

VIDEO-DATENBANKEN

&

VIDEO-ON-DEMAND

EINE EINFÜHRUNG

Von

Michael Hurler

Student der Wirtschaftsinformatik am Fachbereich 1 der
Technische Hochschule Darmstadt

Im Rahmen des Seminars
„Information und Kommunikation“
des Information Transfer Office
im Wintersemester 1996/97

Betreuer: **Thomas Kunkelmann**

Zusammenfassung

Diese Arbeit gibt einen Überblick über Anwendungsgebiete, Technologien, Implementierungen und Probleme von Videodatenbanken und die damit verbundene Funktion Video-on-Demand, d.h. der Möglichkeit Video-Informationen in (nahezu) Echtzeit über einen Rechner-Netz abzurufen. Dabei wird bei den einzelnen Aspekten von Videodatenbanken und Video-on-Demand exemplarisch auf Implementierungen von speziellen Systemen eingegangen.

Zusätzlich wird auf den Aspekt der Indizierung und somit des späteren Recherchierens von Filmen bzw. Teilen von Filmen eingegangen. Dieser Bereich scheint von vielen Entwicklern eher vernachlässigt zu werden, obwohl ihm eine wichtige Rolle zukommen muß, denn eine große Videodatenbank mit unter Umständen vielen tausend Stunden Filmmaterial ist nutzlos, wenn es nicht möglich ist aus dieser Menge Material die gesuchten Szenen herauszusuchen.

Inhaltverzeichnis

1	Einführung	1
1.1	Einsatzgebiete von VoD	2
1.2	Anforderungen an die Qualität des VoD-Dienstes	4
1.3	Die „Natur“ von Videodaten.....	6
2	Video-Datenbanken & Video on demand.....	8
2.1	Eine System-Architektur.....	8
2.1.1	Die verteilte hierarchische Architektur.....	9
2.1.2	Das „Compound Media Object Model“ (CMO-Model).....	10
2.1.3	Das „Distributed Hierarchical Storage Management“	12
2.1.4	Zugriff auf die Daten	14
2.1.5	Cache-Management	16
2.1.6	Parallelisierung von Servern.....	17
2.2	Sicherung der Übertragung	17
2.2.1	Kontinuität	17
2.2.2	Gleichmäßige Verteilung der Netzlast.....	21
3	Indizierung von Videodaten	23
3.1	AMPHORE.....	23
3.1.1	Das Daten-Schema.....	24
3.1.2	Daten-Erfassung und -Recherche	26
3.2	Indizierung im BDVS	27
3.2.1	Datenbank-Schema	28
3.2.2	Datenerfassung und -recherche.....	32
3.3	Vereinfachung der Indizierung in der Zukunft	32
4	Versuchssysteme	33
4.1	Full Service Network.....	33
4.2	BT Interactive Television	34
5	Resümee	35

Literaturverzeichnis

Video-Datenbanken & Video on demand

- [BT94] **BT to Trial Interactive TV**
BT News release (NR9479), 15. November 1994
- [BT96] **BT Successfully Completes Interactive Television Trial**
BT News release (NR9653(R)), 16. Juli 1996
URL: <http://www.bt.com/newsroom/document/nr9653.htm>
- [FED94] **A Distributed Hierarchical Storage Manager for a Video-on-Demand System**
Autoren: Craig Federighi, Lawrence A. Rowe (University of California at Berkeley)
für: Storage and Retrieval for Image and Video Databases II, IS&T/SPIE Symp. on Elec. Imaging Sci. & Tech., San Jose, CA, Februar 1994
- [FEN96] **Providing VCR Functionality in a Constant Quality Video-On-Demand Transportation Service**
Autoren: Wu-chi Feng, Farnam Jahanian, Stuart Sechrest (University of Michigan)
in: IEEE - Proceedings of MULTIMEDIA '96, S. 127-135
- [HOL96] **A Buffer-Triggered Smooth Adaptation Technique for Time-Dependent Media**
Autoren: Silvia Hollfelder, Achim Kraiß, Thomas C. Rakow
in: Arbeitspapiere der GMD Nr. 1002 - Juni 1996
- [HUO96] **Video on Demand: A Survey**
Autoren: Jani Huoponen, Thorsten Wagner
für: EE 4984: Telecommunication Networks - Project 1, 31. April 1996
- [MOU96] **Issues in the design of a storage server for video-on-demand**
Autor: Antoine N. Mourad (AT&T Bell Laboratories)
in: Multimedia Systems (1996) 4: S. 70-86, © Springer-Verlag 1996
- [ORA1] **Oracle Video Option™**
Oracle Informationen im WWW
URL: <http://www.oracle.com.sg>
- [ORA2] **Oracle Video Option™**
An Oracle Q&A, July 1996
URL: <http://www.oracle.com>
- [RAK95] **Multimedia Database Systems - The Notions and the Issues**
Autoren: Thomas C. Rakow, Erich J. Neuhold, Michael Löhr
für: Arbeitspapiere der GMD, St. Augustin, Februar 1995
- [RAJ95] **Video on Demand**
Autoren: Harindra Rajapakshe, Derek Paul Quek
Juni 1995
URL: http://www-dse.doc.ic.ac.uk/~nd/surprise_95/journal/vol4/shr/report.html
- [ROW95] **The Berkeley Distributed Video-on-Demand System**
Autoren: Lawrence A. Rowe, David A. Berger (University of California at Berkeley)
URL: <http://www-plateau.cs.berkeley.edu/papers/NEC95.html>
- [ROW95a] **A Distributed Hierarchical Video-on-Demand System**
Autoren: Lawrence A. Rowe, John S. Boreczky, David A. Berger, David W.

Brubeck, J. Eric Baldeschwieler (University of California at Berkeley)
 URL: <http://www-plateau.cs.berkeley.edu/papers/ICIP95.html>

- [TEW96] **Design and Performance Tradeoffs in Clustered Video Servers**
 Autoren: Renu Tewari, Rajat Mukherjee, Daniel M. Dias, Harrick M. Vin
 (IBM Research Division, University of Texas in Austin)
 in: IEEE - Proceedings of MULTIMEDIA '96, S. 144-150
- [TW96] **Time Warner Cable's Full Service Network™ SERVICES**
 Dokument vom Time Warner-Full-Service-Network-WWW-Server
- [TW96a] **Time Warner Network Architecture Q&A**
 Dokument vom Time Warner-Full-Service-Network-WWW-Server

Indexierung von Videodaten

- [HAR] **Motion Video Annotation and Analysis: An Overview**
 Autor: Harpreet S. Sawhney (IBM Almaden Research Group)
- [SÜL96] **AMPHORE - Ein Arbeitsplatz zur Filmdokumentation**
 Autor: Klaus Süllow
 in: nfd - Nachrichten für Dokumentation, Jg. 47 (1996), Nr. 2, S. 67-74
- [ROW94] **Indexes for User Access to Large Video Databases**
 Lawrence A. Rowe, John S. Boreczky, Charles A. Eads (University of California at Berkeley)
 für: Storage and Retrieval for Image and Video Databases II, IS&T/SPIE
 Symp. on. Elec. Imaging Sci. & Tech., San Jose, CA, Februar 1994
 URL: <http://www-plateau.cs.berkeley.edu/papers/VodsDB94.html>

Zusätzliche Literatur:

- [ATS95] **Digital Audio Compression Standard (AC-3)**
 Advanced Television Systems Committee
 Doc. A/52 vom 20. Dezember 1995
- [ATS95a] **ATSC Digital Television Standard**
 Advanced Television Systems Committee
 Doc. A/53 vom 16. September 1995
- [PAT] **Performance of a Software MPEG Video Decoder**
 Ketan Patel, Brian C. Smith, Lawrence A. Rowe (University of California at Berkeley)
- [ROW94a] **Video Compression - *What to do when everything is changing?***
 Autor: Lawrence A. Rowe (University of California at Berkeley)
 für: Invited Talk Usenix 1994
- [SMI] **Fast Software Processing of Motion JPEG Video**
 Autor: Brian C. Smith (Cornell University)

Abkürzungsverzeichnis

AS	Archiv Server
ATV	Advanced Television
BDVS.....	Berkeley Distributed VoD-Server System
CMO	Compound Media Objekt (Zusammengesetztes Medien-Objekt)
DVD.....	Digital Versatile Disk (Designierter Nachfolger der CD)
HDTV	High Definition Television
OID	Object Identifier (Objekt-Bezeichner)
PAL-TV	PAL-Fernseh-Norm
PAL-VHS.....	PAL-VHS-Videorecorder-Norm
VDB	Video-Datenbank-Browser
VFS	Video-Fileserver
VoD.....	Video on Demand

1 Einführung

Die Themen Video-Datenbanken und Video-on-Demand, im folgenden kurz als VoD zusammengefaßt, sind in den letzten Jahren zu einem populären Thema in Forschung und kommerzieller Anwendung geworden. Dies hängt mit den Vorteilen elektronischer Speicherung und Verarbeitung von Informationen zusammen, wie z.B. dem schnelleren Zugriff auf die Daten, der leichteren Weiterverarbeitung dieser Daten, den Möglichkeiten von Interaktion zwischen dem Empfangenden von Bild/Ton-Informationen und dem Aussendenden dieser Informationen und vielem anderen mehr.

Wie VoD technisch beim Nutzer (Client) realisiert wird, hängt außerdem vom Einsatzzweck ab. Als Client kann ein Computer dienen, sei es eine UNIX-Workstation, eine PC mit Windows oder ein DEC Alpha-Rechner. Für den reinen Konsumenten wird sich jedoch häufig ein System anbieten, daß auf einer Set-Top-Box (einem speziellen Microrechner für den VoD-Anwendungszweck) basiert, die zwischen den, wie auch immer gearteten, Netzanschluß und den Fernseher / Monitor geschaltet wird.

Es muß schon hier erwähnt werden, daß VoD noch auf Jahre hinaus in vielen Bereichen eher auf Near-VoD hinauslaufen wird, als daß immer alle Informationen sofort verfügbar sind. Man wird also auf Grund technischer Gegebenheit damit leben müssen, bestimmte selten referenzierte Informationen beispielsweise erst 30 bis 120 Minuten nach Ihrer Anforderung vorliegen zu haben.

Auch gibt es bereits existierende VoD-Einrichtungen in verschiedenen Staaten, in denen Unternehmen die Einsatzmöglichkeiten und auch die technische Durchführbarkeit von VoD untersuchen. Eine der ersten Versuchsanlagen war vermutlich der Test von Time Warner und Silicon Graphics in Orlando, Florida in der USA. Auf dieses Projekt und ein Projekt der British Telecom wird später noch etwas genauer eingegangen werden.

Die Deutsche Telekom hat auf dem Gebiet VoD bisher nur Fehlschläge verbuchen können. Ein erstes Projekt in Berlin wurde noch in der Planungsphase wegen verschiedener Probleme abgebrochen. Ein zweites Projekt in München wurde Ende Oktober 1996 ebenfalls als Fehlschlag wegen technischer Probleme eingestellt. Dabei waren zuletzt nur noch 3 Teilnehmer an dem Projekt beteiligt.

Mit der Entwicklung von VoD einher geht auch der Bedarf, das Auffinden von gespeicherten Filmdaten mit möglichst umfangreichen Recherche-Möglichkeiten zu unterstützen. Mit diesem Aspekt befaßt sich der zweite Teil dieser Arbeit.

Abschließend ist zu erwähnen, daß sich die hier genannten Fakten für VoD auch auf fast alle anderen Informationen übertragen lassen, die sich durch kontinuierlich Datenströme auszeichnen, d.h. z.B. auf den Bereich von „Audio-on-Demand“. Solche „On-Demand“-Diensten sind jedoch im Vergleich zu VoD eher unproblematisch, da die auftretenden Datenmenge erheblich geringer sind als bei VoD.

1.1 Einsatzgebiete von VoD¹

VoD kann in vielen Bereichen unseres täglichen Lebens vorteilhaft eingesetzt werden.

Diese Anwendungen drehen sich, wie nicht anders zu erwarten, um den Bereich von Video-Anwendungen. Zentral sind hier die naheliegenden Einsatzgebiete wie Video-Verleih zu sehen. Dabei ist mit dem Verleih von Videos primär auf den Endkunden, d.h. den Konsumenten, zu zielen (Spielfilme, Dokumentarfilme und Musikvideos). Wie angenehm wäre es zum Beispiel an einem regnerischen Abend nicht mehr zur Videothek gehen zu müssen um sich einen Film auszuleihen, sondern statt dessen an einer Fernbedienung nur ein paar Knöpfe zu drücken, damit einen Film auszuwählen und diesen sofort ansehen zu können.

¹ Siehe hierzu speziell [RAJ95], [FED94]

Verallgemeinert man diese Sicht des Video-Verleihs, kommt man sehr bald auf die Anwendung der Videodatenbank als dem umfassenderen Dienst. Hiermit könnte sowohl der Bereich des Endverbraucher-Dienstes als auch der Bereich der Speicherung von ganzen Filmarchiven von Unternehmen, Fernsehanstalten, Produktionsfirmen usw. zusammengefaßt werden. Benötigt nun jemand bestimmte Filmaufnahmen, so kann er sie dort abrufen und dann, unter Umständen gegen Entrichtung eines entsprechenden Obolus, verwenden. Hierdurch wäre es möglich auf die kostspielige Neuproduktion bereits existierender Szenen, von deren Existenz in irgendeinem Archiv man heute nicht einmal etwas weiß, zu verzichten.

Ein weiterer Aspekt wäre der Einsatz von VoD in Bildungseinrichtungen. Man könnte zum Beispiel Unterrichtsstunden bzw. Vorlesungen als Video für die Studenten zum Abrufen bereitstellen. Dies würde eine Nachbereitung, ein Nachholen einer verpaßten Vorlesung oder einfach das nochmalige Ansehen schwieriger Passagen erlauben. Würde das ganze noch mit zusätzliche Elementen wie Diagrammen, Animationen und zusätzlichen Filmbeiträgen zum Thema ergänzt, ergebe sich die Möglichkeit für eine ganz neue, höhere Qualität von Lehre und Lernen.

Faßt man die Einsatzgebiete von VoD zusammen, so ergeben sich folgende Punkte:

- zur Verfügung stellen von Videofilmen aller Art auf Abruf, der VoD-Kernbereich
- Ausbildung und Fernlehrgänge
- Nachrichten und Wettervorhersagen
- Spiele, Musik und Unterhaltung
- Home-Shopping (Einkauf von zu Hause aus) und andere Konsumentendienste
- Bankgeschäfte

1.2 Anforderungen an die Qualität des VoD-Dienstes²

Als Idealanforderungen an einen VoD-Dienst, der auch vom Endverbraucher genutzt werden kann, sind folgende Punkte zu sehen:

² Siehe hierzu speziell [HUO96]

a) Komfort und Qualität

- **Navigations- & Recherche-Dienste**, die in eine leicht zu bedienende Benutzeroberfläche eingebaut sind
- **Zeitliche Unabhängigkeit**, d.h. der Benutzer sollte nach Möglichkeit jeden Dienst jederzeit ohne Wartezeiten nutzen können
- **Schnelle Antwortzeiten** bei Interaktion durch den Benutzer (z.B. Aufruf einer bestimmten Informationsseite, anhalten der Wiedergabe oder „spulen“ eines Videofilms)
- die **Qualität der übertragenen Filme und Bilder** muß mindestens die heutigen Qualitätsstandards halten (also z.B. PAL-TV oder PAL-VHS). So ist es für die meisten Anwendungen nicht akzeptabel, daß durch die Wahl einer zu hohen Kompressionsrate (bei MPEG-2 immerhin bis 100:1) Artefaktbildungen in den wiedergegebenen Videos beobachtet werden können. Außerdem muß gewährleistet sein, daß zukünftige Übertragungsnormen (beispielsweise ATV oder HDTV) übertragen werden können.
- der **Preis** muß im Bereich heutiger alternativer Produkte liegen; bei Endverbraucher-Diensten also auf dem Kostenniveau von Pay-per-view-Fernsehen oder der Videothek, bei kommerziellen Nutzern auf dem Preisniveau vergleichbarer heutiger Dienste.

b) Privatsphäre und Sicherheit

- **Zugangssicherung zu Diensten**; es muß z.B. für Eltern die Möglichkeit geben den Zugriff von Kindern auf bestimmte Filme und Dienstleistungen zu kontrollieren bzw. zu sperren.
- **Zugangssicherung zum Netzwerk**; der Dienstanbieter muß garantieren können, daß niemand unerlaubt Zugriff auf die Dienste nehmen kann indem er sich einer fremden Identifikation bedient, oder sich in eine fremde Zugangsleitung einklinkt.
- **Sicherung von Benutzerdaten**; Informationen bezüglich der Benutzer eines Dienstes müssen vor fremden Zugang geschützt werden.

1.3 Die „Natur“ von Videodaten³

Videodaten entsprechen einem Datenstrom. Dieser enthält zum einen die Bildinformationen und zum anderen meist auch Toninformationen.

Bei heutigen Qualitätsmaßstäben in Ton und Bild bedeuten dies für Daten unter MPEG-2-Kodierung von Bild und Ton - die Toninformationen werden im Wechsel mit den Bildinformationen übertragen (interlaced) - eine Stromgröße von 500 - 1000 Kilobyte pro Sekunde. MPEG-2 erreicht bei Erhaltung einer guten Bildqualität eine Kompression von ca. 26:1 bei den Bildinformationen und 7:1 bei den Toninformationen.

Schon diese Stromgrößen bedeuten bei heutigen Netzwerken und Rechnerarchitekturen eine nicht unerhebliche Datenmenge, vor allem wenn man an sehr viele gleichzeitige Ströme denkt.

Für die Zukunft ist allerdings mit erheblich größeren Datenmengen zu rechnen, wobei natürlich auch eine Verbesserung der Bild- und Tonkompression nicht auszuschließen ist. Jedoch sind der Kompression von Daten Grenzen gesteckt, so daß nicht klar ist, wie weit die Kompressionsrate gesteigert werden kann.

Als Beispiel für in Zukunft zu erwartende Datenraten sei hier folgendes Beispiel angeführt. An der Qualität des neuen amerikanischen Standards für digitales Fernsehen ATSC ATV (Advanced Television Systems Committee Advanced Television), jetzt HDTV genannt, wird sich VoD in Konsumenten-Bereich messen lassen müssen, wenn ATV erst einmal auf breiter Front eingeführt worden ist bzw. wenn neue Medien wie digitale Videosysteme oder die DVD (auch wenn dies vielleicht noch 10 Jahre dauert) am Markt etabliert sind.

ATV sieht als eine mittlere Bildauflösung im 16:9-Format 1280x720-Bildpunkte vor. Als Ton-System wurde das fortschrittliche digitale Dolby-Surround-Ton-Kompressionsverfahren AC-3 gewählt. Die Datenrate, die daraus resultiert berechnet sich dann wie folgt:

³ Siehe hierzu speziell [RAK95], [RAJ95], [ATS95], [ATS95a]

- Bild:
 - 1280 x 720 Punkte
 - 3 Byte pro Punkt
 - 25 Bilder pro Sekunde
 - $1280 \times 720 \times 3 \times 25 = 65,91 \text{ Mbyte / s}$
 - selbst bei einer Kompression von 50:1 ergibt dies immer noch 1,318 MByte / s
- Ton:
 - 6 Kanäle
 - 48 kHz Sampling-Frequenz
 - 18 Bit Auflösung
 - $6 \times 48.000 \times 18 = 632,81 \text{ kByte / s}$
 - AC-3 arbeitet mit Datenströmen von 39 - 78,125 kByte / s - gute Tonqualität ergibt sich schon bei ca. 50 kByte / s

Daraus resultiert eine Gesamtdatenrate von hier 1,368 MByte / s.

Ein solche Rate ist für heutige Systeme nur dann zu bewältigen, wenn nicht zahlreiche Benutzer gleichzeitig darauf zugreifen.

2 Video-Datenbanken & Video on demand

Im folgenden wird auf die Realisierung von VoD-Systemen bzw. Systemkomponenten eingegangen. Dies geschieht an Hand von ausgewählten Beispielen verschiedener existierender oder geplanter Systeme.

2.1 Eine System-Architektur⁴

Eine mögliche und sinnvoll erscheinende Architektur eines VoD-Systems wird hier anhand des „Verteilten VoD-Systems“ der Universität von Kalifornien in Berkeley dargestellt (Berkeley Distributed Hierarchical VoD Server System, kurz: BDVS).

In Berkeley wurde das Hauptaugenmerk auf den verteilten hierarchischen Speicher-Manager gelegt. Als reine Video-Server werden kommerzielle Video-Server (z.B. von DEC, HP, IBM, Microsoft, SGI, nCube ...) eingesetzt, anstatt einen eigenen Video-Server zu entwickeln. Man ist in Berkeley zu der Auffassung gelangt, daß ein reiner Video-Dateiserver (VFS - Video-Fileserver) keine noch besonders zu erforschenden Herausforderungen birgt.

Tatsächlich beschäftigt sich die Forschung im Bereich von VFSn hauptsächlich nur noch mit der Art der Speicherung einer Videodatei auf mehreren Platten und ähnlichem, ansonsten scheinen hier keinen größeren Probleme mehr zu bestehen. Und selbst die Unterschiede in diesem Bereich erscheinen nur sehr geringfügig. Prinzipiell muß ein VFS in der Lage sein, möglichst viele gleichzeitige und kontinuierliche Datenströme an verschiedene Empfänger zu übermitteln. Der Datenstrom darf nicht unterbrochen werden und das System muß möglichst zuverlässig sein. Ob man die Daten in speziellen Datenbanken oder in normalen Dateisystemen speichert ist für die Anwendung eher nebensächlich und es gibt wohl in dieser Richtung auch verschiedene Ansätze. Schließlich scheint heute das Hauptproblem eher darin zu liegen, Rechnerarchitekturen zu entwickeln, die einen genügend hohen Datendurchsatz besitzen, um die gigantischen Datenmengen von VoD-Anwendungen zu bewältigen.

⁴ Siehe hierzu speziell [FED94], [ROW95], [ROW95a]

Ein Beispiel für einen VoD-Server ist z.B. die Erweiterung der Oracle Datenbank Version 7, die „Oracle Video Option“. Diese Software ist eng an die „normale“ Oracle-Datenbank gekoppelt, erfordert jedoch für die Videowiedergabe einen eigenen Server, der ausschließlich für die Videowiedergabe zuständig ist. Mögliche Rechnerarchitekturen für die „Video Option“ sind heute, am unteren Ende, Intel-PCs mit Pentium-Prozessoren bis hin zu Hewlett Packard HP9000 Systemen.

Das BDVS ist von seiner Konzeption her als ein System ausgelegt, das sowohl lokalen (LAN) als auch wide-area (WAN) Zugriff auf das Video-Material erlaubt. Dabei wird die Qualität der Übertragung abhängig von der erreichten Übertragungsgeschwindigkeit skaliert. Das System ist konzipiert um verschiedenste Arten von Multimedia-Dokumenten zu verwalten und es existiert ein Konzept zur Speicherung von Informationen bezüglich der Videos, d.h. bibliographische, inhaltsbezogene und strukturelle Informationen.

2.1.1 Die verteilte hierarchische Architektur

Das BDVS basiert auf zwei Arten von Servern. Diese dienen dem VoD Storage Manager zur Organisation von Zugriffen auf angeforderte Videodaten. Die Archivserver (AS) sind mit tertiären Speichersystemen ausgestattet, d.h. im Normalfall mit automatischen Bandwechsellanlagen (Jukeboxes) und dienen als zentrale Datenspeicher (Repositories). Außerdem verwalten die AS die Metadaten-Indizes, die für die Recherche nach Videos dienen. Die AS werden an wenigen zentralen Punkten aufgestellt.

Haupteigenschaften von Video-Servern scheinen zum einen die massiv parallele Architektur (bis mehrere zehntausend Prozessoren), die Verwaltung von Terabyte an Festplattenkapazität und der Einsatz von hunderten Gigabyte an Hauptspeicher zu sein. Als Trend ist dabei zu erkennen, daß man in diesen Systemem einen bzw. wenige Nutzer fest einem Prozessor zuweist, um den Verwaltungsaufwand für die Prozessoren niedrig zu halten.

Die VFS stellen den „verteilten Teil“ des BDVS dar. Für die AS kann eine „beliebige“ Anzahl von VFS zur Verfügung stehen, die nahe dem Kunden (Client) aufgestellt sind. Sie dienen als „Daten-Cache“, indem sie Videodaten von ASn auf Sekundär-Speichern, also insbesondere Festplatten, speichern. Durch eine Erhöhung bzw. Absenkung der Anzahl der VFS ist das System skalierbar.

Im BDVS nicht vorgesehen, aber sicherlich auch denkbar ist die Verwendung von großen RAM-Speicherbänken (sog. Silicon-Discs) für besonders stark frequentierte Daten, also z.B. besonders beliebte Filme.

Abbildung 1 zeigt die Architektur des BDVS.

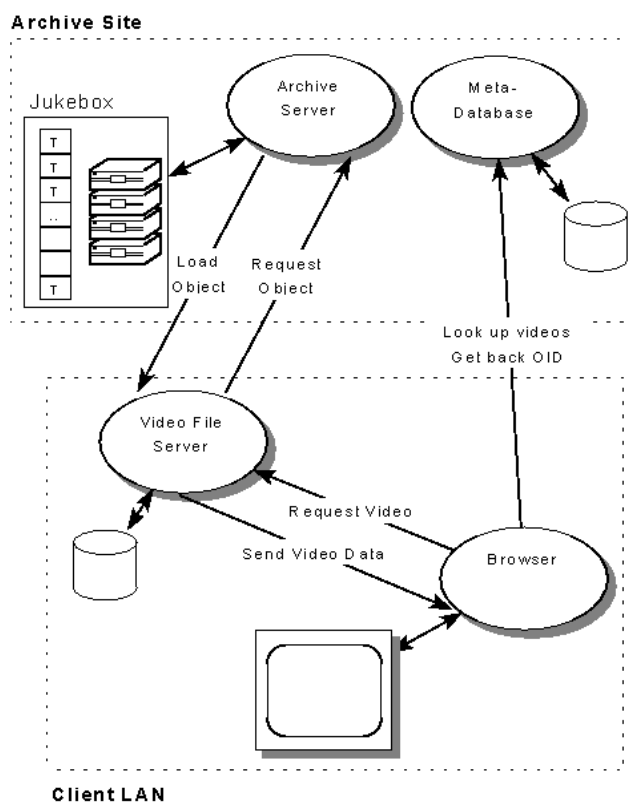


Abbildung 1 Die Architektur des BDVS

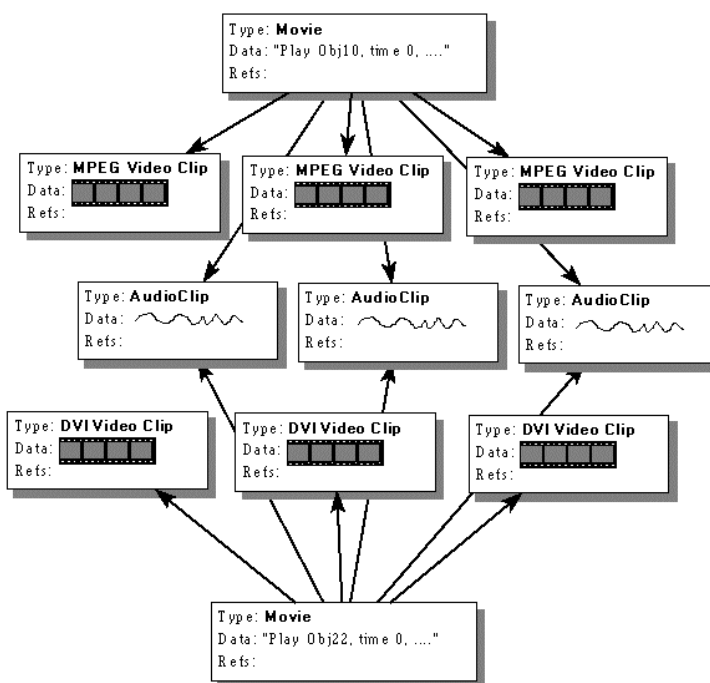
2.1.2 Das „Compound Media Object Model“ (CMO-Model)

Dem BDVS liegt zur Verwaltung der Medien-Objekte (nicht nur Video-Daten sondern z.B. auch Multimedia-Präsentationen, die hier jedoch nicht näher betrachtet werden sollen) das CMO-Modell zugrunde. Es dient dazu, eine Vielzahl von Multimedia-Dokument-Typen in einer Vielzahl von Repräsentationen und Codierungen zur Verfügung zu stellen und anspruchsvolle Anfragen bezüglich dieser Dokumente zu unterstützen. Es ist möglich, Daten in verschiedene Formaten und Codierungen

und in verschiedenen Bildgrößen, Bildraten, Farbtiefen und Tonqualitäten zu speichern. Außerdem unterstützt das System das Hinzufügen oder Ändern von Datenströmen um z.B. Untertitel hinzuzufügen oder aber den Ton in verschiedenen Sprachen wiederzugeben. Das CMO-Modell unterstützt außerdem das Extrahieren von kleinen Ausschnitten zur Wiedergabe oder deren Verwendung in anderen Dokumenten.

Damit bildet das CMO-Modell eine Abstraktionsschicht innerhalb des BDVS, durch die der VoD-Storage-Manager eine einheitliche Schnittstelle zu den gespeicherten Daten hat.

Jedes CMO besteht aus einem einzigartigen Objekt-Bezeichner (OID), einem Typnamen, einem oder mehreren Bitströmen (Dateien) und, optional, einer Liste von externen Referenzen auf andere CMOs. Eine CMO stellt sich nach außen hin als ein generischer Container dar, von dessen inneren Aufbau der Storage Manager nichts weiß, bis auf eine von außen zugängliche Liste von externen Objektreferenzen.



**Abbildung 2 Beispiel für die Verwendung von CMOs:
Zwei Filme-Objekte verwenden die selben Audio-
Objekte**

Beispielsweise denke man an den Ton- und den Videostrom eines Filmes. Diese seien in mehrere CMOs vom Typ „Tonabschnitt“ und „Videoabschnitt“ aufgeteilt. Ein anderer CMO vom Typ „Film“ könnte nun diese Abschnitte referenzieren und die Anordnung dieser Abschnitte beim Abspielen beschreiben. Der Storage-Manager

selbst braucht die Struktur des CMO nicht zu kennen. Es reicht, wenn er weiß, welche CMOs zur Wiedergabe notwendig sind, so daß er deren Speicherung und Verteilung veranlassen und unterstützen kann.

Die Aufteilung von CMOs in andere CMOs hat verschiedene Vorteile, darunter z.B. den, daß es nicht nötig ist, einen in mehreren Filmen verwendeten Ausschnitt mehrfach zu speichern oder aber den, daß man die Sub-CMOs über mehrere verschiedene VFS verteilen kann. Dadurch ist es möglich mehrere VFS in die Wiedergabe eines Films zu involvieren und so die Last des einzelnen auf mehrere Server zu verteilen. Auch von Vorteil ist, daß für ein Suchergebnis nur ein kleiner Ausschnitt zurückgeliefert werden kann, anstatt des kompletten Films. Abbildung 2 zeigt ein Beispiel für die parallele Verwendung von CMOs in anderen CMOs.

2.1.3 Das „Distributed Hierarchical Storage Management“

Dieser Abschnitt geht genauer auf das verteilte hierarchische Speichermanagement ein.

Auf der untersten Ebene stellt sich VoD als ein Problem dar, daß an verteilte Dateisysteme erinnert. Dabei gibt es zwei Aufgaben:

1. Lokalisierere das gesuchte Material und stelle es zur Wiedergabe beim lokalen Rechner bereit. Dabei ist für die Anlieferung von Daten beim Client eine strenges Echtzeit-Scheduling und eine entsprechende Bandbreite notwendig.
2. Zur kosteneffizienten Speicherung ist die Verwendung von tertiären Speichern notwendig. Diese tertiären Speicher bieten jedoch für interaktiven Zugriff auf die dort gespeicherten Daten keine ausreichenden Zugriffszeiten. Außerdem dürften diese Speichersysteme, wie z.B. automatische Bandwechsler mit einer beschränkten Anzahl von Leseköpfen, einen potentiellen Engpaß beim Zugriff auf Daten darstellen.

Auf Dauer wird es notwendig sein, auch das Scheduling von Netzwerken sowohl im Nahbereich (LANs) als auch im Fernbereich (WANs) auf Echtzeit-Erfordernisse anzupassen.

Eine, in Hinsicht auf die Übertragungsbandbreite, interessante Netzwerktechnologie wird in den kommenden Jahren sicherlich das ATM-Netzwerk sein. ATM bietet die Bandbreite an, die eine Übertragung von vielen Video-Strömen zumindest realisierbar erscheinen läßt. Außerdem ist eine obere Grenze für die Bandbreite von ATM noch nicht gänzlich abzusehen.

Schließlich ist noch zu berücksichtigen, daß selbst bei noch schnelleren Netzwerken, unter der Voraussetzung eines schlechten Systemdesigns, die Massenspeichergeschwindigkeit, z.B. von Festplatten, einen Flaschenhals darstellen kann. Hier helfen z.B. Techniken wie das „Striping“, d.h. das Verteilen von Daten über mehrere Platten, oder sogar Server, die die Bandbreitenproblematik an dieser Stelle eher wieder in den Hintergrund treten lassen.

All diese Probleme und Aspekte sprechen dafür, daß die Videodaten zum Client nur innerhalb eines LAN übertragen werden, was auch für die Skalierbarkeit des Systems gut ist, während die Gesamtheit der Daten an wenigen lokalen Stellen mit tertiären Speichern, für große Datenmengen, aufbewahrt werden. Dies alles führte in Berkeley dazu, ein verteiltes hierarchisches Speichermanagement-System zu entwickeln.

Grundidee ist, wie schon erwähnt, eine Speicherhierarchie durch Caching der Videodaten auf sekundären Speichern zu erzeugen und durch den Aspekt der Verteilung zu erweitern, indem man mehrere Caches parallel zuläßt. Erreicht man durch gute Cache-Strategien, daß ein Großteil der Zugriffe auf Daten durch die Caches behandelt werden kann, so folgt hieraus eine gute Performance und eine verhältnismäßig geringe Belastung des WAN.

Dieses Konzept berücksichtigt insbesondere, daß die Zugriffe von Benutzern auf Videodaten eine hohe Lokalität aufweist. Denn zu einem Zeitpunkt werden meist nur wenige Programme besonders populär sein.

Außerdem ermöglicht es das BDVS, dem Storage-Manager „Hinweise“ auf zu erwartende Zugriffe zu geben, so daß dieser diese Hinweise verwenden kann, um die Daten auch direkt und schnell verfügbar zu halten.

2.1.4 Zugriff auf die Daten

Wie läuft nun der Prozeß von der Suche nach einem Video bis zur Wiedergabe ab? Mittels eines „Video-Datenbank-Browsers“ (VDB) greift der Benutzer auf die Metadaten-Datenbank der verfügbaren AS zu. Dort kann mittels Suchkriterien nach dem gesuchten Video recherchiert werden. Ein oder mehrere VFS stehen dem Benutzer für die Echtzeitwiedergabe in seinem LAN zur Verfügung. Wählt der Nutzer nun einen Film aus, so überprüft der VDB zuerst, anhand der OID, ob das Video bereits bei einem der lokalen VFS gepuffert ist. Ist dies der Fall, kann sofort, ohne Zugriff auf das Archiv, mit der Wiedergabe begonnen werden. Ansonsten kann der Client den Film vom Archiv anfordern. Er muß dann von dort geladen werden. Weiß der Benutzer schon vorher, daß er auf den Film zugreifen will, dann kann er dies dem Storage-Manager frühzeitig mitteilen, so daß dieser das Video zur gewünschten Zeit zur Verfügung stellen kann.

Anhand der OID ist es möglich, den Server zu ermitteln, bei dem das CMO zuerst veröffentlicht wurde. Dieser „Home-Server“ besitzt die garantierte Eigenschaft, daß er entweder das gesuchte CMO gespeichert hat, oder aber weiß, wo das CMO gespeichert ist.

Ist das CMO lokalisiert, kann es auf den lokalen VFS geladen werden. Auf Grund der langen Zugriffszeiten der Jukeboxes auf Bänder (Suchzeiten, Bandtauschzeiten, Belegung aller Leseinheiten) ist für das Laden von einem Band ein adäquates Scheduling anzuwenden, so daß unnötige Bandwechsel und unnötiges Spulen innerhalb eines Bandes vermieden werden. Je mehr und je frühzeitiger der Storage-Manager also über die zu ladenden Filme bescheid weiß, umso besser kann er das Scheduling organisieren.

Das Protokoll zur Anforderung der Daten enthält neben der OID zusätzlich auch noch eine Prioritätsangabe und / oder eine „Deadline“, d.h. die Angabe eines spätest möglichen Zeitpunkts, zu dem das CMO vorliegen muß. So berechnet der AS anhand dieser Angaben regelmäßig eine aktuelle Priorität, anhand der statischen Priorität, der Deadline und der Zugriffskosten, die sich aus der Zeit ergeben, die benötigt wird um das CMO zu laden unter der Bedingung des aktuellen Zustands der Leseinheiten der Jukebox. So wird also immer die Anforderung mit der höchsten Priorität privilegiert behandelt.

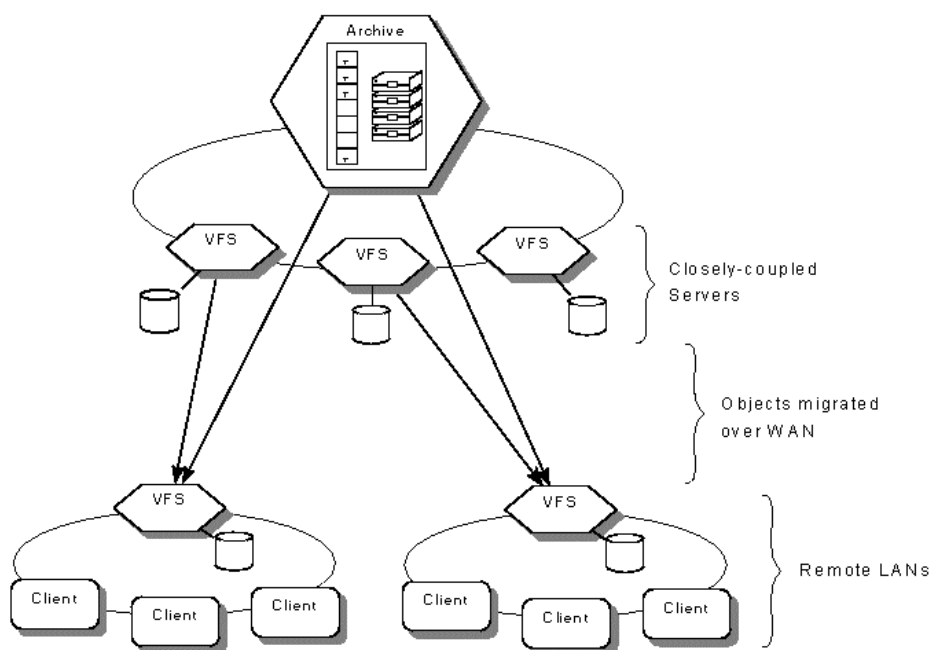


Abbildung 3 - Das Prinzip der Closely-coupled Server

Die Hierarchie des Systems kann sogar noch verstärkt werden, indem zwischen die AS und die VFS noch zusätzliche spezielle VFS (sog. Closely-coupled Servers, siehe Abbildung 3) eingeklinkt werden, die, sofern sie das angeforderte Video gespeichert haben, die Übertragung der Daten an den VFS des Clients übernehmen. Damit wird die Last der AS weiter gesenkt.

Trotz all dieser Techniken ist es jedoch immer noch möglich, daß eine Anforderung erst nach Stunden erfüllt werden kann, je nach Zustand des Systems.

Sobald ein AS den ungefähren Zeitpunkt kennt, zu dem die angeforderten Daten beim VFS des Clients vorliegen, kann er ihm diesen Zeitpunkt mitteilen, so daß der Anfordernde sich darauf einstellen kann.

2.1.5 Cache-Management

Bei aller Optimierung auf der Seite des AS, ist die Systemperformance in erheblichem Maße von der Fähigkeit abhängig, Benutzeranfragen schnell zu erfüllen. Dies erfordert eine möglichst effiziente Cache-Strategie, über die die Ersetzung von Daten im Cache gesteuert wird.

Ein Aspekt dabei ist die Fähigkeit des Cache-Managements, Anfragen in der Zukunft vorausszusehen. Um Vorteile daraus zu ziehen, hat man in Berkeley die bereits ange-deuteten „Cache-Anweisungen“ eingeführt.

Ersetzungsstrategien versuchen anhand von Zugriffen auf Objekte in der Vergangenheit zu erraten, auf welche Objekte in der Zukunft zugegriffen werden könnte. Diese Verfahren sind leider relativ unsicher. Daher ist es den Benutzern zu ermöglichen, ihre zukünftigen Wünsche dem System mitzuteilen. Das System kann dann unter Zuhilfenahme dieser Informationen besser entscheiden, welche Daten ersetzt werden sollen.

Für Video-Daten wären z.B. folgende Cache-Anweisungen denkbar:

- Ich will heute Abend um 20.30 Uhr „Star Trek II - Der Zorn des Kahn“ sehen.
- Halte in der Zeit vom 20.01.1997 bis 20.02.1997 alle Filme zum Thema „Video-on-Demand“ bereit.
- Behalte alle Filme im Cache, die in der letzten Woche von mindestens 10 Personen gesehen wurden.
- Speichere das Video zum Vortrag „Video-on-Demand“ vom 31.01.1997 bis zum Tag des jüngsten Gerichts.

Anhand der Datenbank mit den Zugriffsdaten und den Cache-Anweisungen kann nun jeder VFS für die bei ihm gecachten Daten Prioritäten bestimmen, die es ihm ermög-

lich zu entscheiden, welche Daten bei der nächsten neuen Anforderung gelöscht werden sollen und welche Daten zu Zeiten niedriger Systemlast vorsichtshalber schon einmal geladen werden.

2.1.6 Parallelisierung von Servern

Als letztes Konzept zur Steigerung der Leistung des BDVS wurde die Parallelisierung von mehreren VFS, von denen ein Benutzer Daten laden kann, eingeführt. Das Prinzip dabei ist, daß man einen Art von Software-RAID einführt. Man verteilt dazu die Daten eines Filmes auf mehrere Server. Dadurch ist gewährleistet, daß auch dann, wenn sehr viele Leute den gleichen Film sehen wollen, der entsprechende Server nicht überlastet wird. Der Storage-Manager des BDVS unterstützt ein CMO-orientiertes Striping über mehrere VFS. Damit dies möglich ist, muß man nur einen Film in mehrere CMOs mit einer Länge von wenigen Megabyte unterteilen. Die einzelnen Teile können dann vom Storage-Manager auf das VFS-Array verteilt werden. Während des Abspielens eines Films werden vom System jeweils die Startzeiten für die Übertragung der Daten von den verschiedenen VFSn zum Client berechnet und an die VFS übermittelt. Ergeben sich im vorberechneten Zeitplan Änderungen, z.B. durch Benutzerinteraktion, so muß der Zeitplan korrigiert und den entsprechenden Servern, die an der Wiedergabe beteiligt sind, mitgeteilt werden.

2.2 Sicherung der Übertragung

2.2.1 Kontinuität⁵

Dieser Abschnitt soll eine der Ideen vermitteln, die sicherstellen kann, daß eine Übertragung eines Video-Stroms zwischen einem Video-Server, im Sinne des BDVS wäre dies ein VFS, und einem Client nicht unterbrochen wird. Daher liegt die Idee zu Grunde, für den Fall, daß die zur Verfügung stehende Bandbreite im Netzwerk nicht mehr ausreicht, die Übertragungsqualität von Ton und Bild zu skalieren, d.h. dyna-

⁵ Siehe hierzu speziell [HOL96]

misch an die zur Verfügung stehende Bandbreite anzupassen. Dabei ist der Tonübertragung (so vorhanden) immer Vorrang zu gewähren, denn der Zuschauer reagiert empfindlicher auf Unterbrechungen des Tons, als auf Unterbrechungen des Bildes.

Bei der Übertragung von Datenströmen gilt es zwei Aspekte besonders zu berücksichtigen: der erste betrifft die Intra-Media-Synchronisation, d.h. die Aufrechterhaltung der Synchronisation der Daten innerhalb des Stroms, z.B. die Aufrechterhaltung der Übertragung mit der korrekten Bildrate. Der zweite Aspekt betrifft die Inter-Media-Synchronisation, d.h. die Synchronisation mehrerer Ströme bei der Übertragung, z.B. die Synchronisation des Video-Stroms mit dem Audio-Strom, so daß der wiedergegebene Ton zum wiedergegebenen Bild paßt.

Eine bei der GMD entwickelte Methode ist die „Buffer-Triggered Smooth Adaptation Technique for Time-Dependent Media“, die die Aufrechterhaltung einer brauchbaren Übertragung von Datenströmen über ein Netzwerk zum Ziel hat. Dabei wird eine Anpassungstechnik zusammen mit einer Zugangskontrolle verwendet.

Anpassungstechnik (adaptation technique) , auch Skalierungsmethode oder Fluß-Management genannt, bedeutet, daß nach Erkennung eines Übertragungseinganges die Daten nur noch mit einer niedrigeren Wiedergabequalität übertragen werden, so daß durch die reduzierte Datenmenge eine kontinuierliche Wiedergabe beibehalten werden kann. Es werden zwei Arten von Anpassungstechniken unterschieden: **zeitliche Anpassung** bedeutet dabei, daß z.B. die Bildrate bei der Wiedergabe gesenkt wird, indem einfach Bilder weggelassen werden. Dadurch wird die ursprüngliche Wiedergabegeschwindigkeit beibehalten, allerdings führt diese Methode zu teilweise abrupten Sprüngen im Bild. Problematisch ist hierbei, daß das Weglassen von Bildern bei MPEG-Kodierung nicht trivial ist, da die einzelnen Bilder nicht unabhängig voneinander sind. Beim Ton ist es nicht möglich eine zeitliche Anpassung vorzunehmen, da dies zu Unverständlichkeit des übertragenen Tons führen kann. **Räumliche Anpassung** bedeutet, daß die Auflösung bzw. die Größe des Datenstroms geändert wird. Hierbei kann man wählen, ob der Videoservert bei Bandbreiten-Problemen zur Laufzeit die vorliegenden Daten in Daten geringerer Auflösung und Größe umrechnen soll, oder aber z.B. mehrere Versionen der Videoinformationen in verschie-

denen Auflösungen und mehrere Versionen der Audioinformationen mit verschiedenen Auflösungen, Sampling-Raten oder mit unterschiedlich vielen Kanälen gespeichert werden. Die erste beschriebene Variante wird offensichtlich bereits von MPEG-2 für die Bildinformationen unterstützt, weshalb sich deren Verwendung anbietet.

Zugangskontrolle bedeutet bei solchen Übertragungsverfahren normalerweise, daß Ressourcen wie Prozessoren, Magnetplatten und Netzwerk reserviert werden. Dabei werden im Normalfall eine Unter- und Obergrenze für die benötigte Bandbreite spezifiziert. Ist die Anforderung durch die momentan zur Verfügung stehende Bandbreite nicht zu erfüllen, so wird der Zugang verwehrt. Das GMD-Verfahren verwendet, soweit Ressourcen-Reservierung überhaupt möglich ist, nur eine untere Grenze. Nur diese Bandbreite wird fest reserviert. Dadurch sinkt im Vergleich zu herkömmlichen Strategien die Anzahl der abgelehnten Anfragen. Bei Abweichungen in der verfügbaren Bandbreite von der oberen Grenze, die auch der besten Übertragungsqualität entspricht, wird die Anpassungstechnik verwendet.

Das Konzept schlägt ein Netz von zeitlichen und räumlichen Anpassungstechniken vor. Das Verfahren ermöglicht erstens die Spezifikation von sich überlappenden Intervallen für eine Änderung der Wiedergabequalität und zweitens Anweisungen für die Auswahl der wegzulassenden Bilder. Diese Spezifikationen ermöglichen eine kontinuierliche Wiedergabe durch nur kleine Qualitätssprünge beim Wechseln der Wiedergabequalität. Dynamische Verbesserungen in der Übertragungsrate bzw. der Systemlast resultieren in einer verbesserten Wiedergabequalität, indem bereits gepufferte Objekte mit den gleichen Objekten einer höheren Wiedergabequalität ersetzt werden.

Der Ansatz der GMD beruht darauf, daß auf der Seite des Clients Daten gepuffert werden. Anhand des Füllgrades dieses Puffers wird erkannt, ob ein Übertragungsengpaß vorliegt, da dieser Puffer als ein Indiz für Änderungen in der Übertragungsperformance dienen kann. Sinkt die Durchsatzrate, so sinkt der „Füllstand“ des Puffers und der Client muß einen Anpassungsmechanismus initiieren um die zu übertragende Datenmenge zu reduzieren. Steigt dagegen der Systemdurchsatz, kann der Cli-

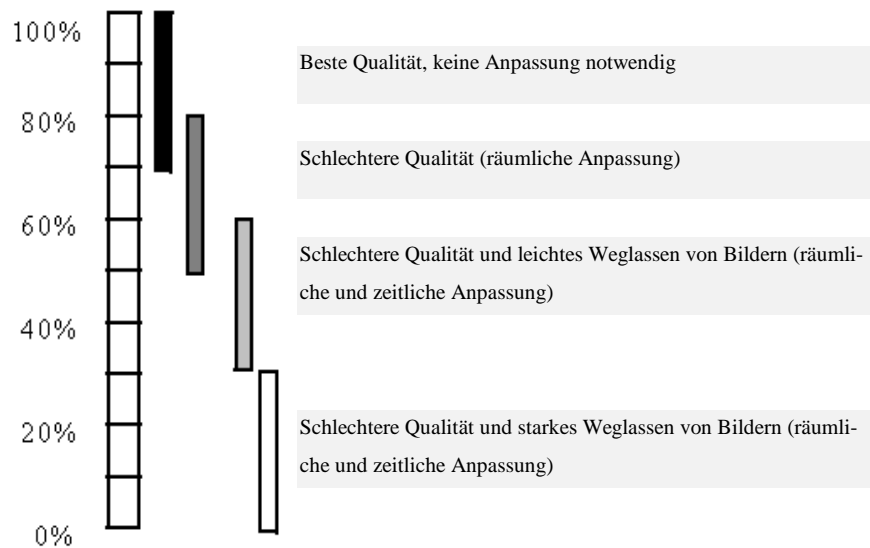


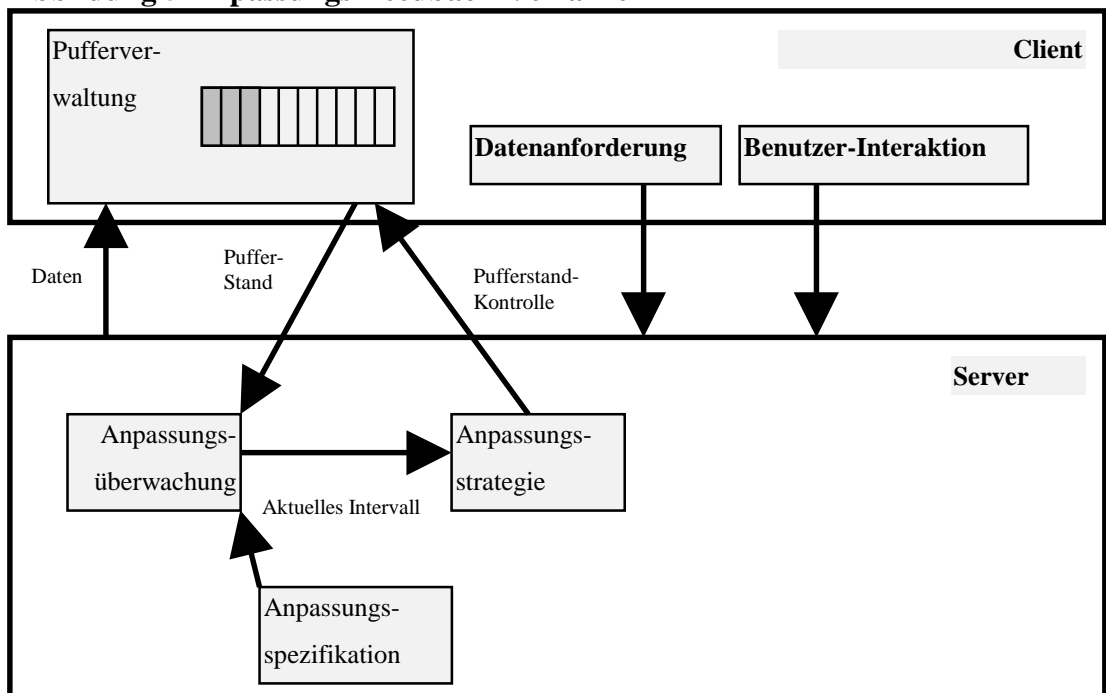
Abbildung 4 Puffer-Nutzung und Anpassungsintervalle

ent den Grad der Ressourcenbelegung steigern und zu einer besseren Wiedergabequalität zurückwechseln.

Während der Wiedergabe überprüft der Videosever kontinuierlich den Pufferfüllstand und bestimmt, anhand der Anpassungsspezifikationen, die anzuwendende Anpassungsstrategie.

Um tatsächlich eine relative Sanftheit bei den Wechseln der Anpassungsstrategie zu erreichen, ist es möglich und beabsichtigt, diese für bestimmte, sich überlappende

Abbildung 5 Anpassungs-Feedback-Verfahren



Intervalle zu definieren. Dadurch kann es nicht zu ständigen, hektischen Wechseln der Anpassungsstrategie kommen.

Die Möglichkeit der Angabe der bei zeitlicher Anpassung wegzulassenden Bilder dient zum einen dazu, bei normaler Wiedergabe eine zeitliche Anpassung vorzunehmen, sie dient jedoch auch dazu, z.B. bei schnellem Vorlauf („Vorspulen“) des Videos die wegzulassenden Bilder auswählen zu können.

Abbildung 5 zeigt das Regelungsverfahren für die Anpassungs-Steuerung.

2.2.2 Gleichmäßige Verteilung der Netzlast⁶

Ein weiteres Verfahren, das der Optimierung der Performance eines VoD-Systems dient und hier kurz vorgestellt werden soll, ist die „Bandwidth Smoothing Technique“ (Bandbreiten-Glättungs-Technik).

Grundlage dieses Verfahrens ist die Verwendung eines (möglichst großen) Puffers auf Seite des Clients. Dieser ermöglicht es dem Server, die zur Übertragung notwendige Bandbreite möglichst gering zu halten und dies um so besser, je größer der Puffer ist.

Die Idee dahinter ist die folgende: Der Server kennt die Daten, die er an den Client schicken muß. Da der Server weiß, wann er welche Daten an den Client schicken muß und da in einem Videostrom (speziell bei MPEG-Codierung) teilweise große Schwankungen in der benötigten Bandbreite auftreten, kann der Server den zu übertragenden Datenstrom glätten. Dabei kann er z.B. nach folgenden Kriterien verfahren:

1. Minimiere die Anzahl der Bandbreiten-Anstiege
2. Minimiere die maximalen Bandbreitenspitzen
3. Maximiere die minimal benötigte Bandbreiten

Ein Algorithmus, der dies erfüllt, ist der „Critical Bandwidth Allocation“ (CBA).

Wird zusätzlich, unter Kenntnis der Puffergröße des Clients, auch noch die Anzahl der Veränderung der benötigten Bandbreite minimiert, so spricht man von der sogenannten „Optimal Bandwidth Allocation“ (OBA). Beide Verfahren sind auf Grund der a priori Kenntnisse über die Video-Daten anwendbar.

Um aber sicherer zu sein, daß man bei einer begonnenen Übertragung nicht plötzlich doch noch in einen Bandbreiten-Engpaß läuft, kann man Ressourcen-Reservierungs-Schemata einsetzen, die voraussetzen, daß man sowohl die benötigte Bandbreite kennt, als auch die Dauer, für die diese Bandbreite benötigt wird. Eine solche sogenannte „vorverlegte Reservierung“ („advance reservation“) wird dann in 2 Phasen bearbeitet. Die erste Phase besteht aus einer Zugangskontrolle. Hier wird geprüft, ob das Reservierungs-Schema erfüllt werden kann. Ist dies nicht der Fall, wird die Reservierung abgelehnt. In der zweiten Phase, der Durchführungsphase, wird dann zu den entsprechenden Zeitpunkten, die entsprechende Bandbreitenreservierung durchgeführt. Auch hier kann es allerdings immer noch zu Problemen bei der Wiedergabe kommen, wenn sich der Zeitplan verschiebt, z.B. durch Benutzerinteraktion („Pause der Wiedergabe“). Um diesem zu begegnen, kann man nun noch eine maximale Verzögerungszeit einführen. Dies schränkt zwar zum einen den Benutzer ein, bietet aber dem System die Möglichkeit, für Verzögerungen entsprechende Ressourcen zurückzuhalten, so daß es zu keinen Engpässen kommen kann.

Abschließend sollte noch erwähnt werden, daß ein entsprechender Puffer zusätzlich auch noch die Möglichkeit bietet, in einem beschränkten, von der Größe des Puffers abhängigen Zeitfenster, Videorecorder-Funktionen (Zurückspulen, Vorspulen) ohne zusätzlich Zugriffe auf den Server zur Verfügung zu stellen. Dabei muß man dann von diesem Puffer einen bestimmten Teil dafür benutzen, daß man bereits wiedergegebene Daten vorrätig hält.

⁶ Siehe hierzu speziell [FEN96]

3 Indizierung von Videodaten

Wie in den vorangegangenen Kapiteln schon mehrfach erwähnt, ist die Indizierung von Video-Daten eine wichtige Voraussetzung, damit große Video-Datenbanken überhaupt eine sinnvolle Verwendung finden können, denn die größten Videoarchive sind sinnlos, wenn es keine effiziente Methode gibt, nach bestimmten Filmen, Aufnahmen oder Szenen zu suchen.

Problematisch ist die Erstellung der entsprechenden Indizes. Sie erfordert heute noch einen außerordentlich hohen personellen Aufwand. Daher scheitert eine solche Indizierung häufig an fehlenden finanziellen Mitteln.

Im folgenden soll auf zwei Ansätze eingegangen werden. Der erste Ansatz wurde bei der GMD entwickelt, um für das IWF, das Institut für den wissenschaftlichen Film, eine Erfassungsmöglichkeit für deren Filmmaterial zu schaffen. Das System heißt **AMPHORE** (für **A**udio-visual **M**edia **P**latform for the **H**ighlighting, **O**rganization and **R**etrieval of **E**ntities).

Danach soll noch auf das Indizierungsverfahren eingegangen werden, daß bei der Universität von Kalifornien in Berkeley entwickelt wird und für die Zusammenarbeit mit dem BDVS konzipiert wurde.

3.1 AMPHORE⁷

AMPHORE ist bereits in einem ersten Prototypen realisiert. Dieser Prototyp enthält die drei zentralen Komponenten des Gesamtsystems:

⁷ Siehe hierzu speziell: [SÜL96]

- eine Datenbank zur Verwaltung der Filmdokumentation. Dabei handelt es sich um ein kommerzielles, SGML⁸-fähiges Dokumentenverwaltungssystem, in dem die Dokumentation in Form von SGML-Dokumenten gespeichert wird.
- ein System zur dokumentarischen Erfassung von Filmen
- ein Recherchesystem

Zur Bearbeitung eines Films wird dieser in Teilsequenzen zerlegt. Dafür kann schon heute eine automatische Schnitterkennung für „harte“ Schnitte, d.h. Schnitte mit deutlichem Motivwechsel, eingesetzt werden. Diese Teilsequenzen sind immerhin schon eine erste Näherung für zu indizierende Teilsequenzen.

Ziel von AMPHORE ist es, Filmmaterial, Begleitmaterial und dokumentarische Informationen einheitlich, d.h. digital, zu speichern. Nur so ist eine effiziente Verwertung von Recherchen möglich. Dabei ist es auch möglich, Querbezüge zwischen den gespeicherten Daten zu erstellen, d.h. explizit anzugeben, oder implizite Querbezüge zu verwenden, also auf Grund von geeigneten Datenbankabfragen ähnliches Material zu recherchieren. Beispielsweise könnte man von einem bestimmten Textabschnitt im Begleitmaterial einen Verweis auf den entsprechenden Abschnitt im Film speichern und umgekehrt.

3.1.1 Das Daten-Schema

Um eine einigermaßen aussagekräftige und eindeutige Beschreibung der Sequenzen zu unterstützen, unterstützt das System den Benutzer mit zwei Funktionen: Zum ei-

⁸ SGML (Standard Generalized Markup Language) - eine abstrakte Dokumentenbeschreibungssprache. Das bekannte HTML-Format des WWW ist z.B SGML-konform. In SGML kodierte Informationen gestatten es, daß die Informationen aus Objekten relativ problemlos von mehreren Anwendungen verwendet werden können. Dabei ist das Format herstellerunabhängig und international standardisiert (ISO 8879) und dient als Grundlage für dokumentenbasierte Informationen in einigen offenen Systemumgebungen. Vorteile des Einsatzes von SGML liegen vorwiegend im Inventionsschutz, da die Speicherung und Verarbeitung von Informationen nicht von einem bestimmten Hersteller abhängig ist. Vielmehr kann man sich unter den Applikationen verschiedener Hersteller, die zur Verwaltung, Bearbeitung und Speicherung von Informationen dienen die für die eigenen Zwecke geeignetsten auswählen. Schließlich ist auch ein Wechsel zwischen Applikationen verhältnismäßig unproblematisch.

nen bietet es eine explizite Trennung von beschreibenden und interpretierenden Aussagen über die Sequenz. Die Wirksamkeit dieser Funktion hängt allerdings stark von den Bemühungen des Anwenders ab, diese Unterscheidung auch vorzunehmen. Zum anderen werden die Beschreibungen mittels Thesauri kontrolliert. D.h. es können nur Wörter verwendet werden, die das System kennt und zu denen das System unter Umständen Synonyme für die Recherche kennt.

Zur Beschreibung der Sequenzen werden ein oder mehrere SPO-Konstrukte, bekannt aus dem Englischunterricht, verwendet. Eine Szene wird mit Subjekt, Prädikat und Objekt beschrieben. Dabei ist es zwingend erforderlich das das Subjekt angegeben wird, das Prädikat und die, bis zu vier, Objekte sind optional. Werden für eine Sequenz mehrere SPO-Konstrukte angegeben, so spiegelt die nachträglich veränderbare Reihenfolge die Gewichte der entsprechenden Aussagen wieder.

Beispiele für solche SPO-Konstrukte:

- („Fuchs“, „fangen“, „Kaninchen“)
- („Elster“, „fliegen“)
- („Revolverheld“, „schießen“, „Sheriff“)
- („Superman“, „Lex Luthor“)

Zusätzlich zur Dokumentation der Sequenzen werden folgende Daten, sofern sie vorliegen, gespeichert:

- Bibliographische Angaben: Regisseur, Titel, Herstellungsort(e), kurze Inhaltsangabe
- Begleitinformationen: Drehbuch, Kommentare, wissenschaftliche Abhandlungen (insbesondere bei wissenschaftlichen Filmen)
- Dialoge und Sprechertexte
- Technische Informationen: Länge der einzelnen Informationen, deren Qualität (z.B. Farb- oder Schwarzweisaufnahme), drehtechnische Eigenschaften (Nacht-aufnahme, Großaufnahme, Zeitlupe / Zeitraffer usw.)
- das Filmmaterial selbst in digitaler Form
- das digitalisierte Tonmaterial (sofern es getrennt vom Bildmaterial vorliegt)

Für die Speicherung der Daten wurde eine, auf SGML basierende, Dokumenten-Typ-Definition entwickelt. Diese schreibt vor, daß zu jedem Film ein Dokument angelegt werden muß, das aus einem Pflicht- und vier optionalen Teilen besteht, nämlich

- Element **DOKUMENT**: enthält die bibliographischen Informationen. Dieser Teil ist verpflichtend.
- Element **FILMDAT**: hier wird eine Referenz auf den gespeicherten Film abgelegt. Dies ist notwendig, da in der SGML-Datenbank die Filmdaten nicht zu speichern sind.
- Element **AUDITEXT**: enthält die Dialogtexte und Kommentare des Films. Dieses Element ist HTML-abwärtskompatibel modelliert, so daß es mit WWW-Browsern wiedergegeben werden kann.
- Element **BEGLEIT**: enthält die Begleitinformationen oder Verweise darauf
- Element **SEQUENZ**: enthält beliebig viele Elemente der Dokumenttypen SEQ und FRAME, die Sequenzen bzw. Einzelbilder des Films beschreiben und zwar durch die erwähnten SPO-Konstrukte, durch optionalen, HTML-abwärtskompatiblen Freitext und durch produktionstechnische Angabe, wie z.B. Aufnahmewinkel, Beleuchtung und Zeitlupengeschwindigkeit.

3.1.2 Daten-Erfassung und -Recherche

Die Erfassung der einzelnen Daten erfolgt über ein komfortables Dokumentationswerkzeug, das die Eingabe der textuellen Informationen, die Auswahl bzw. Erstellung der Filmsequenzen und die Vergabe der SPO-Konstrukte unterstützt.

Recherchiert wird über ein Recherchesystem, das verschiedene Recherchearten unterstützt. So kann über eine Maske nach den SPO-Konstrukten recherchiert werden. Dabei wird die Benutzereingabe vor Beginn der Suche auf Thesauruskonformität überprüft, so daß nur „genormte“ Suchbegriffe verwendet werden können. Des Weiteren ist eine Volltextrecherche möglich, da die Freitext-Teile der Dokumente unter Berücksichtigung einer Stopwortliste indexiert werden. Weiterhin ist eine traditio-

nelle Suche über die bibliographischen Informationen möglich und schließlich ist eine Recherche über die technischen Informationen möglich.

AMPHORE ist bisher als Prototyp realisiert. Dort werden die allgemeinen Funktionen unterstützt, wenn auch die Filmdaten bisher nur als normale Dateien vorliegen können. Fraglich ist jedoch, ob die Wahl der SGML-Volltext-Datenbank zur Speicherung der Informationen eine weise Entscheidung war. Normalerweise würde ich davon ausgehen, daß eine traditionelle, relationale Datenbank mit einem gut designten Datenmodell eine höhere Anfrage-Performance erreichen müßte als ein Volltext-System, ganz egal, wie optimiert die Volltextrecherche ist.

3.2 Indizierung im BDVS⁹

Ein weiterer, diesmal relationaler, Ansatz, der für das BDVS vorgesehen ist, soll nun noch vorgestellt werden. Dabei ist es interessant die Unterschiede zum Konzept von AMPHORE zu beobachten. Es handelt sich allerdings um eine Pilot-Studie, die noch nicht in allen Details ausgearbeitet ist und an manchen Stellen noch nicht endgültig durchdacht zu sein scheint.

Bei der Klassifizierung von Metadaten, die bei einem VoD-System vorkommen, wurden in Berkeley folgende 3 Informations-Indizes ermittelt:

- Bibliographische Informationen
- Strukturelle Informationen.

Hierbei wird eine Hierarchie von strukturellen Einheiten eines Filmes erzeugt. Die Hierarchieebenen sind dabei (von oben nach unten): Film (Movie) • Segment (Segment) • Szene (Scene) • Einstellung (Shot). Dabei ist ein Element einer höheren Ebene immer aus einem oder mehreren Elementen der nächst niedrigeren Ebene zusammengesetzt. Einen Struktur-Index kann man nicht nur für Bild-, sondern auch für Toninformationen erstellen.

⁹ Siehe hierzu speziell: [ROW94]

- Inhaltliche Informationen.

Benutzer können hierdurch Videos anhand von Inhaltsangaben suchen. Beispiele für solche Inhalts-Indizes sind: 1) Mengen von Schlüsselbildern eines Videos, die den Inhalt des Videos repräsentieren, 2) Schlüsselwort-Indizes, die aus den Dialogen und Kommentaren eines Filmes erstellt werden, und 3) Objekt-Indizes, die Anfangs- und End-Bilder für jedes Erscheinen eines wichtigen Objektes oder einer wichtigen Person kennzeichnen.

3.2.1 Datenbank-Schema

Die folgende Tabelle zeigt die Haupt-Entitäten des Schemas:

Index-Typ	Entität	Enthält einen Eintrag für jede(s)
Dokument	DOCS	Dokument
Bibliographisch	VIDEO_BIB	Video-Dokument, das bibliographische Daten enthält
Strukturell	VSV_SHOT	Einstellung in einem Video
	VSV_SCENE	Szene in einem Video
	VSV_SEGMENT	Segment in einem Video
Objekt	PEOPLE	Person, die zu einem Dokument gehört
	OBJECT	Sache, die in einem Dokument vorkommt
	OBJ_INST	Sache, die in einem Video vorkommt
Schlüsselwort	KW_WORDS	Schlüsselwort in einem Dokument

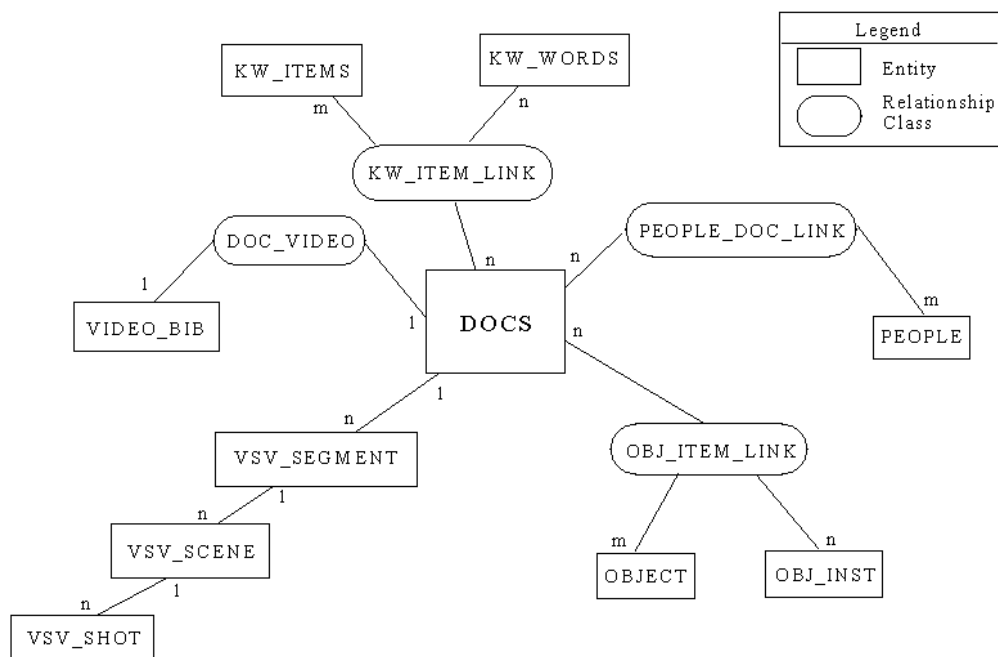


Abbildung 6 Entity-Relationship-Diagramm des Kernbereichs der Indexdatenbank

Abbildung 6 zeigt den Kernbereich des Datenbank-Schemas als ER-Diagramm.

Die einzelnen Entitäten sind, auch in ihrer Menge, noch nicht in allen Details ausgearbeitet.

Die zentrale Entität ist DOCS. Sie enthält für jedes Dokument einen Eintrag. Dort sind Daten wie ein Schlüssel, der Titel, der Name des Dokument-Anlegenden, ein Kommentar, das Datum des Anlegens des Datensatzes usw. gespeichert. Andere Entitäten referenzieren ein Dokument anhand des Schlüssels.

Der Bereich um den bibliographischen Daten-Index VIDEO_BIB soll u.a. folgende Daten speichern: Dokumententitel, Zusammenfassung des Dokuments, Datum der Eintragung des Datensatzes, Kommentare zu den bibliographischen Daten, Veröffentlichungs- oder Drehtag des Videos, Länge des Videos, Genre, Regisseur, Auszeichnungen, usw.

Die strukturellen Indizes dienen dazu, die Struktur des Videos zu speichern. Es gibt jedoch zwei Typen von Video-Strukturen. Der eine orientiert sich an den visuellen Eindrücken, der andere an der Handlung. Es wäre also durchaus denkbar zwei Strukturangaben für ein Video zu speichern.

VSV_SHOT speichert Daten einer Einstellung, d.h. einer ununterbrochenen Sequenz von Bildern, die von einer Kamera aufgenommen wurden. Zu speichernde Daten wären hier: Anfangs- und Endzeit (Format: h:m:s:frames), Beschreibung der Einstellung, Beschreibung der Aufnahmezeit, Schlüssel des zugehörigen DOCS, Übergang in die Einstellung (z.B. Überblendung, Verschieben, harter Schnitt), Übergang aus der Einstellung, Art der Einstellung (z.B. Nahaufnahme, Totale), usw.

VSV_SCENE befaßt sich mit der Speicherung einer Szene, d.h. einer Sequenz von Einstellungen, die sich mit einem oder mehreren bestimmten Objekten befaßt. Zu speichernde Daten wären hier: Anfangs- und Endzeit, Beschreibung der Szene, Beschreibung der Aufnahmezeit, Schlüssel des zugehörigen DOCS, „das interessante Objekt“ (z.B. ein Schauspieler oder eine Lokalität), Schauspieler die in der Szene auftreten, andere vorkommende Objekte, Ort, an dem die Szene spielt, usw.

VSV_SEGMENT befaßt sich mit Segmenten und enthält ähnliche Attribute wie VSV_SCENE.

Schlüsselwort-Indizes dienen zur Unterstützung von Schlüsselwort-Suchen. Damit man auch nach Teilen von Videos suchen kann, ist es notwendig, für jedes Schlüsselwort jedes Vorkommen zusammen mit einem Zeitstempel aufzuzeichnen. Da Schlüsselworte in verschiedenen Daten vorkommen können (Drehbuch, Titel, Zusammenfassung) ist diese Relation schwierig zu modellieren. Die Entität KW_WORDS speichert Informationen über das Schlüsselwort und enthält folgende Attribute: den Wortstamm des Schlüsselwortes, die Anzahl der Dokumente, die das Schlüsselwort enthalten und die Gesamtanzahl der Vorkommen dieses Schlüsselwortes.

KW_ITEMS speichert Informationen über den Ort an dem ein Schlüsselwort vorkommt, also z.B. Titel, Zusammenfassung, Drehbuch. Für jeden Ort, an dem das Schlüsselwort vorkommt, wird ein Eintrag in KW_ITEMS mit folgenden Attributen gespeichert: Schlüssel des Dokuments, Länge des indizierten Eintrags in Worten, Anzahl der Vorkommen in diesem Eintrag, der Erzeuger des Eintrags, Zeiger auf den exakten Eintrag.

KW_ITEM_LINK verbindet schließlich Dokumente und Schlüsselwort-Einträge, was in Abbildung 7 dargestellt ist.

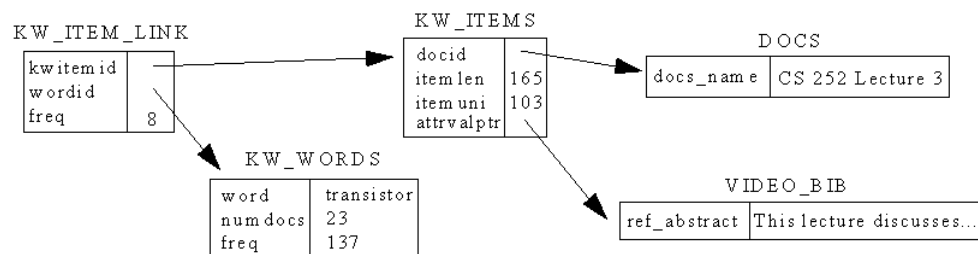


Abbildung 7 Schlüsselwort-Indizierung

Objekt-Indizes dienen zur Indizierung von Objekten im Bild. Dazu zählen Schauspielerinnen und Schauspieler, Objekte die für die Handlung wichtig sind und visuell wichtige Dinge. Für jedes Objekt werden folgende Informationen gespeichert: Art des Objekts, was macht das / geschieht mit dem Objekt, visuelle Charakteristik wie Farbe und Form, einige repräsentative Aufnahmen des Objekts, wann das Objekt im Bild ist, wann das Objekt außerhalb des Bildes, aber immernoch präsent ist.

Es gibt verschiedenen Entitäten, die sich mit Objekten befassen:

PEOPLE enthält folgende Attribute: Name, Geschlecht, Biographie, Hauptbeschäftigung, Anzahl der Dokumente mit denen die Person assoziiert ist, usw.

OBJECT enthält Informationen für jedes wichtige Objekt: Objekttyp (z.B. Tier, Haus, Explosion), Name, Beschreibung, Anzahl der Dokumente, die mit dem Objekt assoziiert sind, usw.

OBJ_INST speichert Informationen über das Vorkommen eines Objektes in einem Film. Es werden u.a. folgende Attribute gespeichert: Position auf dem Schirm, Handlung (was macht das Objekt, was wird mit dem Objekt gemacht), Richtung und

Geschwindigkeit der Bewegung des Objekts und Feld von Zeitintervallen in den das Objekt vorkommt.

3.2.2 Datenerfassung und -recherche

Auch bei diesem Ansatz sollen die Daten mittels eines entsprechender Erfassungs- und Recherchewerkzeuge verwaltet werden. Genauere Vorstellungen scheint es noch nicht zu geben.

3.3 Vereinfachung der Indizierung in der Zukunft¹⁰

In der Zukunft wird man vermehrt an automatischen Indizierungsverfahren forschen. Dies hat den Grund, daß die Indizierungsarbeit sehr personal- und zeitaufwendig ist. Benötigte Verfahren beginnen bei einer zuverlässigen und exakten automatischen Schnitterkennung. Darüber hinaus wird man hochwertige Bilderkennungs- und -analyseverfahren benötigen, die es ermöglichen, die in einem Bild vorkommenden Objekte automatisch zu erkennen und zu indizieren.

Ein sehr ehrgeiziges Ziel ist die dreidimensionale Analyse von Videos. Dabei müßte anhand des zweidimensionalen Bildes die dreidimensionalen räumlichen Verhältnisse und die Bewegungen von Objekten im Raum analysiert werden. Diese Daten könnten dann gespeichert werden.

Ein System mit solchen Eigenschaften würde für wissenschaftliche Untersuchungen (Sport, Medizin, Biologie...) erhebliche Vorteile bringen. Man könnte Bewegungen, Gesten und Verhalten automatisch erfassen lassen und bräuchte hinterher die Daten nur noch zu analysieren. Außerdem könnte man eventuell ganz neue Modell-basierte Kompressions- und Codierungsverfahren erzeugen.

¹⁰ Siehe hierzu speziell: [HAR]

4 Versuchssysteme

4.1 Full Service Network

Das Time Warner / Silicon Graphics Full Service Network™ (FSN) versorgt 4.000 Kabelhaushalte in Orlando, Florida, mit VoD-Diensten. 1996 wurde das Projekt mit dem „The Computerworld Smithsonian Award“ als das Projekt ausgezeichnet, daß die Gesellschaft am stärksten durch den Einsatz von Technologie positiv beeinflusst hat.

FSN hat eine einfach zu bedienende graphische Benutzeroberfläche. Es bietet ca. 200 Movies-on-Demand mit voller Videorekorderfunktionalität (vorspulen, zurückspulen, Pause). Außerdem gestattet es interaktives Einkaufen am Bildschirm, „Nachrichten-on-Demand“, „Sport-on-Demand“ und interaktive Spiele. Es können Informationen über kommunale Ereignisse abgerufen werden und es ist sogar möglich eine Pizza von Pizza Hut zu bestellen.

Geplant sind in Zukunft: Erweiterung der aktuellen Angebote, und Einführung von Musikangeboten, Gesundheits-Angeboten, Home-Banking-Angeboten und Kabel-Modem-Angeboten für Internetzugänge mit hohen Bandbreiten.

Die Hardwareausstattung des Systems besteht aus einem ATM-Netzwerk mit einer Bandbreite von 20 Gigabit / s. Silicon Graphics Inc. stellte den Server und die digitale Hardware, wie z.B. den MIPS R4000-Prozessor für die Home-Communication-Terminals, die Set-Top-Boxen. Von Atari stammt die Hardware für die Spielfunktion des Systems. In den Haushalten werden Farbdrucker von HP verwendet.

Es wurden über 1.000 Meilen Glasfaser-Kabel verlegt. Von 24 Knoten aus führen Koaxial-Kabel mit einer Bandbreite von über 1 GHz in die Wohnungen.

Im System werden soweit möglich Standardprotokolle und -technologien verwendet: SONET, TCP/IP und SNMP.

Die Server werden mit je 8 MIPS-Prozessoren, die in einer verteilten Architektur arbeiten, betrieben. Es können mehrere Terabyte an Daten auf Plattenlaufwerken, mittels einer RAID-ähnlichen Technologie, gespeichert werden. In Zukunft sollen auch Chip- bzw. RAM-basierte Speichertechnologien zum Einsatz kommen.

Video-Daten werden nach dem MPEG-1-Standard kodiert.

4.2 BT Interactive Television

1994 startete die British Telecom (BT) ein Projekt zum Thema „Interaktives Fernsehen“. Bis Mitte 1995 wurden rund 2500 Haushalte an das System angeschlossen. Während der Testphase sollten folgende Dienste angeboten werden: Shopping-on-Demand, Bildungsprogramme für private und schulische Zwecke, Movie-on-Demand, Home-Banking-Dienste, ein Magazin-Service und lokale Informationsdienste.

Das System lief mit einer multimedialen Datenbank und Navigationssoftware von Oracle, nCUBE stellte den massiv-parallelen Video-Server (Leistung äquivalent zu 5000 Pentium-CPU's). Das Kunden-Managementsystem stammte von Sequent und Apple lieferte die Set-Top-Boxen auf Basis des Apple Macintosh LC 475 mit modifizierter Hard- und Software.

Zur Übertragung wurden Kupferkabel mit der Übertragungstechnologie ADSL (Asymmetric Digital Subscriber Loop, eine Art von Spezialmodem-Verfahren) oder Glasfaser verwendet. ADSL bietet über normale Telefonleitungen 2 Mb / s in Richtung zum Kunden und einen Steuerungskanal mit 9,6kb / s. Als Protokoll diente ATM.

Als Videokompression wurde MPEG-1 verwendet.

1996 wurde das Projekt erfolgreich beendet.

5 Resümee

Dargestellt wurden einige Konzepte für eine effiziente Implementierung von VoD-Systemen, was aufzeigt, daß an verschiedenen Punkten der Implementierung von VoD-Systemen noch erheblicher Forschungsbedarf besteht. Vor allem müssen die verschiedenen, vorgeschlagenen Konzepte auch einmal in einem echten Test verwendet werden. Die tatsächlichen Probleme, die in einem großen VoD-System entstehen, dürften in kleinen Testanlage nur unzureichend zu untersuchen sein.

Außerdem müssen diejenigen, die sich mit VoD-Projekten oder -Strategien befassen, einsehen, daß es notwendig ist, dem Kunden eine konstant hohe Übertragungsqualität zu bieten, wenn VoD-Dienste erfolgreich eingeführt und vom Kunden akzeptiert werden sollen. Auch ist es notwendig, daß ein VoD-System für den Kunden nicht teurer sein darf als vergleichbare heutige Dienste.

All dies, in Verbindung mit höheren Netzwerkübertragungsraten und schnelleren Servern, bietet dann eine Basis für die erfolgreiche Verwendung von VoD-Diensten.