

Multimediatdatenbeschreibung am Beispiel MPEG-7

Medieninformatik IL

Andreas Unterweger

Vertiefung Medieninformatik
Studiengang ITS
FH Salzburg

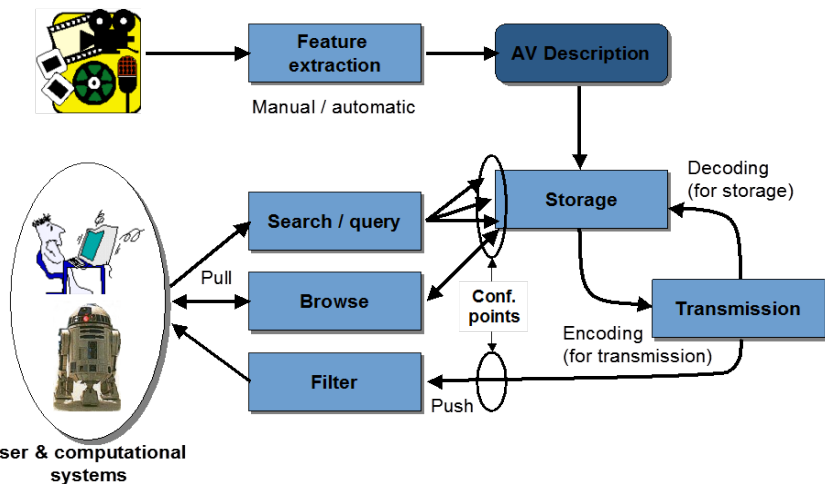
Wintersemester 2019/20

- **Kodierte Multimediadaten sind schwer zu durchsuchen**
 - Bild- bzw. Audiodaten sind stark komprimiert
 - Dekodierte Rohdaten sind zu umfangreich
 - Metadaten beschränken sich auf zur Kodierung notwendiges Minimum
 - Zusatzdaten sind (zumeist) optional
- **Containerformate speichern oft nur rudimentäre Metadaten**
 - Autor
 - Erstellungsdatum
 - Dauer
 - Bildgröße
 - Abtastfrequenz
 - ...

→ Zusätzliche Metadaten­speicherung in relativ einfachem Format

- Durchsuchbarkeit
 - Eigenschaftsspeicherung
 - Ermöglichung gezielter Suche
 - Interpretationsmöglichkeit ohne menschliche Unterstützung
- Einheitliche Informationsspeicherung
 - Maschinenlesbarkeit
 - Austauschbarkeit zwischen Systemen
 - Abbildung verschiedener Metadatenebenen
 - Konvertierbarkeit (z.B. Text in Sprache)
- Formatunabhängigkeit
 - Daten können beliebig kodiert sein
 - Metaebene erlaubt Bezug auf abstrakte Entitäten
- Abdeckung nahezu aller erdenklichen Anwendungsfälle

Überblick MPEG-7 I



Quelle: Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_Overview_v10.zip (12.10.2013), 2004.

- Eigenschaft (englisch *feature*): Datenspezifische Charakteristik mit einer extern zugeordneten Bedeutung
- Deskriptor (englisch *descriptor*): Repräsentation einer Eigenschaft
- Beschreibungsschema (englisch *description scheme*): Zusammenhang zwischen Deskriptoren bzw. Deskriptoren und Beschreibungsschemata
- Beschreibungsdefinitionssprache (englisch *description definition language*): Sprache zur Definition und Erweiterung von Deskriptoren und Beschreibungsschemata
- MPEG-7 Description Definition Language (DDL): Datenbeschreibungsdefinitionssprache
- MPEG-7 Multimedia Description Schemes (DS): Definition generischer Eigenschaften und deren Zusammenhänge

- eXtensive Markup Language (XML)
- An HTML angelehnte Syntax (aber strenger)
 - Tags (immer zu schließen!)
 - Attribute
 - Kommentare

```
<?xml version="1.0" encoding="UTF-8"?>
<Lehrveranstaltung name="Medieninformatik IL">
  <Vortragender>
    <Vorname>Andreas</Vorname>
    <Nachname>Unterweger</Nachname>
  </Vortragender>
  <Raum id="U SE 456"/>
  <Inhalte>
    <!-- Hier stehen die Inhalte -->
  </Inhalte>
</Lehrveranstaltung>
```

Einschub: XML-Schemata

- Definiert XML-Struktur (in XML-Syntax)
 - Erlaubte Tags inkl. Reihenfolge und min./max. Anzahl
 - Erlaubte Attribute
- Erweiterbar (Schema erweitert anderes Schema)

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
            xmlns="MeinNamensraum"> <!-- Optionaler Namensraum -->
<xs:element name="Vortragender">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Vorname" type="xs:string"
                  minOccurs="0" maxOccurs="1"/>
      <xs:element name="Nachname" type="xs:string"
                  minOccurs="1" maxOccurs="1"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:schema>
```

Einschub: XML-Schemata-Instanziierung

- XML-Dokument instanziiert XML-Schema

→ Verifikation möglich

→ Leichter Datenaustausch durch Schema-Konformität

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<Vortragender
```

```
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
```

```
  xsi:noNamespaceSchemaLocation="MeinVortragendenSchema.xsd">
```

```
  <Vorname>Andreas</Vorname>
```

```
  <Nachname>Unterweger</Nachname>
```

```
</Vortragender>
```

- Angabe eines bestimmten Typs möglich (wenn mehrere definiert sind, wie z.B. bei MPEG-7), wenn Namensraum angegeben wird:

```
xsi:type="Vortragender" xmlns="MeinNamensraum"
```


- Basis für DDL: XML-Schemata
 - Erweiterung um MPEG-7-spezifische Elemente
 - DS sind XML-Schemata
- Definition von Basis- und zusammengesetzten Elementen
- Eigene Erweiterungen möglich
- Deskriptoren sind Instanzen von DS
- MPEG-7-Deskriptoren sind in XML repräsentiert
- XML-Werkzeuge wiederverwendbar

XML-Repräsentation II

- Beispiel: 8 Noten aus „Moon River“ von Henry Mancini
- Repräsentation als XML über MPEG-7-Deskriptoren



Quelle: Westermann, U., Klas, W.: *An Analysis of XML Database Solutions Concerning the Management of MPEG-7 Media Descriptions*, Technical Report TR-2002302, Universität Wien, <http://eprints.cs.univie.ac.at/1127/1/xmlsurvey.pdf>, 2002.

- Zu repräsentierende Elemente:
 - Takt (englisch *meter*)
 - Notenzeitpunkte
 - (Noten-)Tonhöhen (relativ)

XML-Repräsentation III

```
<AudioDescriptionScheme
  xmlns="http://www.mpeg7.org/2001/MPEG-7_Schema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:type="MelodyType">
  <Meter>
    <Numerator>3</Numerator>
    <Denominator>4</Denominator>
  </Meter>
  <MelodyContour>
    <!-- Distance between two notes -->
    <Contour>2 -1 -1 -1 -1 -1 1</Contour>
    <!-- Beat position of notes -->
    <Beat>1 4 5 7 8 9 9 10</Beat>
  </MelodyContour>
</AudioDescriptionScheme>
```

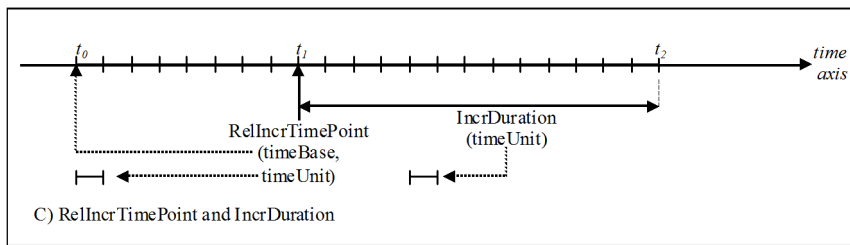
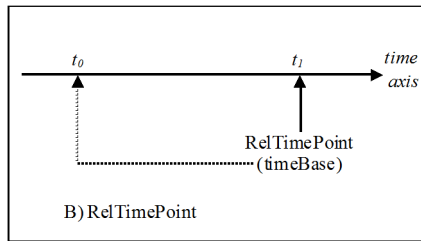
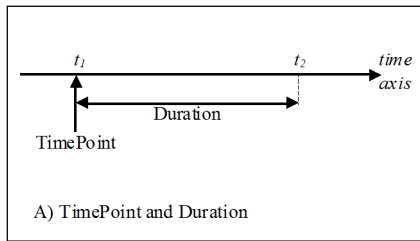
Adaptiert aus Westermann, U., Klas, W.: *An Analysis of XML Database Solutions Concerning the Management of MPEG-7 Media Descriptions*, Technical Report TR-2002302, Universität Wien, <http://eprints.cs.univie.ac.at/1127/1/xmlsurvey.pdf>, 2002.

XML-Repräsentation IV

```
<complexType name="MelodyType">
  <complexContent>
    <extension base="mpeg7:AudioDSType">
      <sequence>
        <element name="Meter" type="mpeg7:MeterType"
          minOccurs="0"/>
        <element name="MelodyContour"
          type="mpeg7:MelodyContourType"
          minOccurs="0"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>
```

Quelle: Westermann, U., Klas, W.: *An Analysis of XML Database Solutions Concerning the Management of MPEG-7 Media Descriptions*, Technical Report TR-2002302, Universität Wien, <http://eprints.cs.univie.ac.at/1127/1/xmlsurvey.pdf>, 2002.

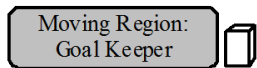
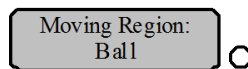
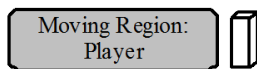
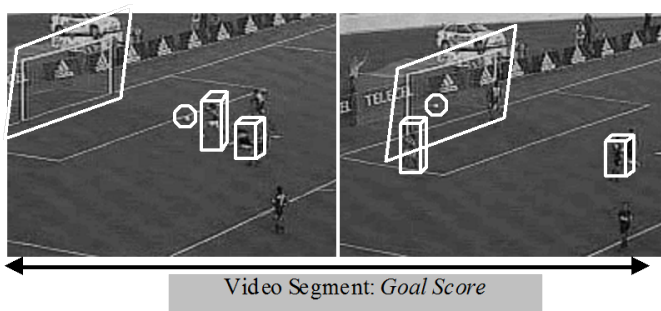
Vordefinierte Grunddatentypen (Beispiel)



Quelle: Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_Overview_v10.zip (12.10.2013), 2004.

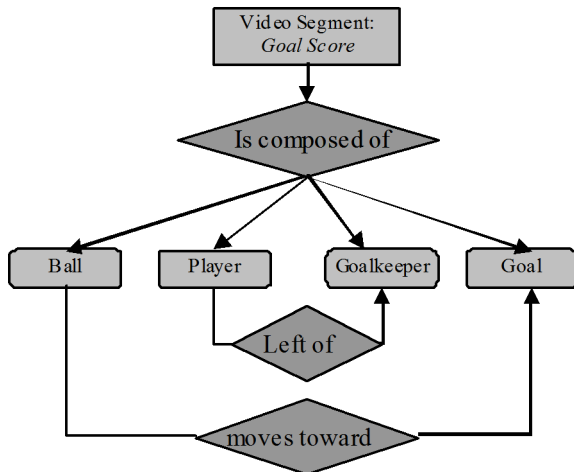
- Für Audio:
 - Durchschnittliche Lautstärke
 - Anzahl und Länge von Sequenzen mit Stille
 - Frequenzspektrum (logarithmisch)
 - Grundswingungsfrequenz
 - Oberswingungsverteilung
- Für Bilder:
 - Dominante Farbe
 - Farbverteilung (Histogrammdeskriptor)
 - Ortsfrequenzverteilung
 - Dominante Kantenrichtungen
 - Verteilung von Kantenrichtungen
- Für Objekte:
 - Position (in absoluten oder relativen Koordinaten)
 - Form

Vordefinierte Strukturdatentypen (Beispiel)



Adaptiert aus Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_Overview_v10.zip (12.10.2013), 2004.

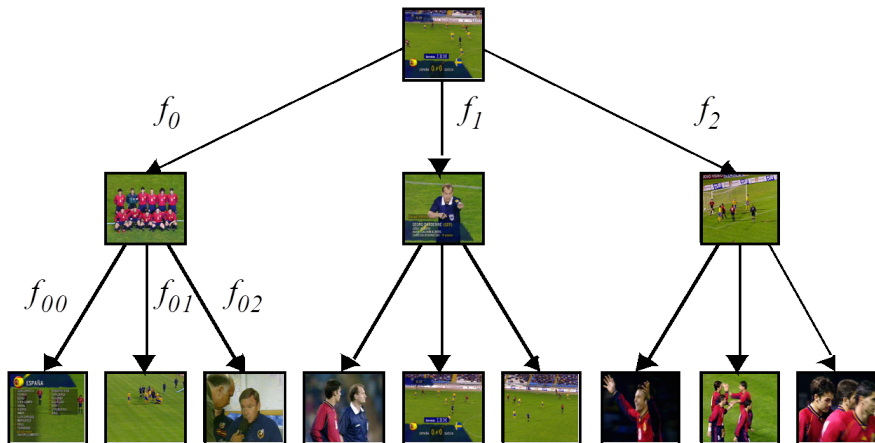
Objektbeziehungen (Beispiel)



Adaptiert aus Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_Overview_v10.zip (12.10.2013), 2004.

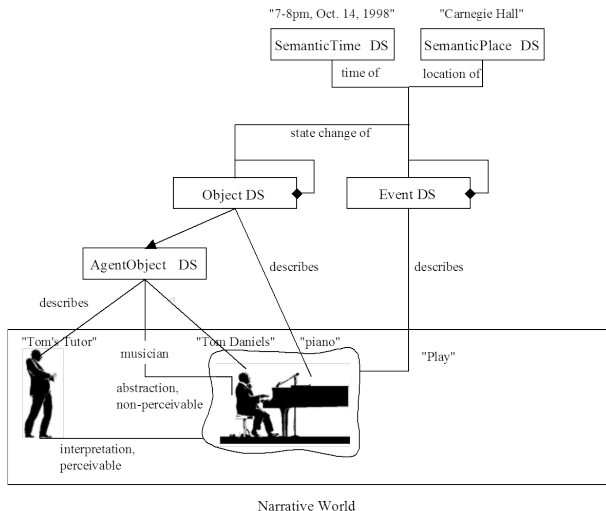
Beschreibung auf struktureller Ebene (Beispiel)

- $f_{i(j)}$: Granularitätslevel der Zusammenfassung



Quelle: Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_Overview_v10.zip (12.10.2013), 2004.

Beschreibung auf konzeptioneller Ebene (Beispiel)



Quelle: Martínez, J. M.: *MPEG-7 Overview (version 10)*. Veröffentlicht als ISO/IEC JTC1/SC29/WG11 N6828 unter http://mpeg.chiariglione.org/sites/default/files/files/standards/docs/w6828_mp7_0overview_v10.zip (12.10.2013), 2004.

- Für Low-Level-Beschreibungen:
 - Ähnlichkeitssuche
 - Datenabgleich
 - ...
- Für High-Level-Beschreibungen:
 - Objekterkennung und -zuordnung
 - Auswertung von Überwachungskameraaufnahmen
 - Langzeitspeicherung von Zusammenhängen
 - Genrezuordnung
 - Alterseinstufung
 - Zugangsbeschränkungsüberprüfung
 - Semantische Suche
 - ...

Fragen?