

# **Multimedia in der Hochschullehre - Ein Überblick über die Grundlagen und aktuelle Konferenz-Werkzeuge**

Julia Wunner

Wolfgang Rosenstiel

WSI-99-7

Technische Informatik  
Wilhelm-Schickard-Institut  
Universität Tübingen  
D-72076 Tübingen, Germany  
Tel ++49-(0)7071-29-75940  
Fax ++49-(0)7071-29-5062  
Email {wunner|rosen}@informatik.uni-tuebingen.de  
WWW <http://www-ti.informatik.uni-tuebingen.de>

©WSI 1999  
ISSN 0946-3852



# **Multimedia in der Hochschullehre - Ein Überblick über die Grundlagen und aktuelle Konferenz-Werkzeuge**

Julia Wunner, Wolfgang Rosenstiel

April 1999

## **Kurzzusammenfassung**

Der Einsatz von Multimedia in der Lehre, speziell in der Hochschullehre, schafft völlig neue Möglichkeiten: Die Zeit- und Ortsgebundenheit von Lehre und Lernen kann aufgehoben, sowie Präsenzlehre von Routineaufgaben entlastet werden. Es müssen nicht mehr alle Hochschulen und Fakultäten ein vollständiges Lehrangebot im jeweiligen Fach bereithalten. Hochschulinterne wie hochschulübergreifende Kooperationen ermöglichen, das Studienangebot zu vervollständigen oder gemeinsame Angebote zu entwickeln. In der vorliegenden Arbeit werden die Grundlagen der Netzwerktechnik und Echtzeit-Kommunikation erläutert und aktuelle Konferenz-Werkzeuge vorgestellt.



# Inhaltsverzeichnis

<b>Abkürzungen</b>	<b>3</b>
<b>1 Einleitung</b>	<b>5</b>
<b>2 Grundlagen</b>	<b>6</b>
2.1 Netzwerktechnik . . . . .	6
2.2 ITU Konferenz-Standards . . . . .	26
2.3 Multimedia-Dateiformate . . . . .	31
2.4 Multimedia-Einsatz in der Lehre . . . . .	41
2.5 Anforderungen . . . . .	51
<b>3 Stand der Technik</b>	<b>56</b>
3.1 PictureTel . . . . .	56
3.2 VTEL . . . . .	58
3.3 Cisco IP/TV . . . . .	60
3.4 Intel ProShare . . . . .	62
3.5 Microsoft NetMeeting . . . . .	64
3.6 Sun ShowMe TV . . . . .	65
3.7 MBone Tools . . . . .	67
<b>4 Fazit und Ausblick</b>	<b>77</b>
<b>Literaturverzeichnis</b>	<b>79</b>



# Abkürzungen

ATM	Asynchronous Transfer Mode, S. 12
CIF	Common Interchange Format, S. 39
CSCW	Computer Supported Cooperative Work, S. 41
CSDN	Circuit Switched Digital Networks, S. 29
fps	frames per second, S. 38
FtF	Face to Face, S. 43
FTP	File Transfer Protocol, S. 15
GCC	Generic Conference Control, S. 30
GIF	Graphical Interchange Format, S. 37
GUI	Graphical User Interface, S. 53
HTML	Hypertext Markup Language, S. 33
IGMP	Internet Group Management Protocol, S. 23
INRIA	Institut National de Recherche en Informatique et en Automatique
IP	Internet Protocol, S. 15
ISDN	Integrated Services Digital Network, S. 29
ISO	International Standards Organization, S. 13
ITU	International Telecom Union, S. 26
JPEG, JPG	Joint Photographic Experts Group, S. 37
LAN	Local Area Network, S. 6
LBNL	Lawrence Berkeley National Laboratory
MBone	Multicast Backbone, S. 23
MCU	Multipoint Control Unit, S. 27
MCS	Multipoint Communication Services, S. 29
MPEG	Motion Picture Expert Group, S. 36

NACK	Negative Acknowledgement, S. 21
NIC	Network Information Center, S. 10
OSI	Open Systems Interconnection, S. 13
PCM	Pulse Code Modulation, S. 32
PDF	Portable Document Format, S. 34
PiP	Picture in Picture, S. 66
POTS	Plain Old Telephone System, S. 29
PS	PostScript, S. 34
PSDN	Public Switched Digital Networks, S. 29
PSTN	Public Switched Telephone Networks, S. 29
QCIF	Quarter CIF, S. 39
QoS	Quality of Service, S. 12
RAS	Registration Admission Status, S. 27
RFC	Request For Comments
RIPE	Réseaux IP Européens, S. 10
RTF	Rich Text Format, S. 33
RTP	Realtime Transport Protocol, S. 17
SAP	Session Announcement Protocol, S. 25
SDP	Session Description Protocol, S. 25
SGML	Structured Generalized Markup Language, S. 33
SMTP	Simple Mail Transfer Protocol, S. 15
TCP	Transmission Control Protocol, S. 16
TTL	Time To Live, S. 16
UCB	University of California, Berkeley
UCL	University College London
UDP	User Datagram Protocol, S. 16
WWW	World Wide Web, S. 33



# Kapitel 1

## Einleitung

Der Einsatz von Multimedia in der Lehre, speziell in der Hochschullehre, schafft völlig neue Möglichkeiten: Die Zeit- und Ortsgebundenheit von Lehre und Lernen kann aufgehoben, sowie Präsenzlehre von Routineaufgaben entlastet werden. Es müssen nicht mehr alle Hochschulen und Fakultäten ein vollständiges Lehrangebot im jeweiligen Fach bereithalten. Hochschulinterne wie hochschulübergreifende Kooperationen ermöglichen, das Studienangebot zu vervollständigen oder gemeinsame Angebote zu entwickeln.

Im Rahmen des Projekts "Multimedia-gestützte Studiengänge an Hochschulen - Bioinform@tik"<sup>1</sup> wurden zunächst die Grundlagen der Echtzeit-Kommunikation über Internet zusammengetragen. Dies beinhaltet eine Einführung in Netzwerktechnik und -architektur, die verschiedenen gebräuchlichen Medienformate und die Diskussion möglicher Lehr- und Lernszenarien an Hochschulen (Kapitel 2).

Für den zweiten Teil (Kapitel 3) wurden aktuelle Konferenz-Werkzeuge, die für die Abwicklung des Audio-, Video- und sonstigen Datenverkehrs zwischen den Teilnehmern benötigt werden, auf ihre Eignung in der multimedial gestützten Lehre untersucht. Abschließend werden die Ergebnisse zusammengefaßt und das Fazit gezogen, daß für die effiziente Durchführung von synchronen multimedialen Lehrveranstaltungen noch ein universelles Konferenz-Werkzeug benötigt wird, das die Vorteile der Multimedia-Kommunikation wie Ortsungebundenheit und Flexibilität effizient nutzt und mögliche Nachteile wie hohen technischen, personellen und organisatorischen Aufwand minimiert (Kapitel 4).

---

<sup>1</sup>Bioinform@tik, <http://www-ra.informatik.uni-tuebingen.de/bioinformatik/welcome.html>

# Kapitel 2

## Grundlagen

In diesem Kapitel werden die für das Verständnis nötigen Grundlagen aus den Gebieten Netzwerktechnik, Multimedia und Videoconferencing vorgestellt. Im Bereich Netzwerktechnik werden Netzwerk-Architekturen und die für Video- und Datenkonferenzen relevanten Protokolle, insbesondere Echtzeitprotokolle, vorgestellt. Der Abschnitt Multimedia beschäftigt sich mit der Kompression der besonders bei Audio- und Videodaten anfallenden großen Datenmengen und präsentiert die wichtigsten Multimedia-Dateiformate. Das Kapitel wird durch einen Abschnitt über Methoden und Anforderungen von Videokonferenzen und ihren speziellen Einsatz in der Hochschullehre abgeschlossen.

### 2.1 Netzwerktechnik

In diesem Abschnitt werden die zum Verständnis notwendigen Begriffe aus dem Bereich der Netzwerktechnik vorgestellt. Dabei wird zu Beginn definiert, was genau ein Netzwerk oder Rechnernetz ausmacht, welche verschiedenen Formen solche Netze annehmen können und welche Datenübertragungsarten existieren. Es werden anschließend die im Internet gebräuchlichen Adressierungs- und Routingmechanismen und die dabei eingesetzte Paketvermittlung (im Gegensatz zur Leitungsvermittlung) erläutert. Software, die für die Kommunikation zwischen Rechnern eingesetzt wird, unterteilt man auf Grund ihrer hohen Komplexität in Schichten. Es wird das OSI und das TCP/IP Modell vorgestellt. Für Konferenzanwendungen sind speziell die Echtzeitprotokolle RTP und LRMP als Protokolle der Transportschicht interessant, sowie der Multicast Backbone (MBone) als Basis für Multimediaanwendungen auf der Applikationsschicht. Diese Themen werden in eigenen Abschnitten ausführlicher behandelt.

#### Definition

Die Begriffe **Netzwerk** oder **Rechnernetz** werden im folgenden als Synonym für miteinander verbundene, eigenständige Rechner, die untereinander Daten austauschen [70], benutzt. Beispiele sind lokale Rechnernetze ( "Local Area Networks", LANs) oder das weltumspannende Internet. Dabei bezeichnet man die einzelnen Rechner als **Knoten** des Netzes,

die physikalischen Leitungen zwischen den Rechnern als **Verbindungen**. Die charakteristische Eigenschaft einer Verbindung ist ihre **Bandbreite**, also die Datenmenge, die pro Zeiteinheit verschickt werden kann und üblicherweise in MBit/s angegeben wird.

Rechner werden aus unterschiedlichen Gründen zu Netzen zusammengeschlossen: Bei einem Datenverbund können alle Rechner gegenseitig auf ihre Datenbestände, beispielsweise Bibliotheksdatenbanken, zugreifen. Bei einem Betriebsmittelverbund können Betriebsmittel wie schnelle Rechner, Drucker oder ähnliches von allen angeschlossenen Rechnern genutzt werden. In einem Lastverbund kann die benötigte Rechenleistung von mehreren Rechnern gemeinsam erbracht werden. Schließlich kann durch den Austausch von Text, Bild, Audio und Videodaten ein Rechnernetz einen **Kommunikationsverbund** bilden und so das herkömmliche Telefonnetz ersetzen und sogar übertreffen.

### Topologien von Rechnernetzen

Je nach der Verteilung der Knoten eines Netzes und der Anzahl und Anordnung seiner Verbindungen unterscheidet man verschiedene Netztopologien, die jeweils spezielle Eigenschaften besitzen [70]. Bei einer **Vollvermaschung** erreicht man die höchste Ausfallsicherheit und die kürzesten Übertragungszeiten zwischen den Knoten. Allerdings steigt hier die Anzahl der Verbindungen und damit die Kosten überproportional mit der Anzahl der Knoten.

Eine Vollvermaschung ist in der Praxis selten sinnvoll, da die Verbindungen erfahrungsgemäß nie voll ausgelastet sind. Wenn die Daten vom Sender zum Empfänger von dazwischenliegenden Knoten weitergeleitet werden, kann man viele Verbindungen einsparen. Bei solchen nicht vollständig vermaschten Netzen unterscheidet man drei Grundstrukturen. Übernimmt nur ein zentraler Knoten die Weiterleitung, so spricht man von einem **Stern**. Diese kostengünstige und gut ausgelastete Topologie hat allerdings den Nachteil, daß der Ausfall des Vermittlungsknotens einen Totalausfall des Netzes bedeutet. Ausfallsicherer ist die Anordnung der Knoten in Form eines **Rings**, siehe Abbildung, da hier die Kommunikation zwischen zwei Knoten auf zwei disjunkten Strecken in die zwei möglichen Richtungen erfolgen kann. Die Ring-Topologie ist ebenfalls kostengünstig, allerdings können die Verbindungszeiten je nach Richtung stark schwanken, und die Überlastung von Teilstrecken kann erheblichen Einfluß auf die Gesamtleistungsfähigkeit haben. Ähnliche Eigenschaften wie ein Ringnetz, allerdings ein ungünstigeres Ausfallverhalten, besitzt die Anordnung der Knoten in einer Reihe in einem **Bus**.

Reale Rechnernetze, deren Struktur entscheidend durch die geographische Lage der Rechner und die jeweilige Aufgabenstellung geprägt ist, nutzen Vorteile und vermeiden Nachteile durch die Kombination der oben genannten Grundstrukturen, gegebenenfalls in mehreren Hierarchieebenen. So wurden im Internet leistungsfähige Hochgeschwindigkeitsverbindungen (**Backbones**) geschaffen, an die lokale Netze an bestimmten Übergabepunkten an-

geschlossen sind<sup>1</sup>. Auf diese Weise werden Daten nicht über zu viele unbeteiligte Knoten oder Teilnetze verschickt.

### Kommunikationsformen

Im Internet werden Daten normalerweise von genau einem Sender(-Rechner) an genau einen Empfänger(-Rechner) verschickt. Diese 1:1-Kommunikation nennt man **Unicast** oder Punkt-zu-Punkt-Kommunikation. Tauschen gleichzeitig mehr als zwei Rechner untereinander Daten aus, so unterscheidet man mehrere Arten der Gruppenkommunikation [77], siehe Abbildung 2.1. Die Kommunikation zwischen einem Sender und mehreren Empfängern (1:n-Kommunikation) wird als **Multicast** bezeichnet und eignet sich beispielsweise für die Übertragung von Vorträgen oder das Verteilen von HTML-Seiten an eine Gruppe. Dabei werden die Daten nicht einfach n mal per Unicast versendet, sondern ressourcen-schonend nur einmal, und nur bei solchen Zwischensystemen (dazwischenliegenden Knoten oder Teilnetzen) an den Ausgängen dupliziert, an denen auch Empfänger vorhanden sind. Multicast zeichnet sich daher durch eine sehr gute **Skalierbarkeit** auch bei großen Teilnehmergruppen aus. Einfacher, aber Ressourcen-vergeudend und daher nur in Nahverkehrsnetzen praktikabel ist der Versand von einem Sender an alle erreichbaren Empfänger. In diesem Fall spricht man von **Broadcast**. Im Vergleich zu Multicast benötigt man beim Broadcast keine Unterscheidung zwischen Gruppenangehörigen und Nicht-Gruppenangehörigen.

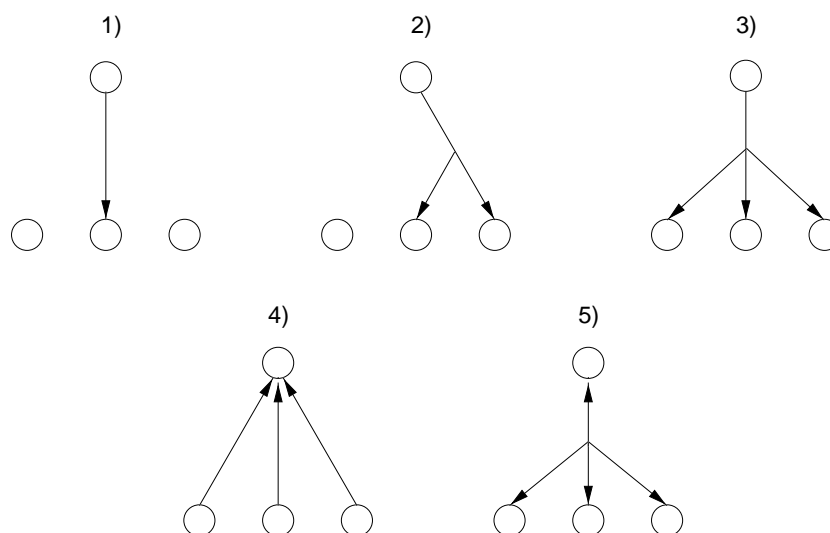


Abbildung 2.1: Kommunikationsformen: 1) Unicast, 2) Multicast, 3) Broadcast, 4) Concast, 5) Multipeer

Umgekehrt spricht man von einer m:1-Kommunikation oder **Concast**, wenn mehrere Sender Daten an einen Empfänger schicken. Im Lehrbereich liegt dieser Fall beispielsweise

<sup>1</sup>In Deutschland existiert z.B. das deutsche Wissenschaftsnetz WIN mit 34 Mbit/s Bandbreite, oder der europäische Ebone, der 72 kommerzielle Netze in 31 Ländern miteinander verbindet.

vor, wenn mehrere Studenten Ergebnisse an einen Tutor senden. Konferenzen, in denen alle Teilnehmer ohne Einschränkungen Daten austauschen, bezeichnet man als m:n-Kommunikation oder **Multipeer**. Die technische Realisation dieser beim Tele-Learning oft wünschenswerten Mehrpunkt-Kommunikation ist sehr schwierig, sie läßt sich aber durch mehrere gleichzeitige Multicast-Verbindungen bewerkstelligen. Durch die Emulation kann es jedoch zu Situationen kommen, in denen die Empfänger auf Grund von unterschiedlich starker Netzbelastung die versendeten Daten in unterschiedlicher Reihenfolge erhalten. Wird dies nicht gewünscht, so müssen zusätzliche Maßnahmen zur **Ordnungserhaltung** ergriffen werden.

### Gruppenkommunikation

Neben den oben erwähnten Eigenschaften der Ordnungserhaltung und Skalierbarkeit sind für die Kommunikation in Gruppen auch noch andere Aspekte entscheidend [77]. Hierzu gehört zum einen die **Offenheit**. Bei offenen Gruppen können auch nicht der Gruppe angehörige Sender Daten an die Gruppe schicken, bei geschlossenen Gruppen geht dies nicht. In der Lehre sind beide Versionen sinnvoll, beispielsweise bei Diskussionen. Ein nächster Aspekt ist die **Dynamik** einer Gruppe. Statische Gruppen bleiben während der gesamten Kommunikation in ihrer Teilnehmeranzahl gleich, beispielsweise bei einer virtuellen Projektbesprechung. Können während der Kommunikation Teilnehmer ausscheiden und neue hinzukommen, beispielsweise in einer großen Konferenz, spricht man von einer dynamischen Gruppe. Auch die **Sicherheit** spielt bei der Gruppenkommunikation eine wichtige Rolle. Einzelne Datenströme wie Audio oder Video können phasenweise oder ständig verschlüsselt versendet werden um sie nur einem bestimmten Teilnehmerkreis, beispielsweise Studenten eines Kurses, zugänglich zu machen. Die Verschlüsselungsmechanismen müssen zudem von der üblichen Punkt-zu-Punkt-Kommunikation an die Gruppenkommunikation mit mehreren Teilnehmern angepaßt sein.

Hinsichtlich der **Lebensdauer** kann man permanente und transiente Gruppen unterscheiden. Permanente Gruppen bestehen im Gegensatz zu transienten Gruppen auch dann fort, wenn keine aktiv (sendenden) Teilnehmer mehr vorhanden sind. Die **Bekanntheit** einer Gruppe bezieht sich auf die Identität der einzelnen Gruppenteilnehmer. In anonymen Gruppen sind die Teilnehmer der restlichen Gruppe nicht (ständig) bekannt, anders als in bekannten Gruppen. Daneben unterscheidet man Gruppen nach ihrer **Heterogenität**, also ob sich die Teilnehmer unterschiedliche Eigenschaften besitzen, beispielsweise in einem Tutor-Student-Szenario, oder ob alle Teilnehmer gleich sind, beispielsweise in einer freien Diskussion. Heterogenität kann sich aber auch durch unterschiedliche physikalische Anbindungen an die Gruppe ausbilden, beispielsweise wenn ein Teil der Gruppe über Hochleistungsnetze kommuniziert, ein anderer Teil über langsame Modemverbindungen.

Speziellere Aspekte der Gruppenkommunikation sind **Zuverlässigkeit**, **Flußkontrolle** und **Gruppenverwaltung** [77]. Der zuverlässige, also korrekte und geordnete Empfang von Da-

ten läßt sich bei der Punkt-zu-Punkt-Kommunikation durch Quittierung der empfangenen Daten realisieren. In größeren Gruppen kann es dadurch jedoch zu Quittungsimplosionen kommen und infolgedessen schlimmstenfalls zum Erliegen der gesamten Kommunikation. Bei manchen Datenströmen, beispielsweise Audio und Video, sind geringe Datenverluste tolerabel, sie können also auch unzuverlässig an eine Gruppe gesendet werden. Andere Daten wie Bild oder Text müssen jedoch halb- oder voll zuverlässig übertragen werden. Hier kommen verschiedene, meist empfangerbasierte Mechanismen zur Anwendung, die weiter unten genauer erläutert werden. Mit der Zuverlässigkeit gekoppelt ist eine Flußkontrolle zur Koordination heterogener Gruppen mit unterschiedlicher Netzanbindung.

Die Adressierung in Gruppen benötigt ebenfalls eigene Mechanismen, da nicht mehr ein einzelner Empfänger (Unicast) oder alle Empfänger (Broadcast) angesprochen werden, sondern eine dedizierte Gruppe. Möglich ist dabei die Adressierung über eine Liste mit Empfängeradressen oder alternativ über eine spezielle Gruppenadresse.

### **Adressen, Routing, Ports**

Im Internet sind Sender(-Rechner) und Empfänger(-Rechner) von Daten durch ihre weltweit eindeutige **IP-Adresse** (von "Internet Protokoll", **IP**, siehe weiter unten) identifizierbar und ansprechbar [70]. In der Version 4 des Internet Protokolls ist eine IP-Adresse eine 32 Bit große Zahl, die meist in der sogenannten "Dotted Decimal Notation" (vier durch Punkte getrennte Oktets) angegeben sind, beispielsweise 134.2.14.27. Der besseren Merkbareit halber können die Rechner auch durch einen oder mehrere Namen wie `struppi.informatik.uni-tuebingen.de` angesprochen werden. Die Umsetzung von Namen in IP-Adressen wird dabei von speziellen Rechnern, sogenannten "Name Servern", durchgeführt.

Im maximal möglichen Adreßraum von 0.0.0.0 bis 255.255.255.255 sind einige Adressen bereits für bestimmte Zwecke reserviert, beispielsweise `x.x.x.255` für einen Broadcast an alle Rechner im lokalen Netz. Die Adressen sind der besseren Übersichtlichkeit halber in fünf Klassen unterteilt, die durch die ersten Bits des ersten Bytes charakterisiert werden, siehe Abbildung 2.2. Bei den folgenden Bits unterscheidet man zusätzlich Netzwerk- und Rechneradresse.

Organisationen oder Firmen, die ihren Rechnern IP Adressen zuteilen wollen, weil sie Daten im Internet versenden möchten, bekommen vom jeweiligen "Network Information Center" (NIC)<sup>2</sup> eine Netzwerk-Adresse zugewiesen. Die Rechneradressen können von den Organisationen oder Firmen beliebig auf die vorhandenen Rechner verteilt werden. Class A Adressen waren für sehr große Organisationen gedacht, die über 16 Millionen Rechner besitzen. Sie werden auf Grund ihrer Größe nicht mehr vergeben. Üblich sind Class B Adressen für größere Organisationen mit maximal etwa 65.000 Rechnern und Class C Adressen für kleinere

---

<sup>2</sup>In Europa ist dafür RIPE ("Réseaux IP Européens") zuständig, <http://www.ripe.net>.

Organisationen mit bis zu 256 Rechnern. Class D Adressen sind für Multicast Adressen reserviert, die nicht einem einzelnen Rechner, sondern einer sich dynamisch ändernden Gruppe von Rechnern zugeordnet sind, wobei die teilnehmenden Rechner natürlich noch zusätzlich eine "richtige" IP Adresse besitzen. Das Konzept der Multicast Adressen wurde bereits 1985 entwickelt [11][12] und wird seit 1992 mit wachsender Verbreitung im Internet eingesetzt. Class E Adressen sind für eine zukünftige Nutzung reserviert.

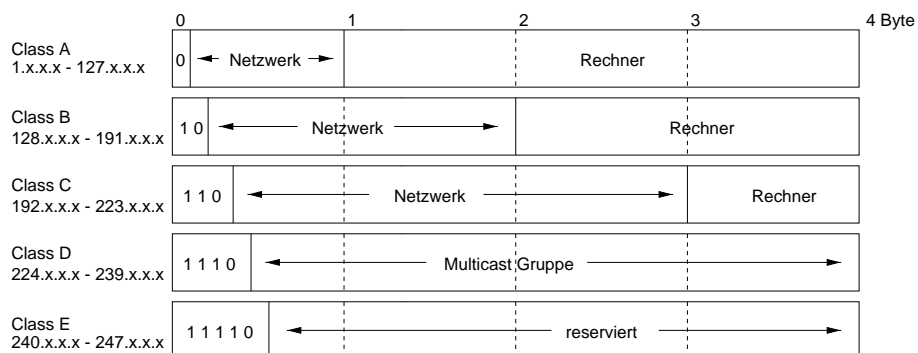


Abbildung 2.2: IP Adressklassen

Das weltumspannende Internet ist auf Grund seiner Größe in einzelne logische Bereiche, sogenannte Autonome Systeme (AS), unterteilt, die selbständig administriert werden, also für die verwendeten Protokolle, Standards, Software und deren Wartung und den Datenaustausch mit anderen AS an definierten Punkten zuständig sind [70, 43]. Da im Internet jeder Rechner mit einem beliebigen anderen kommunizieren kann, müßten theoretisch in einem Rechner die Wege zu allen anderen Rechnern gespeichert sein. Da dies wegen der sehr großen Zahl von angeschlossenen Rechnern nicht möglich ist, setzt man für die Wegewahl spezielle Vermittlungsrechner (**Router**) ein, die in sogenannten Routing Tabellen die Wege zu den zur Zeit etwa 45.000 Routern gespeichert haben. Es gibt verschiedene Strategien für das Routing innerhalb eines und zwischen AS. Man unterscheidet statisches Routing, bei dem die Wegewahl fest vorgegeben ist und das aufgrund seiner Unflexibilität heute veraltet ist, und dynamisches Routing, bei dem in der Tabelle alternative Wege mit angegeben werden können. Es existieren verschiedene Routing Algorithmen, die im einfachsten Fall auf der minimalen Anzahl der "Hops", also der Anzahl der durchlaufenen Router zum Ziel, beruhen.

Moderne Rechner können gleichzeitig Sender und Empfänger von mehreren Datenströmen sein. Das bedeutet, daß die IP Adresse alleine zur Identifikation der Endpunkte eines Datenstroms nicht mehr ausreicht. Jedem Datenstrom im Internet wird daher zusätzlich an jedem Endpunkt ein logischer **Port** zugewiesen. Rechner verfügen über  $2^{16} = 65.536$  Ports, wobei die Ports 0 bis 1023 bereits für bekannte Dienste wie Dateiübertragung und Email reserviert sind [59]. IP-Adresse und Port des Senders und Empfängers stellen zusammen jeweils einen Kommunikationsendpunkt, auch **Socket** genannt, dar. Im Rechner steht jedem

Socket ein reservierter Speicherplatz als Kommunikationspuffer zur Verfügung, in dem die zu sendenden oder empfangenen Daten zwischengespeichert oder abgeholt werden können.

### Vermittlungs- und Verbindungsarten

Bis jetzt wurde nur allgemein von Daten geredet, die zwischen Rechnern verschickt werden. Man unterscheidet jedoch zwei grundlegend verschiedene Vermittlungsarten. Bei der **Leitungsvermittlung** (“Circuit Switching”), die nach dem Prinzip des analogen Telefonnetzes funktioniert, wird für die Dauer der Verbindung eine eigene, vom Sender zum Empfänger durchgehende Leitung mit einer konstanten Übertragungsrate geschaltet, was einerseits Echtzeitkommunikation in mehr oder weniger guter Qualität ermöglicht, andererseits aber auch sehr aufwendig und unflexibel ist. Für die bei Multimedia-Anwendungen anfallenden sehr unterschiedlichen und meist sehr großen Datenmengen ist eine Leitungsvermittlung also nicht geeignet. Flexibler ist im Gegensatz dazu die **Paketvermittlung** (“Packet Switching”), die nach dem Prinzip der Briefpost arbeitet. Dabei werden die Daten in kleinere Datenstücke (**Pakete**) fester oder variabler Länge aufgeteilt, die anschließend einzeln verschickt werden und auch auf verschiedenen Wegen zum Empfänger gelangen können. Diese Methode ist gut skalierbar, hat jedoch den Nachteil, daß die unbekannte Wegewahl, das Aufspalten in Pakete beim Sender und das Wieder-Zusammensetzen beim Empfänger eine beträchtliche Zeitverzögerung nach sich ziehen kann, die jedoch durch geschickte Routing-Mechanismen und eine geschickte Wahl der Paketgröße minimiert werden kann.

Neben der Leitungs- und Paketvermittlung unterscheidet man außerdem **verbindungsorientierte** (“connection-oriented”) und **verbindungslose** (“connection-less”) Verbindungen, wobei bei den ersteren ein Zusammenhang zwischen den übertragenen Paketen wie eine einzuhaltende Reihenfolge oder die Wahl eines bestimmten Weges durch das Netz besteht, bei letzteren nicht. Des weiteren differenziert man zwischen **ungesicherten** und **gesicherten** beziehungsweise **garantierten** Verbindungen, die Mechanismen zur Verfügung stellen, die garantieren, daß alle Daten unverfälscht und auf jeden Fall beim Empfänger ankommen. Wird zusätzlich eine maximale Zeitverzögerung garantiert, spricht man von einer Garantie der “Quality of Service” (**QoS**).

Der Vollständigkeit halber muß erwähnt werden, daß mit dem Asynchronen Transfermodus (**ATM**) eine Vermittlungstechnik existiert, die funktional zwischen Paket- und Leitungsvermittlung liegt, dadurch sehr schnell ist und eine QoS garantieren kann [3][57]. Diese 1991 von verschiedenen Telefongesellschaften und der Computerindustrie realisierte schnelle Paketvermittlung arbeitet mit Paketen einer festen kleinen Größe und reserviert vor der eigentlichen Datenübertragung beim Verbindungsaufbau die benötigten Netzressourcen. Durch die kleine Paketgröße können die Puffer in den Routern ebenfalls klein gehalten werden, wodurch sich die Vermittlungsverzögerung weiter verringert. ATM ist also sehr gut für die Übertragung von Echtzeitdaten und hohen Datenmengen geeignet. Allerdings hat sich die Verwendung von ATM bis jetzt lediglich für Backbones durchge-



setzt. Die große Mehrzahl der Endbenutzer im Internet benutzt jedoch nicht ATM sondern das normale Internet Protokoll, das daher in den nächsten Abschnitten näher beleuchtet wird.

## Schichtenmodelle

Auf unterster Ebene funktioniert das Internet, wie oben erwähnt, paketorientiert. Auf höherer Ebene gibt es je nach Anspruch der Anwendung sowohl die Möglichkeit einer verbindungsorientierten Datenübertragung mit eventuell garantierter QoS, wie auch die ungesicherte verbindungslose Datenübertragung. Die in allen Fällen benötigte Kommunikationssoftware ist so komplex, daß sie hierarchisch strukturiert und in logische Schichten oder Ebenen (“Layers”, “Levels”) unterteilt wird. Jede Schicht ist für spezielle Aufgaben zuständig und stellt der jeweils höheren über eine Schnittstelle (“Interface”) primitive Funktionen, sogenannte **Dienste** (“Services”) zur Verfügung, beispielsweise Dienste zur Verbindungsnachfrage und dem Verbindungsaufbau [70]. Den eigentlichen Daten wird optional von jeder Schicht ein Datenkopf (**Header**) und Anhang (**Trailer**) hinzugefügt, siehe Abbildung 2.3, der die Identifikation der zu einer Schicht gehörenden Informationen enthält und die Datengrenzen festlegt. Schichten kommunizieren über sogenannte **Protokolle** miteinander, die Kommunikationsregeln für das Format und die Bedeutung der ausgetauschten Daten festlegen, wie beispielsweise die Länge und Bedeutung eines Headers. Protokolle werden zur Implementierung der Dienste benötigt. Eine Menge von Protokollen und Schichten bezeichnet man als eine Netzwerkarchitektur oder **Schichtenmodell**.



Abbildung 2.3: Header und Trailer eines Datenpakets in einem Zwei-Schichtenmodell

Bevor auf die bei Video- und Datenkonferenzen benötigten Protokolle eingegangen wird, müssen zunächst zwei Schichtenmodelle vorgestellt werden: Einerseits das von der International Standards Organization (ISO) 1983 vorgeschlagene “Open Systems Interconnection“ Referenzmodell (**OSI Modell**) [10] mit sieben Schichten, das sich zwar in der Praxis nicht durchsetzen konnte, sich aber sehr gut für theoretische Überlegungen eignet; andererseits das im heutigen Internet verwendete und in der Praxis entstandene **TCP/IP Modell** (“Transmission Control Protocol/Internet Protocol”) mit vier Schichten, das erstmals 1974 von Mitbegründern des ARPANET definiert wurde [56].

## OSI Modell

Das OSI Modell unterscheidet sieben Schichten. Die unterste Schicht 1 ist die Bitübertragungs- oder **Physikalische Schicht** (“Physical Layer”), die für die Übertragung des aus Nullen und Einsen bestehenden Bitstroms über das physikalische Medium, zum Beispiel

Kabel oder Funk, zuständig ist. Das zugehörige Protokoll legt beispielsweise fest, welche Spannungen Nullen und welche Einsen repräsentieren.

Schicht 2 ist die **Sicherungsschicht 2** (“Data Link Layer”), in der typischerweise mehrere Hundert Bits in sogenannte **Frames** gepackt und als Einheit verschickt werden und optional mit einem Header und Trailer versehen. Beim Empfänger ist die Sicherungsschicht für die Identifikation der einzelnen Frames beziehungsweise Frame-Grenzen im Bitstrom zuständig. Die Sicherungsschicht behebt außerdem Probleme, die durch das physikalische Medium entstehen können. Dazu gehören Fehler bei der Übertragung, die durch Prüfsummen erkannt oder durch fehlerkorrigierende Codes berichtigt werden können, aber auch eine Flußkontrolle zum Drosseln eines zu schnellen Senders.

Die darüberliegende die Vermittlungs- oder **Netzwerkschicht 3** (“Network Layer”) wählt für Pakete, die aus mehreren Frames bestehen, den Weg zum Ziel. Der Header dieser Schicht enthält typischerweise die Sender- und Empfängeradresse. Führt der Weg eines Pakets über einen oder mehrere dazwischenliegende Router, so wird das Paket beim Router bis zur Netzwerkschicht ausgepackt, die Wahl des Weges auf Grund der Empfängeradresse im Header der Netzwerkschicht festgelegt, und das Paket je nach Fall zum nächsten Router oder zum Empfangs-Rechner weitergeschickt, siehe Abbildung 2.4. Neben dem Routing ist die Netzwerkschicht auch dafür verantwortlich, daß die Pakete so schnell wie möglich weitergeleitet werden und sich insbesondere nicht zu viele Pakete anstauen.

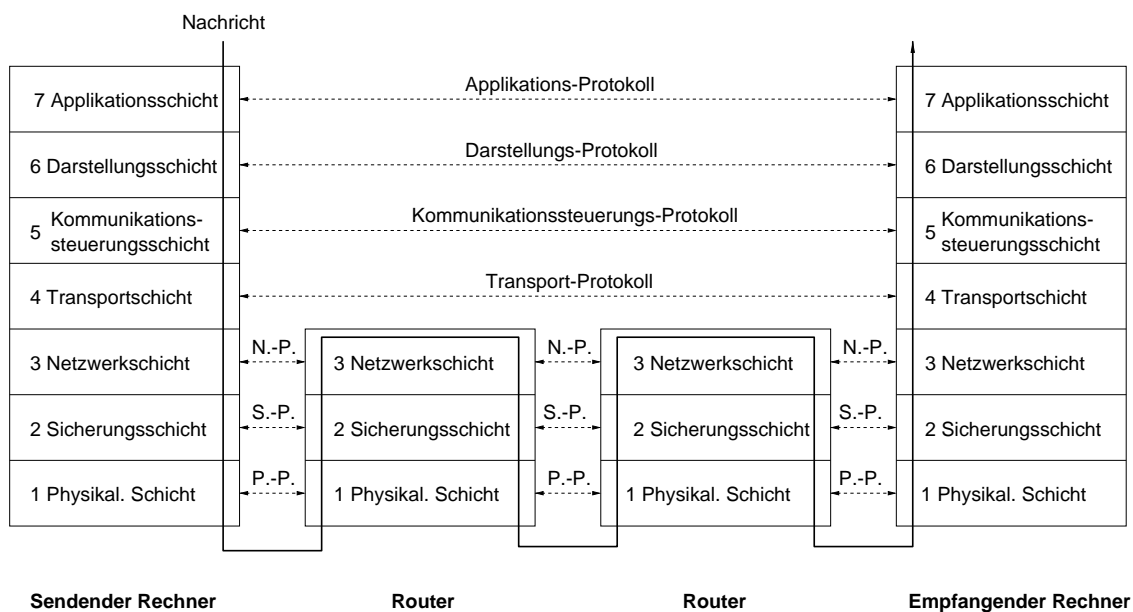


Abbildung 2.4: Das OSI Modell und der Weg einer Nachricht von einem Rechner zu einem anderen über dazwischenliegende Router

Die **Transportschicht 4** (“Transport Layer”) erledigt den zuverlässigen und effizienten

Datentransport vom Sender zum Empfänger. Ab dieser Schicht beginnt die virtuelle Kommunikation von Endpunkt zu Endpunkt, die für die eventuell dazwischenliegende Router nicht sichtbar sind. Im Normalfall erzeugt die Transportschicht für jede von der darüberliegenden Schicht Verbindung eine eigene Netzwerkverbindung. Dafür ist eine Flußkontrolle zwischen den beiden End-Rechnern nötig.

Die **Kommunikationssteuerungsschicht** 5 (“Session Layer”) organisiert und steuert den Dialog zwischen den End-Rechnern. Sie stellt erweiterte Dienste wie beispielsweise eine Dialogkontrolle zur Verfügung. Ein weiterer wichtiger Dienst ist die Synchronisation. Dafür werden in den Datenstrom von Zeit zu Zeit Marken eingefügt, um die Daten bei einer eventuellen Übertragungsstörung nicht von Anfang an, sondern nur von der letzten Marke an noch einmal zu übertragen.

Die **Darstellungsschicht** 6 (“Presentation Layer”) ist für eine einheitliche Informationsdarstellung, die Datenformate und deren Interpretation zuständig. Im Gegensatz zu den vorherigen Schichten geht hier also erstmals die Semantik und Syntax der übertragenen Nachricht ein. Die Darstellungsschicht konvertiert die vom sendenden Rechner benutzte Datendarstellung wie beispielsweise eine Zeichenkette im ASCII<sup>3</sup> Format in eine Standard Netzwerk-Darstellung, und konvertiert sie im Zielrechner in die dort verwendete Darstellung, beispielsweise Unicode.

Die oberste Schicht 7, die **Anwendungsschicht** (“Application Layer”) unterstützt direkt Anwendungen und den Verbindungsaufbau zum Kommunikationspartner und enthält Protokolle für viele gebräuchliche Funktionen, wie den Dateitransfer mittels des “File Transfer Protocols” (FTP), den Email-Transfer mit dem “Simple Mail Transfer Protocol” (SMTP), virtuelle Terminals und viele mehr.

Das OSI Modell ist ein rein theoretisches Modell, das sich in der Praxis nicht durchsetzen konnte. Es ist jedoch sehr gut geeignet zum Analysieren praxisrelevanter Modelle wie das im nächsten Abschnitt vorgestellte TCP/IP Modell.

## TCP/IP Modell und UDP

Im ARPANET und heutigen Internet wird das TCP/IP Modell benutzt, das aus lediglich vier Schichten besteht. Die Physikalische und die Sicherungsschicht des OSI Modells sind im TCP/IP Modell nicht näher spezifiziert, sondern nur allgemein in einer **Rechner-zu-Netzwerk Schicht** (“Host-to-Network Layer”) zusammengefaßt, siehe Abbildung 2.5.

Die Anforderungen des ARPANET führten zur Definition des Internet Protokolls (**IP**) in der **Internet Schicht**. Die Pakete oder Einheiten dieser Schicht werden in Anlehnung an den Telegrammdienst der Briefpost auch als **Datagramme** bezeichnet. Die IP Schicht ent-

---

<sup>3</sup>American Standard Code for Information Interchange

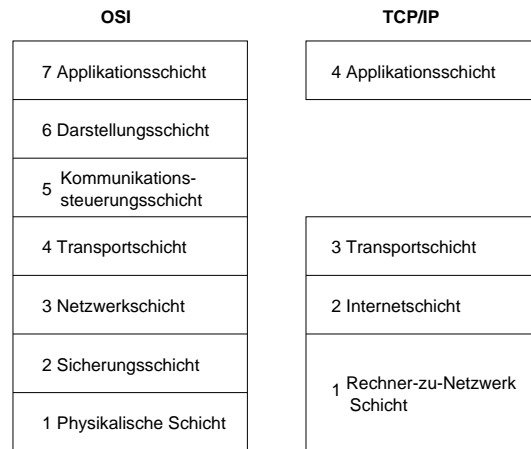


Abbildung 2.5: Gegenüberstellung der Schichten des OSI und des TCP/IP Modells

spricht in etwa der Netzwerkschicht des OSI Modells und ist hauptsächlich für das Routing der Datagramme zuständig. IP arbeitet paketvermittelt, verbindungslos und ungesichert, das heißt, es gibt keine Ende-zu-Ende-Fehlerkontrolle bei unterwegs veränderten oder verlorengegangenen Datagrammen. Der Header der IP Schicht ist 20 Byte groß und enthält unter anderem die IP Adressen des Senders und Empfängers und die Lebensdauer des Datagramms ("Time To Live", TTL), siehe Abbildung 2.6. IP Pakete können insgesamt 576 bis maximal 65.536 Byte (64 kByte) groß sein.

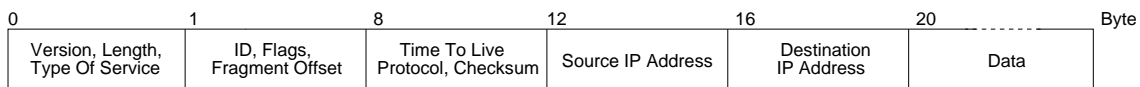


Abbildung 2.6: Aufbau eines IP Pakets

Die **Transportschicht** des TCP/IP Modells entspricht in etwa der des OSI Modells. Es gibt zwei Ende-zu-Ende-Protokolle: Das "Transmission Control Protocol" (**TCP**) ist ein zuverlässiges verbindungsorientiertes Protokoll, das dafür zuständig ist einen Datenstrom fehlerfrei vom Sender zum Empfänger zu übertragen, was zum Beispiel beim Datentransfer mittels FTP der Fall ist. Beim Verbindungsaufbau findet ein Drei-Wege-Handshake statt: Der Sender schickt eine Verbindungsanfrage an den Empfänger, der eine Bestätigung zurückschickt. Daraufhin sendet der Sender wiederum eine Bestätigung der Bestätigung an den Sender. Während der Verbindung werden Pakete, deren Empfang nicht bestätigt wird, noch einmal verschickt. Zu TCP gehört außerdem eine Flußkontrolle, die zu schnelle Sender drosseln kann. TCP besitzt einen mindestens 20 Byte großen Header, auf den hier aber nicht näher eingegangen werden soll.

Da die Zuverlässigkeit von TCP aufwendig ist und Zeit beansprucht, existiert ein alternatives Protokoll, das verbindungslose, ungesicherte "User Datagram Protocol" (**UDP**),

das bei Applikationen angewendet wird, bei denen es hauptsächlich auf eine schnelle Übertragung der Pakete ankommt, während kleinere Übertragungsfehler und Paketverluste toleriert werden, wie es bei Audio- und Video-daten in verteilten Konferenzen der Fall ist. UDP besitzt einen lediglich 8 Byte großen Header, der einen Sender- und Empfängerport, die Länge und eine Prüfsumme zur Fehlerkorrektur enthält, siehe Abbildung 2.7.

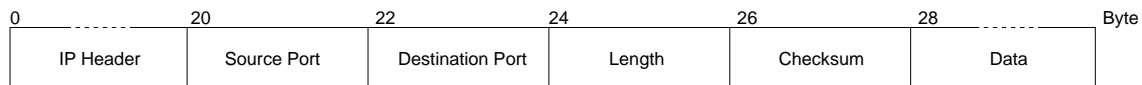


Abbildung 2.7: Aufbau eines UDP/IP Pakets

Das TCP/IP Modell verfügt über keine Kommunikationssteuerungs- und Darstellungsschicht, die sich in der Praxis aber auch als unnötig für die meisten Applikationen herausgestellt haben. Oberhalb der Transportschicht befindet sich beim TCP/IP Modell deshalb gleich die **Applikationsschicht**, die Protokolle wie das oben erwähnte FTP, SMTP, und viele mehr enthält. In den folgenden Abschnitten werden die bei Video- und Datenkonferenzen im Internet eingesetzten Protokolle der Transport- und Applikationsschicht vorgestellt.

### Echtzeitprotokoll RTP

Das Echtzeit-Transportprotokoll **RTP** ("Real Time Transport Protocol") wurde von der IETF<sup>4</sup> Audio/Video-Arbeitsgruppe entwickelt und im Januar 1996 als Internet Standard akzeptiert [62]. Hauptaufgabe von RTP ist das Unterteilen eines kontinuierlichen Datenstroms, wie er bei Live-Audio- und Videoübertragungen auftritt, in einzelne Pakete ("Framing"), und das Versehen der Pakete mit einem Zeitstempel ("Timing") und dem Format des Datenstroms ("Typing"). RTP stellt Ende-zu-Ende Netzwerkfunktionen wie die zeitliche Rekonstruktion, das Entdecken von Paketverlust und in gewissem Rahmen Sicherheit, sowohl über Unicast als auch Multicast zur Verfügung, beinhaltet allerdings keine Ressourcen-Reservierung und gibt keine Garantie für eine QoS. Die Bezeichnung "Transportprotokoll" ist etwas irreführend, da RTP Algorithmen nicht als eigene Schicht implementiert, sondern direkt in die jeweilige Anwendung integriert werden. Obwohl RTP relativ neu ist, gibt es bereits zahlreiche Realisierungen, beispielsweise in den Mbone Tools *vic* und *vat*, siehe Kapitel 3.7. In Zukunft ist auch eine Integration von RTP in Webbrowser geplant [52].

Zu RTP gehört ein Kontrollprotokoll ("Real Time Transport Control Protocol", **RTCP**) zur Unterstützung von Live-Konferenzen mit beliebiger Teilnehmerzahl. Etwa alle 20 RTP-Pakete wird ein RTCP-Paket generiert, in dem zusätzliche Informationen über den Sender wie seine Email-Adresse, die Anzahl der bei ihm eingegangenen RTCP-Pakete enthalten sind, was unter anderem Rückschlüsse auf die Qualität der Datenübertragung ziehen

<sup>4</sup>Internet Engineering Task Force

läßt. RTP-Pakete werden dabei über einen geradzahigen Port gesendet, die zugehörigen RTCP-Pakete über den darüberliegenden ungeraden. RTP und RTCP sind unabhängig vom darunterliegenden Netzwerk- und Transportprotokoll, obwohl es ursprünglich hauptsächlich für UDP/IP gedacht war.

Sicherheitsmechanismen zur Garantie von Authentizität und Integrität der Daten sind in RTP nicht enthalten, da ein Verschlüsselungsmanagement benötigt werden würde, das in Zukunft von den darunterliegenden Schichten zur Verfügung gestellt wird. Vertraulichkeit, also die Garantie, daß nur der richtige Empfänger die Pakete entschlüsseln kann, ist allerdings mit dem "Data Encryption Standard" (DES) möglich. Angriffe auf RTP-Übertragungen werden zusätzlich erschwert durch zufällig gewählte Werte, beispielsweise für die Intervalllänge der RTCP Pakete<sup>5</sup>

RTP und RTCP Pakete besitzen spezielle Header. Der RTP Header ist ein Vielfaches von 4 Byte, mindestens aber 12 Byte groß. Die festen Felder des RTP Header sind in Abbildung 2.8 dargestellt. Die ersten vier Byte sind dabei für verschiedene Flags reserviert. Im einzelnen sind dies 2 Bit für die Versionsnummer V, die aktuell 2 beträgt, daran anschließend 1 Bit zur Anzeige, ob zusätzliche Füll-Bytes ("Padding", P) am Ende des Datenteils angehängt wurden, z.B. für bestimmte Verschlüsselungs-Algorithmen, die feste Blockgrößen benötigen. Das folgende Bit gibt an, ob der Header um (genau eine) Header-Erweiterung ("Extension", X), die ein geradzahiges Vielfaches von 4 Byte groß ist, verlängert ist.

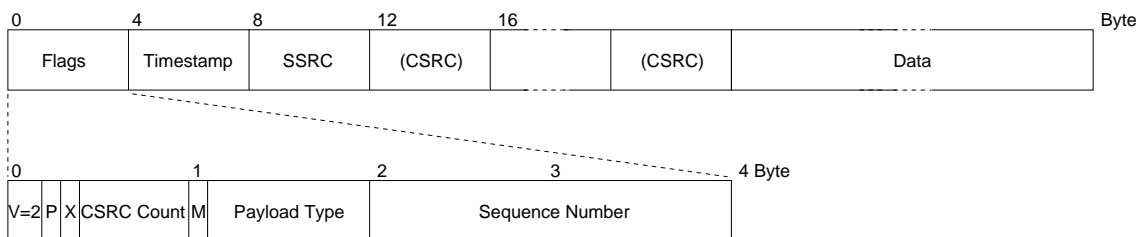


Abbildung 2.8: Aufbau eines RTP Pakets

Als **RTP Session** bezeichnet man eine Menge von Teilnehmern, zusammen mit den RTP Datenströmen, die sie sich gegenseitig senden, beispielsweise Video und Audio. Innerhalb einer RTP Session wird der Sender eines RTP-Datenstroms, die sogenannte Synchronisierungsquelle ("Synchronization Source", SSRC), durch eine eindeutige 32-Bit Zahl identifiziert. Diese Zahl muß zur Identifikation der Teilnehmer und ihrer einzelnen Synchronisierungsquellen innerhalb einer RTP Session eindeutig sein. Die SSRC wird beim Sendebeginn erzeugt, deshalb besteht die größte Gefahr für nichteindeutige SSRCs, wenn viele Teilnehmer gleichzeitig zu senden beginnen. RTP beinhaltet Mechanismen zur Be-

<sup>5</sup>Zufallswerte werden dabei z.B. durch Systemaufrufe wie `getdomainname` und (nicht stille) Mikrofon- oder Videoeingänge generiert.

handlung dieses Falls: Ein Teilnehmer beendet die Session und beginnt kurz darauf noch einmal mit dem Senden, die Pakete werden anhand des unterschiedlichen Identifikators in den Kontrollpaketen unterschieden, oder ein Empfänger verwirft Pakete einer der beiden SSRCs.

Zudem bietet RTP die Möglichkeit, RTP Pakete von einer oder mehreren Quellen ("Contributing Sources", CSRC) zusammenzufassen, beispielsweise alle Datenströme in einem lokalen Netz, die an entfernte Empfänger geschickt werden. Dabei können die Datenströme, die eventuell verschiedene Formate haben, in ein einziges konvertiert oder auch verschlüsselt werden und als ein einziger Datenstrom weitergeschickt werden. Ein solches Zwischensystem bezeichnet man als **RTP Mixer**. Der Datenstrom erhält dann eine neue SSRC Nummer, wobei die SSRC Nummern der zusammengefaßten Datenströme als Liste von CSRC Nummern mitgeschickt wird. Ein 4 Bit großes Feld im Header gibt die Anzahl der in einem RTP Paket enthaltenen zusammengefaßten Quellen (null bis maximal 15) an ("CSRC Count"). Ein anderes Zwischensystem ist der **RTP Translator**, der im Gegensatz zum RTP Mixer aber nur den Datenteil und nicht die Struktur eines RTP Stroms verändert und RTP Pakete mit intaktem SSRC weiterschickt, beispielsweise, wenn nur ein Videostrom verschlüsselt werden soll, oder um Daten durch eine Firewall zu schicken. Mit Hilfe von Tabellen und speziellen Algorithmen wird die Duplizierung von Daten und Kontrollinformationen, sogenannte "Loops", wie sie unter anderem durch RTP Mixer oder Translator entstehen können, verhindert beziehungsweise die Duplikate entfernt.

Mit dem Marker-Bit M lassen sich spezielle Ereignisse wie z.B. Frame-Grenzen markieren. Das darauffolgende 7-Bit Feld beschreibt die Art der Nutzdaten ("Payload Type") und damit die Interpretation der Daten, die auf den Header folgen. Die IANA hat bereits einer Reihe von Audio- und Video-Formaten feste Typnummern zugeteilt [59]. Dabei unterscheidet man Audio-, Video- und kombinierte Audio-Video-Datenströme, bei denen sich Audiodaten und Videodaten abwechseln und gemeinsam abgespielt werden, die Frequenz und die Anzahl der Audiokanäle. Ein eigener Internet Draft definiert verschiedene Profile für die Übertragung von Audio-Video-Daten über RTP [63]. Das dritte und vierte Byte des RTP Headers sind für eine Sequenznummer reserviert, die zu Beginn der Session zufällig gewählt und im Laufe der Session paketweises hochgezählt wird. Durch die Numerierung kann beim Empfänger die ursprüngliche Reihenfolge der Pakete wiederhergestellt bzw. das Fehlen von Paketen entdeckt werden. Die folgenden 4 Byte enthalten eine Zeitmarke ("Timestamp"), die angibt, zu welchem Zeitpunkt das erste Byte der Nutzdaten erzeugt wurde. Sie dient der (relativen) zeitlichen Einordnung der Pakete beim Empfänger. Es kann mehrere RTP Pakete mit der gleichen Zeitmarke geben, beispielsweise, wenn die Pakete zum gleichen Video-Einzelbild gehören. Die erste Zeitmarke wird zu Beginn der Session genau wie die Sequenznummer zufällig erzeugt und mit der für die jeweilige Anwendung nötigen Genauigkeit hochgezählt. In den nächsten 4 Byte ist die SSRC Nummer enthalten, und darauf folgend sovieler 4-Byte-Einträge von CSRC Nummern wie im Feld "CSRC Count" angegeben. Die folgenden Bytes enthalten die Nutzdaten. RTP enthält keine In-

formationen über die Länge des Nutzdatenteils, auch keine Beschränkung einer maximalen Paketgröße, sondern überläßt dies ganz den Protokollen der darunterliegenden Schichten.

Der 4 Byte große Header eines RTCP Pakets ist ähnlich strukturiert wie der RTP Header, siehe Abbildung 2.9. Er enthält die RTP Version (aktuell V=2), ein Füllbit P, einen Empfangs- oder Sendereport- oder Quellenzähler SC (“Source Count”) beziehungsweise RC (“Reception Report Count”), den Nutzdatentyp und die Angabe der Länge des folgenden Textes. Dieser Text ist aus praktischen Gründen ein Vielfaches von 4 Byte groß und enthält je nach Nutzdatentyp verschiedene RTCP Elemente (“Items”).

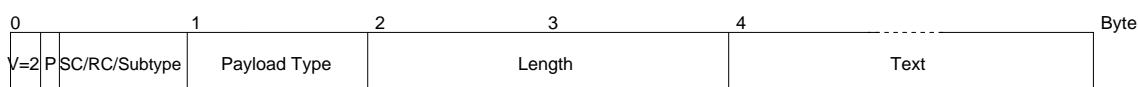


Abbildung 2.9: Aufbau eines RTCP Pakets

So gibt es Sendereportpakete SR mit Nutzdatentyp 200 oder Empfängerreportpakete RR mit dem Nutzdatentyp 201, die im Textteil Zeitmarken, Quellennummern, Anzahl der erwarteten und empfangenen Pakete und weitere Informationen für statistische Zwecke enthalten. So ist sowohl eine im Internet eindeutige Zeitmarke im NTP Format (“Network Time Protocol” [51]) enthalten, als auch eine relativ dazu feste RTP Zeitmarke. NTP Zeitmarken bestehen aus einer 64 Bit großen Fixpunktzahl, die die Sekunden seit 1. Januar 1900, 0 Uhr, angibt, beziehungsweise den ganzzahligen Anteil der Fixpunktzahl in den ersten 32 Bit und die Sekundenbruchteile in den folgenden 32 Bit. Diese “Weltzeit” des Internet kann von sogenannten “Time Servern” abgefragt werden. Durch den Vergleich mit Empfängerreports kann man dann die (zeitliche) Entfernung (“Round-Trip Time”) zu diesen Empfängern berechnen. Außerdem können durch den Vergleich der NTP Zeitmarken der RTCP-Daten eines Audio- und Videodatenstroms von einem Empfänger diese Daten lippensynchron abgespielt werden. Der Textteil von SDES (“Session Description”) Paketen (Nutzdatentyp 202) besteht aus 4 Byte großen SSRC/CSRC Nummern und danach folgend mehreren 4 Byte großen Elementen wie CNAME (“Canonical End-Point Name”), NAME (“User Name”), EMAIL (“Email Address”), PHONE (“Phone Number”), LOC (“Geographic User Location”) oder NOTE (“Notice/Status”). Ein BYE Paket (Nutzdatentyp 203) zeigt an, daß der Sender des Pakets die jeweilige(n) Anwendung(en) beendet hat. Der Textteil enthält in diesem Fall die betreffenden SSRC Nummern und optional den Grund für die Beendigung. Zusätzlich lassen sich anwendungsspezifische APP RTCP Pakete (Nutzdatentyp 204) definieren. Das Feld “Subtype” gibt dabei den Untertyp innerhalb des anwendungsspezifischen RTCP Pakets an.

Der Textteil ist nach dem ISO Standard 10646 UTF-2-kodiert. US-ASCII ist eine Untermenge dieser Kodierung und benötigt daher keine zusätzliche Verschlüsselung. RTCP Pakete bestehen aus mehreren, mindestens aber zwei aneinandergeschlossenen RTCP Ele-



menten. Das Minimum ist ein SDES Paket, das das CNAME Element enthält, zusammen mit einem Empfängerreport RR. Da RTP speziell auf Skalierbarkeit hin entwickelt wurde, wird das Zeitintervall, nach dem ein RTCP Paket gesendet wird, dynamisch geändert, um die zwischen allen Teilnehmern aufgeteilte Bandbreite bei vielen Teilnehmern nicht nur für Kontrollpakete zu verschwenden. Als Richtwert gilt 5-20% aller RTP Pakete, maximal jedoch alle 5 Sekunden.

RTP ist sehr gut für Audio- und Videoübertragung sowohl bei Unicast- als auch Multicast-Konferenzen im Internet geeignet. Für Datenkonferenzen, deren Daten zuverlässig übertragen werden müssen, benötigt man andere Protokolle. Im nächsten Abschnitt wird ein an RTP angelehntes zuverlässiges Protokoll für Datenkonferenzen vorgestellt.

### **Zuverlässige Echtzeitprotokolle SRM und LRMP**

Nicht alle Echtzeitdaten bestehen aus kontinuierlichen Datenströmen wie Live-Audio- und Videoübertragungen. Bei der Verteilung von HTML-Seiten, Zeichnungen oder Text zwischen mehreren Teilnehmern fallen vergleichsweise geringe Datenmengen in kleineren oder größeren Zeitabständen an. Die Datenübertragung ist auch nicht primär zeitkritisch, die Daten müssen aber im Gegensatz zu Audio und Video zuverlässig und in manchen Fällen sogar geordnet bei den Empfängern ankommen. Im Falle von zwei oder sehr wenigen Teilnehmern kann man die Daten einfach über TCP übertragen. Je mehr verteilte Teilnehmer es aber gibt, desto höher ist die übertragene Datenmenge, und desto besser eignen sich spezielle Multicast Protokolle. Eine der ersten Definitionen für ein zuverlässiges Multicast Protokoll war SRM ("Scalable Reliable Multicast Protocol") [19], das teilweise im Mbone Whiteboard `wb` [39] implementiert ist.

SRM beruht auf der Idee, daß alle Teilnehmer gleichberechtigt und unabhängig Daten senden und empfangen können und dabei jeweils lokal eine Kopie der gesamten bereits an die Gruppe gesendeten Daten besitzen. Aufgrund der Durchnummerierung der Daten eines Senders kann jeder Teilnehmer feststellen, ob ihm Daten fehlen und diese explizit anfordern ("Negative Acknowledgement", NACK). Anders als bei TCP werden hier also nicht die empfangenen Daten bestätigt, was einen zu großen zeitlichen und mengenmäßigen Aufwand bedeuten würde, sondern die nicht empfangenen. Alle Teilnehmer, die die angeforderten Daten besitzen, schicken nach einer zufällig gewählten Zeitspanne die gewünschten Daten noch einmal an die Gruppe, oder verwerfen die Sendung, wenn die Daten inzwischen von einem anderen Teilnehmer übertragen wurden.

Die Beteiligung aller Teilnehmer an einer Wiederübertragung ist jedoch ungünstig hinsichtlich der Skalierbarkeit bei großen Gruppen. Basierend auf SRM und RTP/RTCP wurde am INRIA<sup>6</sup> daher **LRMP** ("Light-weight Reliable Multicast Protocol") entwickelt [47][48]. Jeder Teilnehmer ordnet die anderen Teilnehmer nach ihrer Entfernung in verschiedene Grup-

---

<sup>6</sup>Institut National de Recherche En Informatique Et Automatique, Frankreich

pen, sogenannte “Recovery Domains”, ein. Wenn ein Datenverlust entdeckt wird, sind bei LRMP dann nicht mehr alle Teilnehmer beteiligt, sondern zunächst nur diejenigen, die sich in der nächsten Recovery Domain befinden. Nur, wenn die Teilnehmer dieser Gruppe nicht weiterhelfen können, wird die Anfrage an die nächsthöhere Recovery Domain weitergeleitet. LRMP ist daher gut skalierbar, erzeugt wenig Overhead und ist vielseitig einsetzbar bei dennoch vernünftiger Performanz. Zudem ist mit LRMP die Begrenzung der benutzten Bandbreite, der Grad der Zuverlässigkeit und die Möglichkeit einer geordneten Übertragung der Daten einstellbar. LRMP wurde in Java implementiert und kann von INRIA als Java Archivdatei heruntergeladen werden<sup>7</sup>.

Die eigentlichen Daten werden bei LRMP als normale RTP Pakete verschickt. Das Erweiterungs-Bit X im RTP Header wird nicht benutzt und ist daher auf 0 gesetzt. Die RTP-Mixer-Funktionalität ist in LRMP nicht zugelassen, so daß das CSRC Feld ebenfalls immer 0 und somit der RTP Header stets genau 12 Bytes groß ist, siehe Abbildung 2.10. Die gesicherte und geordnete Übertragung der LRMP/RTP Pakete wird durch die Kontrollpakete erreicht. Diese werden bei LRMP als fünf unterschiedliche RTCP Pakete verschickt: ECHO, ECHO ACK, SYNC, NACK und SYNC ERROR. Sie sind wie das anwendungsspezifische RTCP Paket aufgebaut, siehe Abbildung 2.10, mit Nutzdatentyp 206.

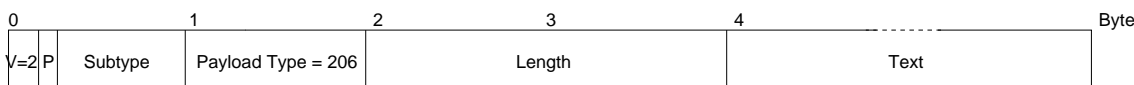


Abbildung 2.10: Aufbau eines LRMP RTCP Pakets

ECHO LRMP/RTCP Pakete (Untertyp 0) dienen zur Bestimmung der Datenlaufzeit von einem neu hinzugekommenen Teilnehmer zu den bereits vorhandenen, die mit einer Bestätigung (ECHO ACK, Nutzdatentyp 1) antworten. SYNC Pakete (Untertyp 2) zeigen bei aktiven Datensendern an, welche Sequenznummer im nächsten RTP Paket benutzt wird. Bei reinen Zuhörern geben sie die Startsequenznummer an. Falls durch die SYNC-Pakete festgestellt wurde, daß eine oder mehrere Sequenznummern nicht angekommen sind, wird ein NACK Paket gesendet, um den Verlust anzuzeigen und den Sender zum nochmaligen Senden eines sogenannten “Repair Packets” aufzufordern, das an alle Teilnehmer geschickt wird. SYNC ERROR Pakete können zusätzlich noch den Grund für einen fehlenden Empfang wie einen Pufferüberlauf bekannt machen, wobei aber keine Bestätigung oder Reaktion erwartet wird.

Die bei LRMP einstellbaren Optionen werden wie folgt realisiert: Die geordnete Übertragung benutzt einen Cache, der eingehende “Repair Packets” speichert und bei Bedarf an die Anwendung weitergibt. Zuverlässigkeit erreicht man durch Senden von NACK

<sup>7</sup>Light-weight Reliable Multicast Protocol, <http://monet.inria.fr/lrmp/>

Paketen. Die benutzte Bandbreite läßt sich durch statistische Auswertung und spezielle Algorithmen an die Netzverhältnisse und die jeweilige Anwendung anpassen.

### **Internet-Multimediasystem MBone**

Mit RTP und LRMP wurden Protokolle vorgestellt, die bei Audio-, Video- und Datenkonferenzanwendungen benutzt werden. In der Applikationsschicht setzen die meisten Anwendungen auf dem **MBone** ("Multicast Backbone") auf, dem digitalen Multimediasystem des Internet, das die bandbreitenschonenden Multicast-Übertragung benutzt. In diesem Abschnitt wird die grundlegende Funktionsweise und der Anwendungsbereich von MBone erläutert.

Das 1985 entwickelte Konzept der Multicast Adressen [11][12] wird seit etwa 1992 mit wachsender Verbreitung im Internet eingesetzt. Mit MBone bezeichnet man eine auf dem Internet aufgesetzte virtuelle Netzstruktur, die aus allen Routern und Rechnern besteht, die mit Multicast Adressen umgehen können. Da noch nicht alle Router multicastfähig sind, müssen sogenannte **Tunnel** zwischen multicastfähigen Routern etabliert werden. Zwischen den Tunnel-Endpunkten werden Multicast-Pakete als normale IP-Pakete mit der IP Adresse des Tunnelendpunkts verpackt und können so von den dazwischenliegenden nicht-multicastfähigen Routern korrekt zum anderen Tunnelendpunkt weitergeleitet werden. Multicastfähige Router sind im Gegensatz zu normalen Routern in der Lage Multicast Adressen zu interpretieren und Pakete gegebenenfalls für verschiedene Ausgänge zu duplizieren. Tunnel sollten insgesamt baumförmig und je Kontinent sternförmig angeordnet sein um den Datenverkehr zu minimieren. Multicast Daten werden dann beispielsweise von einem Rechner in Amerika nur einmal nach Europa übertragen und erst an den Verzweigungspunkten zu den Ästen hin dupliziert, an deren Enden Empfänger eines Multicast-Datenstroms sitzen. In Abbildung 2.11 ist ein Ausschnitt aus dem Tunnelsystem des MBone zu sehen. In nicht allzu ferner Zukunft werden allerdings alle Router multicastfähig sein und damit die Tunnel überflüssig werden.

Auch Rechner, die an Multicast Sitzungen teilnehmen wollen, müssen zur Teilnahme am MBone spezielle Voraussetzungen erfüllen. Sie benötigen die "Multicast Kernel Extensions", eine Betriebssystemerweiterung, die das Senden und Empfangen von Multicast Paketen ermöglicht und Funktionen zum Beitreten und Verlassen von Multicast Gruppen bereitstellt. Multicast Router versenden regelmäßig über das "Internet Group Management Protocol" (IGMP), das integraler Bestandteil des IP Protokolls ist, Anfragen ("Group Membership Queries") im Subnetz, welche Rechner Mitglieder welcher Multicast Gruppen sind. Für die Anfragen wird die spezielle feste Multicast Adresse 224.0.0.1 (`all-systems.mcast.net`) benutzt, die von allen multicastfähigen Rechnern unterstützt wird. Die Rechner antworten dann mit der Angabe der Multicast Gruppen, denen sie angehören oder beitreten wollen ("Group Membership Report"). Diese Informationen werden vom Multicast Router in einer Tabelle eingetragen. Benachbarte Multicast Router benach-

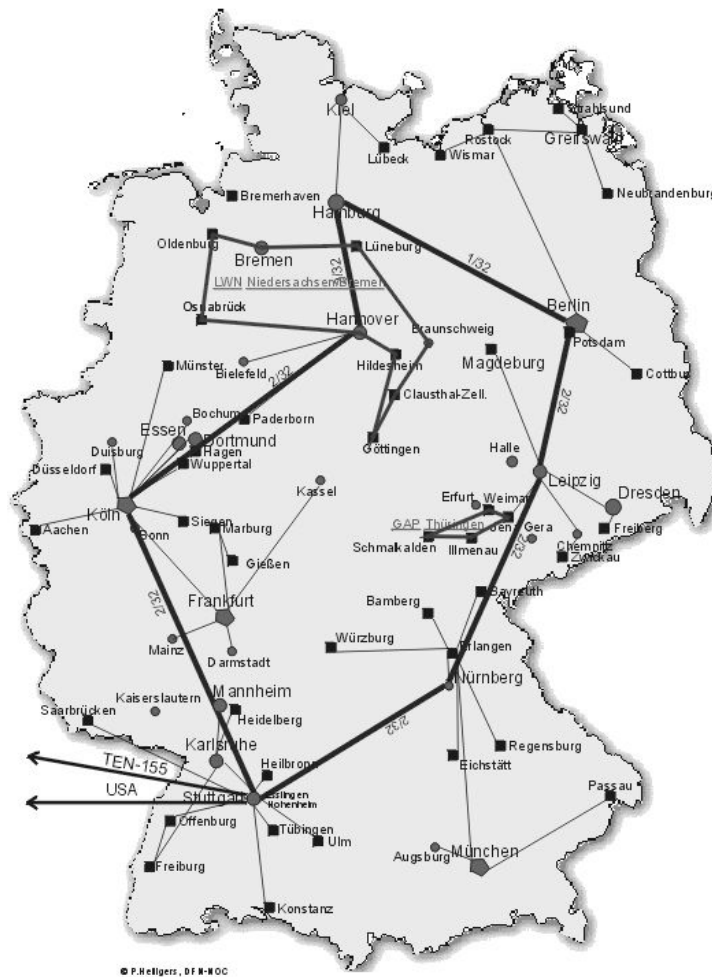


Abbildung 2.11: Ausschnitt aus dem MBone: Das Tunnelnetz des deutschen Wissenschaftsnetzes B-WiN, Quelle: [28]

richten sich gegenseitig regelmäßig über die Zugehörigkeiten der an sie angeschlossenen Rechner über die spezielle Multicast Adresse 224.0.0.4. (`dvmrp.mcast.net`).

Für das Routing von Multicast Paketen können die üblichen Routing Algorithmen mit geringen Adaptionen angewendet werden. Grob unterscheidet man zwei Varianten des Multicast Routing: Einerseits gibt es das Konzept der kürzesten Pfade (“Shortest Path Trees”), bei denen Multicast Daten entlang eines (Tunnel-)Baumes weitergeleitet werden, der dem kürzesten Weg vom Sender zu jedem einzelnen Gruppenmitglied entspricht; dabei muß der Rückweg nicht identisch sein. Der kürzeste Weg wird bestimmt durch die Anzahl an “Hops”, das ist die Anzahl der dazwischenliegenden Router. Realisiert ist dieses Konzept beispielsweise in den Routing Protokollen DVMRP (“Distance Vector Multicast Routing Protocol”), MOSPF (“Multicast Open Shortest Path First”) und DM-PIM (“Dense Mode Protocol Independent Multicast”) [70][43]. Meist wird in Verbindung mit diesen

Protokollen auch "Flood-and-Prune" eingesetzt, was bedeutet, daß Multicast Pakete neu hinzugekommener Sender zunächst in alle Richtungen weitergeleitet werden ("Flooding"), und Router, die keine Abnehmer an ihren Ästen haben, melden sich explizit ab ("Pruning") [70]. Auf der anderen Seite existiert das Konzept der gemeinsam benutzten Pfade ("Shared Trees"), das beispielsweise im CBT Protokoll ("Core Based Trees") realisiert ist, bei dem ein Baum aufgebaut wird, der von allen Teilnehmern gemeinsam benutzt wird. Meistens wird dies durch den kürzesten Pfad zu einem zwischen allen Teilnehmern liegenden Router, dem sogenannten "Core" oder "Rendevous-Point" realisiert. Schließlich gibt es noch hybride Verfahren wie beispielsweise SM-PIM ("Sparse Mode Protocol Independent Multicast") [70].

Um das Internet vor dem kompletten Überfluten mit Multicast Paketen zu schützen, wird zusätzlich das TTL-Feld ("Time To Live") im IP Header zur Reichweitensteuerung ("Scoping") benutzt. Der Sender des Pakets vergibt dabei einen ein Byte großen Wert zwischen Null und 255, der bei jedem durchlaufenen Tunnel um einen festen Wert, das "Gewicht" des Tunnels, dekrementiert wird. Ist die TTL schließlich Null, wird das Paket fallengelassen. Grobe TTL-Richtwerte sind 127 für eine weltweite Verteilung, 64-48 für die Verteilung innerhalb eines Kontinents, 32 innerhalb einer Region, 16-1 innerhalb eines LANs und 0 für das Verbleiben auf dem sendenden Rechner. Dieses Konzept wird auch als Konzept der "Truncated Tunnels" oder "TTL Scoping" bezeichnet.

Da TTL Scoping und Pruning aber schlecht ineinander greifen, kann es zu einem unnötig hohem Datenaufkommen oder auch Datensinken kommen. Die Reichweitensteuerung wurde daher vor kurzem um das sogenannte "Administrative Scoping" [49] ergänzt, bei dem nun auch sich überlappende Zonen möglich sind. Der Adreßraum der administrativ verwalteten Multicast Adressen ist 239.x.x.x. Dabei wird zusätzlich eine lokale und organisationsinterne Reichweitensteuerung unterschieden. Damit Administrative Scoping sinnvoll eingesetzt werden kann, müssen zudem die Multicast Router an den Organisationsgrenzen entsprechend konfiguriert sein.

Sogenannte "Session Directory" Programme zeigen Ankündigungen von zukünftigen und laufenden Multicast Sitzungen an. Veranstalter von Multicast Sitzungen schicken dafür Ankündigungen über ein spezielles Protokoll ("Session Announcement Protocol", **SAP**) [22] in unregelmäßigen Abständen, etwa alle 5-15 Minuten, mit der gleichen TTL-Wert wie die Sitzung selbst auf den Port 9875 an eine bestimmte Multicast Adresse. Bei globalen Sessions ist dies 224.2.127.254 (`sap.mcast.net`), bei Sessions innerhalb eines Administrationsbereichs 239.x.255.255. Beim Aufruf eines Session Directory Programms wird man Mitglied dieser Multicast Gruppe und empfängt die über Multicast versendeten Ankündigungen, die zudem lokal in einem Cache gespeichert werden. Das Format der Ankündigungen ist in einem separaten Protokoll ("Session Description Protocol", **SDP**) [23] festgelegt und enthält Einträge wie den Namen der Sitzung, der Uhrzeit und dem Datum, den Veranstalter, die Angabe der benutzten Medien wie Audio oder Video und deren Protokolle,

die verwendeten Multicast Adressen, UDP Ports und TTL, wobei pro Medium genau eine Adresse und ein UDP Port benötigt werden.

Die Allokation der weltweit eindeutigen globalen beziehungsweise organisationsintern eindeutigen administrativen Multicast Adressen ist in SAP integriert und kann auf verschiedene Arten erfolgen [24]: Zunächst kann aus allen Multicast Adressen zufällig eine gewählt werden. Dies ist zur Zeit wegen der noch relativ geringen Anzahl von Multicast Sessions mit einer geringen Kollisionsgefahr möglich. Um in Zukunft Kollisionen zu vermeiden, kann man erstens die im Cache des Session Directory Programms gespeicherten und damit bekannten Multicast Adressen ausklammern. Zusätzlich kann man verschiedenen Reichweiten bestimmte Multicast Adreßbereiche zuweisen. Im MBone Werkzeug `sdr` wird beispielsweise eine Einteilung in TTL-Bänder mit Werten 2, 16, 32, 48, 64 und 128 ("Informed Partitioned Random Allocation with 6 allocation bands", IPR-6) benutzt; eine neue Adresse wird dann zufällig innerhalb ihres TTL-Bandes vergeben. Die Anzahl der Adressen pro TTL-Band kann auch adaptiv erfolgen, wenn die bereits als benutzt bekannte Adressen in die Berechnung einbezogen werden ("Adaptive IPR", AIPR).

Während die Unterhaltungs- und Elektronikindustrie schon lange plant interaktives Fernsehen und "Video on Demand" auf dem konventionellen Fernseh-Kabelnetz einzuführen und dies erst in Ansätzen verwirklicht hat, hat sich mit MBone unbemerkt im Internet ein Multimediasystem entwickelt, mit dem die genannten Dienste bereits seit längerem möglich sind [70]. Beispielsweise werden im MBone Audio- und Videoaufzeichnungen von Konferenzen live übertragen, Radio- und Fernsehsendungen eingespeist und verschiedene "Video on Demand" Dienste angeboten. Viele Universitäten übertragen Vorlesungen oder Seminare über MBone, an denen Teilnehmer weltweit teilnehmen können. Die Kommunikation erfolgt dabei mit Hilfe von meist frei verfügbaren Audio-, Video- und Textwerkzeugen, die in Kapitel 3.7 genauer vorgestellt werden. MBone und die verwendeten Werkzeuge sind gerade dabei den Experimentierstatus hinter sich zu lassen und zum Standardangebot von Netzanbietern zu werden.

Parallel zu den Entwicklungen im Internet haben die Telefon- und Telekommunikationsgesellschaften eigene Standards für Multimediasdienste auf dem analogen und digitalen Telefonnetz entworfen. Da diese Konzepte und Medienformate zum Teil an MBone-Konzepte angelehnt sind, zum Teil auf Grund ihrer Standardisierung von MBone-Werkzeugen übernommen wurden, werden sie im folgenden Abschnitt näher erläutert.

## 2.2 ITU Konferenz-Standards

Die internationale Telekommunikations-Vereinigung **ITU** ("International Telecom Union") ist ein Zweig der Vereinten Nationen, der für die Entwicklung internationaler Standards zuständig ist. Die Untergruppe ITU-T, die wiederum in fünfzehn Forschungsgruppen un-

terteilt ist, kümmert sich speziell um Telefonstandards. Die dort definierten Konzepte und Medienformate finden zum Teil Verwendung im Mbone, weshalb im folgenden der Standard H.323 für paketbasierte multimediale Kommunikationssysteme und der Standard T.120 für Datenprotokolle von Multimedia-Konferenzen vorgestellt wird.

### Audio-Video-Konferenzstandard H.323

Die ITU hat 1996 in der Serie H ("Verbindungen über Nicht-Telefon-Signale") unter der Kennung H.323 Empfehlungen für "Visuelle Telefon-Systeme und Zubehör für lokale Netzwerke, die keine QoS garantieren" herausgegeben. Die aktuelle Version 2 des Standards H.323 wurde erweitert und trägt den Titel "Paketbasierte multimediale Kommunikationssysteme" [37]. Der Standard ist hardware- und betriebssystemunabhängig und definiert die folgenden vier Hauptkomponenten und ihre Interaktion: Endpunkte, Gatekeeper, Gateways und Kontrolleinheiten ("Multipoint Control Units", MCU), siehe Abbildung 2.12.

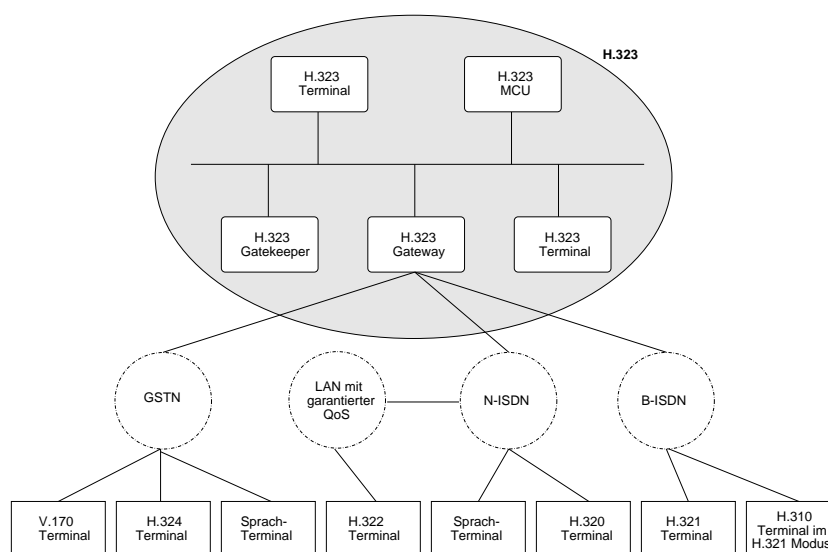


Abbildung 2.12: Architektur des H.323-Standards (nach [13])

**Endpunkte** sind die Klienten eines H.323 Netzwerks, typischerweise Audio/Video- oder sonstige Multimedia-Konferenzwerkzeuge, die sowohl in Rechner integriert oder als eigenständige Geräte wie Videotelefone implementiert sein können. Der Standard H.323 besagt, daß die Endpunkte gewisse Funktionen und Codecs beherrschen müssen. Es wird mindestens Unterstützung für die Echtzeitübertragung über RTP/RTCP, der Sprachkompressionsstandard G.711, das Kontrollprotokoll H.245 für Medien zwischen H.323 Endpunkten, das Verbindungsauf- und -abbauprotokoll Q.931 und ein spezieller RAS Kanal ("Registration Admissions Status", ) Kanal zur Kommunikation mit einem Gatekeeper gefordert. Diese Minimalanforderungen beschreiben also dieselbe Funktionalität wie ein normaler Telefonanruf. Darüberhinaus können H.323 Endpunkte optional Video und zusätzliche Audio

Codecs, Multicast, QoS und Datenkonferenzen über T.120 unterstützen.

**Gatekeeper** sind für die Verwaltung der Calls von Anrufern und Angerufenen und die Bandbreitenallokation im H.323 Netzwerk verantwortlich. Sie können optional eingesetzt werden. Sind jedoch in einem H.323 Netzwerk Gatekeeper vorhanden, so muß davon genau einer benutzt werden. **Gateways** stellen die Verbindung zwischen H.323 Endpunkten und anderen Endpunkten über B-ISDN (H.310), N-ISDN (H.320), LANs mit garantierter QoS (H.322) POTS (H.324) oder einem analogen Telefonsystem usw. her und sind unter anderem für die Audiokompression, Audio-Transkodierung, Rufweiterleitung und das Multiplexing verantwortlich.

Der H.323 Standard definiert die zwei möglichen Konferenzmodi Punkt-zu-Punkt bei zwei Teilnehmern oder Multipoint bei mehr als zwei Teilnehmern. Im Multipoint-Betrieb wird pro H.323 Netz eine Hardware- oder Software-MCU benötigt, die als eine Art Server der H.323 Endpunkte fungiert. Die MCU mischt alle Audio-Ein- und -ausgänge zusammen, schaltet das Video um (beispielsweise zu dem Teilnehmer, der am lautesten spricht) und ist für den Auf- und Abbau der Konferenz verantwortlich.

Der H.323 Standard wird von allen ISDN-Konferenzwerkzeugen implementiert, beispielsweise den VTEL und PictureTel Produkten (siehe Kapitel 3). Einzelne im Rahmen von H.323 verwendete Verfahren wie die Audio- und Videokompression nach H.261 und G.711 beherrschen jedoch auch viele internet-basierte Werkzeuge wie beispielsweise die Mbone Tools oder Sun ShowMeTV (siehe ebenfalls Kapitel 3).

### **Daten-Konferenzstandard T.120**

Die von der ITU in der Serie T ( "Terminal characteristics and higher layer protocols for telematic" ) herausgegebenen Empfehlungen enthalten unter anderem den Standard T.120 ( "Data protocols for multimedia conferencing" ) [38], der eine Reihe von Kommunikations- und Anwendungsprotokollen für Multipoint-Datenkonferenzen beschreibt. Bei solchen kollaborativen Konferenzen kann es sich sowohl um Whiteboard-Konferenzen, als auch Multi-User-Anwendungen oder Multi-Player-Spiele handeln. Die auf dem Gebiet der Datenkonferenzen führenden internationalen Firmen wie Cisco Systems, PictureTel und Intel, aber auch die Deutsche Telekom, haben sich bereit erklärt T.120-konforme Anwendungen zu implementieren, so daß sich daneben wohl keine anderen Standards werden halten können.

T.120 bietet eine elegante Abstraktion für Multipoint-Datenkonferenzen, unterstützt Multicast und Unicast - beides auch gleichzeitig in einer Sitzung - ist netzwerk-, topologie- und plattformunabhängig und skalierbar in Bezug auf die Rechnerleistung vom einfachen PC bis hin zur Mehrprozessor-Maschine. Die T.120 Architektur besteht aus mehreren Schichten, siehe Abbildung 2.13, von denen die unteren Schichten (T.122, T.123, T.124, T.125) anwendungsunabhängige Mechanismen für die Multipoint-Datenkommunikation zur Verfü-



gung stellen, die oberen Schichten (T.126, T.127) dagegen Protokolle für spezielle Konferenzanwendungen wie Whiteboard oder Datentransfer darstellen. T.120-konforme Anwendungen, die nur die unteren Schichten unterstützen, nennt man netzwerk-interoperabel. Bei solchen Datenkonferenzen müssen alle Teilnehmer genau das gleiche Produkt benutzen um miteinander kommunizieren zu können. Anwendungen, die auch die oberen Schichten implementieren und somit auf der Anwendungsschicht interoperabel sind, können mit beliebigen anderen Produkten kommunizieren. Zusätzlich sind nicht-standardisierte Protokolle in einer Implementation erlaubt.

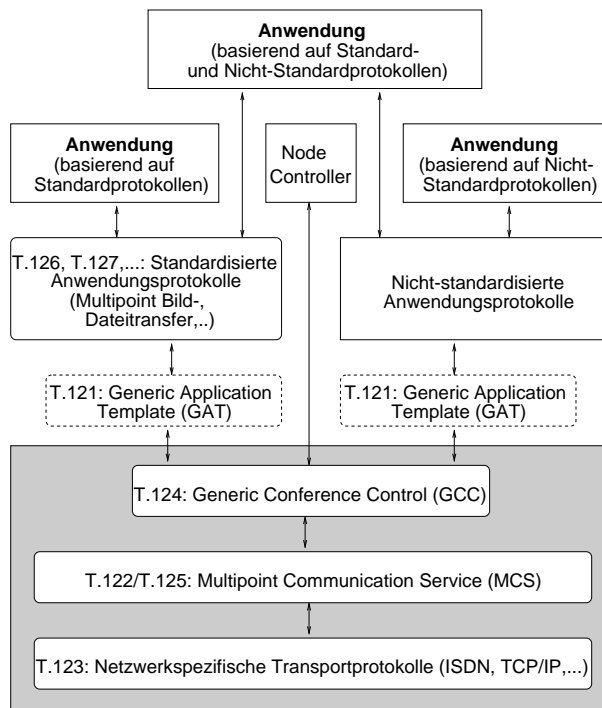


Abbildung 2.13: Die Schichten des T.120-Standards (nach [14])

Im einzelnen spezifiziert T.123 Transportprofile für folgende zuverlässige Transportprotokolle: das öffentliche Telefonsystem ("Public Switched Telephone Networks", PSTN, beziehungsweise "Plain Old Telephone System", POTS), das digitale Telefonsystem ("Integrated Services Digital Networks", ISDN), die leitungs- und paketvermittelten Netze ("Packet Switched Digital Networks", PSDN, "Circuit Switched Digital Networks", CSDN), TCP/IP und Novell Netware IPX. T.122 definiert die in Anwendungen vorhandenen Dienste ("Multipoint Communications Services", MCS), die das Design beliebiger Multipoint-Anwendungen erleichtern, und T.125 das zugehörige Datenprotokoll. Eine Konferenz kann aus mehreren MCS Knoten bestehen, beispielsweise Teilanwendungen oder einzelnen Kanälen für die Daten-, Bild- und Kontrolldatenübertragung, die unterschiedliche Priorität besitzen können. MCS Knoten sind mit einer höheren Kontrolleinheit ("Multipoint Control Unit", MCU) verbunden, von denen mehrere wiederum mit einer noch

höherestehenden MCU verbunden sind. Eine solche Baumstruktur kann jedes beliebige Anwendungsdesign beschreiben. Wenn eine MCU die Sitzung verläßt, werden alle sich im Baum darunter befindenden MCS beendet. Über MCS können beliebig große Datenmengen transferiert werden. Zusätzlich ist der Einsatz von Tokens möglich, beispielsweise um nur dem Leiter einer Konferenz das Senden von Daten zu erlauben, oder um anzuzeigen, daß alle Teilnehmer eine bestimmte Aktion, beispielsweise das Beantworten einer Frage, ausgeführt haben. Token können daher genommen, weitergegeben, gesperrt, freigegeben und angefordert werden und daneben exklusiv oder nicht-exklusiv sein.

Die in T.124 beschriebene Konferenzkontrolle ( "Generic Conference Control" , GCC) ist für die Initiierung und Verwaltung elektronischer Konferenzen zuständig. Die GCC erhält oder fragt Informationen über Konferenzen ab und speichert sie in einer Datenbank. Anwendungen benutzen Mechanismen der GCC um Konferenzen zu erzeugen, ihnen beizutreten, sie zu verlassen oder um andere Teilnehmer einzuladen. Die Anwendungen melden sich mit ihrem eindeutigen Anwendungs-Schlüssel an um herauszufinden, zu welchen anderen MCUs sie kompatibel sind. GCC ermöglicht nämlich die Kommunikation zwischen Anwendungen verschiedener Versionsnummer und Hersteller, indem kompatible Befehls-Untermengen bestimmt werden. GCC ist außerdem für die Sicherheit von Datenkonferenzen zuständig, beispielsweise indem Paßwortschutz zur Verfügung gestellt wird. Darüberhinaus unterstützt GCC das Konzept der Konferenzleitung durch einen Sitzungsleiter. T.124 wurde 1998 um zusätzliche Funktionen zur besseren Skalierbarkeit erweitert, beispielsweise um eine Liste der teilnehmenden MCS Knoten.

T.121 beschreibt das "Generic Application Template" (GAT), eine Vorgabe zum Ressourcen-Management, die für standardisierte Anwendungsprotokolle bindend ist. GAT stellt ein konsistentes Modell zur Verwaltung der von der Anwendung benötigten T.120 Ressourcen zur Verfügung, wobei der Entwickler der Anwendung spezifische Funktionalität hinzufügen kann. GAT verwaltet Kanäle, Tokens und Funktionen auf der Anwendungsseite.

T.126 ("Still Image Exchange and Annotation") ist ein Protokoll der oberen Schicht zum Austausch von nichtbewegten Bildern und Anmerkungen, wie sie beispielsweise bei einem verteilten Whiteboard auftreten. Es ist für die Übertragung von der Bilder, auch in komprimierter Form, und einfachen Anmerkungselementen wie Linien, Text oder Rechtecken oder einem Cursor zuständig, wobei die verschiedenen Objekte in mehreren virtuellen Ebenen übereinanderliegen, siehe Abbildung 2.14. T.126 stellt allerdings keine explizite Anwendungs-Schnittstelle zur Verfügung. Ähnlich spezifiziert T.127 ("Multipoint Binary File Transfer") die Übertragung von Dateien zwischen mehreren Konferenzteilnehmern, wobei vor der Übertragung das Komprimieren der Daten möglich ist und zusätzlich Prioritätsstufen angegeben werden können. Die Daten werden dabei vom Sender in der MCU-Baumstruktur so weit hochgereicht wie nötig und dann entlang der absteigenden Äste bis zum Empfänger.

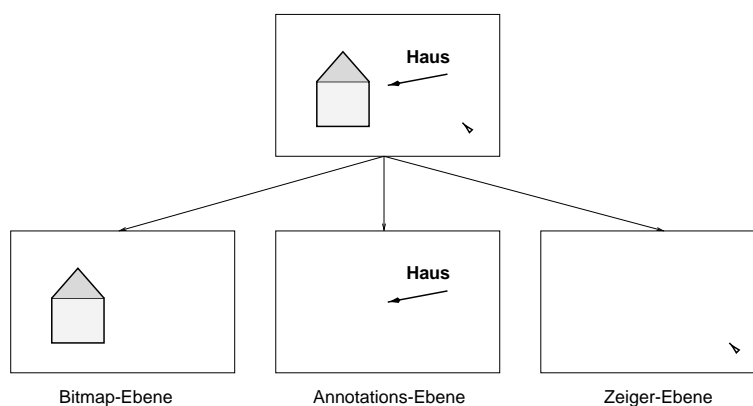


Abbildung 2.14: Die virtuellen Ebenen des T.126-Standards (nach [14])

Der T.120 Standard wird von vielen ISDN-Konferenz-Produkten, beispielsweise den Datenkonferenzwerkzeugen von Intel ProShare und Microsoft NetMeeting, unterstützt (siehe Kapitel 3). Internetbasierte Werkzeuge wie die MBone Tools benutzen meist eigene Protokolle oder führen Dateitransfers mit Hilfe gängiger Unix-Programme wie FTP durch.

## 2.3 Multimedia-Dateiformate

In diesem Abschnitt sollen nun die eigentlichen Daten, die bei Konferenzanwendungen in der Applikationsschicht erzeugt und übertragen werden, genauer beschrieben werden. Dabei unterscheidet man zwischen Audio-, Video- und Text- und sonstigen Daten, deren beliebige Kombination man auch als Multimediadaten bezeichnet. Die Daten können in verschiedenen Kompressionen und Kodierungen vorliegen, von denen die gebräuchlichsten vorgestellt werden.

### Multimedia, Kodierung und Kompression

Multimedia ist aus den lateinischen Wörtern “multi” (viele) und “media” (Medien, Kommunikationsmittel) zusammengesetzt und bedeutet somit wörtlich “viele Kommunikationsmittel”. In der Informatik versteht man unter Medien beispielsweise die Informationsträger Text, Bild, Audio und Video in ihrer digitalisierten Form [66]. Zeitabhängige Medien wie Audio und Video bezeichnet man als **zusammenhängend**, nicht zeitabhängige Medien wie Text und Bild als **diskret**. Als “echte” Multimedia-Anwendungen gelten dann Programme, die nicht rein quantitativ “viele Kommunikationsmittel” einsetzen, sondern sowohl zusammenhängende als auch diskrete Medien [67]; beispielsweise eine Videokonferenz-Anwendung mit einem integrierten Whiteboard. Solche Multimedia-Anwendungen dienen der technisch unterstützten Kommunikation zwischen Menschen, wenn die biologisch vorgegebenen Möglichkeiten nicht ausreichen oder bewußt verbessert werden sollen. So können interaktive Anwendungen im Internet die räumliche Distanz überbrücken oder nicht-interaktive Anwendungen zu beliebigen Zeitpunkten durchführen. Dabei ist sowohl

eine Individual-, als auch Massenkommunikation möglich.

Mit Digitalisierung bezeichnet man die Umwandlung von analogen und sich kontinuierlich verändernden Werten, wie sie meist vorliegen, in diskrete Werte. Der analoge Wert wird dafür in regelmäßigen Zeitpunkten abgetastet und der jeweils abgelesene Wert als diskreter Wert genommen; dieser Vorgang wird als **Sampling** oder **Pulse Code Modulation** (PCM) bezeichnet. Die Häufigkeit der Abtastung bezeichnet man als **Samplingrate** und gibt sie in Hertz (Hz), also der Anzahl der Samples pro Sekunde, an. Die Genauigkeit beziehungsweise die Anzahl der möglichen diskreten Werte bezeichnet man als **Quantisierung** und gibt sie in Bit an. Je höher also die Samplingrate und die Quantisierung, desto besser ist die Qualität der Digitalisierung, aber auch die anfallende Datenmenge. Die in der Informatik verwendeten Medien unterscheiden sich in Komplexität und Datenmenge enorm, siehe Abbildung 2.15. Sollen die Mediendaten dann noch in Echtzeit übertragen werden, erfordern insbesondere bei Audio- und Videodaten sehr hohe Übertragungsraten (**Bitraten**), während Texte, Bilder und gesprochene Sprache relativ unkritisch sind.

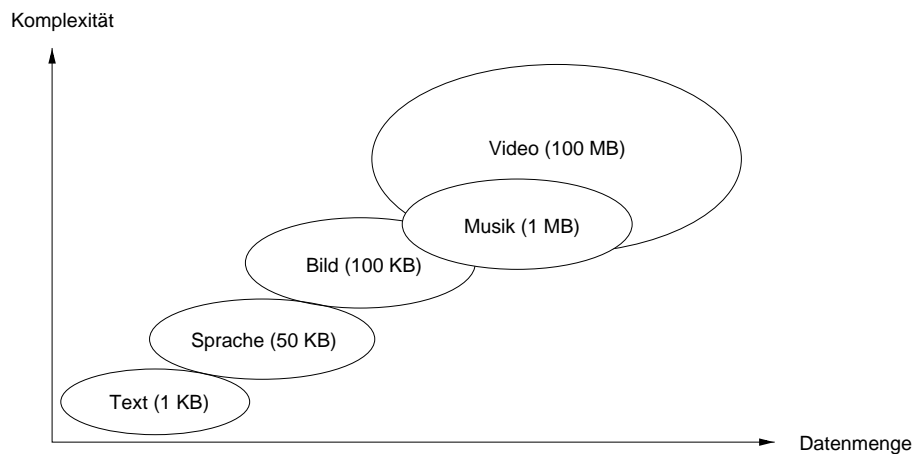


Abbildung 2.15: Komplexität und Datenmenge verschiedener Medientypen, nach [66]. Die Mengenangaben sollen nur grobe Richtwerte darstellen.

Als **Kodierung** oder Datenformat eines Mediums bezeichnet man den Wert der Samplingrate zusammen mit der Quantisierung und eventuell die Kompressionsmethode. Die **Kompression** von Multimediadaten, speziell von Audio und Video, ist auf Grund der großen Datenmengen und trotz der heute im allgemeinen zur Verfügung stehenden hohen Bandbreite in Rechnernetzen unerlässlich. Dabei wird der Datenstrom vor dem Versenden per Hardware oder Software komprimiert und beim Empfänger wieder dekomprimiert. Ist der Aufwand für Kompression und Dekompression gleich hoch, spricht man von synchronen Kompressionsmethoden; asynchrone Verfahren, bei denen meist die Kompression rechenintensiver als die Dekompression ist, werden bei Daten verwendet, die nicht in Echtzeit übertragen werden, sondern beispielsweise auf CD-ROM vorliegen. Man unterscheidet außerdem verlustfreie und verlustbehaftete Kompression [67]. Das verlustfreie "Entropy

Coding" betrachtet die zu komprimierenden Daten als binären Datenstrom, bei dem die Daten mit Varianten der bekannten Huffman- oder LZW-Kodierung komprimiert werden. Bei der Huffman-Kodierung werden im Bitstrom Zeichen, die direkt hintereinander mehrfach auftreten, durch ein einziges Zeichen zusammen mit einem Kodewort, das die Anzahl der auftretenden Zeichen angibt, repräsentiert. Die LZW-Kodierung kann auch komplexere Redundanzen wie mehrfach auftretende Zeichenfolgen suchen. "Source Coding" betrachtet dagegen die Semantik der Daten und kann verlustbehaftet sein. Die verlustbehaftete Kompression benutzt Transformationen, Vorhersagen oder mehrschichtigen Kodierungen und nutzt dabei Eigenheiten der menschlichen Wahrnehmung aus, indem Daten, bei denen die menschlichen Sinne nicht so empfindlich sind, von vornherein entfernt oder besonders stark komprimiert werden. So ist das menschliche Auge bei Helligkeitsunterschieden sensibler als bei Farbunterschieden, und das menschliche Ohr kann Unterschiede innerhalb sehr hoher oder sehr tiefer Frequenzen nicht genau wahrnehmen.

Es existieren auch hybride Verfahren ("Hybrid Coding"), die sowohl Entropy Coding als auch Source Coding implementieren. In den folgenden Abschnitten werden die gängigsten Multimedia-Kodierungen und ihre speziellen Kompressionsverfahren vorgestellt.

## Text

Text ist die einfachste Form von digitaler Information, da er naturgemäß aus einer begrenzten Menge von Symbolen zusammengesetzt ist, die sich leicht digitalisieren lassen, indem jedes Symbol mit einer bestimmten Anzahl von Bits dargestellt wird. Der **ASCII** Code, der für das Kodieren von englischem Text entworfen wurde, und der auch unter der Bezeichnung "International Alphabet No. 5" bekannt ist, benutzt beispielsweise 7 Bit. Die Größe eines digitalisierten Textdokuments ergibt sich dann aus der Anzahl der Zeichen des Textes, multipliziert mit der Anzahl an Bits für die Kodierung. Text läßt sich auf Grund vieler wiederholt vorkommender Zeichen sehr gut mit der Huffman-Kodierung komprimieren. Neben dem ASCII Standard, den alle Rechner verstehen, existieren unzählige Standards für viele weitere Sprachen. Zur Zeit wird von einem internationalen Konsortium am 16-Bit Unicode Standard gearbeitet, der universell für alle Sprachen eingesetzt werden kann [72].

Für Text mit Formatierungen, also beispielsweise den physischen Textauszeichnungen "fett" oder "kursiv", existieren ebenfalls einige Standards oder Quasi-Standards, beispielsweise das "Rich Text Format" (**RTF**) oder applikationsspezifische Formate wie Adobe Frame-Maker oder Microsoft Word. Diese setzen allerdings voraus, daß Sender und Empfänger die gleiche Applikation benutzen. Im "World Wide Web" (WWW) wird seit seiner Einführung 1989 die "Hypertext Markup Language" (**HTML**) [4] für den internationalen Dokumentenaustausch und die Darstellung von formatiertem Text in sogenannten Web-Browsern benutzt. HTML ist eine vom World Wide Web Consortium entwickelte spezielle Anwendung der ISO-Norm 8779 ("Structured Generalized Markup Language", SGML) mit spezieller Ausrichtung auf Hypertext-Funktionalität, also dem virtuellen Verknüpfen ("Linken") meh-

rerer Dokumente. HTML ist eine plattformunabhängige Dokumentbeschreibungssprache, die die Struktur des Dokuments, also Elemente wie Kapitel, Absätze und Grafiken, aber auch physikalische Textauszeichnungen in einer abstrakten Form beschreibt. Dabei werden spezielle Auszeichnungen (‘‘Markups’’) verwendet, beispielsweise `<I>Text</I>` für kursiven Text. HTML Dateien bestehen aus ASCII Text. Web-Browser können diese Dateien interpretieren und die Auszeichnungsbefehle in optisch ansprechender Form am Bildschirm darstellen.<sup>8</sup>

Da HTML Dateien von jedem Browser anders interpretiert werden können und auch von den persönlichen Einstellungen des Benutzers abhängen, hat sich für den Austausch von Dokumenten, die genau so aussehen sollen, wie der Autor sie gestaltet hat, im Internet Adobe PostScript (**PS**) und die Weiterentwicklung **PDF** (‘‘Portable Document Format’’) durchgesetzt. PostScript wurde 1985 entwickelt und ist eine plattformunabhängige physische Beschreibungssprache, die das genaue Positionieren von Text und Bildern ermöglicht [2] und von vielen Druckern und Betrachterprogrammen interpretiert werden kann. PostScript Dateien bestehen aus ASCII Text, der spezielle Anweisungen und Kommentare wie die Seitengröße, die Schriftart oder die Positionierungsdaten enthält. PDF ist eine Erweiterung von PostScript um die Unterstützung verschiedener Kompressionsverfahren, sowohl für Text wie auch für Bilder, und das Einbinden von Hyperlinks [1].

## Audio

Viel komplexer als Text ist das Medium Audio. Zunächst einige Vorbemerkungen zur menschlichen Wahrnehmung: Das Ohr nimmt Schwingungen in der Luft und den sich damit verändernden Luftdruck als Geräusch beziehungsweise Ton wahr. Die Schwingungen können als sich periodisch wiederholende Wellenform (‘‘Schallwelle’’) dargestellt werden, siehe Abbildung 2.16. Die Anzahl der Perioden pro Sekunde wird als **Frequenz** der Schwingung bezeichnet und in Hertz angegeben - je höher die Frequenz, umso höher ist der Ton. Die **Amplitude** der Schwingung, also die Stärke des Luftdrucks, wird subjektiv als Lautstärke empfunden - je größer die Amplitude, umso lauter ist der Ton. Für das Prozessieren von Audiodaten benötigt der Rechner spezielle Hardware. Die Soundkarte ist für die Digitalisierung von eingehenden analogen Signalen, beispielsweise über Mikrofon oder Line-In, und das Umwandeln von digitalen Signalen in Klang, beispielsweise für einen angeschlossenen Lautsprecher, Kopfhörer oder Line-Out, verantwortlich.

Bei der Digitalisierung übliche Werte für Samplingrate und Quantisierung von Schallwellen sind in Tabelle 2.1 angegeben. Die 44.1 kHz bei der Audio-CD für die höchste hörbare Audio-Qualität ergeben sich aus dem Nyquist-Theorem: Für verlustfreie Digitalisierung

---

<sup>8</sup>Flexibler als HTML ist die vom World Wide Web Consortium (W3C, <http://www.w3c.org>) entwickelte XML (‘‘Extensible Markup Language’’), eine Einschränkung von SGML auf die wichtigsten Befehle [20]. XML erlaubt im Gegensatz zu HTML die Definition eigener Auszeichnungen und Dokumententypen und ist somit viel flexibler. Allerdings befindet sich XML noch in der Entwicklung und es existieren nur einige wenige Web-Browser, die XML interpretieren können, so daß hier nicht weiter darauf eingegangen wird. In Zukunft wird XML jedoch eine wichtige Rolle spielen.

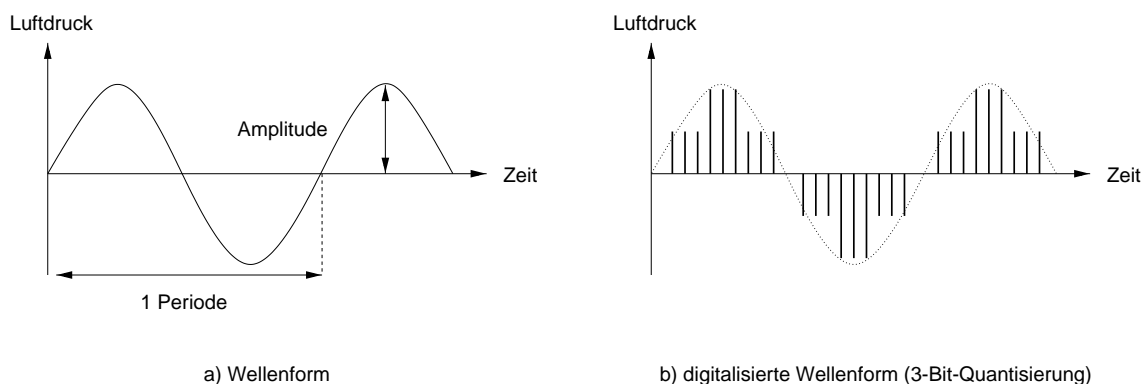


Abbildung 2.16: Darstellung von Schall als Wellenform und digitalisierte Wellenform

Tonqualität	Samplingrate	Quantisierung	Kanäle	Bitrate
DAT	48 kHz	16 Bit	2 (stereo)	etwa 1,5 Mbit/s
Audio CD	44.1 kHz	16 Bit	2 (stereo)	etwa 1,4 Mbit/s
Digitalradio	32 kHz	16 Bit	2 (stereo)	etwa 1 Mbit/s
1/2 Audio CD	22.05 kHz	16 Bit	1 (mono)	etwa 352 Kbit/s
1/4 Audio CD	11.025 kHz	16 Bit	1 (mono)	etwa 176 Kbit/s
Telefon	8 kHz	8 Bit	1 (mono)	etwa 64 Kbit/s

Tabelle 2.1: Gängige Werte für verschiedene Tonqualitäten

muß die Samplingrate mindestens doppelt so hoch sein wie die maximale Grenzfrequenz, im Fall des menschlichen Ohrs  $2 \times (20 \text{ kHz} - 20 \text{ Hz}) \approx 40 \text{ kHz}$ , da es nur Frequenzen zwischen 20 Hz und 20 kHz wahrnehmen kann.

Für Stereoklang verdoppelt sich die Informationsmenge und damit die entsprechende digitale Datenmenge, die dem analogen Klang entspricht. Sie kann sich vervielfachen, wenn zusätzlich bis zu fünf Audio-Kanäle für Surround Sound benutzt werden. Allerdings existieren auch sehr effiziente Kompressionsalgorithmen für Audio. Im folgenden werden die im Internet gebräuchlichsten Audioformate und Kompressionsmethoden vorgestellt.

Bei der Pulse Code Modulation wird bei beliebiger Quantisierung und Samplingrate jeder diskrete Wert einzeln gespeichert. Solche unkomprimierten PCM Daten mit 16 Bit Quantisierung und 44.1 kHz Samplingrate befinden sich beispielsweise auf Audio-CDs. Auch das von NeXt und Sun benutzte "Sun Audio Format" (**AU**) benutzt die PCM-Kodierung. Ursprünglich wurde nur eine Digitalisierung mit 8 Bit bei 8kHz und logarithmischer  $\mu$ -law Kompression unterstützt, inzwischen sind aber auch Kodierungen mit 16 Bit bis 44.1 kHz. Das beim Amiga eingesetzte und von Apple und Silicon Graphics zum "Audio Info File Format" (**AIFF**) weiterentwickelte Audioformat kodiert Daten mit 16 Bit bis 44.1 kHz und unterstützt zusätzlich Kompression (AIFC). Wenn nicht jeder diskrete Wert einzeln gespeichert wird, sondern nur die Differenz zum vorherigen, spricht man von

“Difference/Delta Pulse Code Modulation” (DPCM) und im adaptiven Fall von “Adaptive Delta Pulse Code Modulation” (ADPCM) und benötigt pro Sample bis zu 4 Bit weniger zu speichernde Daten, ohne daß ein Qualitätsverlust eintritt. Das von Microsoft entwickelte Wellenformen-Format **WAV** ist eine abgewandelte Form von AIFF in Little Endian Byte Order. Es unterstützt unter anderem die Kompression mit ADPCM.

Neben den genannten Wellenformen-Formaten existieren auch zeitunabhängige Bitstrom-Formate in verschiedenen verlustbehafteten Kompressionsstufen. Der **MPEG** ISO Standard (auch bekannt als MPG, MPA, MP2, MP3) wurde von der Moving Pictures Expert Group 1988 ursprünglich als kombiniertes Audio-Videoformat entwickelt, es sind aber auch reine Audio- oder Videodatenströme möglich. Eine besonders effektive Komprimierung ist die vom Fraunhofer Institut entwickelte Layer-3-Kodierung, die mit Hilfe psychoakustischer Modelle bei einer Kompression von 12:1 nahezu CD-Qualität erreicht.

Auch die ITU hat einige Empfehlungen speziell für die Sprachkodierung herausgegeben, die heute nicht nur bei der Kommunikation über das Telefonnetz, sondern auch für die Echtzeitkommunikation im Internet verwendet werden. Dazu gehört beispielsweise der Standard G.711 [32], der die PCM-Kodierung für einen 64 kbit/s breiten ISDN-B-Kanal beschreibt, G.722 [33], der die PCM-Kodierung bei 7KHz mit 64 kbit/s definiert, oder G.726 [34] für die ADPCM-Kodierung bei 16, 24, 32 und 40 Bit. Für Mobiltelefone sind die Datenraten jedoch immer noch zu hoch, außerdem wollte man ein in Europa einheitliches System einführen, so daß die **GSM** (“Groupe Special Mobile”)<sup>9</sup>, eine Forschungsgruppe der europäischen Telefongesellschaften, 1989 den GSM-Standard für die zu ISDN kompatible Audioübertragung für Mobiltelefone mit 13 kbit/s entwickelt hat [65]. Die Kompression wird mit Hilfe von Vorhersage-Algorithmen erreicht. Implementationen des GSM-Standards sind zum Teil frei im Internet verfügbar<sup>10</sup> und werden deshalb häufig in Internet-Anwendungen eingesetzt.

Programme zum Abspielen und/oder Aufnehmen von Audiodaten beherrschen meist verschiedene Audioformate, so daß die Kodierung und Qualität flexibel an die Übertragungsgegebenheiten und Möglichkeiten des Empfängers angepaßt werden können.

## Bild

Ein Bild kann, vereinfacht ausgedrückt, durch die Lichtintensität an jedem Punkt der zweidimensionalen Bildfläche dargestellt werden. Bei der Digitalisierung, beispielsweise beim Scannen von Fotos, wird der Wert der Lichtintensität in regelmäßigen Intervallen abgetastet, so daß man eine Matrix von Bildelementen (“Picture Elements”, **Pixel**) erhält [66][67]. Das digitale Bild, eine sogenannte “Bitmap”, siehe Abbildung 2.17 links, wird

<sup>9</sup>Inzwischen werden die Standards auch international eingesetzt, so daß GSM nun für “Global System for Mobile communications” steht.

<sup>10</sup>The GSM 06.10 lossy speech compression library and its applications, <http://kbs.cs.tu-berlin.de/jutta/toast.html>



dann einerseits durch das räumliche Ausmaß spezifiziert, das aus dem Produkt der Pixel in der x- und y-Richtung besteht, beispielsweise 320x240 Pixel, andererseits durch einen Farbwert in Bit pro Pixel, beispielsweise 8 Bit. Die Größe eines Bildes ist dann die Anzahl der Pixel multipliziert mit der sogenannten Farbtiefe, hier im Beispiel also etwa 614 KByte. Größere Bitmaps und Bitmaps mit mehr Farben, beispielsweise im "True Colour" Modus mit 24 oder 32 Bit Farbtiefe, können also sehr groß werden. Im Internet haben sich verschiedene Komprimierungsmethoden für verschiedene Anforderungen etabliert. Alle Formate bestehen aus einem Header, der das Dateiformat enthält, und einem komprimierten Datenteil.

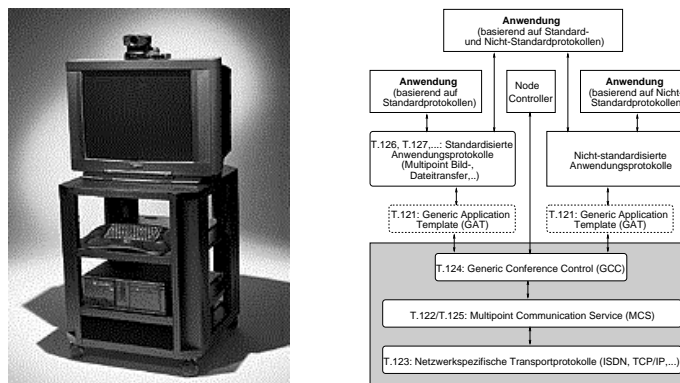


Abbildung 2.17: Digitale Bilder: Bitmap (links) versus Grafik (rechts)

**JPEG** ist ein von der Joint Pictures Expert Group entwickeltes Bildformat, das bis zu 24 Bit Farbtiefe unterstützt und auf Wunsch verlustfrei oder verlustbehaftet komprimiert. Die Komprimierung arbeitet mit Transformationen in den Frequenzbereich und der anschließenden Reduzierung niedriger Frequenzanteile. Sehr stark komprimierte Bilder wirken deshalb unscharf. Das JPEG Format eignet sich aber sehr gut für Fotos und Bilder mit Farbverläufen, da ein Qualitätsverlust von etwa 20% vom menschlichen Auge nicht erkannt wird. Ein Spezialfall ist das "Progressive JPEG" Format, bei dem das Bild nicht zeilenweise gespeichert wird, sondern in verschiedenen Verfeinerungsstufen. Dies ist praktisch bei langsamen Übertragungen im Internet, da man sofort einen groben Eindruck vom Bild erhält, das sich nach und nach verfeinert. JPEG wird deshalb von allen gängigen Web-Browsern und Bildbetrachtern unterstützt.

Das "Graphical Interchange Format" (**GIF**) unterstützt lediglich 8 Bit Farbtiefe, komprimiert aber verlustfrei und kann ähnlich wie Progressive JPEG als "Interlaced GIF" in verschiedenen Verfeinerungsstufen angezeigt werden. Da es mit der LZW-Methode komprimiert, eignet es sich besonders für Strichzeichnungen oder Bilder ohne Farbverläufe, die dann kleiner als bei der JPEG Komprimierung werden. Außerdem können im Unterschied zu JPEG bestimmte Bildbereiche transparent gehalten werden und sogar Anzeigebefehle im Header untergebracht werden. Dadurch sind animierte Bilder ("Animated GIF")

möglich, bei denen in bestimmten Abständen jeweils ein anderer Bildausschnitt eines Bildes angezeigt wird. GIF ist aus diesen Gründen das zweite Format, das ebenfalls von allen Web-Browsern und fast allen Bildbetrachtern unterstützt wird.

Eingescannte Bilder werden meist unkomprimiert im "Tagged Image File Format" (**TIFF**) abgespeichert und sind daher sehr groß. Sie müssen zum Anzeigen in spezielle Bildbetrachterprogramme geladen oder zur Anzeige im WWW nach JPEG oder GIF konvertiert werden. Es ist auch möglich Bilder im PostScript Format zu speichern, was allerdings zu sehr großen Dateien führt, da die Bilder in der vorgegebenen Auflösung Pixel für Pixel ohne Komprimierung gespeichert werden.

Im Gegensatz zu Bitmap Bildern lassen sich grafische Zeichnungen, die aus Elementen wie Text, Kreisen und Quadraten bestehen, siehe Abbildung 2.17 rechts, platzsparend als textuelle Beschreibung des Bildes speichern. Das Format ist dabei von Anwendung zu Anwendung verschieden. Auch PostScript kann grafische Elemente in Form von Vektoren darstellen und dadurch sehr kleine Dateien aus grafischen Zeichnungen erzeugen.

Digitale Bilder erhält man aber nicht nur beim Scannen oder durch Malprogramme, sie werden auch zur Darstellung von digitalem Video benötigt, wie im nächsten Abschnitt näher erläutert wird.

## Video

Film und Video bestehen aus einer schnellen Folge von Einzelbildern, die man auch **Frames** nennt. Frames können von Zeichenprogrammen oder von einer Videokamera wie in Abbildung 2.18 oder einem anderen Video-Eingang erzeugt werden, wo sie von der speziellen Video-Hardware oder -Software digitalisiert und meist gleich komprimiert werden. Durch die Trägheit des menschlichen Auges wird bei etwa 25 bis 30 Bildern pro Sekunde ("Frames per Second", **fps**) die Illusion eines kontinuierlich bewegten Bildes erzeugt, Filme mit unter 15 fps können etwas ruckelig wirken. Die Datenrate eines unkomprimierten digitalen Videos berechnet sich aus dem Produkt der Größe eines unkomprimierten Einzelbildes und der sogenannten **Framerate** in fps; bereits ein kleines Videobild von 120x160 Pixel, 8 Bit Farbtiefe und 15 fps erzeugt eine Datenrate von etwa 2,2 Mbit/s. War die Komprimierung schon bei digitalen Bildern unerlässlich, so gilt dies erst recht für digitales Video. Im folgenden werden die am häufigsten benutzten Formate im Internet vorgestellt.

Die Moving Pictures Expert Group (**MPEG**) entwickelt seit 1988 Standards für die verlustbehaftete Kompression von Audio, Video und kombinierten Audio/Video-Datenströmen ("System Streams"). Das MPEG-1-Verfahren reduziert ein 320x240 großes Video mit bis zu 30 fps auf eine Datenrate von maximal 1,5 MBit/s und wurde damit speziell auf das Abspielen des Videos von einem CD-ROM-Laufwerk zugeschnitten. Die Nachfolge-Standards MPEG-2/4/7 komprimieren Videodaten in Fernsehqualität und unterstützen



Abbildung 2.18: Desktop-Kamera

Szenenintegrationen und Untertitel. Die Besonderheit des MPEG-Formats ist die Verwendung verschiedener Einzelbild-Typen. I-Frames (‘‘Intra-Coded Frames’’) enthalten die Information eines kompletten Bildes und werden im JPEG Format komprimiert, wohingegen P-Frames (‘‘Predictive-Coded Frames’’) lediglich aus den Pixel-Blöcken bestehen, die sich zum vorhergehenden I-Frame unterscheiden und B-Frames (‘‘Bi-Directionally Predictive-Coded Frames’’) lediglich Pixel-Blöcke enthalten, die sich sowohl zum vorhergehenden als auch nachfolgenden I- oder P-Frame verändern. Mit einer geschickten Frame-Reihenfolge, beispielsweise [IBBPBBPBBPBB], kann digitales Video sehr gut komprimiert werden, weshalb das MPEG-Format häufig bei Internet-Anwendungen eingesetzt wird. Beliebt ist auch das Realvideo-Format (**RM**), das ein ähnliches Kompressionsverfahren wie MPEG benutzt, allerdings meist schlechtere Qualität erzeugt.

Das zunächst ausschließlich für die Apple Plattform entwickelte QuickTime Movie-Format (**MOV**) findet man ebenfalls häufig in Form von kleinen Videoclips oder Animationen auf dem Internet. Auf Grund der Verbreitung wird das Format inzwischen von Playern auf allen Plattformen unterstützt. Das QuickTime-Format legt nur die Verknüpfung der im Datenstrom enthaltenen Audio- und Videodaten fest, die Kodierung und die Kompressionsmethode bestimmen dagegen sogenannte **Codecs** (‘‘Coder/Decoder’’). Bekannte Video-Codecs sind zum Beispiel **MJPEG** (‘‘Motion JPEG’’), Intel Indeo oder Radius Cinepak. Ähnlich wie QuickTime ist das von Microsoft entwickelte **AVI** Format (‘‘Audio Video Interleaved’’) aufgebaut, das ebenfalls zu den häufiger eingesetzten Formaten im Internet gehört.

Auch die ITU hat verschiedene Video-Standards entwickelt, die zwar auf die Übertragung über ISDN zugeschnitten sind, sich aber auch im Internet und besonders in der Echtzeitkommunikation bewährt haben. Der Video-Codec **H.261** wurde 1990 speziell für Datenraten entwickelt, die ein ganzzahliges Vielfaches von 64 kBit/s betragen und damit genau einem ISDN-B-Kanal entsprechen [35]. Der Kodieralgorithmus benutzt wie MPEG Zwischenbilder-Vorhersage, Transformationen und Lauflängenkodierung zur Kompression und sogenannte ‘‘Motion Vectors’’ zur Bewegungskompensation. Die Datenrate des Kodieralgorithmus kann zwischen 40 kBit/s und 2 MBit/s liegen. H.261 unterstützt zwei feste Bildgrößen: CIF (‘‘Common Interchange Format’’) mit 352x288 Pixel, und QCIF (Quarter

CIF) mit 176x144 Pixel. Die Weiterentwicklung von H.261 wurde 1995 unter dem Namen H.263 definiert [36]. H.263 unterstützt mehr Datenraten und Bildgrößen (SQCIF, QCIF, CIF, 4CIF, 16CIF) und benutzt einen verbesserten Kodieralgorithmus.

Die meisten Programme zum Anzeigen und oder Aufnehmen von digitalem Video unterstützen mehrere der genannten Formate, so daß wie bei den Audioformaten die Kodierung und Qualität flexibel an die Übertragungsgegebenheiten und Möglichkeiten des Empfängers angepaßt werden können. Einige Videoformate sind kombinierte Audio-Videoformate. Das führt auf den Begriff der Synchronisation verschiedener Medien und wird im nächsten Abschnitt genauer erläutert.

## Synchronisation

Nachdem nun die wichtigsten Medien beschrieben worden sind, ist ein grundlegender Begriff im Zusammenhang mit Multimedia die Synchronisation der verschiedenen Medien [67]. Synchronisation ist beispielsweise erforderlich beim gleichzeitigen Abspielen eines Audio- und Videodatenstroms. In den erwähnten kombinierten Audio-Videodatenströmen befinden sich abwechselnd Audio- und Video-Samples. Bei Stereo-Audio besteht ein Audio-Sample wiederum aus zwei aufeinanderfolgenden Samples für den rechten und linken Kanal. Der eingehende kombinierte Datenstrom wird dann gepuffert und von der Applikation so abgespielt, daß Audio und Video gleichzeitig abzulaufen scheinen.

Bei zwei separaten, aber zusammengehörenden Datenströmen für Audio und Video müssen die Daten- und Frameraten aneinander angeglichen werden. Dabei verwendet man meistens ein Master-Slave-Konzept, bei dem beispielsweise der Videostrom mit 15 fps abgespielt wird und der Audiostrom daran angepaßt wird.

Beschreiben läßt sich Synchronisation auch durch den Begriff der sogenannten **Medienzeit** im Verhältnis zu einer kontinuierlich fortschreitenden Zeit, beispielsweise der Echtzeit. Die Medienzeit beginnt beim Abspielen eines Datenstroms und schreitet relativ zur Echtzeit mit einer vorgegebenen Frame- oder Datenrate voran. Wird das Abspielen angehalten, so stoppt die Medienzeit, während die Echtzeit weiterläuft. Die Medienzeit kann nach folgender Formel berechnet werden [69]:

$$\text{Medienzeit} = \text{Medienstartzeit} + (\text{Echtzeit} - \text{Echtzeitstartzeit}) * \text{Rate}.$$

Bei Multimedia-Anwendungen im Internet kann das "Network Time Protocol" (NTP) [51] zur Synchronisation benutzt werden. Es stellt eine Internet-weite eindeutige und kontinuierlich fortschreitende Zeit zur Verfügung, die von im Internet verteilten sogenannten "Time-Servern" abgefragt werden kann. Multimedia-Daten können dann beim Verschicken mit einer NTP Zeitmarke versehen werden. Audio- und Videodaten, die gleichzeitig beim Sender digitalisiert werden, können so beim Empfänger mit Hilfe dieser Zeitmarken lippen-synchron zusammengesetzt werden. Das Echtzeitprotokoll RTP benutzt zu diesem Zweck

NTP Zeitmarken in seinem RTCP-Kontrollprotokoll. Natürlich können nicht nur Audio- und Videodaten synchronisiert werden. Auch nicht-zeitabhängige Daten wie einzelne Seiten in einer Whiteboard-Sitzung können mit Zeitmarken versehen werden, um die Seiten in einer Art Diashow “abspielen” zu lassen.

Nachdem nun die wichtigsten Multimedia-Dateiformate und -Konzepte vorgestellt wurden, sollen im letzten Abschnitt die Anwender von Multimedia-Applikationen, speziell Lernende und Lehrende, und die Anforderungen an Applikationen in diesem Umfeld näher betrachtet werden.

## 2.4 Multimedia-Einsatz in der Lehre

Neben den bisher beschriebenen technischen Grundlagen müssen nun noch die soziologischen Aspekte von verteilten Multimedia-Anwendungen, die in der Fachliteratur allgemein als “Groupware” oder Anwendungen für “Computer Supported Cooperative Work” (**CSCW**) [64], bezeichnet werden, betrachtet werden. Der Einsatz von Multimedia in Lehrveranstaltungen an der Hochschule stellt dabei noch spezielle Anforderungen. Im folgenden werden Szenarien und Vor- und Nachteile multimedial aufbereiteter Lehre aufgezeigt und daraus ein Pflichtenheft für ein Konferenz-Werkzeug aufgestellt, das die Vorteile der multimedialen Lehre maximiert und die Nachteile minimiert.

### Mögliche Lern- und Lehrszenarien

In der herkömmlichen Hochschullehre findet man grob gesagt vier Szenarien vor: In **Vorlesungen** trägt ein Dozent den Lernstoff meist vor einer großen Gruppe von Studenten vor, meist unterstützt durch Tafelanschrieb, Folien oder Beamer-Projektion, wobei zwischen Dozent und Studenten nur selten direkte Kommunikation stattfindet. Zu den Vorlesungen gehören manchmal **Übungen**, in denen der Lernstoff durch praktische Anwendung vertieft wird. Die Studenten bearbeiten dabei alleine oder in kleinen Gruppen meist schriftlich Aufgaben, die vom Dozenten korrigiert werden. Anschließend werden die Lösungen dann meist vor einer mittelgroßen Gruppe von Studenten vorgetragen und diskutiert (“Präsenzübung”). In **Seminaren** erarbeiten die Studenten selbständig einen etwa halbstündigen Vortrag über ein vorgegebenes Thema, der dann vor einer mittelgroßen Gruppe von Studenten mit Hilfe von Folien, Beamer oder Tafelanschrieb vorgetragen wird. Außerdem wird vom Student eine etwa 20-40 Seiten lange Ausarbeitung angefertigt. In **Praktika** werden in kleineren Gruppen vorgegebene Aufgaben gelöst. Neben den genannten Veranstaltungsformen existiert noch das selbständige Studium des Studenten zum Wiederholen oder Vertiefen des Lehrstoffs. Dies geschieht entweder alleine oder in Lerngruppen.

Multimedia kann bei diesen Lehr- und Lernszenarien nun auf unterschiedliche Art und Weise eingesetzt werden. Vorlesungen können per **Live-Audio-Videokonferenz**, also durch

die Übertragung des Bildes und des Tons des Dozenten im Internet mehr oder weniger gleichzeitig an beliebige Orte verteilt werden, wobei im Prinzip eine reine Ausstrahlung ohne Interaktionsmöglichkeit ausreicht. Der in Vorlesungen übliche Tafelanschrieb oder die Folienpräsentation kann entweder als eigenes Videobild oder in Form einer Datenkonferenz, die die Folien als Bilder anzeigt, übertragen werden. Nützlich ist, wenn bei der Datenkonferenz neben den Folien auch Annotationen versandt werden können, beispielsweise ein Zeiger zum Markieren des aktuell Angesprochenen. Datenkonferenz-Anwendungen dieser Art bezeichnet man oft als **Whiteboard** [61]. Die Kommunikation zwischen Teilnehmern oder zwischen Dozent und Teilnehmern kann in gewissem Rahmen auch rein schriftlich erfolgen. Nahezu gleichzeitig können sich mehrere Personen in einem sogenannten **Chat** oder Talk-Programm unterhalten, einer Minimalversion einer Datenkonferenz, in der lediglich kurze Textbotschaften ausgetauscht werden.

In Übungen können Live-Audio-Video-Datenkonferenzen geographisch verteilte Studentengruppen beim gemeinsamen Bearbeiten von Aufgaben unterstützen. Für die Zusammenarbeit und Vorbereitung der Übung seitens der Studenten eignet sich aber auch die schriftliche Kommunikation per **Email**. Da diese Kommunikation nicht zeitgleich erfolgt, bezeichnet man sie auch als asynchron. Wird eine Mail an einen festen Personenkreis geschickt, spricht man von einer "Mailing-Liste". Präsenzübungen können mit oder ohne Interaktionsmöglichkeit als Live-Konferenz ausgestrahlt werden. Das Gleiche gilt für Seminare. Praktika können zwar per Live-Konferenz ausgestrahlt werden, sie bieten sich allerdings nicht direkt zur multimedialen Aufbereitung an, da das praktische Arbeiten im Vordergrund steht, außer wenn das Praktikum die Arbeit am Computer beinhaltet. In diesem Fall können Programme beispielsweise in einer speziellen Datenkonferenz verteilt ausgeführt werden (**Application Sharing**).

Beim Selbststudium oder Vertiefen des Lernstoffs in kleinen Gruppen können Live-Konferenzen zur Kommunikation zwischen geographisch entfernt lebenden Studenten dienen. Auch Chats oder die asynchrone Kommunikation über Email bieten sich an. Aufgezeichnete Konferenzen, seien es Audio-, Video-, Daten- oder kombinierte Konferenzen können zu einem beliebigen Zeitpunkt abgespielt werden (**Video on Demand**) und bestimmte Passagen darin beliebig oft wiederholt werden. Nach dem Prinzip eines Schwarzen Bretts funktionieren **News**, bei denen Textnachrichten in verschiedenen Diskussionsforen für eine gewisse Zeit veröffentlicht werden, und auch das Antworten auf einzelne Nachrichten oder Diskussionsstränge ("Threads") innerhalb eines Forums möglich ist.

Für das Selbststudium oder Vor- und Nachbereiten des Lernstoffs eignen sich aber auch sehr gut Lehrmaterialien in digitaler Form. Vorlesungsmitschriften oder Lehrbücher können als Textdokumente oder im PostScript und PDF Format zum Betrachten oder Ausdrucken vorliegen, oder als HTML-Seiten im **WWW**, in die Links, Bilder, Animationen und auch Audio- und Videodaten eingebaut werden können. Die Materialien können zum beliebigen Herunterladen im Internet zur Verfügung stehen oder per CD-ROM an die Teilnehmer verteilt werden.

In Tabelle 2.2 sind die oben beschriebenen Medien noch einmal mit ihren Charakteristika bezüglich Medientyp, Orts- und Zeitabhängigkeit, Gruppengröße, Interaktion und Eignung für bestimmte Lehrveranstaltungen zusammengestellt. Dabei wird deutlich, daß die Größe der Gruppe bei Live-Veranstaltungen den Grad der Interaktion bestimmt. So werden Vorlesungen mit vielen Studenten und wenig Interaktion durchgeführt, und dagegen Übungen mit wenigen Studenten und hoher Interaktion. Chat, Email und News können zur Kommunikation zwischen beliebig vielen Teilnehmern, also beispielsweise zur Betreuung von Studenten oder der Kommunikation zwischen Studenten, benutzt werden, wobei beim Chat die Teilnehmer gleichzeitig anwesend sein müssen, bei Email und News nicht. Konferenzaufzeichnungen und andere in digitaler Form vorliegende Lehrmaterialien werden meist von Einzelpersonen zum Erarbeiten oder Vertiefen des Lernstoffs benutzt. Hier findet deshalb auch keine Interaktion zwischen Teilnehmern statt. Allen multimedialen Elementen ist jedoch die Ortsunabhängigkeit gemein, da die Daten im Internet und damit überall zur Verfügung stehen.

Zu jedem Medium gehören auch Konzepte zur **Unterstützung von Lehrveranstaltungen**, da bei der Kommunikation über den Computer viele Informationen einer Kommunikation von Angesicht zu Angesicht ("Face to Face", **FtF**) verloren gehen, beispielsweise Gestik und spontane Reaktionen [29]. Dazu gehört auch die Möglichkeit zur Bildung und Verwaltung von Teilnehmergruppen und die Zuteilung von Ressourcen an die Gruppenteilnehmer oder sogar einzelner Teilnehmer, beispielsweise das Erlauben oder Sperren von Audio-Beiträgen oder schriftlichen Beiträgen in einer Datenkonferenz. Sinnvoll ist auch die Möglichkeit, bestimmten Teilnehmern Privilegien zuordnen zu können, beispielsweise dem Dozenten, und so verschiedene Teilnehmer-Rollen einsetzen zu können. Auch die Konferenzkontrolle sollte verschiedenen Lern- und Lehrszenarien angepaßt werden können, beispielsweise eine explizite Sitzungskontrolle für synchrone Konferenzen. Daneben wäre es praktisch, Reaktionen wie das "Melden" von Teilnehmern, das in FtF-Diskussionen meist über Handzeichen geschieht, anzeigen zu können, beispielsweise durch einen für alle Teilnehmer sichtbar blinkenden Meldeknopf. Auch Abstimmungen unter allen Teilnehmern einer Konferenz, wie sie in der FtF-Diskussion öfter vorkommen und dann über Wortmeldungen oder Handheben durchgeführt werden, sollten mit Hilfe spezieller Bedienelemente zum Abstimmen möglich sein. Für den Dozenten ist auch eine Aufmerksamkeitsüberwachung der Studenten sinnvoll, beispielsweise in Form einer Statistik von aktiven Audio- oder Text-Beiträgen.

Bis jetzt wurde lediglich aufgezeigt, wie sich traditionelle Lehrszenarien multimedial aufbereiten lassen. Es ist naheliegend, daß sich durch den Einsatz von Multimedia die Lehr- und Lernmethoden selbst verändern und sich neue Konzepte herauskristallisieren. Die Forschungen auf diesem Gebiet befinden sich jedoch gerade erst am Anfang und unterliegen auf Grund der sich ständig ändernden technischen Gegebenheiten auch einer stetigen Wandlung. Im nächsten Abschnitt werden daher statt allgemeingültiger Aussagen einige

<p><b>Live-Konferenz, Whiteboard, Application Sharing:</b></p> <ul style="list-style-type: none"> <li>• Audio-, Videokonferenz, einzeln oder kombiniert mit Datenkonferenz</li> <li>• ortsunabhängig, synchron</li> <li>• bei reiner Ausstrahlung beliebige Gruppengröße</li> <li>• bei Interaktion kleine oder mittlere Gruppengröße</li> <li>• geeignet für Ausstrahlung und Aufzeichnung von Demonstrationen und allen Arten von Lehrveranstaltungen, insbesondere Vorlesungen und Seminaren</li> <li>• bei kleinen und mittleren Gruppen Interaktion und Diskussion möglich</li> </ul>
<p><b>WWW, Video on Demand, Aufzeichnungen, digitale Lehrmaterialien:</b></p> <ul style="list-style-type: none"> <li>• verwendete Medien: Text, Bilder, Audio, Video</li> <li>• ortsunabhängig, asynchron</li> <li>• geeignet für das Selbststudium, Interaktion nur innerhalb vorgegebener Lernpfade</li> </ul>
<p><b>Chat:</b></p> <ul style="list-style-type: none"> <li>• verwendetes Medium: Text</li> <li>• ortsunabhängig, synchron</li> <li>• geeignet für Diskussionen mit kurzen Textbeiträgen in kleinen bis mittleren Gruppen, beispielsweise Übungen, Seminaren, Praktika, Lerngruppen</li> </ul>
<p><b>Email:</b></p> <ul style="list-style-type: none"> <li>• verwendetes Medium: Text</li> <li>• ortsunabhängig, asynchron</li> <li>• geeignet für Diskussionen mit beliebig langen Textbeiträgen in beliebig großen Lerngruppen, oder zum technischen und inhaltlichen Informationsaustausch bei der Vor- und Nachbereitung von allen Lehrveranstaltungen</li> </ul>
<p><b>News:</b></p> <ul style="list-style-type: none"> <li>• verwendetes Medium: Text, aber auch Einbinden von Bild und Ton möglich</li> <li>• ortsunabhängig, asynchron</li> <li>• geeignet für Diskussionen mit beliebig langen Textbeiträgen in beliebig großen Lerngruppen, oder zum technischen und inhaltlichen Informationsaustausch bei der Vor- und Nachbereitung von allen Lehrveranstaltungen</li> <li>• Möglichkeit zu Diskussionssträngen innerhalb eines Diskussionsforums ("Threads")</li> </ul>

Tabelle 2.2: Charakteristika von Medien und ihre Eignung für Lehrveranstaltungen



Erfahrungen mit dem Einsatz von Multimedia in der Hochschullehre vorgestellt.

### **Erfahrungsberichte**

Im folgenden soll von Erfahrungen beim Einsatz von Multimedia in der Hochschullehre berichtet werden. Am Rensselaer Polytechnic Institute in Troy, New York, wurden 1996 von Joseph B. Walter, einem Professor der Literatur, Sprache und Kommunikation, Experimente und Untersuchungen über die Auswirkungen von computerbasierter Kommunikation im Gegensatz zur FtF-Kommunikation von in studentischen Arbeitsgruppen durchgeführt [75]. Die allgemeine Annahme, daß computerbasierte Kommunikation weniger sozial und unpersönlich ist, konnte widerlegt werden. In Tests wurde gezeigt, daß die Studenten sich genauso sozial, eher noch gefühlsbetonter verhalten als bei der FtF-Kommunikation. Außerdem stellte sich heraus, daß die Gruppenarbeit mittels Computer demokratischer ablief: Alle Studenten waren gleichberechtigt, auch schüchterne Personen kamen zu Wort, denn Wortführer, die in normalen Diskussionen durch Lautstärke und Gestik dominieren, konnten sich hier nicht so leicht durchsetzen. Da allerdings die nichtverbale Kommunikation, also Gestik oder Verhaltensnormen fehlten, und man seinem Gegenüber bei der Kommunikation über Email "nicht ins Gesicht schauen" muß, ist die Hemmschwelle für Streitereien und Wortgefechte herabgesetzt und es kommt leicht zu sogenannten "Flames", die seit Erfindung der Email und News in den 60er Jahren bekannt sind. Flaming wird allerdings auch nicht so ernst genommen wie harte Wortgefechte in der FtF-Kommunikation. Doch dadurch und durch die rein schriftliche Kommunikation dauerte es länger als in einer normalen Gruppenarbeit, bis die Studenten in einer Diskussion zu einer Entscheidung kamen. Es stellte sich heraus, daß computerbasierte und nicht-computerbasierte Gruppenarbeit im Endeffekt gleich ist, nur daß die computerbasierte Gruppenarbeit mehr Zeit benötigt.

Während im obigen Fall lediglich eine Datenkonferenz per Email stattfand, gibt es jedoch auch erste Projekte, Audio-Video-Konferenzen in der Hochschullehre einzusetzen. An der Universität Hohenheim wurde 1996 von Prof. Krcmar ein verteiltes Seminar durchgeführt, bei dem Studenten der Universitäten Arizona, Tucson, und Hohenheim gemeinsam ein Thema aus dem Bereich Videokonferenzen vorbereiteten und anschließend vortrugen [40]. Dabei kamen in der Vorbereitungsphase meist Email, WWW und Audio-Video-Konferenzen zum Einsatz, beim Vortrag eine Audio-Video-Daten-Konferenz. Aus Zeitmangel, kostengünstigere Alternativen zu testen, entschied man sich für ein Intel ProShare VideoSystem für Microsoft Windows und ISDN (siehe Kapitel 3), inklusive Kamera und Mikrofon-Kopfhörer, und für Novell LAN Workplace für die Datenkonferenz über ISDN unter Microsoft DOS. Die Konferenzen wurden in einem speziell eingerichteten Gruppendiskussionsforschungslabor abgehalten, siehe Abbildung 2.19. Während die Kommunikation über Email während der gesamten Zeit nach Anfangsproblemen gut funktionierte, gab es beim Aufbau der Videokonferenzverbindung größere und langanhaltende technische Probleme, die zum Teil dazu führten, daß die Vorträge lediglich über Telefon übertragen werden konnten. Der technische Aufwand bedeutete ein erhebliches Mehr an Organisation

und Zeitaufwand von Seiten der Veranstalter als auch der Studenten. Insbesondere war für die Vorträge immer noch eine zweite Person (“Chauffeur”) nötig, die dem Vortragenden die technische Seite abnahm und Programme, Kamera und Mikrofone bedienten, während sich der Student auf seinen Vortrag konzentrieren konnte. Aufgrund der Zeitverschiebung mußte in Kauf genommen werden, daß die synchrone Kommunikation zum Teil zu ungewöhnliche Uhrzeiten stattfand. Trotz der technischen Schwierigkeiten und des Mehraufwandes fand das Seminar bei allen Beteiligten großen Anklang.



Abbildung 2.19: Gruppendiskussionsforschungslabor der Universität Hohenheim

Ein anderes Projekt ist die 1997 durchgeführte Tele-Vorlesung “Informatik und Gesellschaft” von Prof. Schinzel an der Universität Freiburg [61]. Dabei wurde die Vorlesung vor den Freiburger Studenten konventionell gehalten und an baden-württembergische Universitäten als Audio-Videokonferenz über Mbone ausgestrahlt. Die Vorlesungsausarbeitung, die Übungen und zusätzliche Materialien wurden in digitalisierter schriftlicher Form im WWW abgelegt. Darüber hinaus wurde an den einzelnen Universitäten die Bildung selbständiger Arbeitsgruppen angeregt, die über eine Mailing-Liste mit den Gruppen der anderen Universitäten kommunizieren sollten. Bei der Durchführung bestand das größte Hindernis in den technischen Problemen der Mbone-Verbindung. In vier von neunzehn Veranstaltungen fiel die Übertragung komplett aus. Auch die frei verfügbaren sogenannten “Mbone Tools”, die Programme zur Durchführung der Audio-, Video- und Datenkonferenz, erwiesen sich als unausgereift, software-ergonomisch schlecht durchdacht, und behinderten durch häufige Abstürze den Vorlesungsfluß. Dadurch kam es zu einem verlangsamten Vortragstempo im Vergleich zu einer konventionellen Vorlesung. Die Beteiligung der entfernt teilnehmenden Studenten bei der Vorlesung und in der Email-Kommunikation war sehr gering, was wohl an der erhöhten Hemmschwelle vor den anderen unbekanntem Teilnehmern lag oder weil die Studenten nicht einschätzen konnten, wann es opportun

war eine Frage zu stellen. Daneben bestand bei Studenten, die die Vorlesung per Video-konferenz verfolgten, die starke Neigung zu Nebentätigkeiten wie Surfen im WWW oder Email-Lesen. Gegen Ende der Vorlesung wurde daher mit Erfolg versucht, durch eingestreute kurze Wissensabfragen die Studenten direkt in die Übertragung mit einzubeziehen, sei es in Form von schriftlichen Beiträgen auf dem MBone Whiteboard, oder in Form von Audio-Beiträgen, wofür aber alle Teilnehmer ein Mikrofon brauchten. Insgesamt wurde festgestellt, daß zumindest in dieser Form die Vorlesung nur mit großem echnischen Aufwand und hochmotivierten Studenten erfolgreich durchgeführt werden kann. Aufgrund der Erfahrungen sieht Prof. Schinzel den Einsatz von computergestütztem Lernen am ehesten im Rahmen entsprechend modifizierter Lehr- und Lernformen, beispielsweise in Tele-Übungen, in denen zuvor von den Studenten zu lesende Texte gemeinsam diskutiert werden.

Die MBone Tools wurden 1997 ebenfalls bei einem verteilten Seminar zwischen den Universitäten Tübingen (Prof. Rosenstiel, Prof. Zell) und Stuttgart (Prof. Rothermel) eingesetzt. Die Betreuung der Studenten wurde auf konventionelle Art in persönlichen Gesprächen vor Ort durchgeführt, lediglich die Vorträge wurden per Audio-Video-Datenkonferenz an alle Standorte übertragen. Zur Kommunikation zwischen allen Teilnehmern und Veranstaltern hat sich eine Mailing-Liste und eine zentrale WWW-Seite als äußerst praktisch erwiesen. Um die Studenten mit den nicht sehr intuitiv bedienbaren MBone Tools vertraut zu machen, wurden vor den eigentlichen Vorträgen zwei Konferenz-Termine zum Üben festgesetzt. Diese waren dringend notwendig, da es durch Fehlbedienungen und technische Probleme immer wieder zu Netzausfällen und Abstürzen kam. Gegen Ende des Seminars hatten sich die Teilnehmer aber bereits so gut an die Bedienung gewöhnt, daß äußerst produktive Vorträge und Diskussionen zustande kamen. Um das Interesse der Studenten während der Vorträge wach zu halten, wurden die Vortragenden angehalten, mindestens zwei bis drei "interaktive Elemente" in den Vortrag einzubauen, beispielsweise durch eine Folie mit Wissensstandabfragen im Whiteboard, zu denen die Studenten sich per Audio-Beitrag oder schriftlich äußern konnten. Es fiel auf, daß schüchternere Studenten selbstbewußter auftraten, da sie sich durch die computerbasierte Kommunikation, in der sich keine Wortführer durchsetzen konnten, sicherer fühlten. Trotz des erheblichen Mehraufwands bezüglich Technik, Betreuung und Organisation war das Seminar sehr erfolgreich und wurde auch in den nächsten Semestern fortgesetzt. Der technische Aufwand nahm aufgrund der fehlenden Weiterentwicklung der MBone Tools im Laufe der folgenden Seminare jedoch leider nicht ab.

Im nächsten Abschnitt sollen die aus den Erfahrungen zu ziehenden Schlüsse noch einmal zusammengefaßt und die Vor- und Nachteile von Multimedia und insbesondere Telekonferenzen in der Hochschullehre aufgezeigt werden.

## Vor- und Nachteile

Aus den in den letzten Abschnitten geschilderten Szenarien und Erfahrungen lassen sich die Hauptargumente für und gegen den Einsatz von Multimedia in der Lehre herausarbeiten. Eindeutigen Vorteilen stehen dabei auch Nachteile gegenüber, die man sorgfältig gegeneinander abwägen muß. In Tabelle 2.3 sind die wichtigsten Punkte aufgeführt. Aus den letzten Abschnitten geht hervor, daß die **Ortsunabhängigkeit** wohl den entscheidendsten Vorteil darstellt, da über Internet und mit Hilfe von Telekonferenzen und anderen multimedialen Anwendungen weltweite Kooperationen und Lehrveranstaltungen möglich sind, und so ein viel größerer Teilnehmerkreis angesprochen wird [16].

Die Ortsunabhängigkeit kann bei Präsenzveranstaltungen und großer geographischer Entfernung zwischen den Studenten wegen der Zeitverschiebung aber auch nachteilig sein. In diesem Fall kann die Veranstaltung aufgezeichnet werden und zu einem beliebigen späteren Zeitpunkt noch einmal abgespielt werden. Hier geht zwar die Interaktionsmöglichkeit bei einer Präsenzveranstaltung verloren, sie ist aber bei Vorlesungen sowieso nicht stark ausgeprägt ist (siehe Tabelle 2.2) und bei anderen Veranstaltungen durch die zusätzliche Kommunikation über Email oder News ersetzt werden kann. Ein Problem kann bei Audio-Videokonferenzen aber die fehlende oder bei geringer Übertragungsqualität nicht so ausgeprägt sichtbare Gestik und Mimik sein. Eine rein schriftliche Kommunikation über Email und News kann ebenfalls nachteilig sein, da sie länger dauert und vielleicht nicht alles "mit Worten" gesagt werden kann. Hier müssen Live- und computerbasierte Kommunikation sinnvoll abgewechselt werden und zum besseren Kennenlernen der Teilnehmer einer Telekonferenz beispielsweise "echte" Treffen veranstaltet werden. Untersuchungen haben gezeigt, daß anschließend weniger Probleme auf Grund der fehlenden Gestik und Mimik auftreten [75]. Bei bestimmten Kommunikationsformen kann die schriftliche Form aber durchaus auch Vorteile haben, zum Beispiel bei der Verteilung von längeren Textdokumenten an eine große Studentenzahl. Die Wahl der Medien muß also an die Lehrform angepaßt werden und es sollten zudem Kurse angeboten werden, die die Studenten oder auch Dozenten in der effizienten Kommunikation über das jeweilige Medium ausbilden. Die Fähigkeit, auf elektronischem Weg kommunizieren zu können, bezeichnet man auch als "Medienkompetenz" [41].

Bei einem weltweiten Teilnehmerkreis können auf Grund der verschiedenen Sprachen Probleme auftreten. Als "Lingua Franca" des Informationszeitalters hat sich jedoch die englische Sprache durchgesetzt, so daß man sich bei internationalen Kursen gut auf Englisch einigen kann. Da das Beherrschen von Fremdsprachen und insbesondere Englisch auch im späteren Berufsleben unerlässlich ist, kann die universitäre Lehre, die Fremdsprachen selbstverständlich einsetzt, hier einen sinnvollen Beitrag zur Ausbildung leisten.

Der zweite große Vorteil von Multimedia ist die **Zeitunabhängigkeit**, da aufgezeichnete Lehrveranstaltungen und Lehrmaterialien in digitaler Form jederzeit über das Internet

<b>Ortsunabhängigkeit</b>	
+ weltweite Teilnahme an Lehrveranstaltungen über Internet möglich	– eventuell große Zeitverschiebung zwischen den Teilnehmern bei Präsenzveranstaltungen → Aufzeichnen
+ weltweite Kooperationen zwischen Universitäten und Teilnehmergruppen	– Probleme mit der fehlenden Gestik bei Audio-Videokonferenzen und Probleme mit der rein schriftlichen Kommunikation bei asynchroner Kommunikation → Kurse zur effizienten Nutzung des jeweiligen Mediums, Mischung aus Telekonferenzen und “echten” Konferenzen
+ höhere Breitenwirksamkeit durch potentiell weltweiten Teilnehmerkreis	– eventuelle Sprachbarriere → Einigung auf eine Sprache im Kurs; Fremdsprachenkenntnisse nützen auch im späteren Arbeitsleben
<b>Zeitunabhängigkeit</b>	
+ Abspielen von Aufzeichnungen, beliebiger Zugriff auf Materialien in digitaler Form über das Internet	– eventuell längere Wartezeit auf Materialien in Stoßzeiten → Vormerken und automatisiertes Übertragen von Materialien zu Nicht-Stoßzeiten
<b>Flexibilität</b>	
+ freier Zugriff für alle und jederzeit auf Materialien in digitaler Form über das Internet	– die Studenten ertrinken in einer Informationsflut → Dozenten werden zu Trainern und leisten Hilfestellung beim Setzen von Prioritäten; Vortragende oder Teilnehmer werden durch Chauffeur von technischen Problemen abgeschirmt
+ qualitative Verbesserung der Lehre; den Studenten steht eine große Auswahl an Universitäten und Lehrveranstaltungen zur Verfügung	– Leistungen werden auf Grund fehlender Übereinkommen und unterschiedlicher Anforderungen eventuell nicht anerkannt → Abkommen zwischen Universitäten sind nötig
<b>Kosten</b>	
+ durch Orts- und Zeitunabhängigkeit entfallen Reisekosten zum Vortragsort	– zur Teilnahme an Multimedia-Veranstaltungen ist zum Teil teure Software und Spezial-Hardware nötig → Einsatz frei verfügbarer Software und Medienformaten, die auch günstige Hardwarekomponenten beherrschen
+ in elektronischer Form vorliegende Materialien können schnell und einfach aktualisiert werden	– die Umsetzung bisheriger Lehrmaterialien und die ständige Aktualisierung ist zeit- und personalaufwendig → Aufbau von Lehrmaterialien in leicht aktualisier- und wiederverwendbarer Form, Verwenden von Skripten

Tabelle 2.3: Vorteile, korrespondierende Nachteile und Lösungsvorschläge von Multimedia in der Lehre

zur Verfügung stehen. Größere Datenmengen, wie sie bei Audio-Videokonferenzen anfallen, können beispielsweise durch Vormerken automatisch zu Zeiten übertragen werden, in denen das Netz nur wenig ausgelastet ist, also nachts oder an Wochenenden. Automatisierte Vormerk- und Datenübertragungsverfahren können die Wartezeiten und den Aufwand einer manuellen Übertragung minimieren.

Durch den beliebigen Zugriff auf Multimedia-Daten im Internet kann die Lehre und das Lernen sehr gut an einzelne Studenten und Studentengruppen angepaßt werden. Diese **Flexibilität** ermöglicht den Studenten dann zwar eine sehr große Wahlfreiheit. Multimedia erlaubt es, daß nicht mehr alle Hochschulen und alle Fakultäten ein vollständiges Lehrangebot im jeweiligen Fach bereithalten müssen, um einen Studiengang mit abschließender Verleihung eines Grades anzubieten [76]. Es besteht aber gleichzeitig die Gefahr, daß die Studenten in der "Informationsflut ertrinken" oder über die technische Seite und Bedienungsfragen die Inhalte einer Lehrveranstaltung in den Hintergrund rücken. Hier müssen die Dozenten neue Aufgaben übernehmen und als Trainer oder Moderatoren Hilfestellung beim Setzen von Prioritäten und beim Erlernen von Medienkompetenz leisten. "Chauffeure" können Vortragende und Studenten die technische Seite abnehmen [40]. Durch die große Auswahl an angebotenen Materialien und Kursen können die Studenten sich aber auch die besten Kurse von verschiedenen Universitäten heraussuchen und erhalten so eine qualitativ bessere Ausbildung. Die große Konkurrenz wird die Veranstalter auch zu besseren Lehrveranstaltungen zwingen, da sich schlechte Kurse mangels Teilnehmerzahlen nicht halten können. Leider gibt es zum jetzigen Zeitpunkt nur wenige Abkommen zwischen Universitäten bezüglich der gegenseitigen Anerkennung von Lehrveranstaltungen. Hier sind noch allgemeingültige Anforderungen zu entwickeln und Abkommen zu treffen. Erste Ansätze gibt es mit der Bestrebung der Einführung von Punktsystemen und einheitlichen universitären Abschlüssen wie dem Bachelor [76].

Daneben spielen aber auch die **Kosten** beim Einsatz von Multimedia eine wichtige Rolle. Durch Orts- und Zeitunabhängigkeit entfallen die teuren Reisekosten zum Vortragsort. Allerdings wird zur Teilnahme an Multimedia-Veranstaltungen häufig spezielle Soft- und Hardware vorausgesetzt, die sehr teuer sein kann, was für Universitäten ein großes Problem sein kann. Hier sollte vorrangig frei verfügbare Software eingesetzt werden, auch wenn sie eventuell nicht so viel Funktionalität wie ein kommerzielles Produkt bietet. Es sollten auch nur weit verbreitete Medienformate zum Einsatz kommen, die von vielen Multimedia-Anwendungen und günstiger Standard-Hardware unterstützt werden.

In der Forschung und Lehre können sich Themen und Wissen schnell ändern, und insbesondere Lehrbücher, Fachliteratur und Vorlesungsskripten schnell veralten. Liegen die Lehrmaterialien in digitaler Form vor, so lassen sich sie kostengünstig und einfach aktualisieren. Allerdings ist die erstmalige Umsetzung von herkömmlichen Materialien in die digitale Form und die ständige Aktualisierung zeit- und personalaufwendig. Der Aufwand kann minimiert werden, wenn auf einen modulhaften Aufbau der Materialien geachtet wird,

bei denen sich Teile wiederverwenden und leicht austauschen lassen. Skripten können Aktualisierungsvorgänge automatisieren.

Wenn man nun die Vorteile von Multimedia in der Lehre voll ausnutzt und die Probleme wie aufgezeigt eliminiert oder minimiert, erzielt man auf jeden Fall eine Verbesserung der Lehre, eine höhere Breitenwirksamkeit und eine bessere Informationsverarbeitung. Im nächsten Abschnitt sollen Anforderungen für ein Werkzeug definiert werden, das universell für verschiedene Lehrveranstaltungen und Benutzer eingesetzt werden kann und dabei die Vorteile von Multimedia optimal ausnutzt und die Nachteile minimiert.

## 2.5 Anforderungen

Welchen Anforderungen müssen nun Anwendungen genügen, die möglichst viele Lernszenarien und Benutzergruppen bei der multimedialen Kommunikation unterstützen? Das Pflichtenheft reicht von Aspekten, die die Software und Hardware betreffen, bis hin zu Anforderungen, die sich aus der Lehrveranstaltungsform ergeben. Die Anwendungen sollten möglichst alle Formen von **Lehrveranstaltungen unterstützen**, also sowohl synchrone, als auch asynchrone Kommunikation, und auch das Selbststudium mit multimedialen Daten ermöglichen. Die höchsten Anforderungen stellen synchrone Audio-Video- und Datenkonferenzen. Asynchrone Kommunikation kann über Standard-Email- und News-Programme und WWW Browser erfolgen beziehungsweise sollte von einer Anwendung für synchrone Konferenzen ebenfalls unterstützt werden.

Für Live-Konferenzen benötigt man zunächst eine Anwendung, die Multicast Adressen für die Konferenz auswählt und diese den anderen Teilnehmern bekannt macht. Dabei sollte auch das einfache Einladen von anderen Teilnehmern unterstützt werden. In synchronen Konferenzen, bei denen ein Teilnehmer länger vorträgt, also in Vorlesungen oder Seminaren, sollten zumindest der Dozent und seine Vortragsmaterialien wie die Folien oder der Tafelanschrieb an alle Zuhörer übertragen werden. Für Wortmeldungen sollten alle Standorte zumindest das Senden von kurzem Text, besser Audio, unterstützen. Bei kleineren Teilnehmergruppen ist auch das Senden eines Videobildes pro Standort oder Teilnehmer sinnvoll. Die Datenkonferenz sollte ebenfalls Interaktion erlauben, beispielsweise das gemeinsame Bearbeiten eines Textes oder einer Zeichnung, und sie sollte einen virtuellen Pointer und andere primitive Malelemente zum Markieren besitzen.

Lehrveranstaltungen werden unterstützt durch die Möglichkeit Gruppen und Untergruppen zu bilden und Teilnehmern verschiedene Rollen wie Student oder Dozent mit unterschiedlichen Privilegien zuzuweisen. Daneben sind virtuelle Mauszeiger, Annotationen und Abstimmungen oder Leistungsabfragen auf Knopfdruck sinnvoll. Um im Falle von Vorlesungen und Seminaren den Dozenten nicht stören zu können, sollten die Zuhörer oder bestimmte Teilnehmer in der Audio-Video-, aber auch Datenkonferenz "stummgeschaltet"

werden können und die Sitzungskontrolle umschaltbar sein zwischen expliziter und impliziter Kontrolle. Nicht direkt zum Vortrag gehörende nebenläufige Kommunikation zwischen allen Teilnehmern, beispielsweise für Fragen zur technischen Bedienung, kann etwa über ein Chat-Programm erfolgen, das als zusätzliche Datenkonferenz nebenher läuft.

Die asynchrone Kommunikation kann wie erwähnt von existierenden Standard-Anwendungen übernommen werden. Um nicht zu viele einzelne Anwendungen bedienen zu müssen, wäre es allerdings praktisch, diese in eine Anwendung für Live-Konferenzen zu integrieren. So ist es sinnvoll, Namen und Email-Adressen der Teilnehmer einer synchronen Konferenz in einer Art Adreßbuch zu speichern und die Daten für die anschließende asynchrone Kommunikation zu benutzen. Für das Selbst- oder Gruppenstudium und die Nachbereitung von Live-Konferenzen sollten die Konferenzdaten aufgezeichnet und möglichst mit der gleichen Anwendung, mit der sie aufgenommen wurden, wieder abgespielt werden können.

Die bei synchroner und asynchroner Kommunikation benutzten **Medienformate** sollten aus Kompatibilitätsgründen gängige Formate sein, die auch von anderen Anwendungen unterstützt werden. So sollten Audiodaten bei Live-Konferenzen im PCM oder GSM Format übertragen werden, gespeicherte Audiodaten können zusätzlich im WAV, AU, RA oder MPEG Format vorliegen. Selbstverständlich sollten Lautstärke, Mikrofonpegel und die Wahl der Audio-Ein- und Ausgänge leicht einstellbar sein. Für Live-Video sollte dementsprechend H.261 oder M-JPEG und für aufgezeichnete Videodaten zusätzlich das MPEG, AVI und RM Format in verschiedenen Größen und Frameraten unterstützt werden. Bei vielen Teilnehmern, die Videobilder senden, ist es praktisch, die Bilder als kleine Vorschau anzuzeigen, die auf Wunsch aber auch vergrößert werden kann. Bei Datenkonferenzen sollte Text als ASCII oder HTML Datei vorliegen oder im RTF, PS oder PDF Format, digitalisierte Bilder sollten das verbreitete JPEG oder GIF Format benutzen.

Die in multimedialen Lehrveranstaltungen benötigte **Hardware** muß hauptsächlich kostengünstig sein, da Bildungseinrichtungen meist nur über geringe Mittel verfügen. Grundvoraussetzung sind Rechner mit Internetanschluß. Damit ist zumindest der Empfang von Live-Audio-Video-Konferenzen, die Interaktion bei Datenkonferenzen und natürlich das Herunterladen von digitalisierten Lehrmaterialien möglich. In dieser Minimalkonfiguration können beispielsweise auch Studenten von Zuhause aus mit ihren Rechnern über Modem an multimedialen Konferenzen teilnehmen. Für das Senden von Live-Audio wird zudem eine Soundkarte benötigt, an die Mikrofon und Kopfhörer beziehungsweise ein Headset (Kopfhörer mit integriertem Mikrofon) angeschlossen werden können. Die Soundkarte muß keine Studioqualität besitzen, da meistens nur Sprache übertragen wird, sollte aber Full Duplex zum gleichzeitigen Senden und Empfangen von Ton unterstützen. Für Teilnehmergruppen in einem Raum sind (Aktiv-)Lautsprecher eine nützliche Hardware-Ausstattung.

Zum Senden von Video benötigt man entweder eine Digital-Kamera oder eine analoge Kamera zusammen mit einer Videokarte. Videokarten sind zwar relativ teuer, erledigen



aber die Kompression der Videodaten und entlasten somit die CPU, so daß man auch ältere Rechner verwenden kann. Als Kamera kann man entweder die teureren Camcorder benutzen, die man dann auch bei anderen Gelegenheiten einsetzen kann, oder die günstigeren Desktop Kameras, die meist stationär oben auf dem Monitor angebracht werden, siehe Abbildung 2.18.

Auch die Wahl des **Betriebssystems** spielt eine wichtige Rolle. An Universitäten werden häufig Sun OS beziehungsweise Solaris, Linux und andere Unix-Derivate, sowie Microsoft Windows und Windows NT eingesetzt, meist nebeneinander, manchmal aber auch ausschließlich. Um nicht von vornherein ganze Lehrstühle oder Universitäten auszuschließen, sollte die Software auf möglichst vielen Plattformen laufen, also plattformunabhängig sein. Werden jedoch für die multimediale Lehre neue Rechner angeschafft, sollte man Unix-Rechnern den Vorzug geben. Sie sind erfahrungsgemäß leistungstärker, ausfallsicherer und schon seit Jahrzehnten auf Multitasking und den Mehrbenutzerbetrieb ausgelegt [42]. Allerdings ist PC-Hardware meistens günstiger. Die insgesamt beste Lösung sind daher Linux-PCs.

Die **Software** selbst sollte auf Grund der chronisch schlechten Geldsituation staatlicher Ausbildungsstätten ebenfalls möglichst umsonst sein. Sie sollte die Anzeige oder das Senden der oben genannten Medienformate beherrschen und zur Minimierung der Netzbelastung bei Live-Konferenzen die Echtzeitprotokolle RTP und LRMP und die Verteilung über Unicast und Multicast (Mbone) unterstützen. Um mit anderen Anwendungen, die zur Übertragung nicht das Internet benutzen, kompatibel zu sein, sollte die Software auch die ISDN-Konferenzstandards H.323 und T.120 beherrschen.

Aus Entwicklersicht sollte die Software zur leichten Erweiter- und Wartbarkeit in einer objektorientierten **Programmiersprache** geschrieben sein. Günstig wäre eine direkte Unterstützung der Kommunikation über das Internet und die Unterstützung der gebräuchlichsten Medienformate. Vorteilhaft wäre außerdem der einfache Entwurf einer **Benutzeroberfläche** ("Graphical User Interface", GUI), denn die Software sollte neben der reinen Funktionalität auch intuitiv bedienbar sein, und zwar sowohl von Anfängern, als auch von Fortgeschrittenen, Studenten und Dozenten. Die Benutzerfreundlichkeit drückt sich insbesondere in einer durchdachten und ansprechend gestalteten grafischen Benutzerschnittstelle einer Anwendung aus [68].

Äußerst sinnvoll wäre eine umfangreiche Hilfefunktion, die die verschiedenen Funktionen und Bedienelemente der Benutzeroberfläche erklärt, sei es in Form eines Hilfe-Menüs oder als sogenannte "Balloon Help", die beim Bewegen der Maus über einem Bedienelement erscheint. Praktisch wäre auch kontextsensitive Funktionalität, beispielsweise ein angepaßtes Menü für die rechte Maustaste, je nachdem, über welchem Teil der Anwendung sich der Mauszeiger gerade befindet. Dabei und bei der Beschriftung der Bedienelemente wie zum Beispiel Knöpfen ("Buttons") wäre es hilfreich, wenn verschiedene Sprachen einstellbar

wären, zumindest aber Deutsch und Englisch. Günstig wäre auch die individuelle Einstellbarkeit der dabei verwendeten Schriftarten und -größen.

Die einzelnen Anwendungen für die Audio-, Video- und Datenkonferenz sollten über einheitliche Bedienparadigma verfügen, also beispielsweise einheitlich aufgebaute und angeordnete Menüleisten und Rollbalken besitzen, und bei einem Mausklick ähnliche Funktionen anbieten. Die Benutzerfreundlichkeit wird deutlich erhöht, wenn nicht für jede Konferenz- und Kommunikationsform eine andere Anwendung benötigt wird, sondern möglichst alle Anforderungen von einer einzigen kombinierten Anwendung erfüllt werden. Da die höchsten Anforderungen von Live-Konferenzen gestellt werden, sollte die kombinierte Anwendung gleichzeitig Audio-, Video- und Datenkonferenzen, Konferenz-Beitritt und Chat unterstützen. Asynchrone Kommunikation und das Selbststudium sollte mit der gleichen Anwendung ebenfalls möglich sein. Die Videobilder und die nötigsten Einstellungen wie Lautstärke und Senden von Audio und Video sollten der besseren Bedienbarkeit und Übersichtlichkeit auf einem normal großen Monitor mit einer Auflösung von etwa 1024x768 Pixel Platz finden und ergonomisch geschickt angeordnet sein [67]. Auf Grund der Aufmerksamkeitsverteilung sollten sich daher beispielsweise wichtige Elemente wie das Videobild des Sprechers und eventuell anderer Teilnehmer rechts oben befinden, der nebenläufige Chat in der Ecke mit der geringsten Aufmerksamkeit links unten. Das Whiteboard findet dann unten rechts Platz und die Einstellungen für das eigene Video und Audio und die Video-Vorschaubilder links oben und in der linken Mitte. Die Einstellungen für das Empfangen und Senden von Audio und Video, beispielsweise die Framerate oder das Medienformat, sollten auch von unerfahrenen Benutzern getroffen werden können, zum Beispiel durch Vorgabe mehrerer Standardbenutzerprofile. Für fortgeschrittene Benutzer sollten sich weitere Funktionen in einer Menüleiste oder hinter speziellen Bedienelementen befinden. Die Anwendung sollte also einerseits einfach bedienbar, aber gleichzeitig auch voll ausgestattet sein.

Die Anforderungen an eine universelle Konferenz-Software sind in Tabelle 2.4 noch einmal stichwortartig zusammengefaßt. Im nächsten Kapitel werden nun die gängigsten existierenden Konferenzanwendungen auf die hier beschriebenen Anforderungen für die multimediale Lehre hin untersucht.

<b>Unterstützung von Lehrveranstaltungen:</b>
<ul style="list-style-type: none"> <li>• Unterstützung synchroner Live-Audio-Video-Datenkonferenzen</li> <li>• Unterstützung von Selbststudium und asynchroner Kommunikation mittels Email, News und WWW</li> <li>• Unterstützung der Teilnehmerverwaltung in Gruppen</li> <li>• Unterstützung verschiedener Teilnehmerrollen und -privilegien</li> <li>• Unterstützung impliziter und expliziter Sitzungskontrolle</li> <li>• Annotations-, Abstimmungs-, Meldemöglichkeit</li> <li>• Optionale Aufmerksamkeitsüberwachung</li> <li>• Unterstützung von Wissensabfragen mit automatischer Auswertung</li> </ul>
<b>Unterstützte Medienformate:</b>
<ul style="list-style-type: none"> <li>• Live-Audio: PCM, GSM; Audio: WAV, AU, RA, MPEG</li> <li>• Live-Video: H.261, M-JPEG; Video: MPEG, AVI, RM</li> <li>• Text: ASCII, HTML, RTF, PS, PDF</li> <li>• Bild: JPEG, GIF</li> </ul>
<b>Hardware:</b>
<ul style="list-style-type: none"> <li>• möglichst einfacher und günstiger Rechner</li> <li>• benötigt lediglich Modem oder Netzwerkkarte für Internet-Anschluß</li> <li>• benötigt Standard-Soundkarte, Kopfhörer, Mikrofon, optional Lautsprecher</li> <li>• besitzt optional Videokarte und Desktop-Kamera</li> </ul>
<b>Betriebssystem:</b>
<ul style="list-style-type: none"> <li>• möglichst Multitasking-/Multiuser-fähig, möglichst günstig</li> <li>• Unix/Linux bevorzugt</li> </ul>
<b>Programmiersprache:</b>
<ul style="list-style-type: none"> <li>• möglichst frei erhältlicher Compiler</li> <li>• objektorientiert wegen besserer Wart- und Erweiterbarkeit</li> <li>• unterstützt möglichst Internet-Kommunikation und gebräuchlichste Medienformate</li> <li>• unterstützt einfachen Entwurf einer Benutzeroberfläche</li> </ul>
<b>Software:</b>
<ul style="list-style-type: none"> <li>• möglichst plattformunabhängig und frei erhältlich</li> <li>• unterstützt Erzeugen und Beitritt zu Multicast Sitzungen</li> <li>• unterstützt Anzeigen/Senden obiger Medienformate</li> <li>• unterstützt RTP, LRMP, Multicast (MBone)</li> <li>• unterstützt ISDN-Standards H.323 und T.120</li> </ul>
<b>Benutzeroberfläche:</b>
<ul style="list-style-type: none"> <li>• übersichtlich, funktional und intuitiv</li> <li>• umfangreiche Hilfefunktion</li> <li>• Unterstützung für mehrere Sprachen, zumindest Deutsch und Englisch</li> <li>• kombiniert möglichst alle nötigen Anwendungen in einer</li> <li>• mittels vorgegebener Benutzerprofile auch von Anfängern bedienbar</li> <li>• viele Funktionen auch für fortgeschrittene Benutzer</li> </ul>

Tabelle 2.4: Anforderungen an ein universelles Konferenz-Werkzeug

# Kapitel 3

## Stand der Technik

In diesem Kapitel werden aktuelle Konferenz-Werkzeuge vorgestellt und auf ihre Eignung in der multimedial gestützten Lehre untersucht. Man kann grob zwei Gruppen von Werkzeugen unterscheiden: Einerseits Komplettsysteme, bei denen eine Firma sowohl Hardware als auch meist proprietäre Software liefert, andererseits reine Software-Werkzeuge, die teils kommerziell, teils frei erhältlich sind. Die Anwendungen werden hinsichtlich ihrer Kompatibilität mit den in den Abschnitten 2.2 und 2.3 beschriebenen Konferenz-Standards und Medienformaten geprüft. Ebenso wird die Übertragungsart und -qualität, der Funktionsumfang, die Bedienbarkeit, der Preis und die Eignung in multimedialen Lehrveranstaltungen untersucht.

### 3.1 PictureTel

Die Firma PictureTel<sup>1</sup> bietet kommerzielle Komplettlösungen von etwa 2.000 DM bis 100.000 DM für Konferenzen im LAN (TCP/IP, Novell) oder über ISDN an. So besteht das Produkt "Swift Site" (Preis für Swift Site SS700-3 circa 15.000 DM zuzüglich Mehrwertsteuer, Stand Januar 1999, Mohr Telekommunikationssysteme) aus einem portablen Gerät, das für Videokonferenzen lediglich einen ISDN-Anschluß und einen Fernseher benötigt [55] (siehe Abbildung 3.1), per Fernbedienung gesteuert wird und hauptsächlich für Konferenzen zwischen zwei Standpunkten gedacht ist. Das "Live200" Paket (Preis für Live200p-520-4X-Y für einen Arbeitsplatz circa 4.800 DM zuzüglich Mehrwertsteuer, Stand Januar 1999, Mohr Telekommunikationssysteme) beinhaltet eine PCI Karte mit einer ISDN Schnittstelle zum Dekodieren von Audio und Video, einer Desktop Kamera, einem Headset, Lautsprechern, Mikrofon und Kabeln, und der Live200 Software für Windows 95/NT auf CD-ROM, mit der Konferenzen zwischen mehreren Standorten möglich sind [53]. Das Haupteinsatzgebiet sind dabei unternehmensweite Konferenzschaltungen.

Die "PictureTel 330" NetConference Multipoint Server Software (Preis für NC330 24-Port Server Software circa 41.100 DM zuzüglich Mehrwertsteuer, Stand Januar 1999,

---

<sup>1</sup>PictureTel Corporation, <http://www.picturetel.com>



Abbildung 3.1: PictureTel "SwiftSite"

---

**Vor- und Nachteile von PictureTel**


---

- + leicht bedienbare Komplettlösungen (Hardware und Software)
  - + verwendet T.120 Protokoll für Multipoint Data Conferencing
  - + verwendet H.323 Protokoll (H.261 Video, G.711/G.723 Audio)
- 
- läuft nur unter Windows 95/NT
  - teuer
  - bietet keine spezielle Unterstützung für Lehrveranstaltungen
  - nur sinnvoll bei Kommunikation zwischen zwei oder drei Standorten
- 

Tabelle 3.1: Vor- und Nachteile von PictureTel

Mohr Telekommunikationssysteme), die die Funktionalität eines Session Directories besitzt, läuft auf einem Windows NT Rechner und kann Konferenzen koordinieren, die im lokalen Netz zwischen "PictureTel LiveLAN" Multimedia-Rechnern stattfinden, über das "LiveGateway" und ISDN oder das Internet mit weiteren PictureTel Rechnern oder beliebigen H.323/T.120-kompatiblen Systemen verbunden sind. Mit der "LiveShare" Software (empfohlener Verkaufspreis für circa 550 DM zuzüglich Mehrwertsteuer, Stand Januar 1999, Mohr Telekommunikationssysteme) ist Application Sharing und Datenübertragung (T.120 vorbereitet) zwischen Windows 95 Rechnern möglich. Außerdem bietet PictureTel passendes Zubehör für Videokonferenzen von Kameras über Monitore, Tischmikrofone bis zu Fernbedienungen an.

**Fazit**

PictureTel Produkte decken fast alle Wünsche im Bereich Videokonferenzen, sowohl auf dem Hardware- als auch auf dem Software-Sektor, professionell ab. Sie sind allerdings hauptsächlich für die unternehmensinterne Kommunikation mit zwei oder maximal drei

Standpunkten über ISDN gedacht und auf Grund der hohen Anschaffungskosten außerdem nur für größere Unternehmen erschwinglich. Für Lehrveranstaltungen gibt es neben der reinen Audio/Videoübertragung keine spezielle Unterstützung.

## 3.2 VTEL

Die Firma VTEL<sup>2</sup> ist nach eigener Aussage Marktführer auf dem Gebiet Data und Video Conferencing in Amerika und bietet eine breite Palette von Komplettlösungen im LAN und über ISDN, seit neuestem auch über IP und ATM an. Das Angebot reicht von SmartStation Desktopsystemen, die aus einer kombinierten ISDN/PCI Codec Karte, einer Desktop Kamera mit Mikrofon und der VTEL AppsView Videokonferenz-Software bestehen (Preis für SmartStation 128 ISDN/IP für einen Arbeitsplatz circa 2.500 DM zuzüglich Mehrwertsteuer, Stand Februar 1999, VTEL Vertrieb Baden-Württemberg) über Komplettgeräte für Arbeitsgruppen, die aus einem Hochleistungs-Pentium PC mit Multimedia- und ISDN-Ausstattung, Windows 95/NT, einem Monitor und einem Rollwagen bestehen (Preis für WG 500 WorkGroup Conferencing System circa 20.000 DM zuzüglich Mehrwertsteuer, Stand Februar 1999, VTEL Vertrieb Baden-Württemberg), siehe Abbildung 3.2, bis hin zu Raumsystemen für die videokonferenzunterstützte Gruppenarbeit, die zusätzlich mit der PenPal Graphics Anwendung für Folienpräsentationen im Microsoft PowerPoint und JPEG Format und für Annotationen ausgestattet sind und eine Übertragungsgeschwindigkeit bis 128 kbps erlauben (Preis für TeamConferencing System TC1000 circa 35.000 DM zuzüglich Mehrwertsteuer, Stand Februar 1999, VTEL Vertrieb Baden-Württemberg).

Am obersten Ende der Preisskala stehen Raumsysteme, die auch in Vortragssälen einsetzbar sind, eine Übertragungsrate bis zu 1536 Kbps und eine Framerate bis zu 30 fps ermöglichen, über zwei Bildschirme und eine extra Dokumentenkamera und die SmartView Videokonferenz-Software verfügen (Preis für LeadershipConferencing System LC5000 circa 100.000 DM zuzüglich Mehrwertsteuer, Stand Februar 1999, VTEL Vertrieb Baden-Württemberg).

Alle Systeme können mit dem SMART Board um ein "echtes" Whiteboard ergänzt werden, auf dessen Touchscreen mit speziellen Stiften in verschiedenen Farben gemalt, allerdings nichts abgespeichert werden kann [74] (Preis für SMART Board circa 10.000 DM zuzüglich Mehrwertsteuer, Stand Februar 1999, VTEL Vertrieb Baden-Württemberg). Laut VTEL werden die Produkte sowohl in der unternehmensweiten Kommunikation als auch im Bildungswesen gerne eingesetzt, einerseits aufgrund der einfachen Handhabung der Fertiglösung und andererseits aufgrund der Kompatibilität mit anderen ISDN-Konferenzsystemen [73].

---

<sup>2</sup>VTEL Corporation, <http://www.vtel.com>



Abbildung 3.2: VTEL Workgroup System WG500

---

**Vor- und Nachteile von VTEL Produkten**


---

- + benutzt T.120 Protokoll für Multipoint Data Conferencing
  - + benutzt H.323 Protokoll (H.261 Video, G.711/G.723 Audio)
  - + einfache Benutzung der Komplettlösungen
- 
- läuft nur mit proprietärer Hardware oder unter Windows 95/NT
  - keine spezielle Unterstützung für Lehrveranstaltungen
  - nur sinnvoll bei Kommunikation zwischen zwei oder drei Standorten
  - teuer
- 

Tabelle 3.2: Vor- und Nachteile von VTEL Produkten

**Fazit**

VTEL Produkte decken den Videokonferenz-Sektor durch ihre breite Palette an professioneller Konferenz-Hardware und -Software weitgehend ab. Leider wird aber entweder proprietäre Hardware eingesetzt oder Software, die zwar konform zu den ISDN-Konferenzstandards ist, aber ausschließlich unter Windows 95/NT läuft. Auf Grund der hohen Kosten sind die VTEL Produkte auch nur für größere Unternehmen erschwinglich. Für Lehrveranstaltungen gibt es keine spezielle Unterstützung. Insbesondere sind die Produkte sowohl von der Hardware als auch von der Handhabbarkeit der Software auf Konferenzen mit zwei oder maximal drei Standorten ausgelegt; Lehrveranstaltungen mit großer Teilnehmerzahl müssen daher mit einem System pro Standort auskommen, was die Audio-Interaktion und die Interaktion bei der Datenkonferenz erheblich einschränkt und praktisch nur bei einer Ein-Weg-Kommunikation von Dozent zu Studenten sinnvoll ist.

### 3.3 Cisco IP/TV

Cisco Systems<sup>3</sup> ist weltweit führend auf dem Gebiet der Netzwerk-Hardware und -software. Cisco bietet Unternehmen und akademischen Einrichtungen mit der IP/TV Client-Server Software in der aktuellen Version 2.0 (April 1998) eine umfassende Lösung für das Senden und Empfangen von Videodaten im lokalen Netzwerk unter Windows 95/NT 4.0. Dabei kann es sich um Broadcast-Sendungen oder Videokonferenzen handeln, die bei großer Zuschauerzahl bandbreitenfreundlich per RTP/RTSP über Multicast übertragen werden, oder um Video on Demand. Aufgrund der guten Skalierbarkeit und der hohen Qualität der Videoübertragung sieht Cisco die Haupteinsatzgebiete in den Bereichen Firmenkommunikation und Angestellten-Fortbildung im großen Rahmen, und Business TV.

Die IP/TV Software besteht aus den drei Komponenten "Content Manager", "Server" und "Viewer", die entweder als eigenständige Programme laufen oder als JavaScript Programme und Netscape Plug-Ins auch einfach über den Webbrowser bedienbar sind (nur Content Manager und Viewer). Mit dem Manager (Preis circa 2800 \$, US Educational Price, Florida State Contract, Stand Februar 1999) können Programme angekündigt, Benutzer verwaltet und Server konfiguriert werden. Der Manager kann auch selbständig Daten, die zu einem bestimmten Zeitpunkt abgespielt werden sollen, per FTP auf andere Server laden, und zwar zu Zeiten, in denen das Netzwerk nicht so stark belastet ist. Der Server (Preis circa 2450 \$, US Educational Price, Florida State Contract, Stand Februar 1999) ist für die Aufnahme, das Speichern und die Übertragung von Live-Video oder gespeicherten Videodaten zuständig. Bei Live-Video wird der ISDN-Standard H.261 unterstützt. Bei Verwendung von RTP ist IP/TV dann sogar interoperabel mit den Mbone Tools "vic" und "vat" (siehe Abschnitt 3.7 weiter unten). Die Übertragung von gespeicherten Videodaten kann in hoher Auflösung in den üblichen Videokodierungen MPEG, AVI oder QuickTime MOV erfolgen. Der Viewer (Preis für Starter Kit für 20 Lizenzen circa 4550 \$, US Educational Price, Florida State Contract, Stand Februar 1999) öffnet für jedes Programm ein eigenes skalierbares Videofenster mit einfachen Audio/Video-Steurelementen. Bei Video on Demand ist die Kontrolle des Videos ähnlich wie bei einem Videoplayer mit Play/Stop/Pause/Forward/Rewind möglich. Bei Live-Video kann ein Vortragender neben seinem Videobild auch Bilder seiner Folien senden (siehe Abbildung 3.3), allerdings ohne Annotationsmöglichkeit. Ein Minimum von Interaktivität zwischen Zuschauern und Vortragendem ist dadurch gegeben, daß jeder Zuschauer im integrierten "Question Manager" eine textbasierte Frage stellen kann.

Mit der IP/TV Software "Stream Watch" (Preis circa 700 \$, US Educational Price, Florida State Contract, Stand Februar 1999) ist es möglich, Anzahl, Namen und Zeiten der Zuschauer und die Netzbelastung für anschließende Analysen zu speichern und als Tabelle darzustellen [9]. Außerdem können Sendungen verschlüsselt werden, um nur berechtigten Zuschauern die Teilnahme zu gewähren.

---

<sup>3</sup>Cisco Systems, Inc., <http://www.cisco.com>



---

**Vor- und Nachteile von Cisco IP/TV**


---

- + unterstützt Multicast und ist interoperabel mit "vic" und "vat"
  - + einfache Bedienung, da nur wenige Funktionen
  - + einfache Installation und Bedienung der Komplettsysteme
  - + Folienübertragung und textbasierte Fragen möglich
- 
- läuft nur unter Windows und spezieller Cisco Hardware
  - nicht frei erhältlich
  - nur ein Videobild pro Sendung
  - keine Annotationsmöglichkeit auf den Folien
  - keine spezielle Unterstützung für Lehrveranstaltungen
- 

Tabelle 3.3: Vor- und Nachteile von Cisco IP/TV

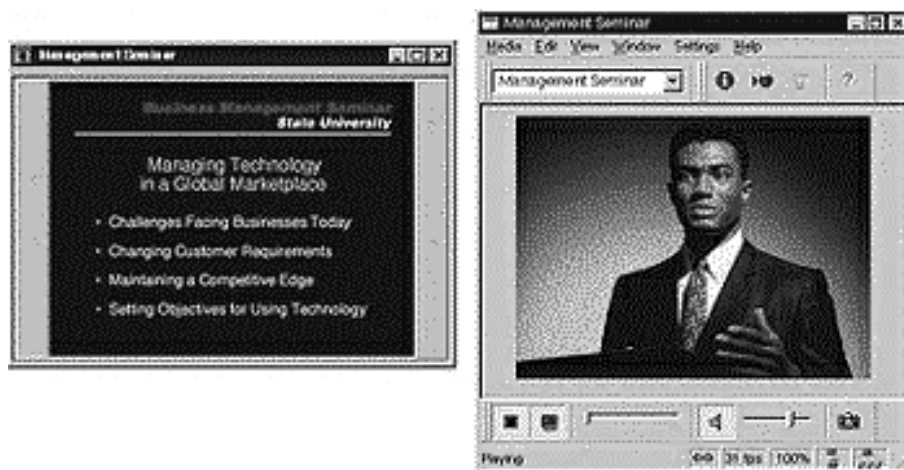


Abbildung 3.3: Cisco IP/TV

Passend zur Software vertreibt Cisco auch spezielle Hardware wie den IP/TV 3410 Control Server (Preis circa 7000 \$, US Educational Price, Florida State Contract, Stand Februar 1999), den IP/TV 3420 Broadcast Server (Preis circa 15400 \$, US Educational Price, Florida State Contract, Stand Februar 1999) und den IP/TV30 Archive Server (Preis circa 16800 \$, US Educational Price, Florida State Contract, Stand Februar 1999), auf denen die IP/TV Software auf Cisco IOS vorinstalliert ist und so eine unkomplizierte Komplettlösung für Netzwerk-Videoubertragungen darstellt [7].

## Fazit

Die Philosophie der IP/TV Software mit der Einteilung in Manager-, Server- und Viewer-Komponente und deren einfache Bedienung klingen überzeugend, ebenso die Möglichkeit

der Interoperabilität mit den MBone Tools "vic" und "vat" bei der Übertragung über RTP, vgl. Tabelle 3.3. Nachteilig ist jedoch, daß die nicht frei erhältliche IP/TV Software auf Windows- und Cisco-Plattformen beschränkt ist, und daß außer einfachsten Textbotschaften keine Interaktivität möglich und insbesondere keine spezielle Unterstützung für Lehrveranstaltungen vorgesehen ist.

### 3.4 Intel ProShare

Intel bietet mit dem ProShare Video System<sup>4</sup> eine Komplettlösung für Multipoint-Video-Konferenzen unter Windows an. Das Video System 500 (Preis circa 949 \$, Kernelsoftware, Stand Februar 1999) besteht aus einer PCI ISDN-Audio-Videokarte mit entsprechenden Anschlüssen, einem Headset, einer Desktop Kamera, der Audio-Videokonferenz-Software "ProShare Conferencing" und der Datenkonferenz-Software "Premier" beziehungsweise Microsoft Netmeeting (Preis circa 2500 DM zuzüglich Mehrwertsteuer pro Arbeitsplatz). Über ISDN können damit H.323/T.120-kompatible Audio-Video-Daten-Konferenzen abgehalten werden, im LAN über TCP/IP lediglich T.120-Datenkonferenzen [31]. Für Audio-Verbindungen im LAN empfiehlt Intel die gleichzeitige Benutzung des Telefons. Neben den recht hohen Anschaffungskosten des Systems entstehen beim Einsatz über ISDN weitere Kosten für Miete und Nutzung der ISDN-Leitung.

Die "ProShare Conferencing" Software enthält den "Konferenz-Manager", mit dem alle ProShare Anwendungen gestartet und Konferenzen verwaltet werden können, siehe Abbildung 3.4 links. So können die Einstellungen eines Benutzers, seine "Visitenkarte", die auch ein Bild von ihm beinhalten können, bearbeitet, die Visitenkarten anderer Teilnehmer in einem Adreßbuch eingesehen, andere Teilnehmer angerufen und Konferenzen in einem Planer ähnlich wie bei einem Session Directory Programm vorgemerkt werden. Neben dem Fenster des Konferenz-Manager wird standardmäßig ein weiteres Fenster für die Anzeige und Steuerung von Audio und Video geöffnet. Dabei werden zwei Videobilder dargestellt: Das eigene und das der Gegenstelle, siehe Abbildung 3.4 rechts. Bei einer Multipoint-Verbindung wird diejenige Gegenstelle angezeigt, deren Teilnehmer am lautesten spricht. Vom Videobild können Snapshots gemacht werden, eine Aufzeichnung von Audio und Video ist jedoch nicht möglich. ProShare unterstützt die Audio-Codecs G.711, G.723 und G.728 und die Audio-Codecs H.261 und H.263 im FCIF/QCIF Format. Die PCI Karte ist bezüglich Audio fullduplex-fähig und besitzt einen eingebauten Echo Cancellor.

Der "Notizblock" ist das Whiteboard der "Premier" Datenkonferenz. Die Teilnehmer können Seiten anlegen, Bilder in einer Vielzahl von Formaten und Text über Menüs, aus der Windows Zwischenablage oder per Drag and Drop hineinladen und auf einer Seite (auch private) Notizen und Zeichnungen anfertigen. Dateien anderer Windows Programme wie Microsoft PowerPoint können als Bilder, die der Teilnehmer mit Hilfe des Pro-

---

<sup>4</sup>Intel ProShare Technology, <http://www.intel.com/proshare/>



Abbildung 3.4: Intel ProShare Konferenz-Manager und Videofenster

Share Snapshot-Tools anfertigt, ebenfalls angezeigt werden. Die Seiten können außerdem neu angeordnet oder mit denen der anderen Teilnehmer synchronisiert werden. Über die "Turnussteuerung" kann optional ein einziger Teilnehmer die alleinige Kontrolle über den Notizblock übernehmen oder sie wieder freigeben. Daneben ist das Abspeichern und Ausdrucken einzelner oder aller Seiten des Notizblocks möglich. (Seit neuestem ist Microsoft NetMeeting integriert und deckt jetzt den Datenkonferenzteil ab.)

Die Datenanwendung "Gemeinsamer Zugriff" erlaubt Konferenzteilnehmern die gemeinsame Bearbeitung von Dateien und Nutzung von Anwendungen, die dabei nur auf einem der beteiligten Rechner installiert sein müssen. Der Zugriff wird dabei ähnlich wie beim Notizblock über die Turnussteuerung geregelt. Die Anwendung "Bildbearbeitung" erlaubt das gemeinsame Betrachten und Sortieren von Bildern hoher Auflösung in den Formaten JPG, GIF, BMP/DIB (Windows Bitmap), TIF (Tag Image File), PCX (PC Paintbrush) und TGA (Targa) in sogenannten Alben. Die Alben und Bilder können synchron oder privat betrachtet, gespeichert oder in den Notizblock eingefügt werden. Mit der Datenanwendung "Dateiübertragung" gemäß Standard T.127 können Dateien beaufsichtigt (mit Bestätigung) oder unbeaufsichtigt (zu einem Standardverzeichnis) an beliebige Teilnehmer verschickt werden.

Aus Erfahrungsberichten [40] geht hervor, daß die Bedienung der Software insgesamt lediglich zufriedenstellend ist. Die Aufteilung der einzelnen Teil-Anwendungen in eigene Fenster und mit unterschiedlichen Benutzerschnittstellen wurde als verwirrend empfunden. Die Audioqualität ließ außerdem bei einer Verbindung zwischen USA und Deutschland trotz ISDN-Leitung sehr zu wünschen übrig.

<b>Vor- und Nachteile von Intel ProShare</b>
+ benutzt T.120 Protokoll für Multipoint Data Conferencing
+ benutzt H.323 Protokoll (H.261 Video, G.711/G.723 Audio)
+ relativ einfache Benutzung der Komplettsysteme
– läuft nur unter Windows
– keine spezielle Unterstützung für Lehrveranstaltungen
– teuer
– nur zwei kleine Videobilder
– unübersichtliche Teil-Anwendungen in verschiedenen Fenstern

Tabelle 3.4: Vor- und Nachteile von Intel ProShare

**Fazit**

Das Haupteinsatzgebiet der Intel ProShare Komplettlösungen dürfte auf Grund des hohen Preises bei gleichzeitig nur einfacher Funktionalität die behelfsweise unternehmensinterne Konferenzschaltungen sein. Unterstützung für Lehrveranstaltungen ist nicht enthalten. Aus den in Tabelle 3.4 aufgeführten Vor- und Nachteilen, wiegt bei letzteren besonders schwer der auf Windows 95/NT und ISDN eingeschränkte Einsatz.

**3.5 Microsoft NetMeeting**

Das von Microsoft für Windows 95/NT entwickelte Kommunikations-Werkzeug NetMeeting ist aktuell in der Version 2.1 (Stand 1998) verfügbar und frei über WWW erhältlich<sup>5</sup>. Mit NetMeeting können Audio- Video- und Chatkonferenzen durchgeführt, Programme oder ein Whiteboard gemeinsam benutzt und Dateien übertragen werden [50], siehe Abbildung 3.5. NetMeeting enthält ILS, den "Internet Locator Server", eine LDAP ("Lightweight Directory Access Protocol") Schnittstelle zum einfachen Beitreten und Verlassen von NetMeeting Sitzungen.

**Fazit**

NetMeeting ist zwar frei erhältlich, der gravierendste Nachteil ist jedoch die ausschließliche Nutzung unter Windows 95/NT. Audio und Video werden zwar in den gängigen Standards H.261 und G.711 über RTP übertragen, das Whiteboard benutzt jedoch ein proprietäres Format und besitzt keine Funktionen zur Unterstützung von Lehrveranstaltungen. In Experimenten hat sich zudem herausgestellt, daß Konferenzen zwischen mehr als zwei Standorten nicht praktikabel sind.

<sup>5</sup>Microsoft NetMeeting, <http://www.microsoft.com/netmeeting/>

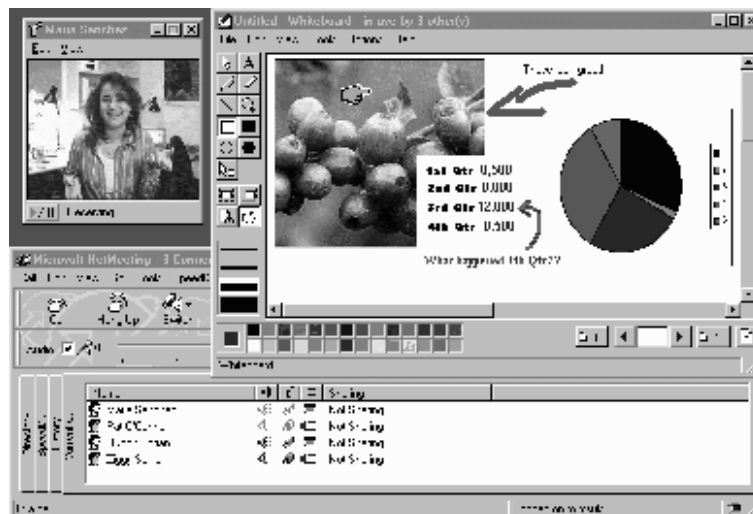


Abbildung 3.5: Microsoft NetMeeting

---

**Vor- und Nachteile von Microsoft NetMeeting**


---

- + frei erhältlich
  - + benutzt T.120 Protokoll für Multipoint Data Conferencing
  - + benutzt H.323 Protokoll (H.261 Video, G.711/G.723 Audio)
  - + benutzt RTP/RTCP
  - + private Audio- und Chatverbindungen innerhalb einer Session möglich
  - + einfacher Aufruf von Windows/Internet Explorer aus
  - + Windows Application Sharing möglich
  - + objekt-orientiertes Whiteboard
- 
- läuft nur unter Windows
  - lediglich Abspeichern von einzelnen Videobildern möglich
  - keine spezielle Unterstützung für Lehrveranstaltungen
- 

Tabelle 3.5: Vor- und Nachteile von Microsoft NetMeeting

### 3.6 Sun ShowMe TV

ShowMe TV ist ein von Sun Microsystems entwickeltes Werkzeug für Audio- und Videokommunikation auf Sun Workstations unter Solaris. Es ist aktuell in der Version 1.2.1 erhältlich und kann kostenlos im Netz heruntergeladen werden<sup>6</sup>. ShowMe TV besteht aus zwei Komponenten: Der "Transmitter" ist ein Server für Audio- und Videodaten. Er ist für das Senden und die Konfiguration von Audio und Video mit einstellbarer Sendezeit, Framerate und Format, und die Netzwerk-Konfiguration mit den einstellbarer Bandbreite, Multicast- und Broadcast-Verteilung zuständig und zeigt den Sendezeitplan in einer Art Fernsehzeitschrift mit einer Programmbeschreibung und die Netzwerkeigenschaften als Statistiken an, beispielsweise die verbrauchte Bandbreite oder die Anzahl der Benutzer. Neben Audio und Video kann auch gleichzeitig Text gesendet werden, der dann über dem

<sup>6</sup>ShowMe Products, <http://www.sun.com/products-n-solutions/sw/ShowMe/>

Videofenster angezeigt wird. Die Benutzer von ShowMe TV werden über das Transmitter-Adressbuch verwaltet, das offene und geschlossene Sendungen erlaubt, wobei ein Email-Programm integriert ist, so daß beispielsweise alle Zuschauer automatisch als Adressaten eingetragen werden können.

Die zweite Komponente von ShowMe TV ist der "Receiver Client", siehe Abbildung 3.6. Mit ihr ist das Einstellen der Lautstärke, von Farbe und Kontrast, der Empfang und auf Wunsch das automatische Aufnehmen von Ton, Bildern und Filmsequenzen des vom Transmitter gesendeten oder lokal gespeicherten Material möglich. Eine besondere Spezialität ist die simultane Darstellung eines zweiten Kanals beziehungsweise kleinen Bildes innerhalb des großen Hauptbildes ("Picture in Picture", PiP). Die Standardbildgröße ist 320x240 (SIF Standard), kann aber auch halbiert oder verdoppelt werden. Vom Receiver aus hat der Benutzer Zugriff auf die Programmzeitschrift. Durch einen einzigen Klick kann eine Mail an den Initiator des Programms geschrieben werden. Eine weitergehende Unterstützung speziell für Lehrveranstaltungen, insbesondere ein Whiteboard, ist jedoch nicht vorhanden.



Abbildung 3.6: Suns ShowMe TV

Die Netzwerkbelastung wird von Sun mit etwa 32 - 176 Kbps für Audio Broadcast und etwa 300 - 900 Kbps für Video Broadcast bei einer Auflösung von 320x240 Pixel und 10-15 fps angegeben, und ist mit diesen recht hohen Werten hauptsächlich für die Übertragung in lokalen Netzen geeignet.

## Fazit

Mit Suns ShowMe TV sind zwar verteilte Audio-Video-Konferenzen über Multicast gut durchführbar, auf Grund des hohen Datenaufkommens allerdings nur in lokalen Netzen.

<b>Vor- und Nachteile von Sun ShowMe TV</b>
+ frei erhältlich
+ Datenverteilung über Multicast möglich
+ unterstützt viele Audio-/Videoformate (PCM, AU, WAV, MPEG, QuickTime/AVI: CellB, JPEG, H.261)
+ Abspeichern von Ton, Bildern und Film ist möglich
+ Benutzerverwaltung mit automatischer Email
+ minimale Text-Datenkonferenz möglich
– läuft nur auf Sun Workstations unter Solaris
– proprietäres Format, inkompatibel zu bestehenden Systemen
– (noch) kein Whiteboard integriert
– hohes Datenaufkommen durch Datenübertragung über TCP/IP

Tabelle 3.6: Vor- und Nachteile von Sun ShowMe TV

Nachteilig ist, daß ShowMe TV nur unter Solaris läuft und durch sein proprietäres Format nicht zu anderen Konferenzwerkzeugen kompatibel ist. Zudem ist mit ShowMe TV keine Datenkonferenz möglich, lediglich das Versenden kurzer Textbotschaften.

### 3.7 MBone Tools

Als MBone Tools bezeichnet man die Menge der frei verfügbaren Werkzeuge zum Abhalten von Multicast-Konferenzen im Internet. Die Werkzeuge stammen von verschiedenen Universitäten und Forschungseinrichtungen, sind aber weitgehend interoperabel. Die MBone Tools sind meist Referenzimplementationen zu den von den jeweiligen Forschungseinrichtungen entwickelten Internet Standards wie RTP oder SAP. Im folgenden werden lediglich die gebräuchlichsten MBone Tools vorgestellt.

#### Session Directory Programme

Mit Session Directory Programmen werden Informationen zu laufenden oder zukünftigen Konferenzen in einem Planer dargestellt, oder selber Sitzungsankündigungen erstellt. Meist ist auch die Einladung von Teilnehmern möglich. Fast immer werden die Standardprotokolle SAP, SDP oder SIP eingesetzt.

Das Session Directory Programm "sdr" ist die Weiterentwicklung des von Van Jacobson (Lawrence Berkeley National Laboratory, LBNL) stammenden "sd", das die Referenzimplementierung der oben genannten Protokolle darstellt. "sdr" wurde im Rahmen der MICE und MERCI Projekte am University College London (UCL) von Mark Handley in Zusammenarbeit mit Bill Fenner (Xerox) und Van Jacobson (LBL) implementiert und ist aktuell in der Version 2.5a5 (Stand August 1998) für verschiedene Plattformen wie Windows und verschiedene Unix Derivate erhältlich<sup>7</sup>. Mit "sdr" ist das Ankündigen und Verschlüsseln

<sup>7</sup>Session Directory, <http://www-mice.cs.ucl.ac.uk/multimedia/software/sdr/>

von Sitzungen und Einladen von Teilnehmern möglich. Durch den Doppelklick auf eine im Hauptfenster aufgelistete Sitzung erhält man ein Informationsfenster, aus dem die benötigten, beliebig wählbaren, Audio-, Video- und Daten-Werkzeuge automatisch gestartet werden können. Ist im Informationsfenster eine URL angegeben, so kann die WWW-Seite mit dem integrierten oder einem beliebigen externen Browser betrachtet werden. Der integrierte "Kalender" bietet einen zeitlichen Überblick über die an einem Tag laufenden Sitzungen. Beim Beenden von "sdr" werden die empfangenen Ankündigungen in Dateien geschrieben. Beim nächsten Aufruf werden die in diesem Cache vorhandenen Ankündigungen dann automatisch wieder angezeigt. "sdr" läuft stabil auf allen Plattformen und ist über die in Tcl/Tk geschriebene Oberfläche relativ intuitiv bedienbar, siehe Abbildung 3.7.

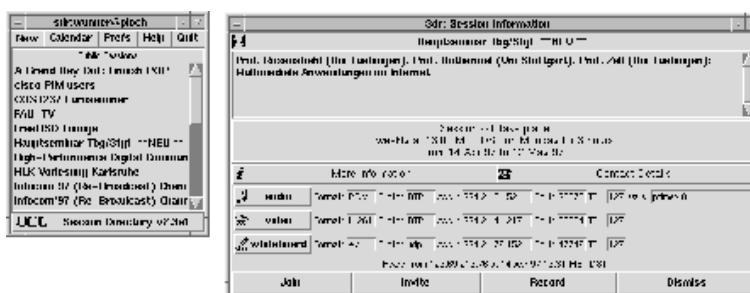


Abbildung 3.7: MBone Session Directory Programm "sdr" mit Sitzungs-Informationsfenster

Weitere MBone Session Directory Programme sind beispielsweise das an der University of California, Berkeley, entwickelte "mash-sdr"<sup>8</sup>, oder der Konferenz-Manager "ConfMan"<sup>9</sup> der Universität Hannover.

## Audio-Werkzeuge

Das zur Zeit am häufigsten für die Audiokommunikation eingesetzte MBone Werkzeug ist das Visual Audio Tool "vat"<sup>10</sup> des Lawrence Berkeley National Laboratory in der Version 4.0b2 (Stand Juni 1996), das Audiodaten über RTP versenden und empfangen kann. Der Grund für die Beliebtheit liegt trotz der fehlenden Weiterentwicklung und der niedrigen Audioqualität von 8 kHz, Mono, in der relativ leichten Bedienung und der überschaubaren Teilnehmerliste, siehe Abbildung 3.8. Eine Weiterentwicklung, die aber weitgehend die gleiche Funktionalität wie "vat" besitzt, ist "mash-vat"<sup>11</sup> vom selben Institut.

Sprechende Teilnehmer werden durch ein ausgefülltes Kästchen neben dem Namen markiert, das erst nach und nach verblaßt. Einzelne Teilnehmer können auch stummgeschaltet werden. Über einen Mausklick mit der linken Taste auf einen Teilnehmernamen erhält man

<sup>8</sup>MASH System and Software, <http://mash.cs.berkeley.edu/mash/software/>

<sup>9</sup>ConfMan Distribution, <http://www.rvs.uni-hannover.de/products/confman/index.html>

<sup>10</sup>LBNL Audio Conferencing Tool (vat), <http://www-nrg.ee.lbl.gov/vat/>

<sup>11</sup>MASH System and Software, <http://mash.cs.berkeley.edu/mash/software/>



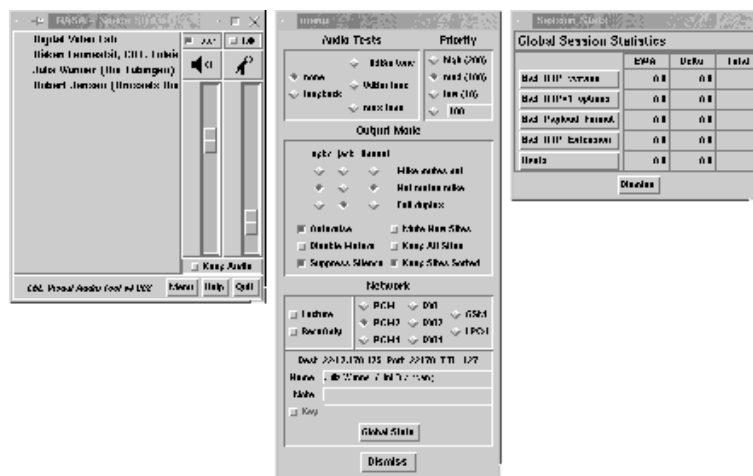


Abbildung 3.8: MBone Audio-Werkzeug "vat" mit Audiokontroll- und Statistikfenster

zusätzliche Informationen wie RTCP Name, Tool und Note, und RTP und Dekoder Statistiken. Man kann zwischen drei Ausgangsgeräten, standardmäßig Lautsprecher, Kopfhörer und Line Out, und zwei Eingangsgeräten, standardmäßig Mikrophon und Line In, jeweils mit Lautstärkeregelung, wählen. "vat" unterstützt Full Duplex, allerdings nur unter Solaris. Technischere Informationen und Kontrollmöglichkeiten verbergen sich für geübtere Benutzer hinter dem "Menu" Button. Dort kann man unter anderem eine der vielen möglichen Audiokodierungen wählen: Standardmäßig PCM2, aber auch PCM/4, DVI/2/4, GSM und LPC4.

Das Robust Audio Tool "rat"<sup>12</sup> des UCL ist zu "vat" kompatibel, besitzt aber eine schlecht aufgebaute Benutzeroberfläche. Dafür kann es ein redundantes Audiosignal benutzen, was besonders bei schlechten Verbindungen noch eine verständliche Übertragung gewährleistet. "rat" ist aktuell in der Version 3.0.24 (Stand April 1998) für viele Plattformen wie Windows und diverse Unix Derivate erhältlich.

Das Audio-Werkzeug "fphone" (Free Phone) von INRIA ist ebenfalls zu "vat" und "rat" kompatibel, bietet aber neben einem redundanten Audiosignal die Möglichkeit in verschiedenen Audio-Qualitäten bis hin zu 48 kHz, Stereo, zu senden. Dabei können verschiedene Redundanzstufen, Frame-Größen und Bitraten eingestellt werden. "fphone" ist in der aktuellen Version 3.5b3 (Stand Dezember 1997) für verschiedene Plattformen wie verschiedene Unix Derivate und Windows95 erhältlich. Es besitzt neben Mikrophon und Line In zusätzlich einen Eingang für ein als CD-Player genutztes CD-ROM-Laufwerk.

<sup>12</sup>Robust Audio Tool, <http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/>

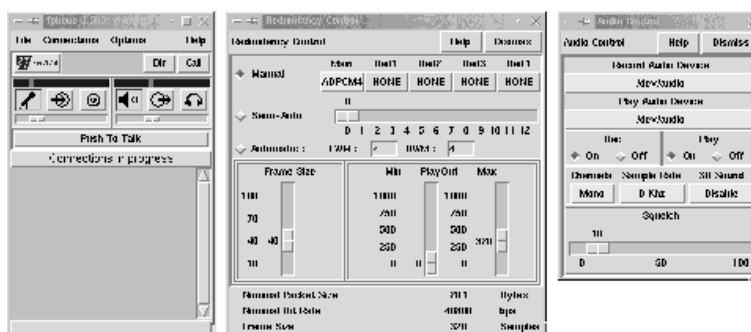


Abbildung 3.9: MBone Audio-Werkzeug "fphone" mit Redundanz- und Audiokontrollfenster

## Video-Werkzeuge

Zum Übertragen von Video wird nahezu ausschließlich das Videoconferencing Tool "vic"<sup>13</sup>, das wie "vat" vom LBNL stammt, eingesetzt. Aktuell ist es in der Version 2.8 (Stand Juli 1997) für viele Plattformen und Videokarten erhältlich. Eine Weiterentwicklung existiert in Form des "mash-vic"<sup>14</sup> der UCB.

"vic" besitzt wie "vat" eine Tcl/Tk-Oberfläche. Im Hauptfenster sind die sendenden Teilnehmer mit kleinen Video-Vorschaubildern und Informationen wie Name, Frame-Rate und Bit-Rate aufgelistet, siehe Abbildung 3.10. Durch einen Klick auf den "Info" Button erhält man wie bei "vat" weitere Informationen wie RTP Name, Tool und Note. Durch einen Doppelklick auf das kleine Vorschaubild erhält man ein größeres Bild im QCIF Format. Man kann das Bild aber auch in anderen gebräuchlichen Formaten wie PAL/16 oder NTSC darstellen. Eine Besonderheit von "vic" ist die "Voice-Switched" Einstellung. Damit wird in einem großen Videobild immer derjenige Teilnehmer angezeigt, der gerade spricht. Daneben ist auch die zufällig wechselnde Anzeige aller Video sendenden Teilnehmer in einem größeren Videobild anwählbar. Das Senden von Video ist etwas versteckt über den "Transmit" Button im "Menu" Fenster möglich. Dort kann man Einstellungen wie die maximale Frame-Rate, die Qualität, Farbe/Schwarz-Weiß, den Video-Codec, standardmäßig H.261, CellB und Motion JPEG, und den Videoeingang vornehmen oder sich die Namen aller Sender und Zuschauer anzeigen lassen.

Die einzige Alternative zu "vic" ist zur Zeit das nicht mehr weiterentwickelte Network Video Tool "nv"<sup>15</sup>, das wie "sdr" im Rahmen des MICE Projekts entstand. Es besitzt nur einen Teil der Funktionen von "vic" und eine unübersichtliche Benutzeroberfläche, siehe Abbildung 3.11. Allerdings kann man mit "nv" nicht nur Eingänge von Videokameras, sondern auch den Bildschirminhalt übertragen.

<sup>13</sup>UCB/LBNL Video Conferencing Tool (vic), <http://www-nrg.ee.lbl.gov/vic/>

<sup>14</sup>MASH System and Software, <http://mash.cs.berkeley.edu/mash/software/>

<sup>15</sup>NV, <http://mice.ed.ac.uk/mice/archive/nv.html>



Kreis, Pfeil oder auch Text sein. Diese Zeichenbefehle (“Draw Commands”) werden über Multicast und UDP an alle Teilnehmer der Multicast Gruppe versandt.



Abbildung 3.12: Mbone Whiteboard “wb” und Kontrollfenster

Neben den Zeichenbefehlen werden noch zusätzliche Daten übertragen, die die zuverlässige Übertragung garantieren, beispielsweise “Sender Reports”, “Receiver Reports” und “Request” Pakete. Aus den verwendeten Algorithmen, die hauptsächlich auf der Benutzung von negativen Bestätigungen beruhen, hat sich SRM und später LRMP entwickelt [19],[47].

Da “wb” jedoch insgesamt nur sehr eingeschränkte Funktionalität bietet, Gruppenarbeit nicht direkt unterstützt und von Anfängern nur schwer zu bedienen ist, wurden an verschiedenen Universitäten Eigen- und Weiterentwicklungen konzipiert. Das fortgeschrittenste Werkzeug scheint das “Digital Lecture Board” (“dlb”) der Universität Mannheim zu sein. Es kann in der in der aktuellen Version 1.8b8 (Stand November 1998) frei heruntergeladen werden<sup>16</sup>. Es ist in C++ und die Oberfläche in Tcl/Tk geschrieben und läuft zur Zeit unter Linux, Solaris und Irix. Die Daten werden mit dem eigens entwickelten zuverlässigen Multicast Protokoll SMP [21], das an SRM angelehnt ist, übertragen. Anders als bei “wb” können verschiedene sogenannte Medienobjekte auf eine Seite geladen werden, also auch mehrere PostScript-Dokumente und zusätzlich Bilder im GIF, EPS, BMP, PPM und PCX Format. Daneben ist das Zoomen in einzelne Seiten möglich und die Einbettung zeitabhängiger Medien wie Audio und Video. Die Seiten werden intern in einem an SGML angelehnten Format gespeichert und können so auch aufgezeichnet oder offline vorbereitet werden. Neben den im “wb” vorhandenen grafischen Objekten sind in “dlb” auch gefüllte Polygone möglich. Die Benutzeroberfläche ist übersichtlich gestaltet und integriert das Kontrollfenster mit der Anzeige der Teilnehmer in das Hauptfenster, siehe Abbildung 3.13.

<sup>16</sup>Digital Lecture Board (dlb), <http://www.informatik.uni-mannheim.de/geyer/dlb/download.html>

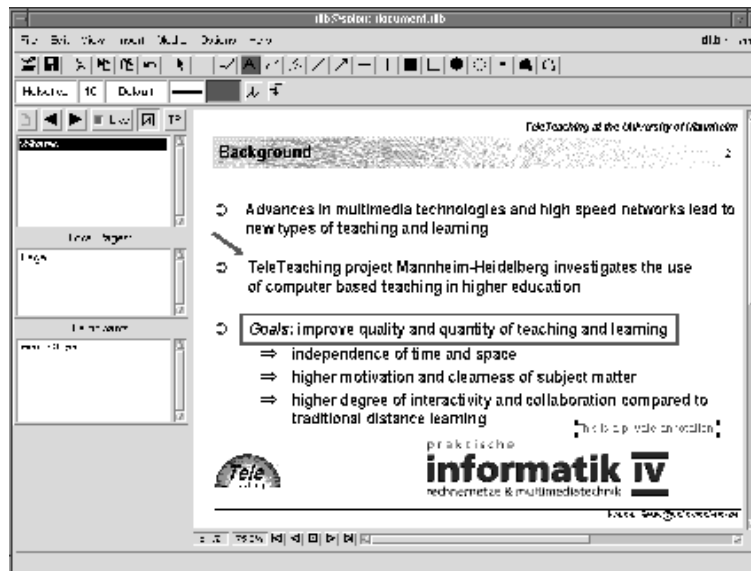


Abbildung 3.13: "Digital Lecture Board (dlb)"

"dlb" wurde speziell für synchrone Lehrveranstaltungen entwickelt und unterstützt daher einen Telepointer, Abstimmungen, Online Feedback und die Aufmerksamkeitsüberwachung. Daneben implementiert "dlb" verschiedene Mechanismen für die effiziente computergestützte Kommunikation, das sogenannte "Collaborative Services Model" [29]. Das Modell beinhaltet die Teilnehmerverwaltung in Gruppen und Untergruppen, die Zuweisung von Ressourcen an einzelne Teilnehmer, Gruppen und Sitzungen, die Unterstützung verschiedener Teilnehmerrollen und -privilegien wie Lehrer oder Studenten, und verschiedene Benutzungsmechanismen wie die implizite oder explizite Kontrolle der Sitzung. Es ist geplant, auch die Audio- und Videokonferenzwerkzeuge "vat" und "vic" in "dlb" und eine Schnittstelle zu einer Datenbank mit Lehrmaterialien herzustellen.

### Sonstige Werkzeuge

Neben den eigentlichen Konferenzwerkzeugen existieren einige MBone Tools, die einzelne kleinere Aufgaben wie das Aufnehmen und Speichern von Konferenzen ermöglichen. Die "RTP Tools" von Henning Schulzrinne sind mehrere frei verfügbare<sup>17</sup> kommandozeilenorientierte Programme, die zum Aufnehmen ("rtpdump") und Abspielen ("rtpplay"), Analysieren ("rtpmon") und Senden ("rtpsend") von RTP Datenströmen, sowohl über Unicast als auch Multicast. Zum Anzeigen der RTP Daten werden beliebige MBone Audio- und Videowerkzeuge benötigt.

Etwas komfortabler ist das am CERN entwickelte und ebenfalls frei erhältliche, komman-

<sup>17</sup>RTP: About RTP and the Audio-Video Transport Working Group, <http://www.cs.columbia.edu/~hgs/rtp/>



besteht aus den drei Komponenten "MVoD Server", der für das Benutzer- und Sitzungsmanagement zuständig ist, dem "MVoD Client", einer grafischen Benutzeroberfläche, siehe Abbildung 3.15, die ähnlich wie "wrtpVoD" funktioniert und in der linken Hälfte eine Art Session Directory Programm beinhaltet, und der "RTP DataPump" zum Aufnehmen, synchronisierten Abspielen und Verwalten der RTP Datenströme.

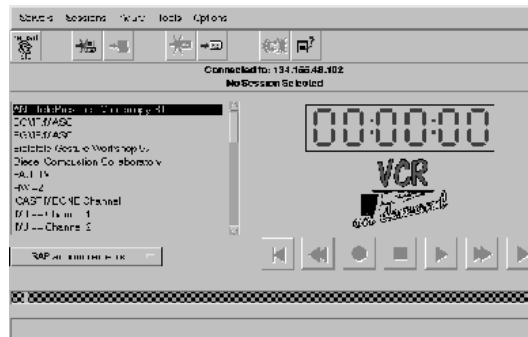


Abbildung 3.15: Aufzeichnen und Abspielen von RTP Sitzungen mit "MVoD"

## Fazit

Die MBone Tools sind für verschiedene, hauptsächlich Unix-Plattformen, frei erhältlich und unterstützen Standard-Hardwarekomponenten. Sie benutzen die Verteilung über Multicast und Echtzeitprotokolle wie RTP oder Eigenentwicklungen wie SMP. Auch das Aufzeichnen und Abspielen von Sitzungen ist möglich. Grundsätzlich sind die MBone Werkzeuge also sehr gut für verteilte Video- und Datenkonferenzen einsetzbar. Nachteilig ist jedoch, wie auch aus Tabelle 3.7 ersichtlich wird, daß für jedes Medium und auch zum Aufzeichnen und Abspielen andere Werkzeuge eingesetzt werden müssen. Bei einer synchronen Audio-Video-Datenkonferenz, die aufgezeichnet wird, sind dies vier verschiedene Werkzeuge, die aus insgesamt vier bis acht Einzel Fenstern bestehen, was äußerst unübersichtlich und benutzer-unfreundlich ist.

Die MBone Tools sind bis auf Ausnahmen nicht für eine Benutzung von Laien gedacht und haben oft sehr technische, verwirrende Benutzerschnittstellen. Erschwerend kommt hinzu, daß man bei der Kombination verschiedener Werkzeuge zwischen verschiedenen Bedienparadigma wechseln muß insgesamt hat man auch keine tatsächliche Plattformunabhängigkeit, da einzelne Werkzeuge nur für wenige Betriebssysteme erhältlich sind. Die MBone Tools sind zum Teil auch noch nicht ausgereift und stürzen oft ab [61]. Nur das Datenkonferenzwerkzeug "dlb" unterstützt Lehrveranstaltungen und die effiziente computerbasierte Gruppenkommunikation.

<b>Vor- und Nachteile der MBone Tools</b>
+ frei erhältlich
+ unterstützen viele Plattformen (Windows, Unix)
+ unterstützen gängige Audio- und Video-Hardware
+ unterstützen viele Audio-/Videoformate
+ benutzen RTP und zuverlässige Echtzeitprotokolle
+ effiziente Datenverteilung per Multicast
– viele einzelne Werkzeuge erschweren Bedienbarkeit
– Werkzeuge sind zum Teil zu "technisch" für Laien
– Werkzeuge sind zum Teil nicht ausgereift, stürzen oft ab
– nur in einem Fall Unterstützung für Lehrveranstaltungen

Tabelle 3.7: Vor- und Nachteile der MBone Tools



## Kapitel 4

### Fazit und Ausblick

In Kapitel 2 wurden die Grundlagen der Echtzeit-Kommunikation über das Internet, angefangen von Netzwerk-Architekturen, über die verschiedenen Medienformate bis hin zu Lehr- und Lernszenarien an Hochschulen vorgestellt. Dabei wurde deutlich, daß der Einsatz von Multimedia in der Lehre eine Orts- und Zeitungebundenheit, erhöhte Flexibilität und insgesamt eine qualitative Verbesserung oder zumindest Ergänzung der herkömmlichen Lehre darstellt. In Kapitel 3 wurden aktuelle Konferenz-Werkzeuge vorgestellt und hinsichtlich ihrer Eignung für den Einsatz in der multimedial gestützten Lehre untersucht.

Dabei wurde deutlich, daß es zwar etliche professionelle Software- und Hardware-Lösungen wie PictureTel, VTEL, IP/TV und ProShare gibt, diese aber teilweise mit proprietären Formaten arbeiten, nur auf Windows Plattformen laufen, die an Hochschulen nicht sehr verbreitet sind, und fast ausschließlich die Übertragung über ISDN unterstützen, was hohe laufende Kosten verursacht. Die Lösungen sind daneben hauptsächlich auf die Kommunikation innerhalb größerer Unternehmen ausgerichtet, was einerseits bedeutet, daß sie für öffentliche Ausbildungsstätten nahezu unerschwinglich sind. Andererseits haben die Produkte den eklatanten Nachteil, daß sie neben der reinen Audio-, Video- und einfachen Datenübertragung keine Unterstützung von Lehrveranstaltungen bieten, also beispielsweise die Gruppenverwaltung der Teilnehmer oder die Möglichkeit verschiedener Teilnehmerrollen und explizite oder implizite Sitzungskontrolle.

Es wurden daher auch frei erhältliche Werkzeuge wie NetMeeting und ShowMe TV untersucht. Diese sind aber ebenfalls auf bestimmte Plattformen wie Windows und Solaris beschränkt und bieten noch weniger Funktionalität als die professionellen Lösungen. Schließlich wurden die MBone Tools vorgestellt, die in großer Auswahl und für viele Plattformen, hauptsächlich Unix Derivate, frei erhältlich sind und eine IP-basierte Übertragung im Internet benutzen. Nach Untersuchung der aktuellen Konferenzwerkzeuge kommt man zu dem Schluß daß für Lehrveranstaltungen und die effiziente computerbasierte Kommunikation noch am ehesten eine Kombination verschiedener MBone Tools sinnvoll und praktikabel ist. Der Einsatz mehrerer Werkzeuge gleichzeitig in einer Sitzung ist jedoch äußerst benutzerunfreundlich, unübersichtlich und kommunikationshemmend. Auch sind viele MBone

Werkzeuge nur universitäre Testentwicklungen, die häufig noch nicht ausgereift sind.

Für multimedial unterstützte Lehrveranstaltungen existiert also noch kein Werkzeug, das flexibel wie die vorhandenen MBone Tools ist, aber alle Funktionalität in einem einzigen Werkzeug vereinigt. Dieses Werkzeug sollte auf möglichst vielen Plattformen lauffähig und leicht erweiterbar sein und eine intuitiv bedienbare Benutzeroberfläche besitzen. Lehrveranstaltungen sollten nicht nur innerhalb von Datenkonferenzen, sondern insgesamt besser unterstützt werden, beispielsweise durch die Unterstützung bei der Gruppenverwaltung und Sitzungskontrolle. Erste brauchbare Ansätze wurden im "dlb" MBone Tool verwirklicht, das jedoch keine Audio- und Video-Fähigkeit integriert. Ein universales Werkzeug für multimedial gestützte Lehrveranstaltungen wird daher noch benötigt.

# Literaturverzeichnis

- [1] Adobe Inc. Portable Document Format Reference Manual. 1996.  
<http://www.adobe.com/supportservice/devrelations/technotes.html>
- [2] Adobe Inc. Postscript Language, Document Structuring, Conventions Specification. 1992.  
<http://www.adobe.com/supportservice/devrelations/technotes.html>
- [3] A. Badach, E. Hoffmann, O. Knauer. High Speed Internetworking: Grundlagen und Konzepte des FDDI- und ATM-Einsatzes. Addison-Weley, Bonn, 1994.
- [4] T. Berners-Lee, D. Connolly. Hypertext Markup Language - 2.0. RFC 1866, November 1995.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1866.txt>
- [5] J. Busse, D. Krause, H. Klaeren. Verbundlehre: Didaktische Synthese von Ferne und Präsenz. Didaktisches Design, Heft 3, 1998, S. 35-54.
- [6] Cisco Systems. Producing Desktop Video. White Paper. June 1998.  
[http://cco.cisco.com/warp/public/732/net\\_enabled/iptv/pdv\\_wp.pdf](http://cco.cisco.com/warp/public/732/net_enabled/iptv/pdv_wp.pdf)
- [7] Cisco Systems. Cisco IP/TV 3400 Series Server. Data Sheet. November 1998.  
[http://cco.cisco.com/warp/public/732/net\\_enabled/iptv/iptv3\\_ds.pdf](http://cco.cisco.com/warp/public/732/net_enabled/iptv/iptv3_ds.pdf)
- [8] Cisco Systems. IP/TV 2.0 Data Sheet. April 1998.  
[http://cco.cisco.com/warp/public/732/net\\_enabled/iptv/iptv2\\_ds.pdf](http://cco.cisco.com/warp/public/732/net_enabled/iptv/iptv2_ds.pdf)
- [9] Cisco Systems. Stream Watch. Data Sheet. June 1998.  
[http://cco.cisco.com/warp/public/732/net\\_enabled/iptv/strmw\\_ds.pdf](http://cco.cisco.com/warp/public/732/net_enabled/iptv/strmw_ds.pdf)
- [10] J.D. Day, H. Zimmermann. The OSI Reference Model, Proceedings of the IEEE, Vol. 71, pp. 1334-1340, 1983.
- [11] S. E. Deering, D. R. Cheriton. Host Groups: A Multicast Extension to the Internet Protocol. RFC 966, 1985.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc966.txt>
- [12] S. E. Deering. Host Extensions for IP Multicasting. RFC1054, 1988.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1054.txt>

- [13] Databeam Corp. A Primer on the H.323 Series Standard. 1998.  
[http://www.databeam.com/pdf/files/h323\\_primer-v2.pdf](http://www.databeam.com/pdf/files/h323_primer-v2.pdf)
- [14] DataBeam Corp. A Primer on the T.120 Series Standard. 1998.  
[http://www.databeam.com/pdf/files/t120\\_primer.pdf](http://www.databeam.com/pdf/files/t120_primer.pdf)
- [15] W. Effelsberg, R. Steinmetz. Video Compression Techniques. dpunkt.verlag, Heidelberg, 1998.
- [16] J. Elkner, S. Geschke, O. Kunert, F.K. Zbrog. MBone - Systemspezifische Gestaltung und erste Erfahrungen mit Videokonferenzen, Technischer Bericht 8, Universität Magdeburg, Juni 1997.  
[http://www.witi.cs.uni-magdeburg.de/iti\\_db/veroeffentlichungen/97/ElkGesKun97.html](http://www.witi.cs.uni-magdeburg.de/iti_db/veroeffentlichungen/97/ElkGesKun97.html)
- [17] H. Fahner, P. Feil. MBone - Multimedia-Konferenzen im Internet. dpunkt.verlag, Heidelberg, 1999.
- [18] D. Flanagan. Java in a Nutshell, Deutsche Ausgabe für Java 1.0. O'Reilly/International Thomson Verlag, Bonn, 1996.
- [19] S. Floyd, V. Jacobson, S. McCanne, C. Liu, L. Zhang. A Reliable Multicast Framework for Light-weight Sessions and Application Framing. IEEE/ACM Transactions on Networking, Volume 5, Number 6, pp. 784-803, December 1997.
- [20] P. Flynn et al. The XML FAQ, Version 1.41. October 1998.  
<http://www.ucc.ie/xml/>
- [21] M. Grumann. Entwurf und Implementierung eines zuverlässigen Multicast-Protokolls zur Unterstützung sicherer Gruppenkommunikation in einer TeleTeaching-Umgebung. Master's Thesis, Lehrstuhl Praktische Informatik IV, Universität Mannheim, 1997.
- [22] M. Handley. SAP - Session Announcement Protocol. Internet Draft, 1997.
- [23] M. Handley, V. Jacobson. SDP: Session Description Protocol. RFC 2327, 1998.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc2327.txt>
- [24] M. Handley. On Scalable Internet Multimedia Conferencing Systems. Thesis, University College London, 1997.
- [25] M. Handley, H. Schulzrinne. SIP: Session Initiation Protocol. Internet Draft, 1997.
- [26] E.R. Harold. Java Network Programming. O'Reilly, Sebastopol, CA, 1997.
- [27] J. Hasenbrock. Multimedia-Psychologie, eine neue Perspektive menschlicher Kommunikation. Spektrum Akademischer Verlag, 1995.
- [28] P. Heiligers. MBone-DE Infrastruktur Map. DFN-NOC, Januar 1999.  
<http://www.mbone.de/maps/mbone-de-01-99.jpg>

- [29] V. Hilt, W. Geyer. A Model for Collaborative Services in Distributed Learning Environments. In: Proceedings of IDMS'97, Technical University of Darmstadt, LNCS 1309, pp. 364 -375, 1997.
- [30] W. Holfelder. Interactive Remote Recording and Playback of Multicast Videoconferences. In: Proceedings of IDMS'97, Technical University of Darmstadt, LNCS 1309, pp. 450-468, 1997.
- [31] Intel Corporation. Benutzerhandbuch Intel ProShare Conferencing Produkte. Hillsboro, OR, März 1996.
- [32] ITU. Recommendation G.711 - Pulse code modulation (PCM) of voice frequencies. 1988.  
[http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g711\\_e\\_70262.pdf](http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g711_e_70262.pdf)
- [33] ITU. Recommendation G.722 - 7 kHz audio-coding withiin 64 kbit/s. 1988.  
[http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g722\\_e\\_70266.pdf](http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g722_e_70266.pdf)
- [34] ITU. Recommendation G.726 - 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM). December 1990.  
[http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g726\\_e\\_70276.pdf](http://www.itu.int/itudocs/itu-t/rec/g/g700-799/g726_e_70276.pdf)
- [35] ITU. Recommendation H.261 - Video codec for audiovisual services at p x 64 kbit/s. März 1993. [http://www.itu.ch/itudocs/itu-t/rec/h/h261\\_e\\_8979.zip](http://www.itu.ch/itudocs/itu-t/rec/h/h261_e_8979.zip)
- [36] ITU. Recommendation H.263 - Video coding for low bit rate communication. February 1998. [http://www.itu.ch/itudocs/itu-t/rec/h/h263\\_e\\_41905.zip](http://www.itu.ch/itudocs/itu-t/rec/h/h263_e_41905.zip)
- [37] ITU. Recommendation H.323 - Packet-based multimedia communications systems. February 1998. [http://www.itu.ch/itudocs/itu-t/rec/h/h323\\_e\\_55639.zip](http://www.itu.ch/itudocs/itu-t/rec/h/h323_e_55639.zip)
- [38] ITU. Recommendation T.120 - Data protocols for multimedia conferencing. July 1996. [http://www.itu.int/itudocs/itu-t/rec/t/t120\\_e\\_72801.pdf](http://www.itu.int/itudocs/itu-t/rec/t/t120_e_72801.pdf)
- [39] V. Jacobson, S. McCanne. Using the LBL Network Whiteboard. Lawrence Berkeley Laboratory, Berkeley, CA, 1995.
- [40] A. Johannsen. Videoconferencing: A Survey of Products, Resources and Research. Group Project Report, University of Hohenheim, 1996.
- [41] M. Kerres. Multimediale und telemediale Lernumgebungen - Konzeption und Entwicklungen. Oldenbourg, München, 1998.
- [42] J. Kirch. Microsoft Windows NT Server 4.0 versus UNIX. February 1999.  
<http://www.unix-vs-nt.org/kirch>
- [43] C. Huitema. Routing in the Internet. Prentice Hall, Englewood Cliffs, NJ, 1995.

- [44] J. Knoll. Kurs- und Seminarmethoden : ein Trainingsbuch zur Gestaltung von Kursen und Seminaren, Arbeits- und Gesprächskreisen. Beltz, Weinheim, 1993.
- [45] D. Krause. Diskussionen mit und ohne Netz. Erfahrungen mit der Netzkomponente im Fernstudium "Informatik und Gesellschaft". Berichte des Wilhelm-Schickard-Institus für Informatik, WSI-98-15, Universität Tübingen, 1998.
- [46] V. Kumar. MBone: Interactive Multimedia on the Internet. New Riders Publishing, Indianapolis, USA, 1995.
- [47] T. Liao. Light-weight Reliable Multicast Protocol as an Extension to RTP. INRIA, 1997.  
[http://monet.inria.fr/lrmp/lrmp\\_rtp.html](http://monet.inria.fr/lrmp/lrmp_rtp.html)
- [48] T. Liao. Light-weight Reliable Multicast Protocol Specification. Internet Draft, October 1998.  
<http://monet.inria.fr/lrmp/draft-liao-lrmp-00.txt>
- [49] D. Meyer. Administratively Scoped IP Multicast. IETF Internet Draft, Juni 1998.
- [50] Microsoft Corporation. Netmeeting Features. Januar 1998.  
<http://www.microsoft.com/netmeeting>
- [51] D.L. Mills. RFC 1305 - Network Time Protocol (Version 3) - Specification, Implementation and Analysis. University of Delaware, March 1992.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1305.txt>
- [52] Netscape. Netscape Livemedia Framework for Internet Multimedia to Integrate Precept Software's RTP Protocol. Pressemitteilung, Palo Alto, Mai 1996.  
<http://www20.netscape.com/newsref/pr/newsrelease136.html>
- [53] PictureTel. Live 200 Data Sheet.  
<http://www.pictel.com/products/download/live200.pdf>
- [54] PictureTel. PicTel 330 Data Sheet.  
<http://www.pictel.com/products/download/pt330.pdf>
- [55] PictureTel. SwiftSite Data Sheet. 1998.  
<http://www.pictel.com/products/download/swiftsit.pdf>
- [56] J. Postel. Transmission Control Protocol. RFC 793, September 1981.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc793.txt>
- [57] M. de Prycker. Asynchronous Transfer Mode - Die Lösung für Breitband-ISDN. Prentice Hall, München, 1994.
- [58] L. Rasmussen. About the Data Format of WB. 1995.  
<http://www.it.kth.se/d90-lra/wb-proto.html>

- [59] J. Reynolds, J. Postel. Assigned Numbers. RFC 1700, October 1994.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1700.txt>
- [60] D. Schick. Netzwerkkunterstützung für verteilte VR-Systeme. Diplomarbeit, Lehrstuhl für Graphische Interaktive Systeme, Wilhelm Schickard-Institut für Informatik, Universität Tübingen, 1997.
- [61] B. Schinzel, V. Wulf. Televorlesungen - ein Mittel zur Bereicherung des Studienangebots? In *it+ti* 39 (1997) 6, S. 35-40, R. Oldenbourg Verlag, München, 1997.
- [62] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. RTP: A Transport Protocol for Real-Time Applications. RFC 1889, 1996.  
<http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1889.txt>
- [63] H. Schulzrinne. RTP Profile for Audio and Video Conferences with Minimal Control. Internet Draft, 1997. [ietf-avt-profile-new-02.txt](http://www.ietf.org/internet-drafts/draft-schulzrinne-02.txt)
- [64] R. Schulz-Zander. Neue Lernkultur durch Vernetzung? In *it+ti* 39 (1997) 6, S. 21-26, R. Oldenbourg Verlag, München, 1997.
- [65] J. Scourias. Overview of the Global System for Mobile Communications. University of Waterloo, Ontario, Canada, 1997.  
<http://ccnga.uwaterloo.ca/~jscouria/GSM/gsmreport.html>
- [66] A. Sloane. Multimedia Communication. Mc Graw-Hill, London, 1996.
- [67] R. Steinmetz, K. Nahrstedt. Multimedia: Computing, Communications and Applications. Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [68] C. Strothotte, T. Strothotte. Seeing between the Pixels: Pictures in Interactive Systems. Springer, Berlin/Heidelberg, 1998.
- [69] S. Sullivan, L. Winzeler, J. Deagen, D. Brown. Programming with the Java Media Framework. Wiley Computer Publishing, New York, 1998.
- [70] A.S. Tanenbaum. Computer Networks. 3rd ed. Prentice Hall, Upper Saddle River, NJ, 1996.
- [71] J. Tranter. Linux Multimedia Guide. O'Reilly, Sebastopol, CA, 1996.
- [72] Unicode Consortium. The Unicode Standard, Version 2.0. Addison Wesley, 1996.  
<http://www.unicode.org>
- [73] VTEL. Product Specification Sheets. 1998.  
<http://www.vtel.com/newsinfo/resource/products/>
- [74] VTEL. Interactive Whiteboard Specification. 1998.  
<http://www.vtel.com/newsinfo/resource/products/whtbor.pdf>

- [75] J.B. Walther. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, pp. 3-43, 1996.
- [76] Pressemitteilung des Wissenschaftsrats. Empfehlungen zur Hochschulentwicklung durch Multimedia in Studium und Lehre. Drs. 3536/98, 1998.
- [77] R. Wittmann, M. Zitterbart. *Multicast: Protokolle, Programmierung, Anwendungen*. dpunkt.verlag, Heidelberg, 1999.