

Vorlesungsscript

Medientechnik
Berufsakademie
Studiengang Informationstechnik

Oliver Mark
o.mark@de.ibm.com
0172-2931376

Inhaltsverzeichnis

1. Einleitung	Seite 6
1.1 Übersicht	Seite 6
1.2 Definitionen	Seite 7
1.2.1 Medium	Seite 7
1.2.2 Darstellungsdimension	Seite 8
1.2.3 Multimedia	Seite 8
1.3 Multimediale Anwendungsfelder	Seite 9
1.4 Medienobjekte	Seite 10
1.4.1 Typen von Medienobjekten	Seite 11
1.4.2 Beziehungen zwischen Medienobjekten	Seite 12
2. Audio	Seite 16
2.1 Grundlagen	Seite 16
2.1.1 Physik	Seite 16
2.1.2 Die menschliche auditive Wahrnehmung	Seite 16
2.1.3 Komponenten der auditiven Informationsvermittlung	Seite 17
2.2 Computerrepräsentation von Audio	Seite 17
2.3 MIDI	Seite 18
2.4 Sprache als Audio	Seite 19
3. Text	Seite 21
3.1 Typographie	Seite 21
3.1.1 Schrift	Seite 21
3.1.2 Dokumente	Seite 23
3.2 Textverarbeitung	Seite 26
3.2.1 Satzsysteme	Seite 26
3.2.2 WYSIWYG/Desktop-Publishing	Seite 27
3.3 Standards	Seite 27
3.3.1 SGML	Seite 27
3.3.2 Rich Text Format (RTF)	Seite 28
3.3.3 Postscript	Seite 28
4. Bild und Grafik	Seite 30
4.1 Grundlagen	Seite 30
4.1.1 Der Begriff Bild	Seite 30
4.1.2 Der Begriff Grafik	Seite 30
4.1.3 Bitmapgrafiken	Seite 30
4.1.4 Vektorgrafiken	Seite 31
4.2 Die Bildsprache	Seite 31
4.2.1 Die Blickrichtung	Seite 31
4.2.2 Das Bildformat	Seite 32
4.2.3 Die Aufteilung der Bildfläche	Seite 32
4.2.4 Gestaltungselemente Linien und Flächen	Seite 33
4.3 Die Farbe	Seite 34
4.3.1 Eigenschaften von Auge und Licht	Seite 34
4.3.2 Licht- und Körperfarben	Seite 35
4.4 Das Licht	Seite 35
4.4.1 Lichtquellen im Computer	Seite 36
4.4.2 Lichtrichtung	Seite 36
4.5 Die Perspektive	Seite 37
4.5.1 Zentralperspektive	Seite 37
4.5.2 Parallelprojektion	Seite 37
4.6 Bildspeicherformate	Seite 37
4.6.1 Graphics Interchange Format (GIF)	Seite 37
4.6.2 Tagged Image File Format (TIFF)	Seite 38
4.6.3 Bitmap (BMP)	Seite 38

4.7 Grafikspeicherformate	Seite 39
4.8 Bildanalyse	Seite 39
4.8.1 Bildverbesserung	Seite 39
4.8.2 Mustererkennung	Seite 39
4.8.3 Szenenanalyse	Seite 39
4.9 Bildmerkmale	Seite 39
4.9.1 Farbe	Seite 39
4.9.2 Textur	Seite 40
4.9.3 Kanten	Seite 40
4.10 Bilderkennung	Seite 40
4.10.1 Formatierung	Seite 40
4.10.2 Konditionierung	Seite 40
4.10.3 Markierung	Seite 40
4.10.4 Gruppierung	Seite 40
4.10.5 Extraktion	Seite 40
4.10.6 Abbildung	Seite 41
5. Video	Seite 42
5.1 Visuelle Repräsentation	Seite 42
5.2 Videokodierung	Seite 42
5.2.1 RGB Signal	Seite 43
5.2.2 YUV-Signal	Seite 43
5.2.3 YIQ-Signal	Seite 43
5.3 Videoformate	Seite 43
5.4 Fernsehformate	Seite 44
6. Computeranimationen	Seite 45
6.1 Klassische Animationstechniken	Seite 46
6.1.1 Dynamik in stehenden Bildern	Seite 46
6.1.2 Der Zeichentrickfilm	Seite 46
6.1.3 Der Animationsfilm	Seite 47
6.2 Die Computeranimation	Seite 48
6.2.1 2D-Technik	Seite 48
6.2.2 3D-Technik	Seite 49
7. Datenkompression	Seite 52
7.1 Grundlagen	Seite 52
7.2 Kodierungsverfahren	Seite 53
7.2.1 Lauflängenkodierung	Seite 53
7.2.2 Nullunterdrückung	Seite 53
7.2.3 Vektorquantisierung	Seite 53
7.2.4 Pattern Substitution	Seite 54
7.2.5 Diatonic Encoding	Seite 54
7.2.6 Statistische Kodierung	Seite 54
7.2.7 Huffman-Kodierung	Seite 54
7.2.8 Arithmetische Kodierung	Seite 55
7.2.9 Transformationskodierung	Seite 55
7.2.10 Prädiktion oder relative Kodierung	Seite 55

1. Einleitung

Medientechnik als Vorlesung beschäftigt sich im wesentlichen mit Inhalten, die allgemeinsprachlich unter Multimedia zusammengefaßt werden. Allerdings ist der Begriff Multimedia in der Öffentlichkeit stark inflationär mißbraucht.

Es werden kaum noch Computer angeboten, die nicht eine Multimedia-Station darstellen. Nahezu jede Anwendung von CD-ROM verspricht ein Multimedia-Erlebnis. Multimedia scheint ein Qualitätsmerkmal geworden zu sein. Für Bill Gates wird "Multimedia wichtiger als alles, was wir bis jetzt gemacht haben". Prognosen sehen ein Marktvolumen von bis zu 700 Milliarden Dollar für die nächsten 25 Jahre, Tageszeitungen berichten über Cybersex, Begriffe werden gesellschaftsfähig, ohne daß sie definiert sind.

Ebenso wie der Begriff PC für Personal-Computer niemals vollständig definiert wurde und sich daher später auf jedem Home-Computer wiederfand, droht der noch so junge Begriff Multimedia für alles und nichts zu stehen, nur ein nichtssagendes Anhängsel zu werden.

Ziel dieses Kapitels ist, den Begriff Multimedia deutlich von bisher dagewesenen Vokabeln zu trennen. Viele (multi) Medien (media) sind keineswegs automatisch ein Multimedia-Erlebnis.

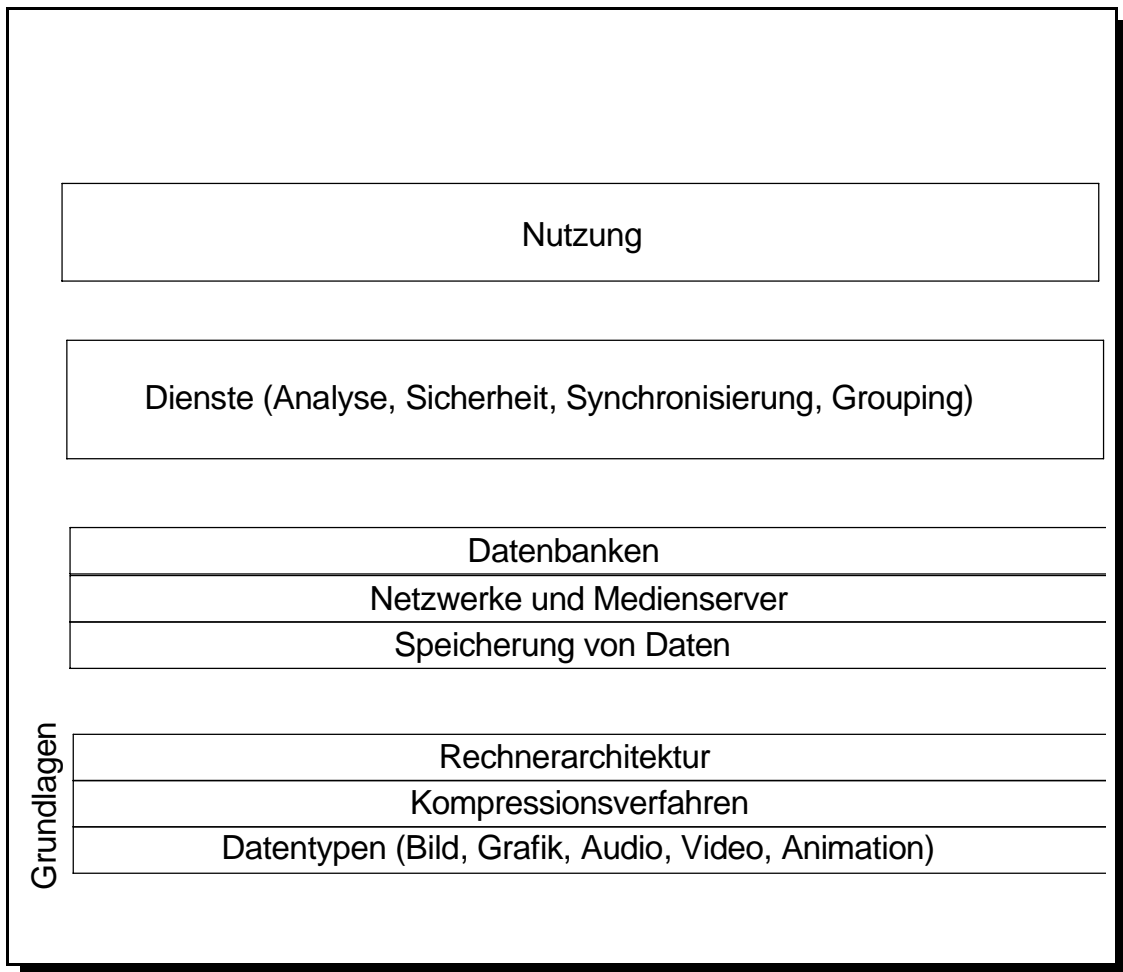
1.1 Übersicht

Medientechnik bzw. Multimedia bezieht seinen Ursprung aus verschiedenen traditionellen Fachrichtungen:

- Telekommunikation: Audio-Übertragung über das Telefon.
- Unterhaltungselektronik: CD-Laufwerke, Hi-Fi-Anlagen
- Studioteknik: Audio und Videobearbeitung für Fernseh- und Rundfunkanstalten
- Verlagswesen: Publikationen in versch. Medien

Strukturieren läßt sich die Medientechnik, oder auch Multimedia folgendermaßen:

- Grundlagen: Betrachtung der einzelnen Medien und technische Implementierungen
- System: Speicherung und Kommunikation der Daten
- Dienste: Aufbereitung und Bearbeitung der Daten
- Nutzung: Art und Gestaltung von Anwendungen, sowie die Schnittstelle zwischen Anwender und System.



1.2 Definitionen

Der Begriff "Multimedia" entsteht durch Komposition der beiden Teile Multi und media. Der Wortbestandteil multi kommt von dem lateinischen Wort "multus" und bedeutet viel, vielfach. Media ist der englische Begriff für Medien (singular: das Medium). Medien sind nach der Brockhaus Enzyklopädie "Vermittlungssysteme für Informationen aller Art, ihre Funktion ist der Transport von Inhalten". Informationen werden in Form von Texten, Bildern, Graphiken, Sprache oder Musik dargestellt.

1.2.1 Medium

Als Medium bezeichnet man ein Mittel zur Verbreitung und Darstellung von Informationen. Beispiele von Medien sind Text, Grafik, Bild, Sprache, Musik. Genauso können nach diesem Verständnis auch Wasser und Luft als Medium aufgefaßt werden.

Eine Unterscheidung kann folgendermaßen vorgenommen werden:

- Das Perzeptionsmedium läßt sich von den menschlichen Sinnen ableiten. Die zentrale Frage lautet: "Wie nimmt der Mensch die Information auf?" Primäre Unterscheidung ist hier "Hören" und "Sehen". Medien in diesem Bereich sind somit Musik, Geräusche (auditiv) und Text, Bild (visuell).
- Das Repräsentationsmedium wird durch die unterschiedliche rechnerinterne Darstellung von Informationen gekennzeichnet. Zentrale Frage ist somit: "Wie wird die Information

im Rechner kodiert ?" So können Text in ASCII, Bilder in JPEG, TIF etc, Audio in PCM, MP3 und Kombinierte Daten z.B. in MPEG abgelegt werden.

- Das Präsentationsmedium bezieht sich auf die Hilfsmittel und Geräte für die Ein- und Ausgabe von Informationen. Zentrale Frage: "Wie wird eine Information in einen Rechner eingelesen oder durch diesen ausgegeben?" Medien wie Papier, Bildschirm und Lautsprecher sind Ausgabemedien, Kamera, Mikrophon, Tastatur sind Eingabemedien.
- Das Speichermedium unterscheidet die verschiedenen Datenträger. Die zentrale Frage: "Wie wird eine Information gespeichert?" Beispiele sind Papier, Festplatte, Band, CD.
- Das Übertragungsmedium charakterisiert die verschiedenen Informationsträger die eine kontinuierliche Übertragung von Daten ermöglichen. Beispiele sind hier Draht, Glasfaser, Funk, Infrarot, Vakuum.

1.2.2 Darstellungsdimension

Jede Darstellung verfügt über eine oder mehrere Darstellungsdimensionen. So besitzt ein Bildschirm zwei Dimensionen (X- & Y-Achse). Die Zeit als Dimension ist gerade im Bereich von Multimedia von zentraler Bedeutung; daher unterscheidet man im wesentlichen zwei Arten von Medien in ihrem zeitlichen Zusammenhang:

- Diskrete Medien: Medien wie Text oder Grafik, die zeitlich unabhängig sind. Die Gültigkeit der Daten hängt nicht von einer zeitlichen Folge ab
- Kontinuierliche Medien: Die Information setzt sich stets aus dem eigentlichen Wert und ihrer zeitlichen Darstellung zusammen. Die Gültigkeit der Daten hängt von ihrer zeitlichen Reihenfolge ab.

1.2.3 Multimedia

Da eine Erklärung wie aus dem Brockhaus nicht befriedigend erscheint, soll eine Definition gefunden werden, die diese neue Technologie von bisherigen unterscheidbar macht. Die charakteristischen Kriterien einer Definition, die Ralf Steinmetz in seinem Buch "Multimedia-Technologie: Einführungen und Grundlagen" nennt, werden einzeln betrachtet und schrittweise zur vollständigen Definition zusammengefügt.

Wie schon in 1.2.2 beschrieben, gibt nicht jede Kombination von Medien eine multimediale Anwendung.

Nach Steinmetz kann erst von einer multimedialen Anwendung gesprochen werden, wenn sie mindestens ein kontinuierliches und ein diskretes Medium verarbeiten kann. Also kann definiert werden:

Ein Multimedia-System ist durch die Verarbeitung von Informationen gekennzeichnet, die in mindestens einem kontinuierlichen und einem diskreten Medium kodiert sind.

Eine Kombination eines kontinuierlichen und eines diskreten Mediums ist z.B. ein Karaoke-Video. Es besteht aus Audio- und Bildinformationen als kontinuierliches Medium und den Texteinblendungen als diskretes Medium. Dies entspricht aber nicht einem intuitiven Begriff einer Multimedia-Anwendung, da die Medien zwar verschiedene Darstellungsformen benutzen (Bild, Ton, Text), sie aber nie wieder getrennt bearbeitbar sind. Die Medien sind gemeinsam auf dem Magnetband gespeichert und damit starr miteinander gekoppelt. Nachträgliches Verändern nur eines Mediums ist nicht möglich. In einer Multimedia-Anwendung müssen die eingesetzten Medien unabhängig voneinander sein. So werden z.B. bei Spielfilmen für Gehörlose die Untertitel separat über Videotext übertragen. Der Videotext kann bei Bedarf einzeln verändert und z.B. in eine andere Sprache übersetzt werden.

Die Definition muß daher folgendermaßen erweitert werden:

Ein Multimedia-System ist durch die Verarbeitung von unabhängigen Informationen gekennzeichnet, die in mindestens einem kontinuierlichen und einem diskreten Medium kodiert sind.

Die Unabhängigkeit der verschiedenen Medien erfordert etwas, was die Verarbeitung koordiniert bzw. die Medien kombiniert. Hierfür ist der Rechner das ideale Werkzeug. Das System kann dabei in einem gewissen Rahmen frei programmiert werden, also genügt die Aufnahme/ das Abspielen durch einen programmierbaren Videorekorder nicht. Die Medien sollen integriert werden, sie sollen eine gemeinsame Funktion erfüllen. Dies wird erreicht, indem zwischen ihnen zeitliche, räumliche oder inhaltliche Beziehungen erstellt werden. Das bedeutet z.B. für eine Textverarbeitung mit Tabellenkalkulation und Grafikdarstellung, daß sie dieser Bedingung erst genügt, wenn die Änderung eines Eintrages in einem Tabellenfeld sich auf die Grafik auswirkt und entsprechende Werte im Text beeinflusst. Diese Beziehung heißt inhaltliche Integrität.

Die Definition wächst dadurch zu:

Ein Multimedia-System ist durch die rechnergesteuerte, integrierte Verarbeitung von unabhängigen Informationen gekennzeichnet, die in mindestens einem kontinuierlichen und einem diskreten Medium kodiert sind.

Der allgemeine Begriff der Verarbeitung von Informationen läßt sich bzgl. der Datenverarbeitung in Informations-Eingabe, Informations-Verarbeitung und Informations-Ausgabe (EVA-Prinzip) gliedern. Im Bereich der Medien entsprechen dem Erzeugung, Manipulation, Darstellung und Speicherung von Informationen.

Der Begriff der Manipulation schließt ein charakteristisches Merkmal einer Multimedia-Anwendung mit ein, welches vielfach im Sprachgebrauch vernachlässigt, aber dennoch meist vorausgesetzt wird: die Interaktion. Für eine "richtige" Multimedia-Anwendung ist es bedeutend, daß der Benutzer den Ablauf und das Erscheinungsbild der Anwendung beeinflussen kann. Er muß in einem gewissen Rahmen selbst entscheiden können, welche Medienobjekte zu welchem Zeitpunkt in welcher Art und Weise dargestellt werden. Da davon ausgegangen werden kann, daß immer mehr Rechner an Kommunikationsnetze angeschlossen sind, wäre eine Einschränkung von Multimedia-Anwendungen auf lokale Rechner ein Rückschritt. Daher kann der Begriff Kommunikation, die dann zwischen den verschiedenen Rechner möglich und sinnvoll ist, zusammen mit den oben genannten in die Definition aufgenommen werden.

Ein Multimedia-System ist durch die rechnergesteuerte, integrierte Erzeugung, Manipulation, Darstellung, Speicherung und Kommunikation von unabhängigen Informationen gekennzeichnet, die in mindestens einem kontinuierlichen und einem diskreten Medium kodiert sind.

1.3 Multimediale Anwendungsfelder

Seit jeher versuchte der Mensch, Kommunikation nicht ausschließlich über ein einzelnes Medium zu vollziehen. Sei es die Kunst der Gestik, Mimik oder allein die Intonation, die als zusätzliches Attribut einer Nachricht beigemischt werden kann. Jedoch erst die neuesten Computer ermöglichen, dieses komplexe Zusammenspiel nachzubilden, da erst sie die technischen Möglichkeiten besitzen, datenintensive Übertragungsmedien wie Ton oder Bild

zu verarbeiten. Innerhalb der Vielschichtigkeit multimedialer Anwendungsfelder können folgende Klassen unterschieden werden:

- **Unterhaltung:** Die Unterhaltungsbranche bildet den weitaus größten Marktanteil von Multimedia-Applikationen, insbesondere in Form von elektronischen Spielen.
- **Kommunikation:** In diesem Bereich vollzieht sich der Wandel von rein textueller elektronischer Post zu einer multimedialen, d.h. der eigentliche Text wird mit Sprach- und Videosequenzen erweitert. Die Entwicklung der Telefonnetze zu digitalen Netzen ermöglicht auch hier die Einbindung von multimedialen Komponenten wie das Bildtelefon und Videokonferenzen.
- **Aus- und Weiterbildung:** Der Einsatz von CBT(Computer-Based-Training)-Anwendungen in schulischer und betrieblicher Ausbildung wird durch zwei bedeutende Argumente gerechtfertigt. Zahlreiche Untersuchungen und Erfahrungen haben gezeigt, daß durch multimedial vermitteltes Wissen ein besserer Lernerfolg, ein besserer Behaltensgrad und eine deutlich höhere Lernmotivation erzielt wird. Dieser Lernerfolg wird in kürzerer Lernzeit erreicht, was zusätzlich zu einer Reduktion der Schulungskosten führt.
- **Werbung, Produkt- und Firmenpräsentation, POI/POS:** Damit ein Produktname im Gedächtnis der Verbraucher haften bleibt, muß die entsprechende Werbung oder Präsentation möglichst effektiv sein und viele Sinne ansprechen. Auch hier werden die Forschungsergebnisse aus der Weiterbildung genutzt. Sogenannte Kiosk-Systeme, Point-Of-Sale/Point-Of-Information, ermöglichen Präsentationen von Produkten rund um die Uhr, also auch nach Geschäftsschluß.
- **Wartung, Reparatur und Produktservice:** Durch das Ansprechen verschiedener Wahrnehmungsorgane können beispielsweise Reparatur-Anleitungen verständlicher gestaltet werden. Langatmige schriftliche Erklärungen mit sparsamer Bebilderung der Reparaturschritte werden durch Videosequenzen ersetzt, die durch Sprache erläutert werden. Auch die umfangreichen Handbücher und Manuals können durch multimediale Bedienungsanleitungen ersetzt werden.
- **Produktkataloge:** Als ursprünglicher Bestandteil der Werbung gewinnen Produktkataloge immer mehr an Bedeutung. Die zum Teil riesige Angebotspalette bestimmter Firmen kann mit Hilfe multimedialer Technologie übersichtlich und ansprechend dargestellt werden. Eventuelle Bestellungen können beim Durchsuchen der Kataloge notiert und die entsprechenden Kosten zwischendurch und endgültig berechnet werden.

Weiterhin sind im Bereich der multimedialen Anwendungen elektronische Publikationen der Buchverlage zu nennen: Enzyklopädien, multimediale Reiseführer, Kochbücher, Kinderbücher, Sprachführer usw. Im Bereich der Computersimulation lassen sich z.B. Flugsimulatoren mit Hilfe von Video und Audio realistischer gestalten, als es allein durch Animation möglich ist.

1.4 Medienobjekte

Multimediale Präsentationen bestehen aus einer Menge an Informationseinheiten. Texte, Zeichnungen, Musikstücke und Videos, in einen bestimmten Zusammenhang zueinandergebracht, werden einem Benutzer in einer zeitlichen Abfolge präsentiert. Die einzelnen Informationseinheiten kapseln jeweils bestimmte Informationen und können, angelehnt an die Terminologie der objektorientierten Programmierung, als Objekte betrachtet

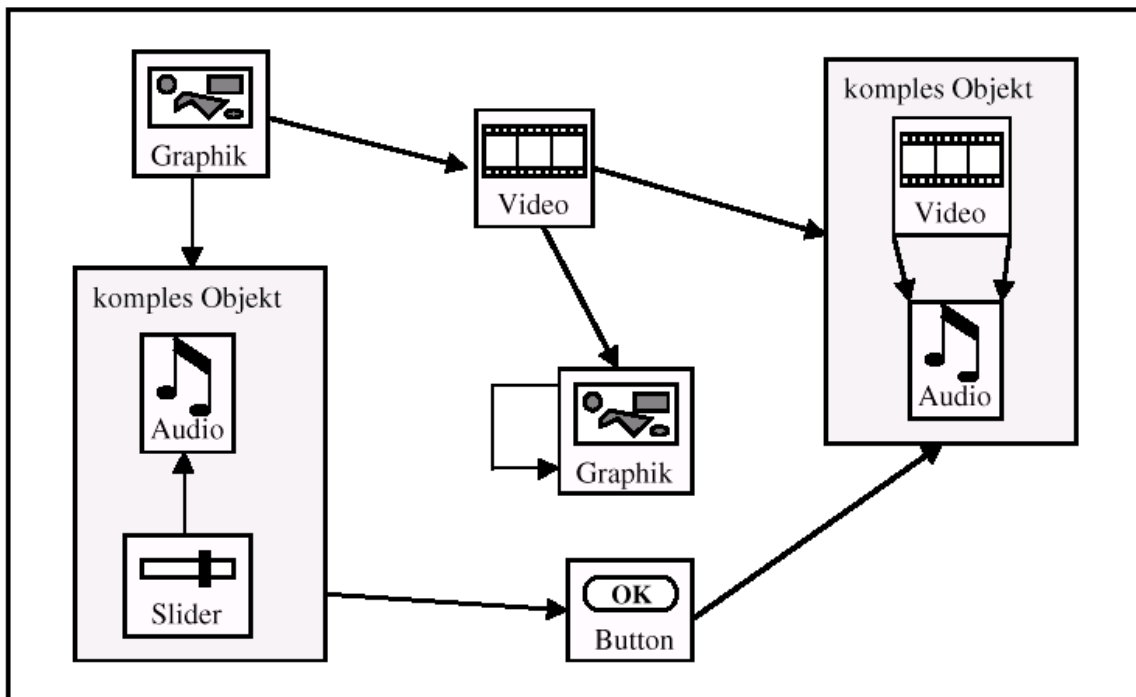
werden. Die Objekte sind von einem bestimmten Typ, genauer gesagt Medientyp, und werden daher als (elementare) Medienobjekte bezeichnet. So sind beispielsweise Texte Instanzen des Medientyps Text und Musikstücke Instanzen des Medientyps Audio.

1.4.1 Typen von Medienobjekten

Von den Ausgabemedientypen, wie Text, Graphik, Audio, Video und Animation, können Eingabemedientypen, auch Interaktionstypen genannt, unterschieden werden, deren Instanzen bestimmte Interaktionstechniken kapseln. Mit Hilfe von Interaktionsobjekten kann eine multimediale Präsentation interaktiv gestaltet werden. Die Rolle eines Benutzers, lediglich ein passiver Konsument der Präsentation zu sein, wird durch die Interaktionsobjekte verändert; der Benutzer wird zu einem aktiven Teilnehmer an der Präsentation.

Ein weiterer Typ von Objekten in multimedialen Anwendungen, auf die im folgenden nicht weiter eingegangen wird, sind die Applikationsobjekte. Sie kapseln bspw. Datenbanken, Wissensbasen oder andere anwendungsspezifische Objekte.

Die Medienobjekte (zusammen mit den Applikationsobjekten) bilden die Knoten eines multimedialen Beziehungsnetzwerks. Sie werden über Kanten, die bestimmte Beziehungen repräsentieren miteinander verbunden. Teilweise ist es sinnvoll, einen Teil des Beziehungsnetzwerks als eine Einheit zu betrachten, beispielsweise einen Film und seine Vertonung. Derartige zusammengesetzte Einheiten können als komplexe Medienobjekte angesehen werden, die durch Komposition von anderen - elementaren bzw. selbst wieder komplexen - Medienobjekten (ihren Komponentenobjekten) und Beziehungen zwischen den Komponentenobjekten gebildet werden. Synonym zum Begriff komplexes Medienobjekt kann auch der Begriff Multimedia-Objekt verwendet werden, insbesondere dann, wenn ein komplexes Medienobjekt mehrere Komponentenobjekte unterschiedlichen Typs beinhaltet. Die folgende Abbildung zeigt ein multimediales Beziehungsnetzwerk, bestehend aus elementaren und komplexen Medienobjekten, die durch Ikonen repräsentiert werden.



1.4.2 Beziehungen zwischen Medienobjekten

Sinn und Zweck einer multimedialen Präsentation ist es, einen Benutzer über einen bestimmten Sachverhalt zu informieren. Dazu werden im allgemeinen viele einzelne Informationseinheiten,

die Medien- und Applikationsobjekte, benötigt, die in einen bestimmten Zusammenhang zueinander gebracht werden müssen. Dieser Zusammenhang läßt sich über Beziehungen, Abhängigkeiten oder Constraints, die zwischen den Objekten definiert werden, herstellen. Im multimedialen Beziehungsnetzwerk repräsentieren die Kanten diese Beziehungen. Die Semantik der Kanten ist allerdings nicht einheitlich. Vielmehr existieren verschiedene Typen von Beziehungen, auf die im Folgenden im einzelnen eingegangen wird.

1.4.2.1 Zeitliche Beziehungen

Über zeitliche Beziehungen wird die zeitliche Abfolge definiert, in der die einzelnen Medienobjekte einem Benutzer präsentiert werden sollen. Zu beachten ist dabei, daß multimediale Präsentationen durch eine inhärente Nebenläufigkeit mehrerer Aus- und Eingabeaktionen gekennzeichnet sind. Beispiele für zeitliche Beziehungen zwischen Medienobjekten sind:

- Ein Video und seine Vertonung sollen gleichzeitig ausgegeben werden.
- Zwei Videos sollen nacheinander präsentiert werden.
- Ein Musikstück soll fünfmal hintereinander gespielt werden.
- Zehn Sekunden nach dem Start eines Videos soll ein Untertitel eingeblendet werden.

Es werden drei Ansätze zur Modellierung zeitlicher Beziehungen zwischen Medienobjekten unterschieden:

- Hierarchische Komposition,
- Komposition auf einer Zeitachse,
- Komposition über Referenzpunkte.

Bei der hierarchischen Komposition werden Bäume gebildet, deren Knoten Medienobjekte darstellen. Die Blätter repräsentieren elementare Medienobjekte. Innere Knoten werden durch Operatoren gebildet, die bestimmte zeitliche Beziehungen, wie beispielsweise den gemeinsamen Start, zwischen ihren Kinderknoten herstellen. Sie können damit als komplexe Medienobjekte angesehen werden. Vorteile der hierarchischen Komposition sind deren einfache Handhabung sowie die Integration von Objekten mit unbestimmter Präsentationsdauer (Live-Videos, Interaktionsobjekte, ...). Nachteile bestehen darin, da dieser Ansatz zur Modellierung bestimmter Sachverhalte nicht mächtig genug ist und daß Beziehungen zwischen Objekten nur über deren Start und Ende hergestellt werden können.

Die in kommerziellen Produkten am häufigsten verwendete Methode zur Modellierung zeitlicher Beziehungen zwischen Objekten ist deren Anordnung auf einer Zeitachse. Probleme treten in diesem sogenannten Timeline-Ansatz allerdings auf, wenn Interaktionsobjekte mit in die Modellierung einbezogen werden sollen, weil hierbei keine Angaben über deren Dauer gemacht werden können.

Den mächtigsten Ansatz stellt die Komposition über Referenzpunkte dar. Hierbei werden Medienobjekte nicht als atomare Einheit angesehen, sondern als eine Zusammensetzung aus Untereinheiten, wie beispielsweise Einzelbilder beim Video. Die Position einer Untereinheit innerhalb eines Objektes wird dabei als Referenzpunkt bezeichnet. Zeitliche Beziehungen zwischen mehreren Objekten werden dadurch hergestellt, da Verbindungen zwischen denjenigen Untereinheiten definiert werden, die gleichzeitig präsentiert werden sollen. Einen

Spezialfall der Komposition über Referenzpunkte bildet die Start-Ende-Komposition, bei der als Referenzpunkte lediglich der Beginn und das Ende eines Objektes dienen.

1.4.2.2 Gestalterische Beziehungen

Medienobjekte beinhalten eine bestimmte Information, d.h. sie definieren, was einem Benutzer präsentiert wird. Während zeitliche Beziehungen festlegen, wann die Information präsentiert wird, werden gestalterische Beziehungen dazu benutzt anzugeben, wie die Information - relativ zu anderen Objekten gesehen - dargestellt wird. Kriterien, die das Präsentationsbild eines Audio-Objektes festlegen sind beispielsweise die Ablaufgeschwindigkeit und die Lautstärke, Texte können unterschiedliche Fonts und Größen haben, die Farbgebung spielt bei Graphik-Objekten eine wichtige Rolle. Derartige Eigenschaften eines Objektes werden im allgemeinen initial vorgegeben. Sie lassen sich aber im Laufe der Präsentation über gestalterische Beziehungen dynamisch verändern. Im Prinzip werden gestalterische Beziehungen nicht zwischen den Objekten selbst definiert, sondern zwischen Attributen der Objekte, deren aktuellen Werte das Präsentationsbild bestimmen. Beispiele für gestalterische Beziehungen zwischen Medienobjekten sind:

- Andert sich die Lautstärke eines Liedes soll auch der Font eines gleichzeitig präsentierten Liedtextes geändert werden.
- Die Farbgebung eines Graphik-Objektes soll abhängig sein von der Größe des Objektes.
- Die Ablaufgeschwindigkeit eines Filmes soll immer gleich der Geschwindigkeit seiner Vertonung sein.

Je nachdem, wann bzw. über welchen Zeitraum gestalterische Beziehungen Gültigkeit besitzen, sind zwei Typen zu unterscheiden: Diskrete gestalterische Beziehungen müssen nur zu bestimmten Zeitpunkten erfüllt sein, beispielsweise zu Beginn der Präsentation. Dahingegen definieren kontinuierliche gestalterische Beziehungen Abhängigkeiten zwischen Medienobjekten, die sich über bestimmte Zeiträume, meistens über die gesamte Präsentation, erstrecken.

1.4.2.3 Räumliche Beziehungen

Räumliche Beziehungen legen den geometrischen Zusammenhang, d.h. Position und Größe, zwischen Medienobjekten auf dem Bildschirm fest. Sie können daher nur zwischen visuellen Medienobjekten definiert werden. Räumliche Beziehungen bilden eine Teilmenge der gestalterischen Beziehungen. Beispiele für räumliche Beziehungen zwischen Medienobjekten sind:

- Zwei Bilder sollen in demselben Fenster präsentiert werden.
- Ein Bild soll über ein anderes Bild gelegt werden, falls sie sich überdecken.
- Die Größe eines Bildes A soll immer gleich der Größe eines Bildes B sein, sodaß eine Größenanpassung erfolgen muß, wenn beispielsweise ein Benutzer die Größe des Bildes B interaktiv verändert.

Von besonderer Bedeutung sind die sogenannten initialen (diskreten) räumlichen Beziehungen. Dies sind räumliche Beziehungen, die nur beim Start des Objektes, zu dem die Beziehung hinführt, erfüllt sein müssen.

1.4.2.4 Gemischte Beziehungen

Beziehungen lassen sich dadurch charakterisieren, da sie in Abhängigkeit des Eintretens einer bestimmten Bedingung eine bestimmte Auswirkung haben. In den vorangehenden

Unterabschnitten wurden Beziehungen betrachtet, bei denen zeitliche bzw. gestalterische Bedingungen auch wieder zeitliche bzw. gestalterische Auswirkungen implizieren. Neben diesen sogenannten reinen Beziehungen sind auch gemischte Beziehungen vorstellbar, bei denen sich der Typ der Bedingung vom Typ der Auswirkung unterscheidet:

- Nach dem Start eines Videos soll eine Zeichnung an eine andere Stelle des Bildschirms positioniert werden (zeitlich-räumliche Beziehung).
- Wird ein Graphik-Objekt auf dem Bildschirm an eine bestimmte Stelle verschoben, soll ein Audio gestartet werden, über das eine Meldung ausgegeben wird (räumlich-zeitliche Beziehung).
- Je lauter ein Musikstück wird, umso größer soll ein bestimmtes Bild dargestellt werden (gestalterisch-räumliche Beziehung).
- Zu Beginn der Darstellung eines Graphik-Objektes soll dieses eine rote Hintergrundfarbe besitzen (zeitlich-gestalterische Beziehung).

Das letzte Beispiel ist ein Beispiel für eine sogenannte Initialisierungsbeziehung. Initialisierungsbeziehungen sind diskrete zeitlich-gestalterische Beziehungen zwischen ein und demselben Objekt, die zu Beginn der Präsentation des Objektes die Initialwerte der Attribute festlegen, die das Erscheinungsbild des Objektes definieren.

1.4.2.5 Interaktionsbeziehungen

Interaktionsbeziehungen sind Beziehungen, die von Interaktionsobjekten ausgehen. Sie kennzeichnen die Auswirkungen, die eine Benutzereingabe mit Hilfe dieses Objekt nach sich zieht. Diese Auswirkung kann sich in Form einer zeitlichen oder gestalterischer Veränderung der Präsentation bemerkbar machen. Beispiele sind:

- Das Drücken eines Buttons soll den Start eines Videos bewirken.
- Mit Hilfe eines Schiebereglers soll die Lautstärke eines Audios reguliert werden.
- Durch die Eingabe eines Koordinatenpaares in ein Texteingabefeld soll ein Bild an die entsprechende Stelle des Bildschirms positioniert werden.

1.4.2.6 Applikationsbeziehungen

Unter Applikationsbeziehungen werden Beziehungen verstanden, die von Applikationsobjekten ausgehen oder zu Applikationsobjekten hinführen. Sie kennzeichnen Auswirkungen von Änderungen im Applikationsobjekt auf die Präsentation bzw. Auswirkungen auf ein Applikationsobjekt durch die Präsentation. Beispiele für Applikationsbeziehungen sind:

- Eingaben eines Nutzer sollen in einer Datenbank gespeichert werden.
- Wird in einer Statistikkomponente ein bestimmter Schwellwert erreicht, soll eine animierte Information ausgegeben werden.

1.4.2.7 Indirekte Beziehungen

Als indirekte Beziehungen werden Beziehungen zwischen Medienobjekten bezeichnet, deren Auswirkungen sich nicht unmittelbar, sondern erst zu einem späteren Zeitpunkt der Präsentation bemerkbar machen. Wird ein Benutzer beispielsweise zu Beginn der Präsentation nach seinen Fähigkeiten befragt und wird in Abhängigkeit von seiner Antwort an einer bestimmten Stelle der Präsentation eine Hilfsmeldung ausgegeben oder nicht, so besteht zwischen dem Interaktionsobjekt, das den anfänglichen Dialog kapselt, und dem Medienobjekt, das die Hilfsmeldung präsentiert, eine indirekte Beziehung.

1.4.2.8 Konfigurable Beziehungen

Einen Beziehungstyp, bei dem Medienobjekte auf einem anderen Abstraktionsniveau betrachtet werden, bilden die konfigurablen Beziehungen. Medienobjekte werden dabei nicht als

atomare Einheiten angesehen, sondern als aus sogenannten Source-, Sink- und Filter-Objekten zusammengesetzte Objekte. Video-Objekte bestehen beispielsweise aus Source-Objekten, die Daten von einer Video-Kamera oder aus einer Datei lesen, und Sink-Objekten, die die Daten auf einen Bildschirm oder in eine Datei ausgeben. Die Daten werden dabei von den Source-Objekten in Form eines Datenstroms an die Sink-Objekte weitergeleitet. Filter-Objekte können zur Manipulation des Datenstroms zwischengeschaltet werden. Konfigurable Beziehungen kennzeichnen bei dieser Sichtweise auf Medienobjekte den Datenfluß zwischen den Teilobjekten. Sie unterscheiden sich von Beziehungen der anderen Typen dadurch, daß der Begriff Beziehung nicht im Sinne von Constraint zwischen zwei Objekten sondern eher im Sinne von "Datenkanal" zwischen zwei Objekten verwendet wird.

Konfigurable Beziehungen werden im weiteren Verlauf dieser Ausarbeitung nicht weiter berücksichtigt, da Medienobjekte als atomare Einheiten angesehen werden sollen. Konfigurable Beziehungen können somit als medienobjektinterne Beziehungen betrachtet werden, die keine Auswirkungen auf andere Medienobjekte haben.

2. Audio

Audio (lat.: " ich höre") kann grob als "Informationsvermittlung durch Ton" charakterisiert werden. Dabei muß jedoch der menschliche Hörbereich berücksichtigt werden, außerdem wird der Begriff Ton mehrdeutig benutzt: Zum einen wird - wie hier - Ton (allgemein) für ein beliebiges Klanggemisch benutzt, zum anderen kann dem (einzelnen) Ton eine Tonhöhe zugeordnet werden. "Klang" ist mit dem allgemeinen Tonbegriff gleichzusetzen und wird im folgenden bevorzugt als Synonym benutzt, um Mehrdeutigkeiten zu vermeiden.

Ein Ton ist ein physikalisches Phänomen, das durch Schwingung von Material verursacht wird. Bei einer derartigen Vibration werden in der Luft, die das Material umgibt, Druckwellenschwankungen ausgelöst. Die Druckwellen breiten sich in der Luft wellenartig aus.

Die Audiotechnik beschäftigt sich mit der Verarbeitung akustischer Signale, die durch den Menschen wahrgenommen werden können. Wesentliche Aspekte sind die psychoakustischen Grundlagen, Verarbeitung von Sprache (Synthese und Analyse) und Verarbeitung von Musik (MIDI).

2.1 Grundlagen

2.1.1 Physik

Ein Ton beruht auf einer Sinuskurve mit der x-achse als Zeit und der y-achse als Druck. Der Kurvenausschlag heißt Amplitude und stellt die Lautstärke dar. Eine Schwingung oder der kompletter Durchgang einer Sinusbewegung wird als Periode bezeichnet. Die Frequenz ist der reziproke Wert der Periode.

Folgende Frequenzen sind im wesentlichen zu unterscheiden:

- Infraschall: 0 - 20 Hz
- Hörschall: 20 Hz - 20 kHz
- Ultraschall: 20 kHz - 1GHz
- Hyperschall: 1GHz - 10 THz

Audio wird als Verarbeitung aller akustischen Signale im Frequenzbereich des Hörschalls betrachtet. Hierbei unterscheiden sich periodische Signale (Töne, Sprache, Musik) und nichtperiodische (Wind, Rauschen, Husten). Die Amplitude drückt die Lautstärke des Signales aus.

2.1.2 Die menschliche auditive Wahrnehmung

Auditive Wahrnehmung wird durch Druckschwankungen der Luft ausgelöst: Das Trommelfell eines Menschen schwingt synchron mit den auftreffenden Schallwellen mit, die Gehörknöchel leiten sie weiter und verstärken sie, bis sie von haarähnlichen Zellen in der Schnecke - einer flüssigkeitsgefüllten, aufgerollten Röhre - verarbeitet werden. Durch Nervenzellen werden die Bewegungen der Härchen in elektrische Impulse umgewandelt, die vom Gehirn interpretiert werden. Dieses ist also für die Beurteilung der akustischen Reize

verantwortlich; so kann das Gehirn die mutmaßliche Position der Schallquelle bestimmen oder auch die Aufnahme von auditiven Daten selektieren (z.B. ist es möglich, sich in einer größeren Menschenmenge auf die Stimme einer bestimmten Person zu konzentrieren).

Die Amplitude der Welle (Stärke der Luftdruckänderung) wird dabei als Lautstärke empfunden. Das menschliche Ohr nimmt die Änderungen logarithmisch wahr, das bedeutet, daß eine Verzehnfachung der Amplitude uns als Verdoppelung der Lautstärke erscheint. Wenn sich innerhalb der Wellenform periodische Wiederholungen befinden, kann eine Tonhöhe erkannt werden. Perioden werden in Sekunden gemessen. Die Anzahl von Perioden in einer bestimmten Zeit wird als Frequenz bezeichnet und in Hertz (Hz) gemessen (Perioden/sec).

Zusätzlich zur Tonhöhe wird Klangfarbe wahrgenommen, die sich aus dem Obertonspektrum ergibt. Tatsächlich besteht in der Regel jeder Ton aus vielen verschiedenen; es klingen leise Nebentöne (Obertöne) mit, die eine Vielzahl der Hauptfrequenz betragen und dieser "beigemischt" sind. Geräusche enthalten keine regelmäßig angeordneten Obertöne, es läßt sich also keine Grundfrequenz und somit keine eindeutige Tonhöhe bestimmen. In der Praxis setzen sich Klänge häufig aus Geräuschen und Tönen zusammen, so ist z.B. das Anblasen einer Posaune sehr geräuschhaft, bis sich der Ton mit seiner Tonhöhe herausbildet, jedoch immer noch mit einem Geräuschanteil.

Das menschliche Ohr unterscheidet weiterhin folgende Kenngrößen der räumlichen Wahrnehmung:

- Interaural Time Difference (ITD): Zeitunterschied zwischen linkem und rechtem Ohr
- Interaural Intensity Difference (IID): Signalunterschied zwischen linkem und rechtem Ohr

2.1.3 Komponenten der auditiven Informationsvermittlung

Auditive Daten können verschiedenen Klassen zugeordnet werden:

- Musik
- Geräusche und
- Sprache.

Dabei kann Musik als Hintergrundbegleitung, als primäre oder als begleitende Information eingesetzt werden. Geräusche bieten sich für kurze Informationen oder als Hilfsmittel für eine realistisch wirkende Anwendung an. Der Einsatz von Sprache eignet sich für Aufforderungen, wichtige Mitteilungen und vor allen Dingen zum Präsentieren textueller Information.

Earcons sind kurze, prägnante Klänge/Sounds/Geräusche/Wörter (in Anlehnung an graphische Icons), die für bestimmte Hinweise (wie Fehlermeldung "Versuch, Grenze zu überschreiten") verwendet werden und in der Regel als Feedback auf eine Benutzeraktion Einsatz finden.

2.2 Computerrepräsentation von Audio

Das Ziel, Natursounds möglichst echt zu synthetisieren oder zu reproduzieren, ist ein fundamentales in der Musikgeschichte - man denke nur an die versuchte Nachbildung von Streicher- und Bläserklängen durch Kirchenorgeln. Mit Sampling ist dieses möglich geworden. Kontinuierliche Signale (elektrische Schwingungen eines Mikrophons oder eines Tonbandgeräts) können mit Hilfe eines Analog-Digital-(A/D-)Wandlers in diskrete Werte

umgewandelt werden. Dabei ist zu beachten, da die Qualität des digitalen Signals von zwei Punkten abhängt:

- Wie oft wird der Wert des analogen Signals digitalisiert (Samplefrequenz)?
- Mit welcher Genauigkeit werden die Werte erfaßt (Auflösung)?

Die Sinuskurve läßt sich nicht direkt im Computer repräsentieren, daher stellt das Resultat nur eine Näherung an das Originalsignal dar. Ein Computer mißt die Amplitude einer Welle in regelmäßigen Abständen und generiert daraus eine Folge von Abtastwerten (Samples). Der Mechanismus wird als Analog-Digital-Konverter (ADC) bezeichnet; die Umkehrung als DAC.

Die Rate, mit der eine Welle abgetastet wird, nennt man Sampling Rate, die gleichfalls in Hz gemessen wird. Aufgrund des Nyquist-Abtasttheorems ist die digitale Abtastrate etwa halb so hoch wie die echte Bandbreite, d.h. bei der Digitalisierung von Klängen gilt die Regel "Samplefrequenz gleich zwei mal Signalfrequenz". Also ist bei einer Signalfrequenz von maximal 20 kHz (Hörgrenze des menschlichen Ohrs) eine Samplefrequenz von mindestens 40 kHz erforderlich. Dabei wird eine Sinuswelle von 20 kHz nur in zwei Stufen digitalisiert, welches nicht mehr einer Welle entspricht. Das menschliche Ohr gleicht solche "Fehler" jedoch aus, so daß kein Hörunterschied zwischen analogen und digitalen Signalen (bei entsprechender Qualität) zu hören ist.

Die Abtastgenauigkeit wird Quantisierung genannt. Sie repräsentiert die möglichen diskreten Werte jeder Abtastung. So entspricht:

- 3-bit 8 Werte
- 8-bit 256 Werte
- 16-bit 65536 Werte (CD)

Inzwischen hat sich eine Auflösung der Samples von 16 Bit (65536 mögliche Werte) durchgesetzt. Je geringer die Auflösung, desto größer ist der Unterschied zwischen dem analogen Wert und dem digitalisierten. Dieser Unterschied führt zu Verzerrungen des Signals, welche sich in Quantisierungsrauschen äußern.

Ein weiteres Problem ist das Aliasing: Wenn das Signal Wellen außerhalb der Hörgrenze enthält (>20 kHz), erscheinen diese im digitalisierten Zustand wie hörbare Wellen. Das ist durch die geringe Samplefrequenz zu erklären. Bekannt ist dieses Phänomen auch aus der Filmindustrie: Wenn sich ein Rad schneller dreht als die Bildwiederholungsfrequenz, scheint das Rad rückwärts zu laufen. Um diesen Effekt zu vermeiden, muß das Eingangssignal ab 20 kHz gefiltert werden, was den Preis für A/D-Wandler in die Höhe treibt. Ist das Signal einmal digitalisiert, kann es mit unterschiedlichen Geschwindigkeiten ausgegeben werden, was zu einer Änderung der Tonhöhe führt. Dadurch ist es möglich, Samples auf einer Tastatur in dem gewohnten Zwölftonsystem zu spielen.

2.3 MIDI

Da es für Keyboarder ein großer Wunsch war, mit einer Tastatur und einer Kontrolleinheit alle Keyboards bedienen zu können, wurde 1983 das Musical Instrument Digital Interface (MIDI) spezifiziert. Die großen Firmen der Musikindustrie (mehr als 10 Hersteller) stellten eine einheitliche Schnittstelle mit einem standardisierten Übertragungsprotokoll vor: MIDI ist

eine serielle Schnittstelle mit 31.25 kBit/s, über die externe Klangerzeuger mit beliebigem Syntheseverfahren angesprochen werden.

Die Definition erlaubt die Übertragung von kodierten Musiksignalen zwischen elektronischen Musikinstrumenten und Rechnern, dabei spezifiziert es sog. Midi-Geräte, die auf mehreren Kanälen zu einem Orchester zusammengefaßt werden können.

10 Minuten Musik sind so in ca 200 KiloByte MIDI-Daten darstellbar.

Komponenten von MIDI sind:

- die physische Implementierung durch Hardware inkl. MIDI-Port & MIDI-Kabel
- ein Datenformat mit folgenden Kennzeichen:
 - die Bezeichnung des Instrumentes
 - Beginn und Ende der Note
 - Grundfrequenz
 - Lautstärke

Der MIDI-Standard spezifiziert:

- MIDI erlaubt eine Kodierung über 10 Oktaven, was 128 Noten entspricht.
- 16 Kanäle (Spuren / Instrumente)
- 128 Instrumente inkl. nichtperiodischer Klänge
- 16 gleichzeitige Noten pro Instrument
- MIDI-Clock zur Synchronisierung
 - 24 Kennungen pro Viertelnote
 - Society of Motion Picture and TV Engineers (SMPTE) Standard = 30 Frames / sec.

2.4 Sprache als Audio

Sprach im computertechnischen Sinn unterscheidet zwischen Spracheingabe und Sprachausgabe. Bei der Sprachausgabe wird versucht eine möglichst naturgetreue Ausgabe von meist fest vorgegebenen Texten zu erreichen. Beispiele sind Ansagetexte in Fahrstühlen, Zügen oder bei Sprachcomputern. Die Spracheingabe hingegen versucht zunächst aus einem analogen Signal eine elektronische Repräsentation zu erstellen und diese nach verschiedenen Methoden inhaltlich zu erkennen.

Wesentliche Merkmale bei der Sprachverarbeitung sind:

- Die Ausgabe gzw. Verarbeitung der Eingabe muss in Echtzeit geschehen.
- Der Wortschatz ist meist begrenzt
- Die Sprachausgabe sollte natürlich klingen, die Eingabe mittels natürlichem Sprechen möglich sein.

Beim Sprachverstehen erfolgt eine Anpassung an Sprecher und Sprechgewohnheiten. Trotz Dialekt und Sprachfarbe (Gefühle) kann Sprache erkannt werden. Das Gehirn ist in der Lage, eine Trennung von Sprache und Störung zu erreichen. Hierzu werden Signale beider Ohren genutzt; im Gehirn erfolgt eine unbewußte Korrektur und Spracherkennungsfehlern durch Kenntnis von Inhalt, Kontext, Grammatik, Phonetik und lexikalischen Wortformen.

Sprache lässt sich grob untergliedern in:

- Phoneme, die sind die kleinsten, bedeutungsunterscheidende, aber nicht bedeutungstragende sprachliche Einheiten. Im Deutschn kennt man etwa 40 Phoneme.
- Allophone kennzeichnen Varianten eines Phonems als Funktion seiner lautlichen Umgebung.
- Morpheme kennzeichnen die kleinste bedeutungstragende sprachliche Einheit, So ist Haus eine Morphem.
- Stimmhafte und stimmlose Laute. So sind F und S stimmlose Laute, M oder W stimmhaft.

Sprachausgabe kann in folgenden Stufen bis zur Verständlichkeit geführt werden:

- Reproduzierend: Die Erklärung durch das Beispiel eines Anrufbeantworters.
- Lautverkettung: Durch eine Verkettung von Lautbausteinen, z.b. durch Aneinanderfügen von Phonemen zu Diphonen und Allophonen. Diphone bestehen (wie der Name sagt) aus zwei Phonemen, dadurch ist oft noch kein Morphem gegeben, wodurch die Verständlichkeit nur mäßig gesteigert wird.
- Bildung von Halbsilben und Silben aus Diphonen zu mindestens morphemen Wörtern.
- Kontinuierliche Darstellung durch gesamtheitliche Darstellung. Der Ansatz alle existierenden Wörter zu speichern, sowie deren klanglichen Abhängigkeiten.

Spracheingabe ist zu unterteilen in verschiedene Teilbereiche:

- Die Sprechereerkennung beschäftigt sich mit der Identifikation einer Person durch seine Sprache. Dies kann bei Kontroll- oder Sicherheitssystemen nützlich sein, steigert aber auch bei Spracherkennung die Qualität der erkannten Wörter und Sätze.
- Die Spracherkennung beschäftigt sich mit der inhaltlichen Deutung von akustischen Signalen der Sprache, also der Zuordnung von Bedeutungen zu einem Phonem, sowie die Beziehung dieser untereinander, um so einen grammatikalischen Ansatz zu verfolgen.
- Die Sprachmustererkennung beschäftigt sich mit Art und Weise der Sprache, mit Deutung einer bestimmten Aussage, z.b. hinsichtlich Ihres Wahrheitsgehaltes, Ironie et. al.

Treffgenauigkeit bei einer computergestützte Spracherkennung lässt sich durch folgende Faktoren kennzeichnen,

- Sprecherbindung
- Ausreichender Wortschatz, mind. 25.000 Wörter kodiert
- Kontexterkenung von mind. 3 Wörtern (im Gegensatz zur Einzelworterkennung)
- Dialekt / Emotionsbeeinträchtigung berücksichtigbar, z.b. durch Adaption des Sprachtrainings
- Störgeräuschbeseitigung, z.b. durch Rauschunterdrückung im Schwellwertbereich

3. Text

Text ist ein wichtiges Mittel zur Informationsvermittlung. Durch Text können Sachverhalte konkret beschrieben werden. Leider ist jedoch durch den täglichen Umgang mit Text ein sicheres und korrektes Gestalten eines Dokumentes noch nicht zwingend gewährleistet. Daher werden zunächst die grundlegenden Begriffe der Schrift vorgestellt. Dies beinhaltet eine Definition des Schriftzeichens, der Zeichenmaße und Serifen. Weiterhin werden Möglichkeiten zur Verbesserung des Schriftbildes dargestellt. Abschließend erfolgt eine Einführung in das Dokumentendesign. Hierzu werden Seitenformat, Satzspiegel, Gestaltungsraster sowie Text- und Stilelemente definiert und erläutert.

Es werden die existierenden Typen von Textverarbeitungssystemen definiert und Satzsysteme sowie WYSIWYG-Systeme behandelt. Ein kurzer Überblick über die heutzutage gängigen und am meisten verbreiteten Standards wie SGML/HTML bzw. SGML/HyTime sowie das Rich Text Format, die Open Document Architecture und Postscript, folgt.

3.1 Typographie

In diesem Abschnitt werden einige grundlegende Begriffe zur Typographie eingeführt. Zunächst wird auf die grundlegenden Termina der Schrift selbst eingegangen, und im zweiten Teil folgt eine Einführung einiger Begriffe zur Dokumentengestaltung.

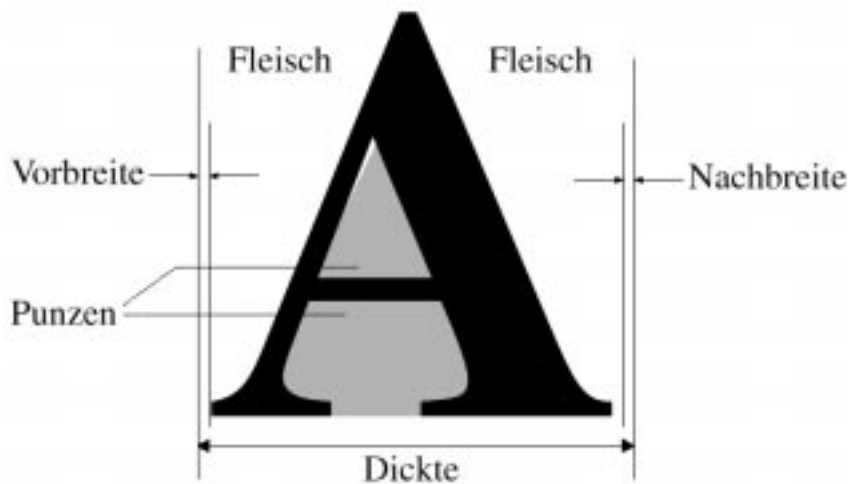
3.1.1 Schrift

3.1.1.1 Das Schriftzeichen

Die Maße und Begriffe der Schriftzeichen leiten sich historisch aus dem Bleisatz ab. So wird im Bleisatz mit dem Schriftkegel der Metallkörper bezeichnet, der das zu druckende Zeichen trägt. Im Desktop-Publishing-Bereich (kurz DTP-Bereich) ist mit dem Schriftkegel der gesamte Bereich gemeint, den ein Zeichen einnimmt. Unter Dichte wird die Breite des Schriftkegels verstanden. Diese ist bei den meisten Druckschriften für die einzelnen Zeichen unterschiedlich. Lediglich in den Schriften, die einer Schreibmaschinenschrift ähneln, haben alle Zeichen die gleiche Dichte. Diese Schriften heißen dichten gleiche Schriften (engl. monospaced fonts). Schriften unterschiedlicher Dichte werden als Proportionalschriften bezeichnet.

Unter dem Fleisch versteht man den freien Raum um das Schriftzeichen. Der schmale Abstand vor und hinter dem Zeichen auf dem Schriftkegel wird als Vorbreite und Nachbreite bezeichnet. Ebenso ist bei den meisten Zeichen oberhalb und unterhalb des Zeichens noch Fleisch vorhanden, um sicherzustellen, dass sich dicht unterinandergesetzte Zeichen nicht berühren.

Natürlich ist es im elektronischen Satz möglich, diesen Abstand ganz zu eliminieren. Der Innenraum des Zeichens schließlich wird Punze genannt.



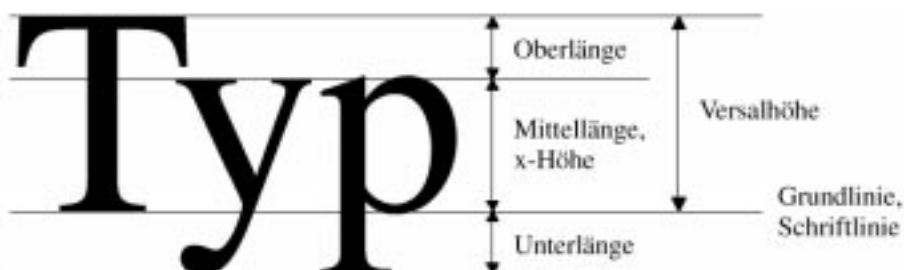
3.1.1.2 Zeichenmaße

In der folgenden Abbildung sind die wichtigsten Größenbegriffe und Maße einer Schrift dargestellt. Die Schriftlinie oder Grundlinie ist die gedachte Linie, an der die Schrift ausgerichtet ist. Man beachte, daß runde Zeichen geringfügig über diese und die im folgenden beschriebenen Linien hinausragen, um optisch die gleiche Höhe wie ein glatt abschließendes Zeichen zu haben. Man spricht dann von einem Überhang.

Die Mittellänge einer Schrift, auch x-Höhe genannt, gibt die Höhe der Zeichen wie a, c, e, m oder x an. Als Unterlänge bezeichnet man die Strecke, um die Kleinbuchstaben wie z.B. p oder q über die Schriftlinie nach unten ragen. Entsprechend dazu ist die Oberlänge die Strecke, um die Zeichen wie h und l über die Mittellänge nach oben hinausragen. Diese Oberlänge darf die maximale Höhe von Großbuchstaben, welche Versalhöhe genannt wird, geringfügig unter oder sogar überschreiten. Die Großbuchstaben werden in Analogie zur Versalhöhe als Versalien oder Majuskeln bezeichnet, während man bei Kleinbuchstaben von Gemeinen oder Minuskeln spricht.

Die Höhe des Schriftkegels (historisch Kegelgröße) ist die Schriftgröße bzw. der Schriftgrad, die sich somit aus der Versalhöhe plus der Unterlänge plus einem kleinen Zwischenraum über und unter den Zeichen zusammensetzt.

Der Begriff Schriftstärke oder Fettegrad definiert die Dicke der Strichstärken einzelner Zeichen. Hierfür werden die Bezeichnungen *leicht*, *mager*, *normal*, *halbfett*, *fett* und *extrafett* verwendet.



Unter dem Schriftschnitt (engl. font) versteht man die geringfügige Variation einer Schrift, beispielsweise die Änderung der Schriftlage, d.h. der Neigung der Schrift. So ist Times New Roman ein Schriftschnitt der Schrift Times. Der Begriff Schriftfamilie faßt alle Schnitte

einer Schrift zusammen. Dazu gehört in der Regel neben der Grundschrift ein kursiver, ein halbfetter und ein fetter Schnitt sowie deren kursive Varianten.

3.1.1.3 Serifen

Die Serifen bei einer Schrift sind die geschwungenen oder rechteckigen Enden der Striche, auch Endstriche genannt. Schriften, bei denen die Serifen extrem hervortreten, heißen serifenbetonte Schriften. Serifenlose Schriften, auch Grotesk-Schriften genannt, werden oftmals als sehr modern empfunden und finden deshalb zunehmende Verwendung. Dabei sollte jedoch beachtet werden, da Serifenschriften meist eine erheblich bessere Lesbarkeit implizieren.

3.1.1.4 Unterschneiden & Liturgen

Bei einigen Buchstabenkombinationen ist es sinnvoll, die Qualität des Schriftbildes nachträglich zu verbessern.

Zum einen bietet sich die Möglichkeit, Buchstaben näher aneinanderzustellen, als es ihren Standarddickten entspricht. Man nennt dies Unterschneiden (engl. kerning). Für die meisten Schriften werden vom Hersteller Unterschneidungstabellen für die wichtigsten entsprechenden Buchstabenpaare mitgeliefert. Professionelle DTP-Programme nutzen diese Tabellen, um das Unterschneiden automatisch durchzuführen.

Unter Ligaturen versteht man das Zusammenfassen von zwei oder drei Zeichen. Typische Ligaturen sind fi, fl, und ft. Einige Ligaturen sind heutzutage sogar fest zu einem Zeichen verschmolzen, wie beispielsweise das lateinische " et2 zum &.

3.1.2 Dokumente

Bei der Gestaltung eines neuen Textdokumentes sollte der Autor grob folgende Reihenfolge in der Vorgehensweise einhalten:

1. Festlegung des Seitenformats
2. Definition des Satzspiegels
3. Festlegen des Gestaltungsrasters und der Textspalten
4. Festlegen von Text- und Stilelementen
5. Anlegen des Dokuments und Eingabe des Inhalts
6. Feinkorrekturen im Umbruch

3.1.2.1 Seitenformat

Der erste Schritt zur Erstellung eines neuen Dokumentes ist die Wahl der Seitengröße. Hier bieten sich zunächst die bekannten Formate DIN A4 oder A5 an. DIN A4 ist für Standardbriefe, Datenblätter und Produktkataloge sicherlich die ideale Wahl, für Bücher oder Gebrauchsanleitungen beispielsweise ist A4 aber zu groß und A5 zu klein. Kann man die Seitengröße frei wählen, so sollte aus ästhetischen Gründen ein Breiten- zu Höhenverhältnis gewählt werden, welches größer als 1:1,414 (entspricht der DIN-Serie A) ist. In der Literatur werden folgende Werte genannt, die dem Leser als besonders harmonisch erscheinen:

- 2:3
- 1:1,538 (Grundlinie und Höhe eines Fünfecks)
- 5:8 (sog. Goldener Schnitt)
- 1: $\sqrt{3}$
- 1: $\sqrt{5}$ (sehr schlankes Format)

Quadratische Formate sollten aus dem oben genannten Grund nur in Ausnahmefällen verwendet werden.

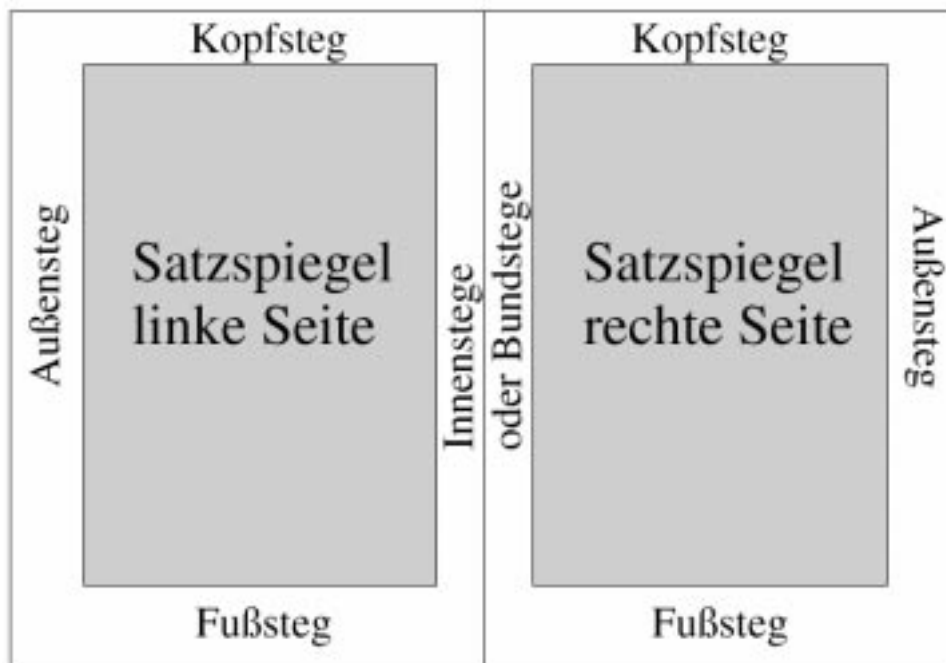
3.1.2.2 Satzspiegel

Der Satzspiegel oder das Layout bezeichnet die Anordnung der Text- und/oder Grafikbereiche auf einer Seite. Dazu gehören auch die Fußnoten und mögliche Kopf- oder Fußzeilen, die sog. Kolumnentitel. Weitere Elemente, die aber nicht zum Satzspiegel gezählt werden, sind die Seitenziffer (Pagina) und die eventuell vorhandene Randspalte (Marginalspalte oder Marginalien), die Kommentare oder Anmerkungen zum Text beinhaltet. Der Satzspiegel wird von vier Randbereichen eingerahmt, den Stegen. Für die Breite der dargestellten Stege sollte man in klassischen Dokumenten folgende Grundregel nicht verletzen: Innensteg < Kopfsteg < Außensteg < Fußsteg

Ein ästhetisch ansprechender Satzspiegel ergibt sich, wenn man die Stege im Verhältnis 2:3:4:5 aufteilt. Dabei sollte der Satzspiegel zwei Drittel und die Stege ein Drittel der Seitenbreite einnehmen.

3.1.2.3 Gestaltungsraaster

Im Anschluß an die Definition des Satzspiegels sollte überlegt werden, wie die Informationseinheiten wie Textspalten und Abbildungen innerhalb des Satzspiegels anzuordnen sind. Das dafür festzulegende Schema nennt man Gestaltungsraaster.



Dazu sollten zunächst Überlegungen zur Aufteilung des Textes in Spalten getroffen werden. Beim Lesevorgang wird der Satz durch den Leser in Blöcken von jeweils 2 bis 3 Worten erfaßt. Erst im zweiten Schritt wird diese Information in Worte zerlegt. Ist ein Wort unbekannt, zu lang oder durch das Schriftbild verfremdet, so wird zeichenweise gelesen. Dies reduziert jedoch die Lesegeschwindigkeit und ermüdet schnell. Gleiches gilt für den Wortabstand; ist dieser zu klein oder zu groß oder die Spaltenbreite zu schmal oder zu breit,

so bereitet die Unterteilung des Textes in Worte Probleme und das Auge erfaßt beim Lesevorgang weniger Text als möglich.

Eine Einteilung des Textes in mehrere Spalten kann also bei geeigneter Wahl der Spaltenbreite die Lesbarkeit des Textes stark erhöhen. Eine optimal gestaltete Textspalte sollte daher in etwa 45 bis 65 Zeichen der verwendeten Schrift aufnehmen können. Weiterhin sollte auch der Abstand zwischen zwei Spalten angemessen gewählt werden, da das Auge bei zu geringem Abstand leicht in die nächste Spalte wandert und bei zu großem Abstand der Eindruck entsteht, daß die Spalten auseinanderfallen. Ein adäquater Abstand stellt in etwa die Breite der Zeichenfolge "mi" dar.

3.1.2.4 Text- und Stilelemente

Die Schriftgröße der Grundschrift bzw. Werkschrift oder Brotschrift eines Dokumentes ist auf die Intention des Dokumentes, der Menge der Zeichen pro Zeile, den Inhalt und den Leserkreis abzustimmen. Ein Dokument für Erwachsene mit normalem Zeilenabstand sollte einen Schriftgrad zwischen 8 und 12 Punkten verwenden. Für Kinder sind Schriftarten mit Größen zwischen 11 und 14 Punkten empfehlenswert.

Auf Overheadfolien sind für die Grundschrift mindestens 14, besser noch 16 und für Schlagworte sogar 18 Punkte zu setzen. Auf Vortragsfolien sollten dabei niemals mehr als 12 Zeilen pro Folie gezeigt werden. Ein gutes Maß sind hier eher 4 bis 7 Stichworte oder Kurzaussagen, da dies der Erinnerungsfähigkeit des Menschen entgegenkommt.

Für Informationen, die aus größerer Entfernung aufgenommen werden sollen, müssen selbstverständlich größere Schriftgrade, sogenannte Schaugrößen bzw. Plakatschriften gewählt werden.

Als Konsultationsschrift bezeichnet man den Schriftgrad, in den Anmerkungen, Hinweise und Fußnoten gesetzt werden. Diese sind in der Regel recht kurz und gehören nicht zum eigentlichen Lesefluß, weshalb sie anderen Größenanforderungen unterliegen. Schriftgrößen zwischen 6 und 10 Punkten stellen hier optimale Werte dar.

Für Überschriften gibt es zwei Alternativen, die Hierarchiestufen optisch hervorzuheben. Zum einen kann man für jede Stufe einen eigenen Schriftgrad verwenden. Hier ergibt sich eine Abstufung 24 Punkt fett - 18 Punkt fett - 14 Punkt fett - 12 Punkt fett - 10 Punkt. Die letzte Angabe bezeichnet hier die Grundschrift. Die Alternative besteht darin, nur zwei Schriftgrade zu verwenden. Die daraus resultierende Abstufung ist beispielsweise 14 Punkt fett - 10 Punkt fett - 10 Punkt fett und kursiv - 10 Punkt kursiv - 10 Punkt normal. Der Abstand oberhalb und unterhalb einer Überschrift sollte so groß gewählt werden, daß die Überschrift optisch aus dem Fließtext heraussticht. Häufig ist Leerraum (in umfangreichen Dokumenten sogar zwischen 4 und 10 Leerzeilen nach der Überschrift) sinnvoller zur Gliederung als die Verwendung mehrerer unterschiedlicher Schriftgrade.

Als letztes Stilelement eines Textes ist die Absatzuntergliederung sehr wichtig. Absätze sind eigenständige Informationseinheiten. Sie sollten weder zu groß noch zu klein gewählt werden, wobei eine typische Absatzlänge etwa 5 bis 15 Zeilen umfaßt. Die erste Zeile eines Absatzes hat in der Regel einen Einzug, um einen neuen Absatz für das Auge des Lesers

leichter erkennbar zu machen. Die Absatzausrichtung schließlich gibt an, wie der Text innerhalb einer Spalte ausgerichtet wird:

- Blocksatz oder abgeschlossener Satz: Standard für Fließtext innerhalb eines Dokumentes;
- Linksbündig gesetzter Text sollte immer dann verwendet werden, wenn man schmale Spalten benutzt;
- Rechtsbündig formatierter Text wird fast ausschließlich in Marginalien, d.h. in Randspalten der linken Seite benutzt;
- Zentriert gestellter Satz wird in Überschriften oder auf Deckblättern verwendet.

Durch den automatischen Absatzumbruch können sogenannte Hurenkinder, d.h. die letzte Zeile eines Absatzes, die am Anfang einer neuen Spalte oder Seite erscheint, oder Schusterjungen, die erste Zeile eines Absatzes am Ende einer Spalte oder Seite, entstehen. Diese beiden Situationen sollten selbstverständlich durch eine manuelle Nachkorrektur des Textes beseitigt werden.

3.2 Textverarbeitung

In diesem Abschnitt werden die zwei existierenden Klassen von Textverarbeitungssystemen charakterisiert. Zum einen sind dies die Satzsysteme, die die Satzattribute für die einzelnen Textelemente (das sogenannte Markup) als Steuerzeichen in einem zu interpretierenden oder zu kompilierenden Textfile erwarten, zum anderen sind es die Programme, die den Text nach dem WYSIWYG-Prinzip (What You See Is What You Get) darstellen.

3.2.1 Satzsysteme

In Satzsystemen wird das Markup quasi als Steuerzeichen in den eigentlichen Text integriert. Dies wird auch als pseudoverbale Darstellung bezeichnet. Das Markup kann durch syntaktische Vereinbarungen wie beispielsweise einen Escape Character von inhaltlichen Text unterschieden werden. Ein Beispiel für ein solches Satzsystem ist T E X und seine Derivate. Bei der T E X- Auszeichnung `\footnote { ...}` etwa besagen die syntaktischen Regeln, daß der Backslash ein T E X-Kommando einleitet, das aus den nachfolgenden Kleinbuchstaben, also "footnote", besteht, und daß das darauf folgende Klammerpaar nicht zum Inhalt, sondern zum Markup gehört und den Fußnotentext begrenzt.

Satzsysteme wie T E X unterscheiden beim Eingeben normalerweise nicht zwischen Markup und inhaltlichem Text, sondern behandeln alle Zeichen gleich. Somit werden Texte meist mit einfachen Texteditoren eingegeben. Dies bedeutet jedoch, daß beim Editieren strukturierte Objekte nicht als Einheit ansprechbar sind, sondern als normale Zeichenketten behandelt werden.

Löscht man zum Beispiel versehentlich in der T E X- Anweisung `\footnote { ...}` nur die letzte geschweifte Klammer, so kommt man in einen fehlerhaften Zustand, der jedoch erst nach der Formatierung des Textes sichtbar wird und unter Umständen eine langwierige Fehlersuche nach sich ziehen kann. Desweiteren sind die Dokumente in der Regel am Bildschirm nicht gut zu lesen, da ihre Darstellungsform sich sehr von den gewohnten Papirausdrucken unterscheidet. Sie kommen jedoch, da sie nur Text darstellen müssen, mit textorientierten Bildschirmen aus, wie sie in UNIX-basierten Umgebungen üblich sind.

Generelle Vorteile von Satzsystemen sind, da man nur das grobe Erscheinungsbild des Dokuments sowie den Inhalt vorgibt und das System die aufwendigen Aufgaben des Formatierens übernimmt. Desweiteren haben sich Satzsysteme wie T E X gerade im wissenschaftlichen Bereich durch die vergleichsweise einfache Integration von mathematischen Formeln in Textdokumenten durchgesetzt.

3.2.2 WYSIWYG/Desktop-Publishing

Markup kann weiterhin auch indirekt, d.h. durch seine Auswirkungen auf die Formatierung am Bildschirm, dargestellt werden. Bei dieser Darstellungsform wird Markup, sobald es vom Benutzer eingegeben ist, sofort vom System interpretiert. Es ist dann nur noch in seiner Auswirkung auf den eigentlichen Text am Bildschirm sichtbar. Systeme, die auf diese Art versuchen, das Dokument dem späteren Ausdruck so ähnlich wie möglich auf dem Bildschirm darzustellen, werden auch als WYSIWYG-Systeme bezeichnet. Ein Beispiel für ein WYSIWYG-System ist Microsoft Word. Möchte man in Word eine Zeichenfolge kursiv setzen, so reicht es aus, den Text am Bildschirm zu markieren und danach das Kommando "Setze kursiv" auszuführen. Die markierte Passage wird danach umgehend kursiv gesetzt.

Moderne WYSIWYG-Systeme führen heutzutage auch Aufgaben aus, die früher als Vorteile von Satzsystemen galten. Dazu zählen zum Beispiel:

- der automatische Zeilenumbruch
- die automatische Numerierung von Überschriften, Aufzählungen und Seiten
- das Erstellen eines Inhaltsverzeichnisses oder Indexes
- eine Einteilung des Dokuments in Spalten
- Einbindung von Grafiken sowie deren Umfließen mit Text

Eine spezielle Form von WYSIWYG-Systemen sind die Desktop-Publishing-Systeme (DTP-Systeme). DTP-Systeme zeichnen sich zusätzlich zu den oben genannten Eigenschaften dadurch aus, da sie die Platzierung eines Textes innerhalb frei definierbarer Rahmen auf einer Seite des Dokuments erlauben. Dies hat den unter bestimmten Umständen gewünschten Vorteil, daß Text und/oder Grafiken millimetergenau im Dokument gesetzt werden können. DTP-Programme werden daher fast ausschließlich im professionellen Bereich zur Herstellung von Zeitungen, Werbeanzeigen, Broschüren etc. eingesetzt, während im normalen Schriftverkehr WYSIWYG-basierte Textverarbeitungsprogramme ausreichen.

3.3 Standards

In diesem Abschnitt soll jeweils kurz auf einige aktuelle Standards im Bereich der Dokumentenverarbeitung eingegangen werden.

3.3.1 SGML

Die Standard Generalized Markup Language (SGML) wurde 1986 unter ISO 8879:1986 verabschiedet. SGML ist ein System zur Definition strukturierter Dokumenttypen und Markup-Sprachen. Das Markup wird durch sogenannte Tags in den inhaltlichen Text eingefügt.

Tags werden durch die spitzen Klammern < und > begrenzt. Dabei wird entweder ein <Starttag> und ein </Endtag> oder ein einzelner <Tag> verwendet. SGML definiert jedoch nur die Syntax und nicht die Semantik der Sprachelemente. Instanzen von SGML sind HTML (Hypertext Markup Language) oder HyTime (Hypermedia/Time-based Document Structuring Language), d.h. in diesen Sprachen wurde die Semantik festgelegt. Jedes HTML- oder HyTime-Dokument ist somit auch ein SGML-Dokument.

3.3.1.1 HTML

HyperText Markup Language, das 1995 in der Version 2.0 verabschiedet wurde, ist zur Zeit der Standard für Dokumente im World Wide Web. Die in HTML verfaßte Eingabedatei wird von einem sogenannten Browser interpretiert und am Bildschirm dargestellt.

3.3.1.2 HyTime

Die SGML-Instanz HyTime wurde 1986 unter ISO 10744 verabschiedet. Es bietet ebenso wie HTML die grundlegenden Konzepte zur Repräsentation von Hypertext-, Hypermedia- und Multimedia-Dokumente. Der Schwerpunkt von HyTime ist jedoch, da abstrakte räumliche und zeitliche Beziehungen zwischen den verbundenen Dokumenten dargestellt werden können.

3.3.2 Rich Text Format (RTF)

Das Rich Text Format stellt eine Möglichkeit dar, Markup eines Dokuments unter Verwendung des druckbaren ASCII-Zeichensatzes im Text zu kodieren. Spezielle Zeichen können ebenso auf diese Weise kodiert werden, obwohl RTF die Verwendung von nicht-druckbaren ASCII-Zeichen nicht verhindert.

Das Markup des Textes wird durch spezielle Befehle in den inhaltlichen Text eingefügt. Der Aufbau eines Befehls ist \Befehl<Begrenzer>, wobei <Begrenzer> folgendes sein kann:

- ein Leerzeichen: das Leerzeichen ist Teil des Befehls
- ein Zeichen, das weder Buchstabe noch Zahl ist: dieses Zeichen beendet den Befehl, gehört aber nicht dazu
- eine Zahl oder ein Bindestrich: die angehängte Zahlenfolge ist der Parameter des Befehls und wird wiederum wie in den beiden oben genannten Fällen begrenzt

Weiterhin besteht in RTF die Möglichkeit, Text und/oder Befehle durch geschweifte Klammern zu Bereichen zu gruppieren. Innerhalb dieser Gruppierungen lassen sich entweder Attribute für den gesamten Bereich festlegen oder aber die Funktion dieses Bereiches läßt sich ändern.

3.3.3 Postscript

Die Dokumentenbeschreibungssprache Postscript wurde 1985 von der Firma Adobe entwickelt. Postscript ist ein grundsätzlich vektororientiertes Format, d.h. alle Bestandteile eines Dokumentes werden durch Kombinationen von grafischen Grundelementen beschrieben. Die Darstellung dieser Grundelemente erfolgt in umgekehrt polnischer Notation, in der die Parameter dem Befehl vorangestellt sind. So führt beispielsweise der Befehl *X Y Radius Alpha Beta arc* dazu, da um den Punkt (X, Y) ein Kreisbogen mit dem angegebenen Radius und Startwinkel Alpha sowie Endwinkel Beta gezeichnet wird.

Erweiterungen von Postscript sind Postscript Level 2 und Encapsulated Postscript (EPS): Diese Erweiterungen wurden im wesentlichen zur Darstellung von Farbe in Dokumenten und zur Verbesserung des Dokumentenaustauschs vorgenommen. Die Erweiterung Display Postscript dient dazu, Postscript zur Laufzeit darzustellen und verändern zu können. Beispielsweise wird die komplette Benutzungsoberfläche des Betriebssystems NEXTSTEP per Display Postscript dargestellt.

4. Bild und Grafik

Transport- und Ausdrucksmittel der Grafik im allgemeinen und der Computergrafik im speziellen ist das Bild. Ein Bild wirkt einerseits durch seine inhaltliche Absicht und andererseits durch seine formale Gestaltung auf den Betrachter.

4.1 Grundlagen

4.1.1 Der Begriff Bild

Ein digitales Bild besteht aus n Zeilen mit m Punkten, den sogenannten Pixeln. Ein Bild kann aus der realen Welt stammen oder virtuell erzeugt sein. Überlicherweise ist ein Bild ein rechteckiger Ausschnitt einer 2-dimensionalen Ebene der realen Welt.

4.1.2 Der Begriff Grafik

Grafiken bestehen aus Primitiven und deren Attribute. Primitive sind unter anderem Linien, Rechtecke, Kreise, Texte, allgemein ein zwei- oder dreidimensionales Objekt. Attribute wie der Stil der Linien, deren Breite und Farbe bestimmen das Aussehen einer Grafik.

4.1.3 Bitmapgrafiken

In der Computergrafik wird zwischen der Vektor- und der Bitmapgrafik unterschieden.

Bei der Bitmaptechnik wird die Darstellung in ein festes Raster von Bildpunkten aufgeteilt, die als Pixel (für picture element) bezeichnet werden. Die Anzahl der Zeilen und Spalten des Rasters bestimmt die Auflösung der Grafik. Eine Bitmapgrafik der Auflösung 1280×1024 verwendet ein Raster mit 1024 Zeilen und 1280 Spalten. Die Anzahl der Farben einer Grafik hängt davon ab, wieviele Bits für das Abspeichern der Farben eines Pixels zur Verfügung stehen. Für einfache Ansprüche reichen meist 8 Bit aus, mit denen 256 Farben auf dem Bildschirm dargestellt werden können. Bei höchstwertigen Ansprüchen werden 24 Bit (ca. 16 Millionen Farben) benötigt.

Ein Hauptnachteil von Bitmapgrafiken ist ihr erheblicher Speicherbedarf. Eine 24-Bit-Darstellung mit 640×480 Bildpunkten benötigt bereits 912 KByte. Für die Ausgabe dieser Darstellung auf dem Drucker (300 dpi) würden 3.8 MByte benötigt. Um den erheblichen Speicherbedarf zu reduzieren, unterstützen einige Bitmap-Datei-Formate wie JPEG, TIFF oder TARGA deshalb verschiedene Kompressionstechniken.

Auch wenn viele Grafikprogramme die Auflösung einer Bitmapgrafik verändern können, bleibt die Abhängigkeit von der Auflösung ein weiterer Nachteil bei der Bearbeitung von Bitmapgrafiken. Grafische Operationen wie Skalieren lassen sich nur mit sehr deutlichen Qualitätseinbußen realisieren.

4.1.4 Vektorgrafiken

Die Alternative zu Bitmapgrafiken sind Vektordarstellungen. Im Gegensatz zur Bitmaptechnik, die jeden Punkt einer Computergrafik beschreibt, wird in der Vektorgrafik ein Objekt durch Linien, Kreise, Rechtecke und Polygone beschrieben. Typischerweise enthält die Beschreibung Merkmale wie Größe, Winkel, Position, Füllmuster. Diese Grafikprimitive werden in eine Bitmapdarstellung umgeformt, die mit der Auflösung des Ausgabegerätes (Drucker, Monitor) übereinstimmt. Vektordarstellungen werden deshalb auch als ausgabe- bzw. auflösungsunabhängig bezeichnet.

Vektorgrafiken haben einen wesentlich geringeren Speicherbedarf als Bitmapgrafiken. Eine ausgefüllter Kreis, der fast den gesamten Computermonitor überdeckt, würde intern nur durch einige Verarbeitungsprimitive repräsentiert werden, wie z.B. zeichne Kreis mit der Farbe grün, dem Radius 200, an Position 300,200. Diese Instruktionen könnten in einigen Byte abgespeichert werden. Eine derartige Darstellung als Bitmap würde einige KByte bis zu einigen Mbyte benötigen, abhängig von der Auflösung und der Anzahl der Farben.

4.2 Die Bildsprache

Ein Bild wirkt einerseits durch seine inhaltliche Absicht und andererseits durch seine formale Gestaltung auf den Betrachter. Der Computergrafiker hat die Aufgabe, Stil- und Gestaltungsmittel zu finden, die auf das Motiv und die mitzuteilende Aussage zugeschnitten sind. Der folgende Abschnitt beschäftigt sich daher mit den formalen Gestaltungsgrundsätzen aus der Bildsprache.

Die Bildsprache entsteht aus der inneren Struktur eines Bildes. Dabei sind die wichtigsten strukturbildenden Elemente Linien und Flächen. Man bezeichnet sie daher auch als "Substantive" der Bildsprache. In der Analogie zur Schriftsprache sind die Farben die "Adjektive" der Bildsprache. Sie "kolorieren" die Substantive. Weiterhin entstehen durch das Zusammenwirken von Linien, Flächen und verschiedenen Farben Kontraste und Konturen. Kontraste und Konturen sind wichtige "Attribute" der Bildsprache. Das Basiselement der Bildsprache ist das Bildformat. Es bestimmt nicht nur die äußere Form, sondern auch den Charakter des Bildes. All diese Faktoren werden in ihrer Gesamtheit auch als Bildschemata bezeichnet.

Von den Bildschemata gehen Symbolwirkungen aus, die über den Wahrnehmungsapparat Assoziationen und Haltungen hervorrufen. Die Bildsprache wirkt auf das Unterbewußtsein des Betrachters. Sie kann beim Betrachter ebenso Gefühle wecken wie auch Spannungen oder Entspannungen erzeugen. Gerade deshalb haben die Mechanismen der Bildsprache in der Computergrafik eine sehr große Bedeutung; unabhängig davon, ob es nun um die Gestaltung von Bildern mit eher künstlerischem Anspruch geht oder ob reine Sachverhalte visualisiert werden sollen. Denn dort, wo unbewußt mit Gestaltungselementen gearbeitet wird, deren Wirkung auf den Betrachter nicht berücksichtigt wurde, können Inhalte und Symbolwirkungen transportiert werden, die entgegengesetzt der beabsichtigten Bildaussage stehen.

4.2.1 Die Blickrichtung

Die Bildbetrachtung kann mit dem Lesen eines Textes verglichen werden. Im europäischen Kulturkreis liest man von links nach rechts und von oben nach unten. Dabei werden Sätze nicht mit einem Blick aufgenommen, sondern es erfolgen Blicksprünge von Wort zu Wort.

Auch bei der Bilderfassung wandert der Blick des Betrachters von links nach rechts und von oben nach unten. Diese Richtungstendenz ist im Gegensatz zur Leserichtung kulturunabhängig. Auch Chinesen und Japaner "lesen" Bilder von links oben nach rechts unten.

Die Betrachtung eines Bildes unterscheidet sich allerdings in einigen Aspekten von dem Lesen einer Textseite. Ein Bild wird nicht immer von links oben nach rechts unten abgetastet. In der Regel konzentriert sich der Blick des Betrachters zuerst auf diejenigen besonderem Interesse sind. Doch das interessierende Bildelement wird wieder von links oben nach rechts unten abgetastet, bevor der Rest des Bildes betrachtet wird. Diese Richtungstendenz hat Auswirkungen auf die Bildgestaltung. Bildelemente oder Kontraste, die eine besondere Wichtigkeit bekommen sollen, sollten im rechten Bildfeld positioniert werden. Der Blick des Betrachters wird dort, nachdem das Bild von links oben nach rechts unten abgetastet wurde, zur Ruhe kommen.

4.2.2 Das Bildformat

Der Bildautor kann zwar das Format des Computermonitors nicht beeinflussen, wohl aber das Format seiner Bilder. Bei den Formatvarianten unterscheidet man zwischen Hochformat, Querformat und quadratischem Format. Je nach verwendetem Format können unterschiedliche Bildwirkungen erzielt werden.

Das Quadrat ist das ausgeglichene aller Formate. Da alle seine Seiten gleich lang sind, verkörpert es Neutralität. Es wird deshalb vom Betrachter als sehr harmonisch und beruhigend empfunden. Ein quadratisches Format symbolisiert visuelle Statik, die sich allerdings auch negativ auswirken kann. So ist z.B. das quadratische Format grundsätzlich spannungsloser als die anderen Bildformate. Nicht zuletzt deshalb wird dieses Format nur in wenigen Fällen angewendet.

Das Querformat ist das natürlichste Format. Es entspricht etwa dem menschlichen Sehfeld. Ein Bild im Querformat wirkt ausladend, panoramaartig und bezieht das Motivumfeld stärker ein. Die Augen können ein Bild im Querformat mit einem Schwenk von links nach rechts abfahren. Die Breitenwirkung kommt deshalb besser zur Geltung. Grundsätzlich eignet sich dieses Format für Motive, bei denen eher die Breitenwirkung, als die Höhenwirkung betont werden soll, wie z.B. bei Landschaftsaufnahmen.

Beim Hochformat werden besonders die vertikalen Motivlinien betont. Die Höhenwirkung des Motivs kommt dadurch besser zur Geltung, während die Breitenwirkung gemindert wird. In einem hochformatigen Bild können Symbolwirkungen wie Größe, Stärke, Übergeordnetheit und Erhabenheit transportiert werden. Ein Motiv im Hochformat wirkt sehr oft eindrucksvoller und beherrschender als in anderen Formaten.

4.2.3 Die Aufteilung der Bildfläche

Ob eine symmetrische oder asymmetrische Aufteilung der Bildfläche entsteht, hängt davon ab, wie Teile des Motivs oder das Gesamtmotiv innerhalb des Bildes angeordnet werden. Betrachter finden in der Regel asymmetrische Aufteilungen interessanter. Sie strahlen mehr

Dynamik und Lebendigkeit aus. Dagegen führen symmetrische Aufteilungen eher zu ruhigen, statischen, geordneten, für den Betrachter sehr angenehmen Bildwirkungen.

Eine wichtige Rolle bei der Aufteilung der Bildfläche spielt der "Goldene Schnitt". Er besagt, daß die Aufteilung einer Fläche oder Strecke im Verhältnis 3:5 als besonders harmonisch empfunden wird. Bezogen auf die Positionierung von Bildelementen bedeutet das, daß das Motiv nicht direkt in die Bildmitte, sondern mehr links bzw. rechts oder leicht oberhalb bzw. leicht unterhalb der Bildmitte positioniert wird. Dadurch wirkt das Bild spannender, als wenn das Motivelement direkt in die Bildmitte plaziert wurde. Hier wird also die Symmetrie aufgegeben.

4.2.4 Gestaltungselemente Linien und Flächen

Noch bevor der Bildbetrachter in einer Grafik einen Gegenstand auch als solchen erkennt, werden unbewußt Linien und Flächen wahrgenommen. Linien und Flächen sind die elementaren Gestaltungselemente der Bildsprache. Ihr Gebrauch löst Stimmungen beim Betrachter aus, die der Computergrafiker kennen sollte.

4.2.4.1 Linien

Linien lenken unbewußt den Blick des Betrachters. Man spricht daher häufig von Führungs- oder Leitlinien. Man kann mit einer Führungslinie direkt auf ein Bildelement zielen. Linien reduzieren dabei ein Bildelement nicht nur auf seine äußere Form, sondern es werden Eindrücke wie Statik oder Dynamik, Spannung oder Ausgeglichenheit, Harmonie oder Disharmonie erzeugt.

Im Bildaufbau wird zwischen imaginären und wirklichen Linien unterschieden. Wirkliche Linien existieren als klare geometrische Konturen im Bild, beispielsweise als Häuserkante, Horizont oder auch als Schatten. Imaginäre Linien dagegen sind nicht direkt im Bild sichtbar. Bei der Betrachtung bildet der Betrachter sie eher unbewußt aus dem Zusammenhang des Bildaufbaus.

Der Grad der Wirkung der Linien auf den Betrachter ist nicht nur abhängig von dem Grad der Deutlichkeit, sondern auch von der Orientierung innerhalb des Bildes. Bei der Orientierung wird zwischen waagerechten, senkrechten und diagonalen Linien unterschieden.

Eine waagerechte Linie unterteilt ein Bild in eine obere und eine untere Bildhälfte. Dem Betrachter werden dabei Eindrücke wie Ruhe, Gleichgewicht, Stabilität, Dauerhaftigkeit und Zuverlässigkeit suggeriert. Senkrechte Linien suggerieren in gewisser Weise Stabilität, doch wirken sie spannungsvoller als waagerechte Linien. Diagonale Linien können als aufsteigend bzw. als absteigend bezeichnet werden. Eine Diagonale drückt also eine Bewegungstendenz aus. Ihre Wirkungen resultieren aus der typischen Blickrichtung von links nach rechts. Aufwärts gerichtete Diagonalen wirken in der Regel am harmonischsten.

4.2.4.2 Flächen

Flächen sind Ruhepunkte in einer Darstellung. Sie verleihen einem Bild optischen Halt. Im Gegensatz zu Linien, die den Blick des Betrachters über das Bild wandern lassen, bilden sie in sich geschlossene Schwerpunkte. Flächen können den Blick des Betrachters stoppen und halten.

Die elementarsten Flächen in der Bildgestaltung sind Dreieck, Kreis, Rechteck und Quadrat. Wie bei Linien muß man auch bei Flächen zwischen wirklichen und imaginären Flächen unterscheiden. Imaginäre Flächen entstehen durch eine entsprechende Anordnung von Bildelementen zu Kreisen, Rechtecken, Dreiecken oder Quadraten. Wirkliche Objekte sind als klare geometrische Objekte im Bild (z.B. Tischplatte, Fenster oder Teller usw.).

Imaginäre Dreiecksflächen stehen in deutlichem Kontrast zum rechteckigen Bildformat. Auf den Betrachter wirken derartige Bildkompositionen sehr spannungsgeladen und dynamisch. Zudem haben Dreiecksflächen die Eigenschaft, den Blick des Betrachters im Bild zu halten - der visuelle Austritt wird sozusagen erschwert.

Kulturell hat der Kreis unter anderem Bedeutungen wie Sonne und Mond und besitzt daher für den Betrachter einen sehr hohen Aufmerksamkeitswert. Symbolisch steht er für Aufgeschlossenheit Stabilität und Ruhe. Kreisflächen haben die Eigenschaft, einen deutlichen Kontrast zum Rechteckformat des Bildes zu erzeugen.

Das Rechteck ist die Form, die am häufigsten vorkommt, z.B. in Gebrauchsgegenständen und in der Architektur. Sehr schmale Rechtecke erzeugen beim Bildbetrachter ähnliche Eindrücke wie senkrechte oder waagerechte Linien. Ansonsten hängt die Wirkung stark von der Größe der Rechteckform im Bild ab. So hat ein großes waagerechtes Rechteck, das mit den Bildrändern korrespondiert, oft den Charakter eines Rahmens. Auf den Betrachter wirken derartige Bildkompositionen sehr ruhig und streng.

Vom Quadrat geht durch seine Form eine neutrale und ruhige, oft aber auch monotone Wirkung aus. Innerhalb von Bildern eignen sich quadratische Formen also dazu, Ruhepunkte für den Blick des Betrachters zu bilden.

4.3 Die Farbe

In der Computergrafik ist die Farbwirkung von besonderer Bedeutung. Durch eine überlegte Farbgestaltung kann der Bildautor bewußt oder auch unbewußt auf den Betrachter wirken, seine Stimmung beeinflussen, seine Gefühle berühren oder auch bestimmte Reaktionen und Assoziationen auslösen.

Bevor allerdings die konkreten Aspekte der Farbgestaltung behandelt werden, soll kurz auf die Eigenschaften von Auge und Licht eingegangen werden.

4.3.1 Eigenschaften von Auge und Licht

Ein Aspekt der bei der Farbgestaltung berücksichtigt werden muß, ist die Tatsache, daß die Linse des menschlichen Auges die Eigenschaft hat, keine "Farbkorrektur" durchzuführen. Die Linse fokussiert beim Scharfstellen nur auf jeweils einen Wellenlängenbereich. Andere Wellenlängenbereiche (Farben) werden vor oder hinter der Netzhaut abgebildet. Die Farben aus anderen Wellenlängenbereichen erscheinen dadurch unscharf. Dieser Effekt erklärt, warum reine Farben in gleicher Distanz zum Auge unterschiedlich scharf erscheinen.

Das für das menschliche Auge sichtbare Licht ist nur ein schmaler Bereich im Spektrum der elektromagnetischen Wellen. Dabei entsprechen die verschiedenen Frequenzen innerhalb

dieses Bereichs den verschiedenen Farben. Die Spektralfarben rot, orange, gelb grün, blau, indigo, und violett decken den Bereich des sichtbaren Lichts von $4 \cdot 10^{14}$ Hz (rot) bis $7 \cdot 10^{14}$ Hz (violett) bzw. von ca. 780 nm bis 380 nm ab, wobei das Auge ungefähr 350.000 verschiedene Farben in diesem Bereich unterscheiden kann.

4.3.2 Licht- und Körperfarben

Die Farben der Gegenstände in der natürlichen Umwelt sind in der Regel Körperfarben. Der Farbeindruck, den der Betrachter von einem Gegenstand erhält, entsteht dadurch, daß ein Teil des Lichts, das auf die Oberfläche des Gegenstands trifft, von dieser reflektiert wird, während der andere Teil absorbiert wird. Dabei bestimmt das übrigbleibende reflektierte Licht die Farbe des Gegenstandes. Ein Objekt, das keine Farbe des Sonnenlichts reflektiert, erscheint daher schwarz. Umgekehrt sieht man einen Gegenstand, der sämtliche Farben reflektiert, in weiß. Körperfarben werden wegen der Absorption (Subtraktion) auch als Subtraktionsfarben bezeichnet.

Alle Farbnuancen der Körperfarben lassen sich auf drei Primärfarben reduzieren: Gelb, Magentarot und Zyanblau. Diese Primärfarben zeichnen sich zum einen dadurch aus, daß sie nicht aus anderen Farben gemischt werden können, zum anderen lassen sich aus den Primärfarben sämtliche anderen Farben gewinnen.

Lichtfarben stellen im Gegensatz zu Körperfarben selbst das Licht dar. Man geht von Teilchen aus, die als Teil der Lichtquelle, selbst Licht abstrahlen. Bilder auf dem Computermonitor oder Fernseher entstehen durch Lichtfarben. Die Anzeige erfolgt hier durch laufendes Aussenden von Elektronen, die auf die Bildschirmoberfläche treffen und dort ein farbiges Aufleuchten des Phosphors bewirken. Lichtfarben entstehen durch additive Farbmischung. Ursprung dabei sind die drei primären Lichtfarben Rot, Grün und Blau. Die Farbe Schwarz entsteht, wenn keine der Grundfarben aktiv ist. Weiß erhält man durch Mischung aller Grundfarben.

Weitere Informationen finden sich unter dem Abschnitt Videoformate.

4.4 Das Licht

Licht als physikalisches Phänomen ist so komplex, daß es nicht mit einem mathematischen Modell vollständig beschrieben werden kann. Statt dessen bedient man sich bei der Analyse und Synthese des Lichts immer nur bestimmter Teilaspekte. So wird Licht je nachdem untersucht als: Welle, Quant, Strahl, Energie oder einfach nur als Information (z.B. Helligkeit).

Rendering-Programme sind also nicht in der Lage, das Licht physikalisch exakt zu simulieren. Vielmehr werden zur Simulation von Licht im Computer stark vereinfachte Modelle gewisser Teilaspekte der Lichterscheinung verwendet. Diese Modelle sind stets nur bestmögliche Annäherung an die tatsächliche Wirkung des Lichts in der Realität. Dem Bildautor steht das synthetische Licht als Gestaltungselement zur Verfügung.

4.4.1 Lichtquellen im Computer

Im folgenden sollen einige typische Lichtquellen, welche besonders in Raytracing und Radiosity-Programmen eingesetzt werden, beschrieben werden. Dabei stehen gestalterische Faktoren im Vordergrund.

Das Umgebungslicht (Ambient-Light) dient zur Festlegung einer gewissen Grundhelligkeit in der Szene. Zu hohes Umgebungslicht kann zu reduzierten Kontrasten führen. Dem Bild wird nicht nur Spannung entzogen, es verliert auch an Tiefenwirkung. Der Raumeindruck kommt nicht mehr in vollem Umfang zur Geltung, wodurch das Bild verflacht und auf den Betrachter langweilig wirkt. Das Umgebungslicht sollte daher nicht als "Helligkeitsregler" verstanden werden, sondern eher als "Kontrastregler".

Punktförmige Lichtquellen werden in Spotlichter (Spot-Lights) und Glühbirnen (Omni-Lights bzw. Point-Lights) unterteilt. Bei flächigen Lichtquellen unterscheidet man zwischen dem Sonnenlicht (unbegrenzt-flächige Lichtquelle) und Lichtwannen (begrenzt-flächige Lichtquellen).

Punkt- und Flächenlichter unterscheiden sich vor allem in den Schattenwürfen. Punktförmige Lichtquellen erzeugen perspektivische Schatten. Die Schatten sind um so größer, je näher sich die Lichtquelle sich am beleuchteten Objekt befindet.

Lichtwannen hingegen erzeugen durch Kern- und Halbschatten einen sanften Hell-Dunkel-Übergang. Beim Sonnenlicht hat die Entfernung zum beleuchteten Objekt keinen Einfluß auf die Größe der Schatten. Denn wegen der großen Entfernung der Sonne zur Erde verlaufen alle Lichtstrahlen nahezu parallel.

4.4.2 Lichtrichtung

Synthetische Lichtquellen können an beliebigen Stellen im Raum plaziert werden. Dabei muß sich der Bildautor über die Wirkungen von Licht aus verschiedenen Richtungen bewußt sein. Die Wirkung der Beleuchtung wird durch die Plazierung der Lichtquellen in Relation zu dem beleuchteten Objekt und der Kameraposition bestimmt.

Frontales Licht verwandelt jeden 3D-Körper in eine zweidimensionale Scheibe. Formen und Strukturen treten kaum noch hervor, wodurch das Motiv verflacht; dies nicht nur, weil die Körperoberfläche an allen Stellen nahezu die gleiche Helligkeit aufweist, sondern auch, weil das frontale Licht keine Schatten auf dem Körper entstehen läßt. Die Trennung zwischen Motiv und Hintergrund wird gewissermaßen aufgehoben.

Soll die Plastizität eines Bildelements hervorgehoben werden, so sollte die Lichtquelle seitlich positioniert werden. Durch Streiflicht kommt eine besonders starke Schattenmodellierung zustande. Die Lichtquelle wird hier im Winkel von 80 bis 90 Grad zu der vorderen Seite eines Objekts positioniert. Bei diesem Verfahren werden die Strukturen und Konturen der Oberfläche des beleuchteten Objekts besonders herausgearbeitet. Werden die Schatten nicht von einer weiteren Lichtquelle überdeckt, so entsteht ein intensiver Hell-Dunkel-Kontrast.

Gegenlicht kann ein Objekt vom Hintergrund lösen. Direktes Gegenlicht, mit einer einzigen Lichtquelle, die in einem Winkel von nahezu 180 Grad in Richtung der Kamera verläuft,

führt in der Computergrafik nicht zu dem gewünschten Ergebnis. Will man beispielsweise sowohl auf der linken als auch entlang der rechten Außenkontur eines Gegenstandes einen Lichtsaum, so benötigt man dazu zwei Gegenlichtquellen, die sich in einem bestimmten, vom Einzelfall abhängigen Winkel überkreuzen.

4.5 Die Perspektive

Damit in einem zweidimensionalen Bild ein räumlicher Eindruck entsteht, muß er perspektivische Merkmale aufweisen. Es gibt verschiedene Methoden, um auf einer Fläche eine perspektivische Darstellung zu erreichen. Die wichtigsten davon sollen im folgenden erklärt werden.

4.5.1 Zentralperspektive

Eine der bekanntesten Darstellungsformen in der Computergrafik ist die Zentralperspektive. Die Zentralperspektive besitzt die Eigenschaft, die abgebildeten Objekte proportional zu ihrer Entfernung von der Bildebene zu verkleinern. D.h. entfernt liegende Körper erscheinen kleiner als näherliegende. Wird der Horizont sehr hoch gewählt, ergibt sich eine Vogelperspektive. Umgekehrt ergibt sich bei sehr niedrig gewähltem Horizont die Froschperspektive.

4.5.2 Parallelprojektion

Eine andere Form der räumlichen Darstellung ist die Parallelprojektion. Bei ihr ist das Projektionszentrum unendlich weit entfernt, wodurch die Projektionsstrahlen parallel zueinander liegen. Dadurch ergibt sich eine Darstellung bei der alle im Raum parallelen Linien auch in der Darstellung parallel liegen. Augenpunkt und Fluchtpunkt gibt es hier nicht.

4.6 Bildspeicherformate

Ein Bild wird durch zwei Parameter spezifiziert, die räumliche Auflösung in einer Pixel-Matrix (x,y) und eine Farbkodierung, die in Bits pro Pixel gemessen wird.

Farbwerte können in unterschiedlichen Werte-Kombinationen abgelegt werden:

- Drei Zahlen für die Werte der Farben Rot, Grün und Blau.
- Drei Zahlen, die Verweise in eine Tabelle von RGB-Werten enthält
- Eine Zahl, die als Verweis auf ein Farbtupel dient

4.6.1 Graphics Interchange Format (GIF)

Das GIF wurde 1987 vom Online-Dienst CompuServe entwickelt, um Bilder plattformunabhängig auszutauschen. GIF zeichnet sich aus durch seine verlustfreie, aber effiziente Kompression und eine mögliche Schachtelung von mehreren Bildern in einer Datei aus.

Grundsätzlich besteht GIF aus folgenden Abschnitte, mit möglicher Wiederholung der letzten vier:

- Header: GIF Erkennungssequenz, Versionsnummer des Algorithmus
- Application: Kodierung von Anwendung und Version des erstellenden Programmes

- Trailer: Beginn / Ende des Datenstroms
- Control: Steuert die Darstellung des folgenden "image"-Blocks. Angaben zur zeitlichen Abfolge und Farbveränderungen sind möglich (Animated GIF).
- Image: eigentliche Bild Informationen (Farbtabelle und Pixeldaten).
- Comment: Kommentar zum Image
- Plain Text: Kommentar zum Bild als ASCII Text im Bild (z.b. Datum im Bild).

Ein GIF -Bild wird als Bitstrom kodiert, ein einzelnes Bild besteht aus dem Angaben von Größe, Position und Art der Farbtabelle, optionalen globalen oder lokalen Farbtabelle und den Werten für die Pixel, die als Verweise auf jene Farbtabelle angegeben werden. Für Pixelangaben werden im LZW (Lempel-Ziv-Welch)-Algorithmus komprimiert.

4.6.2 Tagged Image File Format (TIFF)

Das TIFF wurde 1987 von Aldus Corporation und Microsoft entwickelt, um Portabilität und Hardware-unabhängigkeit eines Bilder zu realisieren. TIFF hat seinen Hauptanwendungsbereich bei Scannern und Faxgeräten.

TIFF unterscheidet den Baseline-Teil und den Extension-Teil. Im Baseline-Teil werden Eigenschaften beschrieben, die Anzeigeprogramme unterstützen sollen; im Extension-Teil werden Funktionen für spezielle Programme beschrieben, wie z.B. die Verwendung von CYMK Farbmodellen.

TIFF unterstützt mehrere Farbmodelle:

- Baseline: Binär, Graustufen, RGB-Palette
- Extension: CYMK, wahrnehmungsorientierte Formate wie YUV

TIFF hat folgenden Aufbau:

- Header: Formatkennung, Versionsnummer
- Directory: Angaben über Verzeichnisse, Farbkodierung, Bildabschnitte
- Structure: Angabe über die Kodierungstechnik und Anzahl der folgenden Tag-Fields
- Tag-Fields: Angabe über Kodierungsblöcke: Zeilen, Objekte, Zellen oder Blöcke
- Data-Fields: Einzelne Elemente entsprechend der Definition in den Tag-Fields

TIFF unterstützt verschiedene Kompressionsverfahren:

- Lauflängenkodierung (auch PackBit oder M-bit)
- LZW
- Fax Gruppe 3 und 4
- Huffman Kodierung

Größter Vorteil von TIFF ist die generische Verwendung und die Möglichkeit Inhalte auf verschiedene Weise zu kodieren.

4.6.3 Bitmap (BMP)

BMP ist eine Microsoft und IBM Entwicklung für OS/2 und dann Windows. BMP unterstützt nur das RGB-Farbformat.

BMP hat folgenden Aufbau:

- Header: Angabe über Größe, Farbtiefe, Farbtabelle und Kompressionsart
- Datenbereich: Serielle Abfolge von Pixelwerten. Die Anzahl der Pixel in den Zeilen eines Bildes werden stets durch 32 teilbar gespeichert, verbleibende Werte werden mit Nullen gefüllt. Als Farbtiefen sind 1,4,8 und 24 zugelassen.

4.7 Grafikspeicherformate

Ein Dateiformat für Vektorgrafiken ermöglicht ausschließlich das Laden und Speichern von Grafiken in einer vektorisierten Darstellung.

Wichtigste Standards hierbei sind:

- IGES: Initial Graphics Exchange Standard zum Transfer von 2D und 3D CAD Daten.
- DXF: Entwickelt von der Firma AutoDesk für das Produkt AutoCAD zur Speicherung von 2D und 3D Grafiken, heute de-facto Standard.
- HPGL: Hewlett-Packard-Graphics-Language zur Ansteuerung von Plottern und zur Darstellung von 2D Daten.

4.8 Bildanalyse

Die Bildanalyse beschäftigt sich mit Techniken zur Extraktion von Beschreibungen aus Bildern, die für Methoden einer Szenenanalyse auf höherer Ebene notwendig sind. Das Wissen um die Position und den Wert eines speziellen Pixels trägt nur unwesentlich zur Information bei, die zur Erkennung eines Objektes, der Beschreibung der Objektform, dessen Position oder Ausrichtung, der Messung jeglicher Distanz zu diesem Objekt oder ob dieses defekt ist, vonnöten ist.

4.8.1 Bildverbesserung

Sie beschäftigt sich mit der Verbesserung der Bildqualität durch Rauschunterdrückung, Kontrastverstärkung oder Farbwandlung.

4.8.2 Mustererkennung

Sie beschäftigt sich mit Entdeckung und Klassifizierung von Standardmustern, sowie der Erkennung von Abweichungen derer. Ein wichtiges Beispiel ist OCR (Optical Character Recognition) oder die Handschrifterkennung.

4.8.3 Szenenanalyse

Sie beschäftigt sich mit Erkennung und Rekonstruktion von 3D-Modellen einer Szene, die aus mehreren 2D-Bildern besteht. Beispiel ist ein Industrieroboter (CNC).

4.9 Bildmerkmale

Die drei am häufigsten verwendeten Merkmale zur Klassifizierung eines Bildes sind Farbe, Textur und Kanten.

4.9.1 Farbe

Üblicherweise werden Bilder im RGB-Format dargestellt, eine gleichgewichtete Darstellung der Farben Rot, Grün und Blau jeweils als 8-bit-Pixel. Weiteres siehe Kapitel Video.

4.9.2 Textur

Unter einer Textur versteht man eine kleinräumige Oberflächenstruktur, egal ob natürlich oder künstlich, regelmäßig oder unregelmäßig. Beispiele sind Holz, Strickmuster, Oberfläche eines Schwammes oder Marmor. Es werden zwei Ansätze unterschieden:

- Bei der strukturellen Analyse sucht man nach kleinen Grundbausteinen sowie einer Anordnungsregel, nach der die Bausteine zu einer Textur gruppiert werden können.
- Bei der statistischen Texturanalyse wird die Textur als Ganzes durch bestimmte Attribute, wie Regelmäßigkeit, Körnung, Orientierung und Kontrast beschrieben.

Zur Texturanalyse wird ein Bild üblicherweise in Grauwerte umgerechnet und betrachtet dann eine signifikante und regelmäßige Variation der Grauwerte in einer kleinen Umgebung.

4.9.3 Kanten

Als Kanten betrachtet man in Bildern, Grauwertsprünge mit gewisser Regelmäßigkeit. Einfachstes Beispiel wäre ein schwarzes Quadrat auf weißem Hintergrund, dessen Kanten eindeutig durch den Farübergang beschrieben wird.

4.10 Bilderkennung

In folgenden werden die notwendigen Schritte zur Erkennung eines Bildes beschrieben.

4.10.1 Formatierung

Zunächst wird die Darstellung eines Bildes in weiterverarbeitbarer Form betrachtet. Dies betrifft Techniken der Aufnahme und Speicherung eines Bildes in digitaler Form (meist explizit, d.h. als Pixel in einer Matrix ohne Komprimierung).

4.10.2 Konditionierung

Die Konditionierung basiert auf einem Modell, das annimmt, daß sich ein beobachtbares Bild aus Informationsmustern zusammensetzt, die von irrelevanten Variationen gestört wird. Die Konditionierung unterdrückt dann "grafisches Rauschen" bzw. normalisiert.

4.10.3 Markierung

Die Markierung beruht auf einem Modell, das annimmt, daß Informationsmuster eine Struktur in Form von räumlichen Arrangements von Objekten haben, von denen jedes eine Menge miteinander verbundener Pixel ist. Die Markierung bestimmt, zu welchem Objekt jedes Pixel gehört. Die Kantenerkennung dient als Grundlage der Markierung.

4.10.4 Gruppierung

Die durch die Markierung erhaltenen Objekte werden zu sinnhaften Gruppen zusammengefaßt, d.h. es werden beispielsweise mehrere erkannte Kanten zu einer geometrischen Figur zusammengefaßt.

4.10.5 Extraktion

Die Gruppierung legt eine Menge von Einheiten fest, die nur unvollständig das Gesamtbild repräsentieren. Die Extraktion berechnet für jede Gruppe eine Liste von Eigenschaften (Mittelpunkt, Fläche, Ausrichtung,...) Weiterhin sind die nicht gruppierbaren "Reste" des Bildes zu betrachten.

4.10.6 Abbildung

Die Abbildung ordnet den Gruppen inhaltliche Bedeutung zu. Beispiel ist hierbei Pattern oder Template Matching, welches erkannte Gruppen mit gespeicherten Mustern vergleicht.

5. Video

Der Beginn der Videotechnik ist im Bereich Fernsehen zu sehen. Basierend auf der CRT (Cathode Tube Ray) Technologie wurden moderne Videotechniken begründet.

Die Bearbeitung eines Videosignales beinhaltet drei Aspekte:

- Visuelle Repräsentation
- Übertragung
- Digitalisierung

5.1 Visuelle Repräsentation

Das zentrale Ziel ist es, dem Betrachter ein möglichst realistisches Bild, bzw. Abbild oder Repräsentation, der Wirklichkeit zu bieten. Wichtige Kenngrößen hierfür sind:

- Horizontale und vertikale Details, d.h. das Bildverhältnis von Breite zu Höhe
- Der Betrachtungsabstand zum Bild
- Die Tiefenwahrnehmung, also die räumliche Dimension
- Die Leuchtdichte oder Intensität einer Farbe
- Der zeitliche Aspekt einer Beleuchtung, bzw. die Farbübergänge im Bewegtbild
- Die Bewegungskontinuität in der Bildfolge (Bildfrequenz und logische Sequenz)
- Die Bildstörungen oder auch Flimmern

Um Bewegungsabläufe als solche zu erkennen, müssen Einzelbilder mit einer bestimmten Geschwindigkeit dem menschlichen Auge vorgeführt werden. Das Minimum stellen dabei 16 Bilder pro Sekunde dar. Beim Film werden normalerweise mit 24 Bildern pro Sekunde gearbeitet, allerdings treten auch hierbei noch starke Flimmererscheinungen auf, die erst ab einer Frequenz von 50 Bildern pro Sekunde nicht mehr als störend empfunden werden.

Da in der Fernsehtechnik 25 Bilder pro Sekunde übertragen werden, wird dort ein kleiner Trick eingesetzt, um dem menschlichen Auge eine höhere Bildwiederholfrequenz vorzutäuschen. Ein Fernsehbild besteht aus 625 Zeilen, das mit der 1. Zeile am oberen linken Bildschirmrand beginnt und mit der 625. Zeile am unteren Bildrand endet; der Aufbau eines kompletten Bildes dauert eine 25stel Sekunde. Da während dieser Zeit die oberen Zeilen des Bildes schon wieder verblassen, die unteren aber gerade in voller Helligkeit auf dem Bildschirm erscheinen, würde es gerade bei besonders hellen Bildern zu starkem Flimmern kommen.

Um dieses nun zu vermeiden, teilt man das Vollbild aus 625 Zeilen in zwei Halbbilder aus je 312,5 Zeilen auf, die dann nacheinander übertragen werden. Es erscheinen somit 50 Halbbilder pro Sekunde, wobei das erste Halbbild alle ungeradzahligen Zeilen und das zweite alle geradzahligen Zeilen umfaßt.

5.2 Videokodierung

Videosignale werden in verschiedenen Kodierungsverfahren aufbereitet:

5.2.1 RGB Signal

Ein RGB-Signal besteht aus den getrennten Signalen für Rot, Grün und Blau. Jede Farbe kann als eine Kombination dieser drei Primärfarben über die Additive Farbmischung kodiert werden. Die Werte sind normiert, sodaß $R+G+B=1$ die Farbe Weiß ergibt.

5.2.2 YUV-Signal

Da die menschliche Wahrnehmung auf Helligkeit empfindlicher als auf Farbinformationen reagiert, unterscheidet das YUV-Verfahren die Leuchtdichte, Luminanz, von der Farbinformation, Chrominanz. Anstatt nach Farben zu teilen, kann man die Helligkeitsinformationen (ausreichend für Schwarz-Weiß-Fernsehen) und die Farbinformationen trennen. Es entstehen folgende Relationen:

- $Y = 0.3 R + 0.59 G + 0.11 B$
- $U = (B - Y) * 0.493$
- $V = (R - Y) * 0.877$

Durch die Relevanz von Luminanz zu Chrominanz, wird das Signal in unterschiedlichen Bandbreiten zur Verfügung gestellt, üblicherweise im Verhältnis 4:2:2. Aufgrund der Abhängigkeit untereinander wird dieses Verfahren manchmal auch als Y,B-Y, R-Y Verfahren bezeichnet.

5.2.3 YIQ-Signal

Eine dem YUV-ähnliche Kodierung, auf der auch das amerikanische NTSC Format basiert, mit folgenden Relationen:

- $Y = 0.3 R + 0.59 G + 0.11 B$
- $I = 0.6 R - 0.28 F - 0.32 B$
- $Q = 0.21R - 0.52G + 0.31 B$

5.3 Videoformate

Das vom einem Rechner verarbeitete Videoformat hängt von der Eingabe- und Ausgabegeräten des Mediums ab. Die Video Digitalisier-Hardware unterscheidet sich bezüglich der Auflösung digitaler Bilder (Frames) und der Bildrate (Frames/sec). Die Ausgabe hängt von der Anzeige-Hardware ab. Am häufigsten kommen hier Raster-Bildschirme zum Einsatz.

Der Video Controller zeigt das Bild an, das im Bildpuffer gespeichert wird, wobei auf den Speicher mittels eines separaten Ports so oft zugegriffen wird, wie dies die Abtastrate erfordert. Aufgrund des Flimmereffektes bewegt sich der Video Controller mit jeweils einer Abtastzeile 60 mal pro Sekunde zyklisch durch den Bildpuffer. Zur Anzeige verschiedener Farben auf dem Bildschirm arbeitet das System mit einer Farbabbildungstabelle einer sog. Color LookupTable (CLUT).

Bekannte Videoformate im Computerbereich sind:

- CGA: Der Color Graphics Adapter besitzt u.a. eine Auflösung von 320x200 Pixel bei gleichzeitiger Darstellung von 4 Farben, damit ergibt sich eine Kapazität von $320 \times 200 \times (2\text{bit}(\text{pixel})/8\text{bit}(\text{Byte})) = 16.000 \text{ byte}$

- EGA: Der Enhanced Graphics Adapter unterstützt bei einer Auflösung von 640x350 Pixel gleichzeitig 16 darstellbare Farben, damit $640 \times 350 \times (4\text{bit}(\text{pixel})/8\text{bit}(\text{Byte})) = 112.000$ byte
- VGA: Der Video Graphics Adapter arbeitet mit einer Auflösung von 640x480 Pixel bei Darstellung von 256 Farben, damit eine Kapazität von: $640 \times 480 \times (8\text{bit}(\text{pixel})/8\text{bit}(\text{Byte})) = 307.200$ byte
- SVGA: Der Super Video Graphics Adapter kann bei einer Auflösung von 1024x768 Pixel 256 Farben anzeigen, damit $1024 \times 768 \times (8\text{bit}(\text{pixel})/8\text{bit}(\text{Byte})) = 786.432$ byte.

Weitere Formate mit entsprechend höheren Auflösungen und Farbtiefen sind im XGA, UGA etc. definiert.

5.4 Fernsehformate

Das Fernsehen ist einer der wichtigen Anwendungsbereiche, allerdings entwickelten sich in den einzelnen Ländern und Regionen unterschiedliche Formate:

- NTSC (National Television Systems Committee) ist der älteste und anzahlmäßig am weitesten verbreitete Standard, entwickelt in den USA. Der Standard arbeitet mit 30 Hz und 525 Zeilen.
- SECAM (Sequential Couleur avec Memoire) wird vorwiegend in Frankreich und Osteuropa, benutzt 25 Hz bei 625 Zeilen.
- PAL (Phase Alternating Line) wird in weiten Teilen Westeuropas eingesetzt.
- HDTV (High Definition Television) wurde ursprünglich in Japan entwickelt mit einer Auflösung von 1440 (bzw. 1920) Zeilen mit jeweils 1152 Pixel, bei 50 oder 60Hz.

6. Computeranimationen

Die Animation eines Sachverhaltes bedeutet wörtlich, diesen "zum Leben zu erwecken". Eine Animation betrifft hierbei alle Änderungen, die einen visuellen Effekt bedingen, wie Bewegung, Formen, Farben, Transparenz, Struktur und Musterung.

Ein sehr altes Beispiel des menschlichen Wunsches, Bewegung in Zeichnungen darzustellen, findet sich in einer Höhle in den Pyrenäen. Die 30.000 Jahre alte Höhlenmalerei zeigt einen Wildebeest mit vier anstatt zwei Beinpaaren. Diese Technik der Mehrfachzeichnung von Gliedmaßen wird auch heute noch in Comics verwendet. Der Comicleser interpretiert dieses intuitiv als Bewegung. Die Technik mehrere Zeichnungen nebeneinander anzuordnen, um so einen zeitlichen Verlauf eines Ereignisses nachzuempfinden, ist schon in japanischen Bildrollen aus dem 13. Jhrh. und in dem ägyptischen Buch der Toten Vorläufer enthalten. Der Wechsel zwischen verschiedenen Darstellungsarten, wie Totale oder Großaufnahme, gehörten schon damals zum Repertoire.

Bevor Animationen in einem Computer eingesetzt werden können, müssen diese in digitalisierter Form vorliegen, sog. Schlüsselbilder. Die Bilder können im Computer selbst erstellt werden oder durch Digitalisierung von Fotos oder Zeichnungen erzeugt werden. Solche Bilder werden als Frames bezeichnet.

Die Animation der Bewegung von einer Position zu einer anderen erfordert die Komposition von Einzelbildern mit Zwischenpositionen zwischen den Schlüsselbildern. Das System erhält Anfangs- und Endposition, die im einfachsten Fall durch lineare Interpolation (Lerping) überführt wird.

Um die Filmtechniken unterscheiden zu können, werden im folgenden die Definitionen und Unterschiede der einzelnen Begriffe beschrieben:

- Zeichentrickfilm. Wie der Name schon sagt, besteht die Animation aus einzelnen Zeichnungen. Dabei ist besonders zu berücksichtigen, daß nur durch Zeichnungen kein dreidimensionaler Eindruck gewonnen werden kann.
- Animationsfilm / Einzelbildtechnik. Die Terminologie ist hier leider nicht eindeutig. In der Literatur wird unter dem Begriff Animationsfilm eine Animation verstanden, die aus Fotos von realen Objekten zusammengesetzt ist.
- Schauspiel im und Video. Eine Videokamera nimmt in gleichmäßigen Abständen Bilder einer Szene mit realen Bewegungen auf. Der Unterschied zum Animationsfilm besteht darin, daß bei der Animation einzelne Fotos einer Szene erst später zu einem Film verbunden werden. Damit lassen sich Bewegungen erstellen, die nie in der Realität stattgefunden haben.

Computeranimationen werden üblicherweise in einer höheren Programmiersprache, wie z.B. LISP, dargestellt. Die Primitiven der Sprache beinhalten Vektoren, Farben, Polygone, Oberflächen, Gruppen verschiedene Sichten, Unterwelten und Beleuchtungsaspekte.

6.1 Klassische Animationstechniken

6.1.1 Dynamik in stehenden Bildern

Der Eindruck eines dynamischen Vorgangs muß nicht immer an ein zeitkontinuierliches Medium gebunden sein. Schon ein einzelnes Bild kann viel über Zustandsänderungen einer Szene aussagen. Einige Techniken, in stehenden Bildern Dynamik zu erzeugen, werden hier vorgestellt:

- Körperhaltung. Der Betrachter kann anhand der Körperhaltung einer Figur deren Bewegung und Bewegungsrichtung erkennen. So kann ein Läufer von einem stehenden Menschen unterschieden werden und die Bewegungsrichtung des Läufers ist erkennbar.
- Deformierung eines Objektes kann z.B. beim freien Fall oder auch bei Kollisionen eingesetzt werden. Aus der veränderten Form kann der Betrachter auf Aufschlagstärke und die vorherige Bewegungsrichtung der Objekte schließen.
- Wellenformen bedeuten für den Betrachter fast immer eine Bewegung oder (periodische) Zustandsänderung eines Objektes. So werden Wasser, Rauch, Flammen oder Dampf immer wellenförmig gezeichnet.
- Explosionen und Blitze sind Beispiele dafür, wie sich komplexe Abläufe durch eine einfache Zeichnung darstellen lassen. Der Betrachter kann bei geschickter Konstruktion die Bewegungsrichtung, Ausmaß und gleichzeitig die Wirkung auf die Figuren ermessen.
- Spuren zeigen den vollendeten Weg eines Objektes an. Dadurch läßt sich z.B. eine Reise dokumentieren.
- Unschärfe oder Verwackeln ist aus der Fototechnik bekannt und vermittelt dem Betrachter den Eindruck, daß sich das unscharfe Objekt im Moment der Aufnahme schnell bewegt hätte.

Beim Erstellen eines Bildes muß der Zeichner natürlich die Gesetze der Perspektive und die Naturgesetze, wie Schwerkraft und Trägheit der Körper, beachten. Dies gilt sowohl in einem Film wie auch in einzelnen Bildern.

6.1.2 Der Zeichentrickfilm

Der Zeichentrickfilm bietet aufgrund seiner inzwischen langen Geschichte ein breites Repertoire an Animationstechniken. Die einfachste Art, eine Szene zu erstellen, besteht darin, jedes einzelne Bild komplett zu zeichnen und abzufotografieren. Die in der Videotechnik vorgegebene Anzahl von 24 Bildern pro Sekunde dient als Richtlinie für Planung und Berechnung der einzelnen Bewegungsabläufe. Dabei genügt es allerdings zwölf Zeichnungen pro Sekunde zu erstellen, die dann jeweils zweimal abfotografiert werden. Dies reicht aus, um dem Zuschauer eine fließende Animation zu bieten.

Bevor mit der Erstellung des Zeichentrickfilms begonnen werden kann, sind Vorarbeiten zu leisten. Der Inhalt des Films wird in einem Drehbuch festgehalten. Das Drehbuch beschreibt jedoch nur rein textuell die Handlung. Für die Regie ist jedoch notwendig einen zeitlichen Ablaufplan mit visuellen Elementen zu erstellen. Aus dem Drehbuch wird darum ein Storyboard erstellt, in dem entlang einer Zeitleiste die Handlungsführung, Szenenwechsel und Kamerabewegungen aufgeführt sind. Anhand der Zeitleiste kann dann genau berechnet werden, wieviele Phasen für eine Bewegung gezeichnet werden müssen. Außerdem werden Charakterzüge und Mimik der Figuren sowie grobe Landschaftsskizzen entworfen. Das Storyboard ist sowohl für die Regie, als auch für die Zeichner und den Kameramann

wichtig. Deshalb wird das Storyboard in ein Layout für die Zeichner und einen Fahrplan für den Kameramann aufgeteilt:

- Das Layout enthält die Unterteilung der Szenen in ständig bewegte, zeitweise bewegte und unbewegte Objekte. Für bewegte Objekte werden Schlüsselzeichnungen erstellt. Schlüsselzeichnungen repräsentieren wichtige Aktionen innerhalb einer Animation. So werden das Auftreten von Charakteren, Bewegungsänderungen in Geschwindigkeit und Richtung oder Änderungen der Perspektive als Schlüsselbild festgehalten. Ist das Herstellungsteam des Films mit dem Layout zufrieden, können die fehlenden Phasen zwischen den Schlüsselbildern gezeichnet werden.
- Der Fahrplan enthält die typischen Kameraanweisungen, wie Fahrt, Zoom, Schnitte, Auf-, Ab- und Überblendungen. Ebenso wird die Beleuchtung detailliert festgehalten.

Die verschiedenen Objekte und Hintergründe werden auf Folien gezeichnet, die beim Abfotografieren auf einem speziellen Tricktisch von hinten beleuchtet werden, um einen höheren Kontrast zu bewirken. Zusätzlich können mehrere Lichtquellen eingesetzt werden, um bessere Farbsättigung und Beleuchtungseffekte zu erzielen. Wichtig ist, daß Folien und Kamera am Tricktisch fixiert sind, damit es keine unbeabsichtigten Verschiebungen und später keine unbeabsichtigten Bewegungen im Film gibt.

Die Folien, die die verschiedenen Phasen einer Animation darstellen, werden hintereinander auf den Hintergrund gelegt und abfotografiert. Außerdem können die Folien dabei verschoben werden, um eine Bewegung zu erzeugen. Neben der Bewegung des Objektes kann auch ein Panorama hinter den Figuren stückweise bewegt werden. Später im Film erscheint es so, als ob sich die Figur entlang einer Landschaft bewegt. Durch die Verwendung mehrerer Folien übereinander läßt sich eine dreidimensionale Szene vortauschen. Eine Folie, der Horizont, bleibt während der Aufnahme unbewegt. Weiter entfernte Gegenstände, z.B. ein Gebäude, bewegen sich langsam. Nahe Objekte wie die Hauptfiguren werden dementsprechend schneller bewegt.

Zeichentrickfilm läßt sich einfach mit Realszenen verbinden. Die Hintergrundbeleuchtung des Tricktisches wird dabei durch einen Spiegel ersetzt, auf dem der Realfilm projiziert wird. Die Folien werden dann auf den Tricktisch fixiert und bilden so eine Einheit mit dem Realbild.

6.1.3 Der Animationsfilm

Um Plastilinpuppen, Einrichtungsgegenstände oder sonstige starre Gegenstände zu bewegen, bedarf es einer genauen Planung der Bewegungsabläufe. Objekte und Kamera müssen millimetergenau für jede Aufnahme verändert werden. Wichtig dabei ist, daß die Veränderungen gleichmäßig geschehen. Im Gegensatz zum Zeichentrickfilm, in dem das Ergebnis schon anhand der Folien im voraus zu kontrollieren ist, liegt die Hauptarbeit hier beim Abfotografieren.

Soll der Film animierte Gegenständen und Schauspielern gleichzeitig enthalten, bedeutet es für die Darsteller eine besondere Anstrengung an Körper und Disziplin. So muß der Darsteller mehrere Minuten in einer bestimmten Stellung verharren und darauf warten, bis Kamera, Beleuchtung und Objekte für die nächste Aufnahme verändert wurden. Verändert der Schauspieler seine Position oder Mimik, muß oft die gesamte Szene von vorne gedreht

werden, da die ursprüngliche Position mit allen Einzelheiten nur schwer wiederhergestellt werden kann.

Ist die Arbeit zwischen Zeichner, Regisseur und Kameramann beim Zeichentrickfilm in zeitliche, voneinander unabhängige Phasen unterteilt, so müssen beim Animationsfilm alle Beteiligten gleichzeitig zusammenwirken. Der große Vorteil dieser Art des Filmens liegt in der vollständigen Dreidimensionalität der Szenen. Allerdings ist die Animation fast nur auf Bewegung beschränkt. So können Farbwechsel, Größen- und Formänderungen starrer Objekte nur schwer realisiert werden.

6.2 Die Computeranimation

Die Computeranimation kann unter verschiedenen Aspekten betrachtet und in Bereiche unterteilt werden. Man unterscheidet zwischen Bewegungs- und Zustandsanimation. Die Animationstechniken unterscheiden sich auch aufgrund der verschiedenen Computergrafikformate. In der 2D-Technik werden Grafiken meist in Pixmaps angelegt. Die 3D-Technik arbeitet nur mit Vektorgrafiken.

In den folgenden Kapiteln werden Pixmap-Animation und Morphing im 2D beschrieben. Anschliessend wird die Modellierung von 3D-Welten und die Computerunterstützte Animation dargestellt.

Die Anwendungsgebiete der Computeranimation sind z.B. Spiele, Präsentationen und Simulationen. Die Begriffe Animation und Simulation werden oft nicht sauber getrennt. Die Simulation ist allgemein eine Nachbildung von Objektverhalten unter bestimmten physikalischen Gesetzen. Die Animation zeigt die wahrnehmbaren Veränderungen der Objekte in einer bestimmten Zeitspanne. Simulationen können allerdings vollkommen ohne Visualisierung stattfinden und bieten nur ein Endergebnis, gehören damit also nicht zur Animation.

6.2.1 2D-Technik

6.2.1.1 Pixelgrafik

Pixmaps werden mit einem Malprogramm erstellt. Zudem besteht die Möglichkeit eine Vorlage einzuscannen und im Computer weiterzuverarbeiten. Malprogramme bieten Funktionen zum Erzeugen und Manipulieren von Bildern. Verschiedene Elemente wie Kreise, Rechtecke, etc. können in das Bild eingefügt werden. Bildausschnitte können z.B. vergrößert, gelöscht oder verfärbt werden. Spezielle Programme zum Bearbeiten von Fotos bieten zudem Funktionen, wie Kontrast- und Helligkeitsänderungen oder Farbiliter an.

Die 2D-Grafiktechnik des Computers ähnelt in vielen Bereichen dem Zeichentrick. Die Trennung von bewegten Figuren und (mehrere) Hintergründen ist genauso wie in der analogen Folientechnik möglich. Diese Technik wird beispielsweise in Spielen eingesetzt. Animierte Pixmaps(Sprites) werden über einen starren Hintergrund bewegt.

6.2.1.2 2D-Morphing

Ein Beispiel für die einfache Erstellung einer Animation mit Hilfe des Computers ist das Morphing. Beim Morphing wird gleichzeitig die Form und die Farbe eines Bildes verändert. Ziel ist es, einen fließenden Übergang zwischen zwei Pixmaps zu erzeugen. Dabei muß der Benutzer markante Punkte in Anfang- und Endbild bestimmen. Zum Beispiel werden bei Porträtfotos Augen, Mundwinkel, Nase, etc. markiert. Die Punkte werden durch ein Netz verbunden, welches die Pixmaps in Bildbereiche ein.

Um eine Veränderung der Gesichtform zu erzielen, werden die Punkte von der Anfangsposition in die Endposition verschoben. Dadurch verändert sich allerdings auch das Netz. Der Computer skaliert die Bildbereiche auf die neue Form. Die Pixel der Zwischenbilder werden aus den Farben des Anfangs- und Endbilds gemischt. Dabei nimmt der prozentuale Anteil des Endbildes zu.

6.2.1.3 Vektorgrafik

Bei der Erzeugung von Computergrafik kann zwischen dem direkten und indirekten Modus unterschieden werden: Grafiksysteme, die Bilder im direkten Modus erzeugen (z.B. Malprogramme) können zwar vordefinierte Formen wie Rechtecke, Kreise oder Polygone in die Pixmap einfügen, der Benutzer kann jedoch nicht eine schon gezeichnete Form auswählen und verändern. Das Programm fügt die Formen direkt in die Grafik ein, ohne die Formen separat abzuspeichern. Manipulationen können auf Pixelebene vorgenommen werden. Eine Pixelmenge, häufig ein rechteckiger Bildausschnitt, kann ausgewählt und beispielsweise in Farbe und Größe verändert werden.

6.2.2 3D-Technik

Den Arbeitsablauf einer Animationserstellung im 3D-Raum kann in vier Bereiche gegliedert werden, die im folgenden genauer erläutert werden:

1. Im Modeller legt der Benutzer Form, Größe und Position der Objekte im dreidimensionalen Raum fest.
2. Der Shader definiert Oberflächenstrukturen für Objekte mit optischen Eigenschaften wie Transparenz- und Reflexionsverhalten. Die einzelnen Objekte werden zu einer Szene zusammengefaßt. Außerdem werden Lichtquellen und Kamera positioniert.
3. Der Renderer berechnet (rendert) aus den Daten des Modellers und Shaders eine Grafik. Je nach verwendeten Renderverfahren werden Schattenwurf der Objekte, und Reflexion und Transparenz der Oberflächen berücksichtigt.
4. Im Animationsmodul kann der Benutzer Bewegungen und Zustandsänderungen der Objekte, der Kamera und der Lichter festlegen.

6.2.2.1 Die Konstruktion im Modeller

Modeller ermöglichen dem Benutzer die Eingabe von Grafikobjekten und deren Platzierung im dreidimensionalen Raum. Für die interne Verarbeitung und Darstellung der Körper im Modeller gibt es zwei vorherrschende Verfahren. 3D-Konstruktionsprogramme werden deshalb generell in zwei Arten unterschieden:

Der Polygonmodeller setzt Objekte aus einer Vielzahl kleiner Flächen (Polygone) zusammen. Polygone sind im Computer leicht zu verarbeiten. Die Speicherung und Berechnung einer Grafik aus Polygondaten ist sehr effizient. Allerdings fehlen oft die nötigen Details. Runde

Formen wie bei Kugeln werden nur näherungsweise dargestellt. Deshalb können diese Körper nicht beliebig vergrößert und verformt werden.

Der Spline-Modeller beschreibt Kurven und Körper anhand von Formeln. Splines berechnen Kurven aus zwei Endpunkten und zwei Kontrollpunkten, die die Stärke der Auslenkung eines Kurvenabschnitts bestimmen. Veränderungen können ohne Qualitätsverlust erfolgen. Außerdem lassen sich Oberflächenstrukturen (Texturen) leicht anpassen. Ein Nachteil ist der längere Aufbau einer Ansicht und Schwächen bei der Durchführung boolescher Mengenoperationen.

6.2.2.2 Der Shader

Das Shading bestimmt die Beleuchtung, Oberflächengestaltung, Kameraposition und die Schattierungsqualität.

- Beleuchtung und Kamera: Der Benutzer kann im Shader zwischen drei Lichtquellen wählen:
 - Ambientes Licht simuliert das natürliche Streulicht oder Umgebungslicht. Es hat keine Richtung, und somit können Objekte keine Schatten werfen.
 - Parallellicht oder gerichtete Lichtquelle trifft immer im gleichen Winkel auf die Objekte. Die Lichtquelle befindet sich sozusagen im Unendlichen. Diese Lichtart simuliert das Sonnenlicht.
 - Lichter mit Position im Raum können verschiedene Arten von Scheinwerfern wie Kugel-, Klappenscheinwerfer, Halogen- und Laserlicht simulieren.
 - Zu der Modellierung von Lichtquellen gehört die Definition der Position, Form (beispielsweise Punkt, Fläche und Kugel), der Intensität, mit der die Lichtquelle strahlt, und dessen Farbe. Für die Kamera kann der Benutzer dessen Position bestimmen. Einige Shader bieten zudem noch unterschiedliche Objektarten (z.B. Weitwinkel) und Farb- oder Nebelfilter an.
- Oberflächengestaltung: Ein Shaderprogramm bietet dem Benutzer die Möglichkeit, Oberflächen für die im Modeller erstellten Objekte zu erzeugen. Oberflächen können Muster aufweisen, jedoch auch spezielle Leuchteigenschaften besitzen. Die Reflexionseigenschaften verschiedener Materialien unterscheiden sich in diffuser und spiegelnder Reflexion. Die diffuse Reflexion verteilt reflektierendes Licht gleichmäßig in alle Richtungen. Das Objekt wirkt dadurch stumpf (z.B. Kreide). Andere Objekte glänzen und reflektieren das Licht in eine bestimmte Richtung (z.B. Spiegel). Dieses nennt sich spiegelnde Reflexion. Hinzu kommt noch, dass ein Objekt lichtdurchlässig (transparent) sein kann. Ein Lichtstrahl kann von einem transparenten Objekt in seiner Richtung abgelenkt werden (Lichtbrechung). Trifft ein Lichtstrahl senkrecht auf eine Oberfläche, weist die Oberfläche oft ein Glanzlicht auf. Die Stärke dieses Effektes lässt sich in den meisten Shading-Programmen steuern.

6.2.2.3 Rendering

Rendering bezeichnet die Berechnung einer Grafik aus den Daten des Modellers und Shaders. Zum Rendering gehören die Ermittlung der vom Betrachter aus sichtbaren Flächen eines Objekts, Schattenwurf und Oberflächenschattierung der Objekte. Die einfachste und schnellste Renderingmethode ist Wireframe. Objekte werden als einfaches Drahtgittermodell dargestellt. Dabei sind auch die Rückseiten eines Objektes sichtbar. Lichtquellen, Schattenwurf und Oberflächenstrukturen werden nicht berücksichtigt. Hidden-Line arbeitet

nach der gleichen Renderingtechnik wie Wireframe. Allerdings werden nur die vom Betrachter sichtbaren Flächen dargestellt.

7. Datenkompression

Einzelbilder sind gegenüber dem Medium Text durch eine nrelativ hohen Speicherplatzbedarf gekennzeichnet. Audio und insbesondere Video haben diesbezüglich noch höhere Anforderungen. Auch die entstehenden Datenraten bei der Verarbeitung und Übertragung kontinuierlicher Medien sind nicht unerheblich.

7.1 Grundlagen

Unkomprimierte Grafiken, Audio- und Videodaten erfordern eine beträchtliche Speicherkapazität, die im Falle unkomprimierter Videodaten nicht einmal mit der heutigen CD- oder DVD-Technologie realisierbar ist. Unterschiedliche Kompressionsverfahren nehmen meist dieselben Aufgaben wahr, teilweise auch komplementäre.

Die Anforderungen an die Kodierung basiert auf den vorliegenden Rohdaten:

- Text: Jedes Zeichen mit einer 8x8 Pixel Matrix auf eine 2 byte großen Codetabelle (Unicode) liefert 4800 Zeichen pro Bildschirmseite bei einem Speicherplatz von 9600 byte.
- Vektorbild: Zur Darstellung eines typischen Vektorbildes geht man von durchschnittlich 500 Geraden aus, die durch Ihre Anfangs- und Endpunkte. sowie einem 8bit großem Attributfeld. Daraus ergeben sich pro Bildschirmseite einer 640x480 Matrix
 $\text{Log}_2(640) = 10 \text{ bit}$
 $\text{Log}_2(480) = 9 \text{ bit}$
2 Punkte a 19 bit zzgl. 8 bit Attribut = 46bit
 $500 \times (46/8) = 2875 \text{ byte}$
- Pixelbild: Ein Pixelbild bei 640x480 und 256 Farben (in einem byte darstellbar) ergibt:
 $640 \times 480 \times 1 \text{ byte} = 307.200 \text{ byte} = 300 \text{ Kbyte}$
- Sprache: Unkomprimierte Sprache in Telefonqualität wurde mit 8kHz abgetastet und mit 8 bit quantisiert, dies ergibt einen Datenstrom von 64Kbit/s und damit einen Speicherbedarf von
 $(64/8) \times (1/1024) = 8 \text{ Kbyte/s}$
- Stereo-Audio: Ein unkomprimiertes Stereo-Audiosignal in CD-Qualität wird mit 44.1kHz abgetastet und mit 16bit quantisiert, dies ergibt eine Datenrate von
 $2 \times 44,1 \times (16/8) = 176400 \text{ byte/s} = 172 \text{ kbyte}$
- Video: Eine Videosequenz besteht aus 25 Vollbilder pro Sekunde mit jeweils 625 Zeilen mit 833 Pixel (PAL), Luminanz und Chrominanz werden getrennt digitalisiert und im Multiplexverfahren übertragen und seien zusammen mit 3 byte codiert. Damit ergibt sich eine Datenrate von $625 \times 833 \times 25 \times 3 = 39046875 \text{ byte / s}$ und ein Speicherplatzbedarf von
 $39046875 / 1024 / 1024 = 37 \text{ Mbyte / s}$.

Mit Hilfe von geeigneten Kompressionsverfahren lassen sich diese Werte jedoch erheblich reduzieren. Diese Techniken sind integrale Bestandteile von Multimediasystemen.

Die Anforderungen an ein Multimediasystem lassen sich folgendermaßen definieren:

- Die Qualität der komprimierten und dekomprimierten sollten möglichst gut sein
- Die Komplexität des Verfahren sollte möglichst einfach sein, um hohe Laufzeiten zu vermeiden und effektive Realisierungen zu ermöglichen
- Die Algorithmen sollten bei der Dekodierung und evtl. auch bei der Kodierung bestimmte Zeitschranken nicht überschreiten, Dekodierung muß sich dem Echtzeitanspruch stellen.

Wobei einige Verfahren Schwerpunkte auf die Kodierung legen (Dialog-Modus, wie z.B. bei Video-conferencing), andere auf die Dekodierung (Abfrage-Modus, wie z.B. bei Auskunftssystemen).

- Das Format sollte unabhängig von Bildschirmgröße und Bildwiederholfrequenz sein. Damit können die Daten auf unterschiedlichen angezeigt werden.
- Für Audio und Video sollten verschiedene Datenraten möglich sein.
- Audio und Video sollten exakt synchronisierbar sein. Eine solche sollte auch mit anderen Medien von einem Systemprogramm realisiert werden können.
- Eine Realisierung sollte einfach und kostengünstig sein.
- Das Verfahren sollte eine Kooperation verschiedener Systems ermöglichen, z.B. um Daten auf anderen Medien zu transportieren.

7.2 Kodierungsverfahren

Die oftmals auf Audio- und Videodaten angewendeten Kompressionsverfahren in Multimediasystemen bestehen selbst wieder aus mehreren unterschiedlichen Verfahren. Die einfachsten Verfahren basieren auf einer Interpolation, bei der man sich Eigenschaften des menschlichen Auges bzw. Ohres zu Nutzen macht. So reagiert das Auge empfindlicher auf Helligkeits- als auch Farbänderungen, damit kann eine RGB-Darstellung in eine YUV bzw. YIQ umgewandelt werden.

7.2.1 Lauflängenkodierung

Viele Daten bestehen aus Folgen von identischen Bytes. Ist deren Anzahl ausreichend groß, so kann eine erhebliche Reduktion erreicht werden, indem die Anzahl der wiederholenden Bytes angegeben wird. Eine spezielle Markierung (m-byte), die selbst nicht als Datenbestandteil auftritt. Als Beispiel sei ein Ausrufezeichen als m-byte definiert; so ist ein einzeln auftretendes Ausrufezeichen ein m-byte, zwei Ausrufezeichen hintereinander als Dateneintrag interpretiert.

7.2.2 Nullunterdrückung

Die Lauflängenkodierung ist eine Verallgemeinerung der Nullunterdrückung. Hier geht man von einem bestimmten, sich oft wiederholenden Zeichen aus. In Texten ist dies beispielsweise das Leerzeichen. Einzelne Leerzeichen bleiben unberücksichtigt; ab drei Bytes in Folge werden Sie durch ein m-byte und eine Anzahl ersetzt. So sind maximale 257 Byte auf 2 Byte reduzierbar.

7.2.3 Vektorquantisierung

Ein Datenstrom wird in Blöcke zu n Bytes unterteilt. Für jeden Block wird in einer Tabelle noch einem dort aufgeführten Muster gesucht, das nach einem festzulegenden Kriterium diesem Block am ähnlichsten sieht. Jedem Muster ist in der Tabelle ein Indexwert zugeordnet. Eine solche Tabelle kann mehrere Dimensionen beinhalten, d.h. der Index ist hier ein Vektor. Der korrespondierende Dekodierer verfügt über die gleiche Tabelle und generiert aus dem Vektor eine Approximation des ursprünglichen Datenstroms.

7.2.4 Pattern Substitution

Ein Verfahren, das sich gut zur Kompressions von Texten eignet, ersetzt häufig vorkommende Muster durch einzelne Bytes. So werden bspw die Terminalsymbole in einer Programmiersprache (If, then, else) ersetzt.

7.2.5 Diatonic Encoding

Eine Variante, die auf der Zusammenfassung von jeweils zwei Datenbytes beruht. Hierfür werden die am häufigsten auftretenden Byte Paare ermittelt und durch einzelne, im Text nicht vorkommenden, Bytes ersetzt.

7.2.6 Statistische Kodierung

Verschiedene Zeichen müssen nicht grundsätzlich mit fester Anzahl von Bits kodiert werden. Wie beim Morsealphabet werden häufig auftretende Zeichen durch kurze Sequenzen und seltener auftretende Zeichen durch längere Sequenzen ersetzt. Die statistische Kodierung richtet sich nach der Häufigkeit des Auftretens einzelner Zeichen oder Folgen von Datenbytes.

7.2.7 Huffman-Kodierung

Die zu kodierenden Zeichen werden mit der Wahrscheinlichkeit ihres Auftretens bewertet und dann die häufigsten Zeichen mit der minimalen Bitsequenz (Codewörter) belegt. Zur Ermittlung eines Huffman-Codes kann man sukzessive einen binären Baum aufbauen. Die Blätter stellen die zu kodierenden Zeichen dar. Alle Knoten beinhalten die relative Wahrscheinlichkeit des Auftretens eines der in diesem Unterbaum befindlichen zu kodierenden Zeichen. Die Kanten werden jeweils mit den Bits 0 und 1 versehen.

Beispiel mit

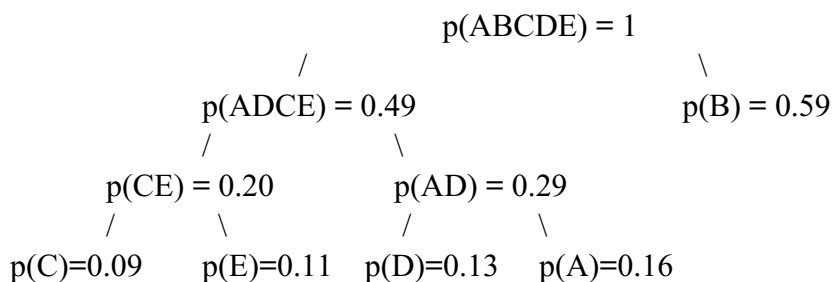
$$p(A) = 0.16$$

$$p(B) = 0.51$$

$$p(C) = 0.09$$

$$p(D) = 0.13$$

$$p(E) = 0.11$$



$$w(A) = 001$$

$$w(B) = 1$$

$$w(C) = 011$$

$$w(D) = 000$$

$$w(E) = 010$$

7.2.8 Arithmetische Kodierung

Wie die Huffman-Kodierung wird eine Statistik gebildet, allerdings nicht über einzelne Buchstaben, sondern in Abhängigkeit zum vorangegangenen Zeichen. Damit muss der kodierte Datenstrom immer sequentiell vom Anfang gelesen werden, erreicht allerdings die gleiche mittlere Kompressionsrate.

7.2.9 Transformationskodierung

Hier werden Daten in einem anderen mathematischen Raum transformiert, der sich besser für eine Kompression eignet. Es muß dabei immer eine inverse Transformation (also eine Ein-Eindeutige Zuordnung) möglich sein.

7.2.10 Prädiktion oder relative Kodierung

Ist bspw. eine Folge von Zeichen deutlich von Null verschieden, jedoch nicht stark von einander abweichend, so kann die Bildung der jeweiligen Differenz zum vorherigen Wert Vorteile bringen. Im Bezug auf ein Bild wirken Kanten als große, Flächen mit ähnlicher Luminanz und Chrominanz als kleine Werte aus. Eine homogene Fläche wäre durch eine Vielzahl von Nullen charakterisiert. Diese kann dann anschließend mit einer Lauflängenkodierung weiter komprimiert werden.