

Video in Vorlesungsaufzeichnungen mit informatikfernen Inhalten am Beispiel Archäologie

Th. Ottmann*, R. Müller*, G. Seitz†, Ch. Steinert*

* Institut für Informatik
† Abteilung für Provinzialrömische Archäologie
Universität Freiburg

Zusammenfassung

Wir berichten über eine langjährige Zusammenarbeit zwischen Informatik und Provinzialrömischer Archäologie, die sowohl Rückwirkungen auf die archäologische Wahrheitsfindung hatte als auch zu neuen Herausforderungen an die Informatik geführt hat. Konkret wurden ein Ausgrabungsfund (die *Villa Urbana* in Heitersheim) mit Mitteln der Computergraphik modelliert, ein ca. 5-minütiger Film eines virtuellen Rundganges durch den Gebäudekomplex berechnet, Bilder und Film in eine Telepräsentation integriert und diese nach dem "Notetaking-Verfahren" in ein multimediales Dokument verwandelt. Es zeigt sich, daß für eine sinnvolle Nutzung neuer Medien in der Archäologie Standardtechniken der Informatik nicht ausreichen, sondern insbesondere im Bereich der Synchronisation von Medienströmen neue Techniken benötigt werden.

Abstract

We report on a long-term cooperation between computer science and archaeology which led to new archaeological insights and new challenges in computer science. In consequence of this cooperation a Roman mansion (the *Villa Urbana* in Heitersheim near Freiburg) was modeled based on the excavation using methods of computer graphics. As a result a movie, approximately 5 minutes long, was rendered describing a round tour through the building. Images and the movie were integrated into a telepresentation which was recorded using a note-taking method and transformed into a multimedia document. It turned out that standard techniques in computer science are often not sufficient when using new media in domains such as archaeology. Especially, new techniques for inter-stream synchronization of media are required.

1 Einleitung

Ziel der archäologischen Forschung ist es, uns heute ein möglichst umfassendes Bild vergangener Kulturepochen zu vermitteln. Wir zeigen hier an einem konkreten Beispiel (der *Villa Urbana* in Heitersheim bei Freiburg), was das bedeutet und wie mit Methoden und Werkzeugen der Informatik ein Beitrag zur Erfüllung dieses Zieles geleistet werden kann. Teilweise kann dabei auf bekannte Techniken zurückgegriffen werden. Im Zuge der Integration alternativer kontinuierlicher Medien, wie etwa Video und Audio, für Inhalte der Archäologie wird jedoch auch klar, daß neue Lösungen gebraucht werden, wenn diese Medien in Forschung und Lehre sinnvoll genutzt werden sollen.

Die traditionelle archäologische Arbeitsweise gliedert sich in die grundlegenden Bereiche Ausgrabung (Datengewinnung), Auswertung (Datenbewertung), Publikation (Datenaufbereitung) und Präsentation (Datendarstellung). Die Zusammenarbeit zwischen Archäologie und Informatik, über die hier berichtet wird, bezog sich fast ausschließlich auf den Bereich *Präsentation*. Anstelle einer Rekonstruktion des Fundobjektes durch eine technische Zeichnung sollte ein *realitätsnäheres* Computermodell erstellt werden.

Während es heute bei der Planung und beim Bau neuer Gebäude vielfach üblich ist, vorab ein computeranimiertes 3-d-Modell zu erstellen, das einen Blick in die später realisierte Zukunft erlaubt, gehört

der entsprechende Blick in die Vergangenheit mit Hilfe virtueller Realität noch nicht zur allgemein üblichen Form der Darstellung archäologischer Forschungsergebnisse. Ausgehend von dem eher spielerischen Wunsch, ein für ein mächtiges 3-d-Modellierungswerkzeug geeignetes Übungsobjekt zu finden, ist so über mehrere Jahre hinweg ein immer wieder revidiertes und mit Hilfe eines professionellen Werkzeugs erstelltes 3-d-Modell der Villa Urbana in Heitersheim entstanden. Die zwei, daraus berechneten Filme von je etwa 5 Minuten Länge bieten virtuelle Rundgänge durch das Gebäude und vermitteln einen der historischen Realität möglichst nahe kommenden Eindruck.

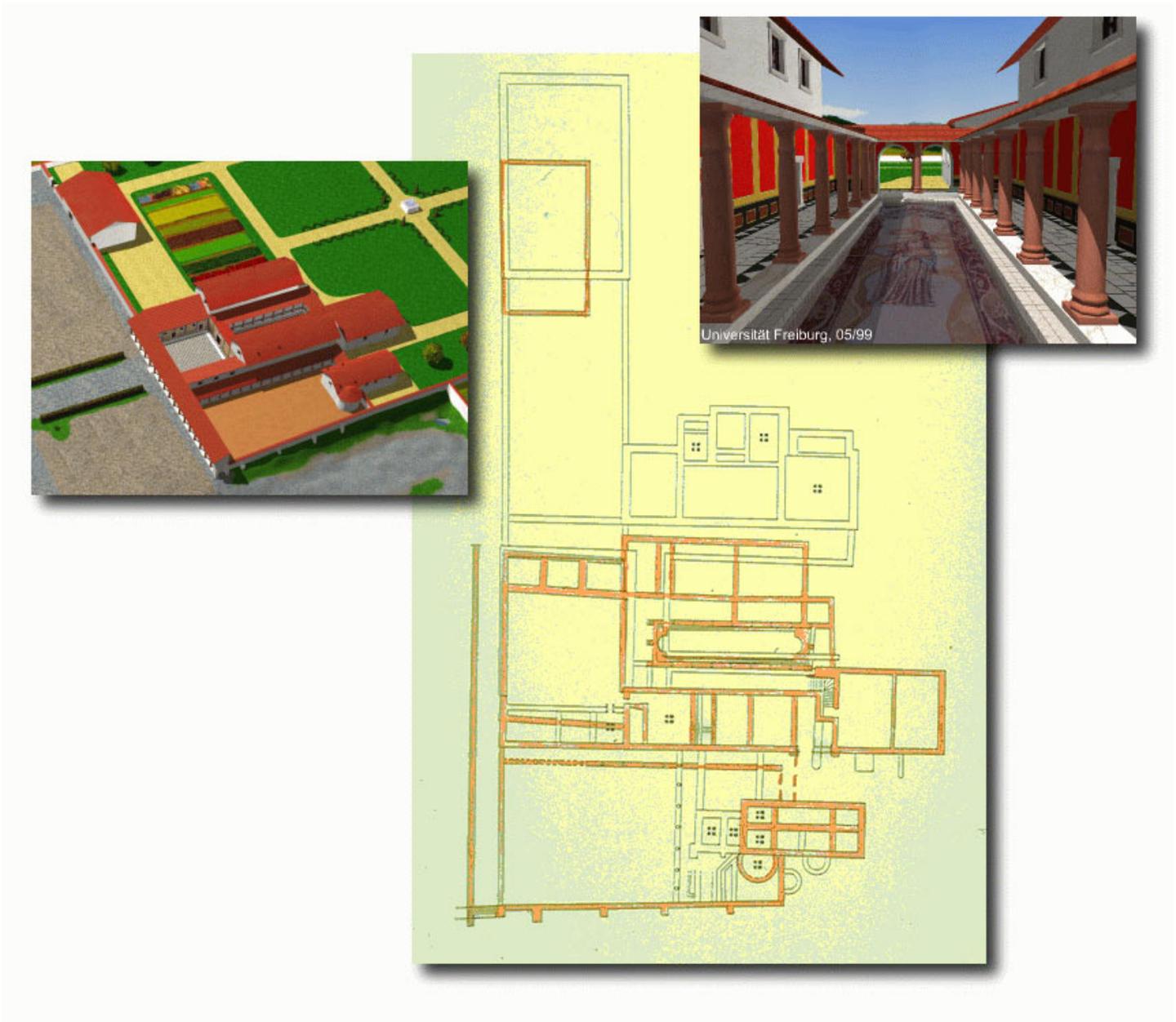


Abbildung 1: Grundriss und Sichten auf die Villa Urbana

Die Filme ebenso wie eine Reihe von Einzelbildern und vielleicht sogar das 3-d-Modell selbst sollen in einer den Ausgrabungsfund dokumentierenden, permanenten Ausstellung in einem Schutzhaus über der Ausgrabungsstätte integriert werden. Darüberhinaus wurden Bildmaterial und Film für eine im Rahmen

des VIROR-Verbundprojekts [20] durchgeführte und nach dem AOF-Verfahren [14] aufgezeichnete Telepräsentation verwendet. Dabei zeigte sich, daß die Nutzung solcher neuen Medien in der archäologischen Lehre von den zur Zeit verfügbaren Werkzeugen nur unzureichend unterstützt wird. Wir skizzieren im folgenden zunächst das Ausgrabungsprojekt selbst. Wir berichten dann über die Modellierung und Filmberechnung und berichten schließlich über die in einem konkreten „Lehrexperiment“ (Telepräsentation eines Archäologen) gewonnene Erfahrung. Wir gehen insbesondere auf die Frage ein, welche Anforderungen an die Annotierbarkeit von Video und die Synchronisation solcher oder ähnlicher Medienströme resultieren.

2 Das Ausgrabungsprojekt

Als ein wichtiger Schritt zur Klärung des Wirkens der Römer am Oberrhein hat sich der Befund einer Villa Urbana in Heitersheim (ca. 20 km südlich von Freiburg gelegen) erwiesen. Dieses Ausgrabungsprojekt wird von der Abteilung für Provinzialrömische Archäologie der Universität Freiburg wissenschaftlich betreut.

Die Ausgrabungen in Heitersheim haben Reste eines bislang einzigartigen Villenkomplexes als Zeugnis römischer Besiedlung aus der Zeit zwischen 30 und 260 n. Chr. rechts des Rheins zutage gefördert. Leider sind nur Teile erhalten geblieben, die unter dem ursprünglichen ebenerdigen Geländeniveau liegen, also Grundmauern und Kellerräume. Hinzu kommen zahlreiche Kleinfunde unterschiedlicher Gattungen (Münzen, Schmuck, etc.). Dennoch war es möglich, aus den Grabungsergebnissen recht genaue Pläne der verschiedenen Gebäude dieses Villenkomplexes zu erstellen und Rückschlüsse auf die Gebäudeart und Bebauung zu verschiedenen Zeitpunkten zu ziehen (vgl. [15] [16] [18] [10]).

Auf der Basis dieser Pläne und Funde wurde nun bereits vor mehreren Jahren damit begonnen, ein 3-d-Modell des Gebäudekomplexes mit Hilfe des Werkzeuges *Softimage* [19] zu erstellen. Ziel war und ist es, eine möglichst realistische Vorstellung dieses einzigartigen Kulturdenkmals zu vermitteln. Dazu mußten natürlich viele nicht mehr vorhandene oder unbekannte Details durch Analogieschlüsse aus anderen Funden derselben Zeit ergänzt werden. Während konventionelle Methoden der Darstellung durch Modelle, Zeichnungen, Rekonstruktion im Maßstab 1:1 sehr material- und kostenintensiv sind und eine zu einem bestimmten Zeitpunkt gültige Forschungsmeinung festschreiben, liegt der Vorteil einer rechnergestützten Modellierung in der größeren Flexibilität. Weil es sich bei dem Heitersheimer Ausgrabungsprojekt um ein noch laufendes Forschungsprojekt handelt, war es immer wieder nötig, das Modell an neue Erkenntnisse anzupassen. Das war zu Beginn des Projektes (im Jahre 1995) so nicht vorhersehbar. Und so wundert es nicht, daß das im ersten Versuch entworfene 3-d-Modell als ad hoc Entwurf den späteren Anforderungen, also den durch Erkenntnisfortschritt notwendigen Anpassungen nicht gerecht wurde. Daher wurde die Modellierung wiederholt, aber von vornherein auf übersichtliche Struktur und Änderungsfreundlichkeit geachtet. Dieses Modell ist nun Schritt für Schritt erweitert und verbessert worden. Es war nicht nur ein interessantes Studienobjekt im Rahmen der Informatikausbildung, um interessierten Studenten ein mächtiges Werkzeug zur 3-d-Modellierung, Verfahren der photorealistischen Computergraphik, virtuelle Realität und außerordentlich rechenintensive Rendering-Verfahren nahe zu bringen. Es hat auch den Archäologen immer wieder neue Einsichten verschafft, weil durch die Teilmodellierung und Einarbeitung von aktuellen Erkenntnissen Widersprüche aufgedeckt oder neue Funde erst richtig eingeordnet werden konnten. Mitte 1997 lag dann ein erster 5-minütiger Film vor, der bereits eine sehr anschauliche Vorstellung der Heitersheimer Siedlungstelle vermittelt. Wir beschreiben im nächsten Abschnitt den aktuellen Stand der Modellierung als Grundlage für die Berechnung eines zweiten Films, die Festlegung der Kamerafahrt und Animation, das Rendering und den Export der Daten in ein für die Videovorführung geeignetes Format.

3 Modellierung und Filmberechnung

3.1 Modellierung

Als Basis für die Modellierung und für den geplanten virtuellen Besichtigungsrundgang dient ein Plan der älteren Steinbauphase des Hauptgebäudes und der weiteren Gebäude des Villenkomplexes. Ferner lag eine Rekonstruktionszeichnung herkömmlicher Art vor. Daraus wurde ein erstes grobes Modell erstellt, das nur die wichtigsten Maße der Gebäude berücksichtigte. Wände, Dächer, Böden, Säulen usw. wurden als

einfache CSG-Bäume (*Constructive Solid Geometry*) realisiert. Dies sind Bäume mit geometrischen Elementarobjekten an den Blättern und Mengenoperatoren an den inneren Knoten. Viele Einzelheiten blieben zunächst offen, die später, jeweils nach Rücksprache mit den Archäologen, ergänzt wurden. Das betraf beispielsweise die Position und Größe von Fenstern und Türen, Texturen für Wände, Dächer und Böden, die Modellierung der Außenanlagen und die Einbettung des Gesamtkomplexes in die Rheinebene vor dem Hintergrund des Schwarzwaldes. Neben der Objektmodellierung bestimmen auch andere Parameter wie Lichtquellen, Schattenwurf, Reflexionsverhalten und Transparenz von Flächen u. a. darüber, ob schließlich ein realitätsnaher Eindruck entsteht. Auch hier wurden immer wieder Korrekturen vorgenommen. Besonderes Augenmerk wurde auf die Modellierung eines vollständig erhaltenen und inzwischen freigelegten Kellerraumes gelegt. Denn hier besteht die Möglichkeit der Prüfung des Modells an der (noch vorhandenen) Realität.

Das Ergebnis ist ein 3-d-Modell, das insgesamt aus ca. 1.300 Elementen, die als CSG-Bäume erstellt wurden, 150.000 Polygonen, 475 verschiedenen Materialien und 100 Grundtexturen besteht.



Abbildung 2: Einzelbild der Villa Urbana als Ergebnis des gerenderten 3-d-Modells

3.2 Kamerapfad und Animation

Für den virtuellen Rundgang mußte nach Abschluß der Modellierung ein "Kamerapfad" festgelegt werden. Er besteht aus zwei Kurven im dreidimensionalen Raum, wobei die erste Kurve die Bewegung der Kamera beschreibt, während die zweite Kurve die Blickrichtung der Kamera zum betrachteten Zeitpunkt angibt. Auf diese Weise können für den Betrachter ganz unterschiedliche optische Eindrücke erzeugt werden, die beispielsweise Ruhe oder Hektik im Filmgeschehen vermitteln.

Die Animationen werden mit dem Programm Softimage durch das Festlegen sogenannter Keyframes erstellt. Ein Frame beschreibt den Zustand der Szene zu einem bestimmten Zeitpunkt. Will man ein Objekt animieren, so weist man ihm zu bestimmten Zeitpunkten Keyframes, d.h. etwa die gewünschte Position

oder Form an diesen Zeitpunkten zu. Die Berechnung zwischen den Keyframes übernimmt dann wieder Softimage. Hier kann mit Hilfe von sogenannten FCurves (function curves) vorgegeben werden, wie der Übergang von der einen vorgegebenen Position/Form zur anderen zu erfolgen hat. Legt man die FCurve z.B. als eine Gerade fest, so geschieht die Transformation linear; das würde beispielsweise eine Bewegung mit einer konstanten Geschwindigkeit erzeugen. Diese Kurven können jedoch beliebig modifiziert werden, um beispielsweise das Abbremsen oder Beschleunigen eines animierten Objekts zu beschreiben.

3.3 Rendering

Für einen Film von 5 Minuten Länge, der mit 25 Bildern pro Sekunde (*frames per second, fps*) dargestellt wird, müssen 7.500 Einzelbilder berechnet werden. Da für die Präsentation des zu berechnenden Films die analogen Formate VHS bzw. S-VHS vorgegeben waren, wurden Einzelbilder im PAL-Format, d. h. im Seitenverhältnis 4:3 mit einer Auflösung von 768×576 und einer Farbtiefe von 24 *bit* bei einer Frame-Rate 25 *fps* berechnet. Hierzu wurde das Programm MentalRay [13], das Teil des Softimage-Paketes ist, benutzt. Das Programm Softimage produziert für jedes Bild eine Szenenbeschreibung, aus der der Renderer das entgültige Bild berechnet. Weil die Berechnung eines einzelnen Bildes ca. 3 bis 5 Minuten Zeit beansprucht, die Einzelbilder aber völlig unabhängig voneinander berechnet werden können, wurden gleichzeitig mehrere Rechner zur Berechnung eingesetzt. Die berechneten Originalbilder mit einem Gesamtdatenvolumen von ca. 9,7 *GByte* wurden auf CDs archiviert.

Es ist offensichtlich, daß eine direkte Wiedergabe einer solchen Sequenz unkomprimierter Originalbilder bei 25 *fps* auch auf leistungsstarken Rechnern mit sehr schnellen Bild- und Plattenspeichern nicht realisierbar ist. Dies würde einen Datentransfer von ca. 25 bis 30 *MByte/s* vom Speichermedium auf den Bildwiederholpeicher des Rechners erfordern. Eine Kompression der Originaldaten ist daher für eine kontinuierliche, rechnerbasierte Wiedergabe unvermeidlich. Dabei ist ein angemessener Kompromiß zwischen der Darstellungsqualität und dem dafür erforderlichen Datenvolumen anzustreben. Der Export in das analoge S-VHS-Videoformat erfolgte über ein spezielles Bildplattenmedium. Das Ergebnis ist optisch zwar akzeptabel, führt aber doch zu sehr deutlichen Qualitätseinbußen im direkten Vergleich mit der Originalbildfolge. Es soll daher untersucht werden, ob ein qualitativ besseres Ergebnis erreicht werden kann, wenn aus der Originalbildfolge einen MPEG-2-Film berechnet und dieser über einen DVD-Player (und Großbildprojektor) abgespielt wird. Berechnet wurde bisher aus den Originalbildern ein MPEG-1-Film mit einer Auflösung von 320×240 und ein Film im AVI-Format mit einer Auflösung von 768×576 .

Die für die Filmberechnung gewählte Auflösung von 768×576 ist für große Ausdrücke von Standbildern (DIN A4 und größer) zu gering. Daher wurden einige Einzelbilder mit sehr hoher Auflösung (3000×2250) zusätzlich berechnet. Hier zeigte sich, daß für die Auswahl von Einzelbildern (Ansichten) die Möglichkeit des Zugriffs auf sämtliche Originalbilder hilfreich, die Möglichkeiten zur Navigation in der Bildfolge, also die Identifizierung einer Ansicht aber nur unzureichend unterstützt wurden.

4 Telepräsentation und Notetaking

4.1 Live-Szenario

Im Mai 1999 fand im Rahmen der Ringvorlesung zum Verbundprojekt VIROR [20] ein Vortrag des das Ausgrabungsprojekt betreuenden Archäologen, Prof. H. Nuber, mit dem Titel „Die Römer am Oberrhein“ statt, der an die an VIROR beteiligten Standorte übertragen und aufgezeichnet wurde. Man kann diese Telepräsentation als durchaus typischen Fall eines vor allem auf Bildmaterial und ein Video abgestützten Vortrages ansehen. Im Vortrag wurde insbesondere auf die Heitersheimer Villa Urbana eingegangen. Der Vortrag wurde mit Hilfe eines SmartBoards (interaktiver Großbildschirm) mit Rückprojektion gehalten und mittels der MBone Tools *vic* und *vat* [12] in Bild und Ton übertragen. Für die spätere Wiedergabe wurde der Vortrag nach dem „Authoring on the Fly“-Prinzip [14] aufgezeichnet. Der Vortragende verwendet dabei die AOF-Whiteboard-Applikation (*AOFwb*, [11]) für das Aufblättern und Annotieren der Bilder und Starten und Stoppen des oben beschriebenen MPEG-Films gehalten. Über die integrierte Aufzeichnungsfunktionalität wird die Stimme des Vortragenden und dessen Aktionen auf dem Whiteboard festgehalten [14]. Die dadurch entstehende Aufzeichnung bezeichnen wir als *AOF-Dokument*. Dieselben Werkzeuge

wurden bisher in nahezu 100 ähnlich verlaufenen (Tele-)Präsentationen eingesetzt, die allerdings nahezu ausschließlich Informatikinhalt hatten. In den meisten Fällen wurde für den Vortrag auch nicht, wie in diesem Fall, das SmartBoard benutzt, sondern der Vortrag direkt am Rechner (mit Bildschirm, Tastatur und Maus) gehalten. Es handelte sich hier also um einen Testfall für einen sehr informatikfernen Inhalt, der von einem Geisteswissenschaftler ohne Routine in der Handhabung der für die Telepräsentation benutzten Werkzeuge gehalten wurde.

Wir erläutern nun, welche Besonderheiten gegenüber entsprechenden Telepräsentationen mit Informatikinhalt sichtbar wurden und welche Folgerungen daraus für die Weiterentwicklung der benutzten Werkzeuge zu ziehen sind.

4.2 Benutzungsschnittstelle für den Dozenten

Ein Vortrag, der vom gesprochenen Wort getragen, und von zahlreichen Bildern (wie Dias) unterstützt wird, führt häufig zu sehr wenig Aktivität auf dem Whiteboard. Zwar wäre in vielen Fällen eine direkte Annotation (Beschriftung, Markierung) eines in der Whiteboard-Applikation aufgeblättern Bildes oder die Verwendung eines Zeigers sinnvoll. Jedoch haben ungeübte, diese Vortragsart nicht gewohnte Benutzer ganz offensichtliche Schwierigkeiten, die zahlreichen von Whiteboard-Applikationen angebotenen Annotationswerkzeuge zu gebrauchen. Zur Verfügung stehen in der Regel neben einfachen geometrischen Objekten auch Telepointer, Freihandlinien und Text. Die Konzentration auf den verbal vorgetragenen Inhalt kann die Nutzung dieser vielfältigen Interaktionsmöglichkeiten deutlich erschweren. Gebraucht wurde in diesem Fall ausschließlich der Telepointer. Doch selbst dessen Handhabung setzt in der Regel eine gewisse Übung bzw. Routine in der Verwendung voraus. So ist im Fall des SmartBoard hinreichender Druck auf das Whiteboard erforderlich, um einen Mouse-Klick zu simulieren und damit den Telepointer zu aktivieren, was ebenfalls leicht vergessen wird. Auch die Auswahl und das Aufblättern der Bilder von der mit einem Rollbalken versehenen Thumbnail-Übersicht am linken Rand der in diesem Fall verwendeten Whiteboard-Applikation AOFwb ist offenbar nicht so intuitiv, wie etwas das Drücken eines „Weiter/Zurück“-Knopfes eines Dia-Projektors. Obwohl die Verwendung des SmartBoards für eine Rechnerpräsentation gegenüber einer Rechnerbedienung mit Bildschirm, Tastatur und Maus dem in diesem Gebiet üblichen und gewohnten Lehrscenario deutlich eher entspricht, zeigte sich doch, daß auch das SmartBoard ohne besondere Anpassung der verwendeten Software noch nicht ideal ist: Die Übertragung der Desktop-Metapher auf das SmartBoard liefert eine Benutzerschnittstelle, die diesem Interaktionsmedium nicht angemessen ist, d. h. Funktionen, die bei Interaktion mit Bildschirm, Tastatur und Maus intuitiv und einfach zu bedienen sind, müssen in anderer Weise realisiert werden. Denn bei einer auf Bildmaterial abgestützten Rechnerpräsentation sind nur ganz wenige, aber einfach und intuitiv nutzbare Annotationsmöglichkeiten (nur Annotation mit Freihandlinien und ggfs. zusätzlicher Telepointer) sinnvoll. Wichtig sind gute Thumbnail-Übersichten und Blättermöglichkeiten in der Whiteboard-Applikation, die durch geeignete Hardware, etwa in Form eines drahtlosen „Vor/Zurück“-Schalters unterstützt werden.

Bei der Vorführung und Diskussion des MPEG-1-Filmes gegen Ende des Vortrags wurde allerdings deutlich, daß hier eine andere Annotationsmöglichkeit sinnvoll einsetzbar gewesen wäre: Zur Erläuterung hätte der Dozent gern auf einer dem laufenden Film überlagerten, transparenten Ebene Annotationen und Markierungen (mit Freihandlinien) vorgenommen. Ferner würde man es sich wünschen, den im Film gezeigten (virtuellen) Rundgang durch die Villa Urbana gleichzeitig auf einer in der Whiteboard-Applikation aufgeblättern Grundrisskarte markieren und verfolgen zu können. Alles das sollte natürlich zeitlich synchron mit dem Audiostrom (also den verbalen Erläuterungen des Archäologen) erfolgen.

Es sollten also neben den bei Telepräsentationen üblichen Datenströmen (Audio, Video, Whiteboardaktionen) noch der MPEG-Strom und ein Strom von Annotationen dieses Stroms aufgezeichnet werden.

4.3 Bedeutung des Videobildes des Dozenten

Bei der (Tele-)Präsentation von Informatikinhalt kommt es vor allem darauf an, die Whiteboard-Aktionen und den Audiostrom des Dozenten zu synchronisieren. Das Videobild („Talking Head“) des Dozenten spielt, wie zahlreiche Versuche zeigen, für das Verständnis kaum eine Rolle, wenn der Dozent seinen Vortrag direkt am Rechner mit Bildschirm, Tastatur und Maus hält. Denn dann sind Bewegungsfreiheit und Gestik des Dozenten ohnehin stark eingeschränkt. Die Situation ändert sich bereits, wenn der Dozent das

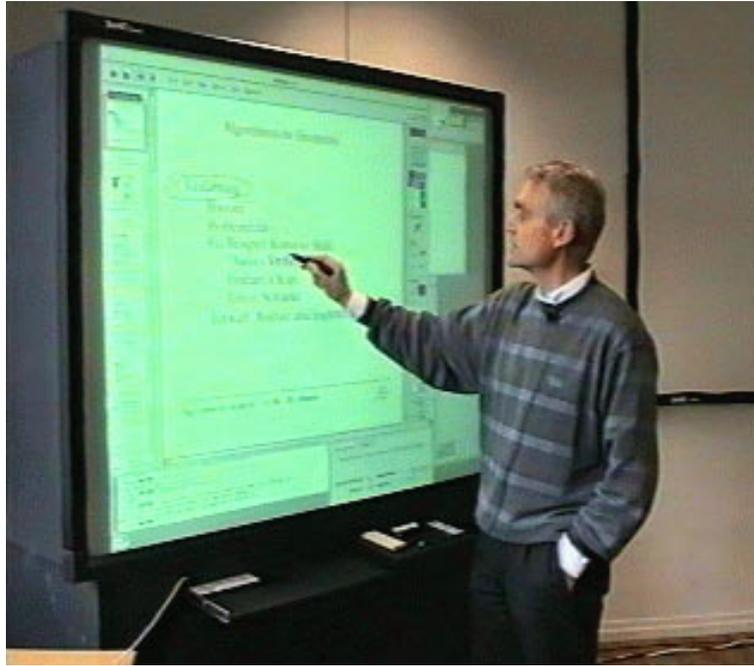


Abbildung 3: SmartBoard: Ein interaktiver Großbildschirm mit Rückprojektion, der die Integration des Videobilds in Vorlesungsaufzeichnungen notwendig macht.

SmartBoard benutzt. Ob in diesem Fall der mit dem Whiteboard- und Audiostrom synchronisierte Videostrom einen wesentlichen Beitrag zu Verständnis von Informatikinhalten liefert, ist bisher allerdings noch ungeklärt. Im Falle des Archäologen kann das Video des Vortragenden offensichtlich durch seine statt der Whiteboard-Werkzeuge zur Erläuterung intuitiv benutzte Gestik sehr wesentliche Bedeutung haben. Sie sollte daher bei der Übertragung an entfernte Standorte sehr gut sichtbar sein und auch in der Aufzeichnung, also bei der Wiedergabe, zur Verfügung stehen. Für die Übertragung des Videobildes des Vortragenden reicht daher ein H.261/H.263-Videostrom (vic [7]) bei geringer Framerate sicher nicht aus. Wir haben dies experimentell beim Vortrag dadurch überprüft, daß eine PAL-Übertragung des Videostromes zusätzlich zur Übertragung des Whiteboard-Stromes in einen dem Vortragsraum benachbarten Hörsaal in Freiburg realisiert wurde. Das Ergebnis ist signifikant besser als eine ausschließlich auf den MBone Tools basierende Übertragung. Zu beachten ist, daß dabei die getrennten Projektionen des Whiteboards und des Dozentenbildes beide in derselben Anordnung erscheinen wie im Vortragssaal. Wenn also der Dozent in der Regel rechts neben dem SmartBoard steht und nach links auf Bilder zeigt, sollte die Videoausgabe auch rechts von der Whiteboard-Ausgabe an den Empfängerstationen angeordnet werden.

Man sieht sofort, daß hieraus zahlreiche Anforderungen an die Übertragungsqualität und die Synchronisierung verschiedener Datenströme resultieren, die für ein skalierbares Multicast-Szenario im Internet gegenwärtig nicht realisierbar sind. Wir verweisen in diesem Zusammenhang auch auf die an der Universität Erlangen-Nürnberg gemachten Erfahrungen [3].

4.4 Die Wiedergabe und der wahlfreie Zugriff

Bei der Integration von Video in Vorlesungsaufzeichnungen, sei es das Videobild des Dozenten oder einzeln eingestreute Videofilme wie etwa der vorher beschriebene MPEG-Film über die Villa Urbana, ergeben sich zusätzliche Anforderungen für die Wiedergabe. Aus der Sicht der Lehrenden (Vortragende, Dozenten, Lehrer) und der Lernenden (Auditorium, Studenten, Schüler) können dabei stets grundlegende Ziele der Präsentationsaufzeichnung festgestellt werden, die zu diesen Anforderungen führen [14]. So soll die Wiedergabe einer Präsentationsaufzeichnung eine möglichst hohe Ähnlichkeit zur Originalveranstaltung, insbesondere also hohe Vollständigkeit des konservierten Informationsgehalts und hohe zeitliche **Synchronität**

der einzelnen aufgezeichneten Medien, aufweisen.

Um die bei der Aufzeichnung einer Vorlesung entstehenden Medienströme (z.B. Audio, Video, Whiteboard) in Echtzeit und synchron zueinander wiedergeben zu können, müssen entweder alle Ströme in einem *Multiplexing*-Schritt in einem gemeinsamen Strom zusammengeführt werden (z.B. *MPEG-1 System*) oder aber jeder einzelne Strom während der Aufzeichnung mit zeitrelevanter Information (Zeitstempeln) versehen werden. Bei kontinuierlichen Datenströmen, wie unkomprimiertes Audio und Video, sind diese Informationen implizit durch Parameter wie Abtastrate und Größe der Abtastwerte (Frames/Samples) gegeben. Andere Datenströme wie etwa der Whiteboard-Strom müssen explizit mit solchen Zeitstempeln versehen werden [2]. Ist die Granularität dieser Zeitstempel fein genug, dann sind die grundsätzlichen Voraussetzungen für eine synchrone Wiedergabe der aufgezeichneten Datenströme gegeben.

Ein weiteres grundlegendes Ziel aus der Sicht der Lehrenden und Lernenden ist die Flexibilität in der Handhabung und der Weiterverwendung solcher Präsentationsaufzeichnungen. So sind abgesehen von der einfachen Wiedergabe (Start, Stop, ggf. sequentielles Vorspulen) einer aufgezeichneten Vorlesung komfortablere Zugriffs- und Navigationsmöglichkeiten für einen weiten Einsatzbereich solcher Aufzeichnungen unerlässlich. Für eine flexible Weiterverwendung sind beispielsweise bei der Strukturierung von Aufzeichnungen in Kollektionen (Archiven), bei der Integration einzelner Aufzeichnungen in beliebige Lehrsoftware bzw. multimediale Lern-/Lehrbibliotheken oder bei der Verknüpfung solcher Aufzeichnungen mit relevanten Zusatzmaterial oder Metadaten (etwa zu multimedialen Lehrbüchern) flexible Integrationstechniken erforderlich. Der schnelle **wahlfreie Zugriff** auf eine zusammengehörige Menge von Medienströmen – also die Verfügbarkeit der zugehörigen Daten aus allen Strömen für einen gegebenen beliebigen Zeitpunkt, sowie deren Darstellung auf dem Bildschirm bzw. hörbare Ausgabe über Lautsprecher innerhalb einer sehr kurzen medienabhängigen Zeitspanne (vgl. [9] für eine genaue Definition) – ist für viele dieser Navigationsmechanismen und Integrationstechniken zwingend erforderlich [14].

Random Visible Scrolling ist ein Beispiel für einen komfortablen Navigationsmechanismus in AOF-Dokumenten, der auf schnellem wahlfreiem Zugriff basiert. Der Benutzer kann hierbei entlang der Zeitachse des Dokument navigieren während gleichzeitig alle visuellen Daten angezeigt werden. Diese Navigation ist unabhängig von Geschwindigkeit und Richtung der Bewegung. Der Benutzer bekommt auf diese Weise sehr schnell einen Überblick über den präsentierten Inhalt und kann gesuchte Stellen leichter lokalisieren. Es erinnert in gewisser Weise an das schnelle Durchblättern eines Buches in der Art eines Daumenkinos. Aber auch der wahlfreie Zugriff auf signifikante Stellen der Aufzeichnung wie Folienwechsel, Startpunkte von Videoausschnitten oder den Beginn eines neuen Abschnitts sind für das Studium einer aufgezeichneten Vorlesung durchaus akzeptanzfördernd und für die Verbindung dieser Dokumente mit anderem relevantem Unterrichtsmaterial in hypermedialen Lernumgebungen sogar zwingend.

Strom-inhärente Kriterien, wie die durchschnittliche Daten- oder Bitrate (z.B. 1,2 Mbit/s bei MPEG) oder eine Kompression entlang der Zeitachse zur Reduktion zeitlicher Redundanzen, also eine δ -Enkodierung (z.B. *Backward/Forward Prediction* bei MPEG), können einen schnellen wahlfreien Zugriff verhindern. Auf der anderen Seite ist aufgrund der anfallenden Datenmengen eine Kompression des Videostroms unerlässlich. So erzeugt beispielsweise ein unkomprimiertes Video in einer Auflösung von 352×288 (CIF), einer Farbtiefe von 24 bit und 25 fps eine Datenrate von ca. $60,8 \text{ Mbit/s}$. Selbst ein MPEG-1-Strom mit einer Datenrate von bis zu $1,5 \text{ Mbit/s}$ erhöht die übliche Datenrate für AOF-Dokumente ohne Video (ca. 130 Kbit/s) ggf. um mehr als den Faktor 10. Hierbei tritt der klassische Zwiespalt zwischen δ -enkodierten Datenströmen und dem wahlfreien Zugriff zu Tage. Auf den ersten Blick erfordert der Zugriff auf die Daten in einem Medienstrom, daß für jeden vorgegebenen Zeitpunkt die vollständige Information, um zum Beispiel den zugehörigen Videoframe darstellen zu können, an einer eindeutig und schnell zu bestimmenden Position im Strom verfügbar ist. Um einen schnellen Zugriff zu gewährleisten sollte die Darstellung der Daten an dieser Position dabei möglichst ohne Rückgriff auf andere vorhergehende oder nachfolgende Daten erfolgen. Das führt zwangsläufig zu einer Redundanz, da Informationen, die über mehrere Abtastpunkte gleich bleiben, jeweils in allen zugehörigen Werten vollständig abgespeichert werden müssen. Dies trifft auf alle kontinuierlichen, visuellen Medien zu, da hier die Abstände zwischen zwei für den Benutzer sichtbaren Veränderungen in der Regel weniger als $1/25$ Sekunde betragen. Man stelle sich als Extrembeispiel ein unkomprimiertes Video vor, bei dem sich über mehrere Sekunden einfach nichts verändert. Jeder Frame entspricht dann dem vorhergehenden. Der Sinn einer δ -Enkodierung liegt nun gerade darin, diese Redundanzen zu beseitigen.

Die mittlerweile in Bezug auf diese Anforderungen, also *zeitliche synchrone Wiedergabe, schneller*

wahlfreier Zugriff und grafische Annotation, von uns gefundenen Lösungen werden für MPEG-1 im Anhang A skizziert.

5 Zusammenarbeit mit der Archäologie

Charakteristisch für das hier beschriebene Projekt war die intensive Zusammenarbeit zwischen Systemfachleuten aus der Informatik und den das Ausgrabungsprojekt betreuenden Wissenschaftlern der Archäologie über einen langen Zeitraum hinweg. Die Archäologen gaben den Modellierern die Vorstellung, wie das Gebäude wohl einmal ausgesehen haben könnte. Weil nur die Grundmauern durch die Grabung verlässlich rekonstruierbar waren, waren für alle weiteren Teile des Gebäudes plausible Vorgaben der Archäologie nötig. Daß die Umsetzung in das Modell wiederum Rückwirkungen auf die archäologischen Vorstellungen hatte, mag folgendes Beispiel belegen. Aus der Freilegung eines Kellerraumes konnte geschlossen werden, daß dieser Raum an den Seiten jeweils zwei ebenerdige Fenster hatte. Daher kann der darüber gelegene Raum nicht ebenfalls ebenerdig auf gleicher Höhe mit allen anderen Räumen gewesen sein. Die Konsistenz des Modells erzwang eine besondere Erklärung. Archäologische Überlegungen ergaben, daß es sich bei diesem über dem Keller liegenden Raum wohl um eine Art Festsaal gehandelt haben muß, also ein Ort für Vergnügungen und Feierlichkeiten des Hausherrn. Dieser Raum enthielt vermutlich ein sogenanntes Triclinium, auf dem Gastgeber und Geladene zu Tische lagen. Dieser Raum war daher gegenüber den anderen um einige Eintrittsstufen höher gelegt.

Auch bei der Modellierung der Umfassungsmauer um den Siedlungskomplex stellte sich heraus, daß die ursprüngliche Vermutung über ihre Höhe revidiert werden mußte: Die Modellierung ergab, daß nur eine die Geländestruktur berücksichtigende und nicht zu hohe Mauer mit der Vorstellung eines Villenkomplexes verträglich ist. Insgesamt zeigte sich an vielen Details, daß auch die Archäologen ihre durch die Ausgrabung begründeten Vorstellungen durch die Computer-Modellierung häufig wesentlich präzisieren und insgesamt plausibler zusammenfügen konnten.

6 Beteiligte und Danksagung

Im Verlauf der Jahre haben an verschiedenen Stadien der Modellierung und Filmberechnung zahlreiche Personen mitgewirkt, deren Vorarbeiten in das nun vorliegende Ergebnis eingeflossen sind. Besonderer Dank geht an Steffen Bunzel, der bei allen Fragen, die das Programm betrafen, stets geholfen hat. Beteiligt waren ferner Jörg Finger, Christian Wetzel, Ronny Fehling, und Thomas Loop. Technische Hilfe leistete auch Frank Dal-Ri. Die Konvertierung des Filmes auf Videoband wurde mit Mitarbeitern am Lehrstuhl für Angewandte Mathematik durchgeführt. Großer Dank gilt auch Oliver Klerx für die Hilfe bei der Implementierung des MPEG-Players.

A Videosynchronisierung unter wahlfreiem Zugriff und grafischer Annotation

A.1 Wahlfreier Zugriff bei MPEG-1

Die δ -Enkodierung oder Inter-Frame-Kodierung eines Videostroms wie bei MPEG hat in der Regel eine variable Bitrate zur Folge. Die Position zu einem gegebenen Zeitpunkt im Strom beim wahlfreien Zugriff läßt sich also nicht einfach aus der Frame-Rate und der Byte-Größe der Frames berechnen.

Ein naheliegender Ausweg, der in sehr vielen Ansätzen in nahezu gleicher Form genannt oder verfolgt wird, liegt daher in der nachträglichen Indizierung des Stroms (vgl. etwa [4][5][6][8]). Dabei wird der Byte-Strom einmalig vor dem Start der Wiedergabe von Anfang bis Ende sequentiell durchsucht und für jeden Frame die Byte-Position im Strom vermerkt. Diese Informationen, in einem Array im Hauptspeicher gehalten, ermöglichen auf den ersten Blick einen sehr schnellen Zugriff auf gesuchte Frames. Der Speicher-Overhead von beispielsweise 4 Byte pro Frame kann hier, angesichts der Bitrate bei Video, vernachlässigt werden. Die Problematik liegt eher in der für die Wiedergabe notwendigen Vorverarbeitung. Wird diese

Vorverarbeitung vor jeder Wiedergabe erneut durchgeführt, der einmal erzeugte Index also nicht gespeichert, ist die Dauer der Vorverarbeitung das Problem. Ist dieser Zeitraum bei kurzen Videofilmen vielleicht noch zu vernachlässigen, so kommt es bei langen Videoströmen, wie etwa einer 60-minütigen Präsentationsaufzeichnung, zu inakzeptabel langen Wartezeiten im Minutenbereich. Greifen gar mehrere Benutzer, beispielsweise innerhalb einer lokalen Lehr- und Lernumgebung, gleichzeitig auf ggf. auch verschiedene Filme zu, führt selbst die normale Wiedergabe der Videoströme über das lokale Netzwerk zu einer überdurchschnittlich hohen Netzbelastung. Die benötigte Zeit für eine vorhergehende Indizierung wird in einem solchen Fall nur unwesentlich unter der normalen Wiedergabedauer eines Stroms liegen.

Um diese Wartezeiten insbesondere bei der Verwendung von Videofilmen während der Aufzeichnung einer Präsentation zu vermeiden, muß ein einmal erzeugter Index permanent zusammen mit dem Videostrom abgelegt werden. Dies erfordert jedoch eine erhöhte Vorbereitungszeit für den Vortragenden, da ggf. der Index eines jeden Videofilms vor Beginn der Präsentation erzeugt werden muß. Zur Vereinfachung der Handhabung werden, wie auch in [17] vorgeschlagen, Index und Strom häufig gekapselt, also in derselben Datei abgelegt. Ein solches proprietäres Format verhindert die anderweitige Nutzung solcher MPEG-Ströme durch gängige Wiedergabewerkzeuge. Für jeden Videofilm müßte also stets ein indizierter und ein nicht-indizierter Strom gespeichert werden, was angesichts der Datenvolumina sehr bedenklich ist. Auch kann ein einmal erzeugter Index eines Stroms in einem solchen Szenario häufig nicht gesichert werden, wenn beispielsweise von "Read-Only"-Speichermedien (CD, DVD, etc.) wiedergegeben wird oder dem Benutzer die Zugriffsrechte fehlen.

Unser Verfahren für den wahlfreien Zugriff kommt ohne Vorverarbeitung und ohne zusätzlichen Speicherplatz aus und basiert auf der Ausnutzung der durchschnittlichen Bitrate, um die Position b_0 des gesuchten Frames zu schätzen. Wenn die Schätzung falsch ist, dann kann die verwendete nicht die bis zu b_0 passende Bitrate gewesen sein. Die vor b_0 liegenden Frames waren also im Schnitt größer oder eben kleiner, also es die ursprüngliche Bitrate prognostiziert hat. Im nächsten Schritt wird nun die zu b_0 passende Bitrate für eine weitere Schätzung verwendet und so weiter. Genauer: Ein MPEG-1-Strom besteht auf *GOP Layer* aus einer Sequenz von *GOPs* (*Group of pictures*). Absolute Zeitinformationen existieren in MPEG-1 nur an Startpunkten von *GOPs*. Einzelne Frames beispielsweise enthalten lediglich eine relative Zeitinformation. Innerhalb einer einzelnen *GOP* kann also – im einfachsten Fall anhand einer sequentiellen Suche – entschieden werden, ob der gesuchte Frame enthalten ist. Daher besteht die Suche nach einem Frame zu einem gegebenen Zeitpunkt in der Suche nach der zugehörigen *GOP*. Die durchschnittliche Bitrate br_0 aus dem Sequenz-Header liefert nun die vermutliche Byte-Position b_0 des gesuchten Frames zu einem vorgegebenen Zeitpunkt. Wenn die zu b_0 gehörende *GOP* nicht die gesuchte *GOP* ist, dann wird dasselbe Verfahren mit der neuen durchschnittlichen Bitrate br_1 rekursiv fortgesetzt. Diese berechnet sich aus dem Zeitstempel (*time code*) der *GOP*, sowie seiner Byte-Position und muß von br_0 verschieden sein.

Offensichtlich kann die Bitrate nicht endlos steigen oder fallen, da dies durch die maximale Bitrate bei MPEG beschränkt ist und die Differenz zwischen zwei aufeinanderfolgenden Bitraten nicht beliebig klein werden kann. Dennoch ist diese Vorgehensweise in dem Moment kein gangbarer Weg, in dem nach dem Steigen ein Fallen der Bitrate folgt. Man sieht schnell an einfachen Beispielen, daß dann das Verfahren nicht zwingend terminiert. In dem Fall, daß dem Steigen ein Fallen der Bitrate folgt (oder umgekehrt) wird für die neue Bitrate schlicht das arithmetische Mittel der letzten beiden Bitraten verwendet.

Die Gewährleistung schnellen wahlfreien Zugriffs durch den oben skizzierten Algorithmus läßt sich nur schwer formal nachweisen, da zuviele unkalkulierbare Randbedingungen (Auslastung des System, Auslastung des lokalen Netzwerks, etc.) miteinbezogen werden müssen. Die mit unserem existierenden Prototyp durchgeführten Experimente haben jedoch gezeigt, daß bei geeigneter Variation der für den wahlfreien Zugriff charakteristischen Parameter im Strom in einer genügend großen Datenbasis durchaus Aussagen darüber getroffen werden können, in welchen Fällen ausreichend schneller wahlfreier Zugriff durch den obigen Algorithmus gewährleistet werden kann. So haben unsere Tests gezeigt, daß insbesondere die *GOP*-Länge ein entscheidender Faktor für die Schnelligkeit des wahlfreien Zugriffs ist. Kann also von einer geeigneten Beschränkung der *GOP*-Längen (z.B. 24 *Frames/GOP*) ausgegangen werden, dann kann der beschriebene Algorithmus ausreichend schnellen wahlfreien Zugriff realisieren. Es sei angemerkt, daß auch bei den zuvor angesprochenen Indizierungsverfahren die Geschwindigkeit des wahlfreien Zugriffs von der *GOP*-Länge abhängt.

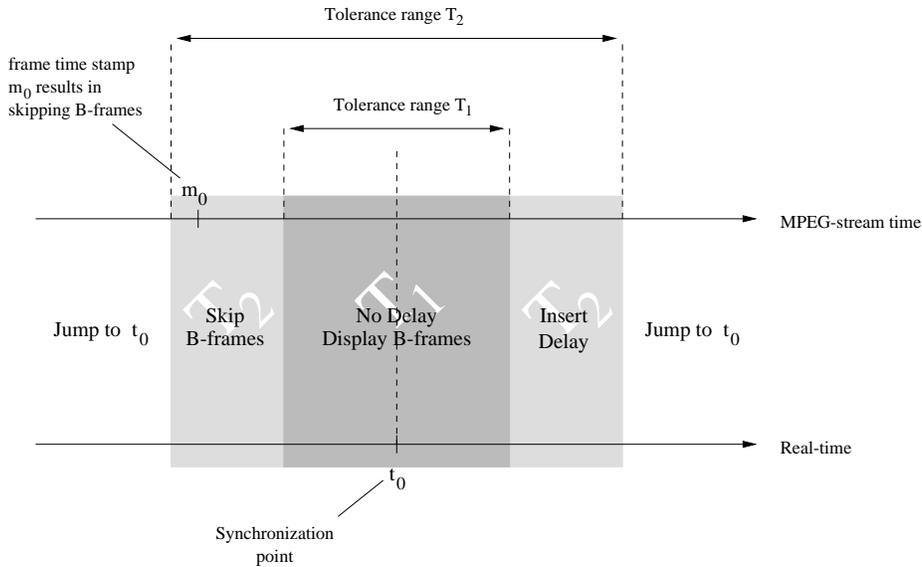


Abbildung 4: Wiedergabe-Modi in Abhängigkeit von der Echtzeit-Abweichung

A.2 Synchronisierung von MPEG-1

Schneller wahlfreier Zugriff kann auch dazu beitragen, den Grad der Synchronität zu erhöhen. In unserem Verfahren wird vor der Darstellung jedes Frames dessen implizit gegebener Zeitstempel mit der Echtzeit verglichen. Solange die Abweichung zwischen Filmzeit und Echtzeit innerhalb eines ersten vernachlässigbaren Toleranzbereichs T_1 liegt (s. Abb. 4), wird dies ignoriert. Für den Fall, daß die Abweichung aus T_1 herausfällt und die Filmzeit der Echtzeit hinterher hinkt, wird die Darstellung der *B-Frames* (*Bi-directionally predictive-coded frames*) abgeschaltet. Es werden also keine B-Frames mehr dargestellt in der Hoffnung, daß die Filmzeit gegenüber der Echtzeit aufholt. Entsprechend wird die Darstellung der B-Frames wieder eingeschaltet, wenn die Abweichung wieder innerhalb T_1 liegt. Für den Fall, daß die Filmzeit der Echtzeit vorausseilt – der Film wird also zu schnell wiedergegeben – wird eine geeignete Verzögerung von einigen Millisekunden zwischen den Frames eingefügt, solange bis die Abweichung wieder in T_1 hereinfällt. Wenn aber die Abweichung der Filmzeit durch die Dekodierung der Frames oder den Zugriff auf die Daten im Dateisystem hervorgerufen wird, reicht auch das Abschalten der Darstellung nicht aus. In diesem Fall kann die Abweichung der Filmzeit durch diese Maßnahme nicht reduziert werden oder erhöht sich ggf. noch weiter. Wenn dann die Abweichung auch aus einem zweiten, T_1 umfassenden Toleranzbereich T_2 herausfällt, wird der wahlfreie Zugriff ausgenutzt und an die gesuchte Stelle t_0 (Echtzeit) gesprungen. Geeignete Größen für T_1 und T_2 haben sich bei uns im Laufe einiger Experimente herauskristallisiert.

Da bei dieser Vorgehensweise im Extremfall (z.B. bei starker Systemauslastung oder schwacher Rechnerleistung) nur von Frame zu Frame gesprungen wird, werden also nur so viele Frames angezeigt, wie es die beschriebenen Faktoren eben zulassen. D.h. die Framerate passt sich automatisch diesen Faktoren an und führt damit zu einem dynamischen Ausgleich zwischen der Anzahl dargestellter Frames und der Systemauslastung. Wir bezeichnen dies mit *Dynamischer Frameraten-Anpassung* (*Dynamic Frame Rate Adaption*). Zugunsten möglichst hoher Synchronität werden hier also einzelne Informationen ausgelassen und nicht angezeigt. Dies Prinzip entspricht der Zielrichtung einer unsicheren Multicast-Übertragung (z.B. MBone), bei der der Verlust von Daten zugunsten einer möglichst verzögerungsfreien Übertragung akzeptiert wird.

Das zuvor beschriebene Verfahren der MPEG-Wiedergabe unter wahlfreiem Zugriff haben wir, wie oben bereits erwähnt, in einem Prototyp eines MPEG-Players implementiert, der bereits für verschiedene Plattformen existiert (Windows95/98, Unix-Derivate). Der Prototyp fungiert als *Helper*-Applikation für das in [14] genauer beschriebene Applikationsmodell, daß ebenfalls bereits implementiert ist und routi-

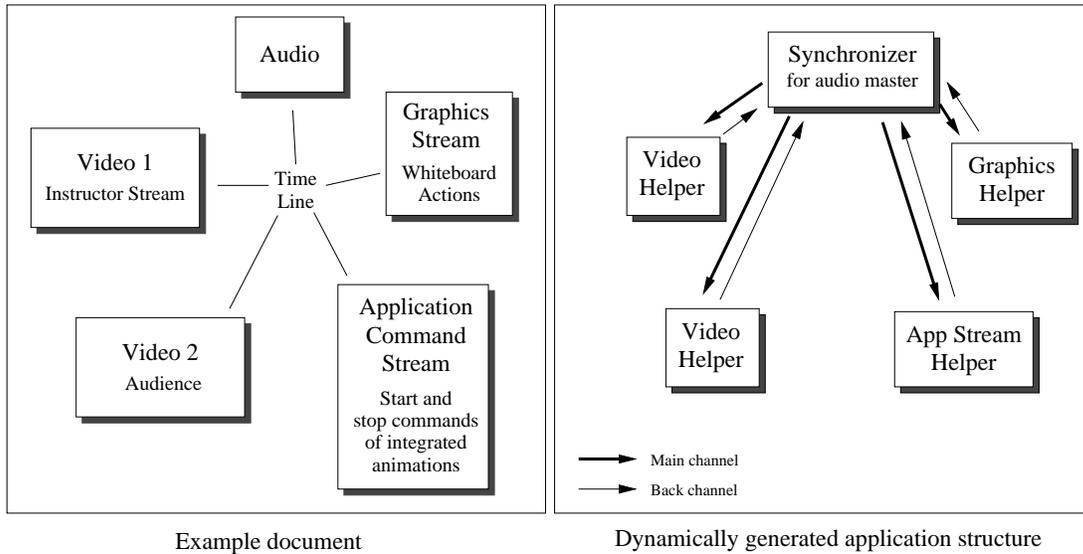


Abbildung 5: Ein Beispiel für ein AOF-Dokument und der zugehörige Aufbau der wiedergebenden Applikation [14].

nemässig für die Wiedergabe von AOF-Dokumenten genutzt wird (vgl. [1],[14]). Dem Applikationsmodell liegt eine Aufspaltung der Wiedergabe auf mehrere nahezu unabhängige Präsentationsmodule (ein *Synchronizer* und verschiedene *Helper*) gemäß der Zusammensetzung der einzelnen Medienströme im AOF-Dokument zugrunde (siehe Abb. 5). Der Synchronizer gibt dabei den Audiostrom wieder und sorgt für die Synchronisierung der einzelnen Ströme. Dazu werden zwischen dem Synchronizer und den einzelnen Helper-Applikationen Kommunikationskanäle geschaltet, über die insbesondere die Zeitstempel für die Synchronisierung übertragen werden (siehe Abb. 6). Da die Helper-Applikationen die ihnen zugeordneten Medienströme selbständig wiedergeben und die Wiedergabe anhand der empfangenen Zeitstempel synchronisieren können, ermöglicht das Applikationsmodell damit die oben genannte dynamische Frameraten-Anpassung.

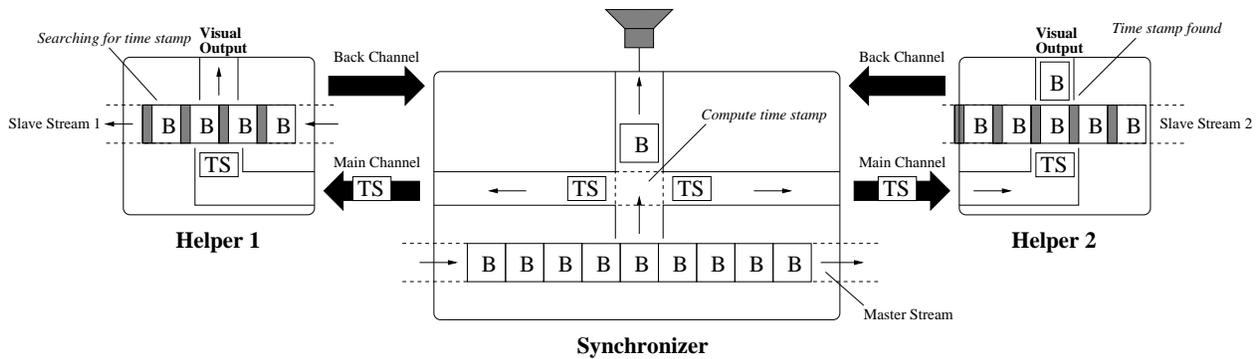


Abbildung 6: Konzeptuelle Darstellung der Funktionsweise des Applikationsmodell für die Wiedergabe von AOF-Dokumenten [9].

Mit Hilfe der Helper-Applikation für MPEG-1 ist also die Integration von Video-Clips, wie auch vom Videobild des Vortragenden in AOF-Dokumente unter Beibehaltung des wahlfreien Zugriffs und der Synchronität zu den verbleibenden Medienströmen der aufgezeichneten Vorlesung realisierbar.

A.3 Grafische Annotation

Bei der grafischen Annotation von in Vorlesungen integrierten Videofilmen soll der Vortragende die Möglichkeit bekommen auf der grafischen Ausgabe der Wiedergabeanwendung (d.h. den Fenstern) zu zeichnen und auf dort dargestellte Objekte zu zeigen. Ähnlich den grafischen Fähigkeiten einer Whiteboard-Applikation sollen hier also Freihand-Linien, einfache grafische Objekte wie Linien, Rechtecke, Ellipsen und eben auch ein Tele/Online-Pointer bereitgestellt werden. Die Annotationen, die der Vortragende während der Aufzeichnung auf den Fenstern der Applikation ausgeführt hat, sollen dann bei der Wiedergabe synchron zum erneuten Ausführen der Wiedergabeapplikationen dargestellt werden.

Eine allgemeine Lösung dieser Aufgabe, die unabhängig von der jeweilig verwendeten Videowiedergabeanwendung ist, ist schwer zu realisieren, insbesondere wenn der Quellcode der Applikation nicht verfügbar ist. Die Realisierung hängt stark vom unterliegenden Fenstersystem (z.B. Windows, X Window System) ab und eine Wiedergabe der aufgezeichneten Annotationen ist damit ggf. nicht plattformunabhängig durchzuführen.

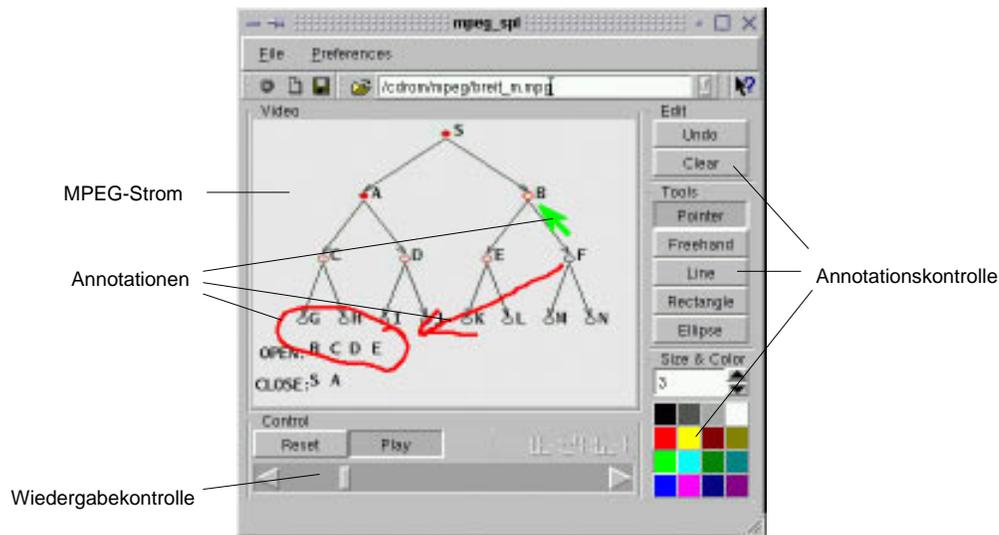


Abbildung 7: Die Aufnahmeumgebung für MPEG-Filme: Annotation und Aufzeichnung eines MPEG-Films einer animierten Breitensuche

Prinzipiell erfordert das Problem die Realisierung eines transparenten Fensters, welches über einen Event/Request-Mechanismus über Veränderungen des dahinter liegenden Bildschirminhalts informiert wird, um die eigenen Annotationen entsprechend aktualisieren zu können. Systemerweiterungen wie beispielsweise die *X Shape Extension* des X Window System oder das *Transparent Blitting* des MS Windows SDK erlauben das Setzen von beliebigen, nicht rechteckigen Clipping-Bereichen und damit die automatische Restriktion des Zeichenbereichs eines Fensters auf etwa eine Freihand-Linie. Alles außerhalb dieser durch Bit-Masken festgelegten Clipping-Bereiche ist nicht Bestandteil des Fensters und damit transparent. Allerdings bringt die Komplexität und wohl auch Zweckentfremdung dieser Mechanismen Geschwindigkeitsprobleme mit sich, insbesondere dann, wenn Videos oder Animationen mit hohen Frame-Raten annotiert werden sollen. Ähnliches gilt auch für den obengenannten Event-/Request-Mechanismus, da dieser bei hohen Frame-Raten nicht schnell genug und zudem asynchron ist.

In dem hier betrachteten Spezialfall eines auf der Basis der Implementierung von Rowe [17] eigens entwickelten MPEG-Player – der Quellcode steht also zur Verfügung – ist die Annotation eines Videos recht leicht realisierbar. Dabei werden, nachdem der aktuell betrachtete Frame dekodiert und zur Darstellung in einen Puffer geschrieben worden ist, die Annotationen des Vortragenden hinzugefügt und erscheinen damit über dem Frame des Videos. Danach wird der gesamte Puffer gemäß *Double-Buffering* für den Benutzer sichtbar in ein Fenster kopiert. Der Aufwand für die zusätzlich hinzugefügten Annotationen ist dann im Vergleich zum Dekodieren und Zeichnen des Frames selbst minimal. Zur Speicherung der grafischen Annotationen wurden die bereits ähnlich im AOF-Whiteboard angewandten Routinen und Datenstrukturen

verwendet [14].

Eine interessante Fragestellung in Bezug auf Annotation bei Video oder kontinuierlicher Medien allgemein ist deren Eignung. Insbesondere in Animationen und Filmen, bei denen die Sichtbarkeitsverhältnisse von Objekten sehr schnell wechseln – Objekte also oft nur sehr kurz sichtbar sind oder sich sehr schnell bewegen – wird Annotation im laufenden Zustand schnell sinnlos. Hier ist es häufig besser, den Videofilm kurzzeitig für die Annotation abzustoppen und nach Beendigung mit der Wiedergabe fortzufahren. Man kann sich jedoch auch Beispiele vorstellen, in denen grafisches Annotieren im laufenden Zustand hilfreich ist. Ein Beispiel ist der Film über die Villa Urbana, wenn etwa der zukünftige Pfad der Kamera durch die Szene beschrieben oder allgemein die anschließend folgende Bewegung von Objekten vorgezeichnet werden soll. Es ist also sicherlich sinnvoll, beide Modi – Annotation im laufenden und im gestoppten Zustand – zur Verfügung zu stellen.

Die in A.2 erwähnte Helper-Applikation für MPEG-1-Ströme wurde unter anderem um den beschriebenen Annotationsmechanismus zu einer Aufnahme-Umgebung erweitert. Während einer Rechner-Präsentation können hiermit beliebige MPEG-Filme eingeladen, zeitlich synchron präsentiert und grafisch annotiert werden. Der hieraus entstehende, kombinierte Strom aus MPEG-Filmen und zugehörigen Annotationen kann ebenfalls synchron und mit Unterstützung des wahlfreien Zugriffs wiedergegeben werden (s. Abb. 7).

Literatur

- [1] Authoring on the Fly – Document Download Area.
<http://ad.informatik.uni-freiburg.de/mmggroup.aof.docdownload>
Letzter Zugriff: 13.01.00.
- [2] Bacher C, Müller R (1998) Generalized Replay of Multi-Streamed Authored Documents. Proceedings of ED-MEDIA '98 Conference of Educational Multimedia and Hypermedia, Freiburg, Germany, June 1998
- [3] Bodendorf F, Grebner R, Langenbach C (1996) The Virtual Lecture Theatre - Practice and Experience. Proceedings of 2. Russian-German Symposium, New Media for Education and Training in Computer Science, Swiridow AP, Widmayer P, Oberhoff WD, Unger H (Eds.), Moscow, Nov. 1996.
- [4] Cen S, Pu S, Staehli R, Cowan C, and Walpole J (1995) A Distributed Real-Time MPEG Video Audio Player. In *Proc. of 5th International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV '95)*, Durham, NH, Apr. 1995.
- [5] Chen Z, Tan SM, Sane A, Li Y, Campbell R, Xie D (1996) Video and Audio: Organization and Retrieval in the WWW. White Paper, Department of Computer Science, University of Illinois at Urbana Champaign, 1996.
- [6] DeDonno T (1996) Open Hypermedia File System. In *Proc. of 2nd Workshop on Open Hypermedia Systems (Hypertext '96)*, Washington, DC, Mar. 1996.
- [7] Floyd S, Jacobson V, Liu C, McCanne S, Zhang L (1997) A Reliable Multicast Framework for Lightweight Sessions and Application Level Framing. *IEEE/ACM Transactions on Networking*, December 1997, Vol. 5, No. 6
- [8] Holfelder W (1998) *Aufzeichnung und Wiedergabe von Internet-Videokonferenzen*. PhD thesis, Universität Mannheim, Germany, May 1998.
- [9] Hürst W and Müller R (1999) A Synchronization Model for Recorded Presentations and its Relevance for Information Retrieval. In *Proc. of ACM Multimedia '99*, Orlando, FL., Oct. 1999.
- [10] Jährliche Grabungsberichte in: Archäologische Ausgrabungen in Baden Württemberg 1991 ff.
- [11] Lienhard J, Maass G (1998) A New Alternative for the Mbone Whiteboard wb. In *Proc. of ED-MEDIA '98, Conference of Educational Multimedia and Hypermedia*, Freiburg, Germany, June 1998.

- [12] Macedonia MR, Brutzman DP (1994) Mbone Provides Audio and Video Across the Internet. IEEE Computer, Vol. 27, No. 4, April 1994
- [13] Mental Ray: Mental images Gesellschaft für Computerfilm und Maschinenintelligenz mbH Co. KG, Berlin
- [14] Müller R, Ottmann T (2000) The “Authoring on the Fly” System for Automated Recording and Replay of (Tele)presentations, to appear in: Special Issue on “Multimedia Authoring and Presentation Techniques” of ACM/Springer Multimedia Systems Journal, 8(3), May 2000.
- [15] Nuber HU (1994) Villenforschung an der Universität Freiburg im Breisgau. In: S. Palágyi (Hrsg.), Forschungen und Ergebnisse. Internationale Tagung über römische Villen. Veszprém 16.–20. Mai 1994. Balácai Közlemények III, 1994 (Veszprém 1995) 159. Die villa urbana von Heitersheim (D). - Ebd.
- [16] Nuber HU (1997) Römische Antike am Oberrhein: Die villa urbana in Heitersheim. Archäologische Nachrichten aus Baden 57, 1997.
- [17] Rowe LA, Patel KD, Smith BC, Liu K (1994) MPEG Video in Software: Representation, Transmission, and Playback. In *Proc. of High Speed Networking and Multimedia Computing, IS&T/SPIE Symp. on Elec. Imaging Sci. & TEch.*, San Jose, CA, Feb. 1994.
- [18] Seitz G (1998) Archäologieprojekt Villa Urbana Heitersheim. In: Expressum. Informationen aus dem Freiburger Bibliothekssystem. Elektronische Dienste für Studium und Wissenschaft (Sonderausgabe 1998).
- [19] Softimage Inc.
<http://www.softimage.com>
Letzter Zugriff: 11.09.00.
- [20] VIROR – Virtual University Upper Rhine Valley.
<http://www.viror.de>
Letzter Zugriff: 02.03.00.