 5.13 in der Computeranimation kann es zu abrupten Geschwindigkeitsschwankungen der Bewegung kommen, da die Stützstellen nicht unbedingt in gleichem zeitlichen Abstand voneinander liegen. Das heißt, daß die Kamera bei ihrer Reise für alle Segmente die gleiche Zeit braucht, auch wenn die Länge der Segmente unterschiedlich ist. BUCKLEY beschreibt ein Verfahren [BUCKLEY 1994], das höhere als C^1 -Kontinuität¹⁸ der Bewegung erlaubt; dies erfordert jedoch Bézierkurven fünf-

¹⁸ Der Begriff der Kontinuität von Kurven beschreibt den Grad der Ableitung, bis zu dem eine zusammengesetzte Kurve „glatt“ ist. Ein einzelnes Stück einer zusammengesetzten Bézierkurve ist in allen Ableitungen stetig. Knicke oder höhere Unstetigkeiten können also nur an den Stellen auftreten, an denen ein Kurvenstück an das andere anschließt. C^0 -Kontinuität bedeutet Positionskontinuität, d. h. die Kurve hat keine Lücke. Dies ergibt sich allerdings schon daraus, daß die Kurven einander berühren. Bei C^1 -Kontinuität ist zusätzlich die Richtung der Kurve an beiden angrenzenden Enden gleich. Die Kurve erscheint durchgängig glatt und es gibt – wichtig für Kamerafahrten – keine abrupten Richtungswechsel. Bei Bézierkurven



Abbildung 5.13: Beide Kurven haben den gleichen Verlauf, sind aber verschieden parametrisiert. Die Punkte entsprechen den t -Werten von 0 bis 1 in $1/8$ -Schritten. Während die linke Kurve eine lineare Parametrisierung aufweist, t also eine konstante Geschwindigkeit hat, wird in der rechten Kurve anfangs beschleunigt und am Ende abgebremsert.

ten Grades, deren Verlauf wesentlich schwieriger zu kontrollieren ist. Dies kann sich besonders im Hinblick auf die Kollisionsfreiheit auswirken.

Um dies zu vermeiden und der Anwendung eine einfache Kontrolle über die Geschwindigkeit zu geben, wird der Zeitparameter t nicht einfach über die Stützstellen, sondern über die Länge der Kurve (angenähert durch die Abstände der Stützstellen) interpoliert (mit $t = 0$ am Anfang und $t = 1$ am Ende) und damit eine annähernd konstante Geschwindigkeit der Bewegung erzielt. Wenn dennoch etwa ein beschleunigtes Anfahren und sanftes Abbremsen gewünscht wird, kann dies durch die Verwendung entsprechender Kennlinien über dem Wertebereich von t ($[0, 1]$) geschehen.

5.4.4 Wahl der Blickrichtung

Das bisher vorgestellte Verfahren plant strenggenommen nur die Position der Kamera während der Bewegung. Der zweite Basis-Parameter, der Blickpunkt bzw. die Blickrichtung,¹⁹ blieb zunächst außen vor. Dabei sind für die Planung einer Blickrichtung die gleichen Eingangsdaten bekannt wie für die Position: Die Blickrichtung zu Beginn der Fahrt und die am Ende.

Für die Zeit der Fahrt müßte die Blickrichtung nun in ähnlicher Weise interpoliert werden wie die Position entlang der Pfadkurve. Im Unterschied zur

erreicht man dies durch eine entsprechende Wahl der Kontrollpunkte (siehe Bild 5.11). C^2 -Kontinuität geht noch weiter und fordert eine Gleichheit der Krümmung auf beiden Seiten. Dieses und noch höhere Kriterien ziehen allerdings weitreichende Einschränkungen bei der Wahl der Kontrollpunkte nach sich [FOLEY et al. 1994].

¹⁹ Für den Zweck einer Kamerafahrt kann dieser Parameter als zweidimensional betrachtet werden: Ein Winkelpaar aus Azimut und Elevation. Ebenso wird die dritte Richtung, der Obenvektor, als konstant angenommen.

Position gibt es allerdings für die Abweichung der aktuellen zur endgültigen Blickrichtung eine Obergrenze, nämlich die entgegengesetzte Richtung.

Einfache Verfahren Im einfachsten Fall könnte nun die Blickrichtung über den Verlauf der Fahrt hinweg entlang eines Großkreises²⁰ in die Zielrichtung überführt werden. Das steht allerdings nicht im Einklang mit der angestrebten Hinweisfunktion der Kamerafahrt, die dadurch unterstützt wird, daß die Kamera auf das Zielobjekt zufährt, also in die Fahrtrichtung schaut. Bessere Ergebnisse erzielt man daher, wenn die Drehung der Kamera schon wesentlich eher (etwa in der ersten Hälfte der Fahrt) abgeschlossen ist. In der zweiten Hälfte kann die Kamera dann entweder entlang der aktuellen Bewegung ausgerichtet werden oder auch auf das Zielobjekt schauen.

Verbesserte Verfahren In den zuvor beschriebenen einfachen Methoden findet allerdings immer noch keine Qualitätskontrolle der Blickrichtungen statt; so ist z. B. nicht gewährleistet, daß das Zielobjekt sichtbar ist, während die Kamera darauf zeigt. Besonders bei längeren Fahrten kann es leicht vorkommen, daß die Kamera zur Kollisionsvermeidung an einer Wand entlangfahren muß, hinter der das Zielobjekt liegt. In diesem Fall ist es besser, die Kamera in Fahrtrichtung zeigen zu lassen. Wird das Zielobjekt sichtbar, kann die Kamera wieder schwenken.

Bei all diesen Überlegungen ist zu beachten, daß häufige Schwenks während der Fahrt zu unruhig wirken und vom Sinn der Kamerafahrt ablenken. Kamerafahrten zur Fokusverlagerung in interaktiven Systemen dürften normalerweise weniger als zehn Sekunden dauern und bieten damit gerade genug Zeit für den unbedingt nötigen Schwenk auf das Zielobjekt.

²⁰ Man kann sich die Blickrichtung als Richtung vom Mittelpunkt einer Kugel zu einem Punkt auf der Kugeloberfläche vorstellen. Ein Großkreis ist ein Kreis auf der Kugeloberfläche, dessen Mittelpunkt auch der der Kugel ist. Längenkreise auf dem Globus sind Großkreise, Breitenkreise nicht. Die kürzeste Verbindung zweier Punkte auf einer Kugeloberfläche ist immer ein Abschnitt eines Großkreises.

5.4.5 Fazit

Das vorgestellte Verfahren zur Planung von Kamerafahrten eignet sich wegen seiner Geschwindigkeit gut für den Einsatz in interaktiven Systemen. Wenn die möglichen Hindernisse (also die gesamte Szene) in Weltkoordinaten vorliegen, sind nur wenige Vorberechnungen zum Bestimmen der *bounding box* nötig. In einer dynamischen Szene, z. B. einer Animation müssen allerdings auch die Koordinaten der bewegten Objekte neu berechnet werden.

Im Unterschied zu einigen existierenden Verfahren aus der Robotik (Roadmap-Methode) wird nicht versucht, zunächst auf einen als „sicher“ bekannten Weg zu gelangen. Stattdessen dient der direkte Weg zwischen Start und Ziel als Planungsgrundlage, von der nur bei Bedarf abgewichen wird.

Durch die anschließende kollisions-sichere Umwandlung des stückweise linearen Pfades in eine Bézierkurve können sanfte Kurven erzeugt werden, ohne daß sich das eigentliche Planungsverfahren verkompliziert.

Während der Fahrt muß zusätzlich zur Verschiebung der Kameraposition auch die Blickrichtung angepaßt werden. Dabei sind übermäßig schnelle Schwenks zu vermeiden. Damit das kommunikative Ziel der Kamerafahrt – die Fokusverlagerung hin zum Zielobjekt – ausreichend klar wird, muß die Kamera möglichst lange, in jedem Fall aber gegen Ende der Fahrt, auf das Fahrtziel ausgerichtet sein.

5.5 Zusammenfassung

In diesem Kapitel wurden Verfahren zur Planung von Kamerafahrten vorgestellt. Eng damit verbunden ist die intelligente Auswahl einer Perspektive als Ziel bzw. Endzustand nach einer Kamerafahrt.

Für die Perspektivwahl muß die intelligente Kamera Beschreibungsformalismen zur Verfügung stellen, die eine inhaltliche Spezifikation der gewünschten Ansicht auf einer höheren Ebene als der von 3D-Koordinaten zulassen. In Analogie z. B. zu den unterstützten Zeichenstilen stellen die Kriterien der Perspektivwahl die Basis für weitere Ausdrucksmittel der Kamera dar.

Bei der Umsetzung der gewünschten Zielvorgaben sind mögliche Zielkonflikte und – wie bereits beim Einsatz von statischen Hervorhebungsmitteln – auftretende Sichtbarkeitsprobleme zu berücksichtigen. Im Falle einer dynamischen, reaktiven Kamerasteuerung kommen weitere Probleme der zeitlichen Kohärenz und der vorausschauenden Planung dazu.

Für die interaktive Planung von Kamerafahrten existieren in der Robotik verschiedene Vorarbeiten. Die dort geltenden Randbedingungen und Kompromisse sind allerdings nicht vollständig auf intelligente interaktive 3D-Umgebungen übertragbar.

Mit der Möglichkeit der Kamerafahrt stellt die Grafikengine ein weiteres Darstellungsmittel zur Verfügung, welches auf Seiten der Anwendung zu Hervorhebungszwecken sowie zur Sichtbarmachung verborgener Objekte genutzt werden kann. Zu den damit gebotenen Freiheitsgraden zählen neben der zeitlichen Steuerung der Fahrt (Dauer, Anfahren, Abbremsen, Verknüpfung mit anderen Fahrten zu einer „Rundreise“) auch die konkrete Auswahl einer nicht in Koordinaten, sondern inhaltlich vorgegebenen Zielperspektive.

Das Umweg-Verfahren zur kollisionssicheren und schnellen Planung von Kamerafahrten ermöglicht in interaktiven Systemen eine geführte Erkundung der Szene. Dadurch daß ausdrücklich versucht wird, Hindernissen auszuweichen, eignet es sich auch für längere Fahrten und komplexere Szenen, ohne daß seitens der Präsentationsplanung auf Schnitte ausgewichen werden muß.

Die nun folgenden Fallstudien verdeutlichen, wie Anwendungen von einer ganzheitlichen Betrachtung aller die Kamerasteuerung betreffenden Aspekte profitieren können. Dabei werden verschiedene grafische interaktive Systeme vorgestellt.

Im VISDOK-System (entstanden im gleichnamigen Projekt¹ Visualisierung in technischer Dokumentation) wurde das Konzept einer Interaktiven Technischen Dokumentation umgesetzt. In ihr wird grafische und textuelle Informationen dynamisch und kohärent vor dem Hintergrund unterschiedlicher Szenarien (Angebot, Wartung) präsentiert. Im Verlauf der Interaktion kann der Nutzer Informationen anhand der präsentierten Bilder und Texte erfragen. Die dabei neu entstehenden Situationen werden durch entsprechend angepaßte grafische Gestaltungsmittel repräsentiert.

Das AGI³LE -System² ist mit VISDOK verwandt, richtet sich aber auf eine völlig andere Wissensdomäne: medizinische Illustration. Daher verlangt es andere Interaktions- und Visualisierungsstrategien als die technische Dokumentation.

¹ gefördert vom Kultusministerium des Landes Sachsen-Anhalt von 1995 bis 1998, Förderkennzeichen: 1986A/0025

² AGI³LE steht für Automatic Generation of Intelligent Interactive Interfaces through Language Engineering.

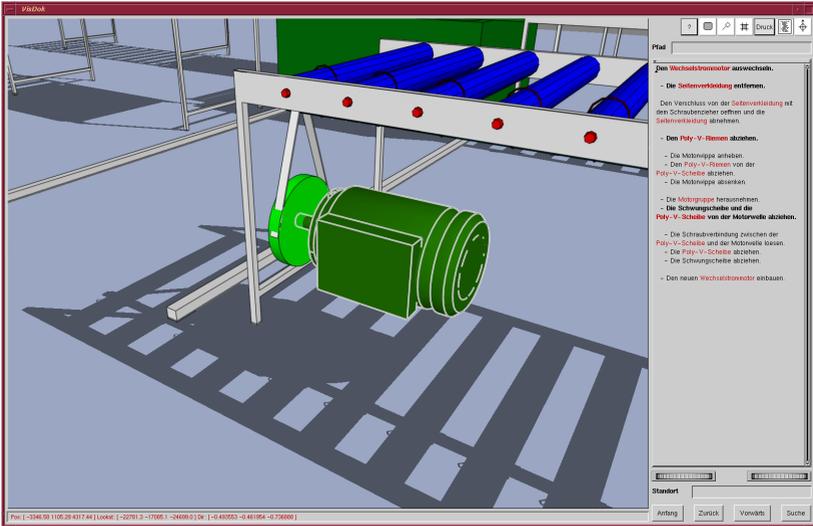


Abbildung 6.1: Eine durch Visdok generierte Präsentation

6.1 Visdok

Im folgenden wird ein technischer Überblick über das im Rahmen des vom Land Sachsen-Anhalt geförderten VISDOK-Projektes entstandene gleichnamige Softwaresystem gegeben.

Eines der wesentlichen Ziele von VISDOK war es, textuelle und visuelle Informationen gemeinsam und transparent zugänglich zu machen und zusätzlich den Betrieb und die Bedienungsabläufe eines technischen Systems durch Animation zu veranschaulichen. Diesem Anliegen folgt auch das gewählte Bildschirm-Layout: VISDOK zeigt neben einer Szenenansicht des technischen Systems erläuternde Texte und ermöglicht das interaktive Erkunden des dokumentierten Produkts (siehe Abbildung 6.1).

Grafische und textuelle Informationen bilden zwei Teile eines Informationsraumes, der von zwei getrennten Softwaresubsystemen verwaltet wird. Dynamisch erzeugte medien- und kontextsensitive Menüs dienen analog zu Hy-

perlinks als Navigationsinstrument im Informationsraum. Gleichzeitig kann der Betrachter seine Position und Blickrichtung in der dargestellten Szene frei wählen. Das System unterstützt also eine Navigation in Bild und Text.

Intelligente Multimediale Präsentationsysteme wie VISDOK, WIP oder PPP [WAHLSTER et al. 1993, ANDRÉ et al. 1996, RIST et al. 1997] verbinden Methoden der Computergrafik mit denen der Künstlichen Intelligenz. Die Symbiose dieser beiden Teilgebiete der Informatik³ ermöglicht Systeme, in denen Information intelligent und adaptiv, d. h. auf den Benutzer und seine aktuelle Situation zugeschnitten präsentiert wird.

Das VISDOK-System (siehe Abbildung 6.1) setzt zahlreiche in den vorhergehenden Kapiteln behandelte Ideen um. Zunächst folgt eine Einführung in das System, seine Architektur und Funktionsweise. Danach wird gezeigt, wie mit Hilfe des Erweiterten Kameramodells die grafischen Darstellungsmittel auf die in VISDOK auftretenden Situationen angewendet werden. Für weitere Informationen zu den wissensbasierten Aspekten des Systems, wie etwa der Interaktionsplanung, Wissensrepräsentation und Textgenerierung, sei auf [HARTMANN 2002] verwiesen.

6.1.1 Technische Dokumentation und Computergrafik

In der technischen Dokumentation spielen neben textuell beschriebenen Sachverhalten wie technischen Daten oder Bedienungsanleitungen auch immer mehr visuelle Mittel eine Rolle, in denen komplexe räumliche Zusammenhänge und der Ablauf dynamischer Vorgänge dargestellt werden. Gedruckte Dokumentationen bieten aufgrund ihrer statischen Natur eingeschränkte Visualisierungsmöglichkeiten und keine Interaktionsmöglichkeiten.

Zudem ist ein hoher Aufwand erforderlich, um die Übereinstimmung zwischen dem Produkt und seiner Dokumentation zu wahren, insbesondere, wenn diese in einer ganzen Reihe von Varianten erstellt werden muß. Diese Varianten ergeben sich aus dem rechtlichen Erfordernis, die Produktdokumentation in der Landessprache des Kunden auszuliefern und durch Anpassung der Dokumentation an den Verwendungszweck, an die angesprochene

³ seit einiger Zeit auch als *Smart Graphics* bezeichnet, siehe z. B. [BUTZ et al. 2001]



Abbildung 6.2: Aufgaben innerhalb eines interaktiven Präsentationssystems

Benutzergruppe (Wartungstechniker, Endkunden, Interessenten) oder an terminologische Vorgaben.

Die Übereinstimmung zwischen einem Produkt und seiner Dokumentation kann gewährleistet werden, wenn die Produktdokumentation automatisch aus einer *formalen Repräsentation* des Produkts abgeleitet wird. Bei VISDOK wird daher davon ausgegangen, daß die vorhandene rechnergestützte Repräsentation des Produktes dieses *hinreichend* genau beschreibt, so daß die daraus abgeleiteten Texte und Grafiken korrekt sind. Dabei bilden die Konstruktionsdaten den wichtigsten Ausgangspunkt. Diese sind um Funktionsbeschreibungen und formale Beschreibungen von Wartungsanweisungen zu erweitern. Im VISDOK-System werden anhand dieser formalen Repräsentation alle mehrsprachigen Varianten automatisch generiert. Abbildung 6.1 auf Seite 110 zeigt das Ergebnis, wenn ein Nutzer die Instandsetzungsanweisungen zu einem in der Szene dargestellten Wechselstrommotor anfordert.

6.1.2 Visdok als multimediales Präsentationssystem

BORDEGONI et. al. extrahieren in einer Analyse verschiedener Beispiele allgemeine Aufgaben intelligenter multimedialer Präsentationssysteme (IMMPS, siehe [BORDEGONI et al. 1997]). VISDOK setzt diese in einer Client-Server-Architektur um (siehe Abbildung 6.2).

Anhand der Aufgaben kann VISDOK in einen Diskursplaner, Medienrealisatoren für Text und Grafik, Darstellungsmethoden für generierten Texte, Szenen und Animationen innerhalb der Medienpräsentation, sowie einen Interaktionsmanager untergliedert werden. Die Diskursplanung umfaßt Präsentations- und Interaktionsplanung. Die Abbildung zeigt, daß die einzelnen Aufgaben nicht unabhängig voneinander sind und sequentiell abgearbei-

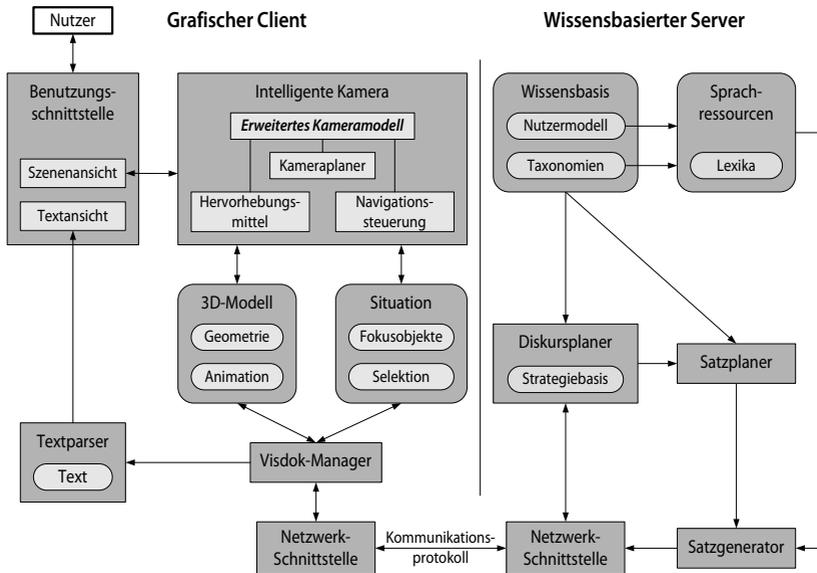


Abbildung 6.3: Systemarchitektur von Visdok

tet werden, sondern in einer engen Wechselbeziehung stehen. Infolge dessen weicht die Systemarchitektur von VISDOK von der logischen Strukturierung nach Aufgaben ab.

6.1.2.1 Verbindung von 3D-Grafik und Wissensbasis

Das VISDOK-System ist als Client-Server-Architektur mit einem grafischen Frontend und einem wissensbasierten Server realisiert worden. Die Komponenten des grafischen Frontends sind in Smalltalk unter Benutzung der OpenGL-Grafikbibliothek implementiert. Die Module des wissensbasierten Servers sind in LISP programmiert. Abbildung 6.3 gibt einen Überblick über die einzelnen Komponenten des VISDOK-Systems.

Das grafische Frontend umfaßt die Grafik- und Animationsrealisatoren sowie den Display- und den Interaktionsmanager. Der wissensbasierte Server bein-

Tabelle 6.1: Server→Client-Kommunikationsprotokoll in Visdok

Nachricht	Parameter	Bedeutung
load-scene	Szenenname	Fordert zum Laden einer bestimmten 3D-Szene auf.
emphasize	Objektnamen	Hebt die genannten Objekte hervor. Kann mit leerer Liste aufgerufen werden, um den Anfangszustand (ohne Hervorhebungen) wiederherzustellen.
show-location	Objektname	Zeigt das Objekt und fordert die Sichtbarkeit. Der Client plant daraufhin eine Kamerafahrt zum Objekt.
visible-objects	keine	Der Server fordert eine Liste der aktuell sichtbaren Objekte an.
animate	Startzeit Endzeit	Client spielt die aktuell geladene Animation teilweise ab. Werden die Zeiten weggelassen, wird die gesamte Animation gezeigt. Der Server kann vor der Animation einen zeitlich annotierten Text senden, der die Animation kommentiert. Im Textbereich des Clients werden dann einzelne Textpassagen während der Animation hervorgehoben.
show	Objektnamen	Zeigt die Objekte an. Das ist der Normalzustand. Die Nachricht wird daher nur nach vorherigem Verstecken der Objekte gebraucht.
hide	Objektnamen	Versteckt die genannten Objekte. Diese werden beim Zeichnen bis auf Widerruf ignoriert.
dominance	Objektname Zahl	Weist einem Objekt seinen aktuellen Dominanzwert zu. Dieser sollte zwischen 0 und 1 liegen.

haltet die Präsentations- und Interaktionsplanung sowie die Textgenerierung. Die netzwerktransparente Kommunikation erfolgt über ein festgelegtes Protokoll, welches die von den beiden Teilsystemen unterstützten Nachrichten definiert (siehe Tabellen 6.1 und 6.2).

Zusätzlich zum Kommunikationsprotokoll zwischen grafischem Frontend und dem wissensbasierten Server muß zwischen den Objekten der in Server und Client verwendeten Wissensbasen und 3D-Modelle vermittelt werden. Die Verbindung erfolgt durch Annotation von Instanzen der Wissensbasis mit Objektnamen des 3D-Modells. Eine Instanz der Wissensbasis kann mit mehreren Objektnamen annotiert werden. Dies ist der Fall, wenn:

Tabelle 6.2: Client→Server-Kommunikationsprotokoll in Visdok

Nachricht	Parameter	Bedeutung
lookupObjName	3D-Name	Erfragt einen aussagekräftigen Namen zu einem Objektbezeichner des 3D-Modells. Wann immer möglich, werden dem Nutzer vollständige Objekt-namen präsentiert. Diese sind sprachabhängig.
set-flag	language, user-group, szenario	Setzt bestimmte Flags im Server, die die Interaktionssteuerung, vor allem aber die Textgenerierung steuern.
get-index	Szenenname	Fordert zur geladenen Szene eine Einstiegsseite bzw. einen Index an.
done	keine	Teilt dem Server mit, daß alle empfangenen Befehle abgearbeitet wurden.
current-scene	Szenenname	Teilt dem Server mit, daß die genannte Szene geladen wurde.
selected-object	Objektname	Informiert den Server, daß der Nutzer ein 3D-Objekt in der Szene selektiert hat. Der Server reagiert gegebenenfalls mit Interaktionsfortsetzungen in Form von Texten oder Menüs.
visible-objects	Objektnamen	Antwort des Clients auf die gleichnamige Anfrage des Servers. Die Menge der sichtbaren Objekte hängt von der derzeitigen Kameraperspektive ab und ändert sich eventuell bei der Navigation.

- die Granularität des 3D-Modells größer als die der Wissensbasis ist,
So kann die Wissensbasis den Begriff *Hand* enthalten, ohne weiter auf ihre Bestandteile wie Finger, usw. einzugehen. Das geometrische Modell dagegen könnte stärker hierarchisch untergliedert sein und jedem Finger einen eigenen Objekt-namen und damit eine Zugriffsmöglichkeit bieten.
- oder aber 3D-Modelle in verschiedenen Detaillierungsgraden vorliegen.

In diesem Fall liegen nicht einzelne Bestandteile der Hand vor, die gleichzeitig dargestellt werden, sondern die Hand existiert in mehreren Versionen, von denen zu jedem Zeitpunkt nur eine aktiv ist. Der Dis-

kursplaner kann so zwischen Modellen verschiedener Detaillierungsgrade wählen.

Umgekehrt können gegebenenfalls mehrfach in der Szene vorkommende Objekte *einer* Kategorie der Wissensbasis zugeordnet sein.

6.1.2.2 Medienpräsentation und Interaktionsmanagement

In der Medienpräsentation werden die medienunabhängig spezifizierten Inhalte dem Benutzer auf einem Ausgabemedium dargestellt. Im VISDOK-System wird diese Aufgabe für Szenen, generierte Texte und Menüstrukturen vom grafischen Frontend übernommen. Desweiteren nimmt das Frontend Benutzereingaben entgegen und arbeitet sie ab bzw. leitet für den Diskursplaner bestimmte Eingaben dorthin weiter.

VISDOK ermöglicht dem Nutzer das freie Navigieren⁴ in der dargestellten Szene. Dadurch kann der Präsentationsplaner nach einer Bewegung in der Szene keine Annahmen mehr über den aktuellen Bildinhalt treffen. Aus diesem Grund hat er die Möglichkeit, die sichtbaren Objekte vom Frontend zu erfragen. Damit ist eine Bezugnahme auf den Sichtbarkeitsstatus von im Text erwähnten Objekten möglich, und es können Maßnahmen zur Sichtbarmachung wichtiger Objekte (siehe Abschnitt 5) eingeleitet werden.

Die Textdarstellung basiert auf einem SGML-Parser. Maussensitive Textstellen werden wie in WWW-Browsern hervorgehoben, so daß mögliche Anknüpfungspunkte im Informationsraum sichtbar werden. An diesen Stellen werden im Text unsichtbar eingebettete Menübeschreibungen aktiviert, die mögliche Diskursfortsetzungen anbieten. Die gleiche Art von Menüs wird bei Selektionen in der 3D-Grafik erzeugt und angezeigt (siehe Abbildung 6.4).

Die Menüeinträge enthalten neben dem Menütext auch eine komplette Beschreibung der entsprechenden Aktionen (siehe Abschnitt 6.1.3), die bei Selektion des Eintrages ausgeführt werden sollen. Die Aktionen können dabei entweder in den Zuständigkeitsbereich der Grafik (also des Frontends) oder

⁴ Die Maus befindet sich bei der Navigation innerhalb der Szenenansicht und nicht an speziellen Buttons. Um die Mausektionen trotzdem eindeutig als Objektselektion oder Navigation interpretieren zu können, ist zur Unterscheidung bei der Navigation ähnlich wie bei VRML-Betrachtern das Drücken der Shift- bzw. Control-Taste erforderlich.

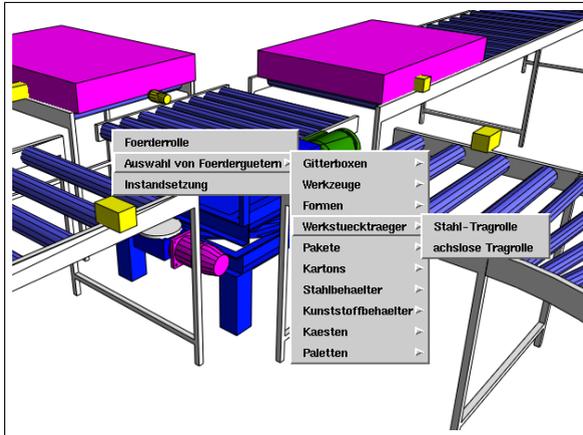


Abbildung 6.4: Menüdarstellung in Visdok

der Wissensbasis fallen. Das grafische Frontend hält eine Liste mit den Aktionsbezeichnern, die es selbst ausführen kann, und schickt alle anderen Aktionen an den Diskursplaner weiter.

6.1.2.3 Medienrealisierung

Diskursplanung und -umsetzung sind in VISDOK getrennt. Die medien-spezifischen Realisatoren bekommen nur noch kommunikative Ziele, die dann in einer für das Medium geeigneten Form umzusetzen sind.

Textrealisierung Innerhalb des Präsentationsplaners wird ein gegebenes kommunikatives Ziel in Teilziele zerlegt, die über rhetorische Beziehungen miteinander verknüpft sind. Diese Beziehungen werden dazu genutzt, die rhetorische Struktur eines zu generierenden Textes [MANN und THOMPSON 1987] festzulegen. Anschließend muß entschieden werden, welche Propositionen in Sätzen zusammengefaßt werden, und es ist die Struktur der zu generierenden Sätze zu bestimmen [KASPER 1989]. Diese Schritte werden als Textplanung, Linearisierung bzw. Satzplanung bezeichnet. Durch sie werden Ausgangsstruk-

turen erzeugt, die an einen Satzgenerator übergeben werden, der sie in natürlichsprachliche Ausdrücke umformt. Satzplaner und -generator beruhen auf Modulen, die innerhalb des TECHDOC-Projektes [RÖSNER und STEDE 1994] entwickelt wurden. In einem weiteren Schritt werden die generierten Texte durch Anweisungen zur Generierung von Diskursfortsetzungen erweitert.

Genauere Beschreibungen der zahlreichen nicht-grafischen Aspekte des Systems finden sich in [HARTMANN 2002].

Grafikrealisierung Die 3D-Szene enthält sämtliche Geometrieinformationen sowie grafische Attribute und die Objekthierarchie. Jedes Grafikobjekt ist durch seine Form (Geometrie), Farbe, Darstellungsattribute und seine Position und Orientierung in der Szene charakterisiert.

VISDOC akzeptiert Szenendateien im 3DS-Format, welche sich mit 3D-Studio, AutoCAD und anderen 3D-Modellierern erzeugen lassen. Für das Anwendungsszenario der technischen Dokumentation wurde davon ausgegangen, daß die benötigten CAD-Modelle ohnehin im Planungsprozeß anfallen. Diese Annahme ist insofern wichtig, als dadurch zusätzliche Kosten vermieden werden, die die Akzeptanz Interaktiver Technischer Dokumentation erschweren.

Animationsrealisierung Die Animation ist aus der Sicht des Präsentationsplaners ein Darstellungsprimitiv, das zur Umsetzung bestimmter Ziele eingesetzt werden kann. Animationssequenzen können als Ergänzung zur textuellen Beschreibung von Vorgängen und Abläufen bei der Benutzung eines technischen Gerätes ausgeführt werden.

Im 3DS-Format lassen sich auch Keyframe-Animationen ablegen. Diese Animationen beschreiben den Zustand⁵ der Szene zu bestimmten Zeitpunkten (*keys*). Für die Animation wird dann zwischen diesen *keys* geeignet interpoliert. Dazu sollte die Szenendatei geeignet strukturiert sein, so daß jedem animierten Kommunikationsziel ein entsprechender Abschnitt in der Keyframesequenz zugeordnet werden kann. Zur Umsetzung eines Kommunikationsziels braucht dann nur dieser Zeitabschnitt animiert werden.

⁵ z. B. Position und Orientierung von Objekten

Dabei kann Animation mit einer textuellen Erklärung koordiniert werden. Dies geschieht, indem zeitgleich mit dem Beginn der Animation ein entsprechender Erklärungstext dargestellt wird, der die einzelnen Schritte im animierten Ablauf beschreibt. Über spezielle Textannotationen wird jedem Abschnitt des Textes ein Zeitabschnitt der Animation zugeordnet, so daß der jeweils aktuelle Teilvorgang im Textfenster hervorgehoben wird, während die Animation läuft.

Kamerafahrten Neben den aus der Szenendatei ausgelesenen Animationen kann VISDOK auch vom System selbst generierte Kamerafahrten (siehe Abschnitt 5) ausführen. Dies ist immer dann erforderlich, wenn der Diskursplaner die Sichtbarkeit eines bestimmten Objektes fordert.⁶ Dazu wird im Frontend ein Bewegungspfad, beginnend mit der aktuellen Kameraposition hin zum hervorzuhebenden Objekt, geplant (zum Planungsverfahren siehe Abschnitt 5.4). Die Planung ist notwendig, da bei der Kamerafahrt keine Objekte durchdrungen werden dürfen.

Wenn der Weg der Kamerafahrt feststeht, sollte auch die Dauer ermittelt werden. Ein mögliches Kriterium berücksichtigte die Komplexität der erzeugten Animation, also Anzahl und Komplexität der während der Fahrt sichtbar bzw. unsichtbar werdenden Objekte. Eine solche Berechnung ist allerdings recht aufwendig. Im VISDOK-System wurde daher ein einfacherer Weg gewählt: Als Komplexitätskriterium wird die Anzahl der Richtungsänderungen während der Fahrt benutzt. Die für die Kamerafahrt veranschlagte Zeit richtet sich nach der Anzahl der zu gehenden Umwege und beträgt etwa fünf bis zehn Sekunden.

6.1.3 Interaktionsplanung

Nach der Vorstellung der Systemarchitektur im vorigen Abschnitt wird im folgenden erläutert, wie Benutzeranfragen in kommunikative Ziele umgewandelt werden, die als Ausgangspunkt für die Präsentationsplanung dienen.

⁶ Im VISDOK-System kann der Benutzer die Erzeugung von Kamerafahrten auch abschalten und erhält so die alleinige Kontrolle über die Kameraperspektive.

Generierung kontext- und medienspezifischer Diskursfortsetzungen

Wird ein Objekt der Szene selektiert, so wird der Interaktionsplaner darüber informiert und legt anhand der über das selektierte Objekt in der Wissensbasis verfügbaren Informationen die Möglichkeiten zur Diskursfortsetzung fest. Je nachdem, welche Informationen dort vorhanden sind, gibt es verschiedene Anfragen, die der Benutzer an das System stellen kann.⁷ Die einzelnen Anfragemöglichkeiten werden dem Benutzer in Form eines Menüs angeboten. Dabei werden zusammengehörende Informationen zu eingebetteten Menüs zusammenfaßt.

Auswahl kommunikativer Ziele Den einzelnen Menüpunkten sind Anweisungen (kommunikative Ziele) zugeordnet, die eine Interpretation der Absichten/Ziele des Benutzers darstellen, wenn der entsprechende Menüpunkt ausgewählt wird. Für jeden Menüpunkt wird also bereits vorher eine Antwort des Systems festgelegt. Diese kommunikativen Ziele dienen als Eingabe für den Präsentationsplaner, der darüber entscheidet, wie diese Ziele umgesetzt werden. Diese Anweisungen können zu Veränderungen der grafischen Darstellung, zur Berechnung bzw. dem Abspielen von Animationen führen, aber auch zu Veränderungen der zugrundeliegenden Wissensbasis, zur Erzeugung weiterer Menüstrukturen, oder zur Erzeugung von Texten. Die Beschreibung der Menüstruktur wird über eine Netzwerkverbindung an die grafische Oberfläche weitergeleitet, die für die Anzeige der generierten Menüstrukturen und die Benutzerinteraktion zuständig ist.

Anhand der Zugehörigkeit der gegebenen Instanz zu einer Klasse der Wissensbasis wird entschieden, welche Diskursfortsetzungen dem Benutzer angeboten werden. Zudem werden globale Parameter, wie die Zielsprache, das aktuelle Szenario oder die Zuordnung des Benutzers zu Benutzerklassen berücksichtigt. Welche speziellen Angebote in die einzelnen Diskursfortsetzungen aufgenommen werden, wird anhand der Relationen, in welchen die gegebene Instanz zu anderen Instanzen der Wissensbasis steht, bestimmt. Dieses Verfahren erlaubt die Beschreibung einzelner Diskursfortsetzungen in klassenspezifischen Methoden. Zur Bestimmung der Gesamtheit von möglichen Dis-

⁷ Wie allgemein üblich, sind nur Fragen erlaubt, deren Antwort dem System bekannt ist.

kursfortsetzungen werden alle anwendbaren Methoden aufgerufen und deren Ergebnisse in einem Menü zusammengefasst.⁸

Auf dem grafischen Interface sind Objekte im Text und in der Grafik mausensensitiv. Es werden dabei die gleichen Methoden zur Generierung kontextsensitiver Menüs benutzt, die aber das Medium berücksichtigen müssen. Falls in der grafischen Darstellung ein im Text erwähntes Objekt vorkommt, wird in die Menüstrukturbeschreibungen zu Textobjekten eine Aufforderung aufgenommen, das sprachlich referenzierte Objekt in der Grafik zu zeigen (show-object). Ist das referenzierte Objekt nicht sichtbar, wird die Kamera so transformiert, daß ausgehend von der aktuellen Betrachterposition eine Kameraperspektive eingenommen wird, welche die Sichtbarkeit dieses Objektes herstellt (siehe Abschnitte 55 und 5). Zudem wird das Objekt hervorgehoben (siehe dazu auch Abschnitte 6.1.2.3 und 3.1.5). In Menüstrukturen, die für selektierte Grafikobjekte erstellt werden, wird der Objektname eingefügt.

Visueller und textueller Index Der Umfang von technischer Dokumentation macht geeignete Zugriffsmechanismen notwendig, um benötigte Information aufzufinden. Dazu werden in VISDOK verschiedene Verfahren benutzt.

Ein Index dient als Einstiegspunkt für die interaktive Erkundung des Informationsraumes. Er wird beim Laden eines Modells automatisch generiert und präsentiert. Auch später während der Sitzung kann der Nutzer durch Drücken eines Buttons der Werkzeug-Leiste jederzeit zu diesem Index zurückkehren.

Im *textuellen Index* werden alphabetisch sortiert alle Objektklassen angeboten, die im aktuell dargestellten Modell benutzt werden und denen Instanzen der Wissensbasis zugeordnet werden konnten.⁹ Durch Selektion der Namen wer-

⁸ Die Festlegung der Reihenfolge der einzelnen Diskursfortsetzungen erfolgt anhand der Spezialisierungshierarchie der Methoden, wobei Diskursfortsetzungen, die durch unspezifischere Methoden erzeugt werden, vor Diskursfortsetzungen, die durch spezifischere Methoden erzeugt werden, angeordnet werden.

⁹ Diese Auslese ist notwendig, da es nicht unbedingt eine 1 : 1-Beziehung zwischen Instanzen der Wissensbasis und den sichtbaren Objekten in der 3D-Szene gibt. So kann die Wissensbasis abstrakte Begriffe enthalten, zu denen keine Vergegenständlichung existiert, oder die Szene ist (da sie ja nicht unbedingt mit dem Ziel der Dokumentation erstellt wurde) viel detaillierter als die Wissensbasis.

den die im Modell sichtbaren Instanzen hervorgehoben. Gleichzeitig werden alle anderen Objekte des Modells deakzentuiert (siehe Abschnitt 3).

Der textuelle Index ermöglicht also das Auffinden von Objekten in der Szene, deren Name dem Nutzer bekannt ist. In umgekehrter Weise ist es auch möglich, zu einem in der Szene gefundenen Objekt den Namen bzw. weitere Informationen abzufragen. Neben dem Textfenster dient also auch die grafische Darstellung als Index. Die frei wählbare Ansicht des 3D-Modells wird dabei als *visuelles Stichwortverzeichnis* zum Einstieg in den Informationsraum genutzt. Hier können Objekte mit der Maus selektiert werden. Daraufhin wird ein Menü generiert, mit dessen Hilfe sich weitere Informationen zu diesem Objekt abrufen lassen. Die Visualisierung bietet dabei den Vorteil, daß die genaue Bezeichnung des Objektes nicht notwendig ist, um an Informationen zu gelangen, die zu diesem Objekt relevant sind.

6.1.4 Beispiel: Eine Visdok-Sitzung

Zu Beginn lädt der Nutzer eine 3D-Szene. Das Frontend fordert ihn dabei auf, ein VISDOK-Backend (der Wissensbasis, Sprachgenerator und Interaktionsplaner bereitstellt) zu wählen. Dies geschieht durch die Angabe eines Rechnernamens und eines TCP-Ports.

Nach dem Laden der Szene sendet das Frontend die Information über den erfolgreichen Ladevorgang an das Backend, welches daraufhin eine passende Wissensbasis aktiviert und eine „Einstiegsseite“ erzeugt. Diese wird an das Frontend geschickt und dort im Textbereich dargestellt. Dieser Zustand entspricht der Ausgangssituation in Bild 6.6 auf Seite 129. Bereits zu diesem Zeitpunkt steht dem Interaktionsplaner die volle Bandbreite der in Tabelle 6.1 auf Seite 114 gezeigten Befehle zur Verfügung. Er kann also auch, falls gewünscht, die Kamera auf ein bestimmtes Objekt lenken, etc. Im Textbereich hat der Nutzer über ein Popup-Menü die Möglichkeit, die gewünschte Sprache für den Textgenerator zu wählen sowie eine Nutzergruppe und ein Szenario festzulegen. Diese Informationen werden bei der Informationsauswahl und der Detaillierung der Texte berücksichtigt.

Die Einstiegsseite sollte so aufgebaut sein, daß sie dem Nutzer neben einer Orientierung auch genügend textuelle Interaktionspunkte in Form von Hyperlinks bereitstellt, so daß die Erkundung des Informationsraumes beginnen

kann. Die Hyperlinks sind in Menüs enthalten, die sich bei Selektion eines Textbereiches öffnen (siehe Abbildung 6.4 auf Seite 117 für eine Darstellung der Menüs auf der Grafikseite; die der Textansicht sind dazu analog gestaltet).

Angenommen, im Text sind bestimmte Maschinen oder ähnliches erwähnt, dann würde sich eine Ausgestaltung der Hyperlinks mit dazu vorliegenden Informationen wie technischen Daten oder Einsatz- und Konfigurationsmöglichkeiten anbieten, die dann in weiteren Textseiten präsentiert werden. Zusätzlich können die Menüs eine grafische Hervorhebung der Objekte oder Kamerafahrten anbieten. Eine vom System aktivierte Hervorhebung entspricht den Zuständen in der rechten Hälfte von Bild 6.6. Eine Kamerafahrt ist in diesem Diagramm ein Übergangszustand, der mit Erreichen des Ziel automatisch in einen der stabilen Zustände übergeht. Welche Interaktionsfortsetzungen in den Menüs bereitgestellt werden, hängt von den Entscheidungen des Interaktionsplaners und dem ihm zugänglichen Wissen in der Wissensbasis ab.

Die 3D-Ansicht bietet in analoger Weise die Möglichkeit, sichtbare Objekte mit Menüs zu verknüpfen. Diese können die gleichen Interaktionsfortsetzungen bieten wie die Textmenüs, müssen aber nicht. So wären z. B. Angebote einer Kamerafahrt zum Objekt wenig sinnvoll, da der Nutzer das fragliche Objekt bereits vor sich hat. Die gebotenen Interaktionsfortsetzungen müssen sich dabei nicht auf das Medium (Bild/Text) beschränken, in dem das Menü aufgerufen wurde. Zusätzlich zum Aufruf der Menüs kann sich der Nutzer in der 3D-Ansicht durch Interaktion mit der Maus frei bewegen. Zur besseren Orientierung erscheinen die Namen von Szenenobjekten am Mauszeiger, wenn sich dieser über einem Objekt befindet.

Damit findet die Arbeit mit dem VISDOK-System in einem Wechselspiel von Bereitstellung von visueller und textueller Information und menügesteuerten Anfragen nach weiteren Informationen statt. Dieser Prozeß ist nicht auf die gerade geladene 3D-Welt beschränkt; über den in Tabelle 6.1 enthaltenen Befehl `load-scene` kann jederzeit der gerade beschriebene Ablauf in Gang gesetzt und eine neue Welt erkundet werden.

Visdok als Intelligente Kamera Das grafische Frontend in VISDOK stellt die 3D-Szene dar und realisiert die Benutzungsschnittstelle. Aus Sicht der

Computergrafik implementiert es dabei eine *intelligente Kamera*, wie sie in Abschnitt.2.3 beschrieben wird. Sie empfängt aus der Anwendung Informationen über den Diskursverlauf bzw. die Interaktionen des Nutzers und leitet daraus entsprechende Darstellungsänderungen ab, die sich auf die Wahl der Kameraperspektive, die gewünschten sichtbaren Objekte, Zeichenstile und andere, von der Grafikengine unterstützte Merkmale beziehen können.

Die folgenden Abschnitte beschreiben die dieser intelligenten Kamera zur Verfügung stehenden Darstellungsmittel und die Situationen, auf die sie angewendet werden müssen. Danach wird das konkrete Zusammenwirken der Darstellungsmittel beschrieben, wie es von der intelligenten Kamera und ihrem erweiterten Kameramodell koordiniert wird.

6.1.5 Hervorhebungen

Sowohl im Text als auch in der grafischen Darstellung werden wichtige Teile hervorgehoben. Die Wahl der Hervorhebungsattribute hängt immer von den zur Verfügung stehenden Realisierungsmöglichkeiten ab. Dieser Abschnitt stellt die in VisDOK unterstützten Darstellungsmittel vor und begründet sie. Zusätzlich werden auftretende Seiteneffekte und Fragen der Kombinierbarkeit mit anderen zeichnerischen Mitteln diskutiert.

6.1.5.1 Schattierte Polygone

Dank guter Erkennbarkeit der geometrischen Form und der schnellen Darstellung werden als Standardmodus schattierte Flächen gewählt. Zusätzlich angebrachte Linien oder deren ausschließliche Verwendung symbolisieren dagegen bereits eine Sonderstellung, wirken also hervorhebend (siehe Abschnitt 3.3.2).

Die benötigten geometrischen Daten (Koordinaten, Farben) werden den CAD-Modellen entnommen. Falls nicht vorhanden, wird die für Schattierungen wichtige Normaleninformation berechnet. Weitergehende Oberflächeneigenschaften wie etwa Glanz oder gar Texturen wurden nicht verwendet. Dies entspricht der Ausrichtung des Systems, welches ja auf vorhandenen CAD-Daten basieren sollte und nicht auf speziell aufbereiteten Modellen, die den Aufwand für die Dokumentation wieder erhöht hätten.

6.1.5.2 Liniengrafik

Liniengrafik existiert in VISDOK in zwei Varianten: Direkt nach dem Laden des Modells können Linien nur mit den gleichen Polygondaten wie beim Schattieren gezeichnet werden. Die Grafikengine versucht allerdings im Interesse einer besseren Darstellungsqualität, Linien geglättet zu zeichnen (*antialiasing*). Das Glättungsverfahren trägt Linien teilweise transparent auf und mischt dabei die Linienfarbe mit dem Untergrund, um Treppeneffekte bei hohem Kontrast von Linie und Hintergrund zu mindern. Manche 3D-Modelle sind allerdings so organisiert, daß gemeinsame Linien benachbarter Flächen nicht ohne weiteres erkannt werden können und daher doppelt gezeichnet werden. Durch das mehrmalige Einmischen der Linienfarbe tritt der Treppeneffekt wieder hervor und verringert die Darstellungsqualität. Hinzukommt, daß detaillierte Modelle eventuell ein zu dichtes Liniennetz aufweisen.

VISDOK generiert daher zusätzlich vorberechnete Linien, die eine wesentlich höhere Darstellungsqualität haben. Allerdings ist der dazu durchgeführte Rechenprozeß recht zeitaufwendig. Da die aktuelle Blickrichtung in die Berechnung eingeht, muß diese nach jeder Navigation neu durchgeführt werden.¹⁰ Als Resultat erhält man Liniendaten, die nur noch sichtbare Silhouettenkanten oder andere, „scharfe“ Objektkanten enthalten [RAAB 1998]. Diese Linien werden automatisch benutzt, sobald sie vorliegen.

Liniengrafik eröffnet u. a. mit der Liniendicke und -farbe noch zusätzliche Parameter, die weitere Informationen transportieren können. Damit stehen in der Kombination bereits drei Grundvarianten zur Verfügung (schattiert mit Umrandungslinien, nur schattiert oder nur Linien), die in ihren Parametern weiter abgewandelt werden könnten.

Kombination mit schattierter Darstellung Besonders Silhouettenlinien betonen Objektkanten und machen dadurch in der Tiefe angeordnete Objekte gut unterscheidbar, auch wenn deren Farben gleich oder ähnlich sind. Die verbesserte Tiefenwahrnehmung ist auch deswegen positiv, weil alternative

¹⁰ Bei einem mittelgroßen Modell mit etwa 20 000 Flächen benötigt eine SGI Octane etwa 60 Sekunden in der verwendeten Smalltalk-Implementation. Aus diesem Grund wird der Algorithmus parallel als Hintergrundprozeß ausgeführt, der niedrig genug priorisiert ist, so daß er die normale interaktive Arbeit nicht behindert.

Techniken wie Stereoskopie oder Tiefenschärfe wesentlich höheren gerätetechnischen Aufwand erfordern.

6.1.5.3 Kamerafahrten

Kamerafahrten werden geplant und ausgeführt, wenn der Benutzer zu einem eventuell nicht sichtbaren Objekt geleitet werden soll. Dies ist besonders dann der Fall, wenn das System auf das Objekt bezogene Erklärungen gibt bzw. der Nutzer das Objekt sucht. Die Fahrt führt jeweils vom aktuellen Standpunkt zu einer freien Position schräg über dem Zielobjekt. Von dort aus kann anschließend wieder interaktiv navigiert werden.

Kombination mit schattierter Darstellung Um die Kamera flüssig zu bewegen, sollte die Szene während der Fahrt oder der interaktiven Navigation in einem möglichst ressourcenschonenden Darstellungsmodus gezeichnet werden. Eine ruckelfreie Bewegung hat dabei Vorrang vor subtilen grafischen Effekten, die erst bei genauerer Betrachtung eines statischen Bildes wirken. Auch wenn sich die Grenzen des Machbaren ständig weiter verschieben, hat sich an dieser prinzipiellen Prioritätensetzung trotz aller rasanten Entwicklung der Grafikhardware nichts geändert.

Da schattierte Polygone in VISDOK am schnellsten dargestellt werden, wird dieser Modus während der Animation ausschließlich benutzt, selbst wenn während einer Kamerafahrt bestimmte Objekte Hervorhebungsstatus haben. Der Hervorhebungseffekt wird durch das spätestens in der zweiten Hälfte der Fahrt offenkundige Anfahren des Zielobjekts klar.

Die 3D-Formerkennung an schattierten Flächen verbessert sich während der Bewegung nochmals, weil die Beleuchtungsgleichungen, die den Farbverläufen zugrundeliegen, auch die Kameraposition berücksichtigen. Zudem wird in VISDOK eine virtuelle Punktlichtquelle zusammen mit der Kamera bewegt, so daß sich dabei auch die Ausleuchtung der Flächen ändert.

Kombination mit Liniengrafik Zur zuvor beschriebenen verbesserten 3D-Wahrnehmung während der Kamerabewegung kann eine Liniengrafik nicht viel beitragen. Zusatzeffekte wie Transparenz oder Liniengrafik werden daher erst nach Beendigung der Fahrt oder Navigation graduell reaktiviert.

6.1.5.4 Schatten

VISDOK generiert zusätzlich zu den Modelldaten eine virtuelle Grundfläche, auf die die Objekte Schatten werfen (siehe Abbildung 3.5 auf Seite 47).

Ein Schatten ist eine Objektsilhouette, gezeichnet aus einer anderen als der aktuellen Blickrichtung. Er gibt damit einen zusätzlichen Hinweis auf die Objektform, der besonders bei Liniendarstellung des eigentlichen Objektes zur Geltung kommt (siehe Abbildungen 3.7 bis 3.8 auf Seiten 49–50). Die Entfernung des Schattens vom schattenwerfenden Objekt zeigt seine relative Höhe an, die ansonsten aus einer 2D-Projektion nicht eindeutig hervorgeht. Außerdem kann die Helligkeit der Schatten als Hervorhebungsmittel wirken, indem fokussierte Objekte mit intensiveren Schatten gezeichnet werden als defokussierte.

Kombination mit schattierten Flächen und Liniengrafik Die Schattendarstellung ändert nichts am Erscheinungsbild der eigentlichen Objekte und kann problemlos mit allen Darstellungsarten kombiniert werden.

Kombination mit Kamerafahrten Bei Bewegungen der Kamera gilt wieder das bereits zuvor festgestellte Primat der Geschwindigkeit. Die in VISDOK verwandte Methode des Schattenwurfs erfordert, daß die Szene mehrfach gezeichnet wird und hat daher (wie andere Zusatzeffekte auch) einen negativen Einfluß auf die Zeichengeschwindigkeit.

6.1.5.5 Sichtbarkeit

Wie bereits zuvor erläutert, können Sichtbarkeitsprobleme in interaktiven 3D-Systemen eine besondere Rolle spielen. Bei geschickter Anwendung der oben beschriebenen Hervorhebungsmittel lassen sie sich jedoch lösen bzw. stark abschwächen (siehe Abbildung 6.5): Zum einen werden die selektierten Objekte durch die veränderte Darstellung kenntlich und unterscheidbar, und zum zweiten wird mit der Liniendarstellung der anderen Objekte erreicht, daß weniger Verdeckungen auftreten, was wiederum die Sichtbarkeit der hervorgehobenen Objekte erleichtert (siehe auch Abschnitte 3.1.5 und 3.2), sofern sie sich nicht hinter der Kamera befinden.

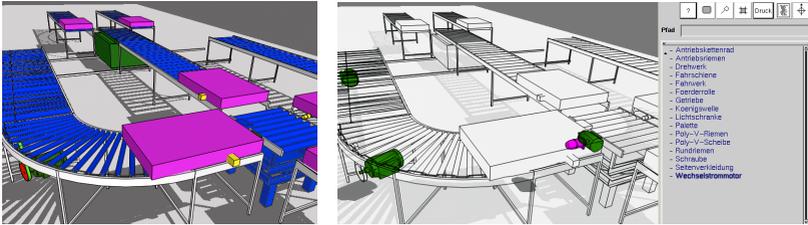


Abbildung 6.5: Visueller und textueller Index in Visdok.

Links: Normalansicht entsprechend des Anfangszustandes aus Abbildung 6.6 (links oben). Es ist keine Auswahl erfolgt und kein Objekt hervorgehoben. Die Sichtbarkeit der Motoren unter den Rollbahnen ist durch Verdeckungen stark beeinträchtigt.

Rechts: Indirekte Hervorhebungen durch Visdok. Selektiert wurde der letzte Indexeintrag, Wechselstrommotor. Das System hebt daraufhin alle Wechselstrommotoren hervor. Die gesamte Szene mit Ausnahme der vier Motoren ist als Liniengrafik dargestellt. Bei flächenhafter Darstellung wie im linken Bild wären nur zwei der Motoren teilweise sichtbar.

6.1.6 Situationen in Visdok

Um Hervorhebungen und Interaktionsmodi gezielt einsetzen zu können, ist die Berücksichtigung der aktuellen Situation, wie sie sich aus den Vermittlungsabsichten des Systems und den Interaktionen des Nutzers ergibt, unerlässlich. Der Situationsvorrat eines Systems variiert natürlich von Anwendung zu Anwendung. In Abbildung 6.6 werden die verschiedenen die Bilddarstellung betreffenden Zustände in VISDOK gezeigt.

VISDOK unterscheidet sowohl intern als auch in der visuellen Darstellung zwischen vom System initiierten *Hervorhebungen* und *Selektionen*, die der Nutzer in der 3D-Ansicht mit Hilfe der Maus ausführt. Auf diese Weise kann der Nutzer zwischen ihm unmittelbar interessierenden und vom System als wichtig erachteten Objekten unterscheiden. *Hervorhebungen* werden dem Diskursfortschritt angepaßt, etwa, wenn das System einen Vorgang textuell erklärt und dabei die erwähnten Objekte in der Grafik hervorhebt. Dazu ist keine Interaktion des Nutzers mit der 3D-Szene nötig.

Selektionen werden durch einen eigenen Linienstil gekennzeichnet und ändern sich nur infolge expliziter Mausbefehle. Gleichzeitig aktive Hervorhebungen bleiben davon unberührt, wie auch das System keine Kontrolle über die Selektion hat.

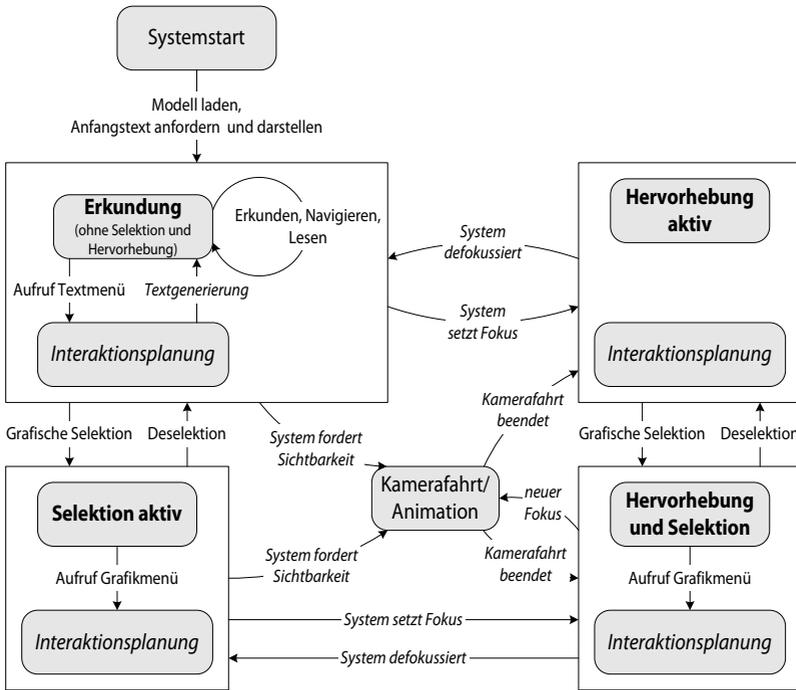


Abbildung 6.6: Situationen (Systemzustände) in Visdok und ihre Übergänge. Die Pfeile mit aufrechter Beschriftung kennzeichnen Nutzeraktionen. Pfeile mit kursiver Beschriftung kennzeichnen Systemreaktionen. Die Interaktionsmöglichkeiten und Reaktionen im Kasten oben links gelten auch in den anderen Situationen, die zum Teil zusätzliche Interaktionen bieten.

Bei einer Selektion wird von der Interaktionsplanung ein Menü angefordert, welches objektbezogene Aktionen anbietet. Der gleiche Mechanismus existiert für die Textansicht. Die Auswahl aus dem dynamisch generierten Menü erzeugt Befehle, die entweder direkt auf die 3D-Szene wirken, weitere textuelle Informationen abrufen oder aber nicht-grafische Informationen zum Objekt ändern. Zu den auf die 3D-Darstellung bezogenen Aktionen zählt insbesondere das Setzen, Rücksetzen oder Ändern der Fokusobjekte, die durch

eine Hervorhebung gekennzeichnet werden.

6.1.7 Das Erweiterte Kameramodell in Visdok

In den vorangegangenen Abschnitten wurden Ziele, Funktionen, grafische Möglichkeiten und auftretende Diskursituationen des VISDOK-Systems beschrieben. Das erweiterte Kameramodell von VISDOK muß diese nun zueinander in Beziehung setzen. Dabei wird außerdem noch auf bisher nicht behandelte Aspekte der interaktiven Navigation eingegangen, die sich ebenfalls aus der beabsichtigten Funktionsweise des Systems ergeben.

Statische Eigenschaften

Basisdaten Kameraposition und -blickrichtung sind frei wählbar. Dies ermöglicht die interaktive Erkundung der Szene unter Ausnutzung von fünf der sieben Freiheitsgrade der virtuellen Kamera (siehe Abschnitt 2.4 und Abbildung 2.3 auf Seite 25). Nicht vorgesehen sind Änderungen des Öffnungswinkels (Zoom) und Kippen der Kamera (Drehen um die Blickrichtung).

Auf die Möglichkeit des Zooms wurde zugunsten der Kamerafahrt verzichtet (zu den Gründen siehe Abschnitt 4.3). Das Kippen der Kamera ist ein meist unerwünschter Nebeneffekt der in interaktiven 3D-Systemen oft benutzten Kamerasteuerung, der in VISDOK nachträglich korrigiert wird.¹¹

Darstellungsstile VISDOK kann schattierte Flächen darstellen, wie sie in den benutzten Modellen beschrieben sind. Zusätzlich können diesen Flächen Linien überlagert werden. Das kann zur Darstellung von Kanten und Silhouetten genutzt werden.

VISDOK nutzt hauptsächlich indirekte Hervorhebungen, d. h. die schattierte Darstellung ist den Fokusobjekten sowie vom Nutzer

¹¹ Zur Steuerung der Blickrichtung wird die Orientierung der Kamera durch Mausbewegungen verändert. Ziel dieser Änderungen ist gewöhnlich ein Neigen der Kamera (Drehen um die waagerechte Bildschirmachse) oder ein Wenden (Drehen um die Senkrechte). Wird dabei der Mauszeiger wie gewöhnlich schräg über den Bildschirm gezogen, ergibt sich zusätzlich ein Kippen um die dritte Bildachse, die Blickrichtung, welches nachträglich korrigiert werden muß.

selbst selektierten Objekten vorbehalten. Sind keine systemseitigen Hervorhebungen aktiv, wird die gesamte Szene schattiert gezeichnet. In Selektionen wird ein kontrastreicherer Linienstil verwendet als für Hervorhebungen (siehe Abbildung 6.7).

Dynamische Eigenschaften

Zeitsteuerung Die zeitliche Steuerung von Animationen wird durch kommunikative Ziele bestimmt. Die Ablaufgeschwindigkeit ist dabei vom System frei wählbar.

Zusätzlich läßt sich vom Benutzer mittels Drehrädern eine bestimmte Animationszeit einstellen, um damit wie im Raum auch

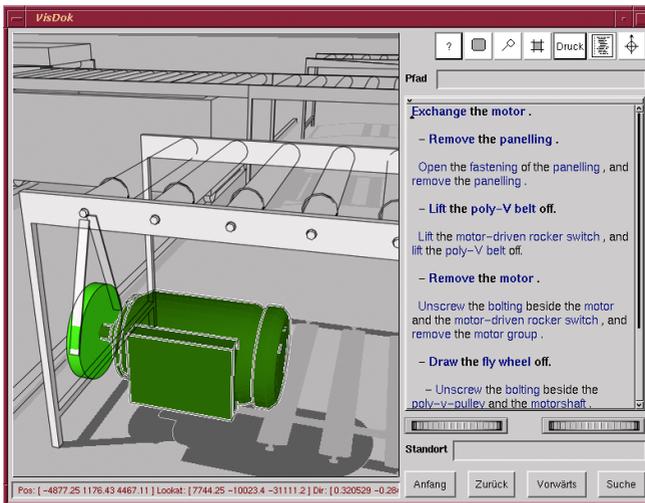


Abbildung 6.7: Gleichzeitige Selektion und Hervorhebung in Visdok. Zur Interaktionsfolge: Der Motor wurde vom Nutzer selektiert. Die Benutzungsschnittstelle informiert den Diskursplaner über dieses Ereignis mit der Nachricht `selected-object`. Dieser antwortet mit der Generierung eines Menüs, über das eine Reparaturanleitung aufgerufen werden kann. Diese wird ebenfalls von der Wissensbasis generiert und im Textbereich rechts präsentiert. Dabei werden die in der Anleitung erwähnten Objekte in der Grafik hervorgehoben. Der Selektionszustand des Motors selbst wird in einem anderen Linienstil hervorgehoben und bleibt von den systemseitigen Hervorhebungen unberührt. Zur Architektur siehe Abbildung 6.3 auf Seite 113. Die Nachrichten sind in den Tabellen 6.2 und 6.1 zusammengefaßt.

eine interaktive Erkundung in der Zeit zu ermöglichen. Damit lassen sich interessierende Phasen eines Vorgangs auch in Eigenregie und aus einer selbst gewählten Perspektive betrachten.

Interaktivität Hier geht es besonders darum, die richtige Balance zwischen

- möglichst freier und intuitiver Erkundbarkeit der Szene in Raum und Zeit und
- Verhinderung bzw. Behandlung von verwirrenden oder unplausiblen Zuständen (Verlaufen, Einnehmen „unmöglicher“ Positionen infolge Durchdringens von Objekten, etc.)

zu finden. Der erste Punkt betrifft sowohl die von VISDOX unterstützte direkte Manipulation der Kamera durch Maus- oder Tastaturbefehle als auch komplexere Aktionen wie die automatisch nach Bedarf geplanten Kamerafahrten.

Eine große Hilfe beim zweiten Punkt ist der in VISDOX benutzte Index (siehe Abschnitt 58), mit dem sich der Benutzer zu bestimmten Orten in der Szene zurücktransportieren kann, wenn er die Orientierung verloren hat. Außerdem erscheint zur besseren Orientierung der Name eines Objektes am Mauszeiger, wenn es überfahren wird. In Situationen, in denen das System bestimmte Erklärungen gibt, kann es den Benutzer in Raum und Zeit durch die Szene führen, ohne daß „störende“ interaktive Eingriffe möglich sind.

In allen Fällen ist jedoch darauf zu achten, daß eventuell nötige Einschränkungen der Bewegungsfreiheit nicht als bevormundend empfunden werden.

Personenbezug In Technischen Dokumentationen ist der Menschbezug verglichen mit Spielen eher schwach ausgeprägt. Der Benutzer tritt in VISDOX nur als Beobachter auf, aber nicht als Agierender. Um in den in Abschnitt 2.8 benutzten Begriffen zu bleiben, kann die Interaktion im VISDOX als Ich-Form betrachtet werden. Es gibt keine Avatare oder andere Bezugspersonen, an denen sich eine *third-person*-Kamerasteuerung orientieren könnte. Damit entfallen auch eventuelle Ansatzpunkte zu personenbezogenen filmischen Gestaltungsmitteln (Einstellungsgrößen, Interaktionen handelnder Charaktere, etc.).

Als nicht personengebundenes filmisches Mittel sind die in VISDOK generierten Kamerafahrten zu betrachten. Diese sind nicht an handelnden Figuren ausgerichtet, sondern dazu gedacht, den Betrachter bei der Navigation in der Szene zu unterstützen.

In dieser Zusammenstellung wird die hohe Spezifik des erweiterten Kameramodells deutlich. Jede der Achsen des Kameramodells (siehe auch Abbildung 2.1 auf Seite 23) steht für eine mögliche Auswahl aus Alternativen bezüglich der Darstellungsmittel, Interaktionsmodi, zeitlicher Abläufe, etc. Die konkreten Entscheidungen und Kompromisse werden nicht nur von den Zeichenfunktionen der Grafikengine und ihren verschiedenen Wechselwirkungen sowie den vorhandenen 3D-Modellen bestimmt, sondern auch von den in der Anwendung vorkommenden Situationen, sowie Fragen der Interaktion, der Zielgruppe und der kommunikativen Absichten. Andere vergleichbare Systeme werden sich in mindestens einem dieser Punkte von VISDOK unterscheiden und andere Darstellungsmittel, Interaktionsparadigmen, etc. bevorzugen. Der Sinn der Erweiterten Kameramodells kann daher nicht darin bestehen, ein allgemeingültiges Regelwerk für den Einsatz grafischer Hervorhebungsmittel zu definieren. Allerdings lassen sich Einflussfaktoren und Kompromisse aufzeigen, die bei Auswahl und Einsatz derartiger Mittel in interaktiven Systemen zu berücksichtigen sind.

Das VISDOK-System erfüllt alle in Abschnitt 2.2 aufgeführten Merkmale eines intelligenten interaktiven Systems. Besonders wichtig erscheinen dabei die Punkte zur Hervorhebungen und Kohärenz. Im Fall von VISDOK als 3D-System sind viele der Darstellungsmittel nicht nur durch die Hervorhebungsabsicht motiviert, sondern auch durch die in interaktiven 3D-Umgebungen essentielle Frage der Sichtbarkeit. Das führte zu Techniken, die weniger wichtige Objekte durchscheinend (durch Transparenz oder Liniengrafik) machen.

Dem wichtigen Ziel der Kohärenz von Bild und Text dienen sowohl die geplanten Kamerafahrten wie auch die dynamische Hervorhebung im Text referenzierter Objekte. Während der erste Punkt sicher auf viele interaktive 3D-Anwendungen übertragbar sein dürfte, stellt sich die Frage der Bild-Text-Kohärenz nur dann, wenn beide Medien präsentiert und koordiniert werden müssen.

6.2 Agi³le

Das AGI³LE -System baut auf VISDOK auf, ist aber auf die Illustration medizinischer Texte gerichtet [HARTMANN et al. 2002]. Dies geschieht wie in VISDOK durch eine intelligente Bild-Text-Kopplung, die geometrische Modelle mit einer symbolischen Repräsentation verbindet, die den Inhalt der Modelle beschreibt (siehe Abbildung 6.8).

Im Unterschied zu VISDOK wird hier aber nicht mit vom System generierten

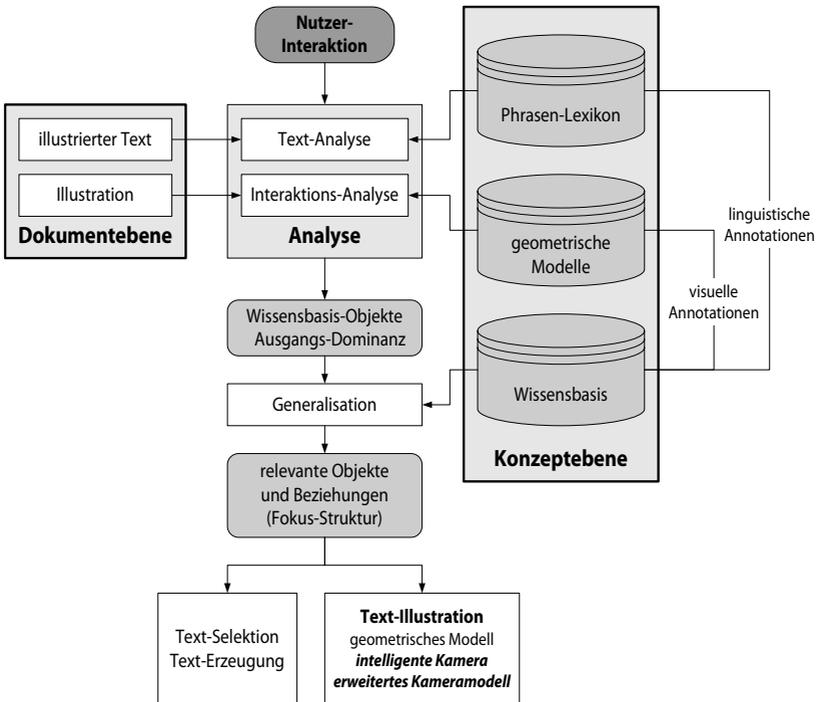


Abbildung 6.8: Die Agi³le -Systemarchitektur nach [Hartmann 2002]. Die Analyse stellt eine Brücke zwischen dem Wissen in der konzeptuellen Ebene und dem Text in der Dokumentenebene her. Die so aufgebaute Fokusstruktur steuert die Textillustration.

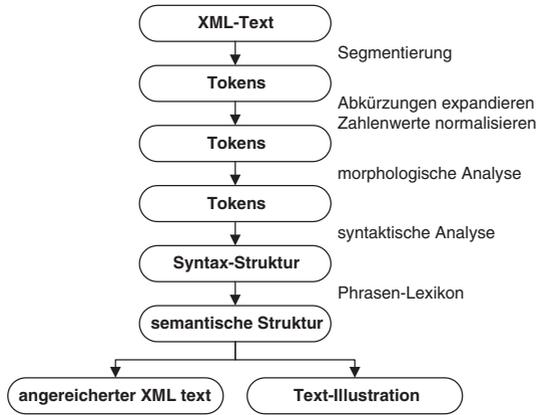


Abbildung 6.9: Linguistische Analyse in Agi³le.

Texten gearbeitet. Stattdessen dienen „handgefertigte“ Texte als *Eingabe*. Als *Ausgabe* entstehen Illustrationen, die den Inhalt des Textes widerspiegeln. Dabei werden ähnlich wie in VISDOK wichtige Objekte hervorgehoben.

6.2.1 Wissensbasierte Textanalyse und Illustration

Um den Text inhaltlich zu erfassen, führt das System eine Textanalyse durch. Im Unterschied zu verwandten Systemen wie dem TEXT ILLUSTRATOR [SCHLECHTWEG und STROTHOTTE 1999] wird aber kein konventioneller *Mustervergleich*, sondern eine morphologische und syntaktische *Analyse* (siehe Abbildung 6.9) durchgeführt. Dadurch werden auch sprachliche Variationen, wie etwa Beugungen oder Synonyme erkannt, die bei einem Mustervergleich verborgen bleiben.

Die bei der Analyse gefundenen Begriffe stehen, da sie ja aus *einer* Satzgruppe stammen, in bestimmten Zusammenhängen zueinander. Diese spiegeln sich auch in der Wissensbasis wider, die verschiedene Begriffe mit Relationen verbindet und damit ein sogenanntes *semantisches Netz* bildet. AGI³LE hebt nun nicht allein die im Text direkt genannten Objekte hervor, sondern auch damit in Beziehung stehende. Dabei wird nicht nur binär zwischen *wichtig*

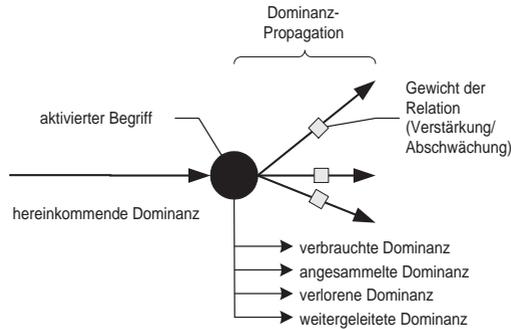


Abbildung 6.10: Weiterleitung der Dominanzen an verbundene Begriffe. Die Relationen in der Wissensbasis sind mit Gewichten versehen, die als Koeffizienten die Weiterleitung der Dominanz verstärken oder abschwächen können, je nach *Stärke* des modellierten Zusammenhangs.

und *unwichtig* unterschieden, sondern jedes Objekt erhält einen *Dominanzwert* zwischen 0 und 1. Die Ermittlung der Dominanz jedes Begriffs geschieht mit einem sogenannten *spreading-activation*-Verfahren, bei dem jeder wichtige Begriff Teile seiner Dominanz an verwandte Begriffe (also solche, zu denen in der Wissensbasis eine Verbindung besteht) weitergibt (siehe Abbildung 6.10).

Dieser Prozeß wird außerdem von einigen Heuristiken gesteuert, die gewährleisten, daß mehrfach aktivierte Knoten (etwa durch Ringbezüge in der Wissensbasis) nicht zu unzulässigen Dominanzanhäufungen führen und die Detailliertheit der Wissensbasis keinen Einfluß auf das Ergebnis hat (zu weiteren Details siehe [HARTMANN 2002]).

Im Ergebnis der Textanalyse entsteht die Fokusstruktur, die beschreibt, welche betrachteten Begriffe in welchem Maße für den Betrachter wichtig sind. Diese Datenstruktur wird für die Illustration geeignet interpretiert. Dabei werden die verfügbaren grafischen Ausdrucksmittel (Beispiele in Abschnitten 3.3 und 3.2) eingesetzt.

6.2.2 Die Intelligente Kamera in Agi³le

Situationen Aufgrund der engen Verwandtschaft zum VISDOK-System ähneln sich auch die in AGI³LE auftretenden Diskurssituationen. Wie in VISDOK ist die vom System oder vom Nutzer bestimmte aktuelle Wichtigkeit der Objekte wesentliches Merkmal der Situation und beeinflusst die Wahl der Darstellungstechniken. Während allerdings in VISDOK Objekte nur entweder ganz oder gar nicht hervorgehoben sein konnten, gibt es in AGI³LE auch Zwischenstufen, die durch eine parametrisierbare Stärke der Hervorhebung realisiert werden.

Die für große Räume noch sinnvollen Kamerafahrten sind für medizinische Objekte nicht immer geeignet. Die Gefahr des „Verlaufens“ ist ebenfalls nicht gegeben. Bedingt durch den oft sehr komplizierten inneren Aufbau von Organen und Körperteilen gibt es in medizinischen Darstellungen häufig Durchdringungen und Überdeckungen etwa zwischen Gefäßen und verschiedenen Gewebsarten. Nur selten existieren Blickrichtungen, die die gesuchten Objekte in ihrer ganzen Ausdehnung unverhüllt zeigen. Das Sichtbarkeitsproblem ist also leichter durch Transparenz als durch automatische Standortwahl zu lösen.

AGI³LE fordert daher keine Kamerafahrten an, sondern überläßt die Navigation komplett dem Benutzer. Der Objektgröße entsprechend ist der Untersuchungs-Modus dabei am geeignetsten (siehe Abschnitt 2.7).

Hervorhebungsmittel Stufenlose Zustandsänderungen der Objekte bedeuten, daß nur parametrisierbare Darstellungsmittel geeignet sind, die sich auch stufenlos variieren lassen. Im Fall von AGI³LE kämen der Transparenzwert gefüllter Flächen, die Linienstärke und -farbe der Umrandung sowie die Intensität des Schattens in Frage (siehe Abbildung 6.11).

In allen Fällen jedoch ist die tatsächlich unterscheidbare Anzahl der Abstufungen begrenzt. So werden sich Transparenzwerte von 75 bzw. 70 Prozent optisch kaum voneinander unterscheiden. Entsprechendes gilt auch für die Schattenprojektion. Während es in VISDOK nur zwei Schattenintensitäten gab (10 und 50 Prozent), bildet AGI³LE eine stufenlose Palette von Grauwerten ab. In der Praxis allerdings dürfte die Anzahl der klar unterscheidbaren

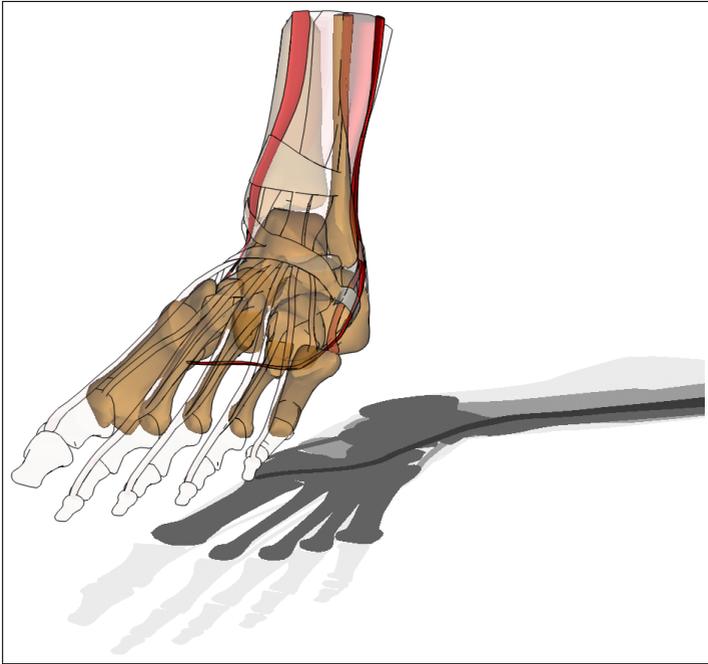


Abbildung 6.11: Beispiel einer Illustration mit Agi³le. Das geometrische Modell enthält Muskeln, Knochen und Sehnen. Der Schwerpunkt liegt auf dem rot gezeichneten Muskel *peroneus longus*. Trotz 100-prozentiger Dominanz des *peroneus longus* ist seine Sichtbarkeit wegen starker Verdeckungen im Bereich einiger ebenfalls dominanter Knochen beeinträchtigt. Die ergänzende Schattendarstellung verbessert daher die Erkennbarkeit des Muskelverlaufs beträchtlich.

Die Zehenknochen und ihre Sehnen dagegen haben mit nur 10 Prozent fast keine Dominanz, dürfen aber nicht weggelassen werden, da erst sie den charakteristischen Gesamteindruck eines Fußes erzeugen.

Graustufen zwischen fünf und sieben liegen. Ebenso ist der sinnvolle Schwankungsbereich für die Liniendicke auf Werte zwischen eins und maximal drei einzugrenzen.

Kameramodell Die zwischen kommunikativen Zielen, grafischen Möglichkeiten und modellabhängigen Freiheitsgraden vermittelnde Instanz ist wiederum das Kameramodell. Wie in VISDOX lassen die Modelle auch hier wichtige Fragen der Darstellung offen. Dazu gehören:

- Flächen- oder Liniendarstellung

Die Objekte werden nur durch ihre Geometrie beschrieben. Es gibt keine gleichzeitige Verwendung beider Zeichenprimitive, so daß die Darstellungsabsicht der Modelle durch die freie Interpretation nicht verfälscht werden kann.

- Transparenz

Die Modelle geben für jedes Objekt eine Farbe an. Diese könnte zwar nicht nur in RGB-Form, sondern auch als RGBA-Wert, also mit Angabe eines Transparenzwertes, erfolgen. Dies wird aber nicht ausgenutzt, so daß wiederum keine Fehlinterpretationen drohen.

- Licht und Schatten

Bereits die schattierte Darstellung der Polygone setzt eine virtuelle Lichtquelle voraus – unabhängig davon, ob sie im Modell enthalten ist oder nicht. Solange der primäre Zweck der Schattierung – die bessere Erkennung von räumlichen Formen – gewährleistet ist, ist die tatsächliche Position der Lichtquelle gleichgültig. Die für den Schattenwurf angenommene virtuelle Lichtquelle muß nicht mit der für die Schattierung identisch sein. Da auch der Schatten letztlich der besseren Formerkennung dient, wäre eventuell sogar eine interaktive Manipulation des Schattenwurfs wünschenswert.

- Kameraperspektive

Angaben über Position und Kamera sind selten in den Modellen enthalten. Ein Festhalten an einer vorgeschriebenen Perspektive hätte auch den Verzicht auf interaktive Navigation zur Folge. Nur für den Fall, daß das gewünscht wird, müssen die Kamerakoordinaten im Modell enthalten sein.

Im Fall von AGI³LE werden folgende Regeln benutzt:

- Dominanzwerte werden auf Transparenzwerte abgebildet, wobei wenig dominante Objekte stark transparent erscheinen.

Damit wird einerseits die Sichtbarkeit der wichtigen Objekte gefördert; andererseits lenkt sich die Aufmerksamkeit klar auf die voll eingefärbten Objekte.

- Alle Objekte erhalten Umriß- und Kantenlinien. Silhouettenlinien werden immer gezeichnet. Von den restlichen Kanten wird ein Prozentsatz gezeichnet, der der Dominanz entspricht.

Das verhindert besonders bei stark transparenten Objekten, daß der Kontext verlorengeht.

- Bei Erreichen eines bestimmten Schwellwertes der Dominanz werden Umrißlinien auch gezeichnet, wenn sie hinter anderen Objekten liegen.

Auf diese Weise wird z. B. der Verlauf von wichtigen Adern oder Knochen unabhängig von ihrer Lage hinter anderen Objekten sichtbar. In Abbildung 6.11 auf Seite 138 betrifft das den Muskel *peroneus longus*.

- Der Schatten bildet die Dominanz analog zur Transparenz über den Grauwert ab.

Dies entspricht der Erwartung, daß transparente Körper einen schwächeren Schatten als undurchsichtige Körper werfen. Zudem macht es neben anderen Vorteilen der Schattens (siehe Abschnitt 6.1.5.4) den trotz Transparenz teilweise verdeckten Verlauf besser erkennbar.

- Grafikobjekte werden nach Dominanz sortiert, so daß transparentere Objekte nach dominanten Objekten gezeichnet werden.

Im Unterschied zu den anderen Regeln ist diese eher technisch motiviert. Der Transparenzeffekt wird dadurch erzielt, daß die transparente Fläche auf eine bereits gezeichnete Fläche aufgetragen wird und sich mit deren Farbe mischt. Würde die vorne liegende transparente Fläche zuerst gezeichnet, so würde sie sich mit dem Hintergrund mischen und die später gezeichnete undurchsichtige Fläche würde gar nicht mehr gezeichnet werden, da sie den Tiefentest nicht besteht (sie liegt ja hinter der transparenten Fläche).

Wie bereits am Beispiel des VISDOX-Systems sichtbar, entstehen auch hier die Ausdrucksmittel, die das System zur Verfügung hat, durch eine *situationsabhängige* Interpretation des vorhandenen Szenenmodells. Die Situation wird im Falle von AGI³LE hauptsächlich durch die aktuelle Fokusstruktur repräsentiert. Das Vorhandensein abgestufter Relevanzwerte erfordert gleichermaßen abgestufte Darstellungsmittel.

Die schon am Beispiel von VISDOX sichtbare hohe Spezifik der intelligenten Kamera und ihres erweiterten Kameramodells führt dazu, daß kein allgemeingültiger Bauplan einer intelligenten Kamera angegeben werden kann. Allerdings lassen sich Einflußfaktoren benennen und aus deren Kenntnis entsprechende Schlußfolgerungen etwa zu Kombinationen von Darstellungsmitteln und möglichen Kompromissen ziehen.

7

Ausblick: Kamerasteuerung in Unterhaltungssoftware

Nachdem sich die beiden Fallstudien mit eher „ernsthaften“ Anwendungen des erweiterten Kameramodells beschäftigen, wendet sich dieses Kapitel den unterhaltenden Systemen zu. In diese Kategorie fallen sowohl die klassischen Computerspiele als auch interaktive Systeme mit didaktischem Hintergrund, bei denen Bezugspersonen auftauchen und die Betätigung spielerischen Charakter aufweist. Umgekehrt nehmen manche Computerspiele für sich in Anspruch, „pädagogisch wertvoll“ zu sein. Wenn also in diesem Abschnitt von Spielen gesprochen wird, sind daher nicht nur die reinen Computerspiele gemeint.

Die Kamerasteuerung in Computerspielen zielt auf die automatische Adaptation grafischer Darstellungsmittel an sich verändernde Spielsituationen. Auch hier beeinflusst der Benutzer durch seine Interaktionen den Fortgang der Handlung. Die sich ergebenden Situationen sind jedoch wesentlich komplexer als im Falle der Interaktiven Technischen Dokumentation und geben Raum für eine wesentlich weitergehende Anwendung filmischer Gestaltungsmittel.

In den Abschnitten 2.8 und 4.2.2 wurden bereits einige auf den Einsatz von Filmtechniken bezogene Unterschiede zwischen technischer Dokumentation

und unterhaltenden Systemen angesprochen. Dabei schienen einige der einfacheren filmischen Gestaltungsmittel durchaus auf Intelligente Interaktive Systeme übertragbar.

Das Gros der filmischen Möglichkeiten aber setzt *Bezugspersonen* voraus, die eher in Spielen als in technischen Dokumentationen anzutreffen sind. Die folgenden Betrachtungen beziehen sich daher hauptsächlich auf Spiele. Allerdings ist zu anzunehmen, daß viele der neu entstandenen Möglichkeiten der Computerspiele durchaus auch für „ernsthafte“ Anwendungen in Frage kommen. Dies ergibt sich aus der in Abschnitt 2.2 festgestellten engen Verwandtschaft von technischen und unterhaltenden interaktiven Systemen.

7.1 Trends in Computerspielen

Seit sich Computer in den späten siebziger Jahren anschickten, den Massenmarkt zu erobern, sind Computerspiele ein Teil des Softwareangebots. In den Anfangsjahren konzentrierte sich die Entwicklung auf eher abstrakte Action- und Geschicklichkeitsspiele à la PACMAN, die entweder auf speziellen Spielekonsolen oder Arcade-Maschinen entstanden; aber letztlich entsprachen auch diese trotz ihrer anderen Ausrichtung und starken Spezialisierung dem technologischen Niveau der zur gleichen Zeit verfügbaren „normalen“ Rechner. Auch zur Zeit scheint sich die Entwicklung bei klassischen Actionspielen wieder mehr auf Konsolen zu konzentrieren, während komplexere Strategiespiele eher auf PCs zu finden sind.

Computerspiele und Innovation Besonders in den letzten Jahren ist zu beobachten, daß sich Computerspiele auch zu einem Motor der technologischen Weiterentwicklung der Computerhardware herauskristallisieren. Längst sind nicht mehr die Erfordernisse der klassischen Büroanwendungen Hauptgrund für Neuanschaffungen, sondern die Anforderungen verschiedenster Multimedia-Anwendungen, darunter auch der Spiele.¹ Besonders die

¹ Selbstverständlich geben auch „rationalere“ Gründe Anlaß zu einer Rechneraufrüstung, wie etwa der, daß die neueste Version des verwendeten Office-Paketes (die leider nur auf dem neuesten und ressourcenhungrigsten Betriebssystem läuft) nun endlich auch die Fehler der Vorversion beseitigt hat.

Weiterentwicklung der Grafikkhardware wird derzeit weniger von den Erfordernissen „professioneller“ Software angetrieben,² sondern von der wachsenden grafischen Detailliertheit und dem komplexen Aufbau dreidimensionaler Spielwelten.

Grafik Die sich durch immer leistungsfähigere Rechnerhardware laufend verbessernden Möglichkeiten für Computerspiele haben zu einer Reihe von neuen Trends in der grafischen Präsentation von Computerspielen geführt. Dazu gehören besonders nicht-photorealistische Darstellungen oder allgemein bildbasierte Verfahren, Skelettbasierte Modelle für die Animationsmodellierung, etc, die den bis dahin vorherrschenden Trend der wachsenden geometrischen Detailliertheit der Modelle abschwächen. Besonders die in neueren Grafikprozessoren vorhandenen flexibel programmierbaren Pixel- und Vertexeinheiten erlauben in Verbindung mit Texturen und *multi-pass rendering* so zahlreiche neue Effekte und Zeichenmethoden, daß ihre Möglichkeiten zur Zeit gerade einmal ansatzweise ausgeschöpft werden.

Simulation Neben der verbesserten Grafikdarstellung gibt es weitere Entwicklungen wie die weg von vordefinierten und vormodellierten Effekten und Charakteren hin zur weitgehenden Simulation (*physically based animation*, z. B. Partikelsysteme, die fließendes Wasser oder Rauch in Echtzeit nachbilden, anstelle vorgefertigter Bildsequenzen). Diese und viele andere Entwicklungen wären vor einigen Jahren aufgrund mangelnder Rechenleistung und Speicherausstattung gar nicht möglich gewesen und zeigen einen Trend zu mehr Realismus bei gleichzeitig verringertem Modellierungsaufwand.³

² Das hängt natürlich auch damit zusammen, daß in der Welt der professionellen Computergrafik vieles von dem schon vorhanden war, was heute im sogenannten *Consumer*-Segment angeboten wird. Insofern ehemals teure Technologie immer preiswerter wurde, zehrt die heute verfügbare Grafikkhardware immer auch von den Früchten der Vergangenheit.

³ An dieser Stelle zeigt sich eine der eingangs erwähnten Gemeinsamkeiten von Spielen und Lernumgebungen: So könnten etwa physikalisch korrekt simulierte Naturphänomene durchaus in Lernsoftware erklärt werden, wobei dem Nutzer anders als im Spiel zusätzlich noch die Kontrolle über diverse beteiligte Parameter (Reibungskoeffizienten, Gewichte, Federkonstanten, Elastizitäten, etc.) überlassen werden könnte, um den Lernerfolg zu fördern.

7.2 Filmtechniken im Computerspiel

In die gleiche Richtung zielt der Einsatz von immer mehr Elementen der Filmsprache anstelle vorab festgelegter Kamerapositionen und -fahrten. Ein Beispiel sind die in vielen Spielen anzutreffenden vorberechneten oder in Echtzeit dargestellten Filmsequenzen zwischen einzelnen Spielabschnitten. Einerseits sind diese Sequenzen wenig mehr als „etwas für's Auge“, haben sie doch mangels interaktiver Eingriffsmöglichkeiten seitens des Spielers keine Auswirkung auf den Spielverlauf. Andererseits zeigen sie durch ihr bloßes Vorhandensein das Bestreben, die Attraktivität des Spieles durch filmische Mittel zu erhöhen.

7.2.1 Abkehr von der Ich-Form

Derzeit wird in einigen Spielen aus dem Action-Adventure-Genre versucht, von der bislang vorherrschenden Ich-Form abzurücken und die Spielfigur bzw. -gruppe aus einer *third-person*-Sicht zu zeigen. Dabei treten allerdings verschiedene neue Probleme auf:

- Zusätzlich zu den eigenen Spielfiguren muß genug von der restlichen Szene und den eventuell darin befindlichen gegnerischen Figuren zu sehen sein, daß z. B. Angriffe (meist durch Anklicken) oder andere Interaktionen gestartet, bzw. einzelne Figuren an bestimmte Orte geschickt werden können.
- Dafür muß der Spieler die Möglichkeit haben, die Kamera zu steuern, ohne gleichzeitig die Figuren zu bewegen; die so veränderten Kameraparameter können entweder dauerhaft oder zeitweise beibehalten werden.
- Wegen der besseren *Übersicht* wird die Kamera meist eine Vogelperspektive einnehmen. Besonders innerhalb von Räumen darf die Kamera aber nicht „an die Decke stoßen“. Ähnliche Sichtbarkeitsprobleme ergeben sich z. B. unter Torbögen.
- Die Kamera soll den Figuren bei ihrer Reise folgen. Bei mehreren Spielfiguren, die sich noch dazu voneinander entfernen können, muß daher festgelegt werden, welche als Bezugsperson der Kamera dienen soll.

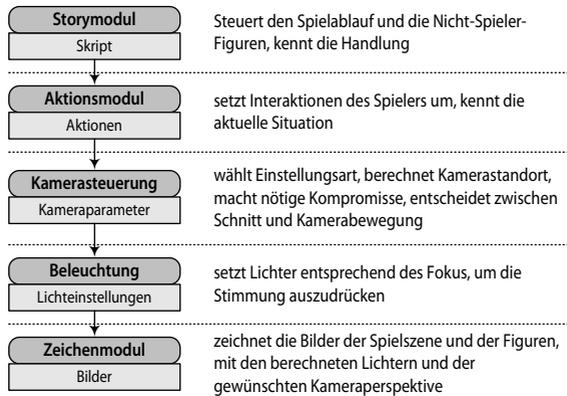


Abbildung 7.1: *Game-Pipeline:* Zusammenhang der unterschiedlichen Module der möglichen Architektur eines Computerspiels im Hinblick auf die Kamerasteuerung und die dafür benötigten Informationen

Eine Kamerasteuerung, die intelligenter agieren soll als die direkt vom Spieler kontrollierte Ich-Form, muß also über Informationen über den Spielverlauf bzw. die Spielsituation verfügen. Diese Informationen bilden die Grundlage für die inhaltliche Beschreibung der Kameraperspektive (siehe auch Abschnitt 5.1). Die intelligente Kamera muß außerdem in der Lage sein, Sichtbarkeitsprobleme selbständig zu erkennen und zu beheben.

7.2.2 Einbettung der Kamerasteuerung

In Abbildung 7.1 wird ein möglicher Informationsfluß skizziert, der die Kamerasteuerung mit den für die Berechnung der Kameraperspektive erforderlichen Daten versorgt.

Das Storymodul ist das Herz der Spielsteuerung. Hier wird die Handlung festgelegt, die Ziele des Spiels definiert. Realisiert wird sie typischerweise durch einen Simulatorekern, der eine Reihe miteinander in Verbindung stehender endlicher Automaten verwaltet, die z. B. Monster und andere nicht vom Spieler selbst kontrollierte Figuren (Nicht-Spieler-Charaktere) repräsentieren. Diese agieren selbst, basierend auf ihrem inneren Zustand, und lösen dabei

Aktionen aus, die wiederum auf ihren eigenen Zustand, andere Automaten bzw. die Spielerfigur einwirken (Bewegen, Schießen, Gegenstand aufnehmen oder benutzen, etc.). Umgekehrt wirkt auch die Spielerfigur auf die Nicht-Spieler-Charaktere ein.

Das Aktionsmodul verwaltet die dabei ausgelösten sichtbaren Aktionen als Konkretisierung der abstrakten Aktionen der Spielsteuerung. Hier wird auch der zeitliche Ablauf der Handlungen festgelegt. Die Interaktionen des Spielers, die ja nicht vom Spiel gesteuert werden, treffen ebenfalls hier ein und werden an die Spielsteuerung weitergeleitet. Im Ergebnis fallen so laufend Ereignisse an, die vom Kameramodul für die zeitliche Steuerung und die Parametrisierung der Einstellungen benötigt werden. Story- und Aktionsmodul definieren die *Spiel-Situation*.

Die Kamerasteuerung setzt die visuellen Ziele, die sich aus der aktuellen Situation ergeben, in Kameraeinstellungen um. Sie trägt außerdem dafür Sorge, daß die generierten Einstellungen auch in ihrer Abfolge zueinander passen. Dabei müssen gegebenenfalls Kompromisse eingegangen werden.

Die visuellen Ziele können zusätzlich durch Beleuchtungseffekte unterstützt werden. Diese haben teilweise ähnlichen Einfluß auf den Betrachterfokus wie die Kameraposition, da die Sichtbarkeit von Objekten auch vom einfallenden Licht abhängt (besonders in dunklen Räumen). Derzeit sind Beleuchtungseffekte⁴ meist bereits in der Szenendefinition enthalten und werden nicht dynamisch mit der Spielsituation variiert.⁵ Neuere Entwicklungen (z. B. DOOM3) ermöglichen unter Rückgriff auf die oben erwähnten Hardwa-

⁴ Für ein realistisches Erscheinungsbild müßten alle Beleuchtungsberechnungen für jeden zu füllenden Bildpunkt einer Grafik ausgeführt werden. Dies ist allerdings sehr rechenintensiv, besonders wenn viele Flächen oder Bilder je Sekunde gezeichnet werden sollen. Daher werden Flächen meist nach GOURAUD schattiert, d. h. Helligkeitsinformation wird über die Flächenkanten hinweg linear interpoliert [GOURAUD 1971]. Die eigentlichen Beleuchtungsberechnungen werden allerdings nur an den Eckpunkten ausgeführt. Da der scheinbare Detailreichtum der Objekte nur über Texturen simuliert wird, können die Objekte mit relativ wenigen Polygonen modelliert werden. Je weniger Polygone und damit Eckpunkte aber vorhanden sind, desto unrealistischer wird ein gouraud-schattiertes Objekt erscheinen. Daher wird versucht, über als *light maps* bezeichnete Hilfstexturen den realistischeren Eindruck einer pixelbasierten Beleuchtung zu erreichen.

⁵ Light maps widerspiegeln die Lichtverteilung in einen bestimmten Zustand der Szene. Bewegen sich Lichtquellen oder schattenwerfende Spielfiguren, so müßten die light maps ebenfalls verändert werden.

retrends dynamische Beleuchtungseffekte wie Schattenvolumen und dynamische *light maps*.

Auf der untersten Ebene finden wir die Grafikengine, die die Szenendaten (Geometrie, Texturen, etc) verwaltet und entsprechend der von den darüberliegenden Modulen definierten Vorgaben darstellt.

7.3 Anforderungen an eine Kamerasteuerung für Computerspiele

Die Arbeit einer intelligenten Kamera unterliegt zahlreichen Einflußfaktoren und Zielen. Je nach Spielgenre (dazu gehören z. B. Adventures, Rollenspiele, Ego-Shooter) werden die im Folgenden aufgelisteten Kriterien unterschiedlich wichtig sein. Es wird daher versucht, keine allzustarke Abhängigkeit vom dem als Motivation dienenden Genre der 3D-Abenteuer mit verschiedenen Spielfiguren zu erzeugen.

Flexibilität Die Auswahl der möglichen Zielvorgaben und ihre Parametrisierung (siehe auch Abschnitt 5.1) legt fest, ob und wie gewünschte Kameraeinstellungen formuliert werden können.

Dabei ist zu beachten, daß insbesondere der Test auf einige Zielvorgaben (z. B. „Objekt A zwischen Objekt B und C“) einen hohen nicht-offensichtlichen Aufwand bedeuten kann. Umgekehrt kann bei wenig ausdrucksstarken Vorgaben ein hoher Aufwand zur Formulierung der gewünschten Einstellungen entstehen.⁶ Als Faustregel kann gelten: „Je mächtiger die Zielvorgabe, desto aufwendiger ihr Test“.

Kohärenz Für den Betrachter unmotivierter Sprünge stören und widersprechen der allgemein erwarteten Kohärenz (siehe Abschnitt 38).

Eine zu starre Kopplung der Kamera an eventuell sprunghafte Veränderungen in der Szene muß vermieden werden. Gegebenenfalls müssen vorübergehende Verletzungen der Zielvorgaben toleriert werden.

⁶ Als Extrembeispiel mag hier die direkte Vorgabe von Koordinaten dienen.

Die nächsten drei Punkte erfordern eine enge Ankopplung der Kamerasteuerung an die höheren Ebenen der Spielsteuerung (siehe Abbildung 7.1 auf Seite 147), da nur dort die entsprechenden Informationen zu erhalten sind.

Lösungsauswahl Nicht immer sind alle Anforderungen umsetzbar, daher muß die Steuerung in der Lage sein, selbständig Kompromisse zu treffen und zu bewerten.

Dies betrifft insbesondere die Sichtbarkeit. Kompromisse können auch vorübergehender Natur sein, wenn etwa bekannt ist, daß ein Objekt bald wieder sichtbar wird. Die Definition von „bald“ hängt hier entscheidend davon ab, wie wichtig das Objekt derzeit ist, bzw. ob daran eventuell Interaktionen ausgeführt werden müssen.

Information Können geforderte Ziele nicht umgesetzt werden, muß die Kamerasteuerung reagieren, um noch eine akzeptable Einstellung zu bestimmen.

Je mehr Informationen über die aktuelle Situation vorliegen, desto leichter lassen sich passende Alternativen finden. Diesem Anliegen entspricht die oben dargestellte Einbettung der intelligenten Kamera in die Spielarchitektur.

Intuition Der Spieler sieht sein Spiel nur durch die Kamera, steuert sie aber nicht selbst.

Dieser Umstand wird im Film zum Erzeugen von Spannung und emotionaler Beteiligung der Zuschauer benutzt, indem etwas wichtiges gerade einmal *nicht* zu sehen ist. Im Spiel jedoch kann er für einige Frustration sorgen. Daher ist es hier noch mehr als beim Film wichtig, daß die Kamera dort ist, wo der Spieler es sich wünscht, sonst „verpaßt“ er das Spiel.

Funktionalität und Kontrolle Anders als im Film ist die Szenenansicht in interaktiven Anwendungen auch Teil einer Benutzungsschnittstelle. Der Nutzer muß also immer in der Lage sein, mit den gerade für ihn wichtigen Interaktionselementen (Spielfiguren, Zielorte für Bewegungen, Gegenstände) zu arbeiten und dazu, wenn nötig, die Kontrolle über die Kamera zurückzuerlangen.

Geschwindigkeit Obwohl die Rechner immer schneller werden, wird es wohl auf absehbare Zeit nie genug Rechenleistung geben, weil gleichzeitig

auch die 3D-Szenen und die verwendeten Algorithmen komplexer werden (siehe Abschnitt 7.1). Da die Kamerasteuerung in Echtzeit ablaufen soll, wird ihr Anteil an der kurzen Zeitscheibe für einen Zyklus *Interaktion – Spielelogik – Szene aktualisieren – Kamera planen – Szene zeichnen* auf absehbare Zeit klein bleiben müssen.

7.4 Architektur einer Kamerasteuerung für Computerspiele

Die im Folgenden skizzierte Architektur versucht die oben besprochenen Anforderungen umzusetzen. Abbildung 7.2 zeigt das Zusammenspiel der einzelnen Komponenten der Kamerasteuerung. Diese zerfällt in zwei Teile: der linke Teil, die Regie, übernimmt alle „höheren“ Kontrollfunktionen, während auf der rechten Seite die tatsächlichen Kamerakonfigurationen berechnet werden.

Das *Aktionsmodul* sendet an die Regie Ereignisse, die die aktuelle Situation und ihr Fortschreiten charakterisieren. Das kann z. B. die Information sein, daß nun ein Dialog zwischen zwei Personen beginnt. Damit bildet die Situation die Grundlage der „mittelfristigen Planung“.

Für die aktuelle Situation können ein oder mehrere Einstellungstypen in einer *Einstellungsbibliothek* vordefiniert sein. Diese enthält Schablonen, die nun mit den entsprechenden Werten gefüllt werden müssen. Das sind insbesondere Angaben zu den beteiligten Akteuren (z. B. ihre Position, Kopfhöhe, Größe, Orientierung).

Zusätzlich muß die *Einstellungsauswahl* auf Abwandlungen der Grundsituationen reagieren. Zum einen sind zu jeder Situation zahlreiche Varianten denkbar; zum anderen helfen leichte Variationen der Parameter, Langeweile zu vermeiden.⁷

⁷ Eine ausreichend große und flexible Einstellungsbibliothek in der Filmsprache ist ähnlich wichtig wie ein guter Wortschatz in der natürlichen Sprache. Einige Filmgenres, wie etwa billig produzierte Sitcoms, kommen mit einem eher geringen Vorrat an Einstellungsarten aus, wogegen Spielfilme das kreative Potential der Kamera weit stärker nutzen.

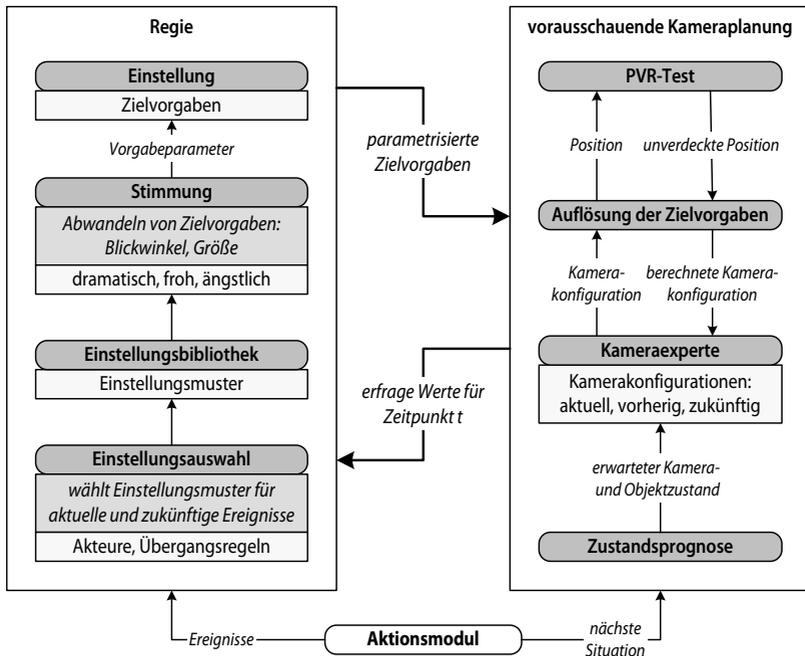


Abbildung 7.2: Architektur der Kamerasteuerung. Die Anweisungen der Regieebene führen zur Auswahl und Parametrisierung von Kamera-Einstellungen, die als Sammlung von *constraints* zur Ermittlung einer geeigneten Kamerakonfiguration dienen. Diese gilt bis auf Widerruf und wird den sich ändernden Positionen der Zielobjekte angepaßt.

Übergangsregeln legen fest, wie der Einstellungswechsel geschieht. Üblicherweise werden hier Schnitte erfolgen, eventuell mit verschiedenen Überblendeffekten. Es sind aber auch gelegentlich andere Stilmittel wie z. B. kurze Kameraschwenks möglich.

Die ausgewählten Einstellungen können in ihren genauen Parametern noch von *Stimmungen* beeinflusst werden. Diese lassen sich besonders gut über die Art der Kamerabewegung transportieren (z. B. Angst: zittrige Bewegung, Freude: Hüpfen, etc). Daneben ließe sich auch die Schnittfrequenz und damit die durchschnittliche Einstellungsdauer variieren, und damit ein Gefühl

von Tempo oder Ruhe erzeugen.

Am Ende liegt eine gewünschte Einstellung in Form von Zielvorgaben einer Kamerakonfiguration vor. Diese Zielvorgaben werden dann im rechten Teil aufgelöst, indem die resultierenden Kamerakoordinaten aus den aktuellen Koordinaten der Bezugsobjekte berechnet werden (siehe Abschnitt 5.1). Für diese Kamerakonfiguration kann mittels des PVR-Verfahrens (*potential visibility regions*, siehe Anhang B) die Sichtbarkeit der Zielobjekte geprüft werden.

Da die Sichtbarkeit durch eine Positionsänderung der Kamera hergestellt wird, benötigt PVR die Angabe eines Gebietes für die Kamera, aus dem alternative Positionen ausgewählt werden. Im einfachsten Fall ist dies eine Kugel um den ursprünglich gewünschten Ort. Es sind aber auch viele andere Varianten möglich, etwa ein flacher Zylinder (Scheibe), wenn die Kamera eine bestimmte Höhe einhalten soll, oder ein dünner Zylinder (Strahl), wenn die Kamera nur ihren Abstand zum Ziel variieren darf. In Abschnitt 38 wurden weitere Fragen der reaktiven und gleichzeitig vorausschauenden Kameraplanung erläutert.

7.4.1 Beispiel Dialog

Im Falle eines Dialogs würde die Regie eine Reihe von Einstellungen vorsehen, zwischen denen dann im Verlauf des Gesprächs hin- und hergeschaltet wird (siehe Abbildung 7.3). Dazu zählt eine einführende Einstellung, die die am Gespräch Beteiligten gemeinsam zeigt, sowie je eine Einstellung, bei der der eine Partner über die Schulter des anderen gezeigt wird. Dazu kommt noch ein sogenannter Gegenschuß (*reaction shot*, auch *reverse shot*), der nur das Gesicht eines der Beteiligten zeigt. Mit diesen fünf vordefinierten Einstellungen wird der gesamte Dialog gefilmt, wobei der jeweils Sprechende im Bild sichtbar sein wird, eventuelle mimische Reaktionen des anderen im Gegenschuß gezeigt werden und hin und wieder auf die Gesamtansicht umgeschaltet werden kann.

Die Kamerasteuerung in Dialogsituationen steht damit vor der Aufgabe, zu Beginn des Dialoges die richtigen Kamerapositionen zu berechnen, die sich aber – sofern die Dialogpartner stillstehen – während des Dialoges kaum ändern werden. Wichtiger ist hier die Kommunikation mit den übergeordneten

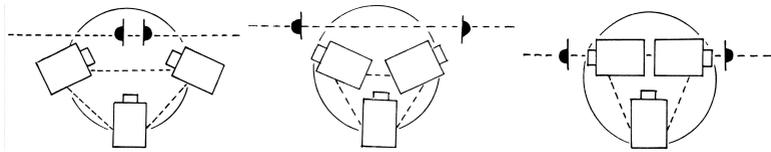


Abbildung 7.3: Kameradreieck für Zweierdialoge aus [Arijon 1976]. Allen Varianten gemeinsam ist die Gesamtansicht am Scheitelpunkt des Dreiecks. Die flache Seite der Figuren symbolisiert die Vorderansicht, der Halbkreis ist hinten. Die gestrichelte Linie (*line-of-interest*, siehe Abschnitt 5.1) verbindet die Figuren und darf von der Kamera nicht überschritten werden.

Links die Variante mit externen Ansichten, die auch Teile der Figur zeigen, die mit dem Rücken zur Kamera steht.

Die Version in der Mitte zeigt die Personen nur einzeln (interne Ansicht). Dabei erscheinen sie größer im Bild als bei externen Ansichten.

Die Variante rechts, die die Personen exakt aus der Perspektive des Gegenübers zeigt und sie direkt in die Kamera blicken läßt, wird seltener angewendet.

Mit der Kombination von internen und externen Ansichten steht damit ein Repertoire von fünf Kameraeinstellungen für den Dialog zur Verfügung.

Aktions- bzw. Storymodulen, die das Timing der verschiedenen Einstellungen ermöglicht.

Als Erweiterung der Grundsituation sind verschiedene Abwandlungen denkbar: So könnte sich z. B. ein dritter Gesprächspartner in den Dialog einschalten oder es könnten wichtige Handlungen im Hintergrund geschehen. Im ersten Fall würden einfach eine oder zwei zusätzliche Kamerapositionen für die dritte Person (analog zu den ersten beiden) berechnet.⁸ Für die Hintergrundhandlung würde die Kameraeinstellung nötigenfalls so variiert werden, daß der Vorgang ins Bild fällt, etwa zwischen den beiden Gesprächspartnern. Den Konventionen des Films folgend wird der Betrachter bereits eine Abweichung von den gewohnten Einstellungsarten als Indiz für eine besondere Situation werten und nach entsprechenden Hinweisen im Bild suchen.

⁸ Für komplette Dreierszenen gibt es noch wesentlich vielfältigere Varianten der räumlichen Anordnung der Personen und dementsprechend kompliziertere Regeln der Kamerapositionierung. Diese müßten dann von Anfang an berücksichtigt werden, weswegen man solche Situationen nicht als Dialogvariante, sondern besser gleich als Dreiergespräch plant.

7.4.2 Beispiel Verfolgung

Häufig soll die Kamera einem oder mehreren Charakteren auf ihrer Reise in der virtuellen Welt folgen. Dabei wird sie versuchen, dem Spieler das gleiche Blickfeld zu bieten wie der Spielfigur, wobei aber – anders als in der Ich-Form – die Spielfigur selbst sichtbar bleibt. Eine Urform dieser Ansicht gab es bereits in den Flugsimulatoren der späten achtziger Jahre, die neben der Innenansicht mit allen Cockpitinstrumenten auch eine Außenansicht boten. Im Unterschied zu erdnahen Szenarien sind in der Luft allerdings keine Sichtbarkeitsprobleme zu erwarten. Genau auf diesem Gebiet liegt aber die Herausforderung einer Außenansicht von Spielfiguren in potentiell engen und verstellten Räumen.

Die Kameraperspektive wird über einfache Zielvorgaben, etwa einen bestimmten Abstand, an die Spielfigur gebunden. Die Blickrichtung wird von der Spielfigur direkt übernommen. Die Problemfälle, die die intelligente Kamera hier abdecken muß, sind:

- errechnete Kameraposition innerhalb von Hindernissen

Dies ist immer dann zu erwarten, wenn die Spielfigur „mit dem Rücken zur Wand steht“. Hier muß der Abstand so angepaßt werden, daß die Kamera immer außerhalb der Verdeckung bleibt. Da das Problem nicht nur an Wänden, sondern auch an anderen Spielfiguren oder im Raum stehenden Gegenständen auftritt, muß sichergestellt werden, daß die resultierende Kamerabewegung kohärent bleibt.

- schnelle Drehungen der Spielfigur

Da der Abstand der Kamera im lokalen Koordinatensystem der Spielfigur gegeben wird, bewegt sich die Kamera auf einer Kreisbahn um die Figur. Die bei schnellen Drehungen entstehenden heftigen Kameschwenks dürften auf Dauer Schwindelgefühle erzeugen und sind zumindest außerhalb der Ich-Form ungewohnt. Abhilfe schafft hier eine verzögerte Nachführung der Kamera – wie bereits im ersten Fall unter zeitweiliger Verletzung der Zielvorgaben.

Damit werfen bereits diese beiden Szenarien einer autonomen Kamerasteuerung viele unterschiedliche Zielkonflikte und Detailprobleme auf. Für die Zukunft zu erwartende reichhaltige Szenen hoher geometrischer Komplexität

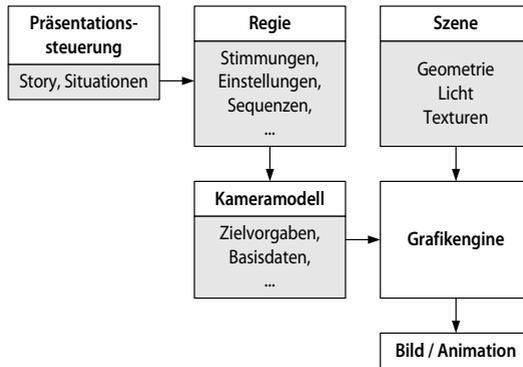


Abbildung 7.4: Zusammenwirken von Kamera und Renderer in Abwandlung von Abbildung 2.2 auf Seite 24. Das Kameramodell vereint hier die unterstützten Zielvorgaben, mit deren Hilfe Einstellungen formuliert werden. Die Auswahl mit Hilfe von Einstellungen und Filmsequenzen ist wesentlich komplexer und unterliegt zahlreichen filmischen Regeln.

verschärfen diese Probleme zusätzlich. Erst deren möglichst vollständige Beherrschung – ohne Ausweichen auf speziell vorbereitete Szenen, die bekannte Konfliktfälle vermeiden – erlaubt eine flexible, modulare Integration der intelligenten Kamera in 3D-Systeme.

7.5 Zusammenfassung

Dieses Kapitel gab einen Ausblick auf eine Kamerasteuerung in spielähnlichen interaktiven Systemen. Im Unterschied zu technischen Anwendungen existiert ein stärkerer Menschbezug und daraus resultierend eine höhere emotionale Beteiligung der Nutzer bzw. Spieler. Zahlreiche Spiele versuchen durch speziell vorberechnete nicht-interaktive Zwischenanimationen Elemente des Films auf das Spiel zu übertragen, ohne solche jedoch in den eigentlichen – interaktiven – Spielablauf zu integrieren.

Der im Abschnitt 5.2 vorgestellte Ansatz einer reaktiven Kamerasteuerung bietet die Möglichkeit, in Echtzeit eine Reihe von Zielvorgaben zur Beschreibung von Kameraperspektiven unter Berücksichtigung der Sichtbarkeit zu prüfen.

Auf dieser Grundlage kann eine intelligente Kamera implementiert werden, deren wesentliches Ausdrucksmittel nicht variable Zeichenstile, sondern Kameraperspektiven bzw. Filmeinstellungen und deren Montage sowie Kamerabewegungen sind (siehe Abbildung 7.4).

Auch in diesem Szenario besteht ein Zusammenhang zwischen den von der Grafikengine gebotenen Ausdrucksmöglichkeiten (hier: den Beschreibungsfomalismen, mit denen Kameraperspektiven spezifiziert und überprüft werden können) und den je nach Anwendung variierenden Situationen.

Selbstverständlich wird sich nicht jede im Film funktionierende Kameratechnik sinnvoll auf interaktive Systeme übertragen lassen. Dieses Kapitel gab daher zunächst zwei Beispiele für häufige, konzeptionell überschaubare Spielsituationen und ihre Darstellung mit filmischen Mitteln. Zukünftige Arbeiten werden zu klären haben, inwieweit die tatsächlich auftretenden Spielsituationen Raum für die unzähligen und variantenreichen Dialogszenarien der Filmkunst geben.

Der wohl wichtigste Ausgangspunkt dieser Arbeit war die Beobachtung, daß viele interaktive Systeme – was ihre grafischen Darstellungen betrifft – als eine Art virtuelle Kamera betrachtet werden können. Deren Eigenschaften werden laufend von anderen Teilen des Systems mit dem Ziel verändert, ein bestimmtes Bild zu erzeugen, welches eine neue Situation symbolisiert bzw. kommuniziert.

Erweiterung der Kamerametapher Die Kamerametapher hat sich seit den Anfangstagen der Computergrafik als nützliche gedankliche Hilfe erwiesen, um computergrafische Abbildungsprozesse zu erklären. Jedoch ist eine virtuelle Kamera nicht nur eine Zusammenfassung von Koordinaten für Position und Blickrichtung. In der Erweiterung der traditionellen Kamerametapher leistet diese Arbeit einen Beitrag zur Computergrafik.

Bereits einfache 3D-Software benötigt mehr als nur diese Angaben, um die Darstellung der 3D-Modelle zu steuern. Die aus der Film- und Fototechnik entlehnte traditionelle Kamerametapher beschreibt mit den geometrischen Parametern der 3D-Darstellung nur einen Teil der Einflußfaktoren, die die Wahl der Kamera, der Linsen und des Films bzw. der Bildsensoren auf das Bildresultat ausübt.

Auch in der Computergrafik wird das berechnete Bild von vielen anderen Faktoren bestimmt, die nicht im 3D-Modell enthalten sind, sondern in anderen Datenstrukturen verwaltet oder ad hoc festgelegt werden. Dazu zählen besonders die von der Grafikengine unterstützten Darstellungs- oder Zeichenstile, die letztlich die Interpretation der Szenendaten beim Zeichnen steuern. Eine Erweiterung der Kamerametapher auch auf diese Aspekte der Bilddarstellung ist daher letztlich eine Konsequenz der unterschiedlichen Bilddarstellungsprozesse in realen Kameras und der Computergrafik.

In interaktiven Systemen kommen zu den auf Einzelbilder bezogenen Faktoren noch Fragen des interaktiven Umgangs mit der Kamera. Hierzu zählen Navigationsmodi, die die Manipulation der Kamera durch den Nutzer beschreiben, aber auch Verfahren, mit denen das 3D-System selbst verschiedene Kameraparameter verändern kann. Speziell Letzteres gibt Raum für Techniken zur effektiveren Steuerung des Betrachterfokus und seiner stärkeren Einbeziehung in das präsentierte Geschehen, wie sie auch im Film genutzt werden.

Diese Techniken bedingen eine Kamerasteuerung auf einer höheren Ebene als der von 3D-Koordinaten. Dazu zählt vor allem ein am gewünschten Bildinhalt orientierter Beschreibungsformalismus für Kameraperspektiven, aus dem sich konkrete Kamerakoordinaten ableiten lassen. Diese Formalismen bilden – in gleicher Weise wie die zeichnerischen Mittel – die Grundlage der Ausdrucksmöglichkeiten einer intelligenten Kamera.

Intelligente Kameras Das erweiterte Kameramodell vereint diese Möglichkeiten mit denen der interaktiven Kamerasteuerung und stellt damit einen Werkzeugkasten der visuellen Darstellungs- und Interaktionsmodi bereit, der von der intelligenten Kamera situationsgerecht koordiniert werden muß.

Die vorgestellten Systeme VISDOK und AGI³LE sind Beispiele für dieses neue Vorgehen. In beiden Systemen werden dem Betrachter grafische Darstellungen präsentiert, deren 3D-Modelle sich verschieden interpretieren lassen. Der Interpretationsspielraum entsteht, weil die benutzte Grafikengine mehr Darstellungsmodi (als zeichnerische Mittel) bereithält als im Modell verlangt werden. So entstehen an verschiedenen Achsen des erweiterten Kameramodells Freiheitsgrade, die zur Anpassung der Darstellung an die aktuelle Situation benutzt werden können.

Leider existiert weder in der technischen Dokumentation noch im Unterhaltungsbereich ein einheitliches Kodierungssystem, welches Gestaltungsmethoden in ihrer Wirkung beschreibt. SCHOLZ bemerkt hierzu:

Das Fehlen eines Gestaltungssystems erschwert aber nicht nur die wissenschaftlich fundierte Herstellung technologischer Bilder, sondern verhindert auch ihre Übertragung in Software-Programme, da keine objektivierbaren Daten und Verweise vorhanden sind. [SCHOLZ 1999]

Damit ist es derzeit unmöglich, allgemeinverbindliche Regeln aufzustellen, nach denen die Verbindung von Darstellungsmitteln (also Parametern im erweiterten Kameramodell) und dem anwendungsabhängigen Situationsvorrat automatisiert werden könnte. Jedoch konnten in dieser Arbeit wichtige Einflußfaktoren in diesem Prozeß identifiziert werden:

- die Szenenmodelle, die das Aussehen und die Geometrie der Szenenobjekte beschreiben und zu deren Darstellung bestimmte Eigenschaften der Grafikengine benötigt werden (z. B. Zeichenstile, Schattierungsmodi, etc),
- die von der Engine insgesamt bereitgestellten Darstellungsmittel und ihre Kombinationsmöglichkeiten
- die damit verbleibenden Freiheitsgrade, d. h. alle Darstellungsmittel, die erkennbar nicht zur „normalen“ Szene gehören können,
- eine durch bestimmte Anforderungen an die entstehenden Bilder beschriebene Menge an Situationen innerhalb der Anwendung,
- Annahmen über etablierte *pop-out*-Effekte und die Erwartungen innerhalb der Zielgruppe,
- sinnvolle Interaktionsparadigmen für die Exploration, d. h. die dynamische Veränderung von Kameraparametern, insbesondere Position, Blickrichtung, etc,
- zusätzliche Anforderungen wie etwa die in multimedialen Dokumentationen geforderte Bild-Text-Kohärenz.

Mit diesen Informationen war es in den Fallstudien möglich, anwendungsspezifische Gestaltungsregeln zu kodieren und in interaktive Systeme zu integrieren. Damit bestätigt sich die eingangs in Abschnitt 1.2 aufgestellte Vermutung,

daß die erweiterte Sicht auf die Kamera als eine intelligente Kamera mit einem erweiterten Kameramodell neue Techniken bei der Vermittlung von Inhalten erschließt.

Die im Ausblick beschriebene Architektur einer intelligenten Kamerasteuerung in Computerspielen setzt diese Gedanken fort. Die Freiheitsgrade bestehen hier nicht so sehr in unterschiedlichen Zeichenstilen oder Licht- und Schatteneffekten, sondern in der Manipulation der Kameraperspektive. Die dafür benötigten Ausdrucksmittel – die Beschreibungsformalismen der Kameraperspektive – werden auch in diesem Szenario von der Grafikengine implementiert und bereitgestellt sowie in das Kameramodell integriert. Die Inhaltsvermittlung wird hierbei durch neuartige filmische Techniken unterstützt, die neben effizienten Algorithmen zur Bestimmung der Sichtbarkeit auch ausgefeilte Planungsstrategien zur Einstellungsauswahl und Montage erfordern.

A

Keramamodelle in 3D- Programmierbibliotheken

Den Keramodellen der verbreiteten Programmierschnittstellen (*application programming interfaces*, APIs) für die 3D-Grafikprogrammierung liegt ein vereinfachtes programmtechnisches Referenzmodell einer Kamera zugrunde. Dieses Modell umfaßt nur die Basisdaten des erweiterten Keramodells¹ und steht in enger Beziehung zu den in der 3D-Grafik üblichen Verarbeitungsschritten für die 3D-Koordinaten, aus denen geometrische Modelle bestehen.

In jeder Programmierumgebung (z. B. OPENGL, PHIGS, Direct3D, sowie den unzähligen anderen Softwareschnittstellen) existieren eigene Basis-Modelle einer virtuellen Kamera. Obwohl sich diese Modelle von denen der anderen APIs unterscheiden, sind sie bis auf kleine Details mathematisch äquivalent. So gibt es z. B. verschiedene Wege, Blickrichtungen zu definieren (Angabe von Blickrichtung oder Blickpunkt).

Dieser Anhang erläutert die wesentlichen Begriffe und Schritte, in denen die 3D-Darstellung abläuft und zeigt dabei, wie das Keramodell diesen Ablauf beeinflußt.

¹ Die anderen Dimensionen des Modells werden (wenn überhaupt) in der 3D-Anwendung implementiert. Auf diese Weise bleiben die APIs schlank und anwendungsneutral.

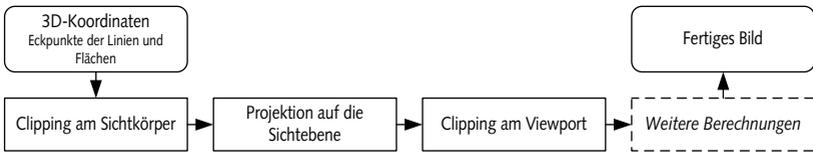


Abbildung A.1: 3D-Pipeline eines Grafik-APIs. Zur Vereinfachung sind hier nur die Stufen enthalten, in denen das Kameramodell eine Rolle spielt.

3D-Pipeline In der 3D-Pipeline (auch Grafik- oder Render-Pipeline) sind sämtliche Verarbeitungsschritte der 3D-Daten von ihrer ursprünglichen Repräsentation in der Anwendung bis zum fertigen Bild zusammengefaßt.

Die 3D-Koordinaten durchlaufen in der Grafikpipeline mehrere Transformationsschritte, in denen sie aus 3D-Weltkoordinaten in die 2D-Koordinaten des jeweiligen Ausgabegeräts (z. B. Bildschirm oder Papier) gebracht werden. Die genauen Parameter dieser Transformationen werden dem Kameramodell entnommen.

Sichtkörper Aus den Parametern des Kameramodells läßt sich der Sichtkörper (siehe Abbildung A.2) konstruieren. Er enthält alles, was für die Kamera sichtbar ist. Alle Punkte des Raumes, die von der aktuellen Position, in der aktuellen Blickrichtung, usw. sichtbar sind, befinden sich also im Sichtkörper. Damit ist zunächst nichts über die Form des Sichtkörpers gesagt. Der Sichtkörper eines menschlichen Auges z. B. dürfte konisch sein, da das Gebiet der Netzhaut, mit dem wir sehen, rund ist. Auch in der Computergrafik und bei Kameras orientiert sich die Form des Sichtkörpers an der der Bildfläche, so daß fast immer Pyramidenstümpfe (bei Perspektivprojektion) oder Quader (bei Parallelprojektion) vorliegen. Die entlang der Blickrichtung verlaufenden Begrenzungsebenen (oben, unten, links und rechts) entstehen zwischen den jeweiligen Projektionsstrahlen.

Koordinatentransformationen Die 3D-Koordinaten liegen als sogenannte Weltkoordinaten vor, die sich auf einen frei gewählten, aber für die Anwendung festen Nullpunkt beziehen (z. B. die Position der Sonne in einem

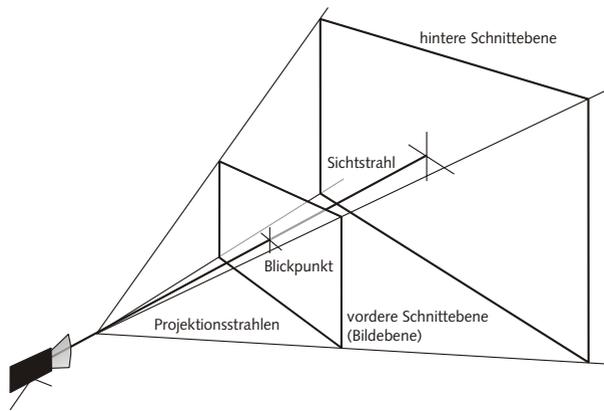


Abbildung A.2: Sichtkörper für Perspektivprojektion. Der Sichtstrahl trifft senkrecht auf die Bildebene. Diese ist wie das Bildschirmfenster rechteckig und begrenzt den Sichtkörper zur Kamera hin. Die hintere Schnittebene ist parallel zur vorderen. Die seitlichen und waagerechten Schnittebenen ergeben sich aus den jeweiligen Projektionsstrahlen von der Kamera zu den Ecken der Bildebene.

virtuellen Planetarium). Sie beschreiben die in der 3D-Szene enthaltenen Flächen, Linien und Punkte (auch Primitive genannt) durch die Angabe ihrer End- bzw. Eckpunkte.²

Im ersten Schritt werden alle Objekte aus Weltkoordinaten in Kamerakordinaten überführt. Die Kameraparameter, die in diesen Schritt eingehen, sind:

Orientierung Die Orientierung ist gegeben durch die Blickrichtung (die auch als Differenz zwischen Zielpunkt und Position berechnet werden kann) und durch einen Drehwinkel um diese Blickrichtung bzw. eine Richtung in der 3D-Welt, die im Bild senkrecht sein soll. Normalerweise steht die Kamera aufrecht, , h. Oben in der Welt ist auch Oben im Bild.

² Besonders im CAD-Bereich werden auch verschiedene Arten von gekrümmten Flächen und Kurven (*splines, subdivision surfaces, etc*) benutzt, aber auch die werden vor der Verarbeitung in der abgebildeten Pipeline mit Liniensegmenten bzw. Drei- oder Vierecken approximiert.

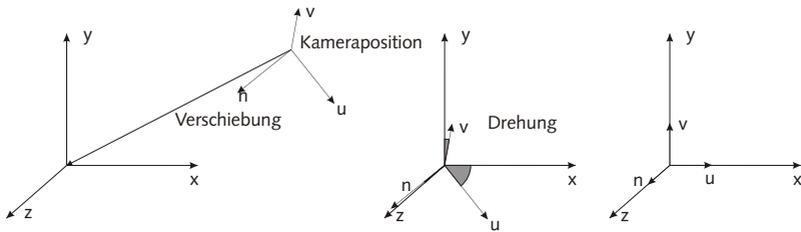


Abbildung A.3: Überführung von Welt- in Kamerakoordinaten. Der erste Schritt ist eine Verschiebung, die die Koordinatenursprünge zusammenführt. Der zweite Schritt dreht die Achsen so, daß sie übereinstimmen. Es wird von jeweils rechtwinklig aufeinanderstehenden Einheitsvektoren ausgegangen.

Position Wird in Weltkoordinaten angegeben und bestimmt zusammen mit der Orientierung das Kamera-Koordinatensystem, welches durch Drehung und Verschiebung aus dem Weltkoordinatensystem entsteht.

Diese beiden Angaben beschreiben das Koordinatensystem der Kamera, zu dem nicht nur ein Koordinatenursprung gehört, sondern auch die Richtung der Einheitsvektoren in diesem Koordinatensystem. Im einfachsten Fall fällt die Position der Kamera mit dem Ursprung der Weltkoordinaten zusammen, und die Kamera blickt aufrecht stehend in die negative Z-Richtung.

Die Transformation (siehe Abbildung A.3) wird aus Effizienzgründen in einem Schritt durchgeführt. Dabei ist T die Transformationsmatrix für die Verschiebung und R die der Drehung (siehe Formel A.1). Der Vektor $\vec{v} = (v_x v_y v_z)^T$ steht für die Oben-Richtung, $\vec{n} = (n_x n_y n_z)^T$ für die Blickrichtung und \vec{vrp} ist die Kameraposition. Die Formeln benutzen die entsprechenden homogenen Koordinaten.

$$T = \begin{pmatrix} 1 & 0 & 0 & -vrp_x \\ 0 & 1 & 0 & -vrp_y \\ 0 & 0 & 1 & -vrp_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad R = \begin{pmatrix} u_x & u_y & u_z & 0 \\ v_x & v_y & v_z & 0 \\ n_x & n_y & n_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (A.1)$$

Die resultierende Transformation V ergibt sich dann aus der Multiplikation beider Teiltransformationen (siehe Formel A.2), was geometrisch einer Nacheinanderausführung der Verschiebung und der Rotation entspricht.

$$V = RT = \begin{pmatrix} u_x & u_y & u_z & -\vec{u} \cdot \overrightarrow{vrp} \\ v_x & v_y & v_z & -\vec{v} \cdot \overrightarrow{vrp} \\ n_x & n_y & n_z & -\vec{n} \cdot \overrightarrow{vrp} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{A.2})$$

Die in den Formeln benutzten (normalisierten) Vektoren \vec{v} und \vec{n} sowie der Punkt \overrightarrow{vrp} stammen direkt aus dem Kameramodell. Der Rechts-Vektor als dritter Einheitsvektor des Kamerakoordinatensystems muß aus der gegebenen Blickrichtung und dem Oben-Vektor berechnet werden (siehe Formel A.3). Dabei wird ausgenutzt, daß auch im Kamerakoordinatensystem die Einheitsvektoren senkrecht aufeinander stehen.³ Da der meistgenutzte Oben-Vektor, die Y-Achse, im Allgemeinen nicht senkrecht zur Blickrichtung ist, muß ein neuer Vektor berechnet werden, der diese Bedingung erfüllt.

$$\vec{u} = \vec{v} \times \vec{n} \quad (\text{A.3})$$

$$\vec{v}' = \vec{n} \times \vec{u} \quad (\text{A.4})$$

Nachdem die Primitive nun in Kamerakoordinaten vorliegen, werden Objekte, bzw. deren Teile, die nicht im aktuellen Sichtkörper (siehe unten) liegen, verworfen oder abgeschnitten. Alle noch übrigbleibenden Primitive werden auf die Bildebene projiziert und erscheinen im Bild, sofern sie nicht von anderen Primitiven verdeckt werden. Die Kameraparameter, die in diesen Schritt eingehen, sind:

Projektionsart Am gebräuchlichsten sind Projektionen, bei denen der vom Blickpunkt kommende Projektionsstrahl senkrecht auf die Bildebene trifft. Das menschliche Auge wie auch reale Kameras liefern *Perspektivprojektionen*, bei denen weiter entfernt liegende Objekte verkleinert

³ Dies setzt voraus, daß Blickrichtung und Oben-Vektor nicht zusammenfallen, also $\vec{n} \neq \vec{v} \cdot \alpha$ für beliebige α ist. Anderenfalls spannen sie keine Ebene auf.

werden und alle Projektionsstrahlen an der Kameraposition zusammenlaufen.

Für technische Zeichnungen und damit auch für zahlreiche 3D-Anwendungen werden oft *Parallelprojektionen* benutzt. Die Parallelprojektion ist ein Grenzfall der Perspektivprojektion für eine unendlich weit entfernte Kamera. Die Projektionsstrahlen verlaufen parallel und der Abbildungsmaßstab der Objekte ist unabhängig von ihrer Entfernung.

Breite und Höhe des Sichtfeldes Diese Parameter legen die horizontale und vertikale Ausdehnung des Sichtkörpers fest.

Für Perspektivprojektionen gibt man meist den Öffnungswinkel der Kamera an. Dabei gilt: $h = 2 d \tan \frac{\alpha}{2}$ (mit h : Höhe des Bildes in der Bildebene, d : Entfernung der Bildebene von der Kamera und α : vertikaler Öffnungswinkel). Im Falle rechteckiger Bildflächen gibt es einen waagerechten und einen senkrechten Öffnungswinkel, die sich entsprechend des Seitenverhältnisses der Bildfläche unterscheiden. Weitwinkelkameras haben einen größeren Öffnungswinkel als Teleobjektive. Für reale Kameras wird anstelle des Winkels meist die Brennweite angegeben, aus der der Öffnungswinkel erst bei Kenntnis weiterer Werte (insbesondere der Größe des Negativs) berechnet werden kann.

Bei Parallelprojektionen wird die Breite und Höhe des Sichtkörpers direkt angegeben.

Tiefe des Sichtfeldes Dieser Parameter hat keine direkte Entsprechung für reale Kameras. Man unterscheidet in der Computergrafik eine vordere und eine hintere Grenzebene. Bei realen Kameras wäre die äußere Linse des Objektivs als vordere Begrenzung anzusehen (auch wenn in dieser Entfernung nichts scharf dargestellt wird). In der Computergrafik wird so verhindert, daß auch Objekte *hinter* der Kamera gezeichnet werden. Die hintere Begrenzung gibt es für reale Kameras nicht. In der Computergrafik wird sie für das hardwareunterstützte Beseitigen verdeckter Flächen (Tiefenspeicher, *z-buffer*) benötigt. Da der Tiefenspeicher nur begrenzte Auflösung (=Wertebereich) haben kann, muß auch die Tiefe des Sichtkörpers begrenzt werden.

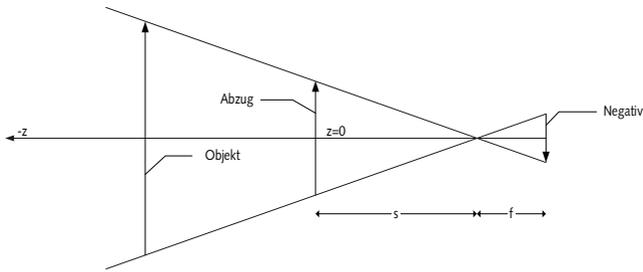


Abbildung A.4: Perspektivprojektion in der Fotokamera (nach [Rogers und Adams 1990]). Das Negativ innerhalb der Kamera steht kopf, das Bild dagegen aufrecht.

Punkte aus dem Objektraum werden nach folgender Gleichung auf das Bild projiziert:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/s & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \\ 1 - z/s \end{pmatrix} \text{ bzw. } \begin{pmatrix} \frac{x \cdot s}{s - z} \\ \frac{y \cdot s}{s - z} \\ 0 \\ 1 \end{pmatrix} \quad (\text{A.5})$$

Zu guter Letzt liegen die Objekte in 2D-Koordinaten vor und werden in ein Bildschirmfenster gezeichnet. Dabei kann das Zeichnen auf einen Teil des Fensters beschränkt werden. Dem entspricht bei Fotoapparaten die Möglichkeit, nicht das ganze Negativbild zu vergrößern, sondern eine Ausschnittsvergrößerung anzufertigen. Wie schon in 3D, werden auch hier wieder Teile der Szene, die außerhalb des sichtbaren Bereiches liegen, verworfen.

Diese Beschreibung des 3D-Zeichnens beschränkte sich auf die Berechnungen, die mit den 3D-Koordinaten angestellt werden. Diese beeinflussen im Ergebnis hauptsächlich die Sichtbarkeit und die sichtbare Form der Objekte.

Der tatsächliche Zeichenprozeß muß in den verschiedenen Pipeline-Stufen noch zahlreiche andere Berechnungen ausführen, insbesondere für die Beleuchtung. Da für die Beleuchtungssimulationen allerdings keine Daten aus

dem Kameramodellen der aktuellen Grafik-APIs herangezogen werden,⁴ können diese Berechnungen hier unbeachtet bleiben.

⁴ Denkbar wäre z. B. Lichtempfindlichkeit eines Filmes mit Möglichkeit der Über- und Unterbelichtung und damit verbundenen Ein- und Ausblendeffekten.

B

Potential Visibility Regions

Das Verfahren der *potential visibility regions* (PVR) erlaubt unter Ausnutzung der Grafikhardware, Regionen zu finden, von denen aus eine freie Sicht auf einen oder mehrere gewünschte Zielpunkte garantiert werden kann. Damit ist es möglich, vorab zu prüfen, ob eine bestimmte Kameraposition Verdeckungen des Zielobjektes hervorruft oder nicht, d. h. PVR erlauben die *Evaluierung einer Kameraperspektive*.

Zunächst soll der Fall betrachtet werden, daß nur *ein* Zielpunkt vorliegt; die Erweiterung auf mehrere Ziele folgt im Anschluß.

Der Algorithmus geht davon aus, daß bereits vorab eine gewünschte Kameraposition berechnet wurde. Dies kann z. B. durch Spezifikation und Auswertung bestimmter Zielvorgaben geschehen (siehe Abschnitt 5.1). Diese lassen jedoch die Frage der Sichtbarkeit offen, die deswegen gesondert beantwortet werden muß.

Sichtbarkeitsregion Das Verfahren basiert auf der Idee einer Sichtbarkeitsregion, einer Umgebung der errechneten Wunschposition. Diese Region ist unscharf begrenzt, d. h. sie repräsentiert einen Wert zwischen 0 und 1 mit dem Maximum an der Wunschposition. Das Verfahren versucht nun, die Position mit dem höchsten Wert innerhalb der Sichtbarkeitsregion zu finden, die nicht von Hindernissen verdeckt ist. Die Grundidee des Verfahrens ist,

dabei die Szene in der entgegengesetzten Blickrichtung, also die Wunschposition aus Sicht des Zielobjektes zu betrachten.

Zuerst werden alle potentiellen Hindernisse – die gesamte Szene mit Ausnahme des Zielobjektes – mit Schwarz gezeichnet. Dadurch liegt außerdem die Tiefeninformation aller potentiellen Hindernisse im Tiefenspeicher (*z-buffer*). Anschließend wird mit hellen Farben die Sichtbarkeitsregion gezeichnet; im einfachsten Fall sind das Kugeln um die berechnete Wunschposition. Dabei wird mit kleinen, weißen Kugeln begonnen. Jeder mögliche Bildpunkt der Kugeln muß den Tiefentest bestehen, also näher an der Kamera (oder in diesem Fall, dem Zielobjekt) liegen als jedes der Hindernisse. Damit werden also immer nur Punkte außerhalb von Hindernissen gezeichnet. Wenn beim Zeichnen dieser Kugeln einzelne Pixel gesetzt werden, und damit den Tiefentest bestehen, wird dies in einem zusätzlichen Puffer, dem Stencilpuffer¹ vermerkt. Der Stencilpuffer wird dafür so konfiguriert, daß einmal markierte Bildpunkte nicht mehr verändert werden können, selbst wenn sie im Ergebnis des Tiefentestes mit einem der Kamera näheren Objekt überschrieben werden müßten.

Wenn also nach und nach immer größere, dunklere Kugeln um die Wunschposition gezeichnet werden, werden nur noch dort Punkte der größeren Kugeln gesetzt, wo nicht bereits Punkte einer kleineren Kugel (die näher am Zielpunkt liegen) den Tiefentest gegen die Hindernisse bestanden haben. Wenn man also die kleinen Kugeln mit der hellsten Farbe zeichnet und die größeren Kugeln immer mehr abdunkelt, erhält man ein Bild wie in Abbildung B.1. In der Anwendung findet dieses zusätzliche Zeichnen der Szene nicht auf dem sichtbaren Bildschirm, sondern in unsichtbaren *off-screen* Puffern der Grafikhardware statt.

Optimalposition Nach dem Zeichnen der Kugeln wird das entstandene *Farbbild* aus dem Bildspeicher ausgelesen und die Stelle mit der hellsten Farbe

¹ Der Stencilpuffer ist ein weiterer unsichtbarer Teil des Bildspeichers. Er ist ähnlich organisiert wie der Tiefenpuffer und speichert je Pixel einen ganzzahligen Wert ab. Dieser kann benutzt werden, um einzelne Bildpunkte zu markieren. Im dazugehörigen Stenciltest kann ähnlich dem Tiefentest entschieden werden, ob ein bestimmter Bildpunkt gesetzt werden darf oder nicht. Darüber hinaus sind weitere Operationen möglich, die den Stencilpuffer für verschiedenste grafische Effekte und Operationen prädestinieren.



Abbildung B.1: Sichtbarkeitsregion nach dem PVR-Algorithmus. Die Szenenobjekte sind als dunkle Schatten auf den konzentrischen Kugeln um die gewünschte Kameraposition zu erahnen. Der hellste Punkt im Bild ist der Punkt, der am nächsten an dieser Wunschposition liegt und noch die Sichtbarkeitsanforderungen erfüllt.

gesucht. Mit Hilfe des an gleicher Stelle im *Tiefenspeicher* gefundenen Tiefenwertes ist eine Rückprojektion des Punktes in den 3D-Raum möglich, so daß am Ende die gesuchte Koordinate vorliegt.

Mehrere Zielobjekte Wenn die Sichtbarkeit von mehr als einem Objekt zu prüfen ist, muß dieser Vorgang mehrmals wiederholt werden. Dabei wird die Szene von mehreren Zielobjekten aus gezeichnet, ohne daß zwischendurch der Tiefenpuffer gelöscht wird. Dabei akkumulieren sich die jeweils gezeichneten Schatten der verdeckenden Körper. Wenn nun die konzentrischen Kugeln der Sichtbarkeitsregion gezeichnet werden, müssen sie den Tiefentest nicht nur gegen die Schatten aus Sicht eines, sondern aller Zielpunkte bestehen, was die Menge der infragekommenden Positionen im Zweifelsfall verringern wird.

Angepaßte Sichtbarkeitsregionen Abgesehen von der Situation, daß einfach die nächstmögliche Stelle von der Optimalposition aus gesucht wird,

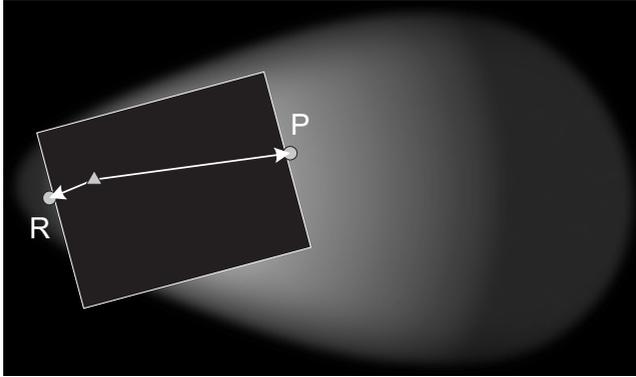


Abbildung B.2: Die Kameraregion wird in Richtung der aktuellen Kamerabewegung verzerrt, damit Positionen, die vor der Kamera liegen, eher berücksichtigt werden als solche dahinter. Hier wird der weit vorne liegende Punkt einem knapp hinter der Kamera liegenden vorgezogen, weil die Helligkeit der Regionen nach hinten wesentlich schneller abfällt.

sind auch Fälle denkbar, in denen die Kamera z. B. eine bestimmte Höhe nicht verlassen oder eine bestimmte Richtung beibehalten soll. Derartige Zusatzanforderungen können leicht durch eine von der Kugel abweichende Sichtbarkeitsregion formuliert werden (siehe Bild B.2).

Soll etwa die Kamera nur ihren Abstand zum Zielobjekt, nicht aber ihren Blickwinkel darauf ändern dürfen, so können anstelle der Kugeln auch immer länger werdende dünne Zylinder gezeichnet werden, die entlang der Verbindungslinie von der Kamera zum Zielobjekt orientiert sind.

Geschwindigkeit Seine hohe Geschwindigkeit erreicht das Verfahren dadurch, daß die Zeichenvorgänge mit voller Hardwareunterstützung stattfinden,² sofern die Grafikkarte das Stenciling unterstützt.³ Dies ist auf moder-

² Außerdem spielen Effekte wie Beleuchtung oder Texturen, die normalerweise beim Zeichnen aktiviert sind und ebenfalls Zeit kosten, keine Rolle, da es ja nur um die Schatten geht.

³ Der Stencil-Puffer ist eine professionelle Technik aus der OpenGL-Welt. Frühere Versionen von Microsoft's 3D-API Direct3D (vor DirectX 7) kannten kein Stenciling. Daher gibt es teilweise noch Grafikkarten, die Stenciling nicht in Hardware unterstützen und auf langsames Zeichnen umschalten, wenn die Anwendung Stenciling fordert.

nen Grafikkarten der Fall. Potentielle Engpässe ergeben sich eher beim Zurücklesen des Bildes, um die hellste Stelle darin zu finden.⁴ Diesen kann man leicht dadurch begegnen, daß die Auflösung des unsichtbaren Hilfspuffers gegenüber der normalen Bildschirmauflösung stark verringert wird. Werte von 256×256 haben sich als ausreichend erwiesen. Dies beschleunigt nicht nur das Zurücklesen der Daten, sondern – in geringerem Maße – auch das Zeichnen. Zusätzlich kann zur Vermeidung allzu komplexer Geometrie auf niedriger detaillierte Hilfsobjekte zurückgegriffen werden, was die teilweise recht aufwendige Geometrieverarbeitung stark beschleunigt.

Die zeitliche Komplexität des Verfahrens ist $O(n \times m)$ mit n als Anzahl der Polygone in der (eventuell vereinfachten) Szene und m als Anzahl der Zielobjekte. Die Größe des Hilfspuffers, die davon abhängende Zeit zum Finden des hellsten Punktes sowie für die Rückprojektion sind konstant.

Ein weiterer Vorteil neben der Geschwindigkeit ist, daß die Kamerasteuerung keine Geometrie der Szenenobjekte verarbeiten oder auch nur kennen muß. Damit eignet sie sich besonders für Szenarien, in denen die Kamerasteuerung mit einer existierenden Grafikengine zusammenarbeiten muß, die die Szenendaten selbst verwaltet (was normalerweise den Datenaustausch erschwert und entweder Zeit kostet oder eine aufwendige doppelte Datenhaltung mit zumindest Konsistenzproblemen erfordert).

⁴ Hierbei bildet nicht wie so oft die Speicherbandbreite zwischen Hauptspeicher und Grafikkarte den Flaschenhals, sondern mangelnde Optimierung dieser eher seltenen Operation seitens der Grafiktreiber. Dieser Punkt verliert allerdings aufgrund besserer Treiber zunehmend an Bedeutung.

Literaturverzeichnis

- ANDRÉ, ELISABETH, J. MUELLER und T. RIST (1996). *WIP/PPP: Automatic generation of personalized multimedia presentations*. In: *Proceedings of the Fourth ACM Multimedia Conference (MULTIMEDIA'96)*, S. 407–408. ACM Press.
- ARIJON, DANIEL (1976). *Grammar of the Film Language*. Communication Arts Books, Hastings House Publishers New York.
- BARES, WILLIAM H. und J. C. LESTER (1999). *Intelligent multi-shot visualization interfaces for dynamic 3D worlds*. In: MAYBURY, MARK, Hrsg.: *Proceedings of the 1999 International Conference on Intelligent User Interfaces (IUI-99)*, S. 119–126, N.Y., New York. ACM Press.
- BARES, WILLIAM H., S. THAINIMIT und S. McDERMOTT (2000). *A model for constraint-based camera planning*. In: *Smart Graphics. Papers from the 2000 AAAI Spring Symposium (Stanford, March 20–22, 2000)*, S. 84–91, Menlo Park. AAAI Press.
- BECKHAUS, STEFFI (2002). *Dynamic Potential Fields for Guided Exploration in Virtual Environments*. Dissertation, Otto-von-Guericke-Universität Magdeburg.
- BECKHAUS, STEFFI, F. RITTER und T. STROTHOTTE (2000). *Cubicalpath – Dynamic Potential Fields for Guided Exploration in Virtual Environments*. In: *Proceedings Pacific Graphics 2000 (Hong Kong, China, October 2000)*, S. 387–395, Los Alamitos. IEEE Computer Society.
- BECKHAUS, STEFFI, F. RITTER und T. STROTHOTTE (2001). *Guided Exploration with Dynamic Potential Fields: The Cubical Path System*. IEEE Computer Graphics Forum, 20(4):201–210.

- BITOMSKY, HARTMUT (1972). *Die Röte des Rots von Technicolor. Kinorealität und Produktionswirklichkeit*. Neuwied/Darmstadt.
- BLINN, JAMES F. (1988). *Jim blinn's corner: Where am I? what am I looking at?*. IEEE Computer Graphics and Applications, 8(4):76–81.
- BORDEGONI, MONICA, G. FACONTI, M. T. MAYBURY, T. RIST, S. RUGGIERI, P. TRAHANIAS und M. WILSON (1997). *A Reference Model for Intelligent Multimedia Presentation Systems*. Computer Standards & Interfaces: The International Journal on the Development and Application of Standards for Computers, Data Communications and Interfaces.
- BUCKLEY, COLM (1994). *Bézier curves for camera motion*. In: *Proceedings of the 2nd Irish Computer Graphics Workshop*, Coleraine.
- BUTZ, ANDREAS (1997). *Anymation with CATHI*. In: *Proc. of the 14th National Conference on Artificial Intelligence and the 9th Conference on the Innovative Applications of Artificial Intelligence (AAAI/IAAI'97)*, S. 957–962, Menlo Park. AAAI Press.
- BUTZ, ANDREAS, A. KRÜGER, P. OLIVIER und M. ZHOU, Hrsg. (2001). *1st International Symposium on Smart Graphics*, New York, USA. Association for Computing Machinery (ACM), ACM Press.
- BUTZ, ANDREAS und A. KRÜGER (1996). *Lean modeling - the intelligent use of geometrical abstraction in 3d animations*. In: WAHLSTER, WOLFGANG, Hrsg.: *Proceedings of the 12th European Conference on Artificial Intelligence (ECAI 96), Budapest, Hungary, August 12–16, 1996*, S. 246–250. John Wiley & Sons, Ltd.
- CHRISTIANSON, DAVID B., S. ANDERSON, L. HE, D. SALESIN, D. WELD und M. COHEN (1996). *Declarative camera control for automatic cinematography*. In: *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, S. 148–155.
- COOK, ROBERT L. und K. E. TORRANCE (1981). *A reflectance model for computer graphics*. In: *SIGGRAPH 81 Proceedings*, Bd. 15, S. 307–316.
- COREN, STANLEY, J. ENNS und L. M. WARD (1994). *Sensation and Perception*. Harcourt College Publishers, 4th Aufl.
- DEUSSEN, OLIVER, J. HAMEL, A. RAAB, S. SCHLECHTWEG und T. STROTHOTTE (1999). *An Illustration Technique Using Intersections and Skeletons*.

- In: *Proceedings of Graphics Interface '99*, S. 175–182, Toronto. Canadian Computer-Human Communications Society.
- DRUCKER, STEVEN M. (1994). *Intelligent Camera Control for Graphical Environments*. Dissertation, MIT Media Lab.
- DUNN, ROBERT M. und B. HERZOG (1977). *Status report of the Graphics Standards Planning Committee of ACM/SIGGRAPH*. *Computer Graphics*, 11(3):I–19 + II–117.
- ECK, MATTHIAS, T. DE ROSE, T. DUCHAMP, H. HOPPE, M. LOUNSBERY und W. STUETZLE (1995). *Multiresolution analysis of arbitrary meshes*. In: COOK, ROBERT, Hrsg.: *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, S. 173–182. ACM SIGGRAPH, Addison Wesley. held in Los Angeles, California, 06–11 August 1995.
- EDWARDS, DEBORAH M. und L. HARDMAN (1989). *Lost in cyberspace: Cognitive mapping and navigation in a hypertext environment*.
- ELM, W. C. und D. D. WOODS (1985). *Getting lost: A case study in user interface design*. In: *Proceedings of the Human Factors Society 29th General Meeting*, S. 927–931, Santa Monica.
- EVERITT, CASS (2001). *Interactive order-independent transparency*. http://developer.nvidia.com/view.asp?IO=Interactive_Order_Transparency.
- FEINER, STEVEN (1985). *APEX: An Experiment in the Automated Creation of Pictorial Explanations*. *IEEE Computer Graphics & Applications*, 5(11):117–123.
- FOLEY, J. D., A. VAN DAM, S. K. FEINER, J. F. HUGHES und R. L. PHILIPS (1994). *Grundlagen der Computergraphik: Einführung, Konzepte, Methoden*. Addison-Wesley, Bonn.
- FUNGE, JOHN, X. TU und D. TERZOPOULOS (1999). *Cognitive modeling: Knowledge, reasoning and planning for intelligent characters*. In: *Proceedings of SIGGRAPH 1999*, Computer Graphics Proceedings, Annual Conference Series, S. 29–38, Los Angeles. ACM SIGGRAPH, ACM Press, New York.
- GOTTSCALK, STEFAN, M. C. LIN und D. MANOCHA (1996). *OBBTree: A hierarchical structure for rapid interference detection*. *Computer Graphics*, 30(Annual Conference Series):171–180.

- GOURAUD, HENRY (1971). *Continuous shading of surfaces*. IEEE Transactions on Computers June 1971, C-20(6):223–228.
- HALPER, NICOLAS, R. HELBING und T. STROTHOTTE (2001). *A camera engine for computer games: Managing the trade-off between constraint satisfaction and frame coherence*. In: CHALMERS, A. und T.-M. RHYNE, Hrsg.: *Proceedings of EUROGRAPHICS*, Manchester.
- HALPER, NICOLAS und P. OLIVIER (2000). *CAMPLAN: A Camera Planning Agent*. In: *Smart Graphics. Papers from the 2000 AAAI Spring Symposium (Stanford, March 20–22, 2000)*, S. 92–100, Menlo Park. AAAI Press.
- HARTMANN, KNUT (2002). *Text-Bild-Beziehungen in multimedialen Dokumenten: Eine Analyse aus Sicht von Wissensrepräsentation, Textstruktur und Visualisierung*. Dissertation, Otto-von-Guericke-Universität Magdeburg, Magdeburg. Published at Shaker-Verlag Aachen, ISBN 3-8322-0206-4.
- HARTMANN, KNUT, S. SCHLECHTWEIG, R. HELBING und T. STROTHOTTE (2002). *Knowledge-Supported Graphical Illustration of Texts*. In: *Advanced Visual Interfaces AVI 2002*, Trento, Italien. ACM Press, New York.
- HEIDRICH, WOLFGANG, P. SLUSALLEK und H.-P. SEIDEL (1997). *An image-based model for realistic lens systems in interactive computer graphics*. In: *Graphics Interface '97*, S. 68–75.
- HELBING, RALF, K. HARTMANN und T. STROTHOTTE (1998). *Dynamic Visual Emphasis in Interactive Technical Documentation*. In: RIST, T., Hrsg.: *Proc. of the Workshop on Combining AI and Graphics for the Interface of the Future*, S. 82–90, Brighton, UK.
- HELBING, RALF und B. PREIM (1998). *Interaction facilities and high-level support for animation design*. In: *Computational Visualization: Graphics, Abstraction, and Interactivity*, Kap. 15, S. 259–276. Springer Verlag, Berlin · Heidelberg · New York.
- HICKETHIER, KNUT (1993). *Film- und Fernsehanalyse*. Metzler, Stuttgart · Weimar.
- HOPPE, AXEL (1998). *Validierung und Nachbearbeitung von gerenderten Bildern*. Dissertation, Otto-von-Guericke-Universität Magdeburg.

- HOPPE, HUGUES (1996). *Progressive meshes*. In: RUSHMEIER, HOLLY, Hrsg.: *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, S. 99–108. ACM SIGGRAPH, Addison Wesley. held in New Orleans, Louisiana, 04–09 August 1996.
- KARP, PETER und S. FEINER (1993). *Automated presentation planning of animation using task decomposition with heuristic reasoning*. In: *Proceedings of Graphics Interface '93*, S. 118–127, Toronto, Ontario, Canada. Canadian Information Processing Society.
- KASPER, ROBERT T. (1989). *A Flexible Interface for Linking Applications to PENMAN's Sentence Generator*. In: *Proc. of the DARPA Workshop on Speech and Natural Language*, S. 153–158, Marina del Rey, USA. USC/Information Sciences Institute.
- KOLB, CRAIG, D. MITCHELL und P. HANRAHAN (1995). *A realistic camera model for computer graphics*. *Computer Graphics*, 29(Annual Conference Series):317–324.
- KOSARA, ROBERT, S. MIKSCH und H. HAUSER (2002). *Focus + context taken literally*. *IEEE Computer Graphics and Applications*, 22(1):22–29.
- KRÜGER, ANTONIO und T. RIST (1995). *Since Less is often More: Methods for Stylistic Abstractions in 3D-Graphics*. In: *Electronic Proceedings of the ACM Workshop on Effective Abstractions in Multimedia*.
- LATOMBE, JEAN-CLAUDE (1991). *Robot Motion Planning*. Kluwer Academic Publishers.
- MANN, WILLIAM C. und S. A. THOMPSON (1987). *Rhetorical structure theory: Description and construction of text structures*. In: KEMPEN, G., Hrsg.: *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*, S. 85–95. Nijhoff, Dordrecht.
- MARCHAND, ÉRIC und N. COURTY (2000). *Image-based virtual camera motion strategies*. In: FELS, S. und P. POULIN, Hrsg.: *Graphics Interface*, S. 69–76, Montreal.
- MONACO, JAMES (1980). *Film verstehen*. Rowohlt, Hamburg.
- MUSGRAVE, F. KENTON (1987). *A realistic model of refraction for computer graphics*. Master's Thesis, Univ. of California at Santa Cruz.

- OLIVIER, PATRICK, J. PICKERING, N. HALPER und P. LUNA (1999). *Visual composition as optimisation*. In: *Proc. of the AISB Symposium on AI and Creativity in Entertainment and Visual Art*, S. 22–30, Edinburgh, UK.
- PHONG, BUI-TUONG (1975). *Illumination for computer generated pictures*. CACM June 1975, 18(6):311–317.
- POTMESIL, MICHAEL und I. CHAKRAVARTY (1982). *Synthetic image generation with a lens and aperture camera model*. ACM Transactions on Graphics, 1(2):85–108.
- PREIM, BERNHARD und F. RITTER (2002). *Techniken zur interaktiven Hervorhebung von Objekten in medizinischen 3D-Visualisierungen*. In: SCHULZE, THOMAS, S. SCHLECHTWEG und V. HINZ, Hrsg.: *Simulation und Visualisierung 2002*, S. 187–200, Magdeburg. Otto-von-Guericke Universität, SCS European Publishing House.
- RAAB, ANDREAS (1998). *Techniken zur Interaktion mit und Visualisierung von geometrischen Modellen*. Dissertation, Otto-von-Guericke-Universität Magdeburg.
- RAAB, ANDREAS und M. RÜGER (1996). *3D-Zoom: Interactive visualisation of structures and relations in complex graphics*. In: B. GIROD, H. NIEMANN, H.-P. SEIDEL, Hrsg.: *3D Image Analysis and Synthesis*, S. 125–132, St. Augustin. infix.
- RIST, THOMAS., E. ANDRÉ und J. MÜLLER (1997). *Adding Animated Presentation Agents to the Interface*. In: *Proc. of the 1997 International Conference on Intelligent User Interfaces*, S. 79–86.
- RITTER, FELIX, B. PREIM, O. DEUSSEN und T. STROTHOTTE (2000). *Using a 3D Puzzle as a Metaphor for Learning Spatial Relations*. In: *Proceedings Graphics Interface 2000 (Montreal, Canada, 15.–17. Mai 2000)*, S. 171–178. Morgan Kaufmann Publishers.
- ROGERS, BRIAN J. und M. GRAHAM (1979). *Similarities between motion parallax and stereopsis in human depth perception*. Vision Research, 22:299–300.
- ROGERS, DAVID F. und J. A. ADAMS (1990). *Mathematical Elements for Computer Graphics*. McGraw-Hill, New York, N. Y., 2nd Aufl.

- RÖSNER, DIETMAR und M. STEDE (1994). *Generating Multilingual Documents from a Knowledge Base: The TECHDOC Project*. In: *Proc. of the 15th International Conference on Computational Linguistics (COLING'94)*, S. 339–346.
- RÜGER, MICHAEL (1998). *Zoom-Techniken zur Benutzerunterstützung*. Dissertation, Otto-von-Guericke Universität Magdeburg.
- SCHIRRA, JÖRG R. J. und M. SCHOLZ (1998). *Abstraction versus Realism: Not the real question*. In: STROTHOTTE, THOMAS, Hrsg.: *Computer Visualization – Graphics, Abstraction, and Interactivity*, Kap. 22, S. 379–401. Springer-Verlag.
- SCHLECHTWEG, STEFAN (1999). *Interaktives wissenschaftliches Illustrieren von Texten*. Dissertation, Otto-von-Guericke-Universität Magdeburg.
- SCHLECHTWEG, STEFAN und T. STROTHOTTE (1999). *Illustrative Browsing: A New Method of Browsing in Long On-line Texts*. In: SASSE, M. A. und C. JOHNSON, Hrsg.: *Computert Human Interaction. Proceedings of INTERACT'99 (Edinburgh, Sep., 1999)*, S. 466–473, Amsterdam. IOS Press.
- SCHOLZ, MARTIN (1999). *Visualisierung in technischen Dokumentationen – prozessuale Bildzeichen*. Dissertation, FB Design, Kunst und Medienpädagogik der Bergischen Universität-Gesamthochschule (BUGH) Wuppertal.
- SELIGMANN, D. und S. FEINER (1991). *Automated Generation of Intent-Based 3D Illustrations*. *Computer Graphics*, 25(4):123–132.
- SHOEMAKE, KEN (1992). *Arbball: A user interface for specifying three-dimensional orientation using a mouse*. In: *Proceedings Graphics Interface 1992*, S. 151–156, Vancouver, Canada.
- STEWART, A. JAMES und M. S. LANGER (1997). *Towards accurate recovery of shape from shading under diffuse lighting*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):1020–1025.
- STROTHOTTE, THOMAS (1998). *Computational Visualisation. Graphics, Abstraction and Interactivity*. Springer Verlag, Berlin – Heidelberg – New York. ISBN: 3-540-63737-0.
- STROTHOTTE, THOMAS, B. PREIM, A. RAAB, J. SCHUMANN und D. R. FORSEY (1994). *How to Render Frames and Influence People*. *Computer Graphics*

- Forum, 13(3):455–466. Proceedings of Eurographics'94 (Oslo, September 16–20, 1994).
- STROTHOTTE, THOMAS und S. SCHLECHTWEG (2002). *Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation*. Morgan Kaufman, San Francisco. ISBN: 1-55860-787-0. 472 pages.
- TOMLINSON, BILL, B. BLUMBERG und D. NAIN (2000). *Expressive autonomous cinematography for interactive virtual environments*. In: SIERRA, CARLES, G. MARIA und J. S. ROSENSCHEIN, Hrsg.: *Proceedings of the 4th International Conference on Autonomous Agents (AGENTS-00)*, S. 317–324, NY. ACM Press.
- TREISMAN, ANNE (1985). *Preattentive processing in vision*. Computer Vision, Graphics, and Image Processing, 31(2):156–177.
- WAHLSTER, WOLFGANG, E. ANDRÉ, W. FINKLER, H.-J. PROFITLICH und T. RIST (1993). *Plan-based Integration of Natural Language and Graphics Generation*. Artificial Intelligence, 63:387–427.
- WARE, COLIN (2000). *Information Visualization – Perception for Design*. Morgan Kaufmann.
- WARE, COLIN und S. OSBORN (1990). *Exploration and virtual camera control in virtual three-dimensional environments*. Interactive 3D graphics, S. 175–184.
- WEIDENMANN, BERND (1993). *Informierende Bilder*. In: *Wissenserwerb mit Bildern: instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen*, Kap. 1, S. 9–59. Verlag Hans Huber, Bern · Göttingen · Toronto · Seattle.
- WILLMANN, THOMAS (2002). *Spiel/film*. <http://www.telepolis.de/deutsch/kolumnen/wil/11958/1.html>.
- WINSTON, PATRICK HENRY (1975). *The Psychology of Computer Vision*, ed.. McGraw Hill, 1975.
- WOLLEN, PETER (1969). *Signs and Meaning in the Cinema*. Secker & Warburg/BFI, London.

Abbildungsverzeichnis

1.1	Strahlverlauf in camera obscura und an dünner Linse	9
2.1	Erweitertes Kameramodell	23
2.2	Abgrenzung von Kameramodell und Renderer	24
2.3	Freiheitsgrade der Kamera aus [DRUCKER 1994]	25
2.4	Kamerapositionen in einer 3D-Szene	33
3.1	Beispiel einer Akzentuierung aus [HOPPE 1998]	39
3.2	Kontinuierlicher Zoom	42
3.3	3D-Fisheye-Zoom	42
3.4	Defokussieren durch Weglassen	46
3.5	Defokussieren durch Transparenz	47
3.6	Einfluß der Transparenzwerte	48
3.7	Sichtbarmachung durch Drahtgitter	49
3.8	Einfluß der Kantenwinkel auf die Drahtgitterdarstellung	50
3.9	Abstraktion mit Flächen	54
4.1	Verhältnis von Filmsprache und Filmstory	60
4.2	Einstellungsgrößen nach ARIJON	61
4.3	Unterschied zwischen Kamerafahrt und Zoom (1)	69
4.4	Unterschied zwischen Kamerafahrt und Zoom (2)	70
4.5	Unterschied zwischen Kamerafahrt und Zoom 3)	70
5.1	Kamerapositionierung bei Zielvorgaben für Blickwinkel	82
5.2	Probleme der reaktiven Kameraplanung (1)	88
5.3	Probleme der reaktiven Kameraplanung (2)	89

5.4	Roadmap und Zellzerlegung	90
5.5	Navigationsobjekte	92
5.6	Planungsmodi	97
5.7	Kollisionstest	98
5.8	Hüllkörper	98
5.9	Ausweichmanöver beim Umgehen von Hindernissen	99
5.10	Planungsalgorithmus	100
5.11	Konstruktion einer Bézierkurve aus Liniensegmenten	102
5.12	Kollision mit einer Bézierkurve	103
5.13	Parametrierung einer Bézierkurve	104
6.1	Eine durch VISDOK generierte Präsentation	110
6.2	Aufgaben innerhalb eines interaktiven Präsentationssystems	112
6.3	Systemarchitektur von VISDOK	113
6.4	Menüdarstellung in VISDOK	117
6.5	Schattierte und Drahtgitterdarstellung in VISDOK	128
6.6	Situationen in VISDOK	129
6.7	Selektion und Hervorhebung in VISDOK	131
6.8	Das AGI ³ LE-System	134
6.9	Linguistische Analyse in AGI ³ LE	135
6.10	Dominanzpropagation	136
6.11	Beispiel einer Illustration mit AGI ³ LE	138
7.1	Game-Pipeline	147
7.2	Architektur der Kamerasteuerung	152
7.3	Kameradreieck für Zweierdialoge	154
7.4	Zusammenwirken von Kamera und Renderer	156
A.1	Vereinfachte 3D-Pipeline	164
A.2	Sichtkörper für Perspektivprojektion	165
A.3	Überführung von Welt- in Kamerakoordinaten	166
A.4	Perspektivprojektion	169
B.1	Sichtbarkeitsregion nach dem PVR-Algorithmus	173
B.2	Anpassung der Kameraregion	174