

Ein Framework zur verteilten Produktion von Präsentationsaufzeichnungen

Von der Fakultät für Elektrotechnik und Informatik
der Universität Hannover
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation
von

Dipl.-Ing. Ralf Einhorn
geboren am 30.03.1969 in Uelzen

2005

Referent Prof. Dr.-Ing. Klaus Jobmann
Korreferent Prof. Dr. techn. Wolfgang Nejd
Tag der Promotion 25.11.2005

Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit der effizienten Transformation von Live-Präsentationen in rechnergestützt abrufbare Repräsentationen. Die multimediale Aufzeichnung von Vorträgen, Vorlesungen, Kursen und Tagungen findet vornehmlich Anwendung in E-Learning-Szenarios.

In den letzten Jahren wurden in den Hochschulen zunehmend Hörsäle und Seminarräume mit Medientechnik für die Aufzeichnung, Übertragungen und Konferenzen ausgestattet. Die Herstellung von anspruchsvollen Aufzeichnungen mit unterschiedlichen Randbedingungen gestaltet sich jedoch personal- und zeitaufwendig.

Existierende Systeme für verschiedene Anwendungsschwerpunkte weisen unterschiedliche Stärken und Schwächen auf; die Arbeit liefert hierzu einen Überblick. Anschließend führt die Analyse der Anwendungsszenarios zusammen mit einer Betrachtung der vorliegenden Produktions-Infrastrukturen zu dem Entwurf eines neuartigen Ansatzes.

Kern des Entwurfs ist ein modulares, netzverteiltes Produktionssystem. Für aktuelle und zukünftige Szenarios können über eine offene Schnittstelle jeweils optimal geeignete Module eingebunden werden. Die Erweiterbarkeit trägt dabei auch sich verändernden Technologien (wie etwa neuen Codecs) Rechnung. Eine Aufzeichnung kann effizient mit Einbindung des vom Dozenten genutzten Präsentationssystems erfolgen, es ist jedoch auch eine Produktion völlig ohne Beeinflussung des Vortragenden (etwa auf Tagungen) möglich. Grundlage hierfür bildet ein auf aktuelle Infrastrukturen zugeschnittenes Aufzeichnungs-Framework. Dieses verteilte System ermöglicht neben einer Lastverteilung und Redundanz auch „light weight authoring“ – die Produktion „on-the-fly“ ohne die Ablenkung oder Einbeziehung des Dozenten.

Neben der effizienten und weit gehend automatisierbaren Produktion steht die geeignete Publikation der erzeugten Repräsentation in Hinblick auf die Distribution im Vordergrund. Das System nutzt die inhaltliche Parallelläufigkeit der in Präsentationen eingesetzten Medien. Mit Hilfe der Verknüpfung von zeitabhängigen und inhaltsbezogenen Metadaten gelingt so die Assoziation parallel verwendeter Medien auf der Ebene von Kapiteln – Aufzeichnungen werden damit navigier- und durchsuchbar, wie es von statischen Dokumenten bekannt ist.

Basis für die Produktionskette bildet ein zur Abbildung von Präsentationen entwickeltes Metadatenmodell. Zudem nutzt das Framework Schnittstellen zu Standardsystemen. Die prototypische Implementierung von ausgewählten Modulen ermöglicht die abschließend durchgeführte Bewertung des Entwurfs anhand realer Szenarios.

Schlagnworte: Verteilte Systeme, Präsentationsaufzeichnung, Medienproduktion, Metadaten, E-Learning

Abstract

This work focusses on the task of efficiently transforming live presentations into computer based distributable *representations*. The main application for multi-medial recording of speeches, lectures, courses and conferences is e-learning.

During the past years many lecture halls have been supplied with media equipment for audiovisual recordings, transmissions and conferences. However, the production of sophisticated representations in different settings is time consuming and causes high personnel costs.

Existing systems for certain utilization purposes have different advantages and weaknesses. This thesis provides an overview of these diverse systems. Moreover the analysis of current application scenarios together with a contemplation of existing production infrastructures lead to a novel design approach.

The core of the designed system is formed by a modular open production framework. Therefor optimized modules can be deployed for each present and future application scenario. The modular approach also suits the need for changing technologies like new video codecs. Recordings can be efficiently made by embedding of the lectures' presentation system. On the other hand a production is also possible without any impact on the lecturer, e. g. while recording speeches at conferences. This can be accomplished by building upon a capturing framework tailored to prevalent infrastructures. Besides redundancy and load balancing this distributed system enables real 'light weight authoring' without any distraction of the speaker.

Efficient and highly automated production is a key issue. Besides stands the suitable publication of the created representation regarding to distribution. The system exploits the content related parallels of the different media utilized in the presentation. By linking time based and content related meta data the association of parallel deployed media leads to a division into chapters – recordings become searchable and can be navigated like static documents.

Basis for the production chain is a meta data model designed to reflect actual presentations. Additionally, the framework uses interfaces for the deployment of standard systems. The prototype implementation of selected modules facilitates the evaluation of the design in real life scenarios.

Keywords: distributed systems, presentation recording, media production, meta data, e-learning

Inhaltsverzeichnis

Kurzfassung	I
Abstract	III
Inhaltsverzeichnis	V
Abbildungsverzeichnis	IX
Tabellenverzeichnis	XI
Abkürzungsverzeichnis	XIII
1 Einleitung.....	1
1.1 Ausgangssituation	1
1.2 Anforderungen an Aufzeichnungssysteme.....	3
1.3 Motivation und Ziele	4
1.4 Aufbau	7
2 Grundlagen.....	9
2.1 Medientechnologie	9
2.1.1 Grundbegriffe.....	9
2.1.2 Medienproduktionskette	9
2.1.3 Repräsentationsmedien	11
2.1.4 Transport.....	14
2.2 Basissysteme	15
2.2.1 Präsentationsmethoden	15
2.2.2 AV-Aufzeichnungssysteme	18
2.2.3 Bearbeitung	24
2.2.4 Distribution	24
2.3 Systemintegration.....	25
2.3.1 Einbindung von AV-Hardware.....	25
2.3.2 Integration von Software.....	26
2.4 Zeit und Synchronisation.....	28
2.4.1 Zeitformat	28
2.4.2 Synchronisation von Uhren.....	28
2.4.3 Synchronisation von Mediendaten.....	29
2.5 Metadatenmodelle	29
2.5.1 Metadaten zur Inhaltsbeschreibung	29
2.5.2 Metadaten zur Beschreibung von Medien	31
3 Problemanalyse	35
3.1 Ausgangspunkt: Anwendung und Kontext.....	35
3.1.1 Anwendungsszenarios.....	35
3.1.2 Ziel der Präsentationsaufzeichnung	36
3.1.3 E-Learning-Kontext	37

3.1.4	Rollen und Begriffe.....	37
3.2	Analyse: Anwendung und Produktionsaspekte.....	38
3.2.1	Produktionstyp.....	38
3.2.2	Veranstaltungsform (Anwendungsszenarios).....	39
3.2.3	Zielgruppe (Anwender von Aufzeichnungssystemen).....	39
3.2.4	Aufzeichnungsvorbereitung.....	40
3.2.5	Postproduktion.....	40
3.2.6	Workflow.....	41
3.3	Analyse: Live-Präsentation.....	42
3.3.1	Vortrag.....	42
3.3.2	Konventionelle und PC-gestützte Präsentationsmedien.....	42
3.3.3	Interaktivität von Medien.....	43
3.3.4	Zeitverhalten von Medien.....	43
3.3.5	Präsentationsmedien: Fazit und Zusammenfassung.....	44
3.4	Analyse: Repräsentation und Distribution.....	44
3.4.1	Vortrag: AV-Aufzeichnung des Dozenten.....	45
3.4.2	Repräsentationsform.....	45
3.4.3	Navigation und Integration: Metadaten.....	47
3.4.4	Distribution.....	47
3.4.5	Medienbruch.....	48
3.4.6	Repräsentationsmedien für Visualisierungen.....	49
3.5	Systemaspekte.....	50
3.5.1	Systemtyp (Flexibilität).....	50
3.5.2	Modularität.....	51
3.5.3	Unimedia/Multimedia.....	52
3.5.4	Integration von Präsentationswerkzeugen.....	53
3.5.5	Standards.....	54
3.6	Existierende Systeme.....	55
3.6.1	Systeme in der Forschung.....	55
3.6.2	Kommerzielle Produkte.....	57
3.6.3	Anwendungsbeispiele.....	57
3.7	Systemklassifikation.....	58
3.7.1	Klassifikationsschema.....	59
3.7.2	Bewertungsschema.....	60
3.7.3	Klassifikation und Bewertung existierender Systeme.....	61
3.8	Anforderungen an ein universelles Framework.....	63
3.8.1	Anwendung und Produktion.....	64
3.8.2	Präsentation.....	65
3.8.3	Repräsentation.....	65
3.8.4	Systemarchitektur.....	66
3.8.5	Weitere Aspekte.....	68
3.8.6	Systemübersicht.....	69

4	Entwurf und Implementierung	71
4.1	Konzepte des Systementwurfs.....	71
4.1.1	Multimedialität.....	71
4.1.2	Modularität: Offenheit, Integration und Synchronisation.....	71
4.1.3	Einsatz von Metadaten.....	73
4.1.4	Verteilte Produktion.....	74
4.2	Produktions-Framework.....	74
4.2.1	Produktionsablauf.....	74
4.2.2	Aufzeichnung.....	76
4.2.3	Nachbearbeitung.....	77
4.2.4	Publikation.....	80
4.3	Präsentationsmodell.....	81
4.3.1	Aspekte einer Präsentation aus Aufzeichnungssicht.....	81
4.3.2	Medienklassen.....	81
4.3.3	Zeitbasen der Medien.....	82
4.4	Metadatenmodell.....	82
4.4.1	Anforderungen an das Metadatenmodell.....	83
4.4.2	Vergleich zu existierenden Metadatenformaten.....	85
4.4.3	Hierarchisches Metadatenmodell und -format.....	86
4.5	Implementierung.....	90
4.5.1	Aufzeichnung.....	91
4.5.2	Import.....	96
4.5.3	Nachbearbeitung.....	97
4.5.4	Export.....	99
4.5.5	Implementierte Module des VACE-Frameworks.....	103
5	Anwendung, Performance und Bewertung	105
5.1	Anwendung des VACE-Frameworks.....	105
5.1.1	Testszenarios.....	105
5.1.2	Durchführung.....	105
5.1.3	Auswertung.....	108
5.2	Performance-Messungen.....	113
5.2.1	Abweichungen der Synchronisation.....	113
5.2.2	Zeitaufwand für die Nachbearbeitung.....	114
5.3	Zusammenfassende Bewertung und Systemvergleich.....	117
6	Zusammenfassung und Ausblick	119
6.1	Zusammenfassung.....	119
6.2	Ausblick.....	120
	Literaturverzeichnis	123

Abbildungsverzeichnis

Abbildung 2-1:	Modell einer Medienproduktionskette	10
Abbildung 2-2:	Schema eines projektorgestützten Präsentationssystems	16
Abbildung 2-3:	Funktionsschema eines Scan-Converters	20
Abbildung 2-4:	Datenflussdiagramm zum Live-Encoding	21
Abbildung 2-5:	Medientechnik im Audimax der Universität Hannover	23
Abbildung 2-6:	Multimedia-Seminarraum im Forschungszentrum L3S	23
Abbildung 3-1:	Blockdiagramm der Produktionskette einer Präsentationsaufzeichnung	36
Abbildung 3-2:	IEEE Learning Technology Systems Architecture [19]	37
Abbildung 3-3:	Elemente einer Live-Präsentation: Vortrag und Präsentationsmedien	42
Abbildung 3-4:	Relevanz von Standardformaten in der Produktion	54
Abbildung 4-1:	Diagramm zum Produktionsfluss im VACE-Framework	75
Abbildung 4-2:	Präsentationsmodell	82
Abbildung 4-3:	Hierarchisches Metadatenmodell zur Repräsentationsbeschreibung	86
Abbildung 4-4:	VaceHxPW, Applikationshauptfenster und Aufzeichnungseinstellungen ..	92
Abbildung 4-6:	VaceBrowser, Applikationshauptfenster und Metadaten-Einstellungen	93
Abbildung 4-5:	Zeitdiagramm zur Ermittlung des Encoding-Startzeitpunkts beim HxPW ..	93
Abbildung 4-7:	VacePPT, Applikationshauptfenster und Dialog für Einstellungen	94
Abbildung 4-8:	VaceEditor, Applikationshauptfenster und Eingabedialog	98
Abbildung 4-9:	VaceEditor, HTML-basierte Online-Hilfe im Windows-Help-Viewer	98
Abbildung 4-11:	Modell der Medien-Synchronisation im HTML-Frameset	101
Abbildung 4-10:	HTML-Repräsentation einer Präsentationsaufzeichnung	101
Abbildung 4-12:	SMIL-Repräsentation einer Präsentationsaufzeichnung im RealPlayer	102
Abbildung 4-13:	Implementierte VACE-Framework-Module	103
Abbildung 5-1:	TestszENARIO 1: Vorlesungsaufzeichnung mit VacePPT	106
Abbildung 5-2:	TestszENARIO 2: Vortragsaufzeichnung mit generischer Folienaufnahme ..	106
Abbildung 5-3:	Assoziation extrahierter Metadaten zur Navigationserzeugung	111
Abbildung 5-4:	Systemaufbau zur Ermittlung der Synchronisationsgenauigkeit	113

Tabellenverzeichnis

Tabelle 2-1:	Medienparameter (Beispiele).....	11
Tabelle 2-2:	Parameter typischer Medienformate	13
Tabelle 2-3:	Funktionen von AV-Plugins	27
Tabelle 2-4:	Dublin Core Elemente (Auswahl).....	30
Tabelle 3-1:	Klassifikation von Präsentationsmedien	44
Tabelle 3-2:	Nutzung von identischen Medien in Präsentation und Repräsentation	48
Tabelle 3-3:	Konventionelle Video- und Präsentationsaufzeichnung.....	49
Tabelle 3-4:	Typische Betrachtungsparameter von PC-Präsentationen	50
Tabelle 3-5:	Klassifikation von Präsentationsaufzeichnungssystemen.....	59
Tabelle 3-6:	Bewertungsschema für Präsentationsaufzeichnungssysteme	61
Tabelle 3-7:	Einordnung von Aufzeichnungssystemen in das Klassifikationsschema	61
Tabelle 3-8:	Bewertung existierender Systeme.....	63
Tabelle 3-9:	Module eines Aufzeichnungs-Frameworks	67
Tabelle 3-10:	Anforderungsprofil eines universellen Aufzeichnungssystems.....	69
Tabelle 4-1:	Metadaten – Funktionen	83
Tabelle 4-2:	Metadatenelemente von „presentation“	87
Tabelle 4-3:	Metadatenelemente von „stream“	88
Tabelle 4-4:	Metadatenelemente von „event“	89
Tabelle 4-5:	Metadatenwerte für „eventtype“	89
Tabelle 4-6:	Übertragung von HTML-Metatags in Elemente der VACE-Events.....	97
Tabelle 4-7:	VACE-Exportformate	99
Tabelle 5-1:	Mehraufwände durch den Einsatz von VACE in den Testszenarios	110
Tabelle 5-2:	Vergleich der Repräsentationen der Testszenarios	112
Tabelle 5-3:	Datenvolumen von typischen Präsentationsaufzeichnungen	114
Tabelle 5-4:	Zeitaufwand für die Datenübertragung in den Testszenarios	115
Tabelle 5-5:	Systemvergleich/Bewertung	118

Abkürzungsverzeichnis

ACM.....	Association for Computing Machinery
ADL.....	Advanced Distributed Learning Initiative (www.adlnet.org)
AV.....	Audio/Video
CD-ROM.....	Compact Disk – Read Only Memory
CHM.....	Compiled Help Module (Dateiformat für Online-Hilfefunktion, Microsoft)
CMS	Content Management System
CSS.....	Cascading Style Sheet [80]
CSV	Comma Separated Value (Textdatenformat, z. B. für Tabellendarstellung)
DC	Dublin Core (www.dublincore.org)
DCMI	Dublin Core Metadata Initiative
DCF77	Funk-Zeitsignal; D=Deutschland, C=Langwelle, F=Frankfurt (Standort), 77,5 kHz Sendefrequenz
DOM.....	Document Object Model [82]
DSL	Digital Subscriber Line (digitale Teilnehmeranschlussleitung)
DV	Digital Video (Kodierungs-, Übertragungs- und Dateiformat, 25 Mbit/s)
DVCAM.....	Professionelle DV-Version (höhere Spurbreite)
DVCPRO.....	Professionelle DV-Version (50 Mbit/s)
DVD	Digital Versatile Disk
EDL.....	Edit Decision List, Schnittliste (Videobearbeitung)
EIA	Electronics Industries Association
EMF.....	Enhanced Metadata Format (Bilddatenformat)
FPS	Frames per second, (Halb-) Bilder pro Sekunde
FTP	File Transfer Protocol (RFC 959)
GIF	CompuServe Graphics Interchange Format (Bilddatenformat)
GPS.....	Global Positioning System (satellitengesteuertes Navigationssystem)
GUI.....	Graphical User Interface, grafische Benutzeroberfläche (eines Programms)
H.323	ITU Norm für Videokonferenzen über IP
HTML.....	Hypertext Markup Language (RFC 1866, 2854)
HTTP.....	Hypertext Transfer Protocol (RFC 2616)
IANA.....	Internet Assigned Numbers Authority
IEEE	Institute of Electrical and Electronics Engineers (www.ieee.org)
IEC	International Electrotechnical Commission (www.iec.ch)
IMS.....	Instructional Management Systems (IMS Global Learning Consortium, Inc.)
IP	Internet Protocol
ISDN.....	Integrated Services Digital Network (digitales Telefonnetz)
ISO	International Organization for Standardization (www.iso.ch)
ITU	International Telecommunication Union (www.itu.int)
ITU-T	ITU-Telecommunication Standardization Bureau; Nachfolger von CCITT (International Telegraph and Telephone Consultative Committee) und CCIR (International Radio Consultative Committee)
ITU-RS601	Digital Video-Standard für Fernsehbilder, Unkomprimiert mit 4:2:2 Subsampling (ehem. CCIR-601)

JPEG.....	Joint Photographic Experts Group (JPEG-Kompression: ISO/IEC 19018)
JS.....	JavaScript
L3S.....	Forschungszentrum L3S (www.l3s.de)
LAN.....	Local Area Network (lokales Rechnernetz)
LMS.....	Learning Management System
LOM.....	Learning Objects Metadata
MBone.....	Multicast Backbone; Netz für die Multicast- (Mehrpunkt-) Datenübertragung
MP3.....	MPEG (-1), Layer-3 (Audiodateiformat, ISO/IEC 11172-3, 13818-3)
MPEG.....	Moving Pictures Experts Group (www.mpeg.org) (Standards: MPEG-1: ISO/IEC 11172, MPEG-2: ISO/IEC 13818, MPEG-4, -7, -21)
MHEG.....	Multimedia and Hypermedia Information Coding Expert Group (www.mheg.org ; MHEG-1: ISO/IEC 13522)
MIDI.....	Musical Instrument Digital Interface (Schnittstellen-Standard)
MIME.....	Multipurpose Internet Mail Extensions (RFC 2045, 2046)
MJPEG.....	Motion-JPEG (Videokodierung)
MM.....	Multimedia
MS.....	Microsoft
NLE.....	Non Linear Editing (Videobearbeitung mit wahlfreiem Zugriff)
NFS.....	Network Filesystem (Netzprotokoll, -dateisystem)
NTP.....	Network Time Protocol (RFC 1305) [56]
OLE.....	Object Linking and Embedding
PCM.....	Pulse-code Modulation
PADLR.....	Personalized Access to Distributed Learning Repositories (L3S-Projekt)
PAL.....	Phase Alternation Line (Fernsehsystem)
PDA.....	Personal Digital Assistant (Handheld-Computer)
PDF.....	Portable Document Format
Pixel.....	Picture Element, Bildpunkt einer Grafik oder Projektionseinrichtung
PNG.....	Portable Network Graphics (Bildformat)
PPT.....	PowerPoint (MS Präsentationsprogramm)
PTZ.....	Pan/Tilt/Zoom: Fernsteuerbarer Schwenk-/ Neige-Mechanismus und Brennweiteneinstellung für Kameras
OCR.....	Optical Character Recognition (Texterkennung)
OS.....	Operating Systems, Betriebssystem
QT.....	QuickTime (AV-Format und -Applikationen, Apple)
RFC.....	Request for Comments (Internet-Standard) (www.rfc-editor.org)
RGB.....	Farbmodell (Rot, Grün, Blau)
RRZN.....	Regionales Rechenzentrum für Niedersachsen
RS-232.....	EIA-Standard für unsymmetrische serielle Computerschnittstelle
RS-422.....	EIA-Standard für symmetrische serielle Computerschnittstelle
RTP.....	Real-time Transport Protocol, RFC 1889 [57], ITU H.225.0
RTSP.....	Real-time Streaming Protocol, RFC 2326 [59]
RM.....	RealMedia (AV-Format und -Applikationen, RealNetworks)
SCORM.....	Sharable Content Object Resource Model
SCP.....	Secure copy (verschlüsseltes Kopieren von Dateien im Netz über SSH)
SFTP.....	Secure FTP (verschlüsselte Dateiübertragung über SSH)
SMB.....	Server Message Block (Netzprotokoll, u. a. zur Dateiübertragung)

SMIL	Synchronized Multimedia Integration Language [84]
SSH.....	Secure Shell (Netzprotokoll für verschlüsselte Verbindungen)
SVG.....	Scalable Vector Graphics (Grafikformat) [83]
TCR.....	Timecode Recording (Aufzeichnung von Zeitinformationen im Video)
TV.....	Television; hier: Videosignal im PAL-Format zur Unterscheidung von PC-Bildsignalen (VGA)
TZD	Time Zone Designator, Angabe der Zeitzone bzw. Differenz zur UTC/GMT
UI.....	User Interface, Benutzeroberfläche eines Programms
URL.....	Uniform Resource Locator (RFC 1737)
URI.....	Uniform Resource Identifiers (RFC 2396)
USB.....	Universal Serial Bus (serielle Computer-Schnittstelle)
UTC.....	Coordinated Universal Time, Kompromiss aus der englischen und französischen Abkürzung (CUT bzw. TUC)
VACE	Video and Audio Capturing and Embedding
VCR.....	Video Cassette Recorder (Videorekorder)
VGA	Video Graphics Array, analoge Grafikkarte ursprünglich 640*480 Pixel; hier: vom PC-Videosignal zur Unterscheidung vom (PAL/TV)-Video
ViDe	Video Development Initiative (www.vide.org)
VNC	Virtual Network Computing, Fernzugriffs-Software (www.realvnc.org)
VR	Virtual Reality
VTR.....	Video Tape Recorder (Videorekorder)
W3C	World Wide Web Consortium (www.w3.org)
WM.....	Windows Media (AV-Format und -Applikationen, Microsoft)
WMF	Windows MetaFile (Bilddatenformat)
WWW.....	World Wide Web
XML.....	eXtensible Markup Language (www.w3.org/XML/)
YUV	Farbmodell: Luminanz (Y), Chrominanz/Farbdifferenz (U, V) (TV)

1 Einleitung

1.1 Ausgangssituation

Mit zunehmender Leistungsfähigkeit der Rechner- und Netzwerktechnik ist das stetige Zusammenwachsen der Bereiche Medien- und Informationstechnologie zu beobachten. War bis vor einigen Jahren z. B. die Verarbeitung von Audio und Video nur mit Hilfe proprietärer Systeme möglich, so ist heute hierfür der Einsatz von PC-basierten Systemen der Standardfall. Durch die Verknüpfung von traditionellen Computer-Datenformaten mit Audio und Video entstand eine Fülle von *multimedialen* Anwendungen sowohl im Unterhaltungsbereich als auch in professionellen Bereichen wie z. B. der Aus- und Weiterbildung.

Die Begriffe *E-Learning* und *Blended Learning* (als Mischform von Präsenz- und Fernlehre) befinden sich seit Ende der 90er Jahre in der Diskussion: Der Studierende verfolgt Vorlesungen und Kurse nicht mehr (nur) an der Hochschule. Mit entsprechender Ausstattung kann er sich diese live oder als Aufzeichnung an einem beliebigen Ort und – im Falle der Aufzeichnung – zu einem beliebigen Zeitpunkt ansehen. Auf diese Weise bietet sich die Möglichkeit, auch ohne Ortswechsel die Vorlesungen anderer Universitäten zu verfolgen und so z. B. Prüfungen in Bereichen abzulegen, die an der eigenen Hochschule nicht angeboten werden. Die Anwendungen hierfür sind vielfältig, speziell im Bereich der professionellen Weiterbildung, wo der Lernende im Gegensatz zum Vollzeit-Studierenden örtlich und terminlich stärkeren Restriktionen unterworfen ist.

Aufzeichnungen können Präsenzveranstaltungen *unterstützen*, indem sie der Nachbereitung des Vorlesungsstoffes dienen. Didaktische und pädagogische Fragestellungen sind jedoch nicht Bestandteil der Arbeit; auch soll nicht die Frage erörtert werden, inwieweit die klassische Präsenzveranstaltung *ersetzt* werden kann.

Im Vordergrund stehen somit die technischen Aspekte der Produktion elektronisch verfügbarer *Repräsentationen* aus Vorträgen, Vorlesungen, Kursen, Seminaren usw. unterschiedlicher Fachbereiche.

Die vom Dozenten im Vortrag präsentierten Inhalte sind möglichst effizient, d. h. mit minimalem Mehraufwand, in eine abrufbare Repräsentation zu transformieren. Denn im Idealfall stellt ein gehaltener Vortrag bereits eine existierende Ressource dar, die lediglich auch in elektronisch übertragbarer, audiovisueller Form nutzbar gemacht werden muss. Entscheidendes Kriterium ist hierbei, den Informationsgehalt der Präsenzveranstaltung in der Aufzeichnung beizubehalten.

Neben Aspekten der Produktion spielt somit insbesondere die Gestaltung der Repräsentation und deren Einbindung in Publikationssysteme eine wichtige Rolle. Im Zuge eines stetig wachsenden Informationsangebotes müssen Inhalte effizient

auffindbar sein, was durch die Bemühungen um das *Semantic Web* und den damit einhergehenden standardisierten Einsatz von Metadaten deutlich wird. Die Einbindung zeitabhängiger Medien (z. B. AV-Aufzeichnungen) in Informationssysteme gestaltet sich im Vergleich zu zeitunabhängigen Dokumenten jedoch komplexer.

Die Eigenschaften eines Präsentationsaufzeichnungssystems werden daher vorwiegend bestimmt durch die im Folgenden erläuterten Punkte Produktionsaufwand, Anwendungsbereiche und Repräsentation.

Produktionsaufwand

Betrachtet man die Prozesskette für die Umsetzung vom Vortrag zur elektronisch übermittelbaren Präsentation, so stellt man fest, dass die einzelnen Arbeitsschritte Vorbereitung, Aufzeichnung, Nachbearbeitung und Publikation je nach Art und Umfang der Umsetzung erheblichen manuellen bzw. personellen Aufwand verursachen können.

Hier ist zu untersuchen, inwieweit Prozessschritte vereinfacht oder automatisiert werden können. Auch ist zu berücksichtigen, dass ein Mehraufwand bei der Erstellung eines Vortrags den Aufwand zur Nachbearbeitung verringern bzw. die Qualität der Umsetzung bestimmen kann.

Anwendungsbereiche

Wie bereits erwähnt, ist eine primäre Anwendung der Präsentationsaufzeichnung im Bereich E-Learning zu sehen. Hier gilt es, den maximalen Nutzen für Lernende und Lehrende zu erreichen. Die Aufzeichnungen können dazu auch als Lerneinheiten in computerbasierten Multimedia-Lernsystemen verwendet werden. Mit Learning-Management-Systemen (speziell auf die Anforderungen von Lerninhalten zugeschnittene Content-Management-Systeme) können die Publikation der Inhalte und die netzgestützte Schnittstelle zwischen Inhalt, Studierenden und Lehrenden realisiert werden.

Grundsätzlich lässt sich jede Art von Präsentation (Vortrag, Vorlesung, Kurs usw.) in eine elektronische Form bringen. Allerdings ist offensichtlich, dass nicht alle Arten von Vorträgen gleich gut umzusetzen sind. So ist z. B. eine Vorlesung, in der eine mathematische Herleitung über mehrere Tafelseiten entwickelt wird, mangels adäquater elektronischer Repräsentation nur unbefriedigend abzubilden. In Vorträgen, bei denen Folien in elektronischer Form eingesetzt werden, ist eine Umsetzung leichter realisierbar. In einigen Fällen können die eingesetzten Medien sogar direkt (d. h. ohne weitere Konvertierung) auch in der aufgezeichneten Version verwendet werden.

Repräsentation

Im allgemeinen Fall müssen die unterschiedlichen Medien des Vortrages in Formen gebracht werden, die für die elektronische Erfassung und das Publizieren geeignet sind. Bild und Ton des Dozenten sind ebenso aufzuzeichnen wie das Tafelbild oder gezeigte Folien. Diese verschiedenen Medientypen stellen unterschiedliche Anforderungen an das zu verwendende (Repräsentations-) Medienformat, z. B. hinsichtlich Auflösung und zeitlicher Varianz. Weiterhin sind

Restriktionen bezüglich des zu nutzenden Transportwegs (z. B. Rechnernetz, Datenträger) zu berücksichtigen.

Abgesehen von den unterschiedlichen Medientypen besteht ein weiteres Problem darin, dass der Betrachter in einer einfachen Vortragsaufzeichnung nicht *navigieren* kann, wie er es z. B. in einem Buch könnte – es gibt lediglich eine zeitliche Achse entsprechend des Ablaufs der Aufzeichnung, aber keine Orientierungshilfen wie ein Kapitelindex oder ein Stichwortregister. Entsprechende Funktionen sind aber gerade für die gezielte Nachbereitung einer Vorlesung hilfreich.

1.2 Anforderungen an Aufzeichnungssysteme

Die grundlegende Aufgabe eines Produktionssystems zur Aufzeichnung von Präsentationen besteht in der Transformation der originären („Live“-) Präsentation in eine Repräsentation für den Zugriff durch den Betrachter.

Erfassung der Informationen einer Präsentation

In jeder Präsentation werden Informationen auf unterschiedliche Weise vermittelt. Zunächst einmal gibt es das gesprochene Wort, die Gestik und die Mimik des Dozenten. Diese können mit bekannten Mechanismen auf Videoband oder Festplatte festgehalten werden.

Schriftliche oder bildliche Informationen werden häufig auf Folien geliefert. Neben einigen rechnerbasierten Systemen wie PowerPoint gibt es auch die Möglichkeit, dass die Folien mittels Overhead-Projektion dargestellt werden, sodass diese zunächst mittels Kamera oder Scanner digitalisiert werden müssen. Vielfach werden schriftliche und bildliche Aufzeichnungen vom Dozenten während des Vortrags entwickelt. Dabei gibt es auch hier neben der klassischen Form des „analogen“ Tafelanschriebs die digitale Form des Whiteboards.

In einigen Fachbereichen werden innerhalb der Präsentation Experimente unterschiedlicher Art durchgeführt, für die möglicherweise die für das Bild des Dozenten verwendete Video-Aufzeichnung qualitativ nicht ausreicht. In Informatik-Kursen werden z. B. Programmierübungen am Rechner vorgeführt.

Man hat es hier mit einer Vielzahl unterschiedlicher Szenarios zu tun, in denen verschiedene Medienarten mit unterschiedlichen Anforderungen an eine Repräsentation eingesetzt werden. Gemeinsam ist den Anwendungen die Audio- und Video-Aufzeichnung des Vortragenden. Für die eingesetzten Präsentationsmedien ist abzuwägen, ob eine jeweils zugeschnittene Lösung gefunden werden muss oder ob ein generischer Ansatz (z. B. das Abfilmen der Folien oder Erfassen des Bildschirminhalts) den Ansprüchen genügt.

Erzeugung der Repräsentationen

Für die elektronische Übermittlung müssen jeweils passende Transportmedien gefunden werden. Zum Teil können diese bereits während der Präsentation erzeugt werden (z. B. Screenshots von rechnergestützten Vortragsfolien), zum anderen Teil stellt das Erfassungsmedium lediglich ein Zwischenschritt dar (z. B. AV-Aufzeichnung auf Videoband), sodass das aufgezeichnete Material konvertiert werden

muss. Obendrein kann es sinnvoll sein, Medienformate unterschiedlicher Qualitätsstufen von derselben Aufzeichnung zu erzeugen, etwa um Versionen für den Internet-Zugriff und die Verwendung auf CD-ROM oder DVD zu generieren.

Einen Sonderfall stellen Formate dar, die sowohl in der Präsentation als auch in der Repräsentation eingesetzt werden können, z. B. Webseiten. Zu deren Aufzeichnung sind lediglich Metainformationen wie die jeweilige Adresse sowie Synchronisationsdaten zu erfassen. Der Betrachter der Aufzeichnung und das Auditorium bei der realen Präsentation sehen somit das identische Material.

Um die Aufzeichnung eines Vortrag möglichst zeitnah publizieren zu können, ist die für die Nachbearbeitung notwendige Zeit zu minimieren. Die Postproduktion muss daher weit gehend automatisiert ablaufen, der manuelle Anteil reduziert bzw. ganz unnötig werden und die Aufbereitung bereits während der Aufzeichnung erfolgen. Übrig bleiben dann eventuell notwendige Formatkonvertierungen, die aber lediglich Rechenaufwand verursachen.

Handhabung des Produktionssystems

Für den Vortragenden, das Produktionsteam und für den Betrachter soll das System leicht zu handhaben sein. Beim Dozenten soll bei der Aufzeichnung kein bzw. nur ein möglichst geringer Mehraufwand entstehen.

Anspruchsvolle Anwendungen erfordern insbesondere bei der Aufnahme eine manuelle Bedienung (z. B. der Kamerasteuerung). Auch für den Betreuer der Aufzeichnung und der eventuell erforderlichen Nachbearbeitung soll sich das System durch eine einfache Handhabung leicht in die Produktion integrieren lassen.

Der Betrachter als Endnutzer soll die Aufzeichnung intuitiv und gewinnbringend verwenden können. Dafür ist neben einfachen Nutzungsvoraussetzungen (z. B. installierter Software) die Repräsentation mit Navigationsfunktionen auszustatten.

1.3 Motivation und Ziele

Es gibt eine Reihe von kommerziellen und im Hochschulbereich entwickelten Lösungen für die oben skizzierte Aufgabenstellung, die unterschiedliche Stärken und Schwächen aufweisen. Teilweise sind sie für spezielle Anwendungen optimiert oder bieten nur eine eingeschränkte Flexibilität, z. B. weil sie auf hersteller-spezifischen Datenformaten aufsetzen.

Betrachtet man die Eigenschaften verfügbarer Systeme in Hinblick auf die Vielfalt verschiedener Anwendungsszenarios, so kristallisieren sich einige gravierende Nachteile heraus:

- Modulare, erweiterbare Lösungen sind nicht verfügbar. Effizient arbeitende und leicht zu handhabende Systeme binden den Dozenten oft an ein vorgegebenes Präsentationsmittel und erfordern zudem z. T. proprietäre Programme beim Dozenten und Betrachter. Erweiterungen oder Anpassungen, z. B. an individuelle Präsentationswerkzeuge oder AV-Codecs, sind nicht möglich.
- Generisch (für den Einsatz mit verschiedenen Präsentationsmitteln) ausgelegte Systeme weisen hingegen andere Nachteile auf. So arbeiten diese z. T. unime-

dial mit ineffizienter Ressourcennutzung und liefern keine für die Navigation und Publikation benötigten Metadaten zu den aufgezeichneten Inhalten.

- Integrierte Systeme schränken zugunsten einfacherer Handhabung das Einsatzspektrum auf Anwendungen ein, die den Dozenten aktiv in die Aufzeichnungsproduktion einbeziehen. Zudem sprechen oft technische und praktische Gründe gegen die Integration sämtlicher Aufzeichnungsfunktionen (Präsentation, AV-Aufzeichnung, Bearbeitung) in einem Gerät.

Aufgrund der Vielzahl unterschiedlicher Anwendungen und Anforderungen kann es *das* optimale System zur Aufzeichnung von Präsentation nicht geben. Mit dieser Feststellung entstand die Idee zur Entwicklung eines modularen Frameworks, dessen grundsätzliche Eigenschaften im Folgenden skizziert sind.

Das zu entwerfende System erhebt nicht den Anspruch, für jede Art der Anwendung eine Lösung zu bieten. Vielmehr besteht das Ziel darin, ein Framework zu entwickeln, dessen Module für ein gegebenes Szenario optimal kombiniert werden können und das bei Bedarf für individuelle Anforderungen erweitert werden kann. Neben der Erfüllung der in Abschnitt 1.2 erläuterten grundlegenden Anforderungen werden bei der Entwicklung des Systems folgende Schwerpunkte gesetzt:

- Flexibler Einsatz durch die Anpassung an diverse Präsentationsszenarios (d. h. mit oder ohne aktive Einbeziehung des Dozenten in die Aufzeichnungsproduktion) sowie an verschiedene Präsentationswerkzeuge und Zugriffsverfahren.
- Integration verfügbarer Systeme in den Produktionsfluss (Präsentation, Aufzeichnung, Nachbearbeitung und Distribution).
- Modulares, verteiltes System, basierend auf offenen Schnittstellen, Standard-Formaten und Metadaten. Erweiterung durch Formate und Werkzeuge.
- Minimierung der manuellen (Nach-) Bearbeitung und weit gehende Automatisierung für kurze Produktionszyklen typischer Anwendungsszenarien.
- Effiziente Nutzung des aufgezeichneten Materials durch Navigation auf Kapitel- bzw. Folienebene. Multimediale Repräsentationen in Standard-Formaten können online und offline genutzt werden.

Einige der dabei zu für die Systemarchitektur und die Repräsentation berücksichtigten Aspekte sind im Folgenden erläutert.

Repräsentation

Aufgrund der unterschiedlichen Parameter von Präsentationsmedien und AV-Aufzeichnung des Dozenten sind für die effektive Übermittlung eines Vortrags unterschiedliche, parallel verwendete Medienformate notwendig.

Die einzelnen Medienströme (z. B. AV-Aufzeichnung und Folien) sind – wie in der Live-Veranstaltung – synchron, *multimedial* darzustellen. Eine Hauptaufgabe des Produktionssystems liegt daher bei der Erfassung von Metainformationen, die bei der Wiedergabe die Synchronisation der Medienströme gewährleisten.

Als „Mehrwert“ im Vergleich zur Präsenzveranstaltung kann die Repräsentation *navigierbar* gestaltet werden. Während der Zuschauer vor Ort den Vortrag von

Anfang bis Ende verfolgt, kann es dem Betrachter der Aufzeichnung ermöglicht werden, gesuchte Inhalte direkt anzusteuern. Er kann sich somit neben der Zeitachse auch wie in einem Buch an einem Index orientieren.

Den Ansatz für die Realisierung dieser Funktion liefert der Aufbau üblicher Vorträge selbst. Typische Präsentationen enthalten Material in textueller Form auf Folien oder im Tafelanschrieb, das die benötigten Metainformationen wie Kapitelnamen oder Stichworte enthält. Gelingt es, zum gesuchten Begriff die passende Folie oder das entsprechende Tafelbild zu assoziieren, so kann aufgrund der Synchronität auch die entsprechende Position auf der AV-Zeitachse gefunden werden. Umgekehrt kann auch der Betrachter, der sich an der Zeitachse der AV-Aufzeichnung orientiert oder sich die Aufzeichnung komplett ansehen möchte, auf das Sekundärmaterial (z. B. Folien) zum gerade erläuterten Inhalt zugreifen.

Systemarchitektur

Dozenten setzen für ihre Vorträge unterschiedliche Werkzeuge zur Präsentation auf unterschiedlichen Plattformen ein, die auch bei der Produktion der Aufzeichnung zum Einsatz kommen. Ferner bestehen abhängig vom Produktionsszenario (z. B. Tagung oder Vorlesung) verschiedene Möglichkeiten des Zugriffs auf das verwendete Präsentationsmittel. Für die jeweils optimale Einbindung dieser Systeme in die Aufzeichnung sind somit unterschiedliche Verfahren unter Ausnutzung geeigneter Schnittstellen zu implementieren.

Typische Werkzeuge für die Darstellung von Folien sind Microsoft Powerpoint, Acrobat Reader (für PDF) oder auch Web-Browser für die Darstellung von HTML-Seiten. Interaktive Systeme wie Whiteboards und deren zeitlich exakte Reproduktion in Aufzeichnungen stehen hingegen nicht im Fokus der Arbeit.

Für die Aufzeichnung und Distribution von audiovisuellen Medien existieren etablierte Systeme und Codecs. Da sich diese in stetiger Entwicklung befinden, und z. B. auch an sich ändernde Netz- und Rechnerleistungen angepasst werden, empfiehlt sich die modulare Anbindung dieser Systeme, um auch zukünftige Entwicklungen nutzen zu können.

Modifikationen und Anpassungen an aktuelle Werkzeuge und AV-Systeme sind wünschenswert bzw. unumgänglich. Um das System flexibel einsetzen zu können, ist es daher auf Erweiterbarkeit mit *offenen Schnittstellen* auszulegen.

Der Einsatz von *Standards* erleichtert die Integration von Modulen. Für den Betrachter haben Standardformate den Vorteil, dass für die Wiedergabe der Repräsentation keine spezielle Software installiert werden muss. Idealerweise werden unterschiedliche Betriebssysteme (Windows, Unix, Mac) unterstützt.

Aufgrund der unterschiedlichen Aufgaben während der Aufzeichnung bietet sich die Realisierung als *verteilt System* an. Auf diese Weise können mehrere, optimal ausgestattete Rechner parallel eingesetzt werden – z. B. ein Laptop für die Darstellung von Folien beim Dozenten und ein mit AV-Schnittstellen ausgestatteter PC für die AV-Aufzeichnung. Hierfür ist jedoch ein geeignetes Mittel zur Synchronisation bereitzustellen.

1.4 Aufbau

Das Kapitel „Grundlagen“ liefert einen Überblick zu verschiedenen Basistechnologien, auf denen das System aufsetzt. Neben Grundlagen zur Medientechnologie (Formate, Transport, Produktion) wird darin speziell auf die für die Aufzeichnung benötigte Infrastruktur eingegangen. Daneben werden die Aspekte Systemintegration, Synchronisation und Metadatenmodelle erörtert, welche die Basis für das später beschriebene Systemdesign bilden.

Im Kapitel „Problemanalyse“ werden zunächst Anwendungsszenarios beschrieben und analysiert. Anschließend folgt die Diskussion und Klassifikation existierender Lösungsansätze bzw. Systeme und Anwendungen. Aus diesen Analysen werden die Ziele und Systemeigenschaften des Capturing-Frameworks herausgearbeitet.

Das Kapitel „Entwurf und Implementierung“ beschreibt den Systementwurf und Aspekte der Implementierung. Hierin wird die Produktionskette (Aufzeichnung, Nachbearbeitung, Distribution) des Systems beschrieben und das aus dem Präsentationsmodell entwickelte Metadatenmodell erläutert. Der Abschnitt „Implementierung“ beleuchtet Aspekte der im Rahmen der Arbeit entwickelten Software-Module.

Im Kapitel „Anwendung, Performance und Bewertung“ folgen eine Systembewertung anhand von beispielhaften Anwendungsszenarios, Betrachtungen zum Aufwand sowie Messungen zum Zeitverhalten einzelner Komponenten.

Das Kapitel „Zusammenfassung und Ausblick“ enthält abschließend eine Übersicht zu möglichen zukünftigen Erweiterungen und Einsatzbereichen.

Der Systementwurf und die Implementierung wurden vom Autor zum Teil am Forschungszentrum L3S im Rahmen eines Moduls des Projektes PADLR (Personalized Access to Distributed Learning Repositories [51]) durchgeführt. Ferner war der Autor an der Universität Hannover u. a. aktiv am Aufbau der im Grundlagenkapitel beschriebenen Medientechnik-Infrastrukturen beteiligt, die die Basis-Plattformen für die thematisierten Systeme bilden.

2 Grundlagen

2.1 Medientechnologie

Dieser Abschnitt liefert eine kurze Einführung in die für diese Arbeit relevanten Grundbegriffe Multimedia, Präsentation und Repräsentation. Anhand der generischen Multimedia-Produktionskette werden die Zusammenhänge zwischen Live-Präsentation, Aufzeichnung, der herzustellenden Repräsentation und den dabei jeweils eingesetzten Medientypen und -formaten verdeutlicht.

2.1.1 Grundbegriffe

Da für einige Begriffe unterschiedliche Definitionen bzw. Auslegungen existieren, wird in der folgenden Zusammenstellung explizit auf die in diesem Kontext verwendete Bedeutung eingegangen.

Den Begriff *Multimedia* definiert Steinmetz wie folgt: „*Ein Multimedia-System ist durch die rechnergestützte, integrierte Erzeugung, Manipulation, Darstellung, Speicherung und Kommunikation von unabhängigen Informationen gekennzeichnet, die in mindestens einem kontinuierlichen (zeitabhängigen) und einem diskreten (zeitunabhängigen) Medium kodiert sind.*“ [69]

Der Autor betrachtet auch ein System, das mindestens *zwei kontinuierliche* Medien parallel verarbeitet als multimedial. Als Beispiel sei hier die Audio/Video-Aufzeichnung eines Vortrags zusammen mit der (synchronisierten) Video-Aufzeichnung der Bildschirmpräsentation angeführt.

Während der Begriff *Medium* im Allgemeinen ein Mittel zur Informationsverbreitung und -darstellung beschreibt, werden in [25] die Begriffe Perzeptions-, Repräsentations-, Präsentations-, Speicher-, Übertragungs- und Informationsaustauschmedium differenziert. *Präsentationsmedien* bezeichnen darin Mittel zur Ein- und Ausgabe von Informationen wie z. B. Kamera und Mikrofon bzw. Bildschirm und Lautsprecher. Das *Repräsentationsmedium* beschreibt die Kodierung bzw. das Format der Informationen, z. B. ein Bild im JPEG-Format oder eine Videosequenz in MPEG.

Aufgrund der Begriffsverwandtschaft mit der eben eingeführten Terminologie werden nachfolgend die Bedeutung der Bezeichnungen Präsentation und Repräsentation für die hier bedeutsame Produktionssicht definiert.

2.1.2 Medienproduktionskette

Während mit *Präsentation* – angelehnt an den allgemeinen Sprachgebrauch – die Live-Veranstaltung (Vortrag, Vorlesung usw.) bezeichnet wird, stellt die *Repräsentation* die daraus produzierte Aufzeichnung dar – sie entspricht damit der Abbildung der Präsentation in digital speicher- bzw. übertragbarer Form.

Somit werden bereits bei der Präsentation *Repräsentationsmedien*, etwa für die Darstellung von Vortragsfolien, verwendet. Diese Inhalte werden auch in der Repräsentation auftauchen – im gleichen oder einem anderen Repräsentationsmedium. So können die ursprünglich bei der Präsentation im PowerPoint-Format verwendeten Vortragsfolien bei der Repräsentation z. B. im PNG-Format vorliegen. Ferner können während der Verarbeitung (weitere) Repräsentationsmedien zum Einsatz kommen, die für die spätere Repräsentation konvertiert werden; z. B. kann die Bildschirmpräsentation in Form von Bitmap-Bildern aufgezeichnet werden, die bei der Verarbeitung in platz sparende PNG-Bilder umgewandelt werden. Umgekehrt wird natürlich auch die Repräsentation wieder über Präsentationsmedien (Bildschirm, Lautsprecher) ausgegeben. Der Zusammenhang der hier genutzten Begriffe ist in Abbildung 2-1 verdeutlicht.

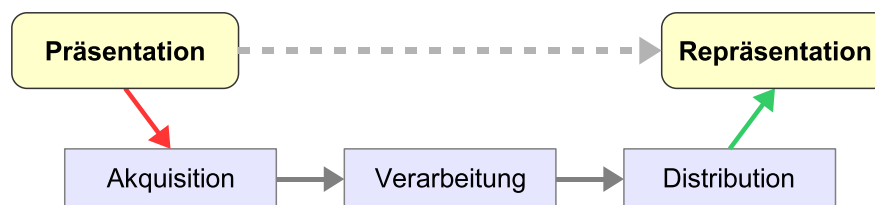


Abbildung 2-1: Modell einer Medienproduktionskette

Das allgemeine Modell gilt für einzelne Medientypen (z. B. Video) wie auch für die Multimedia-Produktion; dort fallen mehrere, parallele Akquisitions- und Verarbeitungs-Schritte an. Das Ergebnis für die Distribution (Verteilung) kann uni- oder multimedial sein. So können etwa ein Video und eine 3D-Animation in eine Multimedia-Präsentation umgesetzt werden oder die 3D-Animation wird – nach der Konvertierung in eine Videosequenz – mit in das Video eingebunden.

Akquisition

Zur Akquisition gehört neben der Aufzeichnung von realem Audio- und Videomaterial auch die Erfassung von rechnergenerierten Bildsequenzen sowie die Erzeugung von Bildmaterial aus Simulationsergebnissen.

Verarbeitung

In den Bereich Verarbeitung (Nachbearbeitung, Postproduktion) fallen sowohl manuell gesteuerte Arbeitsschritte (hier als Post-Production bezeichnet) als auch weit gehend automatisch ablaufende Vorgänge (Post-Processing), wobei die Übergänge nicht immer klar definierbar sind.

Distribution

Die Distribution bzw. Publikation bezeichnet die Aufbereitung der für die Repräsentation erforderlichen Datenstrukturen für den Zugriff des Betrachters.

Im Kontext der Produktionskette entspricht dies der Abbildung der verarbeiteten Live-Präsentation (*Eingabe*) auf die reproduzierbare Repräsentation (*Ausgabe*).

2.1.3 Repräsentationsmedien

Gemäß der in Abschnitt 2.1.1 erläuterten Differenzierung beschreibt dieser Abschnitt Aspekte der Kodierung und Formate von *Repräsentationsmedien*. In dem in Abschnitt 2.1.2 skizzierten Produktionsszenario sind diese sowohl beim Einsatz in der Live-Präsentation durch den Vortragenden als auch bei der Erstellung der Repräsentation von Bedeutung, z. B. für die Auswahl geeigneter Formate zur Abbildung des Originals.

Zeitverhalten

Neben den qualitätsbestimmenden Parametern bezüglich der Auflösung (zeitlich, räumlich, Farbtiefe etc.) ist insbesondere das grundlegende Zeitverhalten eines Mediums (*nicht* die zeitliche Auflösung wie z. B. die Abtastrate) relevant.

Medien wie Bilder oder Texte verhalten sich bezüglich der Zeit *diskret*, Audio und Video *kontinuierlich*. (Zwar erfordert die Digitalisierung stets eine Orts-, Wert- und Zeitdiskretisierung; im Kontext dieser Arbeit wird jedoch die Terminologie nach [69] verwendet.)

Während Audio- und Videoformate ein festes Zeitraster aufweisen (z. B. 44100 Samples oder 25 Bilder pro Sekunde), sind Medientypen ohne feste Abtastfrequenz nicht eindeutig einer Klasse zuzuordnen. So erhält beispielsweise eine (statische) PowerPoint-Präsentation durch die Vorführung des Vortragenden eine zeitliche Komponente, indem den einzelnen Seiten (bzw. Seitenelementen) konkrete Zeitmarken zugewiesen werden. Ein weiteres Beispiel ist das Whiteboard, dessen Inhalte ebenfalls dynamisch, aber ohne festes Zeitraster verändert werden.

Dieses Zeitverhalten gewinnt eine besondere Bedeutung bei der Wahl geeigneter Formate für die Repräsentation dieser Medientypen.

Qualitäts-Parameter

Die Qualität eines Formats wird wesentlich durch die Auflösung bestimmt. Bei diskreten Medien betrifft dies den Wertebereich (z. B. die örtliche Auflösung und die Anzahl unterscheidbarer Signalwerte), bei kontinuierlichen Medien zusätzlich die zeitliche Auflösung. In Tabelle 2-1 sind die wesentlichen Parameter für die Medien Bild, Video und Ton aufgeführt.

Typ	Kriterium	Metrik	Beispiel
Einzelbild	örtliche Auflösung	Pixel	1024*768 (Projektor)
	Farbtiefe	bit/Pixel	8, 16, 24
Video	örtliche Auflösung	Pixel	768*576 (PAL-TV)
	Farbtiefe	bit/Pixel	16 (PAL-TV, 4:2:2)
	Bildwiederholrate	Bilder pro Sekunde	25 (PAL-TV)

Tabelle 2-1: Medienparameter (Beispiele)

Typ	Kriterium	Metrik	Beispiel
Audio	Auflösung (Wert)	bit/Sample	16 (Audio-CD)
	Abtastrate	Frequenz (1/s)	44100 (Audio-CD)
	Kanäle	Anzahl	2 (Stereo)

Tabelle 2-1: Medienparameter (Beispiele)

Durch – verlustbehaftete – Kompression wird die Qualität zugunsten einer geringeren Datenmenge reduziert (Relevanzreduktion). Je nach angewendetem Verfahren ergibt sich hierbei innerhalb des Mediums eine örtlich und/oder zeitlich variable Auflösung, z. B. unterschiedlich scharfe Bildbereiche innerhalb eines Bildes und/oder schwankende Bildwiederholraten in einem Video.

Während sich bei unkomprimierten Formaten die Datenmenge leicht errechnen lässt, erhält man bei komprimierten Formaten eine vom Inhalt abhängige Qualität oder Datenmenge. Bei Videocodecs ist häufig eine konstante Bitrate erforderlich (etwa zur Einhaltung einer Übertragungskapazität), die bei Interframe-Kodierung insbesondere bei dynamischem Material zu verminderter Bildqualität führen kann.

Zusätzlich wird vor der Anwendung von verlustbehafteten Verfahren zur weiteren Reduzierung der Datenrate oft auch die ursprüngliche Auflösung herabgesetzt, z. B. arbeitet MPEG-1 [24] im Vergleich zu CCIR-601 mit einer Halbierung von Höhe und Breite sowie einer anderen Farb-Unterabtastung (Subsampling, 4:2:0 statt 4:2:2 [52]). Weitere qualitätsbeeinflussende Parameter wie z. B. die Verwendung von Zeilensprungverfahren bei Videobildern sollen hier nicht weiter betrachtet werden.

Typische AV-Formate

In der Tabelle 2-2 sind einige der hier relevanten Datenformate zusammengestellt. Aus den aufgestellten Werten der verlustbehafteten Kompression kann kein unmittelbarer Rückschluss auf die subjektive Bewertung der Qualität getroffen werden.

Für Bilddaten wurden hier Rasterbild-Formate wie JPEG [23] aufgeführt, um einen Vergleich mit den sich daraus ableitenden Videoformaten zu ermöglichen. Nicht berücksichtigt sind hier Vektordatenformate wie SVG [83], die eine beliebige Skalierung ohne Verlust der Darstellungsqualität ermöglichen, aber für z. B. mittels Kamera erfasster Bildinhalte nicht anwendbar sind.

Der Vergleich der erforderlichen Datenmengen erklärt die Notwendigkeit der verlustbehafteten Kompression. So ist etwa mit einer verlustfreien Audio-Kodierung lediglich ein Kompressionsfaktor von 1,5 bis 4 erreichbar [18]. Verlustbehaftete Verfahren erreichen Kompressionsraten von mehr als 1:10 bei – abhängig vom Material – unhörbaren bzw. tolerierbaren Qualitätsverlusten.

Im Videobereich sind die Unterschiede noch deutlicher; ohne Bildgrößen-Reduktion sind z. B. mit Hilfe von MPEG-2 [26] Kompressionsraten von 1:28 möglich. Aktuelle Codecs wie AVC ermöglichen noch geringere Bitraten bei vergleichbarer Qualität [71].

Die Qualität im Rahmen eines Kompressionsverfahren ist durch verschiedene Parameter beeinflussbar; neben der Wahl der Bildgröße vornehmlich durch die Bitrate [88]. Abhängig vom Anwendungsfall ist zu untersuchen, bei welchen Parametern die verminderte Wiedergabequalität zugunsten effizienter Ausnutzung von Ressourcen tolerierbar ist [63].

Typ	Format	Kompression	typ. Parameter	typ. Datenaufkommen ^a
Bild	RGB-Bitmap	keine	1024*768 Pixel, 24 bit/Pixel	2360 KBytes
	JPEG	für reale Bilder ausgelegt, ver- lustbehaftet	1024*768 Pixel, 24 bit/Pixel	synthetisches Bild: 100 KBytes (-) reales Bild: 50 KBytes (+)
	GIF	für synthetische Bilder ausgelegt, verlustlos (8 bit)	1024*768 Pixel, 8 bit/Pixel	synthetisches Bild: 90 KBytes (+) reales Bild: 45 KBytes (-)
	PNG	universell	1024*768 Pixel, 24 bit/Pixel	synthetisches Bild: 90 KBytes (++) reales Bild: 1100 KBytes (++)
Video	Referenz: PAL CCIR 601 („TV“)	keine, aber Subsampling 4:2:2	720*576 Pixel, 16 bit/Pixel, 25 fps	166 Mbit/s (1:1)
	DV-PAL	Intraframe, M-JPEG-ähn- lich, 4:2:0	720*576 Pixel, 12 bit/Pixel, 25 fps	25 MBit/s (1:6.6)
	MPEG-2	Interframe, JPEG-ähnlich, 4:2:0	720*576 Pixel, 12 bit/Pixel, 25 fps	6 Mbit/s (1:28)
	MPEG-1	Interframe, JPEG-ähnlich, 4:2:0	352*288 Pixel, 12 bit/Pixel, 25 fps	1,2 Mbit/s (1:140)
	Real, WM, QT, MPEG- 4, AVC, u. a.	verschiedene (Interframe u. a.)	diverse Auflö- sungen, variable Framerate	50 kbit/s..1 Mbit/s (1:3320..1:166)
Audio	Referenz: CD-Audio	keine	16 bit, 2 Kanäle, 44100 Hz	1,4 Mbit/s (1:1)
	Lossless (LPAC u. a.)	verlustfrei	16 bit, 2 Kanäle, 44100 Hz	0,7 MBit/s (1:2)
	MPEG Layer 3 („mp3“)	div. Verfahren, verlustbehaftet	16 bit, 2 Kanäle, 44100 Hz	0,125 MBit/s (1:11)

Tabelle 2-2: Parameter typischer Medienformate

a. Qualität: ++=sehr gut, +=gut, -=schlecht; (1:x): Verhältnis zur unkomprimierten Version

Produktions- und Distributionsformate

Gerade im Videobereich wird der Unterschied zwischen *Produktions-* und *Distributionsformaten* deutlich [37]. Während bei der Produktion primär die Eigenschaften des Formats bezüglich Bearbeitung relevant sind, kommt es bei der Verteilung wesentlich auf eine effiziente Kodierung und die damit verbundene geringe Übertragungskapazität bzw. Speichermenge an. Letztere wird bei Videodaten wesentlich durch die Anwendung von Interframe-Kodierung erreicht, wobei die Daten einzelner Bilder aus vorangehenden und/oder folgenden Bildern berechnet werden. Das erschwert zwar das bildgenaue Editing, stellt für das Abspielen aber keine wesentlichen Nachteile dar (der wahlfreie Zugriff stützt sich auf die Intraframes).

Als Standard für die semiprofessionelle Videoaufzeichnung und -bearbeitung hat sich in den letzten Jahren das DV (DigitalVideo) Format etabliert. Selbst die Consumer- bzw. weit gehend kompatible DVCAM-Variante liefert bei 25 Mbit/s mit M-JPEG-ähnlicher Kompression für die meisten Anwendungen eine gute Qualität und kann mit preisgünstigen Systemen (Camcordern, Recordern, PC-basierten Schnittsystemen) erstellt und verarbeitet werden. Neben der Möglichkeit des verlustfreien, bildgenauen Schnitts bietet es sich als ein Ausgangsformat für die Konvertierung in Transportformate geringerer Qualität und niedriger Bitraten (wie MPEG-1/2/4 oder die proprietären Formate) an.

2.1.4 Transport

Für die Verteilung bzw. den Transport von Daten existieren zwei grundsätzliche Ansätze. Zum einen die Verwendung von Datenträgern, zum anderen die Nutzung von Netzen. Der Aufbau leistungsfähiger und preisgünstiger Rechnernetze führte dazu, dass zunehmend auch größere Datenmengen (z. B. Software-Pakete) online verteilt werden, während vor einigen Jahren CDs verschickt wurden. Aufgrund der Eigenschaften von Multimediadaten (hohe Datenmengen bzw. -raten) sind jedoch Offline-Datenträger hierfür nach wie vor von Bedeutung.

Offline

Neben den generischen Datenträgern wie CD-ROM und DVD-ROM gibt es für AV-Inhalte Derivate, die auch das Abspielen auf Geräten ermöglichen, die keinen PC voraussetzen. Neben den weit verbreiteten Formaten CD-Audio und DVD-Video ist hier noch die Video-CD als Vorgänger der DVD-Video mit geringerer Datenrate bzw. AV-Qualität und geringerem Funktionsumfang zu nennen. Dazwischen rangieren die nicht-standardisierten Formate S-Video-CD und Mini-DVD.

Zwar enthält das DVD-Video-Format Funktionen, die über die einfache Wiedergabe von Ton und Bild hinausgehen (umschaltbare Bildspuren, nahtlose Verzweigungen alternativer Fassungen, mehrere Tonspuren, Menüs etc. [13]). Sie ist aber auf das Ausgabegerät TV-Monitor ausgelegt und bietet daher keine Multimediafähigkeit wie etwa die parallele Darstellung unterschiedlicher Inhalte. Für diesen Zweck ist die Datenvariante einzusetzen – und damit die Benutzung an einem PC vorgegeben.

Speziell für die AV-Produktion ist aufgrund der hohen Datenraten auch das Band ein relevantes Transportmedium, da es die preisgünstige Speicherung großer

Datenmengen bei gleichzeitig einfachem – wenn auch nur linearem – Zugriff mittels Videorecorder ermöglicht. Zudem erlauben digitale Formate die verlustfreie Übertragung zwischen Band und Festplatte.

Online

Während für die Produktion Intranet-Netzprotokolle für verteilte Dateisysteme wie SMB oder NFS analog zu Standard-Fileserver-Anwendungen eingesetzt werden können, bietet sich für die Publikation von Multimediadaten das World Wide Web (WWW) als eine geeignete Plattform an.

In HTML-Dokumenten [62] lassen sich eine Vielzahl von Medienformaten einbinden. Während Bildformate wie PNG, JPEG und GIF nativ von Web-Browsern dargestellt werden, stehen für andere Datenformate (wie AV, 3D u. a.) Plugins zur Verfügung, die als Komponente innerhalb des Web-Browsers laufen und damit eine nahtlose Einbindung von Medienformaten in Web-Seiten ermöglichen. Weitere Datenformate können vom Browser lokal zwischengespeichert und mittels assoziiertem Viewer extern angezeigt werden. HTML-Dokumente lassen sich zudem so gestalten, dass sie auch für die Offline-Nutzung geeignet sind.

Neben dem dateiorientierten Zugriff (*Download*) über HTTP [61] bietet sich insbesondere für AV-Clips mit größerem Datenvolumen die Nutzung von Streaming-Protokollen (Datenübertragung: RTP [57], Steuerung: RTSP [59]) an. Die damit implementierten Zugriffsverfahren ermöglichen zum einen den wahlfreien Zugriff ohne die Notwendigkeit der kompletten Dateiübertragung sowie zum anderen eine Anpassung an aktuelle Netzparameter (Bandbreite, Paket-Verlustrate) und die Leistungsfähigkeit des Client-Systems.

2.2 Basissysteme

Dieser Abschnitt beleuchtet die Soft- und Hardware, auf der die in Abschnitt 2.1.2 skizzierte Produktionskette basiert. Hierzu gehören die lokale Infrastruktur für Präsentation und Aufzeichnung, Systeme zur Nachbearbeitung und zentrale Infrastruktur zur Distribution wie z. B. Streaming-Server.

2.2.1 Präsentationsmethoden

Klassische Präsentationsmethoden für Vorträge beinhalten die Visualisierung z. B. mittels Anschrieb auf Tafel, Whiteboard, Flip-Chart oder Overhead-Folien sowie die Darstellung vorgefertigter Inhalte mit Overhead- oder Diaprojektor.

Für die genannten konventionellen Systeme finden sich elektronische, meist PC-gestützte Pendanten, die zum einen zusätzliche Funktionen bieten, zum anderen aber einige Anwendungen nicht oder nur unzureichend nachbilden können. Hierzu gehören – etwa für aufwendige mathematische Herleitungen – insbesondere der Anschrieb über mehrere Tafelseiten.

Die elektronischen Varianten eignen sich jedoch besonders für die Aufzeichnung, da hier im Vergleich zu konventionellen Präsentationen mit Tafel oder Overhead-Projektion ein direkter Zugriff auf die Präsentationsmedien möglich ist; die Signale liegen bereits in elektronischer oder digitaler Form vor.

Für die oft parallel zur Aufzeichnung stattfindende Live-Übertragung von Präsentationen ist die Nutzung von elektronischen Verfahren obligatorisch.

Ein typisches Live-Präsentationssystem (Abbildung 2-2) besteht aus einem PC mit diverser Software, an den eine Projektion und eine Audio-Anlage zur Ausgabe von Bild und Ton angeschlossen sind. Hinzu kommen, abhängig von der eingesetzten Software, Eingabegeräte zur Steuerung des PCs. Alternativ ermöglicht ein Videopresenter die konventionelle Folienpräsentation an einem Projektor.

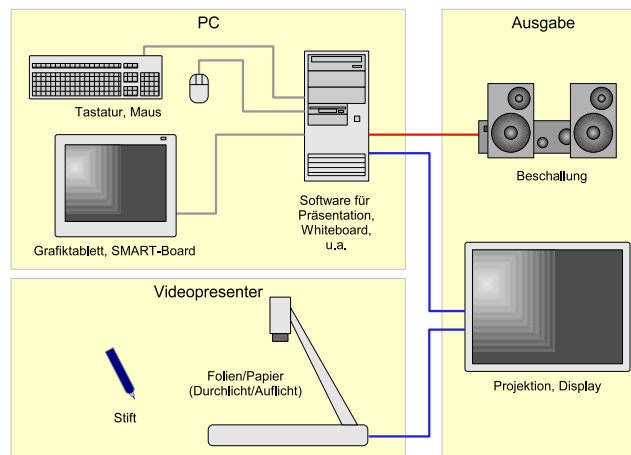


Abbildung 2-2: Schema eines projektorgestützten Präsentationssystems

Projektion

Typischerweise sind Seminarräume und Hörsäle mit mindestens einer an die Raumgröße angepassten Projektionsfläche, die mittels Auf- oder Rückprojektionstechnik von einem Beamer bestrahlt wird. In kleinen Seminarräumen kommen teilweise Plasmabildschirme zum Einsatz, die denselben Zweck erfüllen.

Vergleicht man das Verhältnis von Projektionsdiagonale zum Betrachtungsabstand, so lassen sich hier nicht die Maßstäbe eines Bildschirmarbeitsplatzes anwenden. Während beim Monitor ein Betrachtungsabstand etwa mit der zweifachen Bildschirmdiagonale abgeschätzt werden kann (15-Zoll Display, 70 cm Abstand), ergeben sich in Hörsälen und Seminarräumen oft Betrachtungsabstände vom drei- bis sechsfachen der Projektionsdiagonale. Um eine vergleichbare Lesbarkeit zu gewährleisten, werden Darstellungsgrößen von Schrift und Grafiken oft höher gewählt als für die Darstellung am Monitor mit gleicher physikalischer Auflösung.

Unabhängig von der Größe der Projektionsfläche beträgt die verbreitet eingesetzte Auflösung 1024*768 Bildpunkte; seltener stehen Auflösungen von 1280*1024 oder 1366*1024 zur Verfügung.

Audio

In seltenen Fällen enthalten Bildschirm-Präsentationen Audioinhalte. Wenn das der Fall ist, dann oft in Verbindung mit Bewegtbildern, etwa bei Einbindung eines Videoclips in die Präsentation. Zur adäquaten Wiedergabe ist eine der Raumgröße entsprechende Beschallung einzusetzen.

In größeren Hörsälen (ca. ab 80 Sitzplätzen) steht häufig für Vorträge ein Beschallungs-System für Mikrofone zur Verfügung, in das auch das PC-Signal mit eingespeist werden kann.

Eingabegeräte

Zur Bedienung von Präsentationssoftware sind Standard-PC-Eingabegeräte geeignet. Im einfachsten Fall dienen Tastatur oder Maus zum Weiterschalten der Folien.

Speziell für die Nutzung von Whiteboard-Funktionen bieten sich Geräte zur Eingabe per Stift oder Fingerspitze an. Hierzu gehören Geräte wie Grafik-Tabletts oder berührungsempfindliche Projektionsflächen (z. B. SMART-Boards).

Grafiktablets, insbesondere wenn sie in ein Display integriert sind, ermöglichen die einfache Eingabe von Handskizzen und können damit in einigen Bereichen den klassischen Tafelanschrieb ersetzen. Zwar ist die physikalische Auflösung im Vergleich mit einem Flip-Chart oder einer Tafel deutlich geringer. Die elektronische Lösung bietet jedoch Vorteile, speziell wenn bereits im Rechner vorliegende Bild-daten annotiert oder grafisch ergänzt werden sollen.

Eine berührungsempfindliche Projektionsfläche erfüllt die Funktion eines Grafiktablets mit dem Unterschied, dass der Vortragende direkt auf der projizierten Fläche arbeiten kann. Dies entspricht eher dem klassischen Tafelansatz, allerdings verdeckt der Vortragende in der Regel die Anzeige für die Zuschauer.

Neben der grafischen Eingabe können diese Boards die Funktionen von Maus und Tastatur nachbilden, sodass der Vortragende zur Bedienung lediglich seine Hände einsetzen muss, was abhängig vom Vortragsstil von Vorteil sein kann.

Videopresenter

Als Bindeglied zwischen Overhead-Folien und Projektion kann ein Videopresenter eingesetzt werden, der aus einer Auf- und/oder Durchlichtauflage für Papier- bzw. Folienvorlagen und einer Kamera besteht. Das aufgenommene Bild steht als TV-Video- und/oder VGA-Signal zur Verfügung. Neben dem Anschluss an einen Projektor steht hiermit das Bild auch für die Videoaufzeichnung zur Verfügung.

Video-Presenter kommen hauptsächlich dann zum Einsatz, wenn eine Vielzahl Vortragender mit unterschiedlichen Präsentationsmedien unterstützt werden muss (z. B. bei Tagungen) oder wenn in Räumen keine Fläche für die Overhead-Projektion zur Verfügung steht. Die Benutzung entspricht weitgehend der eines Overhead-Projektors. So können vorgefertigte oder leere Folien auch während des Vortrags beschrieben werden.

Präsentations-Software

Mit 95 % Marktanteil ist zur Zeit die MS-Office-Komponente PowerPoint das verbreitetste Programm zur Darstellung von rechnergestützten Bildschirmpräsentationen [87]. Daneben gibt es eine Reihe von ähnlichen Programmen wie OpenOffice.org/StarOffice Impress für Windows und Linux sowie Apples Keynote für MacOS. Gemeinsam ist diesen Programmen, dass sie die Funktionen einer Diashow bzw. Overhead-Projektion nachbilden. Die Programme bzw. die zugehörigen Formate erlauben die Zusammenstellung von Folien (Dias, Slides), die mit diver-

sen grafischen Animationen angereichert werden können. Darüber hinaus können andere Medientypen wie z. B. Videoclips eingebunden und präsentiert werden.

Neben den spezialisierten Präsentations-Tools kommen aber auch universelle Formate wie HTML oder PDF zusammen mit passenden Anzeige-Programmen (z. B. Webbrowser, Acrobat Reader) zum Einsatz. Insbesondere wenn das zu präsentierende Material auch für andere Zwecke eingesetzt werden soll (etwa zur Nachbereitung in Kursen) oder bereits in diesen Formaten aufbereitet wurde, bietet sich dieses Verfahren an, um Mehraufwand zu vermeiden.

Whiteboard

Das Whiteboard bildet in erster Linie die klassische Wand-Tafel nach und ist insbesondere in Kombination mit den o. g. Grafiktablets oder berührungsempfindlichen Projektionsflächen einzusetzen. Alternativ kann man zur Bedienung von Whiteboard-Applikationen auch die Maus verwenden.

Neben der eigentlichen Funktion als Zeichenbrett zum Skizzieren stehen weitere Funktionen zur Verfügung, etwa das Laden von Bildern in das Whiteboard, die Annotation mit Text sowie diverse aus der Bildbearbeitung bekannte Zeichenfunktionen.

Sonstige Software

Grundsätzlich kann jede verfügbare Software zur Präsentation eingesetzt werden. Sie kann aber auch selbst Gegenstand der Präsentation sein, z. B. bei einem Textverarbeitungskurs. Auch bei Programmierkursen oder -vorlesungen sind z. B. Ein- und Ausgaben einer Text-Shell essenzieller Bestandteil der Veranstaltung.

2.2.2 AV-Aufzeichnungssysteme

Die hier beschriebenen Aufzeichnungssysteme dienen primär der klassischen Audio- und Videoaufzeichnung von realen Vortragsszenarios. Vielfach kommt diese Technik auch zur Aufnahme von Präsentationsdaten wie PC-Bildern, Overhead-Folien u. a. zum Einsatz.

Audio: Mikrofone, Mischung, Akustik

Zur Sprachaufnahme des Vortragenden eignen sich verschiedene Mikrofontypen abhängig vom Vortragsszenario. Wenn sich der Sprecher im Raum bewegt ist ein drahtloses Ansteck- oder Headset-Mikrofon empfehlenswert, um einen konstanten Abstand zwischen Mund und Mikrofon und damit einen gleichmäßigen Pegel bei gleichzeitiger Ausblendung von Umgebungsgeräuschen zu gewährleisten. Wechselt der Sprecher den Standort nicht (etwa an einem Pult), so können auch feststehende Mikrofone verwendet werden.

Treten gleichzeitig mehrere Sprecher auf oder sollen Rückfragen aus dem Auditorium aufgezeichnet werden, so bietet sich die Nutzung mehrerer Mikrofone an. Die einzelnen Signale sind dann mittels Mischpult individuell geeignet auszusteuern. Hier können auch weitere Tonquellen wie z. B. der Ton einer PC-Präsentation mit eingebunden werden.

Vorwiegend bei größeren Hörsälen ist aufgrund der Raumgröße eine Beschallungs-Anlage vorhanden. Das hier vorliegende Signal kann in den meisten Fällen auch zur Aufzeichnung verwendet werden.

Der Einsatz digitaler Audiomixer ermöglicht die gleichzeitige, individuelle Abmischung aller Tonquellen für verschiedene Senken wie lokale Beschallung, Aufzeichnung und Video-Konferenz.

Video: Kameras, Mischung, Beleuchtung

Analog zum Ton gilt der Einsatz von Kameras primär der Aufzeichnung des Vortragenden. Aber auch Experimente, Diskussionen u. a. sind per Videokamera zu erfassen. Einfache Anforderungen wie die Aufzeichnung eines am Pult stehenden Vortragenden sind mit einer einzelnen Kamera zu erfüllen. Sollen weitere Personen oder Versuchsaufbauten erfasst werden, ist der Einsatz weiterer Kameras sowie deren Steuerung nötig. Sobald mehr als ein Quell-Bildsignal vorliegt, ist zur störungsfreien Verarbeitung (Umschnitt, Mischung etc.) ein Videomischer erforderlich, der zudem vielfältige Möglichkeiten zur Bildgestaltung bietet.

Zur Aufzeichnung von Vorträgen werden vorzugsweise fernsteuerbare Kameras aus der Überwachungstechnik oder dem semiprofessionellen Studiobereich eingesetzt. Die zentrale Steuerung (Schwenken, Neigen, Brennweitenänderung etc.) ermöglicht die gleichzeitige Bedienung von mehreren Kameras durch eine Person. DV-Camcorder liefern eine vergleichsweise gute Bildqualität, sind aber mangels verfügbarer Schwenk-/Neigeköpfe nicht in fernsteuerbaren Systemen verwendbar.

Neben der Güte der Kameras ist die Art der Beleuchtung ausschlaggebend für die Bildqualität. Selbst qualitativ schlechte Kameras liefern bei perfekter Beleuchtung vergleichsweise gute Bilder, während in kritischen Beleuchtungssituationen (wie einem abgedunkelten Raum) nur mit sehr hochwertigen Kameras rauschfreie Bilder produziert werden können.

Ansätze zur Automatisierung von Kamera-Steuerung und Bildmischung können mit manuell bedienten Geräten vergleichbare Ergebnisse liefern, erfordern jedoch zusätzlichen technischen Aufwand [64].

Scan-Converter

Scan-Converter dienen der Anpassung diverser Parameter eines Video-Signals wie Bildwiederholfrequenz, Auflösung und Zeilensprungverfahren (Interlacing). Der primäre Einsatzbereich liegt in der Konvertierung eines vom Rechner gelieferten VGA-Bildsignals mit 1024*768 Pixel, 60 Hz, non-interlaced in ein Videosignal in PAL-Auflösung mit 25 Hz, interlaced (Abbildung 2-3).

Zwar hat diese Umwandlung prinzipbedingt eine Verringerung der Bildqualität zur Folge. Das Verfahren ermöglicht aber eine Erfassung des PC-Bildes ohne Zugriff auf den Rechner (bzw. das eingesetzte Betriebssystem) oder die genutzte Software.

Dies ist ein adäquates Verfahren insbesondere zur Aufzeichnung von Konferenzen, bei denen eine Vielzahl unterschiedlicher Laptops sowie ein Video-Presenter für diverse Präsentationen eingesetzt wird.

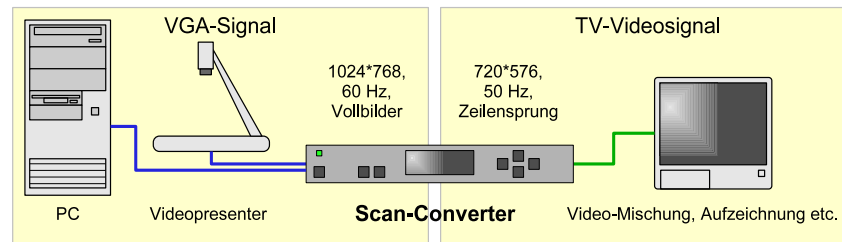


Abbildung 2-3: Funktionsschema eines Scan-Converters

Zudem kann das PC-Bild wie ein Kamerasignal benutzt werden und damit in die Abmischung des Videobildes eingebunden werden, was insbesondere dann notwendig ist, wenn nur dieses Medium für die Videoaufnahme oder -übertragung zur Verfügung steht.

Aufnahme und Speicherung

Während für rechnerbasierende Datenformate wie Präsentationen oder (Einzel-) Bilder per se Rechner als Speicher- bzw. Aufzeichnungsmedium eingesetzt werden, existieren für audiovisuelle Daten diverse Aufnahmesysteme mit unterschiedlichen Vor- und Nachteilen.

Die Aufnahme von Audio und Video wird hier gemeinsam betrachtet, da alle relevanten Verfahren die parallele Verarbeitung von Bild und Ton ermöglichen.

- Für die klassische Aufnahme auf Magnetband steht eine Reihe von Formaten unterschiedlicher Qualität zur Verfügung. Dabei wurde die analoge Aufnahme praktisch vollständig durch digitale Systeme ersetzt. Das im Consumer- und Semi-Profibereich weit verbreitete Format DV (DigitalVideo), arbeitet mit einer M-JPEG-ähnlichen Kompression, die einen Videodatenstrom von 25 Mbit/s liefert sowie wahlweise 2- oder 4-kanaligem PCM-Audio. Daneben gibt es zwei Profi-Derivate (DVCAM und DVCPRO mit 50 Mbit/s) sowie diverse, qualitativ höherwertige Formate.
- PC-Systeme mit Erweiterungen für den Eingang von analogen Audio- und Videosignalen (Sound- bzw. Video-Capturing-Karte) sind abhängig von Rechenleistung und Encoder-Komplexität für die Echtzeit-Kodierung geeignet. Die eingelesenen Daten werden von der Hardware digitalisiert, mit entsprechender Software kodiert und dabei unmittelbar dateibasiert abgespeichert.
- Über die IEEE 1394-Schnittstelle – auch Firewire (Apple) oder i.Link (Sony) genannt – kann das digitale AV-Signal im DV-Format direkt in den PC eingespielt und dateibasiert gespeichert werden. Die Analog/Digital-Wandlung wird hier bereits mit entsprechender Hardware (Wandlerbox, DV-Recorder oder Camcorder) vorgenommen, sodass der Rechner lediglich den Datenstrom von rund 25 Mbit/s (ohne Ton) auf Festplatte schreiben muss.
- Zudem sind kompakte Systeme („Festplatten-Videorecorder“) auf dem Markt, die das DV-Signal PC-unabhängig auf eine integrierte Festplatte ablegen. Außerdem wurden Systeme mit Speicherung auf Festspeichern angekündigt, die im Vergleich zu Festplatten ohne bewegliche Teile auskommen. Derzeit ste-

hen jedoch für dieses bei der digitalen Fotografie etablierte Verfahren noch keine Speicherkarten mit der benötigten Kapazität zur Verfügung.

- Analoge Videorecorder im Consumer-Bereich werden zunehmend durch DVD-Recorder ersetzt, die neben der digitalen Speicherung im MPEG-2-Format mit unterschiedlichen Qualitätsstufen bzw. Bitraten auch kürzere Zugriffszeiten ermöglichen. Derzeit existieren verschiedene Formate, die unterschiedlich gut für eine Weiterverarbeitung geeignet sind.

Diese Übersicht zeigt, dass die Übergänge zwischen klassischer AV- und PC-Technik zunehmend verschwinden. Dennoch werden vielfach – insbesondere aufgrund der geringen Kosten für die Speicherung – weiterhin kassettenbasierte Systeme eingesetzt.

Live-Encoding

Encoding beschreibt die Herstellung einer AV-Datei und/oder eines AV-Datenstroms zur Weiterverteilung auf einem Streaming-Server. Anstelle einer bereits vorliegenden Quelldatei kann dies auch in Echtzeit (*live*) erfolgen, d. h. ein Audio- und Videosignal wird per Hardware digitalisiert und in das gewünschte Format gewandelt (Abbildung 2-4).

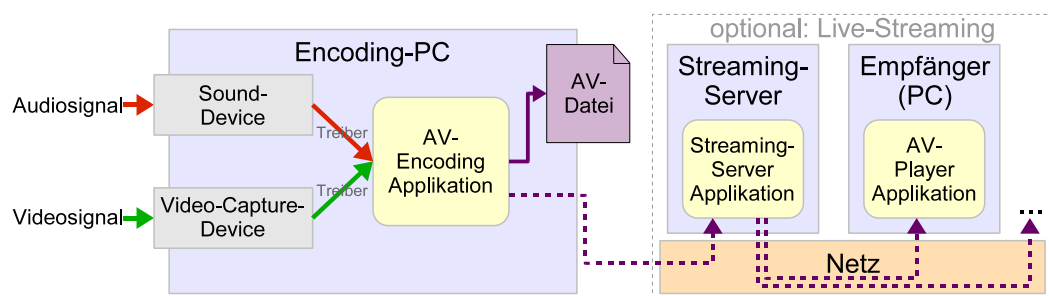


Abbildung 2-4: Datenflussdiagramm zum Live-Encoding

Wenn bereits bei der Produktion feststeht, dass keine hochqualitative Aufzeichnung, sondern lediglich ein Publikationsformat benötigt wird, bietet sich dieses Verfahren an, da keine weitere Konvertierung bei der Postproduktion nötig wird. Optional kann ein paralleler Live-Stream erzeugt und übertragen werden.

Neben speziellen Hardware-Lösungen wird für das Live-Encoding üblicherweise ein PC eingesetzt, der mit AV-Hardware zur Digitalisierung (Videograbber- und Soundkarte) sowie einem Software-Encoder für das gewünschte Format ausgestattet ist. Auch bereits digitalisierte Signale im DV- oder MPEG-Format können über digitale Schnittstellen wie IEEE-1394 oder USB eingelesen werden. Der Rechenaufwand erhöht sich jedoch, da hier der Medienstrom erst dekodiert werden muss, um dann das Zielformat zu erzeugen (*Transkodierung*).

Verbreitete Encoding-Software sind u. a. der RealProducer und der Windows Media Encoder, beide für das jeweils gleichnamige AV-Format.

Screen-Recording

Screen-Recording oder *Screen-Capturing* bezeichnet die Aufzeichnung des PC-Bildschirminhaltes in Form eines Videoclips mit Hilfe einer auf diesem Rechner installierten Software.

Diese Methode wird – zusammen mit paralleler oder nachträglicher Vertonung – häufig zur Software-Schulung eingesetzt. Die Unterschiede des so erzeugten Videos zu konventionellen Videoclips bestehen dabei typischerweise im verwendeten Codec, der geringeren Bildwiederholrate und der höheren Auflösung bis hin zur vollen Bildschirmgröße.

Typische Produkte sind TechSmith Camtasia und die in Microsoft Windows Media Encoder enthaltene Funktion zur Bildschirmaufzeichnung. Beide Programme verfügen über einen für die Anwendung optimierten Codec. Camtasia nutzt dafür den verlustfreien tsc (TechSmith Screen Capture Codec [73]) und erzeugt AVI-Dateien, der WM Encoder generiert WMV-Daten mit dem Windows Media Video 9 Screen Codec [36].

Infrastruktur: Multimediaräume

Die in den vorangehenden Abschnitten beschriebenen Systeme werden in vielen Fällen in dedizierten Umgebungen eingesetzt. Speziell für die Aufzeichnung stellt sich die Aufgabe der Integration der o. g. Systeme unter Berücksichtigung von Aspekten wie einfacher Bedienung, Abstimmung der Geräte aufeinander und auf den Raumeigenschaften wie Licht und Akustik [67].

Räume mit Multimediaausstattung (z. B. AV-Aufzeichnungstechnik für Live-Produktionen, 3D-Projektion) werden in den meisten Fällen nicht nur für eine einzelne Anwendung genutzt. Für den universellen Einsatz bestehen zum Teil gegensätzliche Anforderungen. So wird etwa für die Präsentation z. B. von 3D-Szenen oft ein abgedunkelter Raum bevorzugt, während für den Einsatz von Kameras für Aufzeichnungen oder Videokonferenzen eine möglichst gute Ausleuchtung vorhanden sein sollte. Auf der anderen Seite können Systeme für die Video-Aufzeichnung auch für Tele-Konferenzen genutzt werden; Mikrofone können gleichzeitig für eine eventuell notwendige lokale Beschallung eingesetzt werden.

Das Stichwort *Multimediaraum* deckt eine weite Spanne von Installationen ab. Angefangen vom Videokonferenzraum in Bürogröße bis hin zur Ausstattung eines großen Hörsaals entsteht mit den unterschiedlichen Anforderungen eine Variationsbreite. Orientiert an der Größe lassen sich drei Klassen von Räumen definieren:

- Besprechungsräume mit Platz für bis zu 10 Personen
- Seminarräume für bis zu 50 Personen
- Hörsäle für mehrere hundert Personen

Vereinfacht lassen sich den Klassen primäre Anwendungsbereiche zuordnen. Kleinere Räume eignen sich gut für interaktive verteilte Anwendungen und Videokonferenzen. Große Hörsäle mit einer hohen Anzahl lokaler Teilnehmer eignen sich weniger gut für Telekonferenzen, da eine hochqualitative audiovisuelle Erfassung einzelner Personen Verzögerungszeiten durch den häufigen Wechsel von Kamera-

und Mikrofonpositionen hervorruft und damit einen spontanen Ablauf stört. Größere Räume sind daher primär für Aufzeichnungen bzw. Live-Übertragungen von Veranstaltungen mit geringem interaktiven Anteil wie Vorträge oder Vorlesungen nutzbar [11].

In mittleren und großen Räumen mit umfangreicher Ausstattung sind die Geräte zur Steuerung, Mischung, Aufzeichnung, Signalverteilung usw. in einem separaten Regieraum untergebracht. Damit kann die eigentliche Veranstaltung ungestört von der – meist durch ein oder zwei Personen betreuten – AV-Produktion durchgeführt werden. Beispiele für diese Art von Installationen sind das Audimax der Universität Hannover (Abbildung 2-5, [76]) sowie der Multimedia-Seminarraum „learning space“ im Forschungszentrum L3S (Abbildung 2-6, [31]).



Regieraum: Mikrofonempfänger, Audio- und Videomischung, Vorschaumonitor, Bedienpanel der Kamerasteuerung



Hörsaal: eine von drei fernsteuerbaren PTZ-Kameras (hinten rechts: Regieraum)

Abbildung 2-5: Medientechnik im Audimax der Universität Hannover



Rückprojektionsflächen: zwei SMARTboards, eine 3D-Stereofähige Projektion (mitte), vier fernsteuerbare Kameras, Beschallung



Regieraum: Bild- und Tonmischung, Kamera- und Medienraum-Steuerung, Signalverteilung; im Vordergrund rechts: Encoding-PCs

Abbildung 2-6: Multimedia-Seminarraum im Forschungszentrum L3S

Bei komplexer Multimedia-Ausstattung für verschiedene Anwendungen ist zur einfachen Handhabung der Einsatz einer zentralen Medienraumsteuerung („Mediensteuerung“) sinnvoll. Hiermit werden komplexe Bedienschritte zusammengefasst und abstrahiert, sodass sowohl Laien als auch fortgeschrittene Anwender für ihre Nutzungsszenarios eine effiziente Handhabung vorfinden, z. B. mittels LCD-Touchpanel.

Darüber hinaus gibt es Raum-Konzepte für eine weit gehend betreuungslose Nutzung, wie etwa im Interactive Learning Lab (iL2) des Forschungszentrums L3S realisiert. Standardisierte Anwendungen können hier mit Hilfe vordefinierter Sze-

narios vom Nutzer selbst durchgeführt werden. Abrufbar sind die einzelnen Setups mit Hilfe der mittels PC bedienten Medienraumsteuerung, an die sämtliche Geräte angebunden sind [21].

2.2.3 Bearbeitung

Der Bereich Medienproduktion umfasst eine breite Palette unterschiedlicher Anwendungen, Formate und Systeme. Neben der klassischen Video-Nachbearbeitung (Editing) kommen zunehmend Multimediasysteme zum Einsatz, die verschiedene Formate (Bild, Text, Video, 3D/VR) integrieren. Quellmaterial hierfür stellen sowohl Rechner-generierte Bildsequenzen als auch Aufnahmen realer Objekte (Experimente, Vorträge, Diskussionen) dar. In der Regel erfordern die eingesetzten Editing- und Authoring-Systeme eine interaktive Bearbeitung durch den Benutzer und sind daher nicht für Batch- oder Netzdienste geeignet [9].

Neben der Integration von Videosequenzen in 3D/VR-Medien werden umgekehrt auch Bildsequenzen aus der 3D-Visualisierung in Videoclips gewandelt, da die hierfür notwendigen Abspielsysteme weiter verbreitet sind. Auch sollen AV-Daten oft für verschiedene Publikationswege (z. B. DVD und Internet) unterschiedlich aufbereitet werden, sodass Systeme zur Konvertierung eingesetzt werden müssen.

Die folgende Aufzählung bietet eine Übersicht gängiger Bearbeitungssysteme.

- Für das *Video-Editing* (Bearbeitung, Schnitt) werden heute vornehmlich Rechner-basierte Software-Lösungen wie z. B. Adobe Premiere eingesetzt. Das Ablegen des Videomaterials auf Festplatten ermöglicht im Vergleich zu Bandsystemen den schnellen nicht-linearen Zugriff (NLE-Systeme). Zudem gestattet der PC als Basis den einfache Im- und Export von Bilddaten von und zu anderen Applikationen (z. B. Bildbearbeitung), Encodern und Autorensystemen.
- *Multimedia-Autorensysteme* gehen über die eigentliche Videobearbeitung hinaus, da hier unterschiedliche Medientypen integriert werden. Typische Vertreter dieser Art sind Macromedia Director oder Flash, die Bild, Text, Grafik, Video sowie die Programmierung des Präsentationsablaufs in einem Format zusammenfassen.
- *Formatkonvertierung* kann zum einen ein notwendiger Zwischenschritt zur Adaption an ein benötigtes Format sein oder dient der Erstellung von alternativen Versionen desselben Materials, etwa für unterschiedliche Repräsentationsmedien – siehe Abschnitt 2.1.3.

2.2.4 Distribution

Der letzte Teil der Produktionskette (Abbildung 2-1) besteht in der Distribution (Verteilung) und damit Publikation der erstellten (Re-)Präsentation, indem das Material dem Betrachter zur Verfügung gestellt wird. Die Betrachtung beschränkt sich dabei auf die populäre Verwendung des PCs als Wiedergabe-Plattform, denkbar ist auch die Nutzung von Videokassetten oder DVD-Video als Transportmedium.

Unabhängig von den bereits im Abschnitt 2.1.4 erörterten Transport-Aspekten nimmt die HTML-basierte („Web-basierte“) Distribution für MM-Daten eine herausragende Position ein, die in der weit gehenden Plattformunabhängigkeit, großen Verbreitung und umfangreichen Verknüpfungs- bzw. Integrationsmöglichkeiten für eine Vielzahl von Datenformaten begründet ist. Zudem sind so gestaltete Präsentationen ohne oder mit geringfügiger Modifikation sowohl lokal als auch im Netz einsetzbar. Im Einzelnen ergeben sich folgende Vorteile:

- Native Darstellung einer Vielzahl von Medien.
- Erweiterung für eine hohe Anzahl an Medientypen über Plug-Ins möglich.
- Kopplung von parallelen Medien mittels DHTML/JavaScript.
- Betriebssystem-übergreifende Verfügbarkeit.
- Funktion mit lokalen Daten *und* Daten im Netz.

Bei der Publikation ist ferner die Integration des produzierten Materials in eventuell bereits bestehende Systeme zu berücksichtigen. Exemplarisch ergeben sich folgende Möglichkeiten.

- Definition eines eigenständigen Erscheinungsbildes der MM-Präsentation ohne Berücksichtigung von bestehenden Layouts, Formaten oder Infrastruktur.
- Gestalterische Einbindung in vorhandene Plattformen; z. B. Anpassung von HTML-Seiten an bestehende Layouts.
- Erstellung von Schnittstellen zu Datenbank-basierenden E-Learning- oder Content-Management-Systemen (CMS) zur effizienten Einbindung der MM-Daten.

Neben dem Einsatz von HTML bietet sich für Multimedia-Präsentationen mit parallelen und sequenziellen Medienströmen SMIL als Plattform an. Ähnlich wie in HTML sind dabei unterschiedliche Medientypen integrierbar. AV-Medientypen und Synchronisation werden dabei nativ von der Wiedergabe-Software unterstützt, sodass hier keine zusätzlichen Funktionen z. B. mittels JavaScript implementiert werden müssen. Aufgrund der geringeren Verbreitung ist bei rein auf SMIL basierende Publikationen jedoch mit einer geringeren Reichweite zu rechnen.

2.3 Systemintegration

Aufgrund des komplexen Zusammenspiels verschiedener Module bei der Multimedia-Produktion und -Distribution ist die Integration von existierenden Systemen von besonderer Bedeutung.

Neben der notwendigen Einbindung von AV-Hardware in PC-basierte Produktionssysteme ist auch die Integration von Software bzw. Software-Komponenten ein Schlüssel zur effizienten Implementierung eines neuen Systems.

2.3.1 Einbindung von AV-Hardware

Mediendaten

Essenzielle Voraussetzung für die Produktion ist die Übertragung von Audio- und Videosignalen in einen Rechner. Hierzu steht entsprechende Hardware zur Verfü-

gung, Treiber stellen die Schnittstellen zu den eingesetzten Software-Encodern bereit (siehe Abschnitt 2.2.2, „Live-Encoding“).

Für die Aufzeichnung von Computer-Videosignalen ist zwischen VGA-Signal und Video-Capturing-Karte ein zusätzlicher Scan-Converter notwendig (siehe Abschnitt 2.2.2).

Steuerung

Zur Anbindung von Videorekordern an PC-basierte Schnittsysteme werden z. T. Verbindungen über serielle Schnittstellen (RS-232, RS-422) eingesetzt. Bei DV-Videorekordern werden über die IEEE-1394-Verbindung neben den AV-Daten auch die Steuerungsinformationen übertragen.

Weiterhin ist die Steuerung diverser AV-Geräte wie PTZ-Kameras, Audio- und Videomischer, Projektoren usw. möglich. Neben RS-232 kommen hier MIDI- und USB-Schnittstellen zum Einsatz, die auch von Mediensteuerungssystemen (siehe Abschnitt 2.2.2, „Infrastruktur: Multimediaräume“) genutzt werden.

Mit der Steuerung über standardisierte Schnittstellen ist in Teilbereichen eine Automatisierung von Abläufen sowie die Fernsteuerung möglich, von größerer Bedeutung ist oft aber die einfachere und effizientere Bedienung per PC oder Mediensteuerung.

2.3.2 Integration von Software

Die Einbindung von Software spielt an verschiedenen Stellen des Produktionsflusses eine wichtige Rolle. Hierzu gehören die Steuerung von AV-Encodern, die Einbindung von Präsentations-Tools sowie die Integration unterschiedlicher Software zur synchronisierten Wiedergabe der Repräsentation, speziell im Webbrowser.

Für die Einbindung von Software existieren folgende grundlegende Ansätze mit unterschiedlichen Vor- und Nachteilen.

- *Einbindung des Quellcodes oder von Bibliotheken.* Hier erhält der Nutzer der Quellen die vollständige Kontrolle über das einzubindende Programm und kann bei Bedarf auch Änderungen am Code vornehmen. Alternativ stehen lediglich die Quellen zur Nutzung von Funktionen in Bibliotheken zur Verfügung, die dynamisch in das erstellte Programm eingebunden werden.
- *Nutzung von Komponenten-Software* (z. B. ActiveX Controls, Plugins, Java Beans). Aufgrund zunehmender Komplexität von Software enthalten moderne Programmier-Frameworks Funktionen zur Einbindung von Objekten mittels definierter Schnittstellen, die in verschiedenen Programmiersprachen nutzbar sind. Der Programmierer kennt lediglich diese Schnittstellendefinitionen und kann so ohne Wissen über die Implementierung die Funktionen des Objekts in seinen Programmen nutzen. Unter Einhaltung der Schnittstellen können auf diese Weise auch Komponenten unabhängig voneinander aktualisiert werden.
- *Shell-Aufruf von Programmen.* Hierbei werden lediglich die Mechanismen genutzt, die auch dem Benutzer interaktiv auf der Kommandozeile zur Verfügung stehen. Ein-, Ausgabe- und Fehlerkanal können an eigene Programme

gebunden werden. Die Kommunikation zwischen eigenem und bestehendem Programm ist beschränkt auf die Nachahmung von Tastatureingaben und die Auswertung der Ausgabekanäle.

- *Event-Capturing via Betriebssystem oder Windows-Server:* Applikationen in fensterbasierten Systemen wie MS Windows oder XWindows tauschen Nachrichten über Queues des Message-Handling-System im Windows-Server aus. Über diesen Mechanismus lassen sich z. B. Tastatur- und Mauseingaben simulieren, aber auch laufende Programme beobachten, indem die von ihnen gesendeten Ereignisse (Events) „mitgehört“ werden.

Abhängig von der vorliegenden Applikation und des eingesetzten Betriebssystems stehen fast immer nur einige Methoden zur Verfügung, z. B. liegen nur wenige Programme im Quellcode vor.

Für diese Arbeit ist insbesondere die Anbindung an Komponenten von Bedeutung, da zur Aufzeichnung von Rechner-gestützten Präsentationen die Nutzerinteraktion der eingesetzten Programme erfasst werden soll. Programme wie der Internet Explorer oder sämtliche MS-Office-Programme lassen sich auch als Komponenten einbinden, sodass hier auf einfache Weise die beschriebene Erfassung möglich ist. Der andere Schwerpunkt liegt auf der Einbettung von Applikationen zur Anzeige von Inhalten der Repräsentation, vornehmlich im Webbrowser.

AV-Player: Webbrowser-Plugins

Zur Anzeige von AV-Daten existieren eine Reihe von Programmen, die sich häufig auch als Plugin im Webbrowser nutzen lassen. Auf diese Weise können AV-Daten in HTML-Dokumente eingebunden werden. Zusammen mit JavaScript (JS) ist darüber hinaus eine effiziente dynamische Kopplung von AV-Daten und HTML-Dokumenten möglich.

Die verbreitetsten Player dieser Art sind RealPlayer, Apple QuickTime (QT) und Microsoft Windows Media (WM) [46]. Tabelle 2-3 liefert eine Funktionsübersicht dieser Player zusammen mit den häufig eingesetzten Webbrowsern Microsoft Internet Explorer und der Mozilla-Produktfamilie (Mozilla, Netscape und Firefox) auf dem Betriebssystem MS Windows. Die Ergebnisse beruhen auf entsprechenden Untersuchungen der genannten Software-Komponenten [43].

AV-Plugin	Real	QuickTime	Windows Media
Internet Explorer-Unterstützung	ja	ja	ja
Mozilla-/Netscape-Unterstützung	ja	ja	ohne JS
unterstützte Betriebssysteme	MS Windows, Mac, Linux u. a.	MS Windows, Mac	MS Windows, Mac, Solaris
SMIL-Unterstützung	ja	ja	nein

Tabelle 2-3: Funktionen von AV-Plugins

Für die Einbindung von QuickTime-Medien ist der HTML-Code jeweils individuell für den Browsertyp anzupassen, was jedoch dynamisch mit JS realisiert werden

kann. Um die Unterstützung von anderen Betriebssystemen als MS-Windows ist es (ausgenommen QuickTime für MacOS) oft schlechter bestellt; so hinken die Versionsnummern der Real- und WM-Player den Windows-Pendants hinterher – Codecs der jeweils neuesten Generation stehen nicht zur Verfügung.

2.4 Zeit und Synchronisation

In einem verteilten multimedialen System ist für Aufzeichnung und Wiedergabe die Zeit als Synchronisationsbasis und als Metadatum von zentraler Bedeutung. Dieser Abschnitt beschreibt die relevanten Standards und Mechanismen.

2.4.1 Zeitformat

Das W3C beschreibt in [81] eine Untermenge von den in ISO 8601 definierten Zeitformaten. In der höchsten Auflösung hat die Zeitangabe die Form *YYYY-MM-DDThh:mm:ss.sTZD* mit den folgenden Elementen

`YYYY-MM-DDThh:mm:ss.sTZD`

YYYY = Jahr, vierstellig
 MM = Monat, zweistellig (01=Januar, etc.)
 DD = Tag des Monats, zweistellig (01 bis 31)
 hh = Stunde, zweistellig (00 bis 23) (am/pm ist nicht erlaubt)
 mm = Minuten, zweistellig (00 bis 59)
 ss = Sekunden, zweistellig (00 bis 59)
 s = Sekundenbruchteile, Dezimal, ein oder mehrstellig
 TZD = time zone designator (Z oder +hh:mm oder -hh:mm)

Die internationale Zeitzone wird mit dem TZD (time zone designator) angegeben und trägt den Wert Z (Zulu, entsprechend UTC, Universal Coordinated Time) oder die Angabe der Zeitdifferenz zur UTC in der Form +hh:mm bzw. -hh:mm. So entspricht z. B. 2004-11-22T20:15+01:00 (mitteleuropäischer Winterzeit) der Angabe 2004-11-22T19:15Z.

Die Anordnung der Werte ermöglicht im Vergleich zu der im deutschsprachigen Raum gebräuchlichen Form DD.MM.YYYY eine einfache alphanumerische Sortierung. Im Gegensatz zu den anderen Elementen ist die Anzahl der Nachkommastellen der Sekundenangabe jedoch nicht definiert.

Die Darstellungsform wurde als Datentyp *dateTime* in das XML-Schema Teil 2 aufgenommen [85].

2.4.2 Synchronisation von Uhren

Für zeitabhängige Anwendungen in einem verteilten System sind synchronisierte Uhren von essenzieller Bedeutung. Die Basis für ein seit Jahren etabliertes und standardisiertes Verfahren zur Synchronisation von Uhren über das Netz ist das Network Time Protocol (NTP [56]). Das Protokoll nutzt UDP/IP und ist damit auf jedem Internet-fähigen Rechner einsetzbar. Implementierungen sind für alle gängigen Betriebssysteme verfügbar, z. B. auch für MS Windows [45]. Bei Verwendung im Internet sind Abweichungen zwischen 10 und 100 ms zu erwarten, im lokalen Netz bis zu 1 ms [47][72].

Die Empfänger (NTP-Clients) beziehen die exakte Uhrzeit von einem NTP-Server, der sich wiederum über eine hierarchische NTP-Server-Struktur synchronisiert. Als Quelle für NTP-Server dienen Empfänger, die das Zeitsignal via Satellit (GPS) oder Funk (z. B. DCF77 in Deutschland) empfangen. Alternativ lassen sich natürlich auch beliebige Rechner ohne Netzanbindung auf diese Weise direkt mit der exakten Uhrzeit versorgen.

2.4.3 Synchronisation von Mediendaten

Ein Kernpunkt von Multimedia-Systemen besteht in der Zusammenführung verschiedener unabhängiger Medien in einen zeitlichen Zusammenhang. Zur *Synchronisation* von Medien sind zwei grundsätzliche Verfahren zu unterscheiden [69].

- *Intra-Stream-Synchronisation* bezeichnet die Einhaltung der zeitlichen Beziehung innerhalb eines Medienstroms, z. B. das Zeitraster von 25 Bildern pro Sekunde in einem Video.
- *Inter-Stream-Synchronisation* beschreibt die zeitliche Zuordnung verschiedener Medienströme zueinander, etwa die Präsentation eines Videos mit paralleler Darstellung von einzelnen Bildern.

Für die vorliegende Arbeit ist primär die Inter-Stream Synchronisation zwischen verschiedenen aufgezeichneten Medien (*Streams*) von Bedeutung sowie speziell die Toleranzbereiche möglicher Abweichungen. Untersuchungen ergaben, dass für den Lippensynchronismus (den Versatz zwischen Ton und Bild) ± 80 ms akzeptiert werden, wobei nachlaufender Ton unproblematischer wahrgenommen werden kann als ein nachlaufendes Bild [69].

Das typische Szenario einer Vortragsaufzeichnung (AV mit parallelen Vortagsfolien) ist vergleichbar mit der Synchronisation zwischen Audio und lose gekoppelten Bildern in einer Slideshow-Anwendung. In einem Experiment zur Ermittlung von Akzeptanzgrenzen wurden Toleranzgrenzen von ± 500 ms ermittelt [70].

2.5 Metadatenmodelle

Für die Produktion in einem modularen System werden für die Aufbereitung der Mediendaten zusätzliche Informationen – Metadaten – benötigt. Im Folgenden wird unterschieden zwischen Metadaten im konventionellen Kontext (inhaltsbezogenen Metadaten) und denen zur Erfassung des zeitlichen Ablaufs (Synchronisationsmetadaten). Beide Arten werden für die Produktion von Multimedia-Präsentationen eingesetzt.

2.5.1 Metadaten zur Inhaltsbeschreibung

RDF

Das *Resource Description Framework* dient der Beschreibung und dem Austausch von Metadaten. Eine Ressource ist dabei ein beliebiges, per URI beschreibbares Objekt (z. B. eine Webseite). Ein RDF-Statement besteht aus der Kombination einer Ressource, einer Eigenschaft (Property) dieser Ressource und des Wertes (Value) der Eigenschaft.

Im Vergleich zu XML (das als RDF-Austauschformat genutzt wird) hat RDF den Vorteil der besseren Skalierbarkeit, da RDF-Statements zum einen unabhängig von der Element-Reihenfolge sind und zum anderen keine schwer handhabbaren verschachtelten Datenstrukturen aus Zeichenfolgen, Elementen, Listen usw. zulassen.

Dublin Core

Die Dublin Core Metadata Initiative (DCMI, [7]) stellt eine umfassende Klassifikation für inhaltsbezogene Metadaten bereit, die nicht auf bestimmte Medientypen beschränkt ist, aber auch nicht speziell auf audiovisuelle Medien ausgelegt ist. [60]

Die Video Development Initiative (ViDe) hat sich mit dem Einsatz von DC Metadata im AV-Umfeld befasst und ein Anwendungsprofil für Digitales Video [77] veröffentlicht. Darin sind die DC-Klassen an einigen Stellen erweitert worden. So wird beispielsweise die Erweiterung der Werteliste für das DC-Element *Type* um „video“ und „animation“ vorgeschlagen; DC verwendet bisher den allgemeinen Wert „image“ für alle Arten von Video. Für das Element *Format_Medium*, das den MIME-Type enthält, sind Erweiterungen zur genaueren Spezifikation des Subtypes „mpeg“ vorgeschlagen.

In Tabelle 2-4 werden einige, hier relevante Elemente erläutert. Prinzipiell können zu generierten Präsentationen beliebige inhaltsbezogene Metadaten in Form von allgemein nutzbaren DC-Elementen ergänzt werden.

Bei der Nutzung in multimedialen Systemen ist zu beachten, dass DCMI primär für *einzelne* Ressourcen wie Texte, Bilder oder – speziell die ViDe Erweiterung – Videos ausgelegt ist. Neben dieser klassischen Anwendung von Metadaten (hier: für die gesamte Präsentation) werden dabei auch Metadaten für Ressourcen-Fragmente benötigt.

Bereich	Eigenschaft	Erläuterung
Content	Description	Beschreibung; diverse Subtypen wie Table of Contents, Abstract. ViDe ergänzt eine definierte „genrelist“, die u. a. „classroom lecture“ enthält
	Subject	Stichworte zum Thema bzw. Inhalt der Ressource
	Title	Titel. Zur Kennzeichnung von gesplitteten Videodateien schlägt ViDe vor, den Wert durch „Sequence“ bzw. „Excerpt“ zu ergänzen
Intellectual Property	Creator	Autor
	Rights	Copyright-Informationen in Textform
Instantiation	Date	Datum innerhalb des Lebenszyklusses einer Ressource. DC verwendet das oben erläuterte Format für Datums- bzw. Zeitangaben.
	Format	Internet-Medientyp nach [34] entsprechend MIME-Type [58]

Tabelle 2-4: Dublin Core Elemente (Auswahl)

Dublin Core lieferte einige Anforderungen für das Design von RDF, während RDF wiederum für den Einsatz von DC ein formales zugrunde liegendes Datenmodell bildet. So können DC-Elemente mit Hilfe von RDF formuliert werden (aus [8]):

```
<?xml version="1.0" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description about="http://purl.org/DC/documents/notes-cox-816.htm">
    <dc:title>Recording qualified Dublin Core metadata in HTML</dc:title>
    <dc:description> We describe a notation for recording
    qualified Dublin Core metadata in HTML meta elements. The syntax
    includes recommended usage of the standard HTML syntax to record
    the different classes of qualification needed to represent the
    model.</dc:description>
    <dc:date>1999-08-18</dc:date>
    <dc:format>text/html</dc:format>
    <dc:language>en</dc:language>
    <dc:publisher>Dublin Core Metadata Initiative</dc:publisher>
  </rdf:Description>
</rdf:RDF>
```

LOM

Learning Objects Metadata (LOM) spezifiziert Syntax und Semantik zur Beschreibung von Lernmodulen (*learning objects*) und ermöglicht damit Lernenden und Lehrenden u. a. die Suche, Evaluation und Nutzung dieser Ressourcen [20].

LOM verfolgt damit – spezialisiert auf Lerninhalte – ähnliche Ziele in Richtung Semantic Web wie etwa RDF. Auch LOM lässt sich sowohl in XML als auch in RDF beschreiben. Die LOM-Arbeitsgruppe ist Teil des IEEE Learning Technology Standards Committee (LTSC).

SCORM

SCORM (Sharable Content Object Resource Model) ist als Teil der Advanced Distributed Learning Initiative (ADL) ein Referenzmodell, das verschiedene Standards (u. a. LOM, IMS [16]) zur Beschreibung von Lernobjekten integriert [66].

Die primären Ziele sind dabei die Wiederverwendbarkeit, Interoperabilität, Nachhaltigkeit und Auffindbarkeit von Lernmaterial. Inhalte können so zwischen SCORM-konformen Lernmanagementsystemen ausgetauscht werden. Die Aktivitäten von SCORM gehen dabei über die inhaltliche Beschreibung der Lernobjekte hinaus; im Mittelpunkt stehen die Zusammenstellung von Material zu kompletten Kursen und die Schnittstellen zwischen Lernobjekten und LMS.

2.5.2 Metadaten zur Beschreibung von Medien

Für die Beschreibung und Synchronisierung von Multimediadaten sind neben verschiedenen herstellereigenen Formaten zwei offene Standards verfügbar, die nachfolgend vorgestellt werden.

SMIL

SMIL [84] definiert eine Sprache zur synchronisierten Präsentation von Medienströmen und ist damit ein primäres Zielformat für erzeugte Präsentationen.

Unter anderem ist im AV-System von Real Networks [53] eine umfangreiche SMIL-Implementierung zu finden, aber auch Apples QuickTime Player verfügt über eine SMIL-Unterstützung [3]. Im Folgenden sind die relevanten Elemente (Tags) und Attribute zusammengestellt.

- Das SMIL meta-Tag: Mit dem *meta*-Tag werden in der Form `<meta name="Wert" content="Inhalt" />` verschiedene Informationen zur SMIL-Präsentation angegeben. Die Funktion ist an das gleichnamige HTML-Tag angelehnt. Für das Feld *name* sind die Werte *abstract*, *author*, *base*, *copyright* und *title* definiert. Der Inhalt ist jeweils Text oder eine URL.
- SMIL-ID: Jeder SMIL-Tag kann mit einer *ID* versehen werden. Diese dienen z. B. der eindeutigen Zuordnung parallel genutzter Medienströme in einem SMIL-Browser. Für die Werte gelten zwei Regeln: Groß- und Kleinschreibung wird unterschieden und das erste Zeichen muss ein Buchstabe, Doppelpunkt oder Unterstrich sein.
- Zeitangaben in SMIL: SMIL kennt für die Kennzeichnung von zeitlichen Angaben wie z. B. Startzeit oder Länge zwei Arten. Neben den „Shorthand“ Varianten in der Form „1.5h“, „2.34min“, „5.6s“ oder „300.1ms“ ist ein Zeitformat in der Form „*hh:mm:ss.xy*“ mit Stunden, Minuten, Sekunden und Hundertstel Sekunden definiert. Daneben existiert eine Methode zur Synchronisation von *Abspielvorgängen* mit der realen Zeit (*wallclock-sync*), die Zeitangaben gemäß [81] nutzt, für die Erfassung realer Zeitmarken aber irrelevant ist.

MPEG-7

Während MPEG-4 [38] Mechanismen für die Komposition und Synchronisation von Medienströmen beschreibt, enthalten die in MPEG-7 [39] definierten Beschreibungs-Schemata (description schemes, DS) Metadaten-Strukturen für die inhaltliche Beschreibung und Annotation von MM-Inhalten.

Die Mehrzahl der darin enthaltenen Elemente ermöglicht die Erfassung von „low-level“-Eigenschaften von AV-Inhalten wie z. B. Form, Farbe oder Bewegung. Diese Attribute können mit entsprechenden Werkzeugen automatisch erkannt und extrahiert werden. Einige Schemata dienen dem Content Management und enthalten Elemente zur Beschreibung der Produktion (creation/production), Daten zu verwendeten Medien (media information) sowie Nutzungsrechte (rights).

MPEG-21

Die MPEG-Standards 1, 2, 4 und 7 beschäftigen sich mit der Repräsentation von Multimedia-Daten. MPEG-21 beschreibt ein offenes Framework für die Verteilung und Nutzung dieser Multimedia-Daten. Folglich liegt der Schwerpunkt auf der generischen Beschreibung von Inhalten.

Diese Inhalte (Digital Items) sind Zusammenstellungen von Dokumenten; strukturierte, digitale Objekte, die mit Hilfe des MPEG-21-Frameworks durch eine Stan-

standard-Repräsentation, Identifikation sowie Metadaten gekennzeichnet sind. So kann etwa eine Audio-CD als „MPEG-21 Music Album“ repräsentiert werden, das die Mediendaten (Audiodateien, Cover-Bilder, Texte) sowie diverse diesen Medien assoziierte Metadaten (Identifizierer, Beschreibung) enthält. Das Album kann so als eine Einheit behandelt werden [40].

Ferner bringt MPEG-21 die Digital Items in Zusammenhang mit deren Benutzern (Individuen, Kunden, Organisationen, Firmen usw.), die als Erzeuger, Verbraucher, Rechteinhaber, Anbieter usw. auftreten können [5].

Neben der Integration verschiedener Medientypen und deren Beschreibung ist dabei auch die Sicherstellung deren Interoperabilität sowie die Rechteverwaltung umfasst.

Annodex

Das Annodex-Projekt versucht, die Paradigmen und die Einfachheit von Web-Technologie (http, html, URIs) auf zeitkontinuierliche Medien zu übertragen. Hierzu werden offene Standards und Werkzeuge entwickelt, die eine Einbindung dieser Medien in die Infrastruktur des WWWs einschließlich der Suche, Proxies und Browser ermöglichen [2].

Hierzu gehört die Definition einer Beschreibungssprache für die strukturierte Erfassung von Hyperlinks und Annotationen (Continuous Media Markup Language, CMML), ein Containerformat, das diese Beschreibung und die Mediendaten integriert (ANNODEX) sowie eine Erweiterung der vorhandenen URI Spezifikation für die zeitliche URI-Adressierung (deep hyperlinking). CMML basiert auf XML, als Containerformat für ANNODEX wird Ogg (RFC 3533) eingesetzt.

3 Problemanalyse

Ausgehend von verschiedenen Anwendungsszenarios werden in diesem Kapitel die einzelnen Aspekte von Aufzeichnungssystemen analysiert. Die Gliederung der Analyse orientiert sich dabei an den Bereichen

- Anwendung/Produktion,
- Präsentation (als Ausgangsbasis),
- Repräsentation (als Endprodukt) sowie
- Systemaspekte.

Zusammen mit der Betrachtung existierender Systeme und von Anwendungsbeispielen bildet diese Analyse die Grundlage für ein Klassifikations- und Bewertungsschema. Daraus ergeben sich abschließend die Ziele für die Entwicklung eines neuartigen Ansatzes.

3.1 Ausgangspunkt: Anwendung und Kontext

Den Ausgangspunkt für die Aufzeichnung bildet eine reale Präsentation (siehe Abschnitt 2.2.1), die mittels einer geeigneten Produktionskette (Abschnitt 2.1.2) unter Zuhilfenahme von Aufzeichnungssystemen (Abschnitt 2.2.2) in eine Repräsentation in elektronisch nutzbarer Form abgebildet werden soll.

3.1.1 Anwendungsszenarios

Es lassen sich zwei grundlegende Anwendungsszenarios unterscheiden.

- Der primäre Anwendungsbereich für die Präsentationsaufzeichnung ist die Erfassung von Vorträgen im Rahmen von *Vorlesungen* oder Kursen für den Einsatz als Lehrmittel in Rechner-basierten E-Learning Systemen. Die Aufzeichnung bildet damit den Vortrag der klassischen Präsenzlehre in elektronisch verteilter Form ab. Sie kann mit anderen elektronisch nutzbaren Lehrmitteln verknüpft werden und dient so z. B. der individuellen Offline-Nutzung [30][49].
- Dazu kommt die Aufzeichnung von Vorträgen im Rahmen von *Konferenzen*, die ebenfalls für Lehrzwecke eingesetzt werden können, aber andere Anforderungen und Randbedingungen an die Produktion stellen; der Vortragende kann hier in den meisten Fällen nicht aktiv in die Produktion der Aufzeichnung eingebunden werden.

Zusätzlich kann parallel zur Aufzeichnung auch eine Live-Übertragung (unidirektional oder in Form einer Video-Konferenz) durchgeführt werden. Der Einsatz der audio-visuellen Aufnahmetechnik ist dabei für beide Anwendungen weitgehend identisch. Beide Formen – Aufzeichnung und Übertragung – weisen unterschiedli-

che Vor- und Nachteile für den jeweiligen Anwendungsfall auf, wobei die Aufzeichnung prinzipbedingt eine Mehrfachnutzung ermöglicht [12].

Da vielfach eine klassische Präsenzveranstaltung die Basis für die Aufzeichnung bildet, findet die Aufnahme in einem Hörsaal oder Seminarraum statt, der um die notwendige Ausstattung erweitert wurde, aber auch den Zuhörern vor Ort Platz bietet (vgl. Abschnitt 2.2.2, „Infrastruktur: Multimediaräume“).

Alternativ zum Mitschnitt von Live-Veranstaltungen als Sekundärnutzen können Präsentationen *ausschließlich* für die Aufzeichnung produziert werden. Solche Studio-Produktionen ohne Live-Zuschauer bieten weiter gehende Möglichkeiten – wie etwa die Wiederholung missglückter Passagen – und sind vergleichbar mit einer klassischen Filmproduktion, stehen damit aber nicht im Mittelpunkt der hier betrachteten Szenarios.

Die Einsatzgebiete sind im Bildungsbereich angesiedelt, insbesondere in Hochschulen und im Weiterbildungssektor – speziell für Distance-Learning, aber auch zur Unterstützung der Präsenzlehre. Neben der Produktion von E-Learning-Inhalten sind Aufzeichnungssysteme für weitere Anwendungsbereiche nutzbar, etwa für die Dokumentation oder Nutzungsanalyse von Software, die Protokollierung von Meetings etc.

3.1.2 Ziel der Präsentationsaufzeichnung

Das Ziel einer Präsentationsaufzeichnung ist die Produktion einer optimalen Abbildung der realen Präsentation auf eine abrufbare Repräsentation. Die Repräsentation soll somit den Inhalt des originären Vortrags vermitteln und damit das Original ersetzen oder (z. B. zur Nachbereitung) ergänzen. Ein System zur Präsentationsaufzeichnung stellt eine Konkretisierung der in Abschnitt 2.1.2 erläuterten allgemeinen Medienproduktionskette dar (vgl. Abbildung 2-1 und Abbildung 3-1).

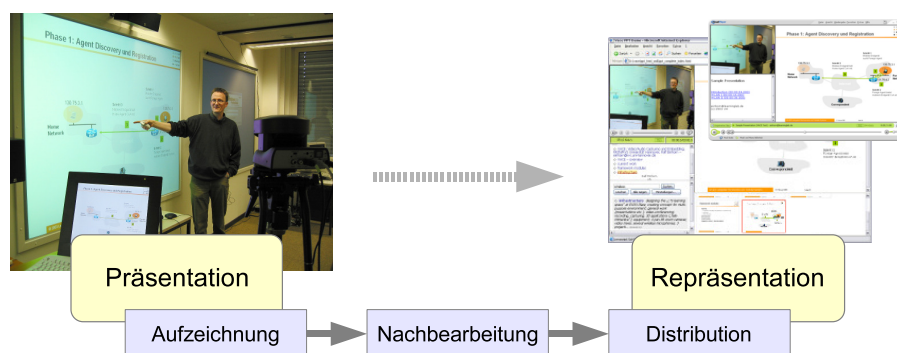


Abbildung 3-1: Blockdiagramm der Produktionskette einer Präsentationsaufzeichnung

Publiziert wird das aufgezeichnete Material äquivalent zur realen Lehrveranstaltung als eigenständige Lehreinheit oder verknüpft mit weiteren Medien, z. B. Lernmitteln oder Kursunterlagen. Darüber hinaus können die Aufzeichnungen mit weiteren Systemen (Diskussionsforen, Online-Prüfungen) verknüpft werden.

3.1.3 E-Learning-Kontext

Als ein zentrales Anwendungsgebiet wurde der Bereich E-Learning identifiziert. Zur Einordnung sei an dieser Stelle der vom IEEE vorgeschlagene *Draft Standard for Learning Technology* und die dort beschriebene *Learning Technology Systems Architecture* herangezogen [19].

Die Abbildung 3-2 stellt die beschriebene Architektur dar. Der Bereich *Content Creation* (zu der auch die Präsentationsaufzeichnung gehört) ist in dem Modell nicht explizit enthalten. Die damit hergestellten Inhalte sind jedoch als *Learning Resources* in der Architektur wiederzufinden. Die Produktion von Vorlesungsaufzeichnungen ist ein Baustein im E-Learning-Umfeld [15].

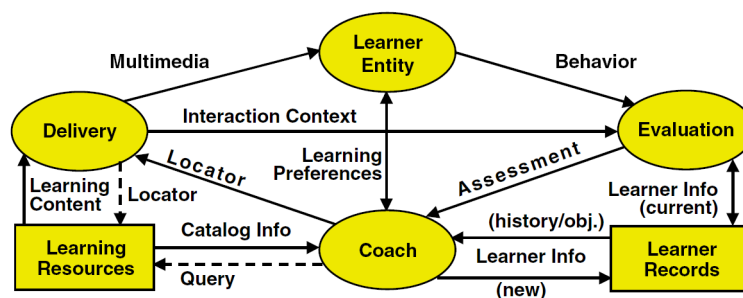


Abbildung 3-2: IEEE Learning Technology Systems Architecture [19]

Deutlich werden in dem Modell die Verbindungen der Lern-Ressourcen zur Verteilung (*Delivery*) sowie zum Lehrenden (*Coach*), die bereits bei der Erstellung von Inhalten berücksichtigt werden müssen. So ist zum einen dafür zu sorgen, dass das Material mit Hilfe von Katalogeinträgen – d. h. Metadaten – gefunden werden kann und zum anderen, dass es leicht distribuiert werden kann. Beides gilt insbesondere für Multimediadaten, die anders als z. B. Text-basiertes Material oft keine Metadaten enthalten und bei der Verteilung andere Anforderungen stellen (vgl. Abschnitt 2.1.4).

3.1.4 Rollen und Begriffe

Das hier thematisierte System zur Präsentationsaufzeichnung, -aufbereitung und -publikation wird im Folgenden vereinfacht (**Präsentations-**) **Aufzeichnungssystem** genannt, obwohl auch die anderen Produktionsschritte umfasst werden. In diesem Kontext treten verschiedene Parteien auf, deren Rollen zur Begriffsdefinition hier kurz erläutert werden:

- Der **Dozent** oder **Vortragende** trägt die Präsentation vor.
- Das **Produktionsteam** (das auch aus einer Person bestehen kann) beschreibt die an der Aufzeichnung und Nachbearbeitung Beteiligten.
- Der **Autor** entspricht dem Dozenten, sofern er maßgeblich an der Produktion der Aufzeichnung beteiligt ist.
- Der **Anbieter** („content provider“, z. B. Universität oder Institut) publiziert die hergestellte Repräsentation.

- Der **Betrachter** ist Konsument der Repräsentation, die ihm der Anbieter zur Verfügung stellt (z. B. der Studierende).

Ist vom *Benutzer* die Rede, so ist damit das Produktionsteam oder der Autor gemeint, das bzw. der das Aufzeichnungssystem einsetzt; *nicht* der Betrachter („Endnutzer“), der die Aufzeichnung abrufen.

Das Objekt der Aufzeichnung ist die **Präsentation**, das produzierte Endprodukt die **Repräsentation** (siehe Abbildung 3-1).

3.2 Analyse: Anwendung und Produktionsaspekte

In diesem Abschnitt werden die in Abschnitt 3.1.1 skizzierten Anwendungen hinsichtlich ihrer produktions- und anwendungsspezifischen Aspekte analysiert. Die Betrachtung der eingesetzten Präsentationsmedien und ihrer potenziellen Repräsentationsformen folgt anschließend im Abschnitt 3.3 und Abschnitt 3.4.

3.2.1 Produktionstyp

Gemäß ihrer Anwendungsschwerpunkte lassen sich zwei grundlegende Produktionstypen identifizieren: „alltägliche“, standardisierte Produktionen sowie Aufzeichnungen von Präsentationen mit außergewöhnlichem, individuellem Aufbau.

Standardisierte Produktionen

Bei Aufzeichnungen von regelmäßig wiederkehrenden Veranstaltungen (z. B. von wöchentlichen Vorlesungen) ergibt sich ein weitgehend identisches Aufzeichnungs-Setup: die Präsentation findet in der gleichen Umgebung und meist unter Einsatz gleicher Präsentationsmedien statt. Da auch die Herstellung der Repräsentation jeweils demselben Muster folgt, ist die gesamte Produktionskette **standardisierbar**.

Bei dieser Art Szenario steht ein kurzer Produktionszyklus, d. h. die zeitnahe Fertigstellung und Publikation im Vordergrund. Eine zeitaufwendige individuelle Gestaltung tritt dagegen in den Hintergrund, die Nachbearbeitung ist auf ein Minimum begrenzt.

Aufgrund der Ähnlichkeit der Produktionsabläufe besteht bei Standard-Anwendungen ein hohes Potenzial für die Vereinfachung bzw. die Automatisierung der Produktionskette, um so auch die Produktionskosten in einem angemessenen Rahmen zu halten.

Individuelle Produktionen

Demgegenüber stehen Non-Standard-Anwendungen, z. B. Aufzeichnungen von außergewöhnlichen Events oder Vorträgen. Hierzu gehören auch verteilte Veranstaltungen (Videokonferenzen) sowie der Einsatz unüblicher Präsentationsmittel, oder die Aufzeichnung von Experimenten.

Bei diesen Anwendungen liegt der Schwerpunkt oft auf hohen technischen Qualitäts-Anforderungen. Damit einher gehen aufwendige Kameraführung, Beleuchtung, Aufzeichnungstechnik oder spezielle Aufzeichnungstechniken, gefolgt von umfangreicher Postproduktion, z. B. dem Zusammenschnitt von Szenen für

bestimmte Zielgruppen. Beispielhaft sei hier die Aufbereitung experimenteller Physik-Vorlesungen für den Einsatz in Schulen genannt.

Die Gruppe dieser speziellen Produktionen steht nicht im Mittelpunkt der hier thematisierten Vortragsaufzeichnung, zumal dort auch der eigentliche Vortrag in den Hintergrund tritt. Dennoch ist auch die Unterstützung individueller Szenarios durch ein Aufzeichnungssystem wünschenswert, etwa um Synchronisationsinformationen für die Postproduktion zu liefern und so die Produktion zu vereinfachen.

3.2.2 Veranstaltungsforn (Anwendungsszenarios)

Für die Anwendungsszenarios lassen sich bezüglich der Veranstaltungsforn zwei grundsätzliche Einsatzgebiete definieren.

Vorlesungsszenario

Vorlesungen, Übungen, Kurse etc. werden von *einem* Dozenten durchgeführt. Der Dozent kann in den Produktionsablauf aktiv einbezogen werden. Das gesamte Produktionsumfeld (technische Ausstattung, Präsentationsmedien etc.) ist *bekannt*.

Konferenzszenario

Präsentationen auf Konferenzen, Tagungen usw. entsprechen im Aufbau grundsätzlich denen des Vorlesungsszenario. Allerdings werden die Vorträge hier von *mehreren*, unterschiedlichen Dozenten gehalten, die im Allgemeinen nicht in die Aufzeichnungsproduktion eingebunden werden. Zudem kommen unterschiedliche Visualisierungsmedien und/oder technische Hilfsmittel zum Einsatz, die teilweise vor Beginn der Veranstaltung unbekannt sind. Somit ist auch die technische Einbindung (z. B. des Dozenten-Laptops) nur beschränkt möglich.

Das Endprodukt der Aufzeichnung unterscheidet sich hingegen in den beiden Szenarios nicht.

3.2.3 Zielgruppe (Anwender von Aufzeichnungssystemen)

Als Anwender von Präsentationsaufzeichnungssystemen waren zunächst in erster Linie *AV-Produktionsteams* anzusehen. Mit zunehmender Vereinfachung oder Automatisierung von Standard-Abläufen können jedoch die Anzahl betreuender Mitarbeiter reduziert werden. In einigen Fällen kann die Aufzeichnung durch den Vortragenden selbst gesteuert werden oder vollständig automatisiert ablaufen. In diesem Fall spricht man auch von Authoring-Systemen – der Dozent wird zum *Autor* der Präsentationsaufzeichnung.

Autoren

Autorensysteme fokussieren als primären Nutzer den Dozenten („Autoren“). Der Ansatz setzt eine leichte Handhabung des Systems voraus, da die Aufzeichnung vom Vortragenden *nebenbei* gesteuert wird. Im Idealfall läuft dieser Prozess vollständig im Hintergrund ab oder die Steuerung beschränkt sich auf das Anklicken von „Aufnahme“ und „Stopp“. Diese Systemart wird daher auch „light weight authoring“ genannt [49].

Demgegenüber stehen eingeschränkte Gestaltungsmöglichkeiten, die manuelle Einflussnahme ist begrenzt. Zudem leidet die gestalterische Qualität der Aufzeichnung bei fehlender (statischer) oder automatischer Kameraführung.

Produktionsteams

Die Ausrichtung auf Produktionsteams als Zielgruppe erlaubt komplexe Systeme mit einer weit gehenden Flexibilität bezüglich des Vortragsszenarios. Damit verbunden ist jedoch ein höherer personeller Aufwand.

Dadurch, dass der Dozent nicht aktiv in die Produktion eingebunden werden muss, sind auch Szenarios wie das im Abschnitt 3.2.3 unter „Konferenzszenario“ beschriebene möglich. Umgekehrt erschwert die fehlende Einbindung des Dozenten sowie seines Rechners den Zugriff auf die Präsentationsmedien. Ohne aktive Einbeziehung des Vortragenden sind jedoch keine Autorensysteme einsetzbar.

Handhabung

In den meisten Fällen handelt es sich bei den Anwendern eines Aufzeichnungssystems (Autoren *und* Produktionsteams) nicht um AV-Produktions-Spezialisten. Neben dem Dozenten, der einen Teil der Produktion unterstützen kann, wird die Produktion zumindest bei alltäglichen Anwendungen oft z. B. von studentischen Hilfskräften betreut.

Daraus ergibt sich die Forderung nach der *einfachen Handhabung* eines Aufzeichnungssystems sowie einer leicht erlernbaren Benutzung.

3.2.4 Aufzeichnungsvorbereitung

Insbesondere Autorensysteme erfordern vom Dozenten eine spezifische Vorbereitung ihrer Präsentation für die Nutzung mit dem Aufzeichnungssystem, etwa die Erstellung oder Aufbereitung der Folien in einem vorgegebenen Format. Diese Eigenschaft ermöglicht eine effiziente Einbindung der Medien, verhindert aber die Verwendung des Systems z. B. in Konferenzen, wo die Dozenten nicht auf diese Weise involviert werden können.

Es hängt somit von der Anwendung ab, ob eine Vorbereitung seitens des Dozenten möglich bzw. erwünscht ist. Das gleiche gilt für die Umsetzung durch Produktionsteams. In beiden Fällen ist eine *Minimierung* anzustreben, sofern damit keine Nachteile bzw. Einschränkungen bei der Produktion oder Repräsentation verbunden sind.

3.2.5 Postproduktion

Postproduktion (Nachbearbeitung) bezeichnet die Abläufe, die das aufgezeichnete Material in die Form der publizierbaren Repräsentation bringen. Man unterscheidet zwischen automatisierbaren Prozessen (*Post-Processing*, z. B. Formatkonvertierung) sowie primär manuell gesteuerter Nachbearbeitung (*Post-Production* im engeren Sinn, z. B. Szenenauswahl bzw. Videoschnitt).

Abhängig von der Gestaltung einer Repräsentation kann die manuelle Nachbearbeitung sehr zeit- bzw. personalintensiv sein, weshalb insbesondere in Standard-Szenarios der Aufwand hierfür weit gehend minimiert werden sollte. An dieser

Stelle ist – neben der Unterstützung der Aufzeichnung – das Potenzial zur Vereinfachung bzw. Beschleunigung der Produktion durch den Einsatz von Präsentationssystemen zu sehen.

Gemäß der Produktionstypen aus Abschnitt 3.2.1 ergeben sich unterschiedliche Anforderungen:

- *Standardisierte Produktionen*: Die individuelle Nachbearbeitung steht bei den hier fokussierten Anwendungen zugunsten einer schnellen Produktion im Hintergrund. Allerdings können auch bei Standard-Anwendungen kleine Änderungen (wie etwa der Austausch einer Folie) nötig sein, die leicht **mit dem System durchführbar** sein sollten.
- *Individuelle Produktionen*: Insbesondere für die Aufbereitung von Aufzeichnungen außergewöhnlicher Veranstaltungen oder Zusammenfassungen ist vielfach ein Editing erforderlich. Das Produktionssystem sollte in diesem Fall diese – manuell durchzuführenden – Arbeitsschritte unterstützen. Das kann über die **Einbindung externer Applikationen** erfolgen, z. B. über Datenexport.

3.2.6 Workflow

Abgedeckte Bereiche der Produktionskette

Ein Produktionssystem kann diverse Arbeitsbereiche von der Präsentationserstellung bis hin zur Publikation der Aufzeichnung abdecken (vgl. Abbildung 3-1). Darüber hinaus ist z. B. bei Tagungen die Integration in ein Scheduling-System zur Aufnahmesteuerung denkbar, etwa um Metadaten bezüglich Autor und Thema mit dem geplanten Ablauf abzugleichen.

In einigen Fällen ist eine Einbindung von Arbeitsschritten nachteilhaft, da sie gewohnte Vorgehensweisen verhindert oder einschränkt, beispielsweise durch die Bindung des Dozenten an ein vorgegebenes Präsentationswerkzeug.

Mit einem modularen, offenen Präsentationssystem ist die flexible Anbindung an spezialisierte Systeme in verschiedenen Abschnitten des Produktionsablaufs möglich.

Automatisierung

In Standard-Szenarios (Abschnitt 3.1.1) wiederholen sich identische Produktionsschritte, die teilweise automatisierbar sind. Nachbearbeitungen wie etwa die Erzeugung von Vorschau grafiken oder der Upload auf einen Server können nach einem vorgegebenen Muster automatisiert werden. Neben der Zeit- und Kostensparnis spricht hierfür auch die Vermeidung von Fehlern bei der manuellen Bearbeitung in wiederkehrenden Abläufen.

Demgegenüber erfordert eine ansprechende Bildführung die manuelle Bedienung von Kamera, Bildmischung und Tonaussteuerung, sodass bei üblichen Anwendungen mindestens eine Person die Aufzeichnung betreut (vgl. Abschnitt 2.2.2).

3.3 Analyse: Live-Präsentation

Im Mittelpunkt jedes Anwendungsszenarios steht als Objekt der Aufzeichnung die reale Live-Präsentation. Die Präsentation zeichnet sich durch den Vortrag eines Dozenten sowie die durch ihn eingesetzten Präsentationsmittel aus (Abbildung 3-3). Die dabei eingesetzten Medien sind auf den konkreten Anwendungsfall abgestimmt und können daher differieren.

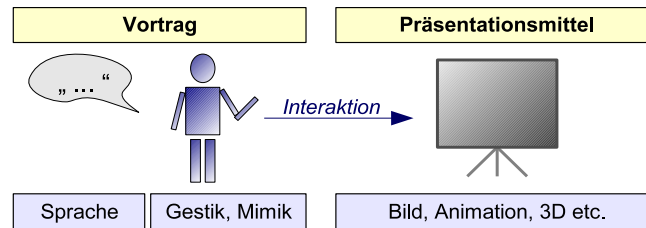


Abbildung 3-3: Elemente einer Live-Präsentation: Vortrag und Präsentationsmedien

3.3.1 Vortrag

Der Begriff *Vortrag* wird an dieser Stelle im engeren Sinn verwendet und bezeichnet die Aktion des Dozenten; die parallel genutzten Präsentationsmittel werden davon nicht umfasst und werden anschließend gesondert betrachtet.

Der Vortrag zu einer Präsentation wird primär durch die verbalen Erläuterungen und die Gestik und Mimik des Dozenten bestimmt. Dazu kommt in einigen Fällen die Interaktion mit den Visualisierungsmitteln (z. B. Zeigen mit Maus oder Laserpointer) sowie mit dem Auditorium (Fragen, Diskussion).

Im Kontext der Aufzeichnung ist entscheidend, dass der Vortrag audiovisuell erfassbar ist – von dem Teilnehmer der Präsentation sowie auch mit Hilfe von AV-Aufnahmegeräten. Für den Ton wird in größeren Hörsälen Mikrofon- und Beschallungstechnik eingesetzt, sodass bereits dem Live-Zuhörer die Sprache des Dozenten indirekt vermittelt wird. Auch das analoge Vorgehen bei der visuellen Komponente – die Darstellung des Dozenten auf einer Videoprojektion – kommt bei Tagungen und Konferenzen in größeren Räumen vermehrt zum Einsatz.

Für die weiteren Betrachtungen in Hinblick auf die Aufzeichnung kann auch der Vortrag als ein *Präsentationsmedium* aufgefasst werden.

3.3.2 Konventionelle und PC-gestützte Präsentationsmedien

Die parallel zum Vortrag eingesetzten Medien weisen eine hohe Relevanz bezüglich der Inhaltsvermittlung auf; die eigentliche AV-Aufzeichnung – die das Videobild des Dozenten zeigt – tritt in den Hintergrund. Diese Gewichtung wird dadurch deutlich, dass auch der Zuhörer des (Live-) Vortrags primär auf den Tafelanschrieb bzw. die Foliendarstellung blickt.

Die verfügbaren Medien für die Live-Präsentation lassen sich zunächst grundsätzlich in zwei Gruppen einteilen: Konventionelle und PC-gestützte Systeme (Überblick siehe Abschnitt 2.2.1).

PC-gestützte Präsentationsmedien

Mit zunehmender Digitalisierung werden vermehrt PC-basierte Systeme für Präsentationen benutzt. Diese bieten im Hinblick auf eine Aufzeichnung den Vorteil, dass das projizierte Bild bereits in elektronischer Form vorliegt.

Neben dedizierten Präsentations-Programmen wie MS PowerPoint, mit denen vorgefertigte Folien (Slides) angezeigt werden, gibt es Software zur Nachbildung der klassischen Tafel (Whiteboard), mit der während der Präsentation Grafiken erstellt sowie Bilder und Texte eingebunden werden können.

Während es sich bei den beiden eben genannten Programmarten um universell einsetzbare Präsentations-Software handelt, werden ferner auch individuell geeignete Programme in Vorträgen verwendet. Hierzu gehören fachspezifische Simulationsprogramme sowie z. B. in Software-Kursen thematisierte Programme, etwa eine grafische Programmier-Entwicklungsumgebung.

Konventionelle Präsentationsmedien

Zu den konventionellen Präsentationswerkzeugen gehören Tafeln, Folien und Dias. Für die Aufzeichnung müssen diese Medien grundsätzlich visuell erfasst und für die Verarbeitung im Rechner digitalisiert werden. Insbesondere beim Tafelanschrieb über große bzw. mehrere Flächen ergeben sich hier für die Abbildung hohe Anforderungen an die physikalische Auflösung.

Als Brücke zwischen konventionellen Folien und vom PC gelieferten Bildern kann ein Video-Presenter eingesetzt werden, der zwar die Nutzung von Folien bzw. Papier gestattet, diese aber mit Auflicht- oder Durchlichtauflage sowie einer Kamera in ein Video- oder VGA-Bild umsetzt und damit den Einsatz von Projektoren sowie die komfortable Aufzeichnung ermöglicht (siehe Abschnitt 2.2.1).

3.3.3 Interaktivität von Medien

Abhängig von der Interaktion zwischen Benutzer und Präsentationsmedium sind zwei Arten von Medien definiert [14].

Interaktive Medien

Die mit interaktiven Medien erzeugten Inhalte werden während der Präsentation dynamisch generiert oder manipuliert, etwa die Visualisierung mittels Anschrieb an eine Tafel, ein Whiteboard oder auf eine Folie, aber auch interaktiv gesteuerte Animationen.

Nicht-interaktive Medien

Nicht-interaktive Medien ändern sich lediglich zeitabhängig, unabhängig von äußeren Einflüssen wie einem Benutzereingriff. Hierzu gehören Audio, Video und (einzelne) Bilder.

3.3.4 Zeitverhalten von Medien

Anhand der Zeitkontinuität (vgl. Abschnitt 2.1.3) lassen sich (Präsentations-) Medien in zwei Klassen einteilen.

Zeitdiskrete Medien

Zeitdiskrete Medien ändern ihren Zustand durch äußere Einwirkung, z. B. per Mausklick durch den Benutzer – nicht aber durch reinen Zeitablauf. Hierzu gehören Präsentationen in diversen Formaten wie Dias, Folien sowie PC-Präsentationen mit Folien (PowerPoint, Impress etc.), HTML oder PDF.

Zeitkontinuierliche Medien

Zeitkontinuierliche Medien verändern sich lediglich zeitabhängig, aber unabhängig von äußeren Einflüssen. Ein typisches Beispiel hierfür ist Audio/Video wie z. B. der Ton und das Bild des Vortragenden.

3.3.5 Präsentationsmedien: Fazit und Zusammenfassung

Einige Präsentationsformen lassen sich nicht eindeutig den zuvor erläuterten Kriterien zuordnen. So kann der Vortragende während der Präsentation Folien per Anschlag ergänzen oder in PowerPoint-Folien können Videos eingebunden sein. Auch der Einsatz des Mauszeigers zur Erläuterung während einer (zeitdiskreten) Folien-Präsentation kann dazu führen, dass das Medium einen – abschnittsweise – zeitkontinuierlichen Charakter erhält.

Die genannten Beispiele verdeutlichen das Problem der eindeutigen Zuordnung verfügbarer Präsentationsmedien in ein Klassifikationsschema, das die Auswahl geeigneter Repräsentationsmedien vereinfachen würde. Es ist daher im Einzelfall zu entscheiden, ob etwa eine Folien-Präsentation aufgrund hoher Interaktivität zeitkontinuierlich oder zeitdiskret repräsentiert werden kann. In Tabelle 3-1 sind die Eigenschaften und möglichen Ausprägungen zusammengestellt.

Eigenschaft	Werte	
System	• Konventionell	• PC-gestützt (Software)
Zeitverhalten	• Kontinuierlich	• Diskret
Interaktivität	• Interaktiv	• Nicht-interaktiv

Tabelle 3-1: Klassifikation von Präsentationsmedien

Abschließend ist anzumerken, dass lediglich vorproduzierte Medien (wie die verbreitet eingesetzten Folien-Präsentationen) uneingeschränkt zeitdiskret repräsentiert werden können.

3.4 Analyse: Repräsentation und Distribution

Für die Herstellung einer elektronisch transportierbaren Repräsentation sind die im vorangehenden Abschnitt 3.3 analysierten Präsentationsinhalte (Vortrag und Visualisierung) aufzuzeichnen und weiterzuverarbeiten. Neben der möglichst optimalen Abbildung sind dabei Aspekte bezüglich der Handhabung von Aufzeichnung und Weiterverarbeitung, der Darstellung beim Betrachter sowie des Produktionsaufwandes zu berücksichtigen. Zentraler Punkt ist das optimale Zielformat (bzw. die optimalen Zielformate) für die Repräsentation.

3.4.1 Vortrag: AV-Aufzeichnung des Dozenten

Im Abschnitt 3.3.1 wurde erläutert, dass der Vortrag audiovisuell aufgezeichnet wird. Während die parallel benutzten Visualisierungsmedien (Folien, Tafelanschrieb) zum Teil für eine Repräsentation wiederverwendet werden können, *muss* der Vortrag in jedem Fall während der Präsentation aufgezeichnet werden.

Eine entsprechende Aufzeichnung wird mit konventioneller AV-Technik erstellt (siehe Abschnitt 2.2.2). Abhängig vom Vortragsstil enthält das Bild primär den Dozenten in Halbtotaler/Nahaufnahme. Bei Interaktivität zwischen Dozent und Auditorium kommen alternativ weitere Einstellungen hinzu, wie z. B. die Totale oder Nahaufnahmen des Publikums. Für die Vermittlung des Inhalts ist insbesondere eine gute Sprachverständlichkeit von Bedeutung. Dies gilt auch für Gespräche zwischen Dozent und Publikum.

Für die Repräsentation kann die AV-Aufzeichnung entsprechend des gewünschten Distributionswegs in das erforderliche Format konvertiert und dabei auf ein akzeptables AV-Qualitätsniveau reduziert werden. Hierbei kann optional auch komplett auf das Videobild verzichtet und nur die Tonaufzeichnung eingesetzt werden. Ist die Zielplattform bereits bei der Aufzeichnung bekannt, so können die AV-Daten bereits während der Aufnahme im Zielformat erzeugt werden (Abschnitt 2.2.2: Live-Encoding).

3.4.2 Repräsentationsform

Ein Aufzeichnungssystem ist dann für eine Anwendung geeignet, wenn mit der erzeugten Abbildung (Repräsentation) die Originalinhalte adäquat vermittelt werden können. Dies hängt wiederum wesentlich von den eingesetzten Präsentationsmitteln der Live-Veranstaltung sowie der Art der Aufzeichnung ab.

Granularität

Unabhängig von der technischen Realisierbarkeit lässt sich eine Folien- oder Tafelpräsentation in verschiedenen Granularitätsstufen aufzeichnen und repräsentieren:

- Darstellung auf der Ebene einzelner grafischer Elemente (Striche, Buchstaben, Texte, Bilder etc.). Die Whiteboard-Aufzeichnung auf **Polygonebene** ermöglicht eine exakte Reproduktion der Original-Präsentation.
- Sequenzielle Darstellung von aufgezeichneten **Einzelbildern** („Snapshots“). Hierzu wird der gesamte Inhalt der Anzeige erfasst. Bei statischen Folien bleiben dabei (abgesehen von einer möglichen Reduktion der Auflösung) sämtliche Informationen erhalten. Zeitabhängige bzw. interaktive Elemente gehen dabei (abhängig von der zeitlichen Auflösung) verloren. Für die Erfassung der Snapshots gibt es verschiedene Möglichkeiten:
 - **Festes Zeitraster**; entsprechend einer Videoaufzeichnung mit niedriger Bildrate (z. B. 2 fps statt der üblichen 25 fps).
 - **Dynamische Aufzeichnung**; ein Snapshot wird jeweils bei der Änderung des Inhaltes (z. B. einem Folienwechsel) aufgenommen.

Zudem sind Mischformen denkbar, etwa die Snapshot-Aufzeichnung ergänzt durch die kontinuierliche Erfassung der Zeigerposition.

Format der Repräsentation

Für die native Speicherung von Whiteboard-Aufzeichnungen werden hersteller- bzw. systemspezifische Formate verwendet, ein Standard existiert derzeit nicht. Zur Betrachtung sind somit spezielle Viewer-Applikationen erforderlich. Eine Ausnahme bildet das zunehmend etablierte SVG-Format, das aber ohne Erweiterungen nicht für dynamische bzw. zeitabhängige Daten geeignet ist [83].

Demgegenüber finden insbesondere Bilddatenformate (PNG, JPG etc., vgl. Abschnitt 2.1.3) eine breite Unterstützung von diversen Applikationen. Der Einsatz dieser Formate erleichtert die Verbreitung der Repräsentation, da die Darstellungsmöglichkeit beim Betrachter vorausgesetzt werden kann.

Abzuwägen ist daher, ob der Einsatz von *proprietären Formaten* zur Vermittlung des Inhaltes notwendig ist, oder ob zugunsten besserer Kompatibilität und einfacher Handhabung *Standardformate* genutzt werden können.

Gewichtung der Präsentationselemente

Bei Gewichtung der einzelnen Medienströme ergibt sich in typischen Szenarios die Reihenfolge

1. Präsentationsdarstellung,
2. Ton (Sprache des Dozenten),
3. Videobild des Dozenten.

Wie bei einer Videokonferenz ist eine minderwertige (Video-) Bildqualität vom Rezipienten eher tolerierbar als ein schlecht verständlicher Ton. Einige Anwender halten das Videobild des Dozenten gar für verzichtbar, wenngleich es der Repräsentation einen persönlichen Charakter verleiht [49].

Abhängig davon, in welchem Grad der Dozent seinen Vortrag auf die eingesetzten Präsentationsmittel (Folien, Tafel) stützt, kann eine Repräsentation durch deren ungenügende Darstellung jedoch unbrauchbar werden, etwa wenn Schriften auf Folien nicht lesbar dargestellt werden.

AV-Qualität

Die erforderliche Abbildungsqualität von Ton, Videobild sowie weiteren eingesetzten Medientypen wird weitgehend durch das Anwendungsszenario bestimmt. Das entscheidende Kriterium ist, ob die erzeugte Repräsentation den Ansprüchen zur Vermittlung der Inhalte genügt.

Die AV-Qualität ist daher eng mit der Art der Repräsentation und den dabei verwendeten Formaten verknüpft. Als Parameter sind primär die zeitliche und örtliche Auflösung von Bilddaten sowie – bei multimedialen Systemen – deren Synchronisation heranzuziehen (vgl. Abschnitt 2.1.3).

Technologieaspekt: Weiterentwicklung von Formaten

Medienformate, insbesondere Software-basierte AV-Codecs unterliegen einer permanenten Weiterentwicklung. Zwar unterliegen auch die hier produzierten Inhalte oft einer kurzen Halbwertszeit (z. B. einige Semester), sodass eine Re-Kodierung

selten notwendig wird. Produktionswerkzeuge sollten jedoch an neue Versionen oder Standards anpassbar sein.

3.4.3 Navigation und Integration: Metadaten

Im Gegensatz zum Beobachter der Live-Veranstaltung, der die Präsentation entsprechend der natürlichen Zeitachse verfolgt, erwartet der Betrachter der Aufzeichnung einen wahlfreien Zugriff auf den Inhalt. Für diesen Zweck sind Orientierungshilfen wie ein Index oder Suchfunktionen zu erstellen. Neben dem gezielten Zugriff auf einzelne Abschnitte muss auch die Präsentation als Ganzes, z. B. im WWW, auffindbar sein. Die erzeugte Repräsentation ist *einzubinden* (siehe auch Abschnitt 3.1.3). Beide Aufgabenstellungen sind mit Hilfe von Metadaten zu lösen:

- Zur Beschreibung der (gesamten) Präsentation sind – analog zu statischen Dokumenten – **globale Metadaten** (Titel, Thema, Autor etc.) bereitzustellen. Diese dienen der Archivierung bzw. dem Nachweis in Bibliotheken.
- Für die Navigation sind Metadaten für einzelne zeitliche Abschnitte zu erzeugen. **Inhaltsbeschreibende und Synchronisationsdaten** müssen korreliert werden, sodass der Betrachter z. B. zu einem gesuchten Stichwort die entsprechenden Abschnitte der Präsentation finden kann.

Während sich für den Bereich der globalen Metadaten verschiedene Modelle etabliert haben (vgl. auch Abschnitt 2.5.1), sind für die Navigation inhaltsbezogene und zeitabhängige Metadaten (siehe Abschnitt 2.5.2) zu kombinieren.

3.4.4 Distribution

Auch die anschließende Distribution (Verteilung, Publikation) der Repräsentation hat Einfluss auf deren Erstellung.

Wie auch bei den Nachbearbeitungssystemen existieren eine Reihe etablierter Distributionsplattformen, mit denen die produzierten Inhalte publiziert werden können. Zur einfachen bzw. automatisierten Einbindung der aufgezeichneten Inhalte in bestehende Plattformen (vgl. Abschnitt 2.2.4) werden entsprechende Schnittstellen zwischen Aufzeichnungs- bzw. Nachbearbeitungs- und dem eingesetzten Publikationssystem benötigt.

Die Alternative zur Integration in existierende Publikationssysteme ist die Generierung von (Re-) Präsentationen, die mittels Standardverfahren per Datenträger oder Netz verteilt werden.

- Als **Transportmedien** (Datenträger und Netz) kommen unterschiedliche Systeme wie z. B. CD, DVD, Modem, ISDN, DSL zum Einsatz (siehe Abschnitt 2.1.4). Durch die Limitierung der Bandbreite oder Datenträgergröße bzw. der dafür entstehenden Kosten entstehen Restriktionen bezüglich Format und Bitrate der Repräsentation abhängig vom Transportmedium.
- Für die Distribution können unterschiedliche **Plattformen** verwendet werden. Neben den AV-Systemen (VCR, Video-DVD) sind hiervon insbesondere rech-

nergestützte Plattformen betroffen. Es besteht sowohl ein Einfluss auf den Anbieter (distributor) als auch den Betrachter.

- **Distributionssysteme beim Anbieter:** Web-Server, CMS/LMS, AV-Streaming-Server (vgl. Abschnitt 2.2.4).
- **Lokales Wiedergabesystem beim Betrachter** (Software, Netzanbindung, Betriebssystem/AV-Player).

Für Transportmedien und Plattformen sind jeweils geeignet angepasste Medienformate und Metadaten (siehe Abschnitt 3.4.3) bereitzustellen. Um eine große Zielgruppe zu erreichen, kann es sinnvoll sein, mehrere alternative Versionen bereitzustellen. Einige AV-Streaming-Formate unterstützen die parallele Kodierung in mehreren Bitraten, sodass der Produktionsprozess vereinfacht wird. Die Anzahl unterschiedlicher Varianten kann ferner reduziert werden, indem verbreitete *Standards* eingesetzt werden und damit eine weit reichende Kompatibilität erzielt wird.

3.4.5 Medienbruch

Ein Medienbruch entsteht, wenn innerhalb der Verarbeitungskette das Medium zur Darstellung der Information wechselt. Bei der Produktion einer Präsentationsaufzeichnung tritt ein Medienbruch auf, wenn das Repräsentationsmedium nicht dem originären Präsentationsmedium entspricht. Die Nutzung des Originalformats (u. U. auch der *Original-Software*) auch für die Repräsentation ist in vielen Fällen jedoch nicht möglich oder nicht praktikabel. Allerdings können die Parameter des Repräsentationsformats (insbesondere die zeitliche und örtliche Auflösung betreffend) optimal an das zu repräsentierende Original angepasst werden.

Theoretisch wird die optimale Abbildung erreicht, wenn für Original und Repräsentation identische Medien zum Einsatz kommen, was zumindest bei PC-gestützten Präsentationsmedien (wie z. B. HTML) realisierbar ist. Neben dem möglichen Problem der fehlenden Zugriffsmöglichkeit (vgl. Abschnitt 3.2.2) sprechen oft weitere Argumente gegen den Einsatz des Originals. In der Tabelle 3-2 sind die Vor- und Nachteile zusammengestellt, wobei die Nachteile abhängig vom eingesetzten Präsentationsmedium auftreten können.

Vorteile	potenzielle Nachteile
<ul style="list-style-type: none"> • kein Informationsverlust • identische Darstellung (ggf. skalierbar) • keine Konvertierungsfehler • geringer Aufwand (keine Konvertierung nötig) • eigene weitere Aufbereitung möglich • Erhaltung ggf. vorhandener Interaktionsmöglichkeit 	<ul style="list-style-type: none"> • ungeeignete Darstellung (Platzbedarf für Darstellung) • große Datenmenge • keine Browser-Einbindung oder nicht-Skript-fähiges Plugin • erschwerte Handhabung beim Betrachter (z. B. SW-Lizenz und -Installation) • kein Kopierschutz • nicht universell einsetzbar

Tabelle 3-2: Nutzung von identischen Medien in Präsentation und Repräsentation

Ein Ansatz zur Lösung dieses Problems besteht in der **Konvertierung des Präsentationsmediums vor der Live-Präsentation**, um so für Vortrag *und* Distribution ein identisches Medienformat zu generieren. Der Medienbruch findet so bereits vor der Präsentation statt. Mit geeigneter Software wird das dabei erzeugte Format vom Vortragenden präsentiert, Zeitmarken aufgezeichnet und bei der Repräsentation mit dem entsprechenden Viewer angezeigt. Dieser Ansatz wird im Projekt VirtPresenter [78] für PowerPoint-Folien verfolgt: Die Folien werden in das vektorbasierte SVG-Format konvertiert, das sowohl beim Vortrag als auch später bei der Distribution zum Einsatz kommt. Die Nutzung eines Vektorformats reduziert zudem den Medienbruch im Vergleich zur Wandlung in Pixel-Bilder: Text und Vektorgrafiken bleiben nativ erhalten.

3.4.6 Repräsentationsmedien für Visualisierungen

Ist die Nutzung des Originalformates für die Visualisierungsmedien nicht möglich oder erwünscht, so muss das Präsentationsmedium bei der Aufzeichnung in ein geeignetes Repräsentationsmedium konvertiert werden (vgl. Abschnitt 3.4.2).

Unterschiedliche Eigenschaften von Standard-Videobildern und Präsentationen

Eine nahe liegende Option zur Aufzeichnung einer Präsentation besteht in der Nutzung klassischer Videotechnik. Konventionelle Videotechnik ist jedoch auf die Verarbeitung realer Bildsequenzen ausgelegt. Die Anforderungen an künstlich erzeugte Bildfolgen von Präsentationsmedien widersprechen diesen in einigen Punkten.

Tabelle 3-3 gibt einen Überblick über die unterschiedlichen Anforderungen. Es scheint zunächst, dass beim Einsatz von Standard-Videotechnik für die Präsentationsoaufzeichnung – abgesehen von der Pixel-Auflösung – die Anforderungen erfüllt werden. Nicht berücksichtigt sind dabei jedoch die Aspekte, die bei der fast immer notwendigen Kompression auftreten (vgl. Abschnitt 2.1.3).

Parameter	Standard-Video	Präsentation
Zeitverhalten	kontinuierlich Bildfrequenz ca. 25 Hz	diskret Bildfrequenz dynamisch, <1 Hz
Auflösung (Pixel)	ca. 160*120 bis 768*576	ca. 320*240 bis 1024*768
Farbumfang	hoch (z. B. 24 Bit/Pixel)	gering (z. B. 8 bit/Pixel)
Kontrast	weich	hart
Audio	fest gekoppelt	meist kein Audio

Tabelle 3-3: Konventionelle Video- und Präsentationsoaufzeichnung

Insbesondere hoch komprimierende Video-Codecs wie z. B. MPEG-2 nutzen die Eigenschaften realer Videosequenzen und liefern daher bei künstlichen Bildern unbefriedigende Ergebnisse. Foliendarstellungen wirken unscharf und Artefakt-behaftet; Wechsel des gesamten Bildinhaltes führen zu Störungen. Auf der anderen Seite wird das feste Zeitraster sowie die hohe Bildwechselfrequenz (MPEG-1 etwa erlaubt keine Bildraten unter 24 fps) nicht benötigt. Zwar ist eine Folie an der

exakten Position im zeitlichen Ablauf anzuzeigen. Der Inhalt der Anzeigen ändert sich in einem Zeitrahmen von typischerweise 30 bis 120 Sekunden hingegen nicht.

Auflösungsanforderung an die Repräsentation von Bildschirmpräsentationen

Wie in Abschnitt 2.2.1 erläutert, kann für die Projektion von PC-Präsentationen in Hörsälen zugunsten der Lesbarkeit oft nicht die Auflösung in vollem Umfang wie bei einem Desktop-Monitor genutzt werden. Hieraus ergibt sich die Möglichkeit, die Pixel-Auflösung für die Repräsentation zu reduzieren.

Das menschliche Auge kann bei einem Betrachtungswinkel von einer Bogenminute ($1/60^\circ$) zwei Punkte voneinander unterscheiden [65]. In Tabelle 3-4 sind die Parameter von typischen PC-Darstellungen aufgeführt. Angegeben ist jeweils der durchschnittliche Abstand vom Betrachter zum Display. Während beim Monitor die maximale physikalische Auflösung von 1024*768 Pixel wahrgenommen werden kann, ist dies in üblichen Projektionsumgebungen nicht der Fall.

Display (1024*768 Pixel)	Display- diagonale	Betrachtungs- abstand	Abstand: Diagonale	Pixel- größe	Pixelgröße (Winkel)
15"-Desktop-Monitor	0,38 m	0,7 m	1,8	0,30 mm	1,47'
Seminarraum (70")	1,80 m	7,5 m	4,2	1,41 mm	0,64'
Audimax	5,00 m	25,0 m	5,0	3,91 mm	0,54'

Tabelle 3-4: Typische Betrachtungsparameter von PC-Präsentationen

Daraus ergibt sich im Vergleich zum PC-Monitor – auf dem die Repräsentation beim Betrachter dargestellt wird – ein um etwa den Faktor zwei verlängerter Betrachtungsabstand. Wird dieser Faktor bei der Erstellung der Präsentation berücksichtigt, lässt sich die Auflösung der aufgezeichneten Projektionsbilder bis auf die Hälfte verringern, um dem Betrachter der Aufzeichnung eine visuell vergleichbare Darstellung zu bieten.

Setzt man voraus, dass der Betrachter über ein Display mit ähnlicher Auflösung verfügt, wie das zur Projektion genutzte, ist die verkleinerte Darstellung auch deshalb notwendig, um die parallele Darstellung weiterer Elemente – wie der Videoaufzeichnung – zu ermöglichen.

3.5 Systemaspekte

Nach der Untersuchung der Ausgangs- und Randbedingungen sowie des Endproduktes stellt sich abschließend die Frage nach deren Auswirkungen auf die Systemeigenschaften eines Aufzeichnungswerkzeuges. Welche Anforderungen bestehen an die Eigenschaften eines Systems, um eine effiziente Umsetzung zu realisieren? Welche Alternativen gibt es?

3.5.1 Systemtyp (Flexibilität)

Um die unterschiedlichen Anforderungen (siehe Abschnitt 3.2 und Abschnitt 3.3) an ein Produktionssystem zu erfüllen, gibt es zwei grundsätzliche Lösungsansätze:

Spezialsystem

Ein Spezialsystem ist jeweils für eine Anwendung mit klar definierten Randbedingungen (Benutzergruppe, Präsentationssystem usw.) ausgelegt. Aufgrund der festen Randbedingungen besteht dabei ein hohes Optimierungspotenzial.

Universalsystem

Ein Universalsystem ist an verschiedene Szenarios anpassbar. Um auf diese Weise ein breites Spektrum von Anwendungsszenarios abdecken zu können, gibt es wiederum zwei Ansätze:

- Ein *generisches System* arbeitet mit beliebigen Anwendungen ohne spezielle Anpassungen, etwa an die eingesetzten Präsentationsmedien. So können z. B. beliebige Visualisierungsmedien eingesetzt werden, indem sie bei der Aufzeichnung grundsätzlich in das Videobild des Dozenten gemischt werden.
- Ein *Framework-System* ermöglicht spezifische Anpassung an verschiedene Szenarios. Das Framework integriert andere Systeme bzw. lässt sich in existierende Systeme integrieren, idealerweise mittels offener Schnittstellen.

Für das zuletzt genannte *Framework-System* gelten die im nächsten Abschnitt erläuterten Aspekte der Modularität.

3.5.2 Modularität

Erweiterbarkeit und Offenheit

In ein erweiterbares Aufzeichnungssystem sind in den einzelnen Produktionsschritten verschiedene Systeme einzubinden.

- *Aufzeichnung*. Integration von AV-Encoder und Präsentations-Software zur Anpassung an unterschiedliche Szenarios (Abschnitt 3.4.1).
- *Nachbearbeitung*. Bei individuellen Produktionen kann eine manuelle Nachbearbeitung der Aufzeichnung notwendig sein (siehe Abschnitt 3.2.1, Abschnitt 3.2.5). Die dafür notwendigen Funktionen können entweder im Aufzeichnungssystem selbst implementiert werden oder es stellt Schnittstellen zu existierenden Produkten zur Postproduktion zur Verfügung – was den Implementierungsaufwand deutlich reduziert.
- *Distribution*. Schnittstellen zu Distributions-Systemen wie z. B. entsprechenden Server- oder Content-Management-Systemen. Neben der Übertragung der Mediendaten ist hierzu üblicherweise eine Anbindung auf Metadaten-Ebene notwendig, etwa um die neu eingespielten Daten dem Distributionssystem bekannt zu machen (Abschnitt 3.4.4).

Bei einer offenen Systemarchitektur können Erweiterungen über entsprechende Schnittstellen vorgenommen werden.

Verteiltheit

Vornehmlich Autorensystem sind nicht zuletzt zugunsten einfacher Handhabung in Form eines *monolithischen Komplett-Systems* implementiert. Präsentations-Software, AV-Aufzeichnung und eine eventuelle Nachbearbeitung finden auf demselben PC statt. Der Rechner des Vortragenden ist somit grundsätzlich mit

einzu beziehen und mit der entsprechenden Soft- und Hardware auszustatten. Daneben sprechen weitere Gründe für den Einsatz eines *verteilten Systems*:

- Die Aufzeichnung von Veranstaltungen wie z. B. Tagungen (vgl. Abschnitt 3.2.2) ist auf diese Weise nicht möglich. Die hierzu erforderlichen parallelen Videoaufzeichnungen von Sprecher und Projektion erfordern den Einsatz mehrerer Systeme.
- Auch die Betrachtung aktuell für die Aufzeichnung eingesetzter Infrastrukturen (Abschnitt 2.2.2, Infrastruktur: Multimediaräume) legt aufgrund der Anordnung von Dozentenplatz und Regie den Einsatz eines verteilten Systems nahe. Zudem entfällt die Notwendigkeit, den für die Präsentation eingesetzten Rechner (oft das Dozenten-Laptop) mit AV-Aufzeichnungstechnik auszustatten.

Der Nachteil parallel verteilter Aufzeichnungen besteht in der höheren Komplexität, insbesondere zur Gewährleistung der Synchronisation. Demgegenüber wird aber auch die Aufzeichnung geographisch verteilter Veranstaltungen möglich.

3.5.3 Unimedia/Multimedia

Für die Repräsentation kann entweder nur ein Medium (meist Audio/Video) eingesetzt werden oder sie kann aus synchronisierten, auf die Anwendung abgestimmten Medien zusammengesetzt werden. Weiterhin können bei der Aufzeichnung mehrere Medienströme erfasst werden, die erst bei der Postproduktion geeignet zusammengefügt werden, etwa die separat aufgenommenen Bilder der Folien und des Vortragenden. Somit ist dies nicht nur ein Aspekt der Repräsentation, sondern auch des Aufzeichnungssystems. Zu unterscheiden sind die beiden folgenden Ansätze:

Unimedia-System

Beim Unimedia-System wird lediglich ein Medium zur Aufzeichnung und Repräsentation eingesetzt, in der Regel eine konventionelle AV-Aufzeichnung. Bild und Ton werden dabei als *ein* (Intra-Stream-synchronisiertes) Medium betrachtet.

Der Vorteil liegt hierbei in der inhärenten Synchronisation und der damit verbundenen leichten Handhabung. Da die Repräsentation lediglich aus AV-Daten besteht ist eine Anpassung an unterschiedliche Distributionsverfahren durch einfache Konvertierungen leicht möglich.

Nachteilig wirkt sich die nicht optimale Abstimmung auf unterschiedliche Präsentationsmedien aus (Abschnitt 3.4.6).

Multimedia-System

Beim Multimedia-System werden für Aufzeichnung und Repräsentation verschiedene, synchronisierte Medientypen eingesetzt, die den jeweiligen Anwendungen angepasst sind. Auch der Einsatz von zwei gleichartigen Medien, etwa separate Videos für Vortragenden und Folien mit entsprechend parametrierter örtlicher und zeitlicher Auflösung fallen unter diese Gruppe. Zudem bietet der Einsatz von multimedialen Systemen die von webbasierten Anwendungen bekannte nahtlose Einbindung weiterer Medien über Hyperlinks.

Fazit: Die Betrachtungen im Abschnitt 3.4.6 sprechen für den Einsatz von multi-medialen Systemen. Auf die Anwendung zugeschnittene Medien liefern bei geringerem Datenaufkommen bessere audiovisuelle Qualität sowie umfangreiche Gestaltungsmöglichkeiten. Demgegenüber steht eine aufwendigere Produktion, die durch ein entsprechendes Aufzeichnungssystem vereinfacht werden kann.

3.5.4 Integration von Präsentationswerkzeugen

Zur Herstellung der Repräsentation müssen die Präsentationsmedien aufgezeichnet werden. Hierzu gibt es – abhängig vom Präsentationsmedium – verschiedene Methoden. Unkritisch für die Aufzeichnung sind Medien wie statische Folien, die z. B. durch Einzelbilder in PC-typischen Auflösungen repräsentierbar sind. Die komplexere Erfassung und Vermittlung umfangreicher Tafelanschriften mit hoher Auflösung kann auf fotografischer Basis realisiert werden [29].

Grundsätzlich ergeben sich drei Möglichkeiten für die Erfassung von Medien:

- **Generisch:** universelle visuelle (fotografische/filmische) Aufnahme des realen Anschriebs (Tafel, Folie) ohne spezifische Unterstützung des Werkzeuges.
- **Proprietäres Werkzeug,** z. B. Einsatz eines elektronischen Whiteboards.
- **Integration** von Präsentations-Software (PowerPoint, Web-Browser).

Der zweite Ansatz führt dazu, dass ein Dozent das vom Aufzeichnungssystem vorgegebene Werkzeug einsetzen und damit von seinem gewohnten Vorgehen abweichen muss. Hinzu kommt, dass verfügbare Whiteboard-Lösungen nicht die Auflösung von mehreren klassischen Tafeln nachbilden können, wie sie z. B. für mathematische Herleitungen benötigt wird.

Zugriffsmöglichkeiten auf rechnerbasierte Präsentationen

Zur Erfassung der mittels PC präsentierten Darstellung existieren diverse grundsätzliche Methoden. Die Aufzeichnungsmethode wird primär durch die Zugriffsmöglichkeiten auf den Dozenten-Rechner bzw. die eingesetzte Software bestimmt.

- Ein Zugriff auf die Mediendaten auf **Applikationsebene** ermöglicht die unmittelbare und damit effizienteste Kopplung. Die eingesetzten Daten können im Originalformat referenziert oder in einem geeigneten Format exportiert werden. Zeitliche Änderungen werden ebenfalls von der Applikation kontrolliert und können damit exakt erfasst werden. Allerdings kann diese Methode nur bei entsprechend vorbereiteten bzw. modifizierbaren Programmen angewendet werden.
- Auf **Betriebssystemebene** besteht die Möglichkeit, Benutzereingaben (Maus, Tastatur) zu überwachen und darüber potenzielle Änderungen an der Ausgabe zu erkennen. Über das dabei genutzte Nachrichtensystem (Message- bzw. Event-Handling) können abhängig von der Applikation weitere Änderungen registriert werden. Weiterhin kann über das Windows-System die Bildschirmdarstellung erfasst und gespeichert werden. Auch die (selten relevante) Audio-Ausgabe kann so aufgezeichnet werden.

- Ein grundsätzlich anwendbares (generisches) Verfahren zur Erfassung der Bildschirmdarstellung ist die Aufzeichnung des **Bildsignals** unter Einsatz klassischer Videotechnik. Mit Hilfe eines Scan-Konverters (siehe Abschnitt 2.2.2) wird ein typisches VGA-Signal in ein Videosignal gewandelt. Das Videobild kann mit einer festen Wiederholfrequenz aufgenommen oder aber in Form von einzelnen Bildern in dynamischen Zeitabständen gespeichert werden. Dieses Verfahren ist universell – unabhängig von der Applikation und des Betriebssystems – einsetzbar und damit besonders bei Konferenzen leicht handhabbar. Durch die Konvertierung (Skalierung, Analog/Digital-, Digital/Analog-Wandlung) entstehen jedoch Bildqualitätsverluste.

Der Zugriff auf Signalebene funktioniert auch bei Nicht-Rechner-gestützten Präsentationswerkzeugen, wenn diese mit Hilfe eines Videopresenters angezeigt werden (Abschnitt 3.3.2).

3.5.5 Standards

Formate

Bei der Produktion (Aufzeichnung, Nachbearbeitung, Publikation) sind je nach Ausprägung eine Reihe verschiedener Medien- und Metadatenformate involviert. Dies betrifft hauptsächlich („Computer“-) Daten, aber auch Hardware-nahe Systeme wie etwa die Videoaufzeichnung. Zu unterscheiden ist hier zwischen der Verwendung von proprietären Formaten und genormten oder de-facto Standards. Relevant ist dies insbesondere bei den Ein- und Ausgabeformaten (Präsentation und Repräsentation), da sie die externen Schnittstellen des Systems betreffen.

Die Relevanz von Standards verhält sich ferner proportional zur Größe der Nutzergruppe (Abbildung 3-4). Ein Dozent ist bei entsprechendem Nutzwert eher bereit, auf die Verwendung von Standard-Präsentations-Werkzeugen (und den damit verbundenen Formaten) zu verzichten. Die große Gruppe der Repräsentations-Betrachter erwartet demgegenüber Standard-Formate zur einfachen Nutzung. Innerhalb der Verarbeitungskette hingegen sind auch (lediglich temporär eingesetzte) Non-Standard-Formate tolerierbar.

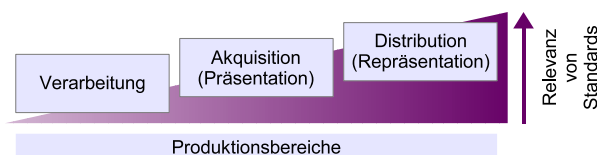


Abbildung 3-4: Relevanz von Standardformaten in der Produktion

In jedem Fall erleichtert bzw. ermöglicht der Einsatz von Standardformaten gegenüber wenig verbreiteten oder proprietären Formaten die Integration von Komponenten (vgl. Abschnitt 3.2.6).

Plattformen

Die Plattform bezeichnet die Hard- und Softwarebasis, die für die Nutzung des Aufzeichnungs-, Nachbearbeitungs- und Wiedergabesystems eingesetzt wird.

Analog zu dem zuvor erläuterten Kriterium der Formate stellt sich insbesondere bei der Präsentation, in noch höherem Maße aber bei der Repräsentation, die Frage nach der Unterstützung relevanter Plattformen, um den Anwendern die Nutzung in gewohntem Umfeld zu ermöglichen.

Speziell im Bereich Verarbeitung spielt für die Software-Integration neben dem Betriebssystem die eingesetzte Programmiersprache und Middleware eine Rolle, um systemfremde Komponenten in die Verarbeitungskette einzubinden.

3.6 Existierende Systeme

Für die Aufzeichnung von Präsentationen wurden und werden einige Systeme in Forschungsprojekten entwickelt. Zudem gibt es eine Reihe von kommerziellen („off-the-shelf“) Produkten. Aus dieser Palette wurden gezielt Systeme ausgewählt, die unterschiedliche Schwerpunkte aufweisen und so einen Überblick über die Bandbreite verschiedener Ansätze ermöglichen.

Die Auswahl basiert auf verschiedenen Systemvergleichen aus der Literatur [32][41][50][74][86] und wurde durch aktuelle Entwicklungen ergänzt. Nicht berücksichtigte Werkzeuge verfolgen entweder ähnliche Ansätze wie die hier vorgestellten oder behandeln jeweils nur Teilaspekte der Produktionskette, wie z. B. das Screen-Capturing. Diese Aspekte wurden in den vorangehenden Abschnitten berücksichtigt.

3.6.1 Systeme in der Forschung

UniTV

Reine Videoaufzeichnungen können mit konventioneller AV-Technik hergestellt werden. Die Adaption des Paradigmas „Fernsehproduktion“ für die Vortragsaufzeichnung ermöglicht den Einsatz erprobter Technologie und Arbeitsabläufe für die Live-Produktion und Nachbearbeitung. Dieser Ansatz wurde beim Münchener *UniTV*-Projekt verfolgt [44]. Das Verfahren stellt hohe Anforderungen an die verwendete Aufzeichnungstechnik und an den Wiedergabekanal, da sämtliche Inhalte mit dem Medium Video transportiert werden. Die Zugriffsmöglichkeiten des Betrachters beschränken sich auf die vom Videorecorder bekannten Mechanismen (Spulen, Zeitmarken), die direkte Einbindung weiterer Medien, wie etwa Vorlesungs-Skripten, ist nicht unmittelbar möglich. Demgegenüber bestehen beim Betrachter keine speziellen Systemvoraussetzungen. Die unimediale Umsetzung ermöglicht eine leichte Handhabung, auch bezüglich Archivierung und Verteilung.

Authoring-on-the-fly (AoF)

Das *Authoring-on-the-fly*-System [4][41] ist ein monolithisches Multimediasystem, das sich insbesondere durch eine exakte Wiedergabe von Whiteboard-Aufzeichnungen sowie eine sichere Synchronisation auszeichnet. Das hierzu entwickelte AoF-Whiteboard steht dabei als Pendant zur klassischen Kreidetafel im Fokus; Bilder, Texte und Animationen können eingebunden werden. Mit Hilfe eines Konverter-Programmes ist der Import von PowerPoint-Präsentationen in das Whiteboard möglich. Hinzu kommt die synchronisierte Aufzeichnung von

(Kamera-) Bild und Ton. Neben der Aufzeichnung ist optional auch die Übertragung mittels MBone-Technologie verfügbar. Eine Nachbearbeitung ist nicht erforderlich; mit dem Linux-Tool *aofedit* können nachträgliche Änderungen (das Editing des Audiodatenstroms und von Whiteboard-Elementen) vorgenommen werden.

Das AoF-System erfordert die Installation spezieller Software sowohl beim Vortragenden – für die Erstellung der Inhalte – als auch beim Benutzer – für das Abspielen der Aufzeichnung. Die Werkzeuge wurden zunächst unter UNIX implementiert. Neue Versionen des Whiteboards (Media Lecture Board, *mlb*) werden für Windows entwickelt, das Wiedergabeprogramm auf Basis des weitgehend Plattform-unabhängigen Java Media Frameworks (JMF).

Der Betrachter kann sich bei vollständiger Synchronisation von AV und Whiteboard-Darstellung wahlfrei in Echtzeit auf der Zeitachse der Aufzeichnung bewegen. Die Distribution erfolgt mittels Datenträger (CD-ROM), Streaming wird nicht unterstützt.

VirtPresenter

Das Projekt *VirtPresenter* der Universität Osnabrück [78] beschränkt sich im ersten Schritt auf die Verarbeitung von PowerPoint-Folien als meist genutztes Präsentationsmittel. Als Besonderheit werden die PPT-Folien vor der Präsentation in das standardisierte SVG-Format [83] konvertiert. Die SVG-Version erhält weitgehend auch die dynamischen Elemente des PPT-Originals, ist aber mit einem Web-Browser mit SVG-Plugin (gegenwärtig wird der MS Internet Explorer mit dem Adobe SVG-Plugin unterstützt) präsentierbar. Für die Präsentation des Dozenten *und* die Darstellung in der Aufzeichnung beim Betrachter wird nun die SVG-Version benutzt. Das Aufzeichnen der Timestamps erfolgt mittels JavaScript im Web-Browser des Dozenten. Die Vektorbasierung des SVG-Formats ermöglicht beim Betrachter eine freie Skalierung der Darstellung; dieselbe Datei kann sowohl für die Vorschau als auch für das Vollbild benutzt werden. Das Projekt befindet sich noch in der Entwicklung und wird im folgenden Vergleich nicht berücksichtigt.

T-Cube

Einen generischen Ansatz zur Aufzeichnung von PC-basierten Präsentationen verfolgt das Projekt T-Cube [32]. Das System setzt zur Aufzeichnung des PC-Bildes die Fernzugriffs-Software VNC [79] ein, die den Bildschirminhalt des Dozenten-PCs auf den Aufzeichnungs-Rechner per Netz überträgt. Alle zwei Sekunden wird ein Einzelbild gespeichert. Die so erzeugten Einzelbilder werden analysiert und unter Ausnutzung zeitlicher und örtlicher Redundanz in einen RealPix Datenstrom konvertiert. Der Vorteil liegt in der verlustfreien Erfassung beliebiger Bildschirminhalte (z. B. auch Annotationen des Dozenten), der Nachteil in der Einbeziehung des Dozentenrechners (VNC-Installation, Netzanbindung) sowie der rein visuellen Aufzeichnung. Das Projekt wurde zu einem Stand-Alone-System „tele-TASK“ („Teleteaching Anywhere Solution Kit“) weiterentwickelt, das sämtliche Funktionen in einem PC-basierten System integriert [6]. Auf dem mittels Ethernet verbundenen Dozenten-PC muss lediglich VNC installiert werden, die Steuerung erfolgt über das Netz mittels eines Control-Panel-Programms.

3.6.2 Kommerzielle Produkte

Real Presenter und Microsoft Producer

Systeme wie *Microsoft Producer for Microsoft Office* [35] oder *Real Presenter* [54] sind in erster Linie auf eine einfache Handhabung optimiert. Beim Live-Präsentationsmedium sind die Programme auf MS PowerPoint beschränkt. Wie auch beim monolithischen System müssen Präsentation und AV-Aufzeichnung auf demselben PC ablaufen. Hierdurch können sich Einschränkungen hinsichtlich der Verarbeitungsleistung und -qualität ergeben, wenn z. B. das zur Präsentation eingesetzte Notebook nicht über Anschlussmöglichkeiten für hochwertige AV-Signale verfügt oder die Rechenleistung nicht für das Echtzeit-Encoding ausreicht.

Real und Microsoft setzen für die Videoaufzeichnung jeweils ihre eigenen Formate RM bzw. WM ein. PowerPoint-Folien werden in Einzelbilddaten konvertiert. Abspielbar sind die Aufzeichnungen mit den jeweils frei verfügbaren Medien-Playern RealPlayer bzw. Windows Media Player.

Lecturnity

Auch in *Lecturnity* [17] ist Microsoft PowerPoint das primäre Präsentationsmedium, optional können aber auch beliebige Aktivitäten am Bildschirm aufgezeichnet werden, indem der gesamte Bildschirm oder ein wählbarer Ausschnitt per Screengrabbing erfasst wird. Ferner sind grafische Annotationen sowie das Einfügen von Texten möglich. Das System unterstützt den kompletten Produktionsfluss bis zur Erzeugung der Repräsentation, die mittels proprietärem Player oder HTML-Browser abgerufen werden kann. Zur Betrachtung der HTML-Version wird der MS Internet Explorer benötigt, sie funktioniert derzeit nicht im aktuellen Netscape 7.x bzw. Mozilla 1.x.

Das Produkt ist als Implementierung aus dem AoF-Konzept (siehe Abschnitt 3.6.1) hervorgegangen und wird in Kooperation zwischen dem Lehrstuhl *Algorithmen, Datenstrukturen und Multimedia* am Institut für Informatik der Universität Freiburg und der Firma *imc F&E* weiter entwickelt.

Gemeinsam ist den erwähnten Produkten, dass sie aus der Struktur der PowerPoint Präsentation einen Index erzeugen, der die Zeitmarken für die Repräsentation bildet. Die Repräsentation ist HTML-basiert und kann mit gängigen Browsern online und offline wiedergegeben werden.

3.6.3 Anwendungsbeispiele

Nachfolgend sind hier einige exemplarische Anwendungen zusammengestellt, in denen die Aufzeichnung und Publikation von Präsentationen (hauptsächlich Vorlesungen) praktiziert wird.

Stanford Center for Professional Development

Das *Stanford Center for Professional Development* (SCPD [68]) produziert seit einigen Jahren Videos von Vorlesungen, die zunächst per TV-Kabel als reine Fernsehsendung zur Mitarbeiterweiterbildung der Firmen des nahe gelegenen Silicon Valley übertragen wurden. Die Produktion wird von Stanford Online (SOL) für

eine Reihe von Fachbereichen betrieben. Neben der klassischen Videotechnik kommen in den letzten Jahren zunehmend Internet-basierte Systeme von Microsoft für die Kodierung und Verteilung zum Einsatz; neben dem Medium AV werden parallel z. B. PowerPoint-Folien eingesetzt.

UniTV

Im Projekt *UniTV* [75] werden diverse Vorlesungen in Sendequalität produziert, die in gekürzter Form im Bildungskanal des Bayerischen Rundfunks BR alpha ausgestrahlt werden, aber auch über einen Real Streaming-Server oder via Mbone live abrufbar sind. Das Projekt enthält verschiedene Aspekte, u. a. die verteilte Video-Produktion; der Live-Schnitt erfolgt beispielsweise an einem entfernten Standort, sodass die entsprechender Technik nicht an jedem Aufzeichnungsort vorhanden sein muss. Hervorzuheben ist die ausschließliche Verwendung der Medien Audio und Video nach vom Fernsehen bekannten Qualitätsstandards.

Sonstige Vorlesungsaufzeichnungen

An vielen Universitäten werden in unterschiedlichen Fachbereichen Vorlesungen für die Nutzung im Campusnetz oder Internet produziert. In den meisten Fällen geschieht dies als Sekundärnutzung klassischer Präsenzveranstaltungen. Es werden aber auch Aufzeichnungen für die exklusive Nutzung als Konserve produziert, etwa vom Institut für Allgemeine Nachrichtentechnik (IANT) der Universität Hannover. Hier werden die Produkte RealPresenter sowie Lecturnity eingesetzt. Die Basis für die Vorträge bilden PowerPoint-Präsentationen [12].

3.7 Systemklassifikation

Nachdem in den vorangegangenen Abschnitten zunächst eine Anforderungs- und Systemanalyse durchgeführt wurde, Anwendungen und existierende Systeme beschrieben wurden, werden im Folgenden die sich daraus ableitbaren Attribute zusammengestellt. Die ausgewählten Attribute liefern Kriterien für die Beschreibung und Bewertung eines Aufzeichnungssystems und bilden damit die Basis für ein Klassifikationsschema.

Aus der Zusammenfassung der bisherigen Abschnitte ergeben sich einige Vereinfachungen, da Aspekte unterschiedlicher Themenkomplexe zum Teil miteinander verknüpft sind.

- Die Aspekte bezüglich der Aufzeichnung und Repräsentation des *Vortrags* (Abschnitt 3.3.1, Abschnitt 3.4.1) sind für die Klassifikation weitgehend irrelevant. Die Umsetzung erfolgt bei allen Systemen mit Hilfe klassischer AV-Technik und unterscheidet sich lediglich in verfügbaren Formaten und der erreichbaren AV-Qualität.
- Die *Granularität* der Visualisierungsrepräsentation und die *Integration von PC-Präsentationsmedien* (Abschnitt 3.4.2, Abschnitt 3.5.4) korrelieren vielfach. Eine Aufzeichnung auf Polygonebene ist lediglich mit einer engen Kopplung von Aufzeichnungssystem und Visualisierungsmittel möglich, die fotografische Erfassung nicht sinnvoll realisierbar. Die Aufzeichnung auf Folienebene ist allerdings sowohl bei Software- bzw. OS-interner Kopplung als

auch bei Kopplung über das Bildsignal möglich. Dieser Aspekt fließt in den Punkt „Präsentation/Unterstützung von Präsentationsmedien“ ein.

- **Workflow**: Fast alle Systeme automatisieren die umfassten Produktionsschritte Aufzeichnung, Nachbearbeitung, Publikation.
- Die **Postproduktion** als Anwendungsaspekt korreliert vielfach mit dem Produktionstyp „individuelle Produktion“. Als Systemaspekt ist hingegen die Art der Einbindung von Postproduktionsabläufen im Bereich Workflow relevant.
- **Medienbrüche** sind aufgrund der verbreitet eingesetzten Präsentationsprogramme wie PowerPoint nicht zu vermeiden; alle Systeme setzen zur Repräsentation andere Formate als bei der Präsentation ein (Ausnahme siehe Abschnitt 3.4.5). Der Betrachtung in Abschnitt 3.4.6 (Auflösungsanforderung an die Repräsentation von Bildschirmpräsentationen) folgend ist die damit typischerweise verbundene Größenreduzierung akzeptabel.
- Formate, Transportmedien und Distributionsplattformen der **Repräsentation** sind voneinander abhängig; so eignen sich bestimmte Formate etwa für AV-Streaming und sind für definierte Plattformen verfügbar.
- **Multimedial/unimedial** als Systemaspekt korreliert in der Regel mit der hergestellten Repräsentation.
- Die im Abschnitt 3.5.5 diskutierten Aspekte zu **Standards** (Formate und Plattformen) fließen in diverse Bereiche implizit ein.

3.7.1 Klassifikationsschema

Tabelle 3-5 enthält die zusammenfassende Darstellung des Klassifikationsschemas, die Erläuterungen finden sich in den jeweils referenzierten Abschnitten.

Kategorie	Eigenschaft	Optionen/Werte
Anwendung und Produktion (Abschnitt 3.2)	Produktionstyp	<ul style="list-style-type: none"> • Standardisierte Produktionen • Individuelle Produktionen
	Veranstaltungsform (Anwendungsszenarios)	<ul style="list-style-type: none"> • Vorlesungsszenario • Konferenzszenario
	Zielgruppe (Anwender von Aufzeichnungssystemen)	<ul style="list-style-type: none"> • Autoren • Produktionsteams
	Aufzeichnungsvorbereitung	<ul style="list-style-type: none"> • nein/ja (Umfang)
	Postproduktion	<ul style="list-style-type: none"> • Systemimmanent • Anbindung
	Workflow	umfasste Produktionsabschnitte, Automatisierung

Tabelle 3-5: Klassifikation von Präsentationsaufzeichnungssystemen

Kategorie	Eigenschaft	Optionen/Werte
Präsentation (Abschnitt 3.3)	Präsentationsmedien, -formate	<ul style="list-style-type: none"> • konventionell • PC-gestützt; Formate
	Einbindung von Präsentationswerkzeugen	<ul style="list-style-type: none"> • proprietäres Tool • Integration • generisch
Repräsentation (Abschnitt 3.4)	Granularität	<ul style="list-style-type: none"> • Polygone • Einzelbilder, dynamische Wechsel • Einzelbilder, festes Zeitraster
	Repräsentationsmedien, -formate	<ul style="list-style-type: none"> • Proprietäres Format • Standardformat(e)
	Navigation	Umfang, Verknüpfung Inhalt/Zeit (j/n) globale Metadaten
	Distribution	Transportmedien Plattformen
System (Abschnitt 3.5)	Systemtyp (Flexibilität)	<ul style="list-style-type: none"> • Spezialesystem • Generisches System • Framework
	Modularität	Erweiterbarkeit (Umfang), Offenes System (j/n) Verteiltes System (j/n) Plattformen
	Unimedia/Multimedia	<ul style="list-style-type: none"> • Unimedia-System • Multimedia-System

Tabelle 3-5: Klassifikation von Präsentationsaufzeichnungssystemen

3.7.2 Bewertungsschema

Die Vielzahl unterschiedlicher Anforderungen erlaubt keinen allgemein gültigen Bewertungsmaßstab. Um dennoch eine Bewertung zu ermöglichen, können verschiedene Einzelaspekte beurteilt werden. Abhängig von deren Relevanz für die vorliegende Anwendung ist so eine Aussage über den Grad der Eignung möglich. Die hierzu geeigneten Kriterien sind im Folgenden erläutert.

Flexibilität

Für welche Szenarios ist das System einsetzbar? Inwieweit ermöglicht das System Anpassungen an individuelle Anforderungen? Wie leicht sind diese zu realisieren? Hierzu sind die Anzahl unterstützter Präsentationsmedien sowie die Systemarchitektur (*Systemtyp, Modularität, offenen Schnittstellen*) zu berücksichtigen.

Benutzerfreundlichkeit

Die Benutzerfreundlichkeit ist für die einzelnen Parteien separat zu betrachten.

- Wie einfach ist die Nutzung des Systems für den **Vortragenden**? Ist eine Aufbereitung der Präsentation für die Aufzeichnung erforderlich? Werden spezielle Formate vorausgesetzt?
- Welche Zusatzaufwände und Erleichterungen entstehen für das **Produktionsteam**? Wie integriert sich das System in verfügbare Infrastrukturen? Welche Aufgaben nimmt das System der Produktion in welchem Umfang ab?
- Wie leicht ist die Repräsentation für den **Betrachter** handhabbar? Hier kommen insbesondere die Punkte Standards (Abschnitt 3.5.5) und Navigation (Abschnitt 3.4.3) zum Tragen.

Abbildungsqualität

Die Abbildungsqualität der Repräsentation wird bestimmt durch die **AV-Qualität** von Vortrag und Visualisierung sowie die **Granularität** der Visualisierung (Abschnitt 3.4.2).

Die genannten Kriterien sind in Tabelle 3-6 zusammengefasst.

Kriterium	Bereich
Flexibilität	<ul style="list-style-type: none"> • Unterstützte Szenarios • Erweiterbarkeit
Benutzerfreundlichkeit	<ul style="list-style-type: none"> • Dozent • Produktionsteam • Betrachter
Repräsentation	<ul style="list-style-type: none"> • AV-Qualität • Granularität • Navigation

Tabelle 3-6: Bewertungsschema für Präsentationsaufzeichnungssysteme

3.7.3 Klassifikation und Bewertung existierender Systeme

In der nachfolgenden Tabelle 3-6 wurden die in Abschnitt 3.6 ausgewählten Systeme anhand des Klassifikationsschemas (Abschnitt 3.7.1) eingeordnet.

Eigenschaft	Real, MS	Lecturnity	AoF	UniTV
Produktionstyp	Standardisierte Produktionen	Standardisierte Produktionen	Standardisierte Produktionen	Individuelle Produktionen
Veranstaltungsform	Vorlesungsszenario	Vorlesungsszenario	Vorlesungsszenario	universell
Zielgruppe	Autorensystem	Autorensystem	Autorensystem	Produktionsteams
Aufzeichnungsvorbereitung	Visualisierungsformat	Visualisierungsformat	Visualisierungsformat	keine

Tabelle 3-7: Einordnung von Aufzeichnungssystemen in das Klassifikationsschema

Eigenschaft	Real, MS	Lecturnity	AoF	UniTV
Postproduktion	Systemimmanent	Systemimmanent	Systemimmanent	Einbindung externer Systeme
Workflow/ Umfang	Aufzeichnung bis Distribution	Aufzeichnung bis Distribution	Aufzeichnung und Produktion	Produktion, manuell
Präsentationsmedien	PC-gestützte Präsentationsmedien	PC-gestützte Präsentationsmedien	Whiteboard, Import von Text und Grafik	konventionell und PC-gestützt
Einbindung von Präsentationswerkzeugen	Integration von PPT	Integration, proprietäres Whiteboard	proprietäres Whiteboard, PPT-Import	generisch (Wandlung in TV-Video)
Granularität	Einzelbilder entsprechend Folienwechsel	Einzelbilder entsprechend Folienwechsel +Whiteboard	auf Polygonebene	festes Zeitraster (in Video integriert)
Format der Repräsentation	HTML+AV(RM bzw. WM)	HTML+AV optional RM oder WM	Proprietäres Format	Video
Navigation	Folien/Überschriften, synchronisiert	Folien/Überschriften, synchronisiert	Folien/Überschriften, synchronisiert, Stichworte	nicht integriert
Distribution	Datenträger, Netz/AV-Streaming	Datenträger, Netz/AV-Streaming	Datenträger	Video (Distribution nicht integriert)
Systemtyp (Flexibilität)	Spezialsystem	Spezialsystem	Spezialsystem	Generisches System
Modularität	keine Erweiterungen, geschlossenes System	keine Erweiterungen, geschlossenes System	Erweiterungen auf Basis von Quellcode	(entfällt)
Unimedia/Multimedia	multimedial	multimedial	multimedial	unimedial

Tabelle 3-7: Einordnung von Aufzeichnungssystemen in das Klassifikationsschema

In Tabelle 3-8 wurden die Systeme aus dem Abschnitt 3.6 anhand des Schema aus Tabelle 3.7.2 bewertet. Die Beurteilungen wurden dazu aus den Systemeigenschaften abgeleitet.

System		Real, MS	Lec- turnity	AoF	UniTV
Flexibilität	Szenarios	-	o	o	+
	Erweiterbarkeit	-	-	o	-
Benutzer- freundlichkeit	Dozent	+	+	+	++
	Produktion				-
	Betrachter	+	+	o	o
Abbildungs- Qualität	Granularität	o	+	+	-
	Audio/Video	o	o	o	+

Tabelle 3-8: Bewertung existierender Systeme^a

a. ++=sehr gut, +=gut, o=mittel, -=schlecht

Die Bewertung der Benutzerfreundlichkeit für die Produktion entfällt bei den Spezialsystemen, da hier die Produktion systembedingt durch den Dozenten durchgeführt wird.

3.8 Anforderungen an ein universelles Framework

Bei der Betrachtung existierender Systeme werden verschiedene Unzulänglichkeiten deutlich, die die Motivation für einen neuartigen Ansatz liefern. Hier zunächst die zentralen Punkte:

- Effizient arbeitende Systeme für standardisierte Produktionen setzen als Visualisierungsmedium MS PowerPoint voraus; es fehlt an der Möglichkeit, alternative Präsentationsmedien zu integrieren. Generisch arbeitende Systeme sind zwar universell einsetzbar, erfordern aber eine zeitaufwendige, weit gehend manuelle Produktion. Zur Unterstützung verschiedener Szenarios (bezüglich Visualisierungsmedium und Zugriffsmöglichkeit) wird daher ein flexibles, adaptives Aufzeichnungssystem benötigt.
- Für die Repräsentation kommt angesichts eines permanent steigenden Informationsangebots der Navigation eine zentrale Bedeutung zu. Die in den PPT-gestützten Systemen aus den Folienwechseln abgeleitete Granularität hat sich dabei als geeignete Basis erwiesen. Allerdings werden lediglich die Folienüberschriften ausgewertet, Erweiterungen auf weitere Metainformationen oder andere Visualisierungswerkzeuge sind nicht möglich.
- Die monolithische Systemarchitektur verfügbarer Autorensysteme ermöglicht zwar eine einfache Benutzung, widerspricht aber den vorhandenen Produktionsumgebungen, die eine verteilte Lösung erfordern. Zudem ist der Einsatz in Konferenz-Szenarios nicht möglich.

Dieser Abschnitt beschreibt die wesentlichen Zielsetzungen für ein universelles System zur Produktion von Präsentationsaufzeichnungen. Er dient damit auch als Anforderungskatalog für den im nächsten Kapitel folgenden Systementwurf.

Die Kategorien des Katalogs orientieren sich an dem eingeführten Klassifikationschema (Tabelle 3-5).

3.8.1 Anwendung und Produktion

Die Betrachtung der Anforderungen (Abschnitt 3.2) zeigt ein breites Spektrum unterschiedlicher Anwendungen und Randbedingungen, die zudem einem Wandel unterliegen. Ziel ist es, diesen Anwendungen nicht mit einer Reihe spezialisierter Systeme zu begegnen, sondern mit einem *universell einsetzbaren Framework* zur Produktion von Präsentationsaufzeichnungen.

Produktionstyp

Das System ist primär auf schnelle, *standardisierte Produktionen* ausgerichtet. Die Palette standardisierter Anwendungen kann zudem durch die Implementierung angepasster Module auf Szenarios erweitert werden, die bisher der Kategorie *individueller Anwendungen* zugeordnet werden mussten.

Veranstaltungsform (Anwendungsszenario)

Das System soll beide Anwendungsszenarios (Vorlesungen und Konferenzen) unterstützen. Eine Nutzung ist damit auch ohne aktive Beteiligung des Dozenten an der Aufzeichnung möglich.

Zielgruppe

Als Zielgruppe werden in erster Linie *Produktionsteams* angesehen. Das System dient primär der Integration bzw. Unterstützung von bisher manuell eingesetzten Systemen und damit der Vereinfachung des Produktionsprozesses (vgl. Abschnitt 2.1.2). Die Basis für die Aufnahme bildet existierende AV-Aufzeichnungstechnik-Infrastruktur (vgl. Abschnitt 2.2.2).

Aufzeichnungsvorbereitung

Eine systemspezifische *Aufbereitung der Präsentationsmedien* für die Aufzeichnung ist *nicht erforderlich*. Für den optionalen Einsatz von Modulen zur Visualisierungsaufzeichnung ist eine entsprechende Software-Installation tolerierbar.

Postproduktion

Für die Postproduktion sind zwei alternative Verfahren möglich:

- Die Mehrzahl der standardisierten (im Idealfall: automatisierten) Produktionen erfordern lediglich rudimentäre Funktionen zur Nachbearbeitung, die im System verfügbar sind. Hierzu gehört die Festlegung von Start- und Endzeitpunkt der Repräsentation bzw. die Extraktion von festgelegten Ausschnitten. Das Endprodukt stellt hier meistens eine Web-basierte Multimedia-Präsentationen dar.
- Aufwendige Produktionen für individuelle Anforderungen werden mit Hilfe existierender Postproduktions-Systeme (Video-Editing, DVD-Authoring) durchgeführt. Für diese System können zur Vereinfachung der Nachbearbeitung

Synchronisationsdaten exportiert werden. Das System selbst enthält jedoch keine weiterführenden Funktionen zur Postproduktion, die von existierenden Produkten (Abschnitt 2.2.3) abgedeckt werden können.

Workflow

Die Produktionskette wird von der *Aufzeichnung bis zur Publikation* der Repräsentation unterstützt. Dabei werden weit gehend bestehende Basis-Systeme eingebunden. Dies betrifft z. B. Präsentationswerkzeuge aber auch AV-Formate und die zum Abspielen der Repräsentation benötigte Software beim Endnutzer. Für die AV-Aufzeichnung werden verfügbare Systeme in gewohnter Weise eingesetzt.

Die manuelle (Nach-) Bearbeitung wird u. a. durch die weit gehende *Automatisierung* minimiert bzw. vereinfacht, um so den personellen Aufwand für Standardaufgaben zu reduzieren und eine zeitnahe Publikation der Aufzeichnung zu ermöglichen.

3.8.2 Präsentation

Geeignete Präsentationsformate

Grundsätzlich können beliebige Präsentationsmedien verwendet werden. Einschränkungen bezüglich verwendbarer Medien ergeben sich lediglich durch die im nächsten Abschnitt beschriebenen Eigenschaften der Repräsentation.

Die Effizienz der Aufzeichnung und Weiterverarbeitung (Bildqualität, Metadatenextraktion) ist dabei abhängig von der Art der Integration des Präsentationswerkzeugs in das Framework.

Integration von Präsentationswerkzeugen

Abhängig vom Produktionstyp können Präsentationswerkzeugen auf verschiedene Arten integriert werden.

- Zur Erfassung von Präsentationsmedien wie z. B. PowerPoint-Folien oder Webseiten werden auf dem Dozenten-PC in das System integrierte Aufzeichnungsmodule eingesetzt. Die Kopplung ermöglicht die effiziente Extraktion von Medien- und Synchronisationsdaten. Mit Hilfe offener Schnittstellen können so auch individuelle Anwendungen in das Framework eingebunden werden.
- Um beliebige visuell erfassbare Präsentationen zu integrieren, erfolgt alternativ die Aufzeichnung generisch als Bilddatenstrom. Diese Methode ist zudem im Konferenz-Szenario einsetzbar, wenn kein Zugriff auf das PC-System (OS, Software) des Vortragenden verfügbar ist.

3.8.3 Repräsentation

Ein Schwerpunkt für die Repräsentationserzeugung liegt auf der Erzeugung von Navigationsmöglichkeiten für den Betrachter. Hierzu eignet sich eine Gliederung in einem Zeitraster von etwa ein bis zwei Änderungen pro Minute.

Granularität

- Entsprechend der angestrebten Aufteilung in Abschnitte bietet sich eine Granularität auf Basis von *dynamisch erzeugten Snapshots* an. Bei den verbreitet ein-

gesetzten folienbasierten Systemen ergibt sich dieses Raster implizit durch die Folienwechsel.

- Alternativ können für die generische Visualisierungserfassung Videoaufzeichnungen mit festem Zeitraster verwendet werden. Die Partitionierung in Kapitel zur Erstellung der Navigation muss dann in der Nachbearbeitung oder unter Zuhilfenahme von parallel aufgezeichneten Annotationen erfolgen.

Stärker interaktive Medien (d. h. Medien mit einer hohen Anzahl von Änderungen pro Zeit wie z. B. ein Whiteboard) werden *nicht* auf Basis grafischer Elemente erfasst, auch aufgrund eines fehlenden Standardformates für die Repräsentation.

Format der Repräsentation

- Das Format der Repräsentation ist abhängig von den eingesetzten Aufzeichnungs- und Wiedergabe-Werkzeugen und somit **grundsätzlich nicht eingeschränkt**. Daher können auch proprietäre Anwendungen und Formate zum Einsatz kommen, wenn diese in das Framework integriert werden.
- Der Fokus liegt jedoch auf der Unterstützung verbreiteter verfügbarer Standard-Programme zur Wiedergabe, um eine einfache Handhabung beim Betrachter sicherzustellen. Hiermit wird der Einsatz von **Standard-Formaten** für die Repräsentation obligatorisch.

Bei proprietären Präsentationssystemen besteht somit die Wahl zwischen der Abbildung (Konvertierung) in ein Standardformat und dem Einsatz des nativen Formats (auch für die Repräsentation) durch Systemeinbindung.

Aufgrund des modularen Konzepts können geeignete AV-Systeme integriert werden. Damit ist auch die Anpassung an neue Technologien wie z. B. effizientere Codecs leicht realisierbar. Die Qualität von Ton und Bild ist (abhängig vom eingesetzten Codec) skalierbar und damit an die vorliegende Anwendung anpassbar.

Navigation und Integration

Wie einleitend erwähnt liegt ein Schwerpunkt des Systementwurfs auf einer effizienten Realisierung der Navigation (Übersicht, Suche) der Repräsentation. Gesuchte Inhalte sollen leicht auffindbar sein. Dieses Ziel korreliert mit der oben genannten Granularität auf Abschnittsebene.

Darüber hinaus soll die Integration der erzeugten Inhalte in Informationssysteme mittels Metadaten bereits bei der Produktion unterstützt werden.

Distribution

Auch für die Verteilung der Repräsentation wird auf verfügbare Standard-Systeme (Web-/Streaming-Server, CD etc.) aufgesetzt. Zudem kann die Unterstützung von E-Learning-Plattformen, z. B. durch Export von Metadaten, realisiert werden.

3.8.4 Systemarchitektur

Das System wird als universell nutzbares, modulares **Framework** gestaltet.

Modularität

Das Aufzeichnungssystem soll aufgrund der unterschiedlichen Präsentations-Szenarios flexibel anpassbar sein. Speziell bei der Aufzeichnung, aber auch in den anderen Produktionsschritten bietet sich dazu ein modulares Konzept an (Tabelle 3-9).

Bereich	Modultypen
Aufzeichnung	<ul style="list-style-type: none"> • AV-Aufzeichnungsmodule für die Erfassung des Vortrags, • Aufzeichnungsmodule zur Integration von Visualisierungsmedien, • generische Aufzeichnungsmodule.
Nachbearbeitung	<ul style="list-style-type: none"> • Module zur Metadatenextraktion, • Medienkonverter, z. B. für die Re-Kodierung von AV-Daten.
Distribution	<ul style="list-style-type: none"> • Exportfilter für verschiedene Medienformate, • Schnittstellen für Distributionsplattformen.

Tabelle 3-9: Module eines Aufzeichnungs-Frameworks

Für den Einsatz in aktuellen und zukünftigen Szenarios muss das Framework **erweiterbar** sein, sodass z. B. aus anwenderspezifischen Präsentationsprogrammen Daten exportiert und aufgezeichnet werden können. Ferner ist die Möglichkeit der Anpassung an Programme für die Verarbeitung neuer Medienformate erforderlich.

Das System zur Aufzeichnung arbeitet gemäß der einleitend und in Abschnitt 3.5.2 diskutierten Aspekte **verteilt**.

- Für verschiedene Anwendungsbereiche können optimal geeignete Systeme eingesetzt werden. Während z. B. die Präsentationsfolien auf dem Dozenten-Notebook erfasst werden, erfolgt die AV-Aufzeichnung auf einem entsprechend ausgestatteten PC im Regiebereich.
- Neben der Umgehung von eventuellen Leistungsdefiziten sind mit der verteilten Aufzeichnung optional redundante Systeme realisierbar.
- Die Verteilung ermöglicht eine flexible Erweiterbarkeit, so können beliebig viele parallele Mediendatenströme aufgezeichnet werden, z. B. individuelle Annotationen.

Aufgrund der Modularität können auch die weiteren Produktionsschritte (Nachbearbeitung, Distribution) verteilt erfolgen.

Uni-/Multimedia

Das System soll entsprechend der in Abschnitt 3.5.3 erläuterten Vorteile **multimedial** arbeiten. Es werden parallel mehrere Aufzeichnungen in geeigneten Formaten angefertigt, z. B. die Videoaufzeichnung des Vortragenden sowie Einzelbildfolgen der Präsentation. Die Daten werden parallel verarbeitet und für die Repräsentation synchronisiert.

Wird für eine Gruppe von Betrachtern eine **unimediale** Repräsentation benötigt (etwa ein leicht zu handhabender AV-Clip), so kann diese bei Bedarf aus den einzelnen Medienströmen der Multimedia-Aufzeichnung extrahiert werden.

3.8.5 Weitere Aspekte

Standards

Entsprechend der zuvor erläuterten Eigenschaften (Nutzerfreundlichkeit, Integration) werden weit gehend Standard-Formate eingesetzt, insbesondere bei der Repräsentation. In den Fällen, in denen dies nicht möglich oder sinnvoll ist, können individuelle Formate (bzw. entsprechende Programme) über Erweiterungen des Frameworks für die Aufzeichnung oder Repräsentation integriert werden.

Wie bei den Formaten stehen auch bei den Plattformen verbreitete und kostengünstige Systeme (Hard- und Software) im Mittelpunkt. Bei der Produktion wird hierzu auf die verbreitet verfügbaren Systeme abgestellt (Abschnitt 2.2). Die primäre Plattform für die Wiedergabe der Repräsentation bilden Web-Browser (Vorteile siehe Abschnitt 2.2.4). Alternativ bietet sich für das Playout von synchronisierten MM-Daten der Einsatz von SMIL-Playern an.

Benutzerfreundlichkeit

- Für den Dozenten ergibt sich die Möglichkeit, seine bisher genutzten Präsentationswerkzeuge weiterhin zu verwenden.
- Für das Produktionsteam vereinfacht sich gegenüber dem manuellen Vorgehen insbesondere die Nachbearbeitung, da fehlerträchtige Standard-Arbeitsschritte (wie die Zuordnung von Tafelbildern zu AV-Abschnitten) entfallen. Die Framework-Architektur benötigt für unterschiedliche Anwendungen lediglich entsprechende Module; die Betreuung unterschiedlicher Systeme entfällt.
- Dem Betrachter wird die Nutzung des produzierten Materials dadurch vereinfacht, dass er lediglich Standard-Software (Web-Browser, Medien-Player) benötigt. Neben einer guten Darstellungsqualität bietet die Repräsentation eine effiziente Navigation.

3.8.6 Systemübersicht

In Tabelle 3-10 sind die gewünschten Eigenschaften der zu entwickelnden Systems in das Klassifikationsschema aus Abschnitt 3.7.1 eingeordnet.

Kategorie	Eigenschaft	Optionen
Anwendung und Produktion (Abschnitt 3.2)	Produktionstyp	Standardisierte Produktionen, z. T. Individuelle Produktionen
	Veranstaltungsform (Anwendungsszenarios)	Vorlesungsszenario <i>und</i> Konferenzszenario
	Zielgruppe (Anwender von Aufzeichnungssystemen)	Produktionsteams
	Aufzeichnungsvorbereitung	nein
	Postproduktion	möglich via Schnittstellen, essenzielle Funktionen integriert
	Workflow	Bereiche: Aufzeichnung bis Distribution Einbindung von Systemen in allen Bereichen Automatisierung möglich
Präsentation (Abschnitt 3.3)	Präsentationsmedien, -formate	beliebige PC-gestützte Präsentationsmedien <i>und</i> Konventionelle Präsentationsmedien
	Einbindung von Präsentationswerkzeugen	Integration mit Modulen, optional: generische Aufzeichnung
Repräsentation (Abschnitt 3.4)	Granularität, vgl.: Integration von Präsentationswerkzeugen	Einzelbilder, dynamische Wechsel und Einzelbilder, festes Zeitraster (Video)
	Repräsentationsformat	Standardformate, Nutzung proprietärer Formate möglich
	Navigation, Metadaten	Globale Metadaten und Verknüpfung Inhalt/Zeit
	Distribution	Standard: Netz- und Datenträgerbasiert
System (Abschnitt 3.5)	Systemtyp (Flexibilität)	universelles Framework verschiedene Plattformen
	Erweiterbarkeit, Offenheit	Erweiterbar auf Ebene von Datenschnittstellen, offenes System. verteiltes System
	Unimedia/Multimedia	Multimedia-System

Tabelle 3-10: Anforderungsprofil eines universellen Aufzeichnungssystems

4 Entwurf und Implementierung

Dieses Kapitel beschreibt die Umsetzung der in Abschnitt 3.8 skizzierten Systemanforderungen an ein modulares Framework zur universellen Produktion von Präsentationsaufzeichnungen.

Zunächst werden dazu die grundlegenden Ideen und Konzepte erläutert. Daran anschließend werden die für den Systementwurf benötigten Basistechniken (Metadatenmodell, Systemarchitektur, Synchronisationsmechanismen) entwickelt und abschließend die Umsetzung in eine prototypische Implementierung beschrieben. Das entworfene Framework wird im Folgenden kurz *VACE* (für *Video & Audio Capturing and Embedding*) genannt.

4.1 Konzepte des Systementwurfs

Dieser Abschnitt liefert eine Übersicht der grundlegenden Strukturen und Ideen, auf deren Basis der Systementwurf aufbaut.

4.1.1 Multimedialität

Entsprechend den unterschiedlichen Anforderungen innerhalb einer Präsentation und angepasst an verschiedene Präsentationswerkzeuge und Zugriffsmöglichkeiten kommen jeweils adäquate Medienformate zum Einsatz. Bei der Mehrzahl von Präsentationsformen bietet sich der parallele Einsatz mehrerer Medien an, in den meisten Fällen für die AV-Aufzeichnung des Vortragenden und die Erfassung von Folien oder Tafelanschrieb.

Neben der effizienten Nutzung geeigneter Medien wird mit der multimedialen Aufnahme auch eine Mehrfachnutzung ermöglicht – nicht zu verwechseln mit der ebenfalls gegebenen Wiederverwendbarkeit. Aufgezeichnete Mediendaten können bei der Nachbearbeitung für verschiedene Publikationsformen (z. B. Intranet, DVD) optimal zusammengestellt werden.

Hierzu müssen verschiedene Medien unabhängig voneinander auf geeigneten Plattformen aufgezeichnet und anschließend zusammengeführt werden. Hieraus ergibt sich die Forderung nach einem *verteilten* System, woraus wiederum die Notwendigkeit der Partitionierung in einzelne Module folgt.

4.1.2 Modularität: Offenheit, Integration und Synchronisation

Voraussetzung für die parallele Nutzung diverser Medientypen in verteilten Prozessen – speziell für die Aufzeichnung – ist die Aufteilung der Produktion auf separat einsetzbare Module.

Diese Module implementieren Schnittstellen zu den eingesetzten Basissystemen oder – sofern keine verfügbar oder einsetzbar sind – bilden generische Tools, z. B. zur allgemeinen Aufzeichnung von Projektionsinhalten.

Offene Standards und Schnittstellen

In allen Produktionsbereichen Aufzeichnung, Nachbearbeitung und Publikation gibt es ein Vielzahl von Systemen bzw. Produkten, die bereits eingesetzt werden oder durch die Anwendung vorgegeben sind (siehe auch Abschnitt 2.2). Diese sollen nicht ersetzt, sondern integriert werden.

Eine Integration ist bidirektional möglich. So können Systeme in das Produktions-Framework integriert werden, wenn sie die entsprechenden Schnittstellen zur Verfügung stellen. Umgekehrt kann eine System in das Framework integrieren werden, wenn es VACE-konforme Daten verarbeitet bzw. erzeugt. In beiden Fällen vereinfacht der Einsatz von offenen Standards die Realisierung.

Einbindung von Systemen

In der gesamten Produktionskette werden weit gehend existierende Standard-Systeme eingesetzt. Neben der Implementierung von Schnittstellen zu diesen Systemen kommen (insbesondere für die Aufzeichnung) zudem generische Tools zum Einsatz.

Darüber hinaus ist es möglich, Nicht-Standard-Systeme (z. B. zur 3D-Visualisierung) einzubinden, die über proprietäre Formate verfügen. Auf diese Weise können individuelle Visualisierungsanwendungen nativ in Aufzeichnung und Repräsentation eingebunden werden [27].

Erweiterbarkeit

Die Erweiterbarkeit wird über Schnittstellen und die Nutzung von Standard-Datenformaten realisiert. Zudem ist bei der Implementierung die leichte Wiederverwendbarkeit von Code (Klassen) zu gewährleisten, um so schnell und effizient neue Module zu entwickeln.

Kopplung

Die einzelnen Module kommunizieren untereinander lediglich mittels Datenaustausch. Diese datenorientierte Umsetzung ermöglicht im Vergleich zur Kommunikation über Programmierschnittstellen die einfache Einbeziehung verschiedener Betriebssysteme und Hardware, da lediglich die Kompatibilität der Meta- und Mediendaten sicher gestellt werden muss. Somit ergibt sich eine leicht zu handhabende Umsetzung des Modulkonzepts. Die lose Kopplung erlaubt den voneinander unabhängigen Einsatz von Modulen in allen Produktionsschritten.

Gleichzeitig entfällt damit aber die – primär die Synchronisierung betreffende – enge Kopplung mit Hilfe von Interprozesskommunikation. Um dennoch die Synchronisation bei der Aufzeichnung zu gewährleisten, wird eine gemeinsame Zeitbasis vorausgesetzt. Realisiert wird diese durch die Synchronisation der Uhren auf allen beteiligten Aufzeichnungsgeräten mit der Echtzeit, wie in Abschnitt 2.4.2 beschrieben.

Synchronisation

Die eben vorgestellte Kopplung für unabhängig voneinander aufgezeichnete Medienströme erlaubt lediglich eine Inter-Stream-Synchronisation (vgl. Abschnitt 2.4.3). Die erreichbare Synchronisationsgenauigkeit ist dabei u. a. abhängig vom Gleichlauf der Uhren auf den parallel genutzten Systemen.

Für die zeitkritische Synchronisation von Ton und Bild bei der Aufzeichnung des Vortragenden (Lippensynchronisation, +/- 80 ms [69]) wäre eine auf diese Art realisierte AV-Synchronisation nicht ausreichend. Daher wird diese dadurch gewährleistet, dass beide Medien als ein gemeinsames – Intra-Stream synchronisiertes – Medium betrachtet werden. Diese Voraussetzung ist zulässig, da alle praxisrelevanten AV-Aufzeichnungssysteme parallel synchron Audio und Video verarbeiten (vgl. Abschnitt 2.4.3 und Abschnitt 2.2.2). Die gemeinsame Verarbeitung von Bild und Ton entspricht außerdem auch den hier relevanten Produktionsumgebungen.

Im Vergleich zu einer nachträglich herzustellenden Synchronisation mittels redundanter Aufzeichnung oder Synchronisationsmarken ist das hier vorgestellte Modell mit wesentlich geringerem Aufwand zu realisieren [42].

4.1.3 Einsatz von Metadaten

Basis für die Produktionskette

Als Basis für den verteilten Produktionsablauf von der Aufzeichnung über die Postproduktion bis zur Publikation dient ein Metadatenmodell, das konventionelle (auf den Inhalt bezogene) sowie Synchronisations-Metadaten für einzelne Elemente der Aufzeichnungen enthält. Damit dient es der Erzeugung von Navigationselementen und zur Synchronisation der separat aufgezeichneten Medien.

Gleichzeitig transportiert das Metadatenmodell optional die zur Steuerung des Produktionsablaufs notwendigen Datenstrukturen.

Realisierung von Navigation in der Repräsentation

Die Kopplung von (konventionellen) inhaltsbezogenen Metadaten und Synchronisationsinformationen für einzelne Abschnitte ermöglicht die Realisierung von Navigation in der Repräsentation und damit die Möglichkeit, die aus textbasierten Medien bekannten Such- und Navigationsfunktionen auch für audiovisuelle Medien zu nutzen.

Während eine automatisierte Indexierung bzw. Metadatenerzeugung aus AV-Daten aufwendig zu realisieren ist, können über die verknüpften Datenströme auch assoziierte AV-Aufzeichnungen eingebunden werden. So kann beispielsweise die auf einer Folie verwendete Überschrift als Metadatum zum Auffinden der mündlichen Erläuterungen in der AV-Datei dienen, da der Vortragende zu diesem Zeitpunkt über dieses Thema gesprochen hat.

Auch können aus z. B. einer Bilddatenanalyse der verwendeten Folien weitere Metadaten gewonnen werden, indem beispielsweise mit Hilfe von Texterkennung (OCR) und Wortlisten Stichworte extrahiert werden.

4.1.4 Verteilte Produktion

Sämtliche Produktionsschritte können verteilt, d. h. auf verschiedenen Systemen erfolgen. Bei der Aufzeichnung können parallel mehrere Mediendatenströme in adäquaten Formaten und mit geeigneten Systemen aufgezeichnet werden. Hierdurch werden Restriktionen aufgrund von Rechenleistung oder Kompatibilität umgangen. Ferner ist die räumliche Trennung beteiligter Aufzeichnungssysteme realisierbar, parallele Aufzeichnungen können auf mehrere, optimal geeignete Systeme verteilt werden. So kann z. B. der Dozent zur Präsentation seinen Laptop einsetzen, während die AV-Aufzeichnung auf einem entsprechend ausgestatteten stationären PC erfolgt.

Neben der Aufzeichnung kann auch die Nachbearbeitung – etwa die nachträgliche Extraktion von Metadaten aus Mediendaten – aufgrund der modularen, datenbasierten Architektur verteilt erfolgen.

Schließlich weist oft auch bei der Distribution ein verteilter Ansatz Vorteile auf; so können AV-Daten über einen Streaming-Server ausgespielt werden, während statische Dokumente wie HTML oder Bilder auf einem Webserver zur Verfügung stehen (vgl. Abschnitt 2.1.4). Ferner kommt vielfach die separate Speicherung von Medien- und Archivierungsdaten in Bibliotheken zum Einsatz.

4.2 Produktions-Framework

Dieser Abschnitt beschreibt den Aufbau des VACE-Produktions-Frameworks. Basis für die Schnittstellen der Framework-Module bildet das VACE-Metadatenmodell, dessen Anforderungen durch die im Folgenden beschriebenen Funktionen des Frameworks definiert werden.

4.2.1 Produktionsablauf

Die Produktion umfasst die im Abschnitt 2.1.2 vorgestellten Bereiche Aufzeichnung, Nachbearbeitung und Distribution. Einleitend wird zunächst eine Übersicht der einzelnen Arbeitsschritte und deren Zusammenhänge skizziert.

Planung (optional)

An die verfügbaren Capturing-Module können Aufzeichnungsaufträge in Form von VACE-Metadaten verteilt werden. Die Daten enthalten eine Beschreibung aller aufzuzeichnenden Streams sowie der gewünschten Parameter (z. B. Bildgröße, Framerate) und den geplanten Aufzeichnungszeitraum. Daneben kann ein Ziel-Verzeichnis (lokal oder im Netz) angegeben werden, in das die fertig gestellten Medien- und Meta-Dateien übertragen werden.

Ohne die Vorbereitung von Aufträgen können sämtliche Aufzeichnungen auch manuell gesteuert werden. Es entfällt lediglich die Automatisierung bezüglich Parameter-Einstellung der Aufzeichnung, Zeitsteuerung und Dateiverwaltung.

Aufzeichnung

Es werden mehrere Streams mit geeigneten Systemen verteilt aufgezeichnet. Hierzu bestehen drei Alternativen.

- A1: Verarbeitung der Aufträge; dabei werden Mediendaten erzeugt, die im Auftrag vorliegenden Metadaten werden aktualisiert und ergänzt.
- A2: Unabhängige, selbstständige Aufzeichnung; Erzeugung von Medien- und Metadaten.
- A3: Aufzeichnung von Mediendaten mit Hilfe von allgemeinen Komponenten, die keine VACE-konformen Metadaten erzeugen (generische Aufzeichnung).

Nachbearbeitung

- Zusammenführung der aufgezeichneten Daten:
 - Zusammenführung von Metadaten aus A1 und/oder A2.
 - Import von Mediendaten aus A3.
- Erzeugung von publikationsfähigen Metadaten (aus VACE-Metadaten).

Optionale Arbeitsschritte in der Nachbearbeitung:

- Erzeugung von Metadaten aus Mediendaten.
- Manuelle Bearbeitung von Metadaten, Festlegen der Präsentations-Timeline.
- Konvertierung von Mediendaten in andere oder alternative Formate.

Distribution

Playout der Mediendaten aus der Aufzeichnung oder Konvertierung zusammen mit den erzeugten Publikationsmetadaten.

Die Abbildung 4-1 zeigt vereinfacht das Schema des Produktionsablaufs. Die Capturing-Module der drei Klassen VACE-Capturing (A1/A2), Generic Capturing (A3) und Annotator (A2) können optional und parallel eingesetzt werden.

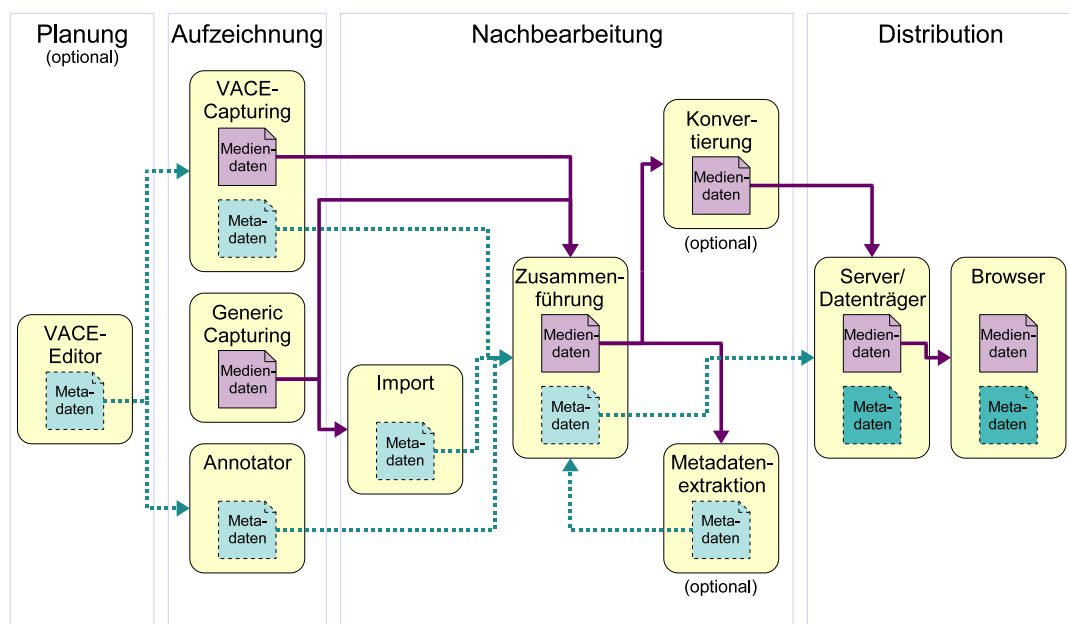


Abbildung 4-1: Diagramm zum Produktionsfluss im VACE-Framework

4.2.2 Aufzeichnung

Das verteilte Capturing-System ermöglicht die parallele, unabhängige Aufzeichnung von Mediendaten in verschiedenen Formaten auf unterschiedlichen Plattformen. Neben den Medientypen sind dabei Aspekte zur Realisierung der Module sowie die Systemsicht zu berücksichtigen.

Medientypen

Die Formate können jeweils adäquat bezüglich Auflösung, Zeitverhalten usw. abhängig vom vorliegenden Präsentationsszenario gewählt werden. Die folgende Liste enthält typische Beispiele, sie ist nicht als vollständig anzusehen.

- Audio und Video zur Aufnahme für die Sprache und das Bild des Vortragenden
- Zusätzliche Videoaufzeichnungen, z. B. für Experimente.
- Tafelbild (dynamisch erzeugt; Whiteboard, Tafel,...).
- Vorgefertigte Folien (Slides) in HTML, PPT usw.; Als Referenz auf das Original oder exportiert als Folge von Bilddateien.
- Annotationen (Kommentare, Notizen, Stichworte), auch von mehreren Benutzern. Diese Echtzeit-Anmerkungen können auch in Form einer Gliederung ausgeführt werden und somit fehlende Metadaten ersetzen. Auch können diese bereits vorbereitet und während der Aufzeichnung lediglich mit Zeitmarken versehen werden.

Optimal für die Erzeugung einer Repräsentation ist in vielen Fällen der Einsatz des bei der Live-Präsentation genutzten Originalformates, was jedoch nur in Einzelfällen realisierbar ist (vgl. Abschnitt 3.4.5). So ist in jedem Fall für die Erfassung von Bild und Ton des Vortragenden eine AV-Aufzeichnung erforderlich. Werden während der Live-Präsentation z. B. HTML-Seiten eingesetzt, ist es möglich, lediglich Synchronisationsdaten (Timestamp und Referenz) aufzuzeichnen, und für die Erzeugung der Repräsentation das Originalmaterial zu nutzen.

Für allgemeine Anwendungen (oder wenn z. B. auf Tagungen ein Zugriff auf die Originaldaten bzw. den Rechner des Vortragenden nicht möglich ist) bietet sich die Verwendung von generischen Tools an, die neben den Mediendaten bereits direkt Metadaten in dem in Abschnitt 4.4.3 beschriebenen Format erzeugen. Weiterhin ist ein Werkzeug für die manuelle Aufzeichnung von Notizen oder Gliederungsinformationen oder auch Anmerkungen zur Produktion hilfreich.

Somit ergeben sich zwei Klassen von Capturing-Modulen.

Capturing-Modulklassen

- Mediendaten-Aufzeichnung von Medien wie z. B. für AV oder Einzelbilder. VACE-konforme Module erzeugen parallel dazu Metadaten, die Informationen zur Synchronisation und zum Inhalt enthalten.
- Reine Metadaten-Aufzeichnung bei referenzierbaren Medien, die als Original weiterverwendet werden können, z. B. URLs von Webseiten oder auch Notizen (Annotationen), die ausschließlich aus Metadaten bestehen.

Realisierung von Capturing-Modulen

Für die Realisierung der Module ergeben sich abhängig von der Anwendung drei verschiedene Typen:

- Einbindung existierender Tools zur AV-Aufzeichnung (z. B. AV-Realtime-Encoding Software) mittels Steuerung dieser Systeme. Für die Integration in das VACE-Framework sorgt eine Wrapper-Applikation, die einen Standard-Encoder steuert und die zusätzlich benötigten Metainformationen (wie etwa den Startzeitpunkt) erzeugt.
- Einbindung existierender Tools zur Live-Präsentation (z. B. Web-Browser, Präsentations-Tools wie PowerPoint) mittels Steuerung oder Überwachung dieser Applikationen. So werden etwa für die Aufzeichnung von Folien lediglich Synchronisationsinformationen aufgezeichnet, mit denen die Zuordnung vom Zeitpunkt zur Foliennummer ermöglicht wird.
- Generische VACE-Tools, z. B. Aufzeichnung von Standbildern (Image-Capturing) für Projektionen, die unabhängig von der eingesetzten Applikation arbeiten. Diese Eigenentwicklungen generieren per se VACE-konforme Metadaten sowie Standard-Metadaten. Ein Sonderfall ist das Annotations-Werkzeug, das die Eingabe von Notizen und Gliederungsinformationen in Echtzeit ermöglicht und damit lediglich Metadaten erzeugt.

Innerhalb dieser Realisierungstypen werden unter Verwendung der in Abschnitt 2.3.2 beschriebenen Mechanismen individuelle Module für die unterschiedlichen Basis-Tools bzw. Anwendungen implementiert. Auf die Einbindung von AV-Hardware (Abschnitt 2.3.1) soll an dieser Stelle nicht detailliert eingegangen werden, sie ist aber prinzipiell realisierbar. Metadaten sind dabei auf dem Aufzeichnungsmedium oder mit Hilfe eines steuernden Rechners zu erfassen.

Capturing-System

Aufgrund der Abhängigkeit von der realen Uhrzeit müssen alle Aufzeichnungssysteme über synchronisierte Uhren verfügen, was im Rechnernetz mittels NTP-Dienstprogrammen (siehe Abschnitt 2.4.2) leicht realisierbar ist.

Für eine automatisierte Produktion verfügen VACE-konforme Capturing-Module neben der impliziten VACE-Metadaten-Erzeugung auch über die Möglichkeit, vorbereitete Aufträge in Form von Metadaten einzulesen und dementsprechend eine Aufzeichnung durchzuführen. Auf diese Weise können Aufzeichnungen geplant und analog einer Videorekorder-Programmierung durchgeführt werden. Zudem können AV-Dateien schon bei der Aufzeichnung mit vorliegenden Metainformationen (Titel, Autor,...) versehen werden.

4.2.3 Nachbearbeitung

Der nächste Schritt nach der Aufzeichnung besteht in der Nachbearbeitung; der Aufbereitung von Medien- und Metadaten für die Publikation der Repräsentation. Hierzu kann ein automatisierbares Post-Processing ausreichen, es können jedoch auch weiter gehende, auch manuelle Schritte zur Post-Production vorgenommen werden. Dieser Aufwand ist abhängig von der Ergiebigkeit der bereits bei der Auf-

zeichnung akquirierbaren Metainformationen und natürlich dem Anspruch an eine optimale Umsetzung.

Metadaten-Extraktion (optional)

Neben den bereits bei der Aufzeichnung erfassten Metadaten und der manuellen Ergänzung können zusätzliche Metainformationen aus den aufgezeichneten Mediendaten extrahiert werden. Hierzu gehört z. B. die Extraktion von Stichworten aus Folien. Möglich sind grundsätzlich alle geeigneten Methoden zur Text-, Bild- oder Spracherkennung.

Diese optionalen Schritte können auch bereits vor der Zusammenführung der einzelnen Medien in unabhängigen Modulen realisiert werden; hierbei werden jeweils die vorliegenden Metadaten ergänzt bzw. aktualisiert. Möglich ist auch die Generierung zusätzlicher Medienströme, z. B. in Form von Annotationen.

Medien-Verknüpfung

Beim Merging der aufgezeichneten Medienströme werden die bei der Aufzeichnung generierten Metadaten zusammengeführt. Damit stehen die benötigten Daten zur Generierung von Repräsentationen zur Verfügung.

So kann z. B. die Aufteilung einer AV-Aufzeichnung anhand von Zeitmarken der Folienwechsel gesteuert werden, eine HTML/SMIL-basierende Präsentation aus den vorliegenden Daten exportiert werden oder der Transfer von Daten z. B. auf einen Streaming-Server automatisiert werden. Standard-Abläufe können mit Hilfe von Vorlagen vereinfacht werden.

Manuelles Editing

VACE ist primär für die schnelle, standardisierte und weit gehend automatisierte Produktion von Präsentationsaufzeichnungen konzipiert. Dennoch sind in vielen Fällen manuelle Eingriffe sinnvoll und müssen somit durch die Realisierung ermöglicht werden. Hierzu gehören die folgenden Punkte.

- *Ergänzungen, Änderungen an Metadaten:* Das Editing der inhaltsbezogenen Metadaten hat i. d. R. den größten Anteil an der manuellen Nachbearbeitung. Kleinere Änderungen bei automatisch extrahierten Metadaten wie Titeln, Stichworten usw. können per Hand sehr schnell und intuitiv durchgeführt werden. Dazu kommt die Erstellung einer Gliederung bei Rohdaten, die keine derartigen Informationen enthalten, etwa reine Videoaufzeichnungen.
- *Modifikationen an Timestamps:* Hierzu gehören Korrekturen wie etwa das Verschieben oder Entfernen einer Folie, die unbeabsichtigt, zu früh oder zu spät während der Live-Präsentation gezeigt wurde.
- *Definition der Präsentations-Timeline:* Durch das Festlegen von Start- und Endpunkt werden für die Repräsentation irrelevante oder unerwünschte Überhänge am Anfang und Ende der Aufzeichnung entfernt. Dies ist eine Standard-Anforderung (Trimmen) an AV-Schnittsysteme, die hier ohne aufwendige Zuhilfenahme von externen Programmen ermöglicht wird.
- *Editing (Schnitt, klassische Nachbearbeitung):* Funktionen, die über die eben beschriebenen hinausgehen befinden sich nicht im Fokus des VACE-Frame-

works. Insbesondere für die stark interaktiv geprägte Videobearbeitung gibt es eine ausreichende Anzahl an Software, die hier nicht repliziert werden soll (siehe Abschnitt 2.2.3). Dennoch gibt es speziell bei außergewöhnlichen Veranstaltungen den Bedarf der individuellen Bearbeitung der von VACE mit Hilfe von geeignetem Metadaten-Export unterstützt wird und somit auch die klassische Videobearbeitung erleichtert.

Medien-Konvertierung (optional)

Die Konvertierung von Mediendaten in andere Formate kann für unterschiedliche Anwendungen erforderlich sein. Die benötigten Medien-Repräsentationen können im Regelfall im Batch-Betrieb mit der Ansteuerung externer Konverter-Software erzeugt werden.

- *Erstellung von alternativen Formaten.* Abhängig von der gewünschten Anwendungen und Verteilung werden Repräsentationen mit individuellen Parametern bezüglich Auflösung, Datenrate, Format, Plattform usw. benötigt (vgl. Abschnitt 2.1.4). So kann etwa eine qualitativ hochwertige Version auf CD/DVD verteilt werden, während eine qualitätsreduzierte Version im Internet zur Verfügung gestellt werden kann.
- *Erzeugung alternativer Darstellungen.* Auch innerhalb einer Repräsentationsversion können unterschiedliche Versionen einer Mediendatei benötigt werden. Typisches Beispiel ist die Erstellung von Thumbnails, der verkleinerten Darstellung von Bildern (z. B. Folien) auf einer Übersichtsseite zur einfachen visuellen Wiedererkennung.
- *Aufsplitten von Mediendaten.* Eine mögliche Anwendung ist hierbei die Verteilung einer Video-Aufzeichnung auf einzelne thematisch assoziierte Webseiten; über einen Link kann so der Betrachter auf Wunsch die audio-visuelle Erläuterung abrufen.
- *Zusammenführung unterschiedlicher Medienströme.* Zur Verringerung der Komplexität der Repräsentation oder zur Anpassung an die Möglichkeiten eines Übertragungsmediums werden separat aufgezeichnete Medienströme in ein Medium gemischt. So können z. B. für eine reine Video-Repräsentation Annotationen als Untertitel in das Videobild eingeblendet werden.

Navigation generieren (optional)

Ein essenzieller Bestandteil der Repräsentation und Mehrwert im Vergleich zu den Rohdaten ist die Navigation (Index, Suche etc.). Erzeugt wird sie mit Hilfe der inhaltsbezogenen Metadaten sowie der aufgezeichneten Zeitstempel. Festgelegt werden muss hierzu, welche Informationen zur Erstellung herangezogen werden sollen, d. h. aus welchen Streams welche Events in den Index aufgenommen werden sollen. So können z. B. die Gliederungsinformationen von Annotationsdaten oder das Zeitraster der Folienwechsel benutzt werden.

Grundsätzlich bestehen verschiedene Möglichkeiten zu Realisierung einer Navigation abhängig von der weiteren Nutzung der Repräsentation bzw. der Publikationsform.

- Werden die Daten in einem *Standard-System* publiziert, so können in geeigneten Medienformaten (HTML, SMIL) Navigationselemente ergänzt werden. Hierzu gehören: Inhaltsverzeichnis, Suchindex, Folienübersicht (Thumbnails).
- Sollen die Daten der Repräsentation in ein definiertes *Publikationssystem* (CMS, Datenbank o. ä.) integriert werden, ist die Herstellung von Navigationselementen nicht Bestandteil der Nachbearbeitung. Statt dessen sind die Navigationsdaten in einem entsprechenden Format zu exportieren; die Nutzeroberfläche ist dann Bestandteil des CMS.

Metadaten-Export (optional)

Neben dem Export von Metadaten für die Indexierung des aufgezeichneten Materials in Datenbanken können in geeigneter Form exportierte Metadaten auch für weitere Anwendungen genutzt werden. Dazu gehört z. B. Erzeugung von Schnittlisten (EDLs) für Video-Editing-Systeme oder das DVD-Authoring.

Um für diese Vielzahl potenzieller Anwendungen keine einzelnen Schnittstellen implementieren zu müssen, bietet sich hier der Export in einem generischen Format an, aus dem dann die relevanten Datenstrukturen extrahiert werden. Neben XML kommen textbasierte Formate wie CSV zum Einsatz.

4.2.4 Publikation

Die Funktionen für die Publikation runden die Produktionskette ab. Die letzten Schritte der Verteilung sind nicht Bestandteil von VACE; hier wird auf Standard-Systeme (primär Web-basierte Plattformen) aufgesetzt.

Ein Teil der Aufbereitung des Materials für die Veröffentlichung im Web oder auf Datenträgern wurde bereits im Abschnitt Nachbearbeitung erörtert, da ein Teil der Arbeitsschritte nicht eindeutig zuzuordnen ist.

Medientransport

Web-basierte Medien können – sofern für die Referenzierung der einzelnen Elemente relative Verzeichnis-Pfade verwendet werden – sowohl auf Datenträgern als auch zwischen Web-Servern und innerhalb der Verzeichnisstruktur eines Servers verschoben werden, ohne dass die Funktion beeinträchtigt wird.

Wenn aus verschiedenen Gründen mit absoluten Pfaden gearbeitet wurde, müssen diese bei einem „Umzug“ angepasst werden. Hierbei werden die vorhandenen Pfadangaben an aktuelle (absolute) Pfade angepasst bzw. durch relative ersetzt. Dieser Vorgang ist in den meisten Fällen automatisierbar.

Einen Sonderfall stellt die Modifikation des Transportmediums bzw. -protokolls dar, der üblicherweise bei der Portierung von Mediendaten auf einen Streaming-Server auftritt. Während bei dem Transport auf einem Datenträger der Zugriff grundsätzlich dateibasiert erfolgt, bestehen für die Verteilung im Netz verschiedene Optionen. Neben der Übertragung mittels HTTP bieten sich speziell bei längeren Videos effizienter arbeitende Streaming-Protokolle an (vgl. Abschnitt 2.1.4). Hierfür sind zum einen die Daten entsprechend zu verschieben und zum anderen Links innerhalb der Präsentation anzupassen.

Präsentationsformat

Wenn die Publikationsplattform nicht durch ein CMS vorgegeben ist oder die aufgezzeichneten Präsentationen unabhängig davon publiziert werden sollen, kann die Gestaltung mit Hilfe des VACE-Systems definiert werden.

Wie im Abschnitt 2.2.4 erläutert, kommen aufgrund der flexiblen Einsatzmöglichkeiten Web-basierten Formaten eine besondere Bedeutung zu. Für die hier relevanten MM-Präsentationen können generische Framesets eingesetzt werden, die mittels dynamischem HTML Navigation, Index und Suche bereitstellen sowie die audiovisuellen Medien einbinden. Zur gestalterischen Anpassung von HTML-Seiten bietet sich die Verwendung von CSSs (cascading style sheets [80]) an. Damit können Design-Elemente zentral gestaltet und modifiziert werden und so an individuelle Layouts angepasst werden.

Neben der Verwendung von HTML-Techniken kann auch das (in Teilbereichen verwandte) SMIL-Format eingesetzt werden. Die Gestaltungsmöglichkeiten, die aktuelle Implementierungen zur Verfügung stellen, sind jedoch im Vergleich zu HTML eingeschränkt (vgl. Abschnitt 2.2.4).

4.3 Präsentationsmodell

Objekt der Produktion ist die Live-Präsentation, die mittels eines Datenmodells repräsentiert (abgebildet) werden muss.

4.3.1 Aspekte einer Präsentation aus Aufzeichnungssicht

Als Vorbereitung zum Entwurf eines Datenmodells zur Abbildung von Präsentationen wird zunächst ein abstraktes Präsentationsmodell vorgestellt, das die hier relevanten Aspekte in einer für die Anwendungsszenarios (vgl. Abschnitt 3.2) geeigneten Weise widerspiegelt.

Eine Präsentation (Vortrag, Vorlesung usw.) besteht aus Aufzeichnungssicht aus n ($n \geq 1$) zeitabhängigen Medien, mit denen parallel derselbe Inhalt vermittelt wird. Auch der Vortragende selbst wird in diesem Modell als Medium betrachtet, das mittels geeigneter Aufzeichnung repräsentiert wird.

Die bei der Präsentation bzw. deren Aufzeichnung differenzierten Medien finden sich in vielen Fällen in der später publizierten Repräsentation unmittelbar wieder. Es können jedoch auch bei der Nachbearbeitung z. B. einzelne Medien zusammengeführt werden. Denkbar ist etwa die Einblendung von separat erfassten Annotationen in ein Videobild.

4.3.2 Medienklassen

Das Präsentationsmodell unterscheidet zwei Klassen von Medientypen (hier mit Medienströmen bzw. *Streams* bezeichnet) anhand ihres Zeitverhaltens. In Abbildung 4-2 sind beispielhaft zwei Streams illustriert. Veränderungen mit zeitlichem Bezug sind als Ereignis (*Event*) gekennzeichnet.

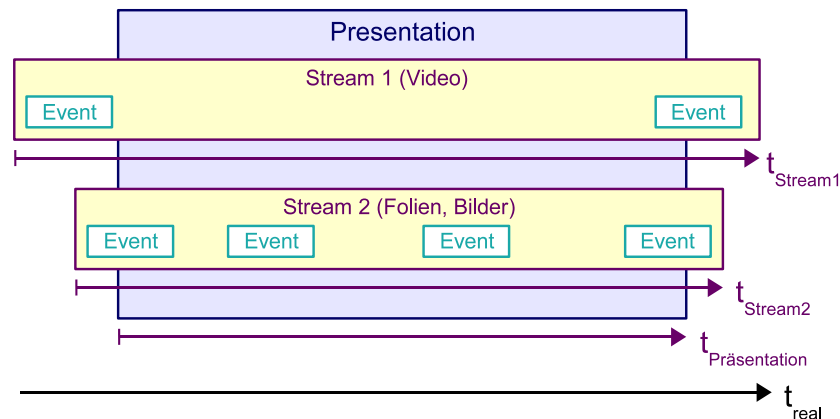


Abbildung 4-2: Präsentationsmodell

Zeitkontinuierlich

Als zeitkontinuierlich werden zeitlich veränderliche Medien mit inhärentem Zeitverhalten angesehen, z. B. die AV-Aufzeichnung des Vortrags. Für die Erfassung der relevanten Zeitinformationen genügt der Start- und Endzeitpunkt der Aufzeichnung; dazwischen wird ein lineares Verhalten des Mediums vorausgesetzt. In vielen Fällen kommt dabei ein äquidistantes Zeitraster zum Einsatz (z. B. 25 Bilder pro Sekunde). Start und Ende entsprechen je einem Event.

Zeitdiskret

Zeitdiskrete Medien weisen kein eigenes Zeitverhalten auf. Ein zeitlicher Bezug ergibt sich erst aus dem Zeitpunkt der Darstellung während der Präsentation. Typische Beispiele hierfür sind (PowerPoint-) Folien, Dias oder Annotationen. Die einzelnen (statischen) Medienelemente erhalten so eine zeitliche Abfolge ohne inhärente Zeitbasis. Somit entspricht jede Änderung (z. B. ein Folienwechsel) einem Event.

4.3.3 Zeitbasen der Medien

In Abbildung 4-2 sind die verschiedenen Zeitbasen (Timelines) der eingesetzten Medien veranschaulicht. Ausgehend vom linearen Zeitverhalten ist zur Synchronisation eine gemeinsame Basis bzw. ein gemeinsamer Bezugspunkt erforderlich.

Neben den Timelines der bei der Aufzeichnung eingesetzten Medien ergibt sich bei der Herstellung der Repräsentation eine weitere, frei definierbare Zeitachse. Im Vergleich zu den mehr oder weniger willkürlich entstehenden Medien-Timelines bietet sich für die Produktion als eindeutiger und natürlicher Bezugspunkt die reale Uhrzeit an.

4.4 Metadatenmodell

In diesem Abschnitt wird das dem VACE-Framework zugrunde liegende Metadatenmodell zur Beschreibung des im letzten Abschnitt erläuterten Präsentationsmodells unter Berücksichtigung der in Abschnitt 4.1.3 aufgeführten Anforderungen entwickelt.

Das Metadatenmodell realisiert eine Abbildung der benötigten Informationen für die in Tabelle 4-1 aufgeführten Anwendungsbereiche.

Bereich	Funktion
Inhaltliche Beschreibung	Globale Beschreibung der Präsentation
	Beschreibung einzelner Präsentations-Elemente
Produktion	Synchronisation der Medien
	Ablaufsteuerung der Produktion

Tabelle 4-1: Metadaten – Funktionen

Wie im Abschnitt 2.5 erläutert, werden für die effiziente Aufbereitung der Mediendaten zusätzliche Informationen in Form von Metadaten benötigt. Dabei wird zwischen Metadaten im herkömmlichen Sinn (inhaltsbezogen) und denen zur Erfassung des zeitlichen Ablaufs (Synchronisationsmetadaten) unterschieden [10].

Beide Metadatenarten werden dabei eng miteinander gekoppelt. Der entscheidende Vorteil der Zusammenführung von Inhalts- und Synchronisations-Metadaten besteht in der Nutzung für die Aufbereitung der Repräsentation. Da die Metadaten paralleler Medien-Aufzeichnungen denselben Inhalt beschreiben, können z. B. die leicht zu generierenden Metadaten einer Folie zum Auffinden der entsprechenden Position in der AV-Aufzeichnung dienen, an der vom Vortragenden der gesuchte Inhalt erläutert wird. Umgekehrt können bei linearer (zeitkontinuierlicher) Betrachtung der Video-Aufzeichnung auch die entsprechenden Folien referenziert werden.

Die Assoziation der Metadaten von Inhalt und Zeit ermöglicht so die Generierung einer effizienten Navigation in der Repräsentation. Dazu gehört die Indexierung, Suche und Integration in Datenbanken bzw. Content-Management-Systemen.

4.4.1 Anforderungen an das Metadatenmodell

Globale inhaltsbezogene Metadaten

Globale inhaltsbezogene Metadaten beschreiben die gesamte Präsentation. Dies entspricht den bekannten Metadatenmechanismen, in denen *einer* Ressource Daten wie Autor, Titel etc. zugeordnet werden. Diese Daten dienen der Einordnung bzw. Auffindbarkeit in Datenbanken oder im WWW und werden analog zu denen anderer Ressourcen wie z. B. einem HTML-Dokument eingesetzt.

Zeitbezogene inhaltsbezogene Metadaten

Ähnlich wie die beschriebenen globalen inhaltsbezogenen Metadaten werden Daten eingesetzt, um die *Fragmente der Präsentation* (repräsentiert durch Events) inhaltlich zu beschreiben. Darin sind z. B. Stichworte zu einer Folie enthalten.

Die relevanten Metadaten-Elemente können dabei der jeweiligen Anwendung angepasst bzw. erweitert werden. Typische Elemente sind auch hier z. B. der Titel oder Stichworte.

Synchronisations-Metadaten

Mit Hilfe der Synchronisations-Metadaten können die verschiedenen, unabhängig aufgezeichneten Mediendaten bei der Postproduktion synchronisiert werden.

Bei der Aufzeichnung sind die Zeitpunkte der Änderungen der eingesetzten Medien zu erfassen. Bei zeitkontinuierlichen Medien wie z. B. einer Videoaufzeichnung genügt – bei Vernachlässigung des Jitters – die Erfassung des Startzeitpunkts der Aufzeichnung sowie die Länge der Aufzeichnung bzw. der Endzeitpunkt. Werden aufgrund von Restriktionen wie etwa der Bandlänge mehrere Aufzeichnungen des gleichen Medienstroms angefertigt, müssen neben der eindeutigen Kennung der Teile die jeweiligen Startzeiten erfasst werden. Bei Medienströmen mit zeitdiskreten Änderungen wie einer Folge von Bildschirmfotos sind auch hier jeweils Zeitpunkt und Inhalt (bzw. eine entsprechende Referenz) aufzuzeichnen.

In einer später erzeugten Repräsentation hingegen hat die ursprüngliche Uhrzeit nur eine geringe Relevanz (als inhaltsbezogenes Metadatum), da für die generierte Repräsentation eine vom Autor festgelegte Timeline gilt, die für den Betrachter bei Null beginnt (vgl. die Darstellung der unterschiedlichen Timelines in Abbildung 4-3). In AV-Aufzeichnungen können wiederum eigene Zeitbasen (TCR) gelten, wobei mit PCs aufgezeichnete Dateien üblicherweise bei $t=0$ beginnen.

Das VACE-System arbeitet bei der Aufzeichnung mit der realen Zeit. Für die Synchronisation z. B. von Audio/Video und Folien ist dabei eine Auflösung im Bereich von 1/100 bis 1/10 Sekunden ausreichend. Audio und Video wird praxisgerecht als ein Medium und damit implizit synchronisiert betrachtet.

Unüblich ist hierbei die Kombination von Echtzeit-Angaben mit der für audiovisuellen Medien benötigten Auflösung. Während für inhaltsbezogene Metadaten in den meisten Anwendungsfällen die Spezifikation von Datum bzw. Uhrzeit im Minutenbereich ausreicht, verwenden typische Applikationen zur Synchronisation eine eigene, von der realen Uhrzeit losgelöste Zeitbasis.

Als Format für Zeit und Datum kann das in Abschnitt 2.4.1 erläuterte Standardformat genutzt werden. Jedoch ist für VACE bei der Aufzeichnung einer lokalen Präsentation die Zeitzone von untergeordneter Bedeutung. Um eine leichte alphanumerische Sortierung zu ermöglichen, wurde die Verwendung einer festen Stellenanzahl mit einer Auflösung von einer Millisekunde gewählt (z. B. „2004-11-22 11:29:05.070“).

Produktionsmetadaten

Da der Ablauf der Produktionskette komplett auf das Datenmodell abgestellt ist, sind auch für diesen Zweck Metadaten-Elemente vorzusehen. Hierbei handelt es sich um Hilfsdaten, die nicht der Beschreibung der Präsentation dienen, sondern lediglich die Automatisierung der Produktion erleichtern.

Hierzu gehören Daten zur Beschreibung von eingesetzten Medienformaten sowie Informationen zu verwendeten Dateibasen, etwa das Verzeichnis, in dem die Mediendaten für den nächsten Produktionsschritt abgelegt werden sollen.

Damit können optional Aufzeichnungen vorbereitet werden, indem bereits Metadaten-elemente wie Medienformate und deren Parameter sowie die Zeitpunkte der geplanten Aufzeichnungen vorgegeben werden. Die so erzeugten Metadaten können an die separat arbeitenden Aufzeichnungssysteme übergeben werden, bei der Aufzeichnung aktualisiert und ergänzt werden und anschließend an die Nachbearbeitung übergeben werden.

4.4.2 Vergleich zu existierenden Metadatenformaten

Im Abschnitt 2.5 wurden verfügbare Metadatenmodelle vorgestellt, die hier im Hinblick auf das Design des VACE-Metadatenmodells einzuordnen sind.

Inhaltsbezogene Metadaten

Mit DC bzw. LOM existieren speziell im Bereich wissenschaftlicher Veröffentlichungen und der Lehre etablierte Metadaten-Modelle, die in diversen Systemen im Bereich E-Learning eingesetzt werden (vgl. Abschnitt 2.5.1). Dabei werden in den meisten Fällen anwendungsspezifische Besonderheiten berücksichtigt, indem z. B. Subsets der verfügbaren Elemente eingesetzt werden.

Synchronisations Metadaten

Im Abschnitt 2.5.2 wurde SMIL als ein Zieldatenformat für Multimedia-Präsentationen identifiziert. Es liegt somit nahe, bereits bei der Aufzeichnung an SMIL angelegte Datenstrukturen zu nutzen. Aufgrund der unterschiedlichen Anwendungsbereiche *Aufnahme/Produktion* (VACE-Capturing) sowie *Wiedergabe* (SMIL) fehlen in SMIL jedoch Elemente für die Erfassung der Produktionsmetadaten. Während für VACE Datenstrukturen zur Erfassung einer realen Präsentation benötigt wird, stellt SMIL ein potenziell nutzbares Datenformat für das Endprodukt dar. Dies gilt auch für die im Annodex-Projekt entwickelten Ansätze, wobei das Annodex-Containerformat im Gegensatz zu SMIL Medien *und* Metadaten der Repräsentation aufnehmen kann.

Der Fokus der MPEG-7-Standards liegt in der Beschreibung von Bildelementen, die über das Ziel der hier entwickelten Anwendung hinausgehen bzw. nicht relevant sind. Zur Vermeidung des Overheads und aufgrund der divergierenden Schwerpunkte wurde MPEG-7 für diese Arbeit nicht eingesetzt.

Ergebnis

Für das Design des VACE-Metadatenmodell wird ein Ansatz gewählt, der Kernelemente von Dublin Core enthält und bei Bedarf beliebig anwendungsspezifisch erweiterbar ist. Die Metadaten zur Synchronisation werden an SMIL angelehnt und dort ergänzt, wo Elemente zur Beschreibung der Aufzeichnung und Produktion (die in SMIL nicht berücksichtigt sind) erforderlich werden.

Das Modell ist bewusst überschaubar aber erweiterbar konzipiert. Da die VACE-Metadaten in der Produktion eingesetzt werden – und damit kein Endbenutzerformat darstellen – wird zugunsten der einfacheren Implementierung auf die Erweiterung eines existierenden Modells verzichtet. Kein verfügbares Modell erfüllt die hier vorliegenden Anforderungen an den Funktionsumfang zusammen mit einer einfachen Handhabung. Schnittstellen zu existierenden Systemen können aufgrund

der Anlehnung an standardisierte Formate (DC, SMIL) sowie des kompakten VACE-Modells mittels einfach zu implementierender Import- und Export-Funktionen hergestellt werden. Auch die Erweiterung auf Klassifikationsschemata wie das ACM Classification Scheme [1] kann sinnvoll sein, wenngleich derzeit z. B. kein bekanntes Präsentationsprogramm entsprechende Metadaten unterstützt; die Eingabe müsste daher manuell erfolgen.

4.4.3 Hierarchisches Metadatenmodell und -format

Modell

Das für VACE entwickelte Metadatenmodell ermöglicht eine unabhängige Aufzeichnung auf synchronisierten Systemen und enthält ferner die Basisdaten für die Post-Produktion und Publikation.

Das Metadatenmodell basiert auf dem in Abschnitt 4.3 vorgestellten Präsentationsmodell, das bei der Live-Präsentation (bzw. deren Aufzeichnung) verwendete Medien in zwei Klassen von Datenströmen aufteilt: zeitkontinuierliche (z. B. die AV-Aufzeichnung des Vortragenden) und zeitdiskrete (z. B. Präsentationsfolien). Eine Präsentation (Presentation) enthält eine beliebige Anzahl von Streams, die wiederum verschiedene Ereignisse (Events) enthalten. Die Ereignistypen sind abhängig vom Stream-Typ; so enthalten z. B. zeitkontinuierliche Videoaufzeichnungen Start- und Stop-Events. Aus diesen Elementen ergibt sich das in Abbildung 4-3 skizzierte hierarchische Metadatenmodell.

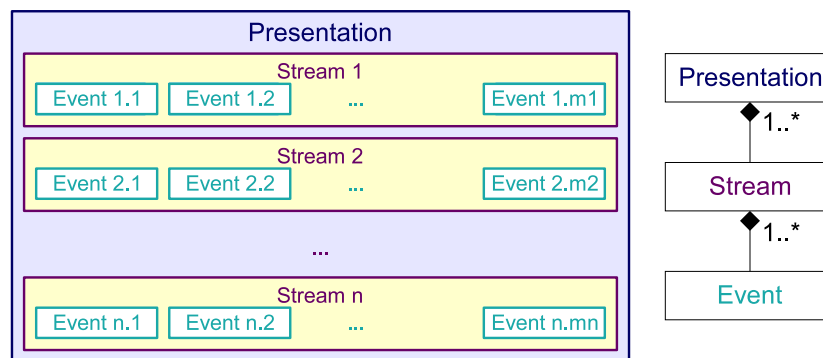


Abbildung 4-3: Hierarchisches Metadatenmodell zur Repräsentationsbeschreibung

- *Presentation* (Präsentation). Neben globalen, inhaltsbezogenen Metadaten aus DC – die für die gesamte Aufzeichnung gelten – enthält die Präsentation n_s ($n_s \geq 1$) Medienströme (Streams).
- *Stream* (Medienstrom). Ein Stream stellt die Repräsentation *eines* Medientyps dar. So ist etwa die AV-Aufzeichnung des Vortragenden und die Folge von Folien jeweils ein Stream. Werden zwei Videokassetten nacheinander für die Aufzeichnung des Vortragenden genutzt, so zählen diese als ein Stream. Die parallele Aufzeichnung eines Versuchsaufbaus bildet hingegen einen weiteren Stream. Ein Stream verfügt über n_e ($n_e \geq 1$) Events.
- *Event* (zeitlich definiertes Ereignis). Ein Event stellt den Zusammenhang zwischen Mediendaten und der Zeit her („*was geschah wann*“, z. B. der Startzeit-

punkt einer Videoaufzeichnung). Abhängig vom Streamtyp sind unterschiedliche Event-Typen definiert.

Format

Für die Elemente des Modells sind jeweils Datenstrukturen zu definieren. Im Folgenden sind sie und die benötigten Subelemente aufgeführt. Als Datenformat wird exemplarisch XML verwendet; zu jedem Element ist das entsprechende XML-Tag angegeben. Da es sich bei den Daten im Wesentlichen um lineare Listen handelt, sind auch andere Datenformate denkbar, die sich z. B. besser zur Nutzung in Streaming-Systemen eignen.

In der Spalte „Eigenschaft“ ist neben dem Werteformat angegeben, ob das Element optional verwendet werden kann oder zwingend angegeben sein muss.

Aus dem Metadatenmodell wurde ein XML-basierendes Datenformat abgeleitet, das für die Beschreibung der Inhalte auf Dublin-Core-Elemente zurückgreift und bei Bedarf erweiterbar ist. Zusätzlich wurden Elemente für die Synchronisation eingeführt.

Für die Datums- und Zeitangaben wird das vom W3C definierte *dateTime* [81] mit einer festen Dezimalstellenzahl und führenden Nullen verwendet (YYYY-MM-DDThh:mm:ss.xyz mit xyz in ms, z. B. 2004-11-22T12:30:03.969). Für die Zeitdifferenz (Dauer) wird *timeDuration* genutzt.

Presentation

Die *Presentation* (Tabelle 4-2) bildet den Container für die Metadaten einer Aufzeichnung. Neben den Stream-Informationen sind hier einige globale Informationen untergebracht, die mit anderen DC-Elementen erweitert werden können.

Element	Name-space	Beschreibung	Eigenschaft	Auftreten
title	DC	Titel der Präsentation	Text	zwingend
creator	DC	Autor, Urheber	Text	optional
rights	DC	Copyright Informationen	Text	optional
description	DC	Beschreibung	Text	optional
subject	DC	Thema/Stichworte, z. B. für Suche	Text	optional
starttime	VACE	Startpunkt der Repräsentation	dateTime	zwingend
endtime	VACE	Endzeitpunkt bzw. Länge	dateTime	zwingend

Tabelle 4-2: Metadatenelemente von „presentation“

Mit dem Start- und Endzeitpunkt wird die Präsentations-Timeline bestimmt. Alle anderen Elemente enthalten Textzeilen (ohne Zeilenumbrüche).

Stream

Ein *Stream* (Tabelle 4-3) repräsentiert jeweils einen Medienstrom und stellt damit einen Container für die entsprechenden Informationen zur Verfügung. Unterschiedliche Medientypen werden mittels MIME-Type wie z. B. *video/mpeg1* definiert. Eine Ausnahme bildet dabei der Typ *annotation*, der keine weiteren Mediendateien benötigt; die Informationen werden direkt in den Metadaten gespeichert. Die IANA-Liste der MIME-Types in [34] enthält nicht alle benötigten Typen, es fehlt z. B. *application/vnd.rn-realmedia* für RealMedia-Dateien.

Element	Name-space	Beschreibung	Eigenschaft	Auftreten
id	VACE	eindeutige Bezeichnung	Text	zwingend
format	DC	Medientyp (Audio, Video, AV, URLs, (Einzel-)Bilder, Annotationen)	MIME Type/ Subtype	zwingend
timeoffset	VACE	Korrekturwert: $t_{\text{stream}} = t_{\text{real}} - t_{\text{offset}}$	dateTime	zwingend
creator	DC	Autor, Urheber	Text	optional
rights	DC	Copyright-Informationen	Text	optional
description	DC	Beschreibung, Schlüsselworte	Text	optional
subject	DC	Thema/Stichworte	Text	optional
filebase	VACE	Verzeichnis der Mediendaten, gemeinsame Basis im Dateisystem	Text	optional
filepattern	VACE	Maske der Dateinamen, z. B. „capture*.png“	Text	optional
bitrate	VACE	Bitrate des Mediendatenstroms	Number	optional
width	VACE	Bildbreite in Pixel	Number	optional
height	VACE	Bildhöhe in Pixel	Number	optional
thumbsdir	VACE	Verzeichnis für Thumbnails	Text	optional

Tabelle 4-3: Metadatenelemente von „stream“

Jeder Stream verfügt über einen – innerhalb der Präsentation – eindeutigen Namen (*id*) sowie einen *timeoffset*, der bei kontinuierlichen AV-Aufzeichnungen dem (ersten) Startzeitpunkt entspricht. Dieser Wert ermöglicht auch die Anzeige der Stream-basierten Zeit im Editor.

Die *filebase* dient der konsistenten Datenspeicherung und gibt an, wo die Dateien im Format *filepattern* für den nächsten Produktionsschritt (z. B. für eine Nachbearbeitung) abgelegt werden.

Event

Zentrales Element eines *Events* (Tabelle 4-4) ist der Zeitpunkt des Ereignisses. Das Zeitfenster einer Präsentation bewegt sich üblicherweise im Bereich von wenigen Minuten bis zu einigen Stunden. Die benötigte Auflösung für diese Anwendung

liegt im Bereich mehrerer Millisekunden – die oft verwendete Sekundenauflösung ist nicht ausreichend.

Element	Name-space	Beschreibung	Eigenschaft	Auftreten
type	VACE	Art des Events	eventtype	zwingend
time	DC	Zeitpunkt des Ereignisses in Realzeit	dateTime	zwingend
title	DC	Titel	Text	optional
description	DC	Textliche Beschreibung des Events	Text	optional
subject	DC	Thema/Stichworte, z. B. für die Suche	Text	optional
file	VACE	Dateiname oder URL der Mediendatei	Text	optional
timebase	VACE	Offset des Startzeitpunkt in der Mediendatei (Standard: 0)	timeDuration	optional

Tabelle 4-4: Metadatenelemente von „event“

Um eine alphanumerische Sortierung zu erleichtern, sind Datum und Zeit zusammengefasst. Die Angabe des Datums ist in den meisten Fällen nicht erforderlich, ermöglicht jedoch eine Monotonität bei Zeitfenstern, die vor 0:00 Uhr beginnen und nach 0:00 Uhr enden.

In `file` ist der Name der Medien-Datei gespeichert, z. B. die Bilddatei eines Bildschirmfotos. Der Speicherort und das Dateinamensformat ist im Stream bestimmt. Alternativ kann auch eine vollständige (absolut referenzierte) URL angegeben werden.

Bei kontinuierlichen AV-Aufzeichnungen entspricht `timebase` der (ersten) `start-time`. Damit können im Editor die `datei-` bzw. `medienbasierten` Zeitstempel angezeigt werden, um so eine leichtere Handhabung zu ermöglichen.

Event-Typen

Abhängig vom Typ des Streams gibt es eine Reihe verschiedener Eventtypes, die in Tabelle 4-5 zusammengestellt sind.

Streamtype	Eventtypes	Bemerkung
video/*	start, end	AV-Länge ergibt sich aus Differenz der Zeit-Angaben
image/*	img	analog zum HTML-Tag
text/html	href	Link auf URI
text/annotation	p, h1, h2,...	HTML-Tags (Headings, Paragraph)

Tabelle 4-5: Metadatenwerte für „eventtype“

4.5 Implementierung

Dieser Abschnitt beschreibt Aspekte der Implementierung und liefert eine Funktionsübersicht der prototypisch fertiggestellten Module.

Aufgrund der benötigten Schnittstellen zu existierenden Programmen, insbesondere zur Präsentation und Aufzeichnung, wurde für den Prototyp als Plattform Microsoft Windows und als Programmierumgebung Microsoft Visual C++ gewählt. Eine Vielzahl der in diesem Umfeld eingesetzten Software läuft primär auf Windows und bietet z. T. entsprechende Programmier-Schnittstellen, sodass eine leichte Einbindung möglich ist.

Da das VACE-Framework wesentlich durch Datenstrukturen (Meta- und Mediendaten) bestimmt wird, sind jedoch auch Aufzeichnungs- und Bearbeitungs-Module auf weiteren Plattformen und Betriebssystemen (z. B. JAVA, Linux) realisierbar.

Die grundlegenden Basisfunktionen des VACE-Frameworks wurden in zwei C++-Klassen, *CVaceDoc* und *CVaceHelper* implementiert. In *CVaceDoc* sind das VACE Metadatenformat sowie sämtliche damit verbundenen Funktionen gekapselt. Die Klasse *CVaceHelper* enthält weitere häufig benötigte Funktionen. In den einzelnen Programmen (den VACE-Modulen) werden diese Klassen eingebunden. Hiermit wird zum einen die Entwicklung neuer Module vereinfacht und zum anderen eine konsistente Nutzung des Metadatenmodells gewährleistet.

Im Prototyp wurden die folgenden Funktionen implementiert.

- Aufnahme von Bild und Ton des Vortragenden sowie universelle Aufzeichnung von Bildschirmhalten via Scan-Konverter mit Hilfe des RealProducers/Helix-Producers.
- Aufzeichnung der Präsentation von Webseiten.
- Aufzeichnung von PowerPoint-Präsentationen.
- Import von Einzelbildern im PNG-, JPEG- und GIF-Format unter Zuhilfenahme der Zeitinformationen des Dateisystems.
- Eingabe von Echtzeit-Annotationen während einer Aufzeichnung.
- Extraktion von Metadaten aus (lokalen) HTML-Dokumenten.
- Zusammenführung und Bearbeitung der VACE-Metadaten.
- Export von HTML-Präsentationen (AV, HTML oder Grafik, Index, Suche).
- Export von SMIL-Präsentationen (AV, weiteres AV oder Grafik, Index).
- Export von Listen zur Analyse (Debugging).

Mit diesen Eigenschaften ist die Produktionskette des VACE-Frameworks für viele Einsatzfälle bereits komplett nutzbar, sodass hiermit ein „proof-of-concept“ durchführbar ist. Als Format für den Export wurden solche gewählt, die zum einen weit verbreitet sind und zum anderen als Web-Browser-Plugin uneingeschränkt nutzbar sind.

Die genannten Funktionen wurden in den vier folgenden Programmen umgesetzt.

- VACE HelixProducer Wrapper
- VACE Web-Browser
- VACE PPT
- VACE Editor

Die Programme nutzen die o. g. Basisklassen. Bei den Aufzeichnungsmodulen wurden exemplarisch zwei unterschiedliche Verfahren zur Integration – ActiveX und Kommunikation via Shell – genutzt. Die erstellten Module dienen somit auch als Vorlage für weitere Komponenten.

4.5.1 Aufzeichnung

Zur Aufzeichnung von Medien und gleichzeitigen Generierung von VACE-Metadaten wurden beispielhaft drei Anwendungen implementiert. Die Anwendungen unterstützen das Laden von vorbereiteten VACE-Metadaten und die Modifikation globaler Stream-Metadaten. Zudem können Medien- und Metadaten nach der Fertigstellung zur Weiterverarbeitung an eine definierte Position im Netz oder lokalen Dateisystem kopiert oder verschoben werden.

Synchronisierte Uhren werden auf den genutzten Systemen vorausgesetzt. Hierzu kann z. B. ein NTP-Client installiert werden.

VACE RealProducer Wrapper

Zur Aufzeichnung von Audio und Video können verschiedenen Systeme eingesetzt werden. Besonders effizient ist die Verwendung von Echtzeit-Encodern, die bereits während der Präsentation das gewünschte Format erzeugen können. Aufgrund seiner Verbreitung und universellen Nutzung wurde für diesen Prototyp die freie Version des Real- bzw. HelixProducers ausgewählt. Dieser erstellt AV-Daten mit unterschiedlichen Qualitätsstufen sowohl bei der Aufzeichnung realer Videobilder als auch bei der Aufnahme von Computer-generierten Präsentationen, die mittels Scan-Konverter erfasst werden. Vereinfacht wird dies durch die Nutzung von Profilen, die auf die jeweilige Anwendung zugeschnitten sind; so lassen sich etwa aus Bildschirmpräsentationen Videos geringer Bildfrequenz zugunsten schärferer Darstellung und höherer Auflösung kodieren. Auf diese Weise können zwei Encoder parallel für die Aufzeichnung von Bewegtbild und die generische Erfassung von PC-Präsentation eingesetzt werden.

Nach der Installation liegt der Producer in zwei Versionen vor. Neben der Variante mit GUI (*Helix Producer Basic.exe*) gibt es eine weitere für die Nutzung auf der Kommandozeile (*producer.exe*). Neben weiteren Dateien steht auch der *rmeditor* zur Verfügung, der zur Modifikation und Analyse von RealMedia-Daten dient. Der implementierte *VACE Helix Producer Wrapper* (*VaceHxPW*, *Abbildung 4-4*) verwendet die beiden letztgenannten Programme. Hierzu werden die Programme vom *VaceHxPW* gestartet und deren Ausgabe ausgewertet. Eine Modifikation oder Einbindung auf Programmiersprachenebene ist nicht erforderlich. Weiterhin können neuere Versionen des Encoders genutzt werden, sofern die genutzten Eingabeparameter und Ausgaben nicht geändert wurden.

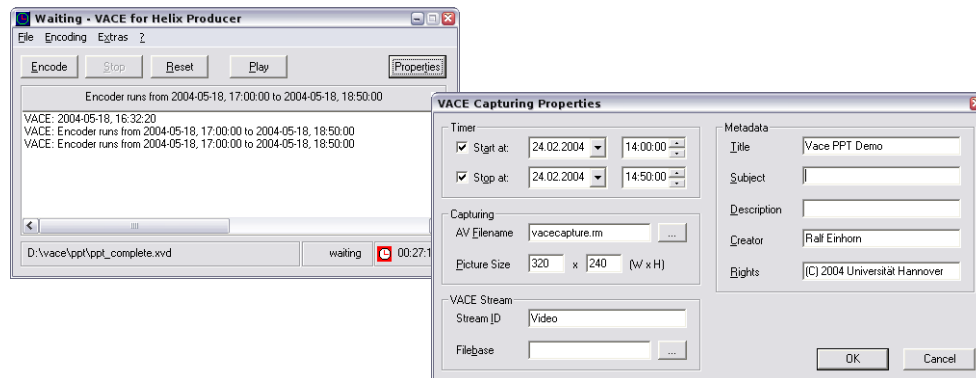


Abbildung 4-4: VaceHxPW, Applikationshauptfenster und Aufzeichnungseinstellungen

Mit Hilfe der Einträge in der Windows Registry findet VaceHxPW die benötigten Programme automatisch bzw. gibt bei fehlender Installation eine Fehlermeldung aus. Die verwendeten Aufzeichnungsgeräte (capture devices, Sound- und Video-karte) müssen einmalig konfiguriert werden.

Zur Aufzeichnung können vorbereitete VACE-Metadaten geladen werden, in denen die Informationen zu Titel, Autor usw. aber auch Bildgröße und Bitrate enthalten sind. Zudem kann mit einem Timer die Aufzeichnung zu einer vorgegebenen Zeit gestartet und beendet werden. Alternativ können die Daten manuell eingegeben werden und der Encoder von Hand gestartet und gestoppt werden.

Einige Video-Capture-Karten weisen die Eigenschaft auf, dass sie bei der ersten Verwendung nach dem Systemstart keine Bilddaten liefern. VaceHxPW enthält hierfür eine Funktion zur Initialisierung nach dem Start, sodass das System auch ohne interaktiven Eingriff arbeitet.

Beim Start des Real-Encoders kommt es zu einer Initialisierungsverzögerung Δt_B von etwa zwei Sekunden. Der Startzeitpunkt $t_{\text{Start/Real}}$ kann daher durch die steuernde Applikation zunächst nicht exakt ermittelt werden. Anders verhält es sich beim Beenden des Encoders: hier ist die Verzögerung Δt_E vernachlässigbar – der Encoder beendet die Aufzeichnung unmittelbar. Mit Hilfe des zum Real-Encoder gehörenden Programms *rmeditor* kann die Länge des erzeugten Videoclips Δt_{Datei} exakt ermittelt werden. Somit ergibt sich das implementierte Verfahren zum Auffinden des Startzeitpunktes als Differenz zwischen (hinreichend exakt bekanntem) Endzeitpunkt und der nachträglich ermittelten Länge (siehe Abbildung 4-5).

In dem Modell sind nicht die systemimmanenten Verzögerungen durch die Video-signalverarbeitung, die Encoding-Hardware und den Treiber berücksichtigt. Ferner kann nicht ermittelt werden, ob der Encoder bei der Herstellung der AV-Datei beim Start oder Ende Bilder bzw. Audiodaten verwirft. Innerhalb der erzeugten Datei treten jedoch nachweisbar keinerlei Schwankungen auf.

Die Messung der mit dem Verfahren erreichten Abweichung ist im Abschnitt 5.2.1 beschrieben. Eine noch höhere Genauigkeit könnte durch direkte Einbindung des Encoders an die Systemzeit erfassende Applikation erreicht werden. Die hierzu notwendige Integration z. B. auf Programmebene (vgl. Abschnitt 2.3.2) ist jedoch nicht mit allen Encodern realisierbar.

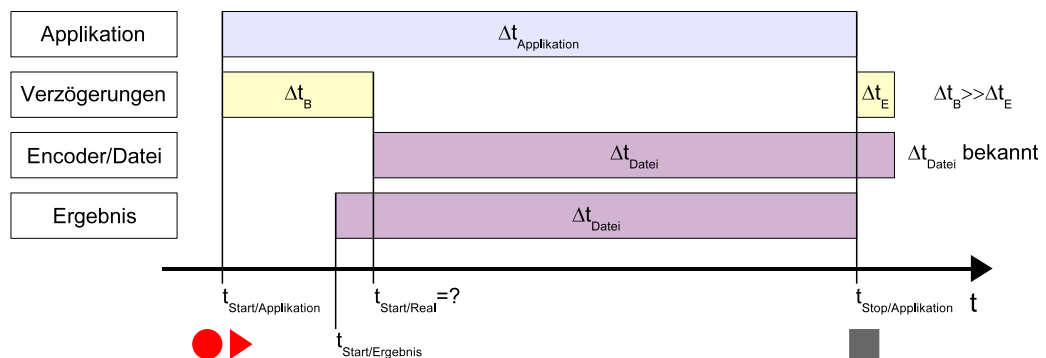


Abbildung 4-5: Zeitdiagramm zur Ermittlung des Encoding-Startzeitpunkts beim HxPW

VACE Web-Browser

Der VACE Web-Browser (VaceBrowser) dient zur Erfassung der Präsentation von lokal gespeicherten oder im Netz verfügbaren Webseiten. Als Medium für die Aufzeichnung sind Web-basierte Inhalte prädestiniert, da sowohl bei der Live-Präsentation als auch bei der Repräsentation dasselbe Medium verwendet werden kann – Zuschauer der Live-Veranstaltung und Betrachter der Aufzeichnung sehen somit identische Inhalte ohne Qualitätsverlust oder Medienbruch.

Der VaceBrowser (Abbildung 4-6) ist ein ausführbares Windows-Programm, das zur Ausführung einen installierten Internet Explorer voraussetzt. Dem Vortragenden steht ein vereinfachtes UI mit allen wesentlichen Funktionen zur Verfügung. Das Fenster kann zur Präsentation auf den gesamten Bildschirm vergrößert werden, die Menüleiste am oberen Rand wird optional ausgeblendet.

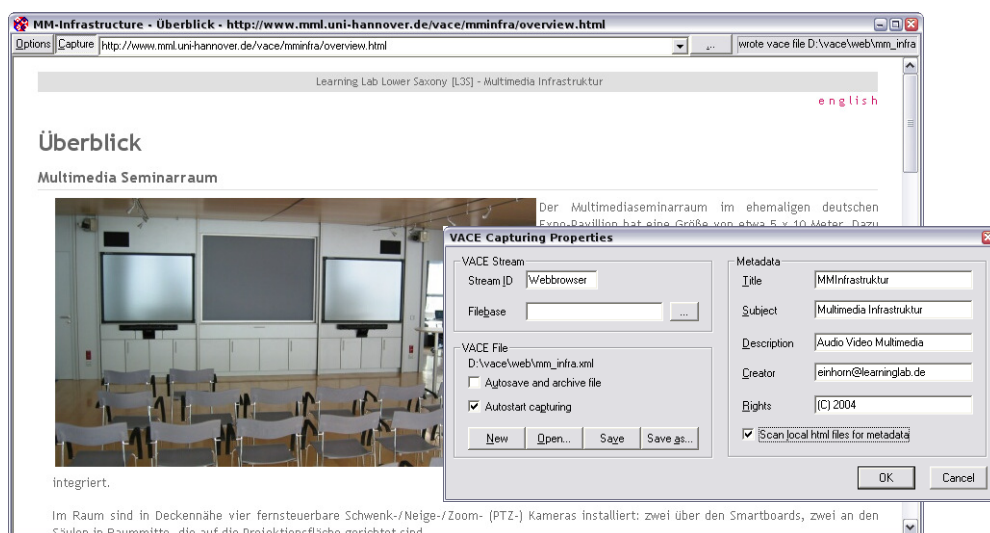


Abbildung 4-6: VaceBrowser, Applikationshauptfenster und Metadaten-Einstellungen

Einschränkungen ergeben sich lediglich abhängig von der Gestaltung der angezeigten Webseiten. So wird die Navigation durch längere Dokumente mittels Scrollbar nicht protokolliert – der Betrachter der Repräsentation muss den relevanten Inhalt selbstständig finden. Umgangen werden kann dieses Manko durch die

Verwendung von Anchor-Tags, die wie Links ausgeführt und damit aufgezeichnet werden.

Der Internet Explorer ist als ActiveX-Control in die Applikation eingebunden. Die vom Control gelieferten Events, z. B. nach vollständigem Laden der Webseite werden ausgewertet und anschließend der Titel im Event mit der bereits erfassten URL und Zeitmarke ergänzt. Lokale HTML-Dateien können zudem nach Metadaten durchsucht werden.

VACE PPT

Mit dem Programm VacePPT wird die VACE-konforme Aufzeichnung von PowerPoint Präsentationen realisiert. Es nutzt MS PowerPoint mittels OLE-Mechanismus und setzt damit die Installation von MS Office voraus. Der kostenfrei erhältliche PowerPoint Viewer bietet zwar ebenfalls eine Automatisierungsschnittstelle, stellt darüber aber nicht alle hier benötigten Funktionen (wie z. B. den Export von Daten) bereit.

Wie auch der VACE Web-Browser ist die Anwendung so konzipiert bzw. konfigurierbar, dass sie von einem Endnutzer ohne Einweisung leicht einsetzbar ist. Der Anwender startet VacePPT und lädt darüber die zu präsentierende PowerPoint-Datei. Die Präsentation wird wie gewohnt im Vollbild dargestellt, im Hintergrund überwacht VacePPT die aktuell dargestellten Folien. Optional lässt sich die Vor- und Zurückschaltung von Folien auch von VacePPT steuern.

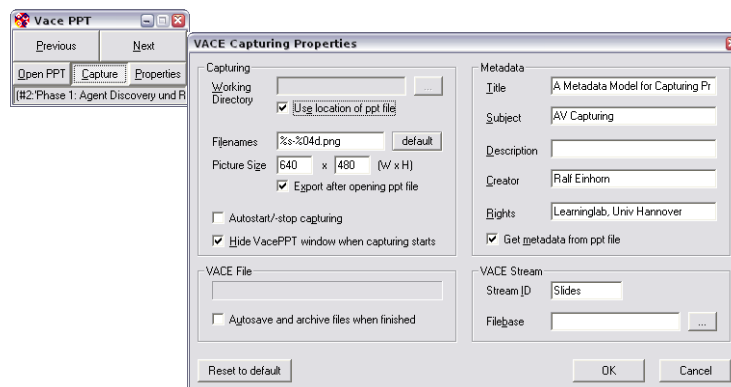


Abbildung 4-7: VacePPT, Applikationshauptfenster und Dialog für Einstellungen

VacePPT extrahiert die in der PPT-Datei verfügbaren Metadaten (Titel, Thema, Autor, Stichworte, Firma) und überträgt diese in VACE-Metadatenstruktur für den Stream. Bei der Anzeige der Folien werden jeweils der Text von Überschrift und Seiteninhalt extrahiert und als Titel bzw. Stichworte zu dieser Seite gespeichert.

Die einzelnen Folien (Slides) werden als Bilddaten wahlweise in den Formaten PNG, GIF, JPEG, EMF oder WMF abgelegt und können so für die Repräsentation weiter verarbeitet werden. Im Regelfall bietet sich der PNG-Standard an, der eine gute Kompression, einen großen Farbraum, Artefaktfreiheit und die Kompatibilität zur Anzeige in Web-Browsern aufweist. Möglich ist auch die Verwendung der vek-

torbasierten Formate EMF oder WMF mit anschließender Konvertierung in ein Publikationsformat wie SVG (vgl. Abschnitt 2.1.3).

Eine native Verwendung von PowerPoint als Repräsentationsmedium ist nur eingeschränkt möglich, da das Web-Browser-Plugin nicht als sicher klassifiziert ist und daher im allgemeinen nur mit lokal gespeicherten Daten funktioniert. Der in PowerPoint ebenfalls verfügbare Webseiten-Export erzeugt Dateien, die lediglich mit dem MS Internet Explorer angezeigt werden können und somit auch nur eingeschränkt einsetzbar sind. Die Nutzung von Einzelbildern stellt daher die universellste Methode dar; zusammen mit dem extrahierten Text wird die Navigation ermöglicht.

Der Export *einzelner* Folien ist in der vorliegenden PowerPoint-Version (Office XP) fehlerhaft implementiert; Schriftarten und Farben werden teilweise ersetzt und damit das Layout zerstört. VacePPT nutzt daher optional die Exportfunktion für die (*gesamte*) Präsentation. Hiermit werden die Bilddaten sämtlicher Folien in einem Durchlauf erzeugt. Werden bei umfangreichen Präsentationen nur wenige Folien gezeigt, führt dieses Vorgehen jedoch zu einem hohen Datenaufkommen nicht benötigter Bilddaten. Nicht berücksichtigt werden in beiden Export-Varianten die in den Folien eventuell enthaltenen Animationen; jede Folie wird als statisches Bild exportiert.

Weitere Browser-Anwendungen

Die Implementierung mit Hilfe von ActiveX-Controls bzw. OLE-Objekten ist sehr effizient, da sich die Programmierung der Browser-spezifischen Funktionen auf den Aufruf der genutzten Funktionen und die Auswertung der benötigten Callbacks beschränkt.

Auf ähnliche Weise können eine Reihe von Applikationen eingebunden werden, u. a. die Microsoft Office Programme. Leider gestaltet sich die spätere Anzeige hier problematischer als etwa bei Web-Seiten, da die Web-Browser-Plugins – wieder als ActiveX-Control, diesmal im Web-Browser eingebunden – nicht als „sicher für Scripting“ eingestuft sind. Der Abruf solcher Repräsentationen erfordert daher beim Betrachter die Veränderung der Sicherheitseinstellungen. Für geschlossene Benutzergruppen ist jedoch auch diese Art der Integration realisierbar; Voraussetzung ist neben der Nutzung des MS Internet Explorers auch ein installiertes MS-Office-Paket.

Echtzeit-Annotationen

Neben der Aufzeichnung von Medien ist in vielen Anwendungsfällen auch die Live-Erfassung von inhaltsbezogenen Metadaten relevant. Auf diese Weise können bereits während der Präsentation Informationen etwa zur Gliederung des Vortrags erfasst werden. Dies ist insbesondere dann sinnvoll, wenn diese Daten nicht durch andere Medien – z. B. Folien – geliefert werden bzw. mit diesen assoziiert sind. Dazu kommen Anwendungen wie die Kommentierung zu Auswertungszwecken.

Als streamtype wurde im Metadatenmodell der Typ *text/annotation* eingeführt, der ein Subset von HTML-Elementen zur hierarchischen Gliederung (die Überschriftsebenen *h1* bis *h4*) sowie Text ohne hierarchische Spezifikation (*p*) enthält.

Für die Echtzeitannotation wurde der VaceEditor, der auch zur allgemeinen Nachbearbeitung der Metadaten dient, um einige Funktionen zur Erfassung von Echtzeit-Annotationen erweitert.

- Annotationen können zu beliebigen Zeitpunkten eingegeben werden. Ihnen wird per Tastendruck oder Mausklick die aktuelle Zeit zugewiesen. Ferner sind manuelle Änderungen der Zeitstempel (das Verschieben eines Events) jederzeit möglich.
- Vereinfacht wird die Eingabe durch vier frei belegbaren Funktionstasten. Bei Anwahl einer Funktionstaste wird ein neues Event mit aktuellem Zeitstempel sowie vorgegebenen Typs und (optional) Inhalt eingefügt. Hiermit ist die schnelle Eingabe von Informationen in verschiedenen Gliederungsebenen möglich. Selbstverständlich können auch diese Werte jederzeit von Hand modifiziert werden.

Die Annotationen werden direkt in einer VACE-Metadatei gespeichert. Eine Mediendatei wird nicht benötigt.

4.5.2 Import

Während die im Abschnitt 4.5.1 beschriebenen Module bereits bei der Aufzeichnung parallel zu den Mediendaten die VACE-konformen Metadaten generieren, fehlen diese Metadaten bei generisch (d. h. mit Hilfe beliebiger Tools) aufgezeichneten Mediendaten.

Um auch solche Medien zu nutzen (*zu importieren*), werden die fehlenden Informationen nachträglich erzeugt. Beispielhaft sind im Folgenden zwei Importfilter beschrieben, die beide im VaceEditor implementiert sind.

Bilder

Neben der Videoaufzeichnung besteht bei der Mehrzahl von Vorträgen die Notwendigkeit einer separaten Erfassung von Präsentations-Folien (*Slides*). Hierzu bietet sich in vielen Fällen die Nutzung von (Einzel-) Bildformaten an. Diese können z. B. mittels spezieller Software auf dem vom Dozenten genutzten Rechner erfasst werden oder universell mit Hilfe eines Scan-Konverters aufgenommen werden.

Für die Einbindung und Weiterverarbeitung von Bildern werden – neben den Mediendaten selbst – folgende Metadaten benötigt.

- Der **Zeitpunkt** der Aufzeichnung des Bildes ist in den Dateisysteminformationen (*timestamp*) enthalten. Es wird vorausgesetzt, dass das Bilderfassungs-Programm das aufgenommene Bild unmittelbar im Dateisystem ablegt. Ein statischer Versatz kann so gegebenenfalls leicht korrigiert werden.
- Der **Bilddateityp** kann aus der Datei selbst oder deren Namenspostfix (z. B. .jpg) ermittelt werden.
- Ist das Format bekannt, so können aus den Header-Daten Informationen zur Bildgröße (**Breite** und **Höhe**) ermittelt werden.

- Die z. B. in SMIL/RealPix benötigte **Übertragungsrate** der Bilddaten ergibt sich aus der Gesamtdatenmenge dividiert durch die Länge der Präsentation

$$\text{Bitrate} = (\sum_{i=1..n}(\text{Dateigröße Bild}_i)) / (t_{\text{Bild}_n} - t_{\text{Bild}_1})$$

Unterstützt werden derzeit die Formate PNG, JPEG und GIF. Zur Bezeichnung der Events wird der Dateiname als Titel genutzt, weitere inhaltsbezogene Metadaten werden nicht automatisch generiert.

HTML-Metadaten

Inhaltsbezogene Metadaten in HTML-Dateien können nachträglich automatisch extrahiert werden. Hierzu wurde ein entsprechender Parser implementiert, der die HTML-Datei nach bekannten Tags in der Form `<meta name="Description" content="Implementierung des VACE-Frameworks">` durchsucht und diese Informationen den VACE-Elementen der HTML-Events ergänzt. Die Tabelle 4-6 zeigt die Zuordnung der unterstützten Elemente. Hinzu kommt das title-Element, das dem HTML-Tag `<title>` entnommen wird.

HTML metatag	VACE event element
description	description
keywords	subject
author	creator
copyright	rights

Tabelle 4-6: Übertragung von HTML-Metatags in Elemente der VACE-Events

Die Funktion ist als Bestandteil von *CVaceHelper* implementiert und kommt im VaceEditor und VaceBrowser zum Einsatz.

4.5.3 Nachbearbeitung

Sämtliche Nachbearbeitungsfunktionen sind – wie auch die Import- und Exportmodule – derzeit im VaceEditor implementiert. Das Windows-Programm (Abbildung 4-8) verfügt über ein skalierbares grafisches Benutzer-Interface (GUI) in den Sprachen Deutsch und Englisch. Die wesentlichen Funktionen werden in den folgenden Abschnitten erläutert, die ausführliche Erklärung sämtlicher GUI-Elemente befindet sich im dazu erstellten Benutzerhandbuch. Das in HTML erstellte Handbuch steht im (HTML-basierten) CHM-Format als mit der Taste F1 aufrufbare Online-Hilfe zur Verfügung (Abbildung 4-9).

Zusammenführung der Daten

Zentraler Arbeitsschritt nach der Aufzeichnung ist das Zusammenführen (*Merging*) der verschiedenen Medien- und Metadaten. Hierzu werden die von den Aufzeichnungsmodulen generierten VACE-Metadaten in Form der gespeicherten XML-Dateien im VaceEditor geöffnet bzw. zusammengefügt.

Die so entstehende Datenstruktur enthält die (inhalts- und zeitbezogenen) Metadaten zu sämtlichen aufgezeichneten Medienströmen. Globale Elemente der Präsen-

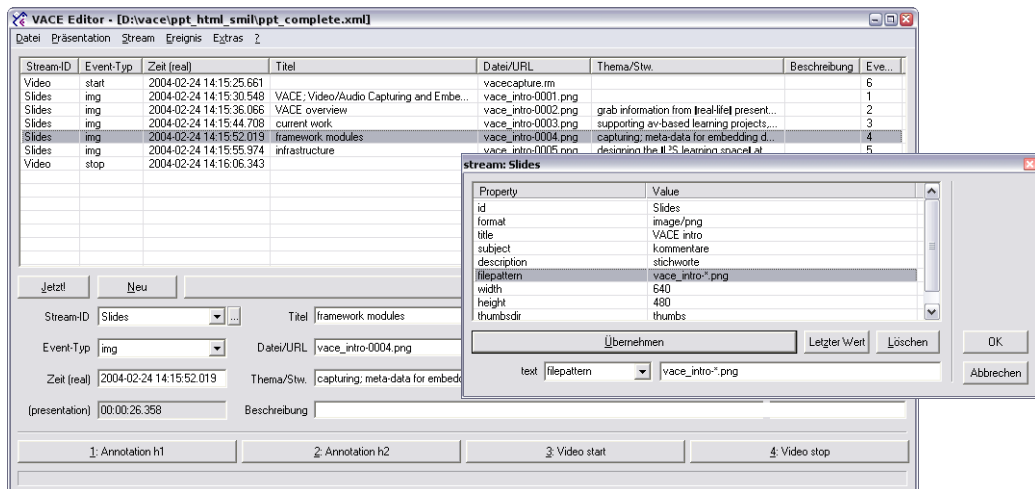


Abbildung 4-8: VaceEditor, Applikationshauptfenster und Eingabedialog

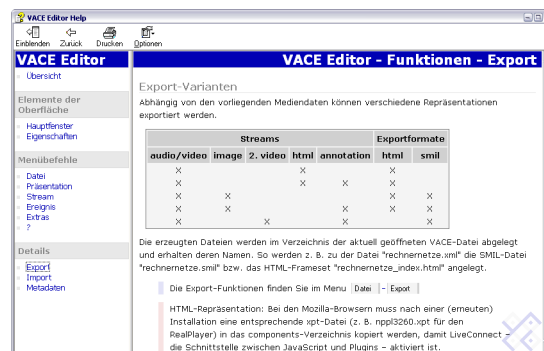


Abbildung 4-9: VaceEditor, HTML-basierte Online-Hilfe im Windows-Help-Viewer

tation – die potenziell unterschiedliche Werte enthalten können – werden dabei entsprechend einer ODER-Verknüpfung zusammengefasst.

Bearbeitung der Metadaten

Alle Elementes des in Abschnitt 4.4.3 eingeführten Metadatenmodells sowie weitere, benutzerdefinierte Elemente lassen sich im Editor bearbeiten. Neben der Modifikation zusammgeführter Daten können Elemente ergänzt und entfernt werden.

Die hierarchische Struktur findet sich in der Nutzeroberfläche des Editors wieder. So sind die Eigenschaften der Elemente Präsentation, Stream und Ereignis (presentation, stream, event) jeweils zusammengefasst. Zentrales Element des GUIs ist die Liste der Ereignisse, die einen schnellen Überblick der zeitlichen Anordnung der einzelnen Abschnitte der Medienströme ermöglicht. Die Werte der Event-Eigenschaften können direkt nach Auswahl des Events bearbeitet werden. Zur Bearbeitung der seltener benötigten Eigenschaften der Präsentation und der Streams öffnet sich bei Bedarf ein weiteres Fenster, in dem Eigenschaften (property) und Werte (value) in Form einer Tabelle dargestellt werden.

Sofern es einen definierten Wertebereich für eine Eigenschaft gibt (z. B. Event-Typ) kann der Wert aus einer Drop-Down-Liste ausgewählt werden. Alternativ können nicht enthaltene Werte von Hand eingegeben werden.

Definition der Präsentations-Timeline

Das Festlegen von Start- und Endzeitpunkt (bzw. Länge) der Präsentation gehört zur Metadatenbearbeitung. Aufgrund der Bedeutung für den folgenden Export sei an dieser Stelle darauf explizit eingegangen.

Der VaceEditor enthält eine Funktion zum automatischen Festlegen von Start- und Endpunkt. Dabei werden der erste und letzte Zeitstempel aus allen Events ermittelt und als *starttime* bzw. *endtime* der Präsentation eingetragen.

Oft ergibt sich bei der Aufzeichnung jedoch ein Vor- und/oder Nachlauf, der in der Repräsentation nicht sichtbar sein soll. Hierzu können Start und Ende manuell eingegeben werden oder der Zeitstempel eines Events (z. B. einer Folie) genutzt werden. Im Editor stehen hierzu entsprechende Funktionen bereit.

Konvertierung

Für die Konvertierung von Mediendaten, etwa die Reduzierung der Bildqualität von Videos oder Einzelbildern zugunsten geringerer Speichermenge, können Module zur Ansteuerung externer Applikationen (Encoder bzw. Bildverarbeitung) eingebunden werden. Damit ergibt sich eine alternative Repräsentation, deren zeit- und inhaltsbezogene Metadaten identisch sind.

Exemplarisch wurde die Erstellung verkleinerter Vorschaubilder (Thumbnails) für die Nutzung in HTML-Seiten implementiert. Hierzu wird das frei verfügbare Programm IrfanView [22] eingesetzt, das neben diversen anderen Funktionen auch die Skalierung von Bilddaten in allen gängigen Formaten ermöglicht.

4.5.4 Export

Im Prototyp des VaceEditors sind verschiedene Export-Funktionen zur Erzeugung einer Repräsentation enthalten. Die Funktionen wurden in der Klasse CVaceExport implementiert und stehen damit auch weiteren Applikationen zur Verfügung.

Abhängig von den vorliegenden Aufzeichnungsmedien können die in Tabelle 4-7 aufgeführten Formate erzeugt werden.

streams	Exportformat
video, text/html	HTML
video, text/html, text/annotation	
video, image	
video, image, text/annotation	

Tabelle 4-7: VACE-Exportformate

streams	Exportformat
video, image	SMIL
video, image, text/annotation	
video, video (slides), text/annotation	
beliebig	Text (nur Metadaten)

Tabelle 4-7: VACE-Exportformate

HTML

Beim HTML-Export wird ein Frameset erzeugt, das die folgenden Elemente enthält.

- Der **AV-Player** umfasst neben der Anzeige des Videobildes die vom Videorekorder bekannten Steuerelemente wie Start, Pause, Spulen etc., einen Schieberegler (Slider) zum wahlfreien Zugriff sowie eine Statusanzeige, die u. a. die aktuelle Position enthält.
- Im **Inhaltsfenster** (Content) werden synchron zur im Videoplayer angezeigten Szene die entsprechenden Medien – meist Vortragsfolien – in HTML oder als Bild dargestellt.
- Der **Index** enthält eine Liste der einzelnen Abschnitte mit den verfügbaren Angaben wie Titel, Beschreibung, Stichworte. Der Betrachter kann diese Bereiche direkt anwählen. Anhand ihrer Zeitstempel werden Video und Inhalt entsprechend angezeigt.
- Wenn als Inhaltsformat Bilder (image/*) verwendet werden können in der optionalen **Vorschau** (Thumbnails) verkleinerte Darstellungen angezeigt werden, die analog zur Navigation direkt angewählt werden können. Auch dieses Fenster wird mit der aktuellen Position synchronisiert, indem das im Inhaltsfenster angezeigte Bild durch eine farbige Umrahmung hervorgehoben wird.
- In dem Rahmen **Suche** kann der Benutzer Begriffe eingeben, um nach Stichworten in den vorliegenden Metadaten zu suchen und so gezielt entsprechende Abschnitte finden.

Die HTML-Repräsentation (Abbildung 4-10) ist mit Hilfe von JavaScript (JS) und dem Document Object Model (DOM, [82]) realisiert, als Videoplayer wird der RealPlayer als Plugin eingesetzt. Das Aussehen der HTML-Elemente wird mit Hilfe eines zentralen Stylesheets (CSS, cascading style sheet [80]) bestimmt und kann so leicht modifiziert werden, um z. B. an das Design bereits existierender Webseiten angepasst zu werden.

Die Synchronisation der Medien bei der Wiedergabe erfolgt durch die JS-Engine des Web-Browsers, als zeitliche Basis dient das Video. Eine JS-Funktion fragt periodisch im (einstellbaren) Interval von 50 ms die Position im Videoplayer ab, wodurch sämtliche Positionsänderungen berücksichtigt werden – durch einfaches Abspielen oder Nutzereingriff, z. B. auch während der Slider bewegt wird. Diese Position wird mit der Datenstruktur im Navigationsfenster verglichen. Falls sich

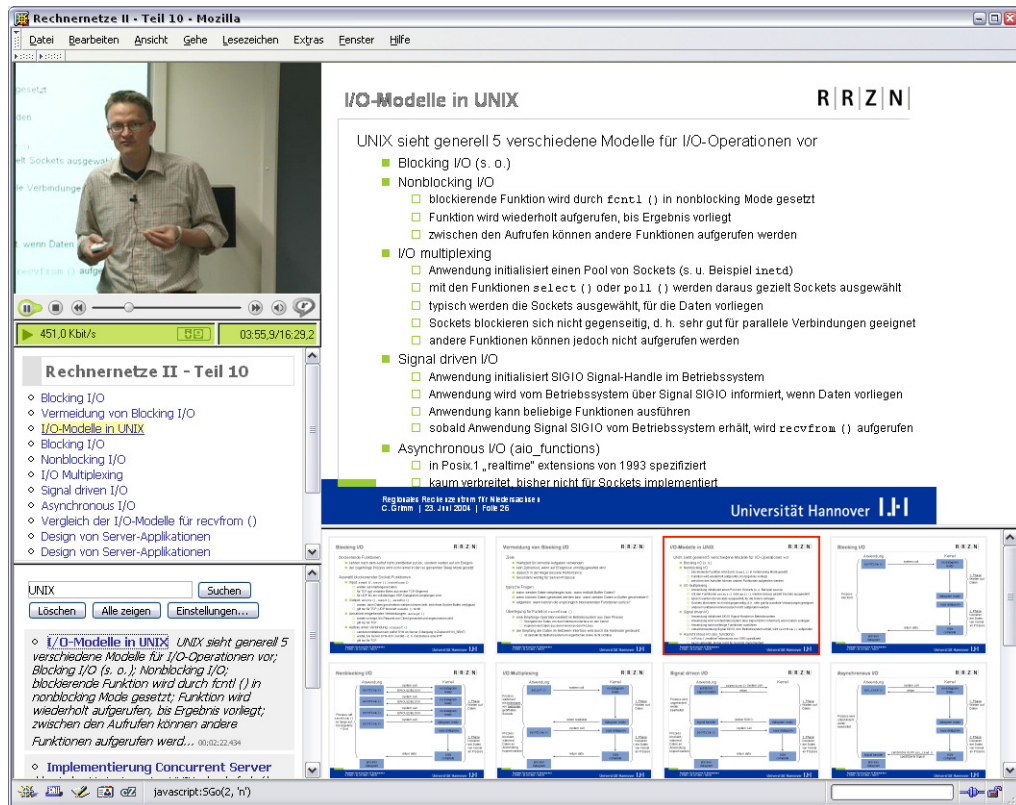


Abbildung 4-10: HTML-Repräsentation einer Präsentationsaufzeichnung

hierbei eine Veränderung des Inhaltsfensters ergibt, wird dieses entsprechend der Liste aktualisiert und der Abschnitt wird in der Navigationsliste optisch hervorgehoben. Wählt der Benutzer einen Verweis in der Navigation oder der Suche, so wird die Position im Videoplayer angesprungen. Mit der Positionsüberwachung folgt mittelbar die Synchronisation des Inhalts (siehe Abbildung 4-11).

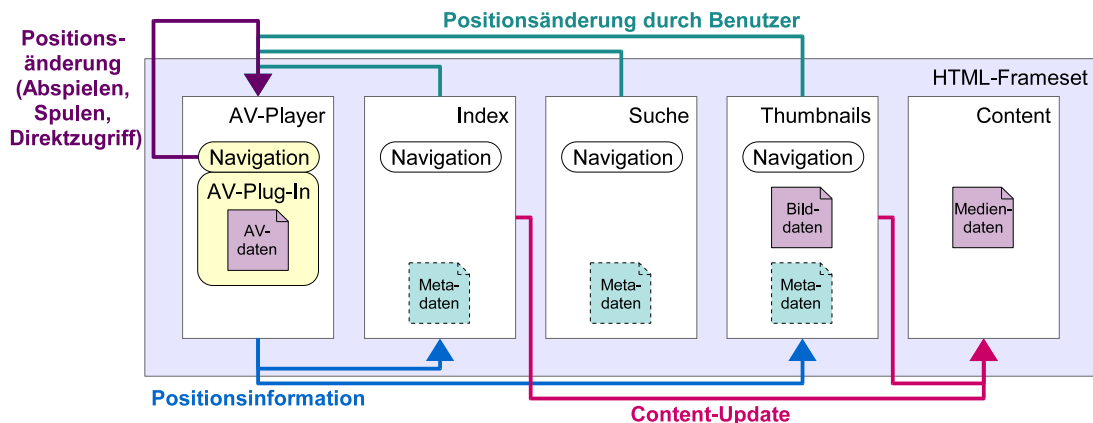


Abbildung 4-11: Modell der Medien-Synchronisation im HTML-Frameset

Beim Export werden Frameset, AV-Player, Index, Vorschau und Suche sowie eine SMIL-Datei zum Zugriff auf die Videodaten dynamisch erzeugt. Die benötigten

JS- und CSS-Dateien sind statisch und werden dazu kopiert. Die Mediendaten im Inhaltsbereich liegen bereits aus der Aufzeichnung vor.

Die so erzeugten Daten können sowohl lokal im Dateisystem (als CD-ROM etc.) als auch auf einem Webserver bereitgestellt werden. Optional kann die Videodatei über einen Streaming-Server ausgespielt werden, wozu lediglich der Link innerhalb der SMIL-Datei anzupassen ist.

Die Systemvoraussetzungen beim Betrachter sind ein aktueller Web-Browser (ab Mozilla 1.5, Netscape 7, Firefox, Internet Explorer 6) mit aktiviertem JavaScript sowie der RealPlayer.

SMIL

Die SMIL-Repräsentation enthält neben dem *Video* eine Inhaltsübersicht in Form einer Liste (*Navigation*), in der gezielt Abschnitte angewählt werden können. Parallel dazu werden z. B. die Folien in Form eines weiteren Videos oder in Form von Einzelbildern dargestellt (*Inhalt*).

Beim Export wird neben der SMIL-Datei, welche die Informationen zur Synchronisation der einzelnen Medien enthält, für die Navigation eine RealText-Datei erzeugt, die ähnlich wie eine HTML-Seite die Links auf die einzelnen Abschnitte enthält. Werden für die zweite Bilddarstellung Einzelbilder verwendet, so ist hierzu die Generierung einer RealPix-Datei erforderlich, in der die zeitliche Anordnung der Bilder beschrieben ist. Die Synchronisation der Medien obliegt hier dem SMIL-Player, im vorliegenden Fall wurde der RealPlayer eingesetzt.

Wie beim HTML-Export können auch hier die erzeugten Repräsentationen (Abbildung 4-12) sowohl lokal als auch über das Netz – vorzugsweise via Streaming-Server – abgespielt werden. Die Systemvoraussetzung beim Betrachter besteht hierbei lediglich in einem installierten RealPlayer.

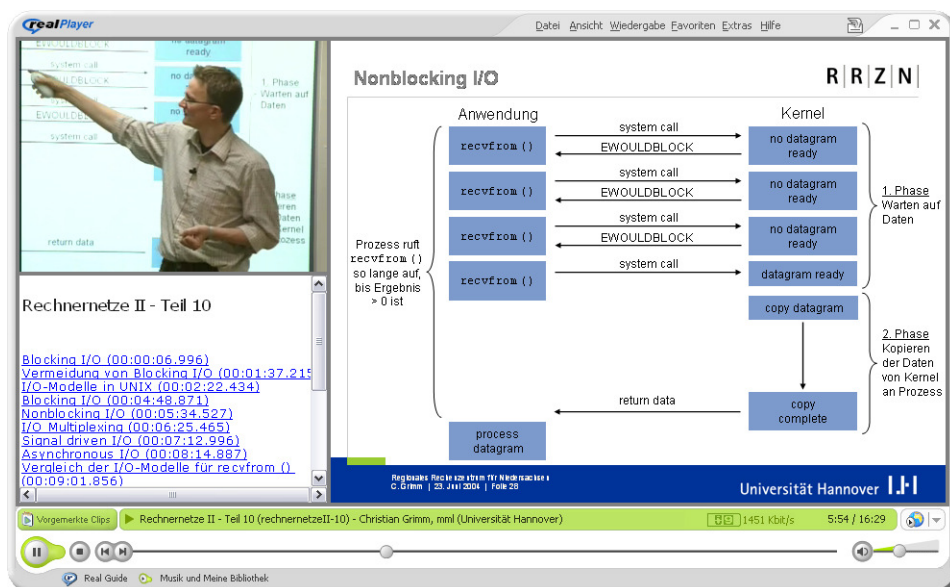


Abbildung 4-12: SMIL-Repräsentation einer Präsentationsaufzeichnung im RealPlayer

Text

Der Metadaten-Export in Text-Form dient vornehmlich zur Fehlersuche (Debugging), eignet sich aufgrund seiner Struktur aber auch zur generischen Weiterverarbeitung in nicht explizit berücksichtigten Applikationen, z. B. als Schnittliste für die Videobearbeitung (EDL).

Es wurden zwei Versionen implementiert, die sich in der Art der Anordnung unterscheiden. Beide enthalten die globalen Daten zur Präsentation. Dazu kommt wahlweise eine Liste der globalen Stream-Daten gefolgt von einer globalen Liste aller Events in zeitlicher Folge oder eine Liste der Streams mit den jeweils enthaltenen Events.

4.5.5 Implementierte Module des VACE-Frameworks

In Abbildung 4-13 sind abschließend die vorangehend beschriebenen implementierten Module zusammengestellt.

Die Darstellung orientiert sich am abstrakten Framework in Abbildung 4-1

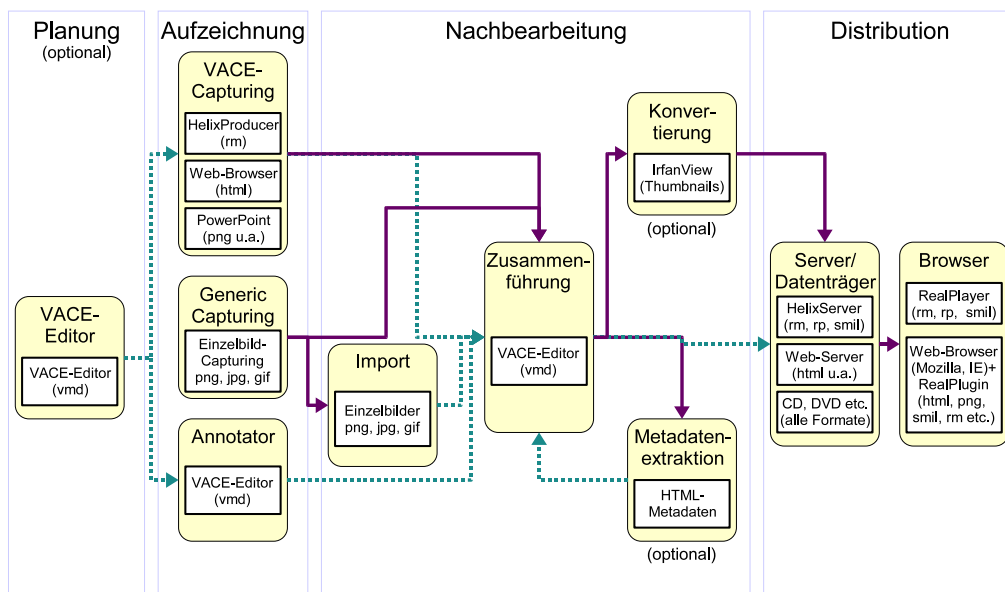


Abbildung 4-13: Implementierte VACE-Framework-Module

5 Anwendung, Performance und Bewertung

5.1 Anwendung des VACE-Frameworks

In diesem Abschnitt wird der Einsatz des entwickelten Systems an Hand des implementierten Prototyps beschrieben. Hierzu wird eine typische Anwendung gemäß Abschnitt 3.8.1 in zwei exemplarischen Varianten unter Produktionsbedingungen durchgeführt und untersucht.

5.1.1 Testscenarios

Als Testscenario für den Einsatz des entwickelten Systems wurde die Aufzeichnung und Verarbeitung einer typischen Vorlesung gewählt. Der Vortragende benutzt zur Präsentation PowerPoint. Die Repräsentation soll eine Streamingfähige Audio-/Videoaufzeichnung mit Bild und Ton des Vortragenden sowie die parallele Darstellung der Folien in adäquater Auflösung enthalten.

Die Aufzeichnung erfolgt im MM-Studio des RRZN, einem mittelgroßen mit professioneller Medientechnik ausgestatteten Seminarraum (vgl. Abschnitt 2.2.2, [55]). Für die Produktion werden neben dem vorhandenen Equipment Teile des implementierten Prototyps (Abschnitt 4.5) aus dem VACE-Framework eingesetzt.

Zur Verdeutlichung der unterschiedlichen Anwendungsfälle *Vorlesung* und *Tagung* werden zwei unterschiedliche Szenarios definiert, die sich durch die Zugriffsmöglichkeit auf die Folienpräsentation unterscheiden.

- **Szenario 1** („*Vorlesung*“, Abbildung 5-1): Für die Produktion kann auf den PC des Dozenten zugegriffen werden. In diesem Fall wird das Programm VacePPT (Abschnitt 4.5.1, „VACE PPT“) für die VACE-konforme Aufzeichnung der PowerPoint-Folien eingesetzt.
- **Szenario 2** („*Tagung*“, Abbildung 5-2): Wie bei Tagungen üblich steht hier lediglich das Ausgangssignal des Dozenten-Laptops zur Verfügung, das auch auf der Projektion angezeigt wird. Die Aufzeichnung der Präsentation erfolgt somit über das VGA-Signal des Dozenten-Rechners. Dieses Setup kann auch bei Präsentationen mit herkömmlichen Folien mittels Video-Presenter eingesetzt werden (vgl. Abschnitt 2.2.1).

Im Folgenden wird jeweils zunächst Szenario 1 erläutert und anschließend zu Szenario 2 die Unterschiede zur ersten Variante beschrieben.

5.1.2 Durchführung

AV-Aufzeichnung

Zur AV-Aufzeichnung wird die vorhandene Infrastruktur eingesetzt. Der Dozent ist mit einem drahtlosen Ansteckmikrofon zur Sprachaufnahme ausgestattet. Das

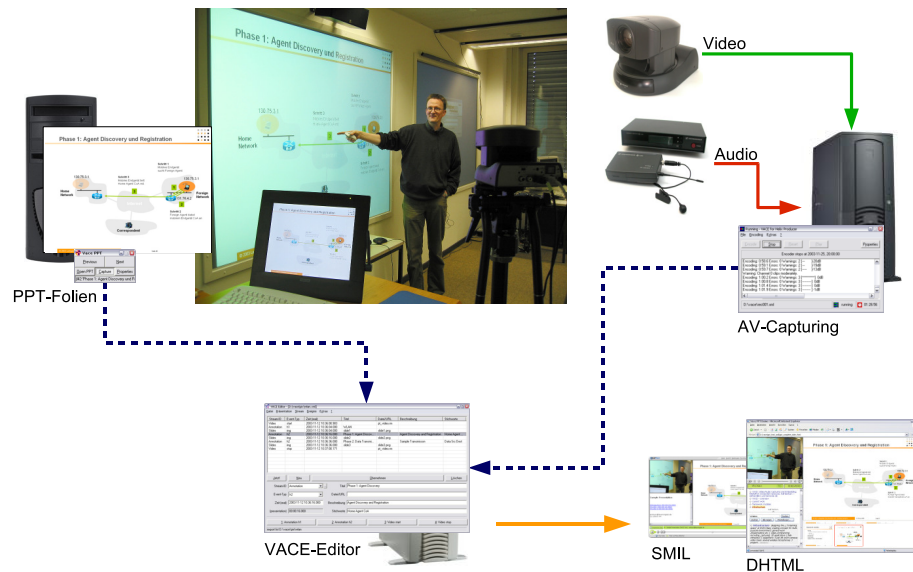


Abbildung 5-1: Testszenario 1: Vorlesungsaufzeichnung mit VacePPT

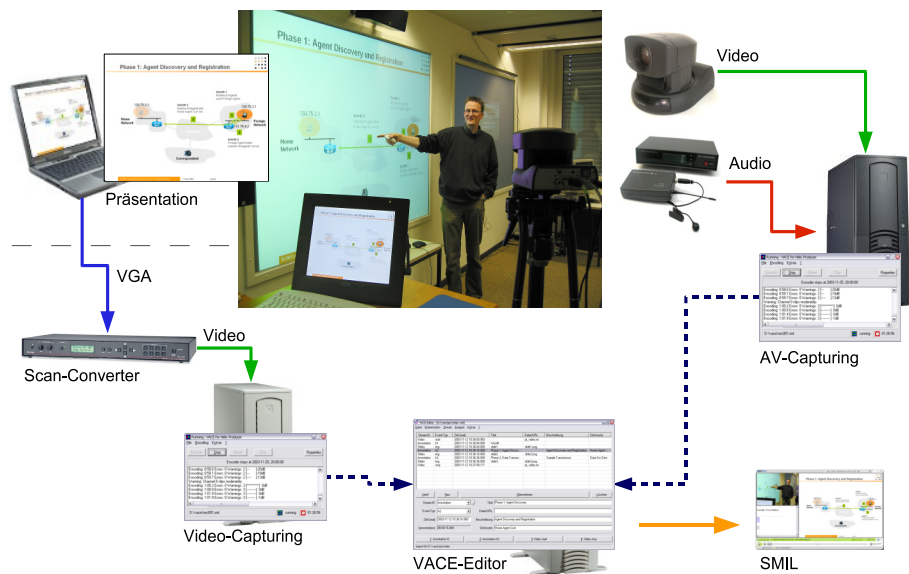


Abbildung 5-2: Testszenario 2: Vortragsaufzeichnung mit generischer Folienaufnahme

Videobild wird mit Hilfe von fernsteuerbaren Kameras erfasst und mittels bildmischer verarbeitet. Die AV-Ausgangssignale werden mit einem entsprechend ausgestatteten PC im Regierbereich in Echtzeit kodiert (vgl. Abschnitt 2.2.2: Live-Encoding). Hierzu wird neben dem installierten RealProducer der VaceHxPW (vgl. Abschnitt 4.5.1: VACE RealProducer Wrapper) eingesetzt. Nach der einmaligen Einrichtung des Wrappers wird das Encoding über VaceHxPW gesteuert; beim RealProducer müssen keine Einstellungen vorgenommen werden. Der PC empfängt über das lokale Netz via NTP die aktuelle Uhrzeit (siehe Abschnitt 2.4.2).

Die AV-Aufzeichnung wird von einem Betreuer im Regierbereich gesteuert, der neben den Kameraeinstellungen, der Bildmischung und der Tonregelung das Encoding manuell startet. Dieses Vorgehen ist in Szenario 1 und 2 identisch und ent-

spricht dem üblichen Vorgehen bei Vorlesungsaufzeichnungen. Der Unterschied zu herkömmlichen Setups besteht lediglich im Einsatz des Wrappers. Am Ende der Aufzeichnung stehen neben der AV-Datei zusätzlich die VACE-Metadaten mit den Synchronisationsinformationen bereit. Beide Dateien können nach der Fertigstellung vom VaceHxPW automatisch zur weiteren Verarbeitung auf einen weiteren Rechner im Netz verschoben werden.

Aufzeichnung der Folienpräsentation (Dozent)

Zur Präsentation der PowerPoint-Folien benutzt der Vortragende einen separaten Rechner im Bereich der Projektion. Über das LAN kopiert der Dozent die Folien auf diesen PC. Im Unterschied zur üblichen Präsentation mit MS PowerPoint startet er zunächst VacePPT und öffnet damit die PPT-Datei (z. B. per Klicken und Ziehen der Datei auf das VacePPT-Fenster). VacePPT öffnet daraufhin die Präsentation mit PowerPoint.

Zum Start der Folienaufzeichnung wählt der Dozent „Capture“; VacePPT läuft daraufhin im Hintergrund. Der Benutzer kann wie gewohnt die Folien wechseln bzw. gezielt auswählen. Am Ende der Präsentation werden die VACE Metadaten neben den exportierten Bildern auf einen PC zur Nachbearbeitung verschoben. Wie auch bei der AV-Aufzeichnung kann dies automatisiert erfolgen.

Szenario 2

Im Gegensatz zu Szenario 1 besteht für den Dozenten bei der Nutzung seiner Präsentationssoftware kein Unterschied im Vergleich zur Verwendung ohne Aufzeichnung. Der verwendete Rechner (meist ein Notebook) wird lediglich an die Projektion angeschlossen.

Zur Erfassung der Folien wird daher das zur Verfügung stehende VGA-Signal der Projektion parallel in einen Scan-Converter eingespeist, der das Bild in ein TV-konformes Videosignal wandelt (vgl. Abschnitt 2.2.2: Scan-Converter). Damit kann das Bild mit gängiger Video-Hardware weiter verarbeitet werden. Hier wird das Bildsignal analog zur vorher beschriebenen Aufzeichnung des Dozenten-AV-Streams mit Hilfe eines weiteren PC-Systems live kodiert. Dabei kommt ebenfalls der VaceHxPW zum Einsatz. Die Aufzeichnungsparameter werden jedoch an die vorliegende Anforderung angepasst: kein Ton, höhere Auflösung, geringere Bildrate („Slideshow“-Modus). Auch hier werden neben den AV-Daten die VACE-Metadaten erzeugt.

Nachbearbeitung

Zur Nachbearbeitung werden die erzeugten VACE-Metadaten mit dem VaceEditor (siehe Abschnitt 4.5.3) zusammengeführt: die beiden XML-Dateien werden hierzu mit dem Programm geöffnet. Für die generierten Einzelbilder der Folien werden Vorschaubilder erzeugt, die in der anschließend erzeugten Repräsentation eine visuelle Übersicht ermöglichen. Sämtliche Aktionen entsprechen jeweils einem Funktionsaufruf im Menü des VaceEditors:

- Zusammenführen (beide XML-Dateien öffnen),
- Thumbnails aus Image-Stream erzeugen,

- globale Metadaten aus Slides-Stream übernehmen,
- Start und Ende der Repräsentation aus Slides-Stream übernehmen,
- Export der Repräsentation.

Für die vorliegenden Streams ist der Export als HTML-basierte und SMIL-Präsentation möglich (vgl. Abschnitt 4.5.4). Manuelle Ergänzungen oder Modifikationen der (inhaltsbasierten) Metadaten werden nicht vorgenommen.

Szenario 2

Der Slides-Stream besteht im Szenario 2 aus einer weiteren Videodatei – nicht wie in Szenario 1 aus einzelnen Bildern. Die Zusammenführung der beiden Streams entspricht dem in Szenario 1. Allerdings liegen hier keine aus der Präsentation extrahierten inhaltlichen Metadaten vor. Auch fehlen die Kapitelinformationen zur Navigation, die sich bei der ersten Variante aus den Folienwechseln ergeben. Titel und Autor werden im VaceEditor manuell eingegeben; die beiden Medien werden in Form einer SMIL-Präsentation exportiert.

Repräsentation

Aus den aufgezeichneten Medien- und Metadaten werden die in Abschnitt 4.5.4 beschriebenen HTML- und SMIL-Repräsentationen erzeugt. Die hierfür benötigten Daten werden vom VaceEditor beim Export generiert.

Beide Versionen eignen sich unmittelbar zum Abruf im lokalen Dateisystem und via Webserver. Zudem können beide Repräsentationen mit minimalen Anpassungen auf einem Real-Streaming-Server abgelegt werden, was sich speziell bei AV-Clips mit umfangreichem Datenvolumen anbietet (vgl. Abschnitt 2.1.4).

Bei der HTML-Version wird der AV-Stream auf einen Real-Streaming-Server abgelegt, die restlichen Daten werden auf einen Webserver kopiert. Zum Abruf der AV-Daten über den Streaming-Server ist die URI anzupassen. Im Prototyp ist diese Funktion noch nicht implementiert, sodass dies manuell mittels Texteditor erfolgt.

Szenario 2

Das Vorgehen bei der Repräsentationserstellung ist grundsätzlich identisch zum Szenario 1, allerdings wird lediglich eine SMIL-Repräsentation erzeugt. Beide Video-Dateien werden zum Abruf auf den Streaming-Server kopiert, die URIs manuell angepasst.

Das Datenvolumen der Folien-Repräsentation ist aufgrund des Mediums Video deutlich umfangreicher als die Einzelbilder in Szenario 1.

5.1.3 Auswertung

Aufwand

Als Vergleichsmaßstab für den Aufwand dient das Vorgehen ohne die Komponenten aus dem VACE-Framework.

Als einmaliger Mehraufwand ist die Installation bzw. erstmalige Einrichtung der VACE-Software-Komponenten anzusehen. Eine Installation ist nicht erforderlich;

die ausführbaren Programme können direkt gestartet werden. Zudem ist auf den beteiligten Rechnern ein NTP-Client zu installieren und zu konfigurieren.

Für den Vortragenden entsteht ein minimaler Mehraufwand durch das Starten der PPT-Präsentation mittels VacePPT. Während der Präsentation besteht keinerlei Beeinträchtigung – die Benutzung von PowerPoint erfolgt wie gewohnt. Der Folienexport für die Repräsentation geschieht optional im Hintergrund (beim ersten Aufruf einer Folie) oder in einem Durchgang (Erläuterung siehe Abschnitt 4.5.1: VACE PPT). Je nach Komplexität der PPT-Präsentation kann dieser Vorgang einige Sekunden in Anspruch nehmen. Zum Ende der Präsentation wird VacePPT explizit gestoppt; die Metadaten werden gesichert und zusammen mit den Bilddaten auf den Rechner zur Nachbearbeitung übertragen.

Ähnlich ist der Mehraufwand bei der AV-Produktion. Auch hier wird die Applikation (in diesem Fall der Encoder) mit Hilfe eines Wrappers gestartet. Allerdings steuert der VaceHxPW den Encoder komplett, sodass die benötigten Parameter (Auflösung, Bitrate, Videomodus) lediglich in VaceHxPW bzw. der vorbereiteten VACE-Metadatei definiert werden müssen. Der Zeitaufwand ist identisch mit der Anwendung der originären Applikation.

Für die oben beschriebene Nachbearbeitung sind die VACE-Metadaten sowie – für die Erstellung der Thumbnails – die Einzelbilddaten der Folien auf einem Rechner zusammenzuführen. Alternativ kann mittels SMB über das Netz auf die Daten zugegriffen werden. Die Durchführung der erläuterten Bearbeitungsschritte dauert weniger als eine Minute; auch der Zeitaufwand für die Thumbnail-Generierung ist vernachlässigbar. Weitere Betrachtungen zu den Zeitaufwänden folgen in Abschnitt 5.2.2.

Szenario 2

Im Gegensatz zu Szenario 1 gibt es in Szenario 2 keinerlei Beeinflussung des Dozenten, sodass an dieser Stelle auch kein Mehraufwand auftritt.

Der Aufwand für die Produktion des zweiten Videodatenstroms mit der Folienaufzeichnung entspricht dem der Erstellung für das Dozenten-Video. Der benötigte Scan-Konverter ist wie der zweite Encoding-PC in die Medientechnik-Infrastruktur integriert, da diese Konfiguration auch beim Live-Streaming von Vorträgen eingesetzt wird.

Auch im Szenario 2 ist der zeitliche Aufwand für die Nachbearbeitung minimal. Die im Szenario 1 automatisch aus der PPT-Datei extrahierten globalen Metadaten (Titel, Autor etc.) sind wahlweise bei der Aufzeichnung oder der Nachbearbeitung manuell einzugeben.

Tabelle 5-1 fasst den durch VACE entstandenen Mehraufwand zusammen. In den Zeitangaben sind lediglich die manuellen Arbeitsschritte, nicht die Dauer für das Kopieren der Daten über das Netz berücksichtigt (siehe dazu Abschnitt 5.2.2).

Mehraufwand	Szenario 1	Szenario 2
Einmalige Vorbereitung der Produktion	Installation von VaceHxPW, Installation von nptime (ca. 10 Minuten)	Installation von VaceHxPW, Installation von nptime auf beiden Capturing-PCs (ca. 15 Minuten)
Einmalige Vorbereitung des Dozenten-PCs	Installation von VacePPT Installation von nptime (ca. 10 Minuten)	–
Aufzeichnung/ Produktion	Starten der Kodierung mit Vace- HwPW (kein Mehraufwand)	Starten der Kodierung mit Vace- HwPW (2x, kein Mehraufwand)
Aufzeichnung/ Dozent	Starten der Präsentation über VacePPT, Start/Stop klicken (<1 Minute)	–
Nachbearbeitung (Produktion)	Zusammenführen der Metadaten, Bearbeitung, Export (<1 Minute)	Zusammenführen der Metadaten, Bearbeitung (globale Metadaten ergänzen), Export (2 Minuten)
Publikation	Transfer der Daten auf Server, manuelle URI-Anpassung (2 Minuten)	Transfer der Daten auf Server, manuelle URI-Anpassung (2 Minuten)

Tabelle 5-1: Mehraufwände durch den Einsatz von VACE in den Testszenarios

Eigenschaften der Repräsentation

Die **HTML**-Version enthält die synchronisierten Elemente in einem D-HTML-Frameset (vgl. Abbildung 4-10). Neben dem AV-Player (in dem das Video des Dozenten angezeigt wird) und den Folien (in Form von PNG-Bildern) erhält der Betrachter verschiedene Mittel zur Navigation. Die Inhaltsübersicht mit den Folienüberschriften und die Folienvorschau ermöglichen die direkte Anwahl eines gewählten Abschnitts. Die gerade im Video erläuterte Folie wird in beiden Übersichten farbig hervorgehoben und bei Bedarf in den sichtbaren Anzeigebereich des Rahmens gebracht. Zudem kann in einem weiteren Rahmen nach Stichworten gesucht werden. Die gefundenen Folientitel werden angezeigt und können direkt ausgewählt werden.

Die automatisch generierten Übersichts- und Suchelemente realisieren umfangreiche Navigationsmöglichkeiten und erleichtern so den gezielten Zugriff auf einzelne Abschnitte. Auch ohne manuelle Nachbearbeitung sind die (hier aus den Folien extrahierten) Metadaten aussagekräftig. Abbildung 5-3 zeigt hierzu entsprechende Beispiele.

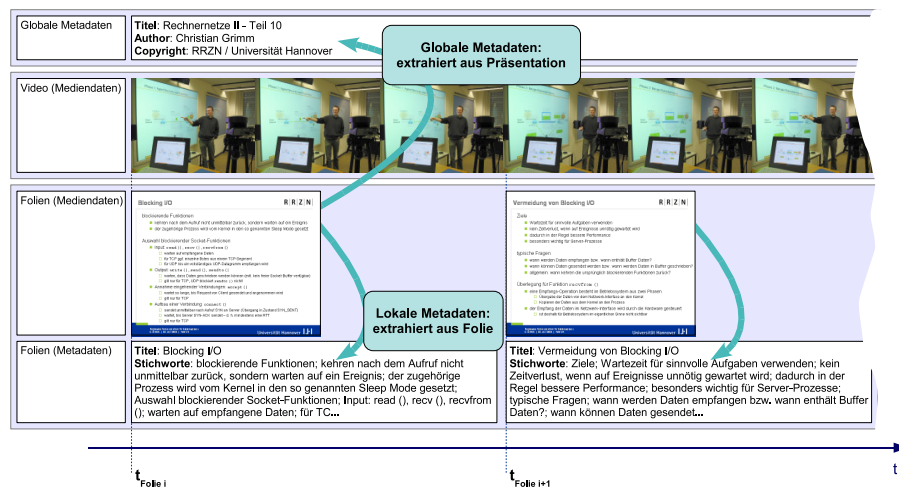


Abbildung 5-3: Assoziation extrahierter Metadaten zur Navigationserzeugung

Sämtliche parallel dargestellten Elemente bleiben unabhängig von der Zugriffsförm (vgl. Abbildung 4-11) synchron. Die Abweichungen bleiben dabei unterhalb von 500 ms (Messungen zur Synchronisationsgüte folgen in Abschnitt 5.2.1).

Die Wiedergabe der HTML-basierten Repräsentation funktioniert mit dem Internet Explorer und den Mozilla-Browsern (Mozilla, Netscape, Firefox), aktiviertem JavaScript und RealPlayer-Plugin.

Die *SMIL*-Version enthält neben dem identischen Dozenten-Video dieselben Folien-Bilddaten als RealPix-Datenstrom. Dazu kommt ein Navigationsfeld mit den Folien-Überschriften (Abbildung 4-12).

Die Wiedergabe erfordert nur die Installation des RealOne Players. Die Darstellung fällt bedingt durch die fehlenden Vorschaubilder und das Suchfeld etwas kompakter aus. Dafür entfällt die damit verbundene Funktionalität.

Szenario 2

Im Szenario 2 besteht die Repräsentation aus den beiden parallelen Videodatenströmen, die mittels *SMIL* synchronisiert abgespielt werden können. Ohne weitere Nachbearbeitung fehlen sämtliche inhaltsbezogenen Navigationselemente. Der Betrachter kann sich somit lediglich an der Zeitachse orientieren.

Prinzipbedingt sind sämtliche Mausebewegungen sowie in der Präsentation eingesetzte Animationen auch in der Repräsentation enthalten. Die Bildqualität der Folien ist aufgrund der mehrfachen Signalwandlung durch Rauschanteile beeinträchtigt. Auch hier verhalten sich beide Medienströme im Rahmen der Toleranzgrenze von 500 ms synchron.

Die in den Repräsentationen enthaltenen Elemente bzw. Funktionen sind in Tabelle 5-2 zusammengefasst.

Repräsentation	Szenario 1		Szenario 2
	HTML	SMIL	SMIL
Dozenten-Video	ja	ja	ja
Folien	Images	Images	Slideshow-Video
wahlfreier Zugriff	ja	ja	ja
Gliederung, Folientitel	ja	ja	nein
Vorschaubilder	ja	nein	nein
Suchfunktion	ja	nein	nein

Tabelle 5-2: Vergleich der Repräsentationen der Testszenarios

Fazit aus der Auswertung

Die Produktion und das Ergebnis in Szenario 1 erfüllen die gestellten Erwartungen in jeder Hinsicht. Größtes Manko in Szenario 2 stellt die fehlende Navigation dar. Zwar besteht die Möglichkeit, die fehlenden Informationen während der Vorlesung mit der Annotationsfunktion des VaceEditors *manuell* zu ergänzen (vgl. Abschnitt 4.5.1: Echtzeit-Annotationen). Aber das VACE-Framework bietet aufgrund seines modularen Designs an dieser Stelle das Potenzial für Erweiterungen, um diese Aufgabenstellung *automatisiert* zu lösen. Hierzu sind an zwei Stellen Modifikationen erforderlich:

- Die Videoaufzeichnung der Folien ist durch die Erzeugung von Einzelbildern zu ersetzen. Anstelle eines kontinuierlichen Videostroms werden dann lediglich einzelne Bilder („Screenshots“) aufgezeichnet, wenn sich der Inhalt ändert. Diese Funktion ist z. B. in verfügbaren Programmen zur Videoüberwachung oder in Video-Schnittprogrammen als Szenenerkennung enthalten. Die erstellten Einzelbilder können in VACE importiert werden; eine entsprechende Funktion ist bereits im Prototyp vorhanden (vgl. Abschnitt 4.5.2). Durch die Zeitmarken der einzelnen Bilder ergibt sich nun bereits eine Gliederung in Zeitabschnitte.
- Mit VacePPT werden Titel und Inhaltsinformationen zu den einzelnen Folien mit Hilfe der Präsentationsapplikation extrahiert. Um analog dazu an inhaltliche Informationen aus den Einzelbildern zu gelangen, können diese mittels Texterkennung (OCR-Software) analysiert werden. Die so erhaltenen Informationen werden dann in Form von VACE-Metadaten gespeichert.

Alternativ kann die beschriebene Funktionalität in einem weiteren VACE-Capturing-Modul implementiert werden. Dessen Ausgabe in Form von VACE-konformen Metadaten und Standard-Einzelbildern entspricht der von VacePPT (vgl. Abschnitt 4.2.2 und Abbildung 4-1).

5.2 Performance-Messungen

Dieser Abschnitt beleuchtet die wichtigsten Performance-Parameter des VACE-Frameworks. Hierzu gehört zum einen die erreichbare Synchronisation der unabhängig voneinander aufgezeichneten Mediendatenströme. Desweiteren wird der Zeitaufwand für die Produktion untersucht.

5.2.1 Abweichungen der Synchronisation

Im Abschnitt 4.5.1 ist das implementierte Verfahren zur Ermittlung der Synchronisationsinformationen von in Echtzeit kodierten AV-Daten mit dem VACE RealProducer Wrapper VaceHxPW beschrieben. Die Exaktheit besitzt eine hohe Relevanz bezüglich der grundsätzlichen Nutzbarkeit des VACE-Frameworks.

Um die Funktionsfähigkeit dieses Verfahrens nachzuweisen, wird die Uhrzeitanzeige einer DCF-Funkuhr (vgl. Abschnitt 2.4.2) mit einer Videokamera aufgezeichnet und die Position (Frame-Nummer) eines Sekundenumsprungs ermittelt. Der damit berechenbare Startzeitpunkt der AV-Aufzeichnung wird mit dem vom Wrapper gelieferten Wert verglichen.

Für die Aufzeichnung kommt ein typisches PC-basiertes Encoding-System zum Einsatz (vgl. Abschnitt 2.2.2, Live-Encoding): Der Pentium-4-PC ist mit einer Brooktree-848-Videokarte, einer Creative-Audigy-Soundkarte sowie einer Fast-Ethernet-Karte ausgestattet. Als Betriebssystem ist Windows XP Professional installiert, als Real-Encoder fungiert der RealProducer Basic in der Version 10. Für die Zeit-Synchronisation mit dem NTP-Server wird ntptime für Windows [48] eingesetzt. Abbildung 5-4 zeigt den Versuchsaufbau.

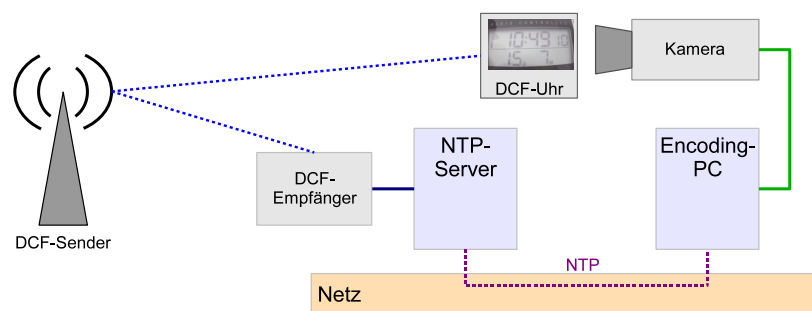


Abbildung 5-4: Systemaufbau zur Ermittlung der Synchronisationsgenauigkeit

Es wurde eine Reihe von 50 AV-Aufzeichnungen mit unterschiedlichen Auflösungen und Bitraten durchgeführt. Da der RealPlayer über keine exakte Anzeige der Position verfügt (es werden lediglich 1/10 Sekunden angezeigt), kommt für die Auswertung der freie Media Player Classic [33] zum Einsatz, der neben einer Standbildanzeige auch die Frame-genaue Positionierung ermöglicht. Durch sukzessives Ansteuern von Frames kann so die genaue Position eines Sekundenumsprungs ermittelt werden. Hiermit kann ferner nachgewiesen werden, dass innerhalb einer AV-Datei keine Abweichungen im Zeitverhalten („Gleichlaufschwankungen“) auftreten.

Die mit dem beschriebenen Verfahren ermittelten Werte liefern Abweichungen von maximal 350 ms bei einem Mittelwert von 225 ms. Die zur Synchronisation von AV-Daten und Folien im Abschnitt 2.4.3 erläuterte Toleranzgrenze von einer halben Sekunde wird somit eingehalten.

5.2.2 Zeitaufwand für die Nachbearbeitung

In vielen Anwendungsszenarios ist eine zeitnahe Fertigstellung der Repräsentation von hohem Interesse (vgl. Abschnitt 3.2.1). Häufig soll eine Standard-Aufzeichnung – wie auch in den Testszenarios in Abschnitt 5.1 – ohne weitere (manuelle) Nachbearbeitung publiziert werden. Abhängig von der Produktionskette stellt sich die Frage nach der Latenz – der Dauer vom Ende der Aufzeichnung bis zum möglichen Abruf durch den Betrachter.

Generell sind dabei drei Bereiche von Zeitaufwänden zu berücksichtigen: Datentransfer, Datenverarbeitung und manuelle Bearbeitung. Die Betrachtung umfasst die Ergebnisse aus den Testszenarios sowie potenziell anfallende weitere Arbeitsschritte.

Datentransfer

Innerhalb der Produktionskette (Abschnitt 4.2.1) werden die Mediendaten üblicherweise zweimal von einem System auf ein anderes kopiert: vom Aufzeichnungssystem zum Postproduktionssystem sowie vom Postproduktionssystem zum Distributionsserver. Dabei kann ein Kopierschritt entfallen, wenn die Nachbearbeitung auf dem Aufzeichnungsrechner stattfindet. In den Testszenarios aus Abschnitt 5.1 wurde beispielsweise der Encoding-PC für die Nachbearbeitung eingesetzt, wodurch lediglich die Daten der Folienaufzeichnung auf diesen Rechner übertragen werden mussten.

In der Tabelle 5-3 ist das Datenaufkommen für zwei typische Präsentationsaufzeichnungen zusammengestellt und auf eine Stunde umgerechnet. Zugrunde liegen hier die Daten aus dem Testszenario im Abschnitt 5.1.1. Nicht berücksichtigt sind die statischen Dokumente (HTML-Frameset, JavaScript- und CSS-Dateien). Der Umfang dieser Dateien beträgt unabhängig von der Gesamtlänge ca. 18 KByte und ist damit im Vergleich zum Gesamtvolumen zu vernachlässigen. Bei den Einzelbilddaten der Folien ergaben sich bei der Testaufzeichnung pro Stunde jeweils ca. 60 Dateien, entsprechend einer Folie pro Minute.

Daten	Inhalt	Bytes/s	Bytes/h	Anteil (%)
Video/Dozent und Einzelbilder (Szenario 1)				
Video (Dozent)	320*240 Pixel, 450 kbit/s	56.700	204.180.300	99,30
Bilder (Folien)	640*480 Pixel, PNG	284	1.024.900	0,50
Vorschaubilder	160*120 Pixel, PNG	96	347.700	0,17
HTML	Navigation, Vorschau, Suche	16	57.700	0,03

Tabelle 5-3: Datenvolumen von typischen Präsentationsaufzeichnungen

Daten	Inhalt	Bytes/s	Bytes/h	Anteil (%)
VACE Metadaten	Folien	7	25.700	0,01
Summe		57100	205.610.600	100
Video/Dozent und Video/Folien (Szenario 2)				
Video (Dozent)	320*240 Pixel, 450 kbit/s	56.700	204.180.300	56,29
Video (Folien)	480*360 Pixel, 350 kbit/s	44.000	158.565.100	43,71
Summe		100.700	362.745.400	100

Tabelle 5-3: Datenvolumen von typischen Präsentationsaufzeichnungen

Es zeigt sich, dass der Hauptanteil des Datenaufkommens bei den Videodaten liegt. Damit wird der Zeitbedarf für das Kopieren primär durch die AV-Daten bestimmt. Nebenbei wird die ineffiziente Verwendung eines Videoformats zur generischen Folienaufzeichnung in Szenario 2 deutlich.

Für die Datenübertragung in einem typischen 100 Mbit/s-Ethernet-LAN wurde für die Datenübertragung in Szenario 1 mit dem SMB-Protokoll eine Transferrate von 8 Megabyte/s ermittelt. Die Rate lag bei den Daten im Szenario 2 aufgrund der größeren Dateien etwas höher (ca. 9,7 Megabyte/s). In beiden Messungen waren Rechner und Netz lediglich durch die Datenübertragung belastet.

Legt man für die durchschnittliche Transferrate eine Abschätzung von 7,5 Megabyte/s zugrunde, so ergibt sich in Szenario 1 eine Übertragungszeit von rund 28 Sekunden, in Szenario 2 rund 48 Sekunden – jeweils für das Material einer Stunde Aufzeichnung (Tabelle 5-4).

Szenario	Datenvolumen	Transferdauer (SMB, 7,5 MByte/s)	Transferdauer (scp, 1 MByte/s)
Szenario 1	206 Megabyte	27,5 s	206 s
Szenario 2	363 Megabyte	48,4 s	363 s

Tabelle 5-4: Zeitaufwand für die Datenübertragung in den Testszenarios

Während der Produktion werden im lokalen Netz typischerweise unverschlüsselte Verbindungen benutzt. Für die Übertragung der AV-Daten auf einen Streaming-Server können jedoch verschlüsselnde Übertragungs-Protokolle eingesetzt werden, die aufgrund des höheren Rechenaufwands deutlich längere Übertragungszeiten verursachen. Die – hier nicht näher betrachteten – scp- oder sftp-Verbindungen arbeiten auf aktuellen Systemen mit Transferraten von ca. 1 Megabyte/s.

Datenverarbeitung (Post-Processing)

Der Aufwand für die Datenverarbeitung nach der Aufzeichnung ist abhängig von den eingesetzten Medienformaten und der Art der Aufbereitung (vgl. Abschnitt 4.2.1). Die Arbeitsschritte lassen sich in drei Bereiche aufteilen:

- **Metadaten-Verarbeitung.** Der Aufwand für die Verarbeitung der VACE-Metadaten besteht primär in der Zusammenführung der Datensätze aus den verschiedenen Aufzeichnungen. Je Event (Start/Stop der Videoaufzeichnung, Folienwechsel, siehe Abschnitt 4.4.3) wird ein Datensatz verarbeitet. Die hierfür erforderlichen Zeiten liegen bei einer typischen Produktion im Bereich von Sekundenbruchteilen und sind gegenüber den Dateitransfer- und Mediendatenverarbeitungszeiten vernachlässigbar.
- **Mediendaten-Verarbeitung.** Hier gibt es drei Bereiche, in denen Aufwände anfallen können:
 - **Metadaten-Extraktion.** Zur nachträglichen Ermittlung von inhaltsbezogenen Metadaten aus den aufgezeichneten Mediendaten können externe Programme eingebunden werden, z. B. OCR-Programme zur Texterkennung. Diese Funktion wurde nicht implementiert, sodass hier keine Messwerte vorliegen. Zur Abschätzung genügt eine separate Betrachtung der Performance der einzusetzenden Programme. Ferner kann die Extraktion bereits während der Aufzeichnung stattfinden, sodass keine zusätzlichen Nachbearbeitungszeiten anfallen.
 - **Konvertierung.** In typischen Produktionen (wie in den Beispielen im Abschnitt 5.1.1) wird bereits bei der Aufzeichnung das Zielformat für die Repräsentation erzeugt, sodass weitere Formatkonvertierungen nicht anfallen. Sollten AV-Konvertierungen nötig sein – üblicherweise zur Erstellung von Versionen mit geringerer Datenrate – so ist der Zeitaufwand abhängig vom Codec und dem eingesetzten Rechner. Erfahrungswerte bewegen sich im Bereich von der 0,2- bis 3-fachen Echtzeitlänge des AV-Clips, sodass hier relativ hohe Post-Processing-Zeiten auftreten können.
 - **Erzeugung von Vorschaubildern.** Für die Repräsentation können für die Gesamtübersicht Vorschaubilder (Thumbnails) erstellt werden. Hierzu werden mit einer Grafiksoftware die erzeugten Einzelbilder der Folien skaliert. Im Beispielszenario benötigte die Thumbnail-Erzeugung für 14 PNG-Bilder weniger als eine Sekunde.
- **Export der Repräsentation.** Auch bei der Erzeugung der Repräsentation in HTML oder SMIL werden lediglich Metadaten verarbeitet. Die VACE-Metadaten werden vom VaceEditor in das entsprechende Format umgesetzt. In den Testszenarios war auch dieser Schritt in weniger als einer Sekunde erledigt.

Zusammenfassend ist festzuhalten, dass nennenswerte Zeitaufwände lediglich bei der Verarbeitung der Mediendaten anfallen (Konvertierung von AV-Daten sowie der nachträglichen Extraktion von Metadaten). Der Zeitbedarf für die Verarbeitung der Metadaten ist hingegen zu vernachlässigen.

Manueller Aufwand (Post-Production)

Da in dem vorliegenden Prototyp noch nicht sämtliche Funktionen zur Automatisierung implementiert sind, wurden im Test einige Schritte manuell ausgeführt (siehe Abschnitt 5.1.2). Da es sich hierbei jedoch lediglich um den Aufruf von Befehlen im Menu des VaceEditors handelt, ist auch die Zeit hierfür vernachlässig-

bar. Die benötigten Funktionen sind in der Online-Hilfe dokumentiert. Diese interaktiv aufgerufenen Arbeitsschritte bewirken keine rechenaufwendigen Operationen, sodass keine Wartezeiten anfallen.

Möglich sind an dieser Stelle manuelle Änderungen der Folientitel und -beschreibungen sowie diverser anderer Metadaten. Mit dem implementierten Editor ist der Aufwand für die Eingabe vergleichbar mit dem Ausfüllen bzw. Bearbeiten von typischen Windows-Formularen (vgl. Abschnitt 4.5.3).

Zusammenfassung

Der größte Zeitaufwand für Standard-Anwendungen (wie den Test-Szenarios) fällt bei der Datenübertragung der AV-Daten an. Typische Präsentationen mit üblichen AV-Daten verursachen dabei Transferzeiten von rund 2 Minuten. Ein Großteil der Anwendungen wird hiermit abgedeckt.

Sollen AV-Daten konvertiert werden, so ist hier mit einem erheblichen Zeitaufwand zu rechnen, der (abhängig von den erläuterten Randbedingungen) im Bereich von einigen Minuten bis hin zu einigen Stunden liegt. Konvertierungen können automatisiert im Batch-Betrieb durchgeführt werden, sodass keine personelle Betreuung notwendig ist.

Der Zeitbedarf für die verbleibende manuelle Nachbearbeitung (wie das Editing der Metadaten zu einzelnen Folien) ist von der individuellen Anwendung abhängig und damit schwer abzuschätzen. Mit den vorliegenden Tools wird der Aufwand jedoch auf die reine Texteingabe beschränkt.

Vergleicht man die Zeitaufwände mit rein manuellem Vorgehen, so sind die Einsparungen erheblich. Für die nachträgliche Zuordnung von Vortragsfolien zu einer AV-Aufzeichnung sowie der Erstellung entsprechender Verknüpfungen von Hand benötigt man etwa die Dauer der Aufzeichnung in Echtzeit.

Im Vergleich zu anderen Produktionssystemen ist festzustellen, dass die systemimmanenten Verarbeitungszeiten gegenüber den Transferzeiten zu vernachlässigen sind. Lediglich wenn AV-Konvertierungen oder Metadaten-Extraktionen notwendig werden, kann sich die Postproduktion nennenswert verzögern.

5.3 Zusammenfassende Bewertung und Systemvergleich

Im Abschnitt 3 wurden Unzulänglichkeiten existierender Systeme identifiziert, die mit einem neuartigen System umgangen werden sollten. Die Tabelle 5-5 bietet einen Überblick über die (in Abschnitt 3.6 ausgewählten) existierenden Systeme im Vergleich mit dem hier entwickelten VACE-System.

Der Vergleich zeigt, dass das entwickelte System die Flexibilität von generischen Ansätzen mit der Effizienz von Autorensystemen verbindet.

- Das VACE-Framework ist aufgrund seines modularen Aufbaus in unterschiedlichen Szenarios einsetzbar. Effizienz, Aufwand und Abbildungsqualität hängen dabei vom konkreten Einsatzfall ab. Keines der verfügbaren Systeme verfügt über offene Schnittstellen für die Erweiterung auf individuelle Anwendungen.

- Sämtliche Systeme sind auf eine hohe Benutzerfreundlichkeit für den Dozenten ausgelegt. Im Falle einer klassischen AV-Aufzeichnung (UniTV) entsteht gar kein Mehraufwand. Die Systeme, die den Dozenten einbinden, erfordern üblicherweise die Nutzung eines „Start-/Stopp“-Applets für die Aufnahmesteuerung. Dies trifft auch auf VACE zu, sofern die entsprechenden Module (VaceBrowser, VacePPT) eingesetzt werden. VACE kann jedoch auch mit Modulen arbeiten, die für den Dozenten im Hintergrund bleiben und so keinerlei Einfluss auf den Vortragenden nehmen. Der Mehraufwand für die Produktion gestaltet sich als minimal (vgl. Abschnitt 5.1.2).
- Besonderer Wert wurde auf die Nutzung von verfügbaren Standards gelegt, nicht zuletzt, um mit hoher Kompatibilität eine hohe Nutzerfreundlichkeit in allen Produktionsabschnitten zu ermöglichen.
- Durch den primären Fokus auf die Navigation der Repräsentation arbeitet VACE mit einer Granularität auf der Ebene von Präsentationsabschnitten. Feinere Auflösungen (wie die einzelnen grafischen Elemente auf einem Whiteboard) werden nicht implizit erfasst. Anwendungen, die diese Funktion zwingend erfordern, sind mit Systemen wie AoF oder Lecturnity besser zu lösen – auf Kosten der Benutzerfreundlichkeit.
- Die AV-Qualität ist mittels der im VACE-Framework eingebundenen Systeme weitgehend skalierbar. Weitere Codecs können bei Bedarf integriert werden.

System		Real, MS	Lec- turnity	AoF	UniTV	VACE
Flexibilität	Szenarios	-	o	o	+	+
	Erweiterbarkeit	-	-	o	-	+
Benutzer- freundlichkeit	Dozent	+	+	+	++	+/++
	Produktion				-	+
	Betrachter	+	+	o	o	+
Abbildungs- Qualität	Granularität	o	+	+	-	o
	Audio/Video	o	o	o	+	o..+

Tabelle 5-5: Systemvergleich/Bewertung^a

a. +++=sehr gut, +=gut, o=mittel, -=schlecht

Die in Abschnitt 3.8 gesetzten Ziele (Tabelle 3-10) konnten mit dem Systementwurf erreicht werden und mit der prototypischen Implementierung in der Praxis nachgewiesen werden.

6 Zusammenfassung und Ausblick

6.1 Zusammenfassung

Ziel der Arbeit ist die Produktion von Präsentationsaufzeichnungen, wobei die effiziente Transformation der Live-Präsentation in eine universell verwendbare Repräsentation im Mittelpunkt steht.

Zunächst wird ein Überblick über die aktuellen Infrastrukturen geliefert, die in den letzten Jahren in vielen Hochschulen zu diesem Zweck aufgebaut wurden. Sie bilden damit die Plattform für die Produktion.

Die Anforderungsanalyse der Einsatzszenarios fördert ein breites Spektrum von Anwendungen zu Tage. Gleichzeitig werden die Defizite existierender Systeme deutlich, wenn es um die Lösung von Anforderungen jenseits der üblichen Standards geht. Diese Betrachtung liefert die Motivation für den Entwurf eines universellen Systems, das sich an unterschiedliche Anwendungen und Rahmenbedingungen anpassen lässt. Diese Anforderungen umfassen:

- Flexibler, adaptiver Einsatz in unterschiedlichen Anwendungsszenarios.
- Effiziente, zeitnahe Produktion einer standardkonformen, multimedialen Repräsentation mit Navigationsfunktionen und Metadaten-Unterstützung.
- Integration von Standards für die Präsentation, Produktion und Distribution.
- Verteiltes, an aktuelle Produktionsinfrastrukturen angepasstes System.

Das in dieser Arbeit zur Erfüllung dieser Anforderungen entwickelte Framework bietet die Basis für die effiziente Produktion der Aufzeichnung von Vorlesungen, Vorträgen, Kursen und Tagungen in E-Learning- und Dokumentations-Anwendungen. Bei der Entwicklung des Frameworks wurden Synergien aus der Konzeption und dem Aufbau der in Abschnitt 2.2 erläuterten Medientechnik und Infrastruktur genutzt.

Die Bewertung demonstriert die Einsatz- und Funktionsfähigkeit des Systems im produktiven Betrieb. Der Ansatz des mittels NTP synchronisierten verteilten Aufzeichnungssystems übertrifft die für die Anwendung notwendige Toleranzanforderung. Alltägliche Aufzeichnungsszenarios können weit gehend automatisiert und ohne personalintensive Nachbearbeitung produziert und zeitnah publiziert werden.

Es werden in allen Schritten der Produktionskette Standard-Systeme integriert, sodass eine flexible Anpassung an aktuelle und zukünftige Anwendungsszenarios und Formate möglich ist. Zudem erlaubt die alleinige Basierung des Produktions-Frameworks auf Standards, Metadaten und synchronisierten Uhren die plattform-unabhängige Implementierung von Modulen.

Im Vergleich zu existierenden Systemen sind insbesondere die folgenden Eigenschaften herauszustellen:

- Die modulare, offene, erweiterbare Systemarchitektur ermöglicht Erweiterungen beziehungsweise Anpassungen auch für individuelle Anforderungen.
- Der verteilte Ansatz integriert sich nahtlos in aktuelle Produktionsumgebungen.
- Die intensiv auf den Metadaten-Einsatz abgestellte Produktion liefert eine effizient verknüpfte multimediale Repräsentation mit umfangreichen Navigationsfunktionen und vereinfacht die Einbindung in Content-Management-Systeme.

Zudem erzeugt VACE die inhaltlichen und Synchronisations-Metadaten, die für den sinnvollen Einsatz integrierender Medienformate wie z. B. SMIL notwendig sind, bisher aber keine Unterstützung durch Aufzeichnungssysteme fanden.

6.2 Ausblick

Neben der Integration weiterer Capturing- und Export-Module bietet das VACE-Framework diverse Erweiterungsmöglichkeiten. Neben systeminternen Funktionen wie einer zentralen oder Event-basierten Steuerung der einzelnen Capturing-Module über das Netz ist auch die Nutzung in neuen Anwendungsgebieten realisierbar. Die folgenden Absätze geben dazu einige Beispiele.

Live-Übertragungen und Video-Conferencing

Eng im Kontext mit Vortragsaufzeichnungen stehen die Bereiche *Live-Übertragungen* (z. B. von Vorlesungen) und *Video-Conferencing* (z. B. für verteilte Veranstaltungen oder „Sprechstunden“). Während bei der Live-Übertragung keine Postproduktion stattfinden kann, wohl aber identische Systeme zur Übertragung eingesetzt werden können, stellt das Video-Conferencing höhere Anforderungen an eine kurze Latenz, um interaktive Kommunikation zu ermöglichen.

Allen Bereichen gemeinsam ist jedoch die Notwendigkeit der Übermittlung von zusätzlichen Datenströmen, da die verfügbare Videoqualität allein nicht ausreicht.

Sind bestimmte Randbedingungen erfüllt, ist auch die Erweiterung des Frameworks für den Live-Einsatz möglich. Hierzu gehört etwa, dass die Medienformate in Echtzeit generiert und übertragen werden können – oder bereits verteilt verfügbar sind, z. B. in Form von HTML-Seiten.

Live-Annotation durch Zuhörer

Eine denkbare Erweiterung des Systems besteht in der Einbeziehung von Anmerkungen durch (Live-) Zuschauer. Da das System beliebig viele verschiedene Medienströme parallel erfassen kann, bietet sich die Möglichkeit, auch Annotationen vom Auditorium als zusätzliche Informationsquelle zu verarbeiten.

So ist es möglich, dass sich ein Studierender (z. B. mit Hilfe eines PDAs) während der Präsenzveranstaltung Abschnitte „markiert“, an denen er sich später die Ausführungen in der Aufzeichnung noch einmal ansehen will. Umgekehrt könnte der Dozent mit Hilfe solcher Informationen erfahren, an welchen Abschnitten noch

zusätzlicher Erklärungsbedarf besteht. Das System könnte damit für Online-Evaluationen ausgebaut werden.

Einbindung individueller Präsentationswerkzeuge

Die Aufzeichnung multimedialer Präsentationen mit nutzerspezifischen Visualisierungswerkzeugen (z. B. ein 3D-Viewer) kann mittels VACE-Capturing-Modulen realisiert werden. Analog zur Nutzung von Vortragsfolien wird die Anwendung in die Aufzeichnung und Wiedergabe nativ eingebunden; in Präsentation und Repräsentation kann somit das identische Medienformat eingesetzt werden [27].

Evaluation verteilter Systeme

Zur Untersuchung des Nutzerverhaltens bei der Anwendung interaktiver verteilter Systeme werden parallele Aufzeichnungen benötigt, die das Verhalten der Benutzer während der Evaluation festhalten. Aufgezeichnet werden hier neben den Interaktionen mit der Anwendung der Ton und das Bild der Benutzer. Hier kann VACE als Basis für die Synchronisation der Medienströme zum Einsatz kommen. Mittels eines angepassten Viewers können so die einzelnen AV-Aufzeichnungen zusammen mit der Aufzeichnung der Anwendungsdaten in der Auswertung synchronisiert wiedergegeben werden [28].

Literaturverzeichnis

- [1] ACM Computer Classification Scheme: <http://www.acm.org/class/>
- [2] Annodex: Open standards for annotating and indexing networked media. <http://www.annodex.net>
- [3] Apple Computer, Inc.: Inside QuickTime – Interactive Movies; The QuickTime Technical Reference Library, Oktober 2002
- [4] Authoring on the fly: <http://ad.informatik.uni-freiburg.de/mmggroup/aof/>
- [5] Burnett, I. et al.: MPEG-21: Goals and Achievements; IEEE MultiMedia 10-12/2003, p.60-70 (<http://www.chiariglione.org/mpeg/events&tutorials/MMpaper.pdf>)
- [6] Chen, T.; Ma, M.; Meinel, C.; Schillings, V.: A Novel Approach to E-Learning Content Creation. m-ICTE 2003 Badajoz (Spain), Proceedings pp. 410-414
- [7] Dublin Core Metadata Initiative: <http://dublincore.org/>
- [8] Dublin Core DCMI Frequently Asked Questions (FAQ) 2001-2004: <http://dublincore.org/resources/faq/>
- [9] Einhorn, R.: Realisierung einer „Online-Preview“-Komponente für die Videoproduktion in Breitbandnetzen. Studienarbeit, Lehrgebiet Rechnernetze und Verteilte Systeme, Universität Hannover, 1996
- [10] Einhorn, R.; Olbrich, S.; Nejdil, W.: A Metadata Model for Capturing Presentations. 3rd IEEE International Conference on Advanced Learning Technologies (ICALT), Athens, 9-11.07.2003, Proceedings pp.110-114
- [11] Einhorn, R., Olbrich, S., Jensen, N., Chmielewski, K.: Aufbau und Entwicklung von Multimedia-Infrastrukturen und -Services für innovative E-Learning-Anwendungen. 17. DFN-Arbeitstagung über Kommunikationsnetze, 11.-13.06.2003, Düsseldorf. GI-Edition „Lecture Notes in Informatics“ (LNI), Köllen Verlag, Bonn, S.541-567
- [12] Esteve, M.; Palau, C.; Guerri, J.C.; Jobmann, K; Niemann, F.; Tuchs, K.-D.: Synchronous and Asynchronous Teleteaching: Comparison of two Experiences. International Symposium on Telemedicine and Teleeducation in Practice, ISTEP 2000, March 2000, Kosice, Slovakia
- [13] DVD, Interactive Functions: http://www.mpeg.org/MPEG/DVD/Book_B/Intfunc.html (Pioneer)

- [14] Hilt, V., Mauve, M., Kuhmünch, C., Effelsberg, W.: A Generic Scheme for the Recording of Interactive Media Streams. International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services 1999 (IDMS'99), S. 291-304, Toulouse, France, 1999
- [15] Hinrichs, R. J: Technology, Learning and Scholarship in the Early 21st Century Microsoft Research, Learning Science and Technology, Microsoft, Inc. 2002
- [16] IMS Global Learning Consortium, Inc. <http://www.imsproject.org>
- [17] Lecturnity, Produkt-Webseiten: <http://www.im-c.de>
- [18] Liebchen, T.: MPEG-4 Lossless Coding for High-Definition Audio. 115th AES Convention, New York, 10/2003.
- [19] IEEE P1484.1/D9: Draft Standard for Learning Technology; Learning Technology Systems Architecture (LTSA). Chair: Tyler, J.; Technical Editors: Farance, F.; Tonkel, J.; 30.11.2001
- [20] IEEE 1484.12.1-2002 Draft Standard for Learning Object Metadata (LOM). <http://standards.ieee.org/reading/ieee/std/learning/1484.12.1-2002.pdf>, http://ltsc.ieee.org/doc/wg12/LOM_1484_12_1_v1_Final_Draft.pdf, 15.07.2002
- [21] iL2 – interactive Learning Lab, Informationsseite. <http://www.learninglab.de/iL2/>
- [22] IrfanView, Bildbetrachter für Windows. <http://irfanview.tuwien.ac.at/>
- [23] ISO/IEC 10918: Information technology – Digital Compression and Coding of Continuous-Tone Still Images
- [24] ISO/IEC 11172-1/2/3: Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Systems, Audio, Video.
- [25] ISO/IEC 13522-1: Information Technology – Coded Representation of Multimedia and Hypermedia Information Objects (MHEG), 1997.
- [26] ISO/IEC 13818-1/2/3: Information processing - Generic coding of moving pictures and associated audio information – Systems, Audio, Video
- [27] Jensen, N.; Einhorn, R.; von Voigt, G.: A Supply-Chain for Computer-Mediated Communication and Visualization. Poster, WSCG Conference 2004, Plzen, Czech Republic, 02.-06.02.2004.
- [28] Jensen, N.; Seipel, S.; von Voigt, G.; Raasch, S.; Olbrich, S.; Chmielewski, K.; Einhorn, R.; Nejd, W.: Heuristic Evaluation of a Virtual Lab System. Technical Report L3S Q4 2003. VASE 3, L3S, University of Hanover
- [29] Kataria, V.: BITE: Blackboard Integration Tool for ELAN. Vortrag, Institut fuer Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover, 25.07.2003

- [30] Krüger, M.; Siegmund, G.: E-Learning gestern und heute. Neue Wege für Training und Weiterbildung. NET – Zeitschrift für Kommunikationsmanagement, 56. Jahrgang, 11/2002, S. 23-25
- [31] Learning-Space: Multimedia-Seminarraum im L3S: http://www.rrzn.uni-hannover.de/fileadmin/multimedia/pdf/l3s_mm_seminarraum.pdf
- [32] Ma, M.; Schillings, V.; Chen, T.; Meinel, C.: T-Cube: A Multimedia Authoring System for eLearning. Proceedings of E-Learn 2003, 7.-11.11.2003, Phoenix, Arizona, USA, S. 2289-2296
- [33] Media Player Classic, guliverkli Project: <http://sourceforge.net/projects/guliverkli/>
- [34] MIME Media Types: <http://www.iana.org/assignments/media-types>
- [35] Microsoft Producer for Powerpoint: <http://microsoft.com/office/powerpoint/producer/>
- [36] Microsoft Windows Media Codecs: <http://www.microsoft.com/windows/windowsmedia/9series/codecs/video.aspx>; Codecs 101: <http://www.microsoft.com/windows/windowsmedia/howto/articles/codecs.aspx>
Übersicht: Microsoft Windows Media: <http://www.microsoft.com/windows/windowsmedia/default.aspx>
- [37] Miranda Technologies Inc.: Digital Television Standards; Production, Compression and Transmission Formats. <http://www.miranda.com/library/en/DigitalStandards/>
- [38] MPEG-4 Overview: ISO/IEC JTC1/SC29/WG11 N4668. 03/2002, Editor: Koenen, R.
- [39] MPEG-7 Overview: ISO/IEC JTC1/SC29/WG11 N4980. Klagenfurt, 07/2002, Editor: Martinez, J. M. (UPM-GTI, ES)
- [40] MPEG-21 Overview: ISO/IEC JTC1/SC29/WG11/N5231. Shanghai, 10/2002, Editor: Bormans, J.; Hill, K.
- [41] Müller, R.; Ottmann, T.: The „Authoring on the Fly“ system for automated recording and replay of (tele)presentations. Multimedia Systems 8 Issue 3 (2000) S. 158-176
- [42] Mukhopadhyay, S.; Smith, B.: Passive Capture and Structuring of Lectures. Proceedings of ACM Multimedia '99, S. 477-487, <http://bmrc.berkeley.edu/projects/lb/lb-mm99.pdf>
- [43] Multimedia Labor, RRZN, Universität Hannover: Beispiele zur Einbindung von AV in HTML: <http://www.mml.uni-hannover.de/embedding/>, 29.09.2003
- [44] Naegele-Jackson, S., Gräve, M., Eschbaum, N., Holleczeck, P.: Distributed TV Productions and Video-on-Demand Services at Universities (UniTV). TERENA Networking Conference 2000, Lisbon, Portugal, 05/2000

- [45] National Standards Institute for Standards and Technology, NIST Physics Laboratory, Time and Frequency Division: Configuring Windows 2000 and Windows XP to use NIST Time Servers. <http://www.boulder.nist.gov/timefreq/service/pdf/win2000xp.pdf>, 26.11.2002
- [46] Nielsen/Netratings (Chan, C.): Nielsen/NetRatings Launches New Web Multimedia Format Report, Tracking Realmedia, Windows Media and Quicktime, http://direct.www.nielsen-netratings.com/pr/pr_020620.pdf), 20.06.2002
- [47] NTP FAQ and HowTo: Understanding and using the Network Time Protocol, Editor: Windl, U. et al. <http://www.ntp.org/ntpfaq/NTP-a-faq.htm>, 09.04.2004,
- [48] NTPTIME for Windows: <http://home.att.net/~Tom.Horsley/ntptime.html>
- [49] Ottmann, T., Pomplun, R.: e-Learning im Hochschulbereich: Kritische Erfolgsfaktoren. 17. DFN-Arbeitstagung über Kommunikationsnetze, 11.-13.06.2003, Düsseldorf. GI-Edition „Lecture Notes in Informatics“ (LNI), Köllen Verlag, Bonn, S.371-378
- [50] Ottmann, T.: Erfahrungen mit VIROR und ULI, Workshop Virtuelle Hochschulen, 21. November 2002, Leipzig. http://www.ft-informatik.de/docs/workshops/ws_virt_hs/ottmann.ppt
- [51] PADLR: Personalized Access to Distributed Learning Repositories, Projekt-Webseiten. <http://www.l3s.de/padlr/>
- [52] Poynton, C.: Chroma subsampling notation. http://www.poynton.com/PDFs/Chroma_subsampling_notation.pdf
- [53] Real Networks: RealSystem iQ Production Guide. <http://service.real.com/help/library/guides/realone/ProductionGuide/PDF/ProductionGuide.pdf>
- [54] Real Presenter: <http://www.realnetworks.com/products/>
- [55] Regionales Rechenzentrums für Niedersachsen (RRZN): Multimedia-Studio, <http://www.rrzn.uni-hannover.de/avstudio.html>
- [56] RFC 1305: Network Time Protocol (V.3), Specification, Implementation and Analysis. Mills, D. L., 03/1992
- [57] RFC 1889: RTP: A Transport Protocol for Real-Time Applications. Schulzrinne, H. et al.; 01/1996
- [58] RFC 2046: Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types. Freed, N.; Borenstein, N.; 11/1996
- [59] RFC 2326: Real Time Streaming Protocol (RTSP). Schulzrinne, H. et al.; 04/1998
- [60] RFC 2413: Dublin Core Metadata for Resource Discovery. Weibel, S.; Kunze, J.; Lagoze, C.; Wolf, M.; 09/1998

- [61] RFC 2616: Hypertext Transfer Protocol HTTP/1.1. Fielding, et al.; 06/1999
- [62] RFC 2854: The 'text/html' Media Type. Connolly; Masinter 06/2000
- [63] Richardson, I. E. G.: Video Codec Design. Wiley, 2002
- [64] Rui, Y.; He, L.; Gupta, A.; Liu, Q.: Building an Intelligent Camera Management System ACM Multimedia 2001 Conference Proceedings, S. 2-11, Ottawa, Canada, 10/2001
- [65] Schierz, C.: Mensch-Computer Interaktion, Sehen und Bildschirm. Zürich, 29.10.2001, http://e-collection.ethbib.ethz.ch/ecol-pool/lehr/lehr_64.pdf
- [66] Sharable Content Object Reference Model (SCORM) Version 1.2. Editor: Dodds, P.; ADL (www.adlnet.org)
- [67] Staatliches Baumanagement Niedersachsen, Oberfinanzdirektion Hannover, Landesbauabteilung (Hrsg.): Multimedia-Räume an den Niedersächsischen Hochschulen – Hinweise zur baulichen Umsetzung. 2., überarbeitete Auflage, März 2002. www.ofd.niedersachsen.de/functions/downloadObject/0,,c1408265_s20,00.pdf
- [68] Stanford Center for Professional Development: <http://scpd.stanford.edu/scpd/default.htm>
- [69] Steinmetz, R.: Multimedia-Technologie: Grundlagen, Komponenten und Systeme. 3., überarbeitete Auflage, Springer, 2000.
- [70] Steinmetz, R.: Human Perception of Jitter and Media Synchronization. IEEE Journal on Selected Areas in Communications, Vol. 14, 01/1996, S. 61-72
- [71] Sühling, K.; Schwarz, H.; Wiegand, T.: Effizienter kodieren. c't Magazin für Computer Technik 2003, Heft 6
- [72] Tanenbaum, A.; van Steen, M.: Verteilte Systeme, Grundlagen und Paradigmen, Prentice Hall, Pearson Studium, 2003
- [73] TechSmith Screen Capture Codec: <http://www.techsmith.de/products/studio/codec.asp>
- [74] Truong, K. N.; Abowd, G. D.; Brotherton, J. A: Who, What, When, Where, How: Design Issues of Capture & Access Applications. Lecture Notes In Computer Science archive, Proceedings of the 3rd international conference on Ubiquitous Computing table of contents, Atlanta, Georgia, USA, 2001, Springer-Verlag, London, S.209-224
- [75] Uni-TV Homepage: <http://www.uni-tv.uni-erlangen.de/>
- [76] Universität Hannover: Multimedia-Ausstattung im Audimax-Hörsaal: <http://www.rrzn.uni-hannover.de/avaudimax.html>
- [77] ViDe User's Guide: Dublin Core Application Profile for Digital Video. Editors: Agnew, G.; Kniesner, D. http://www.vide.net/workgroups/videoaccess/resources/vide_dc_userguide_20010909.pdf

- [78] VirtPresenter: <http://www.virtuos.uni-osnabrueck.de/projekte/virtpresenter.php>
- [79] VNC, Virtual Network Computing: <http://www.realvnc.com>
- [80] W3C: Cascading Style Sheets, level 2, CSS2 Specification, W3C Recommendation 12.05.1998, <http://www.w3.org/TR/1998/REC-CSS2-19980512>; Editor: Bos, B. et al.
- [81] W3C: Date and Time Formats, Note, <http://www.w3.org/TR/1998/NOTE-datetime-19980827>, Wolf, M., Wicksteed, C.
- [82] W3C: Document Object Model (DOM) Level 2 HTML Specification, Version 1.0. W3C Recommendation, 09.01.2003, <http://www.w3.org/TR/2003/REC-DOM-Level-2-HTML-20030109>, Editors: Stenback, J. et al.
- [83] W3C: Scalable Vector Graphics (SVG) 1.1 Specification. W3C Recommendation, 14.01.2003, <http://www.w3.org/TR/SVG/>, Editor: Ferraiolo, J.
- [84] W3C: Synchronized Multimedia Integration Language (SMIL 2.0). W3C Recommendation 07.08.2001, <http://www.w3.org/TR/2001/REC-smil20-20010807/>
- [85] W3C: XML Schema Part2: Datatypes. W3C Recommendation 02.05.2001, Editors: Biron, P. V.; Malhotra, A. <http://www.w3.org/TR/xmlschema-2/#dateTime>
- [86] Wolf, P.; Steinmetz, A.: How to capture a conference – using the LectureLounge system to grasp the conference experience. 2002. <http://www.ipsi.fraunhofer.de/~pwolf/docs/howToCaptureAConference.pdf>
- [87] Zongker, D. E.; Salesin, D. H.: On Creating Animated Presentations. Eurographics/SIGGRAPH Symposium on Computer Animation (2003)
- [88] Zota, V.: Kompressionisten – Aktuelle Video-Codecs im Vergleich. c't Magazin für Computer Technik 2003, Heft 10, Seite 146-159