

# **VO Multimedia 2: Technologien**

**Vortragender: Christian Breiteneder**

## **Zusammenfassung und Beantwortung**

der alten Prüfungsangaben, die auf

**[http://www.ims.tuwien.ac.at/teaching\\_detail.php?ims\\_id=188145](http://www.ims.tuwien.ac.at/teaching_detail.php?ims_id=188145)**

bereitgestellt werden.

(konkret die Prüfungen aus Jun 03, Nov 03, Mar 04, Jun 04, Okt 04, Jan 05, Apr 05, Jul 05, Okt 05, Dez 05)

**Stand: 01. 10. 2006**

### **Aufteilung:**

Die Fragen wurden zusammengefasst und in Kategorien (laut Folien) eingeteilt.  
Die Reihenfolge der Fragen orientiert sich ebenfalls an die Folien.

- 1) Media Server**
- 2) Visual Retrieval**
- 3) Content Description**
- 4) Multimedia Standards**
- 5) Digitales Fernsehen**

### **Arbeitsvorschlag:**

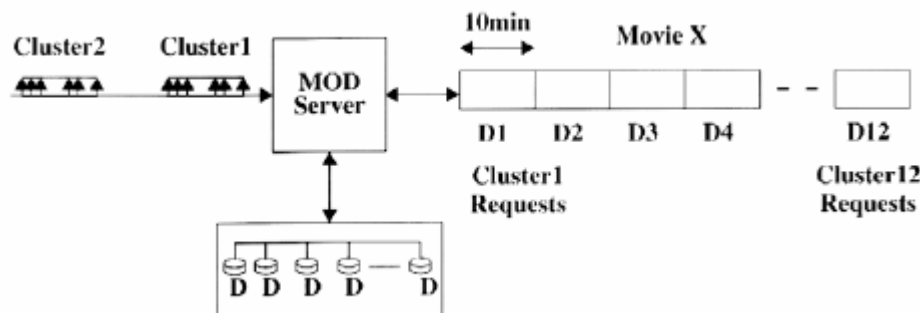
Grundsätzlich sollte es reichen, den Fragenkatalog zu lernen, um eine positive Note zu erhalten. Schaden wird es aber bestimmt nicht, zusätzlich die Folien auf der oben genannten Webseite zumindestens einmal durchzuschauen.

Ich kann natürlich nicht garantieren, dass der Professor bei diesen Fragen bleiben wird, aber er wird das schon nicht nach jahrelanger Praxis von heute auf morgen ändern.

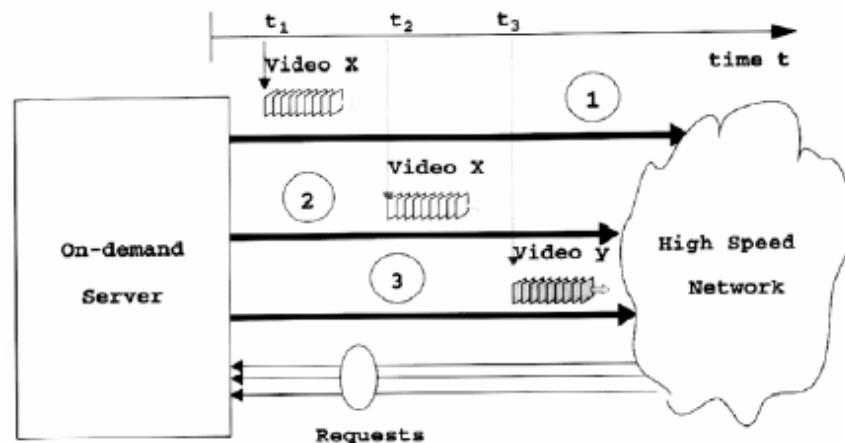
# 1) Media Server

Nennen Sie die wichtigsten On-Demand Service Typen und beschreiben Sie sie kurz.

- **Pay-Per-View (PPV)**  
Zugriffsdauer wird vom Server entschieden. Ausgestrahlt wird zu bestimmten Zeiten.
- **Near Video-On-Demand (NVOD)** bzw. Shared Viewing with Constraints (SVC)  
Zeitpunkt nicht frei wählbar. Es wird in bestimmten Intervallen ausgestrahlt (zB alle 30 Minuten kann man sich in den Anfang des Streams eines Filmes einwählen). Weiters:
  - der Server nimmt Requests in Gruppen auf und bearbeitet sie
  - ein neuer Client kann schwankende Zugangsverzögerungen aufweisen
  - Client wird Teil einer Multicast-Gruppe



- **True-Multimedia-On-Demand (TMOD)** bzw. Dedicated Viewing
  - jeder Zugriff wird unabhängig voneinander vom Server behandelt
  - Musterbeispiel für personalisierte, interaktive Multimedia-Übergabe
  - gravierende Auswirkungen auf den Durchsatz



Was sind die wichtigsten Performance-Metriken für Media Server?

- Gleichzeitigkeit (*concurrency*) – max. Anzahl an Clients, die unabhängig voneinander auf ein Multimedia-Dokument zugreifen können
- Zugriffsverzögerung (und Operationsverzögerung) (*latency*) – die Zeit, die ein Client nach Senden eines Requests warten muss (für eine interaktive Operation); weniger als eine Sekunde (< 1 Sek); sollte unabhängig von der Serverbelastung sein
- Speicherkapazität (*storage capacity*)
- Skalierbarkeit (*scalability*)
- Erweiterbarkeit (*extensibility*) – mehrere Applikations-Szenarien; zur Unterstützung für verschiedene Service-Modelle erweiterbar

Performance-Metriken dienen zum Vergleich und der Klassifizierung von Server-Architekturen.

## Was versteht man unter einem RAID? Beschreiben Sie Architektur eines RAID-Systems ihrer Wahl mit Level > 0.

**RAID (Redundant Array of Inexpensive Discs)** ist eine Technik zur Organisation mehrerer Festplatten.

- wurde in den 1980er Jahren als eine Möglichkeit vorgeschlagen, mehrere kostengünstige Festplatten zu einem logischen Laufwerk zu verbinden um die I/O-Leistung zu verbessern
- heute sind sie im Sortiment der meisten Computerhersteller enthalten
- 2 orthogonale (siehe [Orthogonalität in der Informatik](#)) Konzepte werden verwendet
  - Data Striping um die Leistung zu erhöhen
  - Redundanz um die Ausfallsicherheit zu erhöhen

### Disk Array Basics:

- Data Striping verteilt die Daten erkennbar über mehrere Festplatten, um sie als eine große Festplatte erscheinen zu lassen
- Data Striping erhöht den Datendurchsatz indem Festplattenzugriffe parallel abgewickelt werden können
  - mehr unabhängige Zugriffe: werden parallel auf verschiedenen Festplatten abgewickelt, wodurch die Wartezeiten verkürzt werden
  - einzelner Multiple-Block-Zugriff: wird durch das Zusammenarbeiten mehrerer Festplatten abgewickelt, wodurch sich die Transferrate erhöht
- je mehr Festplatten verwendet werden, desto größer die potentiellen Leistungsgewinne
- leider sinkt mit der Anzahl der Festplatten die Ausfallsicherheit
- Redundanz bedingt 2 orthogonale Probleme
  - Auswählen einer Methode zur Berechnung der redundanten Information (die meisten redundanten Arrays verwenden Parity, einige Hamming- oder Reed-Solomon-Codes)
  - Auswählen einer Methode zur Verteilung der redundanten Information über das Disk-Array (redundante Information auf wenigen Festplatten vs. Redundanz allgemein verteilt über alle Festplatten)
- das Konzept von Data Striping und Redundanz ist einfach, doch die Auswahl zwischen den vielen verschiedenen Modellen erfordert eine komplexe Abstimmung zwischen Ausfallsicherheit, Leistung und Kosten

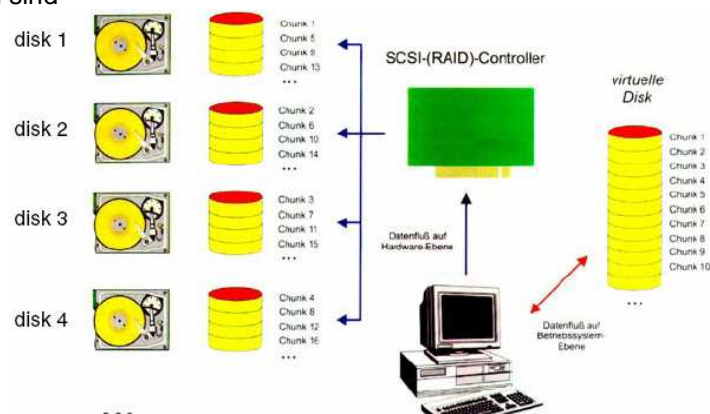
### RAID Level 3

[siehe unten](#)

### Ergänzung: RAID Level 0 und RAID Level 1

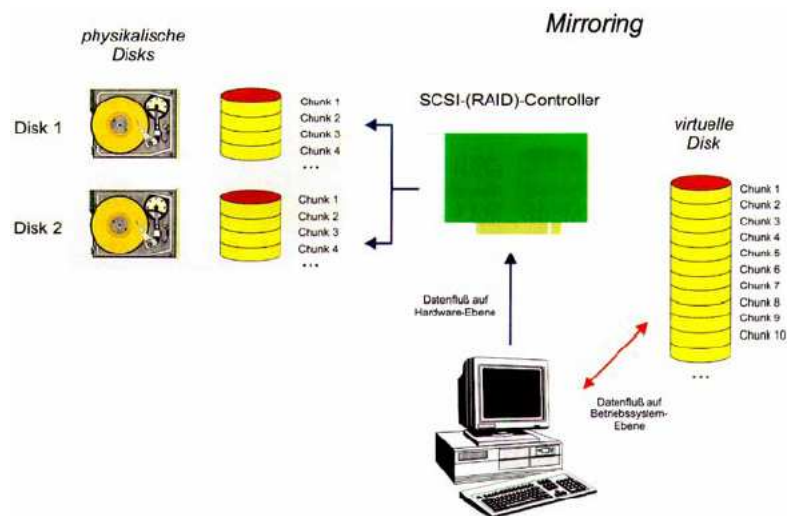
#### RAID 0

- ein nicht-redundanter Disk Array; Konsequenzen:
  - niedrigste Kosten unter allen RAID-Aufbaumöglichkeiten
  - beste Schreibleistung
- überraschenderweise bietet dieser Aufbau nicht die beste Leseleistung; Redundanz-Modelle erzielen eine bessere Leseleistung durch gezieltes Planen der Zugriffe auf Festplatten mit der kürzest erwarteten Suche und abwechselnden Verzögerungen
- größtenteils bei Höchstleistungsrechnern in Verwendung, wo Leistung und Kapazität die Hauptanliegen sind



## RAID 1

- Mirroring oder Shadowing benötigt doppelt soviele Festplatten als ein nicht-redundanter Array
- immer 2 Kopien der Daten
- Mirroring wird häufig bei Datenbank-Anwendungen benutzt, wo Verfügbarkeit und „transaction rate“ wichtiger ist als Speichereffizienz

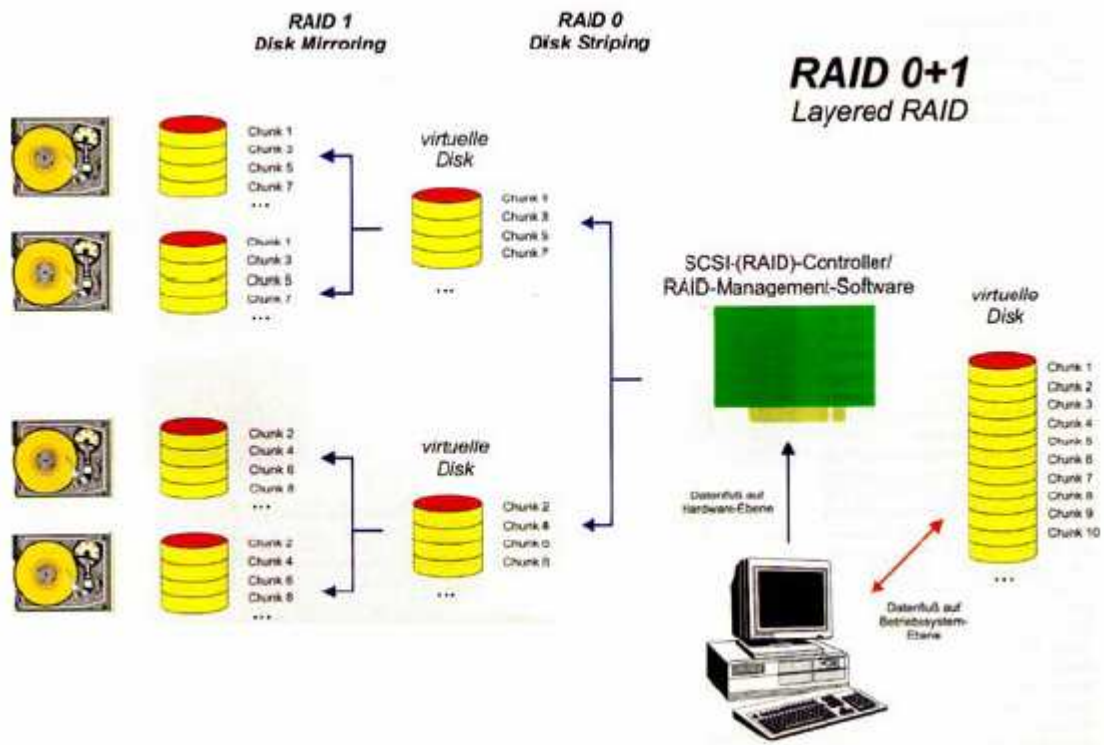


## **Was versteht man unter einer „Layered RAID 0+1“ Architektur?**

Kombination aus RAID 0 und RAID 1.

RAID 0 stellt die untere Ebene dar. Hier wird das Striping angewendet. Darüber liegt RAID 1, das Mirroring anwendet.

So kann man die Vorteile beider RAIDs ausnutzen und hat somit einen Geschwindigkeitszuwachs sowie eine erhöhte Ausfallsicherheit.



## Was versteht man unter einem RAID? In den Ausprägungen welcher 2 Eigenschaften unterscheiden sich alle RAID-Architekturen?

### RAID

[siehe oben](#)

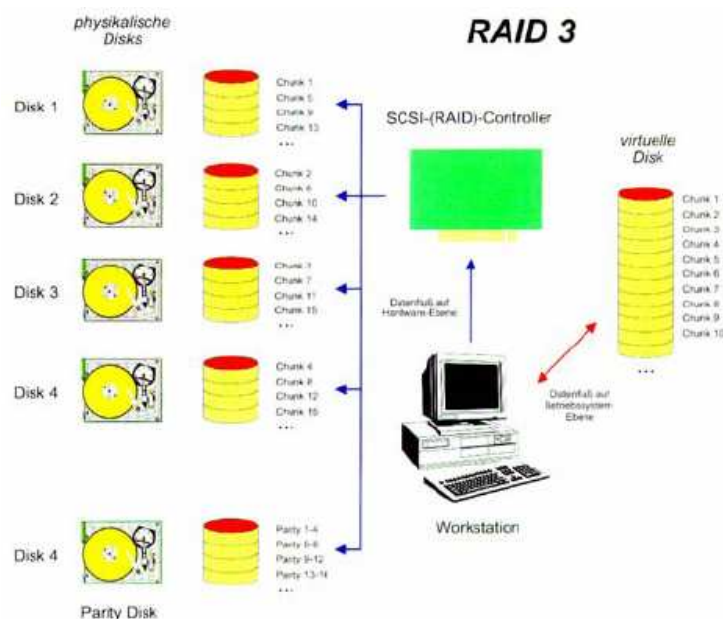
2 Eigenschaften:

- **Ausfallsicherheit (über Redundanz)**
- **Leistung (über Data Striping)**

## Was sind die wichtigsten Eigenschaften einer RAID Level 3 Architektur?

### Bit-Interleaved Parity

- „single parity disk“ – Disk-Controller kann leicht erkennen welche Disk ausgefallen ist
- Daten sind bitweise miteinander über die ganze Daten-Disk verzahnt
- jeder Lese-Request greift auf alle Daten-Disks zu, jeder Schreib-Request greift auf alle Daten-Disks UND die Parity-Disk zu
- Requests können nur einzeln bearbeitet werden
- wird häufig in Anwendungen verwendet, die hohe Bandbreiten benötigen, aber hohe I/O-Raten vernachlässigen können
- leichter zu implementieren als die Level 4, 5 und 6



## Beschreiben Sie kurz 2 Aufgaben des Client-Request-Schedulings.

- VCR-Kontroll-Operationen: man benötigt für ein „resume“ einen neuen logischen Kanal; andere Operationen (zB „fast-forward“) stellen überhaupt unvorhersehbare Anforderungen an den Server.
  - Die Contingency Channel Policy bestimmt ein paar statistisch verteilte Sicherheitskanäle, um unvorhersehbare Anforderungen zu bewältigen.
  - Algorithmen für die Reoptimierung oder das Verringern der hohen Anforderung an eine Ressource, indem man eine andere Variante benutzt
- Allgemeine Data-Streams: teilen sich den gesamten oder Teile des logischen Kanals unter mehreren Clients auf (Batching Policies, multicast)
  - Daten-zentrierter Zugang
  - Serverkapazität im Verhältnis zur Anzahl der Medianobjekte
  - Interval Caching
- zeitvariante Auslastung (*time-varying workload*)

## Was sind die Ziele des Channel Scheduling?

- langfristige Minimierung der Absagewahrscheinlichkeit (*reneging propability*)
- kurzfristige Minimierung der höchsten Absagewahrscheinlichkeit
- Minimierung der durchschnittlichen Wartezeit
- Fairness – eine Scheduling-Policy ist fair, wenn die Absagewahrscheinlichkeit für alle Video-Requests gleich ist
- Wartezeit bis zum Wiederaufnehmen
- Berücksichtigung der VCR-Anforderungen

## Was ist der Grundgedanke von Contingency Channel Policy? Wozu dient das Verfahren?

Policy wie man mit VCR-Kontrolloperationen umzugehen hat.

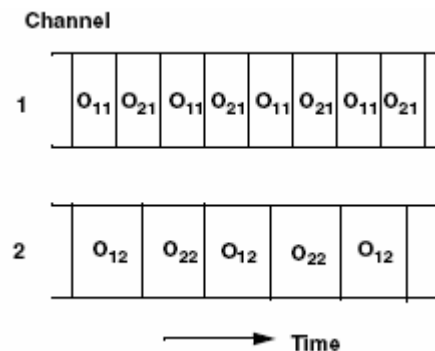
Man benötigt nur allein für ein „resume“ schon einen neuen logischen Kanal, andere Operationen (zB „fast-forward“) stellen überhaupt unvorhersehbare Anforderungen an den Server.

Die Contingency Channel Policy bestimmt ein paar statistisch verteilte Sicherheitskanäle für unvorhersehbare Anforderungen, um diese zu bewältigen.

## Beschreiben Sie kurz Pyramid Broadcasting. Wozu dient das Verfahren?

Mit dem Pyramid Broadcasting versucht man eine höhere Performance zu erreichen, in dem mehrere Anfragen in logischen Kanälen zusammengefasst werden, statt für jede einzelne Anfrage einen eigenen logischen Kanal zu öffnen.

Die max. Wartezeit einfacher NVODs ( $N_{\text{objects}} * T_{\text{play}} / N_{\text{logical.server.chan}}$ ) kann durch das Pyramid-Broadcasting-Verfahren reduziert werden.



- Die Bandbreite des Servers wird in  $N_{\text{pyr.ch}}$  logische Kanäle geteilt
- Jedes Video wird in  $N_{\text{pyr.ch}}$  Segmente geteilt; weiters wird deren Größe um  $\alpha_{\text{pyr.ch}}$  erhöht
- Erhöht sich  $N_{\text{pyr.ch}}$ , verkleinert sich das erste Segment
- das  $i$ -te Segment aller Videos wird wiederholt an Kanal  $i$  gesendet
- um Video  $i$  abzuspielen, startet der Client den Download von  $O_{i1}$ , falls verfügbar, und spielt es ab
- das Programm benötigt soviel lokalen Speicher wie das größte Segment ausmacht
- der Wert von  $\alpha_{\text{pyr.ch}}$ : bevor das aktuelle Segment fertig abgespielt wurde, muss der Download des nächsten Segments gestartet haben um eine kontinuierliche Anzeige zu gewährleisten
- max. Wert von  $\alpha_{\text{pyr.ch}}$ :  $N_{\text{logical.server.chan}} / (N_{\text{objects}} * N_{\text{pyr.ch}})$

## Wozu dienen dynamische Batching Policies? Nennen Sie ein Verfahren und seine Eigenschaften.

NVOD-Policies setzen Kenntnisse von Zugriffsmustern auf Videos voraus.

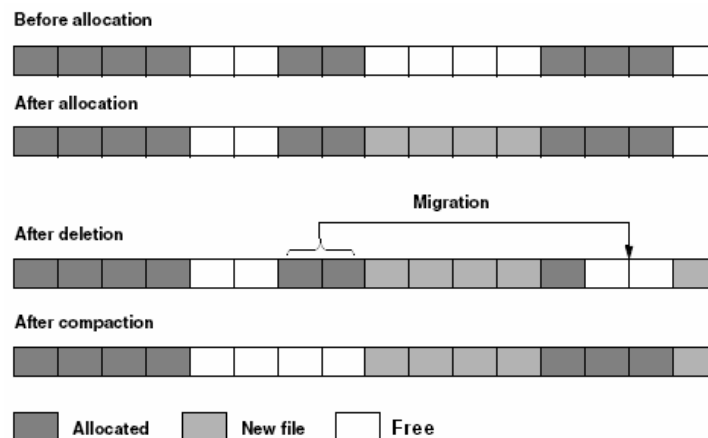
Dynamic Batching Policies passen sich an Veränderungen der Anforderungen an. Sie erkennen populäre Videos und betreuen mehrere User eines einzelnen Streams.

- **FCFS Policy** (first come first serve)
  - einzelne Warteschlange in der man sich einreicht
  - man kann somit eine max. Wartezeit anzeigen
  - fair
  - einfach zu implementieren
- **MQL Policy** (max. queuing length)
  - "greedy" Policy
  - zu jedem Video gehört eine einzelne Warteschlange
  - hohe Abbruchrate bei „kalten“ (wenig benutzten) Videos
  - bearbeitet mehr Anfragen gleichzeitig als FCFS,
  - ist dafür aber unfair
  - Wartezeit kann nicht berechnet werden
- **GGCS-FCFS Policy** (group guaranteed server capacity):
  - versucht die durchschnittliche Wartezeit von Anfragen zu minimieren
  - noch immer Bevorzugung von „heißen“ Videos
  - Wartezeit kann berechnet werden

## Beschreiben Sie kurz die Everest Contiguous Allocation. Wozu dient sie?

Der Sinn der Contiguous Allocation ist, die Dateien auf der Festplatte so gut wie möglich in aufeinander folgende Blöcke aufzuteilen (um eine Fragmentierung zu vermeiden, da eine Defragmentierung sehr aufwendig ist).

Die Everest Contiguous Allocation ist eine Adaption der Methode der Contiguous Allocation Methode, die versucht die Contiguous Allocation beizubehalten und gleichzeitig die Dateien kompakter zu verteilen.



### Funktionsweise:

- Die Dateien werden auf zusammenhängende Segmente der Größe  $w^i$  (Blöcke) aufgeteilt.  $w$  ist ein verstellbarer Parameter,  $i$  eine ganze Zahl.
- Um eine Datei auf  $L_f$  Blöcke zu verteilen, stellt die Methode  $L_f$  als Zahl zur Basis  $w$ ,  $d_k d_{k-1} \dots$  dar, und verteilt  $d_i$  dazugehörige Segmente der Länge  $w^i$ .
- Bsp.: eine Datei mit einer Länge von 5 Blöcken kann folgendermaßen angelegt werden:
  - $w = 2$ : 1 Segment mit 4 Blöcken ( $2^2$ ) und 1 Segment mit 1 Block
  - $w = 3$ : 1 Segment mit 3 Blöcken und 2 Segmente zu je einem Block

- freie Blöcke werden in freie Segmente zusammengefasst und in einer freie-Segmente-Liste beibehalten (eine Liste für jede mögliche Segmentgröße  $w^i$ )
- maximal  $w-1$  Segmente in einer Liste
- sollte es nach dem Löschen einer Datei in einer Liste zu mehr als  $w-1$  freien Segmenten der Größe  $w^i$  kommen, werden die Segmente so verteilt, dass zumindest  $w$  freie zusammenhängende Segmente der Größe  $w^i$  entstehen
- zusammenhängende Segmente werden dann in einem freien Segment der Größe  $w^{i+1}$  zusammengelegt

## Welche Constrained-Placement Policies kennen Sie? Diskutieren Sie kurz ein Verfahren.

Polys nutzen den Vorteil des sequentiellen Zugangs zu Medienstreams um den Seek Overhead zu begrenzen:

- **REBECA** [siehe unten](#)
- **Strand-based allocation**

## Beschreiben Sie kurz REBECA. Wozu dient das Verfahren?

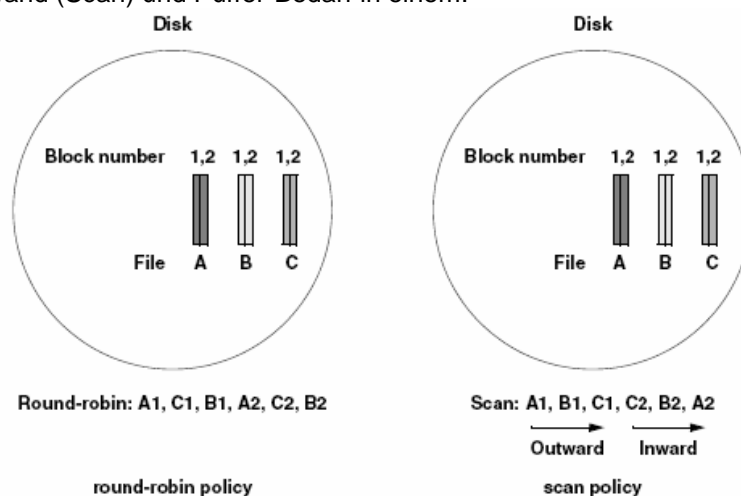
### REBECA (Region based block allocation method)

Kompromiss zwischen dem Seek Overhead (Zeit, die für Suchoperationen verloren geht) und der Startverzögerung neuer Streams

- Speichergerät wird in zusammenhängende Regionen unterteilt
- Zugriffe zu Blöcken in einer Region werden in einem Scan-Schritt behandelt
- erhöht man die Anzahl von Regionen sinkt der Seek Overhead, aber die Startverzögerung erhöht sich
- die Blöcke eines Medienobjekts werden verschiedenen Regionen zugeteilt, in der Reihenfolge in der auf sie zugegriffen wurde

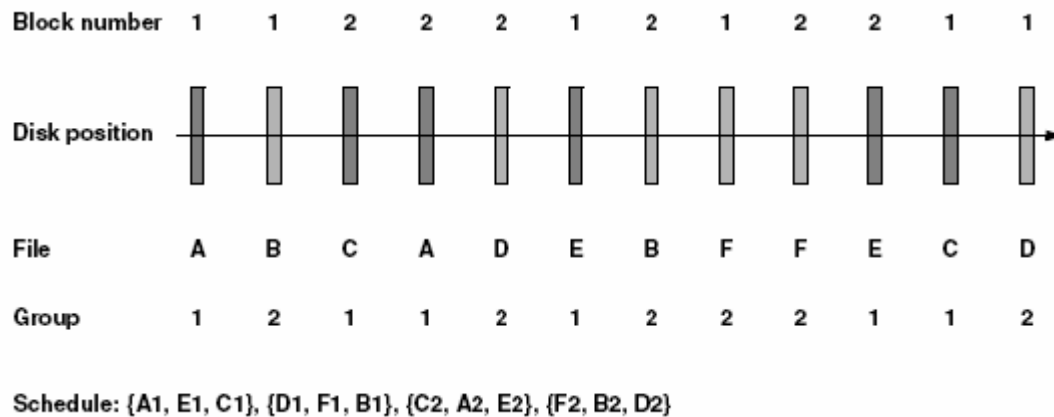
## Beschreiben Sie kurz Group Sweeping Scheduling. Wozu dient das Verfahren?

Kombiniert das Scan- und das Round-Robin-Verfahren, woraus eine Methode entsteht mit einem geringen Suchaufwand (Scan) und Puffer-Bedarf in einem.



1. Teile die  $V$  aktiven Streams in  $M_{gr}$  Gruppen
2. Wende das Round-Robin-Verfahren an den  $M_{gr}$  Gruppen an
3. Wende für die Streams in jeder Gruppe das Scan-Verfahren an

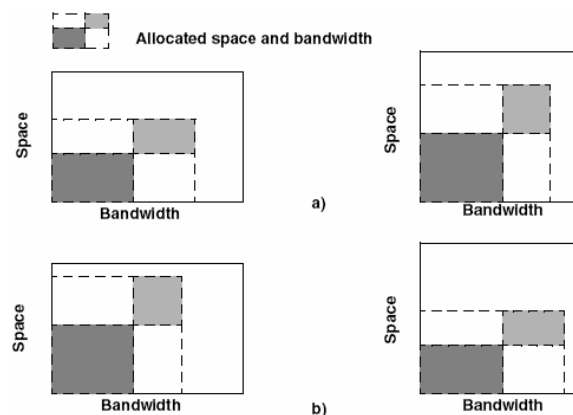




### Beschreiben Sie kurz die Bandwidth-to-Space (BSR) Policy. Wozu dient das Verfahren?

Ziel ist es, die vorhandenen Kapazitäten (Bandbreite, Speicherplatz) durch entsprechende Verteilung der Medienobjekte optimal auszunutzen.

- Das Verfahren versucht „heiße“ und „kalte“, große und kleine Medienobjekte in heterogene Speichergeräte zu vermischen (Striping-Gruppen)
- die Methode kennzeichnet jedes Gerät mit seinem Bandbreite-Platz-Verhältnis
- ähnlich werden Medienobjekte nach der erwarteten Bandbreiten-Anforderung ihrer Kopie und dem benötigten Speicherplatz gekennzeichnet
- das Verfahren erstellt und löscht Kopien nach Bedarf
- Bsp (schlechte Aufteilung): mehrere Objekte mit hohem Platzbedarf, aber geringer Bandbreite.
- Bsp (gute Aufteilung): mehrere Objekte in Summe entsprechend der Charakteristik des jeweiligen Geräts angepasst.



### Nennen und erklären Sie kurz 3 Ziele von Caching.

- **Erhöhung der Serverkapazität** – Speichern aller oder zum Teil oft benutzter MM-Objekte im Serverspeicher
- **Reduzierung der Zugriffsverzögerung** – Zugriffszeit verändert sich mit der Lage der Daten in der Speicherhierarchie; zB können Daten, die im Speicher liegen augenblicklich an die Clients gesendet werden
- **Reduzierung der Anforderungen an die Netzwerk-Bandbreite** – ein großangelegter Server, der aus einem lokalem Server (sendet Daten an Clients) und Remote-Storage-Server (lagert Originalkopien der Daten; lokale Server zwischenspeichern Daten auf lokalen Disks und Speicher um einen „communication overhead“ zu verhindern

## Beschreiben Sie kurz die Viewer Enrollment Window Policy. Wozu dient das Verfahren?

### (a) Einfacher Buffer

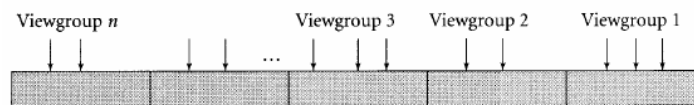
- sollten mehrere Streams auf dasselbe Medienobjekt zugreifen und wurde dem ersten Stream ein großer Buffer zugewiesen, so können die nachfolgenden Streams diesen Buffer auch verwenden
- Buffer werden im Auftrag vom ersten Stream regelmäßig aktualisiert
- die Anzahl der Benutzer wird ständig überwacht, damit durch die Bufferaktualisierung keine Blöcke überschrieben werden, die gerade konsumiert werden
- Während eines Enrollment Windows werden Betrachter eines gleichen Objekts akzeptiert

### (b) Aufgeteilter Buffer

- Verbesserung durch das Aufteilen des Buffers und das Zuweisen jedes Bufferteils an eine Betrachtergruppe
- mehrere Enrollment Windows
- weitere Verbesserungen durch
  - angepasste Aufteilung der Buffer
  - kommt ein neuer Betrachter nachdem alle Enrollment Windows abgeschlossen wurden, wird eine neue Betrachtergruppe erstellt, dem unbenützter Bufferplatz zugewiesen wird



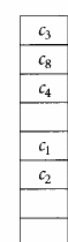
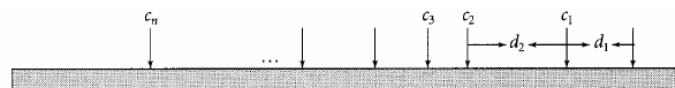
(a)



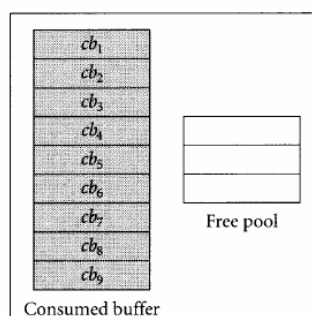
(b)

## Beschreiben Sie kurz die DISTANCE Policy. Wozu dient das Verfahren?

- macht expliziten Gebrauch von den Distanz-Verhältnissen zwischen den Streams
- die Distanz  $d_i$  eines Clients  $c_i$  vom nächstfolgenden Client  $c_{i+1}$  wird gemessen nach dem Abstand in Anzahl von Blöcken
- falls ein Client keinen nächstfolgenden besitzt, wird die Distanz sehr groß angenommen
- die Clients werden nach ihren Distanzen aufsteigend sortiert
- in jedem Service-Durchlauf wird zuerst der verbrauchte Buffer aus dem vorigen Durchlauf,  $cb_i$ , in der aufsteigenden Reihenfolge von  $d_i$  freigegeben.
- geringerer Implementationsaufwand als BASIC, weil die Distanzberechnungen sich nur ändern, wenn ein Client angehalten oder gestoppt wird, oder ein neuer dazu kommt



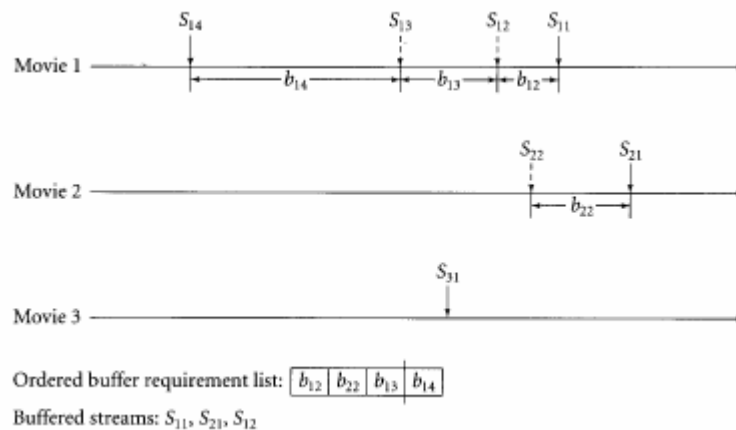
Ordered clients



Consumed buffer

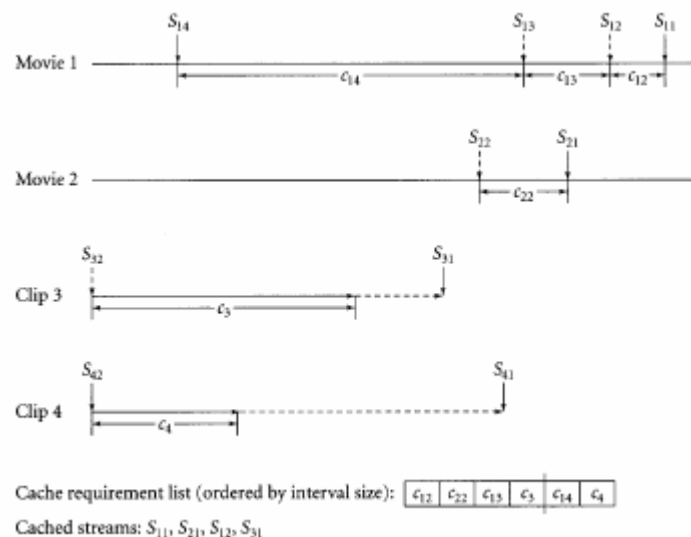
**Beschreiben Sie kurz die General Interval Caching Policy. Worin unterscheidet sie sich vom Interval Caching?**

**Interval Caching:**



- Datenblöcke zwischen zwei aufeinander folgende Streams,  $S_{ij}$  und  $S_{i(j+1)}$  (preceding- und following-Stream), werden als Intervall bezeichnet.
- Durch das Zwischenspeichern des aktuellen Intervalls, dienen die Blöcke aus dem preceding-Stream auch dem following-Stream.
- Die Intervallgröße wird aus der Zeitdifferenz der Zugriffe beider Streams auf denselben Block geschätzt.
- Für den Cache-Bedarf eines Intervalls benötigt man eine Menge an Blöcken.
- Um den Zugriffsquotienten auf den Cache zu maximieren und die I/O zu minimieren, werden die Intervalle nach Größe geordnet und die kürzesten zwischengespeichert.
- Geringe Implementationskosten, der Cache ändert sich nur durch Hinzufügen oder Entfernen von Streams.

**General Interval Caching:**



- IC eignet sich nicht besonders für kleine Medienobjekte, da diese keine Intervalle bilden können.
- Die GIC-Methode integriert in sich nun stream-bewusstes Cache-Management sowie Zeitlagen.
- Bei kleinen Objekten wird ein Zeitintervall zwischen den beiden aufeinander folgenden Zugangspunkten desselben Objekts erzeugt.
- der Cache-Bedarf ist jedoch so groß wie das Objekt; in dem Fall kommt das gesamte Objekt in den Cache
- Die Größe des letzten Intervalls wird als voraussichtliche Intervallgröße angenommen.

## 2) Visual Retrieval

Beschreiben Sie kurz die Architektur eines Image Retrieval Systems.

### 1. Input-Modul

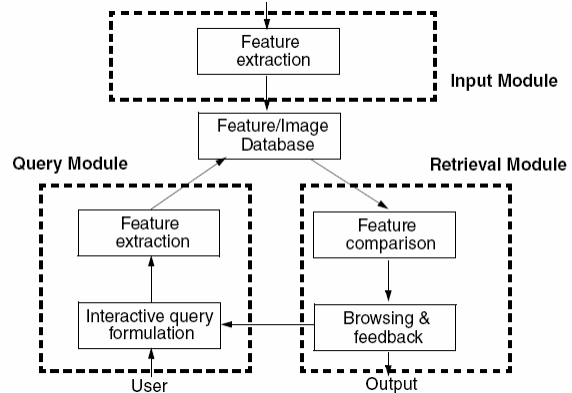
Die entsprechenden Features der Bilder werden extrahiert und zusammen mit ihnen in der Datenbank gespeichert

### 2. Query-Modul

Bei der Suche nach Bildern formuliert der Benutzer interaktiv die entsprechenden Anfragen (RBR, ROA, RSC, ...). Aus diesen Anfragen werden die für die Suche entsprechenden Features in der Datenbank ermittelt und als Basis für die Suche verwendet

### 3. Retrieval-Modul

Die aus der Benutzeranfrage ermittelten Features werden mit denjenigen der Bilder verglichen. Dieser Vergleich kann eine große Anzahl an Bildern in der Ausgabe ergeben. Das Ergebnis kann im nächsten Schritt noch manuell durchsucht und entsprechend weiter eingeschränkt werden. Solange bis das gewünschte Ergebnis erreicht wurde

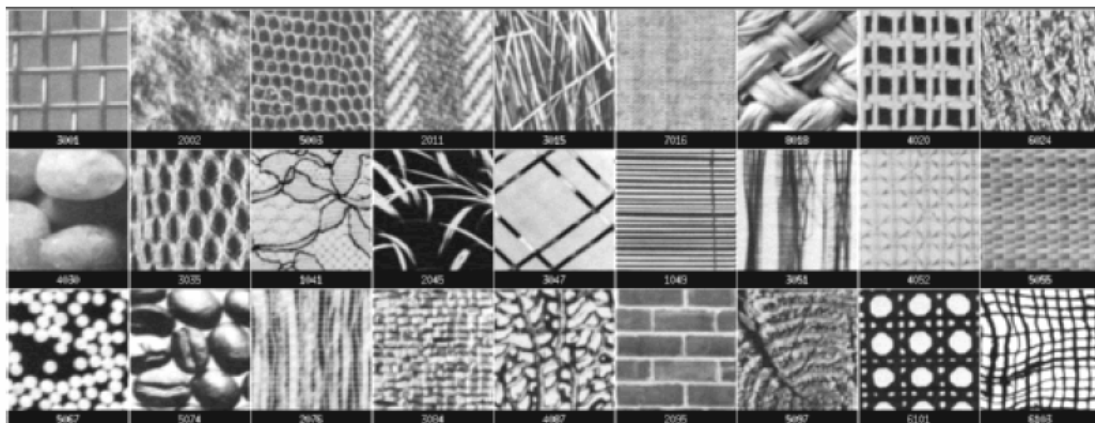


### Zusatzinfo über RBR, ROA, RSC

- RBR (retrieval by browsing) – die Suche erfolgt mittels Browsen durch Bilder oder Thumbnails
- ROA (retrieval by objective attributes) – eine Anfrage wird über Meta-Attribute, logische Attribute oder einer Kombination formuliert
- RSC (retrieval by spatial constraints) – Anfragen basieren auf relativ räumlichen Verknüpfungen unter Bildobjekten, z.B. Richtungs-Verhältnis, Nähe, Überlagerung, Eingrenzung
  - relaxed RSC queries – sucht Bilder, die so viele gewünschte räumliche Verhältnisse wie möglich erfüllen (Klassifizierung notwendig)
  - strict RSC queries – sucht Bilder, die genau die räumlichen Verhältnisse erfüllt, die in der Anfrage angegeben wurden

**Erklären Sie kurz, wie verschiedene auf Texturfeatures beruhende Retrievalverfahren mit Hilfe der Brodatzdatenbank verglichen werden können.**

Die Brodatzdatenbank enthält 112 512x512 Graubilder, von denen jedes eine andere Texturklasse repräsentiert.



Aus der Mitte von jedem Brodatz-Bild werden 9 Subbilder gewonnen. Features werden dann berechnet, indem man den Textur-Algorithmus, der ausgewertet werden soll an jedem der insgesamt 1008 Subbilder anwendet. Danach wird die Distanz jedes Test-Subbildes zu allen anderen berechnet.

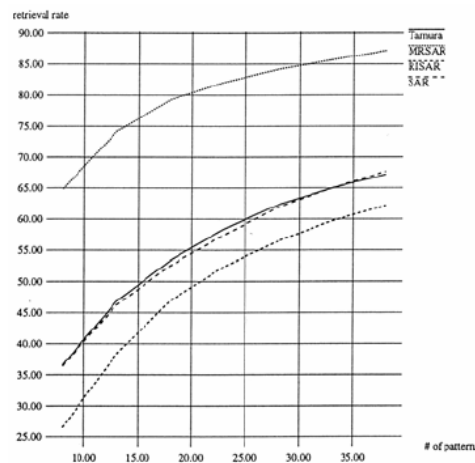
2 Ergebnisse:

- Retrieval Rate: Prozentanteil der n nächsten Subbilder, die zu der selben Gruppe wie das Test-Subbild gehören
- Average Retrieval Rate: wird berechnet indem jedes Subbildes einmal als Test-Subbild errechnet wird

Textur-Algorithmen:

- Tamura
- SAR
- RISAR
- MRSAR

und deren Average Retrieval Rate:



**Erklären Sie kurz die wichtigsten Performance-Evaluierungsmaße und ihren Zusammenhang.**

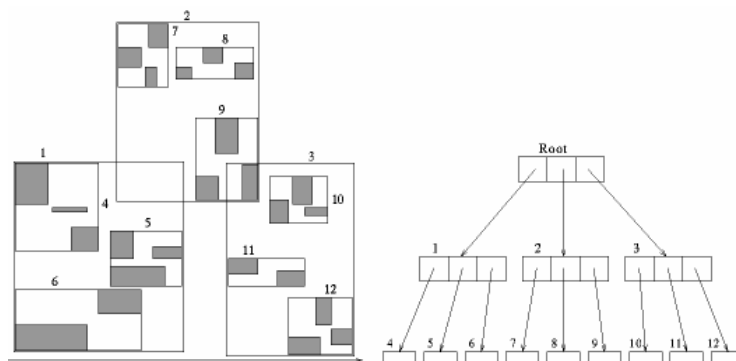
Die üblichen Maße zur Performance-Evaluierung eines Retrieval-Systems sind

- Recall:  
Verhältnis „Anzahl *erhaltener und relevanter* Elemente“ zu „Anzahl *relevanter* Elemente“
- Precision:  
Verhältnis „Anzahl *erhaltener und relevanter* Elemente“ zu „Anzahl *erhaltener* Elemente“

Verbessert man den Recall, verschlechtert sich die Precision und umgekehrt.

Der User muss entscheiden, welches Maß vorzuziehen ist.

**Erklären sie kurz R-Bäume. Was ist ihre Aufgabe?**



- Die Indexstruktur des R-Baumes kann für höhere Dimensionen effizient sein, wenn der Fan-Out > 2 und die Dimensionalität < 20 bleibt.
- ein geometrisches Objekt wird durch sein Minimum Bounding Rectangle (MBR) dargestellt
- Knoten enthalten Einträge der Form (R, ptr), wobei ptr ein Pointer zum Child-Knoten ist und R der MBR, der alle Rechtecke der Childknoten umfasst
- Blätter enthalten Einträge der Form (obj-id, R), wobei obj-id ein Pointer zur Objektbeschreibung ist und R das MBR des Objekts
- die hauptsächliche Neuheit ist, dass Vaterknoten überdecken dürfen. Dadurch kann der R-Baum garantieren, dass zumindestens 50 % des Platzes ausgelastet wird.

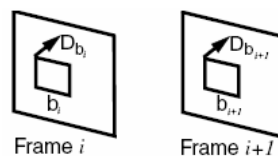
R-Bäume bieten eine mehrdimensional indizierende Technik, die für schnelle Query-Verarbeitungen verwendet wird.

**Nennen Sie 3 Verfahren zur Schnitterkennung in Videos. Erklären Sie kurz die Arbeitsweise eines dieser Verfahren.**

- **Pixelvergleich**  
Die Pixel von zwei aufeinander folgenden Frames werden verglichen. Zwei Pixel werden als unterschiedlich beurteilt, wenn der Unterschied in der Intensität einen gegebenen Schwellwert  $t$  übersteigt.  
Wenn die Anzahl der unterschiedlichen Pixel einen gegebenen Prozentsatz übersteigt, wird der Framewechsel als Schnitt erkannt.
- **Likelihood Ratio**
- **Histogramm-Techniken**

**Was versteht man unter Motion Continuity und wie wird sie berechnet?**

- Optical flow – Vektorfeld in dem jeder Vektor die Geschwindigkeit des entsprechenden Blocks auf der Bildfläche repräsentiert
- Bewegungsvektoren werden durch Block Matching berechnet
- Block Matching: Jeder Block  $D_{b_i}$  in Frame  $i$  wird entlang den Vektor  $b_i$  versetzt



Motion Continuity ist ein Maß für die Motion *Smoothness* – definiert als das Verhältnis von Geschwindigkeit zur Bewegung

- für jeden block  $b_i$ ,  $w_{1i}(b)$  ist definiert als

$$w_{1i}(b) = \begin{cases} 1 & \text{if } |D_{b_i}| > t_s \\ 0 & \text{otherwise} \end{cases}$$

- ähnlich,  $w_{2i}(b)$  ist definiert als

$$w_{2i}(b) = \begin{cases} 1 & \text{if } |D_{b_{i+1}} - D_{b_i}| > t_m \\ 0 & \text{otherwise} \end{cases}$$

- $t_s$  und  $t_m$ ... kleine Schwellwerte
- $\sum_b w_{1i}(b)$  – Anzahl signifikanter Bewegungsvektoren in Frame  $i$
- $\sum_b w_{2i}(b)$  – Anzahl Bewegungsvektoren in Frame  $i$  welche signifikant von ihrem entsprechendem Vektor in Frame  $i+1$  abweichen
- die *Smoothness*  $W_i$  von Frame  $i$  –

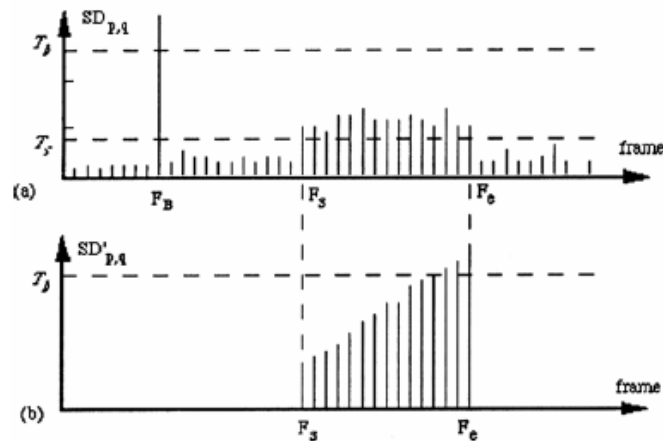
$$W_i = \frac{\sum_b w_{1i}(b)}{\sum_b w_{2i}(b)}$$

wenn  $W_i < T \rightarrow$  Kameraschnitt

Blöcke einer einzelnen Region verschieben sich = Objektbewegung  
Alles verschiebt sich (einheitlich) = Kamerabewegung

**Erklären Sie kurz die Funktionsweise des Twin-Comparison Approach**

- benötigt zwei Begrenzungsschwellwerte:  $T_b$  wird für die Kameraschnitterkennung verwendet,  $T_s$  für die Erkennung eines Special-Effects
- sobald eine Differenz zwischen  $T_s$  und  $T_b$  fällt, wird der entsprechende Frame markiert und mit den nachfolgenden Frames verglichen
- der End-Frame des Übergangs erkannt, wenn die Differenz zwischen auf einander folgender Frames und  $T_s$  fällt, während der Vergleich mit dem markierten Frame auf einen Wert größer als  $T_b$  gestiegen ist



- $SD_{p,q}$ : die Differenz aufeinanderfolgender Frames, die durch das Differenz-Maß definiert wird
- $SD'_{p,q}$ : wird berechnet, wenn  $T_s < SD_{p,q} < T_b$ . die angehäuften Differenzen zwischen dem aktuellen Frame und dem potentiellen Start-Frame  $F_s$  eines Übergangs (Transition).
- $T_s$ : der Schwellwert um den Start-Frame  $F_s$  eines Übergangs zu berechnen
- $T_b$ : der Schwellwert um den End-Frame  $F_e$  und die Kameraschnitte zu berechnen

Immer wenn eine Differenz  $SD_{p,q}$  zwischen  $T_s$  und  $T_b$  fällt, wird der entsprechende Frame als potentieller Start  $F_s$  eines Übergangs markiert und mit den folgenden Frames verglichen. Der End-Frame  $F_e$  des Übergangs wurde erreicht, wenn die Differenzen  $SD_{p,q}$  zu den nachfolgenden Frames kleiner werden als  $T_s$ , während die angehäuften Differenzen  $SD'_{p,q}$  zum Start-Frame  $F_s$   $T_b$  überschreiten.

### Formulieren Sie in Pseudocode einen Algorithmus für den Twin-Comparison Approach.

Für jeden Frame

Differenz zu nachfolgenden Frame berechnen (in  $SD$ )

Wenn  $T_s$  (unterer Schwellwert)  $< SD < T_b$  (oberer Schwellwert)  
 aktuellen Frame als Startframe merken  
 Übergangserkennung aktiv

Wenn Übergangserkennung aktiv  
 $SD$  akkumulieren (in  $SD'$ )  
 Wenn  $SD < T_s$  und  $SD' > T_b$   
 Nächsten Frame als Endframe merken

Wenn Übergangserkennung inaktiv und  $SD > T_b$   
 Frame als harten Schnitt merken

### Beschreiben Sie kurz den Multi-Pass-Ansatz zur Videoschnitterkennung. Weshalb verwendet man diesen Ansatz?

Verfahren zur Reduzierung der Bearbeitungszeit in der Videoschnitterkennung.

#### Erster Durchgang:

- es wird ein niedriges  $T_b$  bestimmt.
- Auflösung wird vorübergehend reduziert (zB werden bei einem Video mit 25 frames/sec nur 5 frames/sec geprüft) → das verstärkte Überspringen erhöht  $SD_{p,q}$  bei Übergängen
- Sowohl Kameraschnitte als auch Übergänge werden erkannt (inklusive Fehlerkennungen)

#### Weitere Durchgänge:

Zur genaueren Lokalisierung der Segmentgrenzen wird vorübergehend eine erhöhte Auflösung verwendet.

**Erklären Sie kurz in welchem Kontext die Erkennung von Kamerabewegungen eine Rolle spielt. Erklären Sie kurz die prinzipielle Arbeitsweise der Erkennung.**

Kamerabewegungen können ähnliche Werte erzeugen wie Übergänge. Aus diesem Grund ist es notwendig Muster zu erkennen, die durch Kamerabewegungen auftreten. Dazu gibt es zwei Verfahren:

**Motion Vector Analysis**

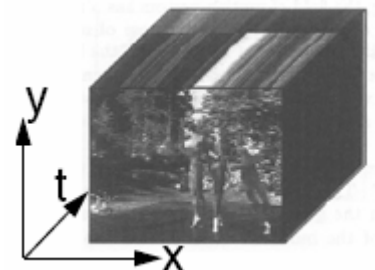
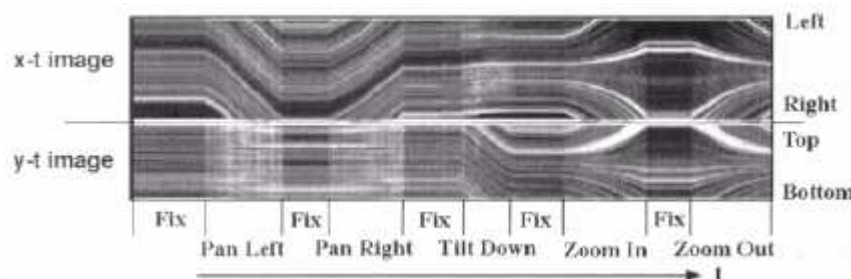
1. Frame  $i$  wird in Blöcke aufgeteilt
2. *block matching* – Jedem Block wird ein Vektor zugewiesen, der den Block versetzt
3. *optical flow* – die Vektoren erzeugen ein Vektorfeld, in dem jeder Vektor die Geschwindigkeit des entsprechenden Blocks auf der Bildfläche darstellt
4. die Verteilung der Bewegungsvektoren in einem Frame kann
  - panning und tilting (Schwenken und Kippen der Kamera)
  - zooming (in und out)
  - und vordefinierte Bewegungsmuster erkennen

**Video X-Ray**

Raum-Zeitliches Bild:

Annahme: Hintergrund enthält Unterscheidungsmerkmale/Auffälligkeiten

- panning – Schräge Linien in der Draufsicht
- tilting – Schräge Linien in der Seitenansicht
- zooming – Linien laufen auseinander oder zusammen



**Erklären Sie kurz 2 Ansätze zur Erkennung der Kameraoption „Zoom“ und nennen Sie Vor- und Nachteile der Ansätze.**

- **Motion Vector Analysis** [siehe oben](#)
  - **Vorteil:** Zoom leicht zu erkennen
  - **Nachteil:** Fokus muss im Frame-Zentrum liegen
- **Video X-Ray** [siehe oben](#)
  - **Vorteil:** flexibler bezüglich Fokus

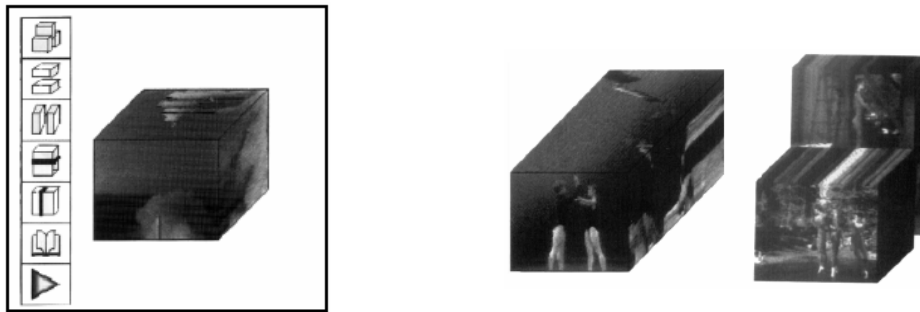
**Beschreiben Sie kurz 3 Werkzeuge für die Inhaltsanzeige (Content Indication) von Video**

Content Indication ist ein Verfahren, um bestimmte Teile eines Contents besser zu erfassen.

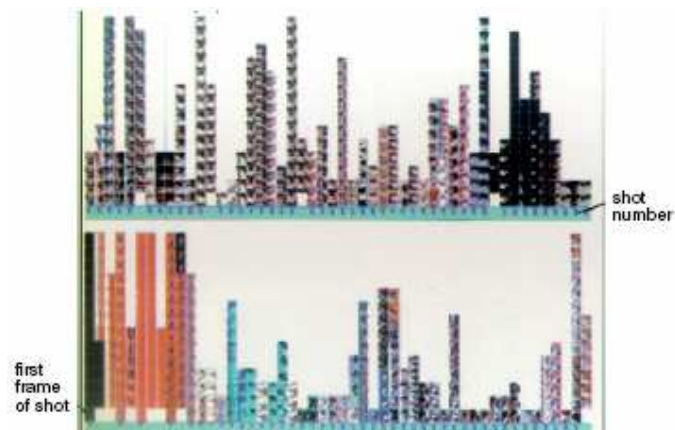
- **movie icon (micon)** Video wird als 3D-Gebilde dargestellt



- **interactive micons** Auswertungs- und Bedienungs-umgebung



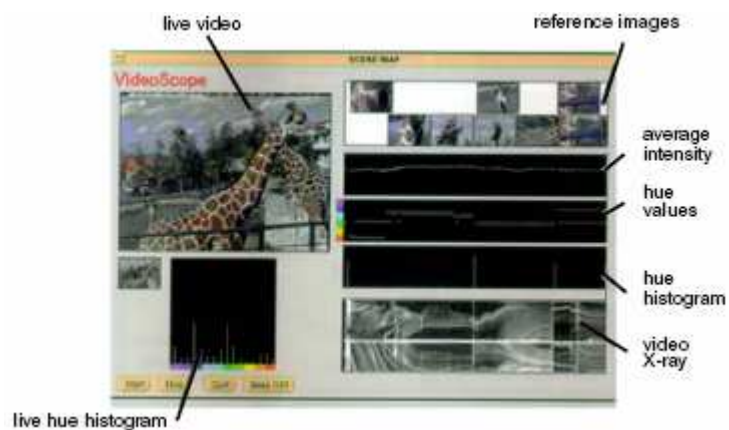
- **paper video** auf Diagrammen basierender Video-Browser



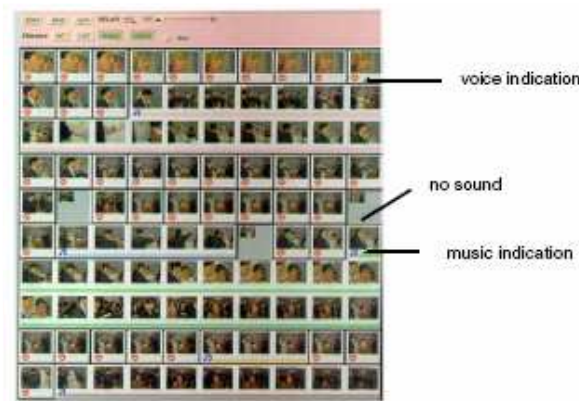
- **video panorama** Anzeige des festgehaltenen Videoraumes



- **videoscope** Analyseprogramm für Video-Content



- **sound browser** zeigt das Vorhandensein des jeweiligen Sounds (Stimme, Musik...)



**Welche Verfahren zur Schnitterkennung in MPEG-komprimierten Videos kennen Sie? Erklären Sie kurz die Arbeitsweise eines dieser Verfahren.**

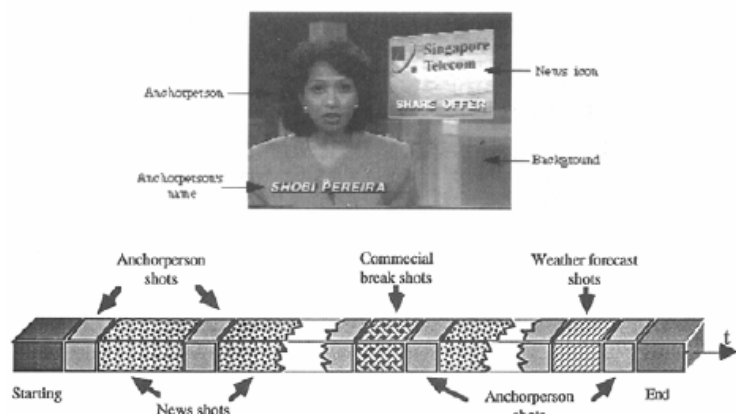
- Algorithmen basierend auf DCT-Koeffizienten
  - DCT Coefficients Correlation
  - DCT Block Comparison
- Algorithmus basierend auf Motion-Vectorer
- Hybrid Approach to Partitioning

#### **Algorithmus basierend auf Bewegungsvektoren**

- Motion-Vecotoren im MPEG-Datenfluss:
  - P-Frames – einzelne Gruppe von Motion-Vektoren
  - B-Frames – zwei Gruppen von Motion-Vektoren (vorwärts und rückwärts)
- Feld von Motion-Vektoren in einem Video:
  - Innerhalb einer Aufnahme – relativ kontinuierliche Änderungen
  - zwischen verschiedenen Aufnahmen – Kontinuität wird gestört
- Definition M:
  - P-Frame: Zahl zulässiger Motion-Vektoren
  - B-Frame: kleinere Zahl zulässiger Motion-Vektoren (vorwärts und rückwärts)
- $M < T_b \dots$  Keraschnitt

**Erklären Sie kurz die Klassifikation von Shots für das Parsing von News Videos.**

- definiere eine Vorlage eines A-Shots (Anchorperson-Shot)
- erstelle Maße für die Ähnlichkeit, die man mit diesen Vorlagen vergleichen kann
- zeitliche Gliederungsvorlage einer gesamten Nachrichtensendung – Erkennung von Sequenzen und Episoden



### 3) Content Description

#### Was versteht man unter einer MPEG-7-Description?

Eine Description besteht aus einem Description Scheme (Struktur) und einer Reihe von Description-Werten (Umschreibungen), die die Daten beschreiben. Eine Description beinhaltet oder zeigt auf ein vollständig oder teilweise umschriebenes Description Scheme

#### Nennen und erklären Sie kurz 3 Anforderungen an die MPEG-7 DDL.

- **Compositional capabilities:** Die DDL liefert die Fähigkeit neue 'Description Scheme'n und 'Descriptor'en zu bilden, wobei 'Description Scheme'n auch aus mehreren 'Description Scheme'n gebildet werden kann. Ein neu erstelltes 'Description Scheme' muss die Erstellung von MPEG-7-konformen 'Description'en erlauben.
- **Transformational capabilities:** Die DDL erlaubt das Wiederverwenden, Erweitern und Vererben von bestehenden 'Descriptor'en und 'Description Scheme'n
- **Unique identification:** Die DDL bietet Mechanismen zur eindeutigen Identifizierung von 'Description Scheme'n und 'Descriptor'en, wodurch diese eindeutig bezeichnet werden können.
- **Data types:** Die DDL bietet einen Satz von einfachen Datentypen, wie z.B. text, integer, real, data, time/time index, version, um zusammengesetzte Datentypen kurz und bündig zu beschreiben. Auch wenn diese vielleicht aus aufbereiteten Digitalsignalen (z.B. Histogramme, Grafen, RGB-Werte) hervorgehen. Weiters muss die DDL einen Mechanismus bieten, um 'Descriptor'en Daten von mehreren Medientypen zuordnen zu können, die eine inhärente Struktur aufweisen (z.B. Audio, Video, Audio-visuelle Präsentationen etc.)

## 4) Multimedia Standards

**Erklären Sie kurz die Begriffe räumliche und zeitliche Komposition. Wie ist das Konzept Komposition in SMIL realisiert?**

Komposition ist die Spezifikation von zeitlichen oder räumlichen Beziehungen zwischen Medienobjekten.

Bsp (räumliche Komposition):

- zwei Bilder sollen in demselben Fenster präsentiert werden
- ein Bild soll über ein anderes gelegt werden – falls sie sich überdecken etc.

SMIL: zB mittels der `<region>`-Tags

Bsp (zeitliche Komposition):

- ein Video und seine Vertonung sollen gleichzeitig ausgegeben werden
- zwei Videos sollen nacheinander abgespielt werden etc.

SMIL: zB mittels der `<seq>`, `<par>` und `<excl>`-Tags

**Geben Sie (in SMIL-Syntax) ein Beispiel für eine der unter der oberen Frage genannten Relationen zweier Medien “Video 1” und “Video 2” an.**

Bsp.: „sequenz“

```
<seq>
  <video id="Video 1" src="video1.mpg" dur="6s" begin="0s" />
  <video id="Video 2" src="video2.mpg" dur="6s" begin="0s" />
</seq>
```

**Was versteht man unter Personalisierung?**

Personalisierung bezeichnet das Anpassen eines Dienstes, eines Programmes oder eines Informationsprodukts an die Bedürfnisse, Abneigungen und Fähigkeiten des Benutzers.

Die Anpassung kann explizit oder implizit erfolgen. Entweder der Benutzer passt zB die Webseite selbst seinen Bedürfnissen an (zB Farbe, Schrift, usw.) oder sie (die Webseite zB) passt sich von selbst dem Verhalten (zB Kaufverhalten bei Online-Shops) an.

**Welche Möglichkeiten zur Personalisierung bietet der SMIL-Standard?**

SMIL bietet nur geringe Möglichkeiten der Personalisierung.  
Diese bietet das `<switch>`-Element.

Bsp.:

```
<switch>
  <audio src="http://www.webseite.com/datei1.rm" system-bitrate="250000"/>
  <audio src="http://www.webseite.com/datei2.rm" system-bitrate="100000"/>
  <audio src="http://www.webseite.com/datei3.rm" system-bitrate="40000"/>
  <audio src="http://www.webseite.com/datei4.rm" system-bitrate="1000"/>
</switch>
```

Das erste Element, dessen Testattribut als wahr erkannt werden, wird ausgeführt. In diesem Fall ist das Testattribut die `system-bitrate` (Übertragungsgeschwindigkeit in bps).

Weitere wären

- `system-captions`: on, Untertitel werden übertragen; off, Untertitel werden nicht übertragen
- `system-language`: Sprachkürzel; mehrere Sprachangaben können durch Kommata getrennt angegeben werden

- *system-overdup-or-captions*: caption, Untertitel werden gesendet; overdup, Sprachsynchronisation wird gesendet
- *system-required*: zur Zeit nicht besetzt
- *system-screen-depth*: Zahl der Farben der Farbpalette in bits
- *system-screen-size*: Bildschirmgrößenangabe nach dem Schema screen-size-val ::= screen-height"X"screen-width

**Nennen Sie einen weiteren Web-Standard zur Personalisierung und beschreiben Sie ihn kurz (Ziele, Aufbau).**

## CC/PP

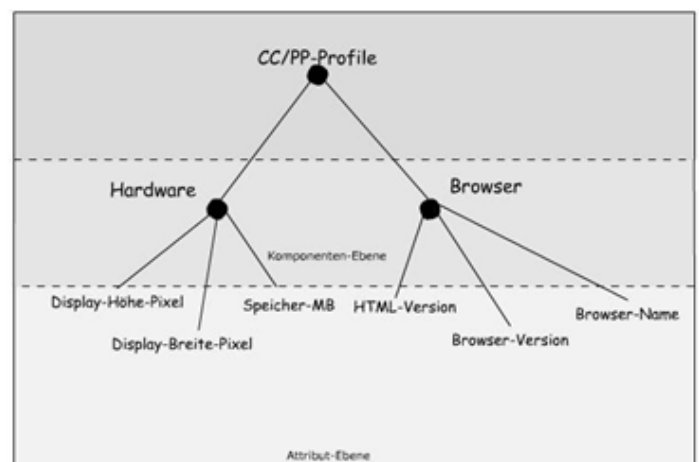
ein **Hauptziel** des W3C: universeller Webzugang:

Unabhängigkeit von

- Hardwareausstattung
- Softwareausstattung
- Netzinfrastruktur
- Sprache
- Kultur
- geologische Lage

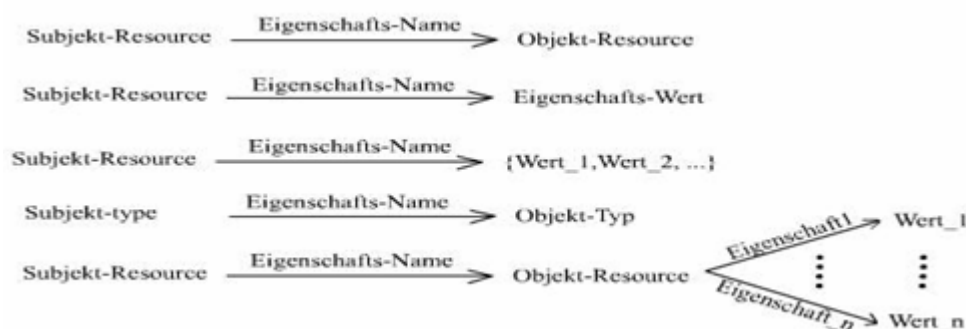
## CC/PP-Dokument-Struktur

- 2-Ebenen-Hierarchie
  - ein CC/PP-Profil hat eine oder mehrere Komponenten  
z.B. Software-Plattform, Browser, ...  
jede Komponente besitzt ein oder mehrere Attribute  
z.B. Versionsnummern, Namen, Werte
  - Zusätzlich: Defaults  
ermöglicht das Setzen von Default-Attributwerten



## CC/PP und RDF

- 2-Ebenen-Struktur ist ein gerichteter Graph
- CC/PP-Dokumente werden mittels RDF/XML beschrieben.



[weiteres unten](#)

## Was ist und wozu dient CC/PP?

CC/PP (*Composite Capabilities / Preference Profiles*) ist eine Standardprofilsprache zur Darstellung von Webinhalten auf unterschiedlichen Geräten

- Beschreibung der Fähigkeiten der Geräte
- Beschreibung der Nutzervorteile

CC/PP ist ein erweiterbarer Rahmen für die Übertragung von Webinhalten zwischen Webserver und Endgerät

- Nutzerbeschreibung
- Kontextbeschreibung
- Gerätebeschreibung

## Welche Abschnitte umfasst der Lebenszyklus eines SMIL-Zeitcontainers? Erklären Sie kurz die Abschnitte.

1. Startup – Erhalten des ersten Intervalls
2. Warten auf den Beginn des aktuellen Intervalls
3. Active time – Intervall abspielen
4. End of an interval – den nächsten Intervall berechnen und Abhängigkeiten melden
5. Post active – führe jegliches „fill“ aus und warte auf einen nächsten Intervall

## Nennen Sie kurz die Vor- und Nachteile von SMIL.

### Vorteile

- Verschiedene Formate können zu einer einheitlichen Präsentation zusammengefasst werden.
- Multimediaobjekte müssen physikalisch nicht auf einem einzigen Server liegen.
- Multilingualität
- Unterstützung unterschiedlicher Bandbreiten.
- SMIL-Präsentationen können sich automatisch an Browser oder Player anpassen.
- einfache Möglichkeit die Präsentationsdauer jedes Multimediaobjekts zu kontrollieren.
- viele Möglichkeiten das Layout einer Multimediapräsentation zu gestalten.

### Nachteile

- mangelnde Unterstützung durch Hersteller
- mangelnde User Agents mit verwirrendem Sprachumfang und Bugs

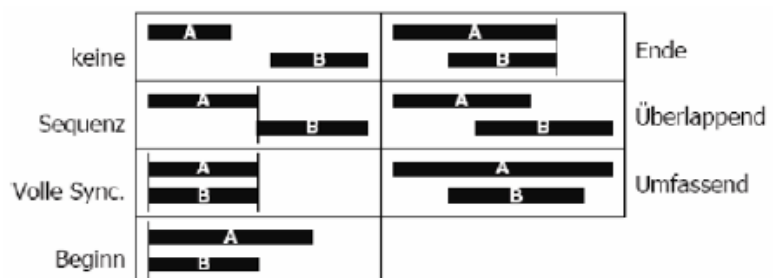
## Welche Elemente und Attribute bietet SMIL zur zeitlichen Synchronisation von Medien?

Elemente: par, seq, excl

Attribute: begin, end, dur, fill, endsync

## Nennen Sie vier mögliche zeitliche Relationen von Medien und beschreiben Sie sie kurz.

1. Sequenz: Medienobjekte werden hintereinander abgespielt
2. Ende: Medienobjekte enden gleichzeitig
3. Überlappend: Medienobjekte überschneiden sich
4. Volle Synchronisation: Medienobjekte beginnen und enden gleichzeitig



## 5) Digitales Fernsehen

### **Beschreiben Sie kurz Aufgabe und Umfang der Multimedia Home Platform.**

Der Multimedia Home Platform-Standard, kurz MHP, wurde vom internationalen DVB-Projekt als Standard verabschiedet und spezifiziert die Übertragung und Darstellung interaktiver Inhalte im Digitalen Fernsehen auf Basis der Programmiersprache Java.

- MHP ist ein offener Standard der auf JAVA basiert
- MHP standardisiert die Schnittstelle zwischen der Software der Set-Top-Box und dem TV-Gerät
- MHP ist unabhängig von Lizenzen
- MHP Set-Top-Boxen ermöglichen den Empfang und die Darstellung von Anwendungen in einem von Anbieter, Urheber und Sender neutralen Framework
- die Applikationen vieler Anbieter werden vollständig kompatibel zu den verschiedensten MHP-Anwendungen sein, und das in einem horizontalen Markt
- MHP will Low-End-Produkte zu High-End-Produkten machen
- MHP besteht aus User-Terminal (Set-Top-Box, PC, zugehörige Peripherie-Geräte), Standard-Middleware und einer Gruppe von APIs, die eine Reihe von Diensten unterstützen können
- MPH basiert auf Internet-Standards und ermöglicht so Kompatibilität und Annäherung zwischen TV und Internet

Mehrwert für die User:

- Erweitertes Fernsehen - zB HTML-basierender Teletext; liefert viel mehr Informationen in einem professionellem Layout
- Interaktives Fernsehen – ermöglicht Interaktivität wie z.B. Televoting (dazu benötigt die Set-Top-Box nur einen Rückkanal)
- Internetdienste – vollständige Verknüpfung zwischen TV-Gerät und Internet (www, E-Mail-Dienste, ...)

### **Was sind die wichtigsten Features von Set-Top-Boxen?**

- Entschlüsselung einkommender digitaler Signale
- Überprüfung von Zugriffsrechten und Sicherheitsstufen
- Bilder in Kinoqualität am TV-Gerät
- Digitaler Surround Sound und
- Ausführen und Rendern von Internetdiensten und interaktiven Leistungen

### **In welchen Kategorien werden Set-Top-Boxen eingeteilt?**

- Broadcast TV
  - nur fähig Daten aus einem MPEG2-Stream zu empfangen
- Enhanced TV – hat zusätzlich einen Rückkanal
  - VOD, e-commerce und Internet-Browsen möglich
- Advanced Services – ähnlich einem MM-Desktop
  - High-Speed-Rückkanal: Zugriff zu Internetdiensten und Interaktiven Leistungen bei sehr hohen Geschwindigkeiten

## **KOMMT NICHT?**

**Fragen, die ausscheiden sollten, da sie nun Stoff von Multimedia I sind (der Vollständigkeit zuliebe trotzdem angeführt):**

- Worin unterscheiden sich Ihrer Meinung nach die im vorgestellten Framework definierten Components vom allgemeinem Begriff Komponenten wie sonst im Software Engineering verwendet? Was haben die beiden Konzepte gemein?
- Nennen und diskutieren Sie kurz zwei Gründe für das Konzept Ableitung (Derivation).
- Nennen Sie 3 Gründe, weshalb die Multimedia Programmierung fast ausschließlich dem objektorientierten Ansatz folgt.
- Worin unterscheiden sich zeitabhängige Medienobjekte von anderen Medienobjekten?
- Was versteht man unter padding und weshalb wird padding eingesetzt?
- Was ist die Aufgabe der Klasse Transform im vorgestellten MM-Framework?
- Erklären Sie kurz die Konzepte Interpretation, Ableitung (Derivation) und Komposition. Wie hängen diese Konzepte zusammen?
- Erklären Sie kurz den Begriff Quality of Service.
- Was ist ein Framework und welche Anforderungen werden an ein Multimediaframework gestellt?
- Beschreiben Sie kurz die Aufgabe der Formatklassen im vorgestellten Framework.
- Nennen Sie 2 Aufgaben der Klasse Component im vorgestellten MM-Framework.
- Wozu dienen die Ports von Komponenten? Unter welchen Umständen dürfen Ports von verschiedenen Komponenten verbunden werden?
- Welchen Zweck dienen function objects? In welchen Klassen des Frameworks werden sie implementiert?
- Nennen und erklären Sie die wichtigsten Eigenschaften, nach denen zeitabhängige Ströme (timed Streams) klassifiziert werden können.
- Was versteht man unter der „dualen Natur“ zeitanhängiger Medien?
- Was versteht man unter Konfiguration und wie wird das Konzept im vorgestellten Framework realisiert?
- Was sind function objects? In welchen Klassen des Frameworks werden sie Implementiert?