

Datenströme in multimedialen Systemen

– Merkmale und Auswirkungen –

Alexander Eichhorn
Technische Universität Ilmenau

Winfried Kühnhauser
Technische Universität Ilmenau

Abstract. Multimediale Anwendungssysteme besitzen einen extremen Bedarf an Rechenressourcen und können vielfach nur unter hohen, für viele Anwendungsszenarien unakzeptablen Qualitätseinschränkungen realisiert werden. Besonders deutlich wird dies in verteilten Szenarien mit kleineren, mobilen Systemen und drahtlosen Kommunikationsmedien, die sich aus diesem Grund heute noch vielen attraktiven Anwendungsfeldern verschließen.

Ein wesentlicher Grund für den enormen Ressourcen hunger multimedialer Anwendungssysteme liegt im Volumen kontinuierlicher multimedialer Datenströme, die in derartigen Applikationen kommuniziert und verarbeitet werden und welches viele Terabit je Sekunde betragen kann.

Dieser Beitrag bestimmt wesentliche Eigenschaften kontinuierlicher multimedialer Datenströme und identifiziert Implikationen für adäquate Systemarchitekturen und Kommunikationsmodelle. Er argumentiert, dass Gründe für den heutigen Ressourcen hunger multimedialer Anwendungssysteme vielfach in den problemfremden Kommunikationsmodellen und Implementierungsstrategien bestehender IT-Infrastrukturen zu suchen sind und liefert Hinweise für problemspezifische, effiziente Architekturen, Modelle und Mechanismen multimedialer Infrastrukturen.

1 Einleitung

Multimediale Anwendungssysteme reflektieren die auf audio/visuelle Kommunikation ausgerichteten menschlichen Sinnesorgane und erschließen Anwendungsszenarien, in denen Effizienz, Mühelosigkeit und Informativität der Kommunikation mit oder mittels IT-Systemen eine besondere Rolle spielen. Die Erwartungen an multimediale IT-Systeme sind hoch, da zahllose Anwendungsszenarien der Kommunikationstechnik, Luft- und Raumfahrt, Telemedizin, Verkehrstelematik und Produktionstechnik erst durch multimediale Techniken überhaupt machbar werden.

Die heute bereits realisierten Systeme machen allerdings leider auch die Leistungsgrenzen existierender IT-Systeme deutlich. Ursache hierfür sind die enor-

men IT-Ressourcen, die für die Verarbeitung und Kommunikation kontinuierlicher multimedialer Datenströme erforderlich sind.

Multimediale Applikationen verbrauchen ein nie gesehenes Maß an Prozessor-, Speicher- und Kommunikationsressourcen. Sie stellen zudem Forderungen hinsichtlich Rechtzeitigkeit und Zuverlässigkeit, denen heutige Kommunikationssysteme bereits architektonisch nicht gewachsen sind. Dies ist auch nicht verwunderlich, waren denn zum Zeitpunkt deren Entwicklung Multimediaapplikationen nicht einmal am Horizont zu entdecken. Hinzu kommen neuartige Kommunikationsmuster, deren Charakteristiken sich sogar innerhalb ein und derselben Applikation mit der Zeit ändern können. Je nachdem was eine Multimediaapplikation gerade tut, sind beispielsweise einmal Rechtzeitigkeit und Zuverlässigkeit gleich wichtig (beispielsweise während des Abspielen eines Videofilms), ein anders Mal interessieren Verluste nur sekundär (schnelles Vorspulen im Film).

Verteilte multimediale Anwendungen haben zusätzlich mit den unangenehmen Eigenschaften realer Kommunikationssysteme (Ausfälle, Reihenfolgenuntreue, Duplikate, Verspätungen, Verzögerungsschwankung usw.) zu kämpfen. Diese ohnehin prekäre Situation wird durch die wachsenden Inhomogenitäten der Prozessoren und Übertragungsmedien noch weiter zugespitzt. So sind die verschiedenen Szenarien geprägt von Systemen mit sehr unterschiedlicher Ressourcenausstattung und bestenfalls punktuellen Unterstützung für Multimediaanwendungen. Hier finden wir Hochleistungsserver und garantieorientierte Hochleistungsnetzwerke (ATM) einerseits sowie spärlich ausgestattete mobile Rechner (PDA's) und drahtlose Netzwerke (WLAN, UMTS) auf der anderen Seite. Insbesondere drahtlose Netzwerke mit ihren äußerst wechselhaften Eigenschaften stellen heutige reservierungsbasierte Qualitätsgarantiemodelle grundsätzlich in Frage. Diese Randbedingungen zeichnen ein sehr inhomogenes Bild hinsichtlich zentraler Systemeigenschaften wie

- Prozessorleistung, Dateisystemleistung
- Verlässlichkeit
- Bandbreite
- Latenzzeiten

Das Papier beschreibt wesentliche charakteristische Merkmale multimedialer Datenströme und zeigt Folgen für das Design und die Implementierung multimedialer Applikationen in heutigen IT-Systemen auf. Es nennt Gründe für den heutigen Ressourcen hunger multimedialer Anwendungssysteme und liefert Hinweise für den Entwurf effizienter Architekturen, Modelle und Mechanismen multimedialer Infrastrukturen und Anwendungssysteme.

2 Multimediale Datenströme

2.1 Datenvolumen

Multimediale Daten sind generell relativ voluminös. Eine einfache 10-sekündige Szene in unkomprimiertem regulären YUV-4:2:2-Format (PAL Auflösung mit

720x567 Pixeln und 25 Bildern/Sekunde) ist ca. 200 MByte groß, dieselbe Szene in HDTV (1920x1080 Pixel, 24-Bit Farbtiefe, 60 Bilder/Sekunde) benötigt 3,5 GByte. Der enorme Bedarf an Plattenplatz für die Speicherung sowie der Bedarf an Bandbreite, CPU-Leistung und Hauptspeicher für die Kommunikation solcher Ströme überschreitet in vielen Fällen die Fähigkeiten heutiger Standard – IT-Systeme.

Die Reaktion auf dieses Datenvolumen ist zunächst die Verwendung komprimierter Audio- und Video-Formate. Daneben kommen meist parallel Skalierungsverfahren zum Einsatz. Die gängigen Komprimierungsverfahren reduzieren Redundanz und Irrelevanz innerhalb von Stromfragmenten. Dies geschieht meist unter Verwendung von Differenzbildern bei Video und psychoakustischen Modellen bei Audiodaten. Insbesondere die verlustbehafteten Kompressionsverfahren machen sich dabei die physiologischen Eigenschaften des menschlichen Auges und Ohres zunutze. Komprimierte Ströme unterscheiden sich von ihren Originalen maßgeblich in

- Größe (bei Videoströmen wird heute eine Reduktion um 2 Größenordnungen erreicht)
- Struktur
- Gleichmäßigkeit der Datenrate

Die Verfahren zur Skalierung reduzieren Ströme in räumlicher bzw. zeitlicher Dimension (Bildgröße, Farbtiefe, Bild- und Abtastrate). Das hat zwar weniger Auswirkungen auf Struktur und Gleichmäßigkeit, beeinflusst jedoch stark die Qualität der im Strom enthaltenen Informationen.

So gut Komprimierungs- und Skalierungsverfahren auch die Datenvolumina reduzieren, sie verschieben einen unkalkulierbar großen Teil des Bedarfs an Speicher- und Kommunikationsressourcen hin zu Prozessorressourcen. Einige der Verfahren sind dabei asymmetrisch ausgelegt, so dass die Kompression mehr Ressourcen verbraucht als die Dekompression (MPEG, [Ben97]). Gerade auf leistungsarmen mobilen Rechnern lassen sich damit keine anspruchsvollen Anwendungen (Videokonferenzen) realisieren. Hier ist denn auch ein Trend in Richtung spezieller Videokompressionshardware zu beobachten.

2.2 Struktur

Multimediale Dokumente werden als zusammenhängende, sequenzielle und hierarchisch strukturierte Datenströme aus diskreten Datenelementen verstanden. Ein typischer unkomprimierter Videostrom (Abbildung 1) setzt sich hierarchisch aus einer Reihe an Szenen zusammen, welche jeweils Folgen aus Einzelbildern enthalten. Jedes Bild ist wiederum strukturiert als Folge einzelner Blöcke (meist Zeilen) und die Blöcke als Folge von Bildpunkten. Bildpunkte werden im RGB-Farbsystem als 3-Tupel aus jeweils einem Byte für die Farben Rot, Grün und Blau repräsentiert. Audioströme besitzen in der Regel eine flachere Hierarchie,

in der einzelne Abtastwerte zu Blöcken zusammengefaßt werden. Derartige Substrukturen besitzen stets eine konstante Größe und wiederholen sich innerhalb eines Stroms regelmäßig.

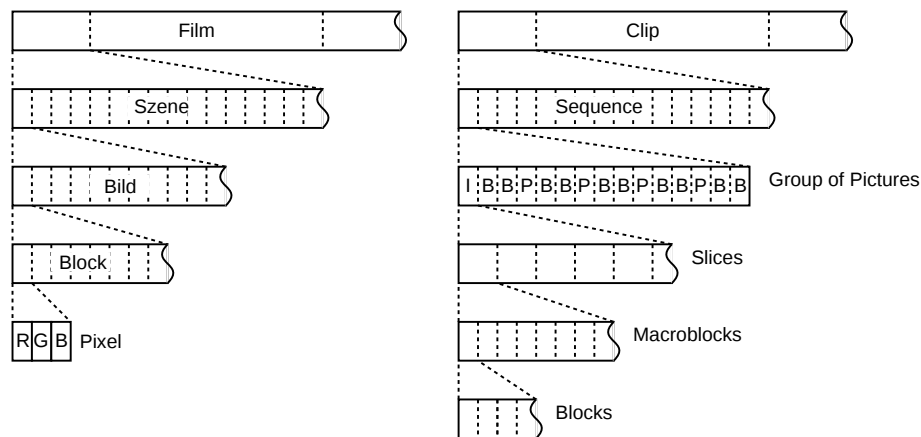


Abbildung 1. Typische hierarchisch strukturierte Datenelemente in Videoströmen (unkomprimierter Strom und MPEG-Strom)

Die Kompression hat auf die Datenstromeigenschaften nicht nur verdichtende Wirkung, sie wirkt auch strukturverändernd. Alle bekannten Verfahren (Digital-Video, MPEG-Familie, H263 usw.) erzeugen formatspezifische Datenelemente. Oftmals entstehen zusätzlich formatbedingte inhaltliche Abhängigkeiten zwischen Datenelementen (Differenzbilder benötigen Vollbilder als Ausgangsbasis). Insbesondere arbeiten solche Formate, die ein günstigeres Verhältnis von Datenvolumen zu Qualität anstreben (MPEG-Familie), mit variablen Blockgrößen. Das Volumen der einzelnen Datenelemente hängt hier direkt vom Inhalt des Ausgangsmaterials ab (Videoszenen mit wenig oder viel Bewegung) und ist somit a priori unkalkulierbar.

Ein direktes Editieren und Skalieren komprimierter Ströme wird durch diese Eigenschaften extrem erschwert; es macht in der Regel eine Dekomprimierung notwendig und erfordert somit erneut CPU- und Speicherressourcen. Die variablen Blockgrößen machen eine explizite Markierung der Datenelemente in Form von direkt in den komprimierten Strom eingebrachten Marken erforderlich, so dass Blöcke später wieder identifiziert werden können. Eine spätere Rekonstruktion der inneren Struktur, die für die Arbeit eines Dekoders unerlässlich ist, macht folglich das Parsen des gesamten Stroms notwendig. Hierzu müssen alle Bits vom Prozessor untersucht werden und im Hauptspeicher liegen. Im Falle von Ressourcenengpässen ist somit das ansonsten mögliche geordnete Überspringen einzelner Datenelemente ausgeschlossen.

Speziell bei einigen Videokompressionsverfahren, die auf der mathematischen

Methode der diskreten Kosinus-Transformation (DCT) aufbauen, entsteht eine neue Dimensionen der Skalierung, der Frequenzbereich. Diese Verfahren zerlegen einen Block aus Bildpunkten in Elementarfrequenzen. Begrenzt man nun die Breite des entstehenden Frequenzspektrums, verändern sich Kontrast und Schärfe des später zu dekomprimierenden Bildes.

Die reduzierte Redundanz und die stärkeren strukturellen Verflechtungen in komprimierten Strömen erhöhen deren Verlustanfälligkeit. Insbesondere entfaltet sich diese negative Wirkung gerade dort, wo es weh tut – in Funknetzen mit schmalen Bandbreiten und hohen Ausfallraten. In paketorientierten Netzen besteht ohnehin die prinzipielle Gefahr von statistisch verteilten Paketverlusten. Bereits Verluste einzelner Fragmente von Datenelementen wirken sich wegen der starken strukturbedingten Abhängigkeiten in der Regel auf weitere Datenelemente aus. Strukturbeschreibende Markierungen sind besonders sensible Informationen. Da Dekomprimierungsalgorithmen in der Regel nicht robust gegen deren Verlust sind, existiert eine grundsätzliche Regel zur Abbildung eines Stroms auf Netzwerkpakete. Beginnt der Inhalt jedes Pakets mit einer Markierung aus beliebiger Hierarchieebene kann auch der Verlust mehrerer aufeinanderfolgender Pakete die Strukturerkennung nur begrenzt beeinträchtigen.

2.3 Ordnung

Datenelemente in multimedialen Dokumenten sind total geordnet. Die Ordnung wird einerseits durch die zeitliche Reihenfolge einzelner Datenelemente (Video-bilder, Audiosamples) bestimmt. Andererseits legt ebenfalls die räumliche Anordnung der Datenelemente verbunden mit einem festen Schema bei Abtastung und Präsentation (Bildpunkt- und Zeilenweise Verarbeitung von Bildern) die Reihenfolge der Elemente im Strom fest. Diese Ordnung ist auch in der Struktur eines multimedialen Stroms deutlich erkennbar. Wird ein Strom kommuniziert und danach weiterverarbeitet, ist die Aufrechterhaltung der Ordnung notwendig. Da Kommunikationssysteme die Reihenfolgentreue nicht in jedem Fall garantieren, müssen vorsorglich geeignete Mechanismen eingesetzt werden. Innerhalb echtzeitabhängiger Applikationen reicht dies alleine aber noch nicht aus, da der zusätzliche Bezug zur realen Zeit ebenfalls das rechtzeitige Eintreffen des passenden Folgelements erfordert.

2.4 Zeit

Die zeitliche Dimension spielt für Datenlemente in multimedialen Strömen eine bedeutende Rolle. Während in herkömmlichen IT-Systemen und Anwendungsszenarien der Faktor Zeit von eher untergeordneter Bedeutung ist, erfüllt die Zeit für multimediale Anwendungen gleich mehrere Zwecke:

- Implementierung der Ordnung und Periodizität unter den Datenelementen
- Kalkulieren des Wertes eines Datenelements
- Synchronisation von Datenströmen (siehe 2.5)

Träger von Information und Semantik in kontinuierlichen Medienströmen sind nicht alleine die Werte der Datenelemente, sondern auch der Zeitpunkt des Auftretens. Die bereits erwähnten Substrukturen innerhalb von Multimediaströmen besitzen untereinander einen zeitlichen Bezug. So besteht zwischen zwei aufeinanderfolgenden Medienelementen, gemessen an der Realzeit, ein fester zeitlicher Abstand, der gleichzeitig die Ordnung der Elemente definiert (z.B. 25 Videobilder/Sekunde in fester Reihenfolge). Deshalb werden solche Ströme auch als streng periodisch bezeichnet. Die Periodizität entsteht bei der Aufnahme durch Kameras bzw. Mikrofone und muß bei der Präsentation zur Vermeidung von Informationsverlusten exakt rekonstruiert werden. Treten Datenverluste auf, so wird die Periodizität unter Verwendung redundanter Strom-Metainformationen (Zeitstempeln) explizit rekonstruiert.

Bei realzeitgebundenen Anwendungen (Luft- und Raumfahrt, Verkehrstelematik, Videokonferenzen, Telemedizin) hängt die Gültigkeit eines Datenelements von dessen rechtzeitigiger Verfügbarkeit ab. Verfrüht eintreffende Elemente erfordern zusätzliche Speicherressourcen, verspätete Elemente sind gänzlich unbrauchbar. Insbesondere interaktive Anwendungen verlangen von Bearbeitungsstufen und Kommunikationsdiensten geringe und kontrollierte Verzögerungszeiten ($<200\text{ms}$). Hohe Schwankungen der Verzögerung beeinflussen die Periodizität negativ und müssen vorbeugend durch Pufferung geglättet werden, wodurch nun wiederum die Gesamtverzögerungszeit negativ beeinflusst wird.

Der Einsatz von Kompressionsalgorithmen kann Periodizität und Rechtzeitigkeit stören. Die Rechenzeitkomplexität der gängigen gut komprimierenden Verfahren ist sehr hoch und zudem stark inhaltsabhängig. Das resultiert einerseits in Schwankungen der Bearbeitungszeit und somit in Verzögerungsschwankungen. Versucht man andererseits den Worst-Case der Bearbeitungszeit durch Vorabreservierung und Zuteilung von Ressourcen einzugrenzen, werden Ressourcen nicht optimal ausgenutzt. Besonders beim Einsatz auf leistungsarmen Prozessoren für mobile Rechner erzielen die Verfahren unangenehm geringe Durchsätze.

Die heute verwendeten optimistischen Fehlererkennungs- und behebevverfahren zur Sicherung der Zuverlässigkeit basieren weitestgehend auf Neuübertragungen. Für die Kommunikation über fehleranfällige Funknetze hat diese Strategie unkalkulierbare Verzögerungszeiten zur Folge, die Rechtzeitigkeit grundsätzlich in Frage stellen. In paketorientierten Weitverkehrsnetzen ist ebenfalls die unkalkulierbare und generell hohe Verzögerung ein limitierender Faktor für interaktive Anwendungen. Grundsätzlich erreichen durchgängig garantierorientierte Kommunikationstechnologien die stabilsten Verzögerungszeiten, allerdings wieder auf Kosten einer optimalen Ressourcenauslastung.

2.5 Synchronität

Nahezu alle multimedialen Anwendungen arbeiten mit mehr als einem Strom. Man denke beispielsweise an die digitale Aufnahme eines Symphoniekonzerts, bei der diverse Mikrofone für einzelne Instrumente und Instrumentengruppen die Audioströme produzieren und diverse Kameras das Orchester aus verschie-

denen Blickwinkeln aufnehmen. Zwischen den Datenelementen der Ströme bestehen enge semantische Beziehungen auf zeitlicher und inhaltlicher Ebene. Die Synchronisation der Datenelemente verfolgt mehrere Ziele:

- Aufrechterhaltung der Periodizität innerhalb eines Stroms
- zeitliche und inhaltliche Synchronisation zwischen unterschiedlichen Strömen (Lippensynchronität, Mischen von Mehrspur-Audioaufnahmen)
- zeitliche und inhaltliche Synchronisation zwischen Strömen und externen Objekten (weitere Medien, z.B. Untertitel, Vortragsfolien)

Die Synchronisierung zwischen Strömen ist nötig, um korrespondierende Datenelemente gleichzeitig zu bearbeiten bzw. zu präsentieren. Dies erfordert meist, dass Elemente eines Stroms auf das Eintreffen von Elementen anderer Ströme warten müssen. Applikationen, denen Rechtzeitigkeit wichtiger ist als Zuverlässigkeit, brauchen zusätzlich sinnvolle Abbruchkriterien für Warteoperationen (Timeouts). Zur Erkennung der Partnerelemente sowie zur expliziten Kennzeichnung von Synchronisierungspunkten innerhalb eines Stroms sind Metainformationen zu Zeit und Inhalt wichtig. Diese finden sich in heutigen Formaten in Form von Zeitstempeln und Synchronisationsmarken, diese allerdings stark formatabhängig.

Besonders Audio-Datenströme stellen sehr strenge Forderungen an die Güte der Synchronität. Das menschliche Gehör ist bei der räumlichen Ortung von Schallquellen sehr leistungsfähig und in der Lage, bereits zeitliche Abweichungen von weniger als einer Millisekunde differenziert zu erfassen. Zur Bewahrung einer als korrekt empfundenen räumlichen Information wird beispielsweise bei Multimikrofonaufnahmen ein Auseinanderlaufen der Audioströme von nicht mehr als eine Millisekunde gefordert. Die Güte der Synchronisation wird durch die Güte der physischen Uhren bzw. der Uhrensynchronisationsschemata der verarbeitenden IT-Systeme begrenzt. Die aus verteilten Systemen bekannten Synchronisationsalgorithmen und Systeme (z.B. NTP unter i386/Linux) sind hier bereits zu schwach.

Kompressions- bzw. Dekompressionsverfahren sowie Kommunikationsdienste beeinflussen die Güte der Synchronität meist negativ. Naive Implementierungen von Synchronisierungsmechanismen innerhalb einer Applikationsarchitektur vergessen häufig die unterschiedlichen Laufzeiten von Dekompressionsalgorithmen zu berücksichtigen. Sind die Synchronisationspartner gar voneinander entfernt, ist die Güte ebenfalls abhängig von Botschaftenlaufzeiten und statistischen Laufzeitschwankungen. Verfahren zur verteilten Uhrensynchronisation können allerdings einige angenehme Eigenschaften der Ströme ausnutzen. Durch die Ordnung der Datenelemente und deren Periodizität ist eine Art getaktete globale logische Zeit recht elegant herstellbar.

3 Zusammenfassung

Thema dieses Papiers waren die hohen Anforderungen multimedialer Anwendungssysteme an elementare informationstechnische Ressourcen wie Rechenleistung, Speichervermögen und Bandbreite von Kommunikationsmedien, die dazu führen, dass viele attraktive Anwendungsszenarien auf heutigen Systemen nicht mit akzeptablen Qualitätseigenschaften realisiert werden können. Das Papier konzentrierte sich hierbei auf eine Betrachtung der Eigenschaften multimedialer Datenströme, die durch ihre Großvolumigkeit bei ihrer Bearbeitung und Kommunikation eine wesentliche Ursache des Ressourcenkonsums bilden.

Drei wesentliche Ergebnisse sind festzuhalten.

Zum Ersten zeigt eine Betrachtung der Struktureigenschaften multimedialer Datenströme komplexe Zusammenhänge zwischen den verschiedenen Ressourcentypen auf: in Ermangelung leistungsfähiger Kommunikationsmedien werden Datenströme komprimiert; jedoch gerade dies erfordert Rechen- und Speicherressourcen für die Kompressions- und Dekompressionsalgorithmen, die wiederum den Informationsgehalt des Datenstroms verdichten und ihm eine andere Struktur geben. Dies wiederum hat Folgen für die Fehlersemantik (die Art und Weise, in der auf Packetverluste reagiert wird) und die Strategien zum Umgang mit Ressourcenengpässen. Quintessenz ist hier, dass diese komplexen Zusammenhänge von den heutigen, auf singuläre Ressourcentypen ausgerichteten Managementstrategien der Betriebs- und Middleware-systeme nicht beherrscht werden und durch ein adaptives, garantiegebendes und holistisches Ressourcenmanagement substituiert werden müssen.

Zum Zweiten weisen multimediale Datenströme Eigenschaften auf, die ein erhebliches Potential für die Implementierung notwendiger Funktionalität darstellen, die aber von heutigen Kommunikationsmechanismen vollkommen ignoriert werden. Als Beispiel wurden die Periodizität, Kontinuität und totale Ordnung angeführt, die als Grundlage eines effizienten und hochpräzisen Uhrensynchronisationsschemas herangezogen werden können. Des weiteren lassen sich Metadaten der Ströme (z.B. noch einmal die Zeitstempel) verwenden, um den relativen Wert eines Datenelementes für die Applikation zu bestimmen und auf diese Weise intelligente, problemspezifische Algorithmen für den Umgang mit Ressourcenengpässen zu implementieren.

Drittens und Letztens weisen heutige Kommunikationmechanismen Lücken in ihrer Funktionalität auf, die im Nachhinein außerhalb des Kommunikationssystems und innerhalb einer Applikation nur sehr ineffizient geschlossen werden können. Als Beispiel hierfür mag an die enge zeitliche Synchronisation einer Vielzahl von voneinander abhängigen Datenströmen innerhalb einer Applikation erinnert werden.

Als Quintessenz sei festgehalten, dass multimediale Anwendungssysteme tatsächlich einen durchaus hohen Bedarf an Ressourcen besitzen, dass aber gleichzeitig eine Vielzahl der beobachtbaren Qualitäts- und Performanzmängel auf die problemfernen Modelle und Mechanismen heutiger Standard-IT-Systeme zurückgeführt werden können.

Deutlich soll gesagt werden, dass dieses Papier keinesfalls als Ruf nach einer

Ablösung der heutigen Kommunikationsmodelle aufzufassen ist. Diese wurden teils vor mehr als 20 Jahren mit Blick auf einen Applikationsmix mit völlig andersartigen Kommunikationsmustern entworfen, und derartige Muster haben heute immer noch für eine Vielzahl von Anwendungssystemen Gültigkeit. Absicht der Autoren ist es vielmehr aufzuzeigen, dass viele wesentliche Merkmale heutiger multimedialer Anwendungssysteme diesen Mustern nicht entsprechen und Anregungen im Sinne einer evolutionären Fortentwicklung zu geben. Einige Ansätze dazu sind in diesem Papier zu finden.

Literatur

[Ben97] Hervé Benoit. *Digital Television; MPEG-1, MPEG-2 and principles of the DVB system*. Arnold, 1997.