

Bridging the Semantic Gap  
Seminar

# MPEG-7

Universität Mannheim  
Lehrstuhl für Praktische Informatik IV  
Sommersemester 2006

Victor Schan  
vschan@rumms.uni-mannheim.de

**Einleitung**

**Überblick**

**Struktur**

**Ausgewählte Deskriptoren**

**Distanzmaß**

**Anwendungen**

**Fazit**

MPEG-7 ist

- **kein** Kompressionsstandard wie seine Vorgänger MPEG 1/2.
- **kein** (semi-)automatisches Verfahren zur Merkmalsextraktion aus Multimediadokumenten.



(Quelle: <http://www.members.cox.net>, 2004)

## **MPEG-7:**

- universelle Beschreibungssprache für multimediale Daten
- Informationen über Informationen (Metadaten)
- Metadaten werden in einem XML-Dokument gespeichert.
- XML-Darstellung kann verlustfrei in eine binäre Darstellung umgewandelt werden (z.B. für Streaming).
- Format und Medium spielen keine Rolle.

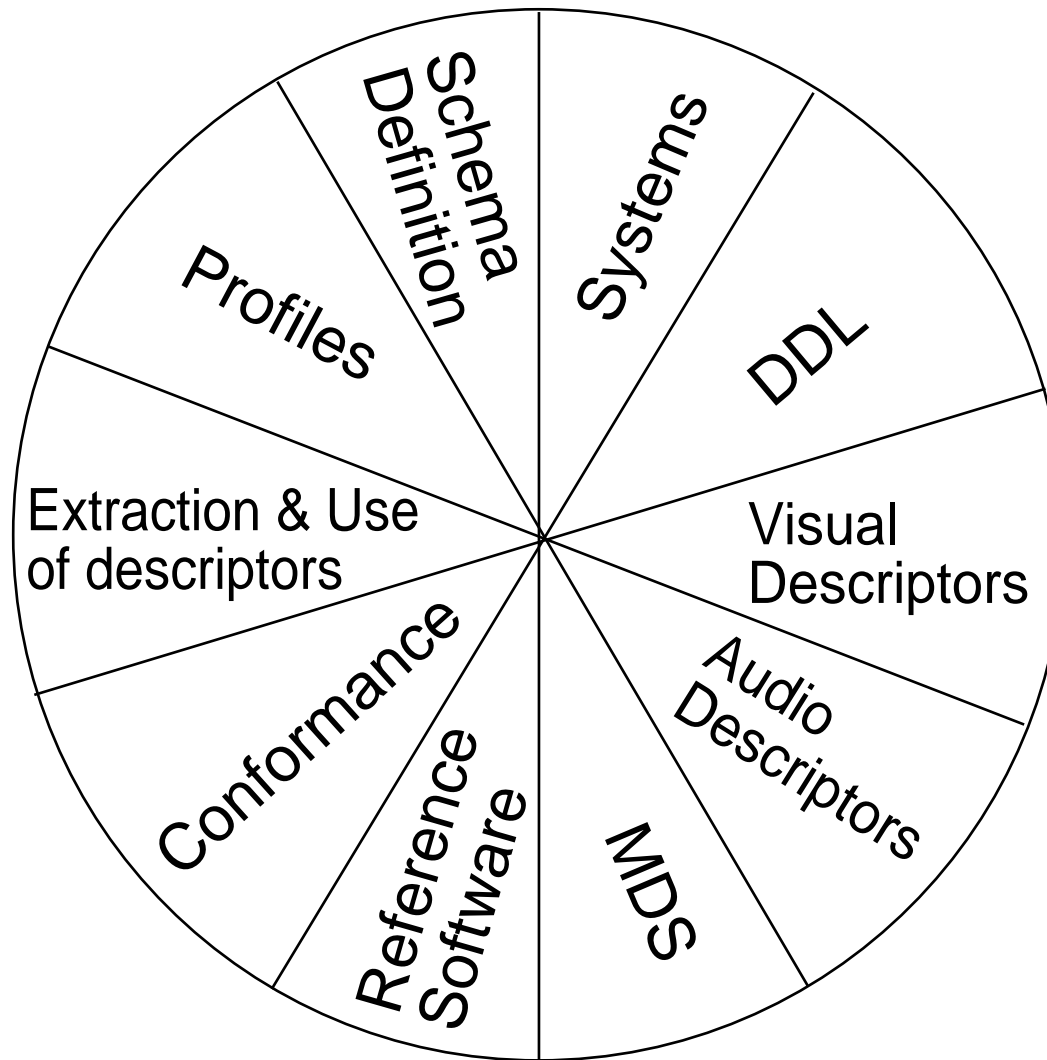
## Motivation:

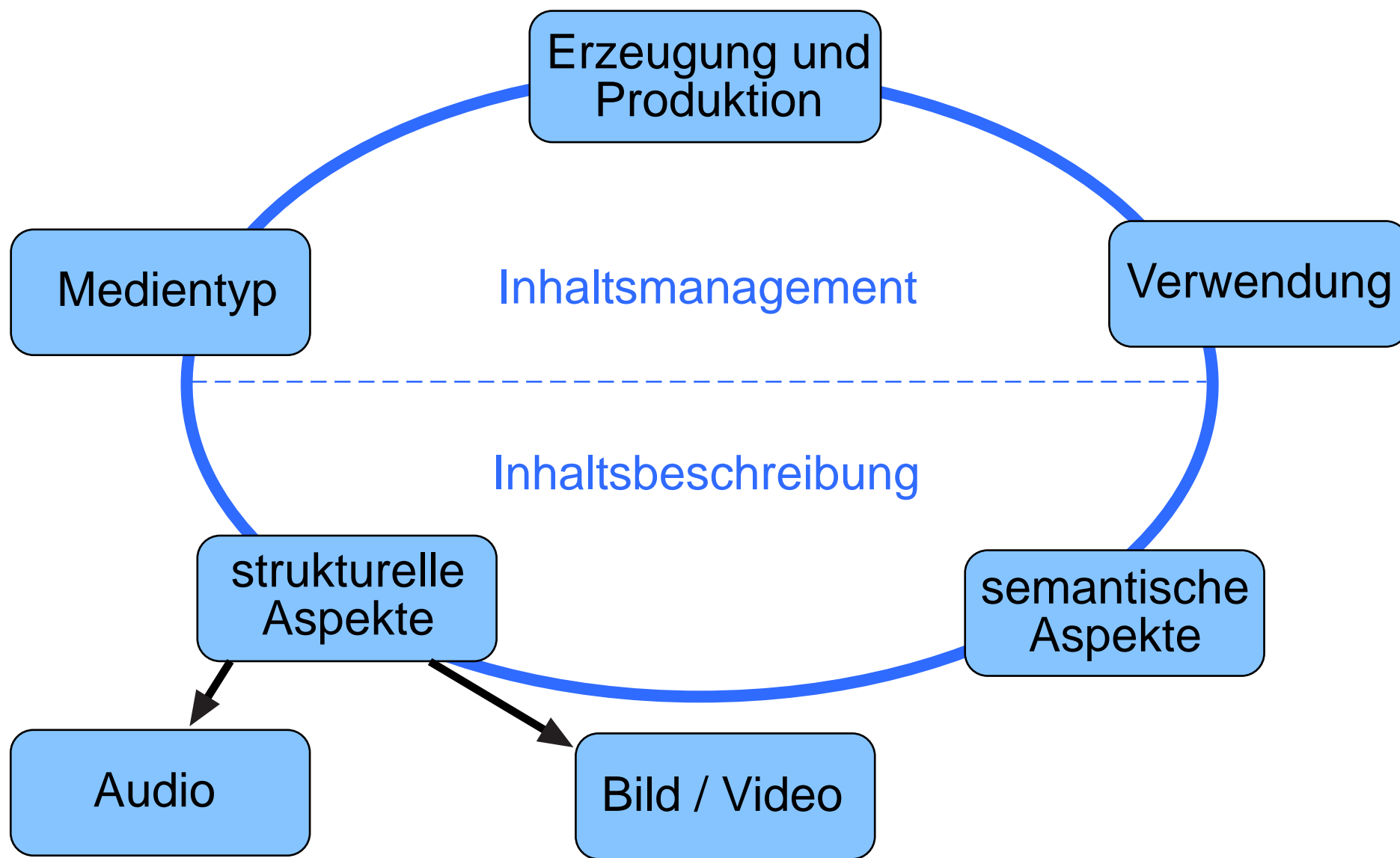


- stetig wachsender Informationsberg
- große Multimediabibliotheken
- effizient suchen
- in Multimediadaten enthaltene Metadaten effizient und intelligent nutzen
- neue Anwendungen

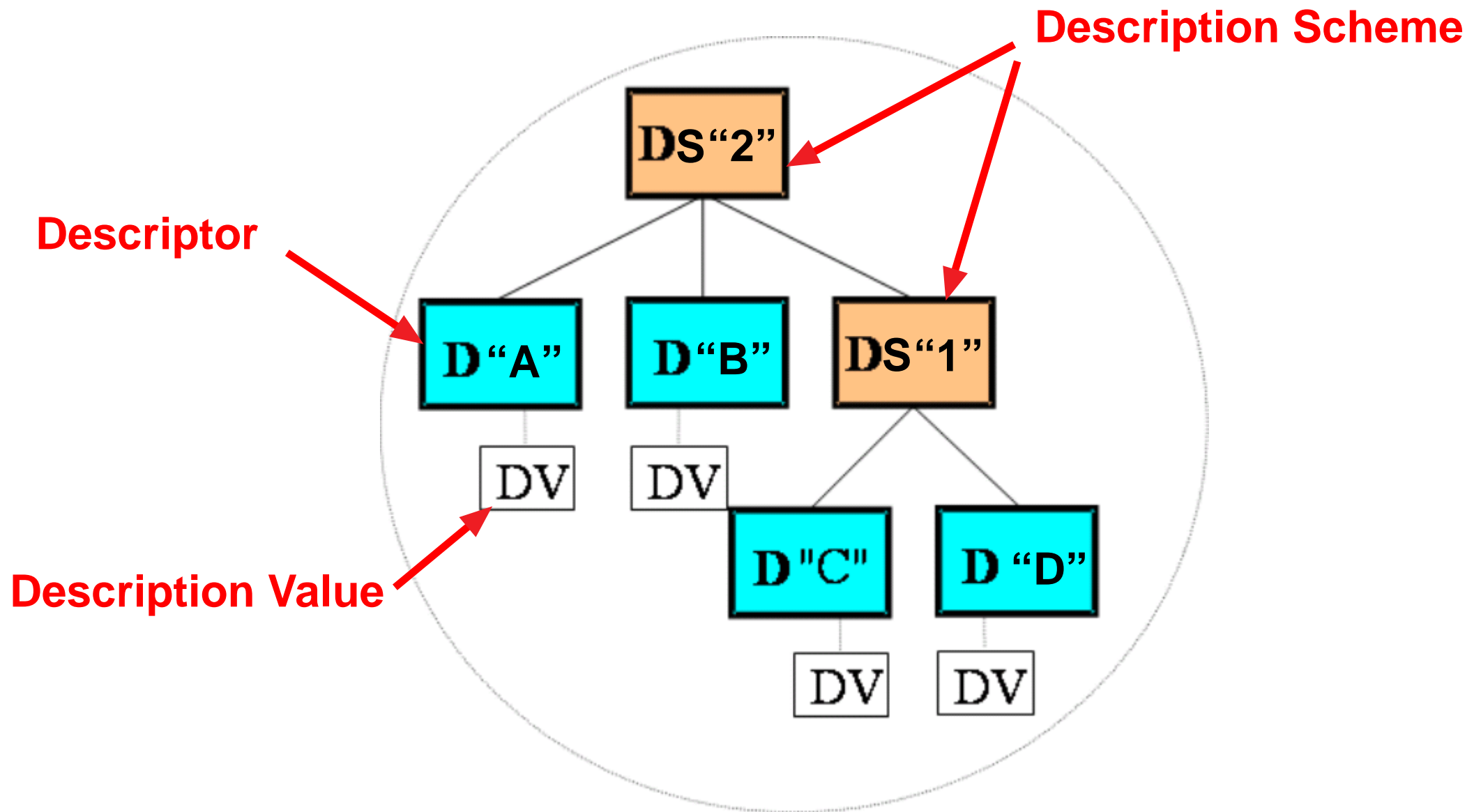
(Quelle: <http://www.ibm.com>, 2004)

## Der MPEG-7 Standard:





(Quelle: Ohm, 2005)



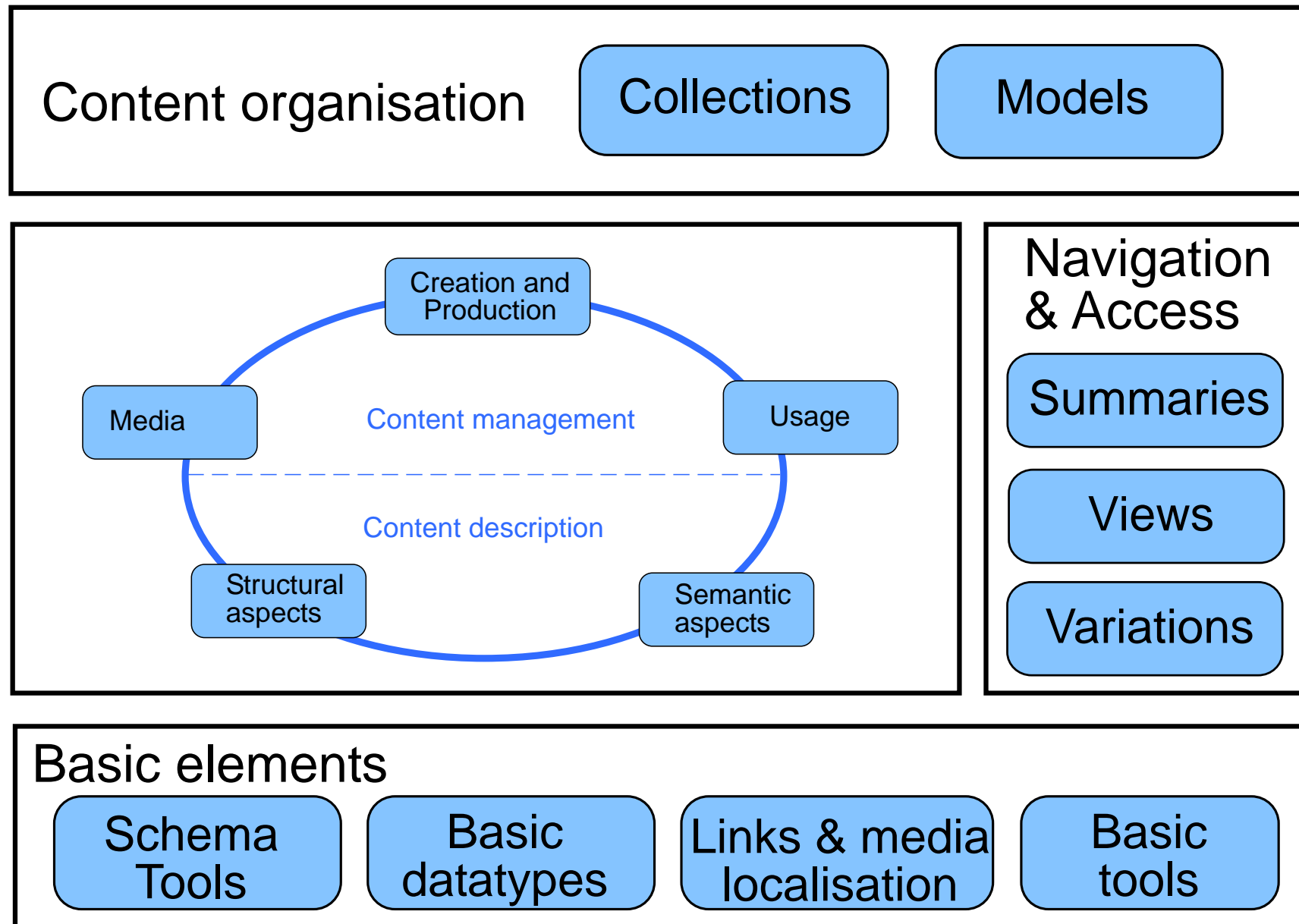
(Quelle: Höynck, 2004)



## Description Definition Language (DDL)

- basiert auf XML
- erweitert XML um Datentypen wie z.B. Arrays, Matrizen, Ländercodes, ...
- erlaubt die Erstellung eigener Descriptions und Description Schemes (DS)
- kann bestehende Descriptions und DS erweitern und modifizieren

## Media Description Schemes (MDS)

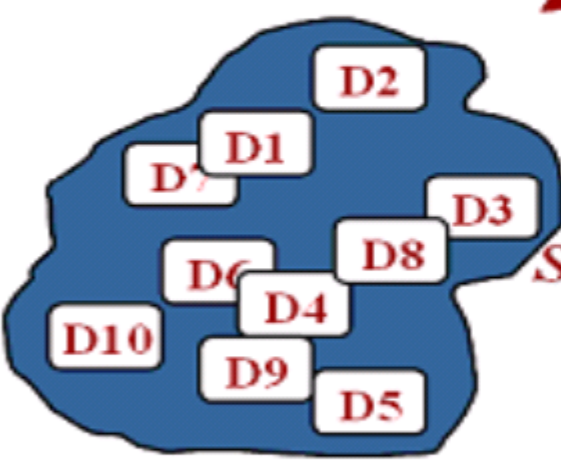


(Quelle: Martinez, 2004)

Description Definition Language

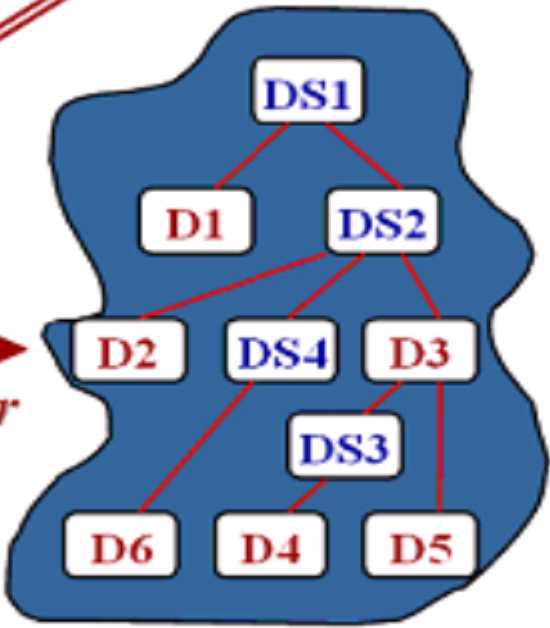
Definition

Tags



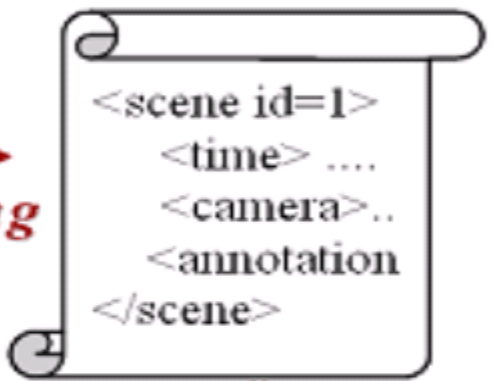
Deskriptoren:  
(Syntax & Semantik  
der Merkmals-Repäsentation)

Struktur

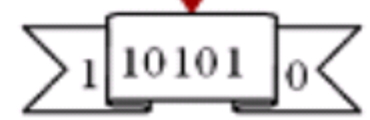


Description Schemes

Instantiierung



kodieren & verschicken



(Quelle: Martinez, 2004)

## Color Structure Descriptor

- unabhängig vom Format
- unabhängig von der Größe

kompakt



sehr unstrukturiert



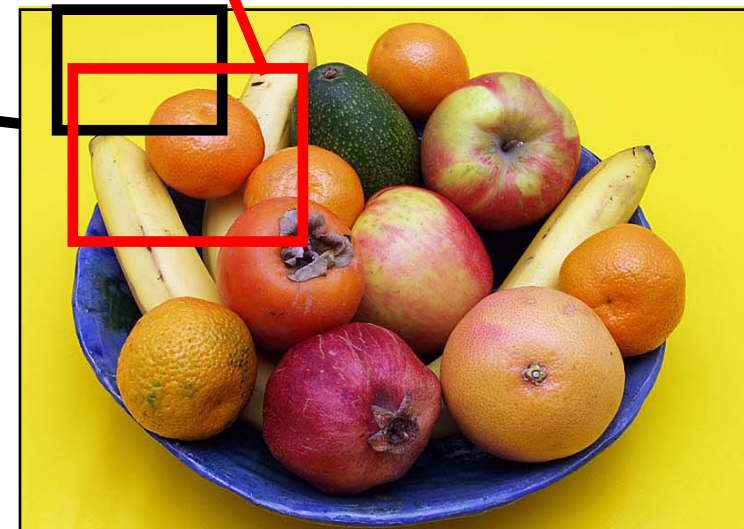
↑ schwarz

← gelb



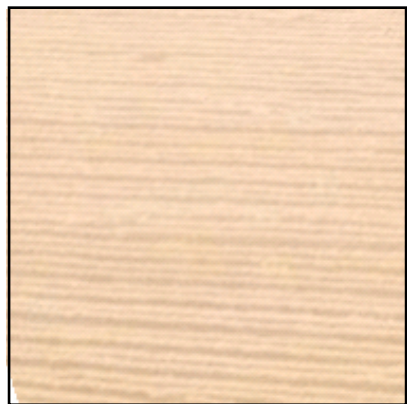
## Dominant Color Descriptor

- charakterisiert einen Ausschnitt anhand der dominanten Farbe

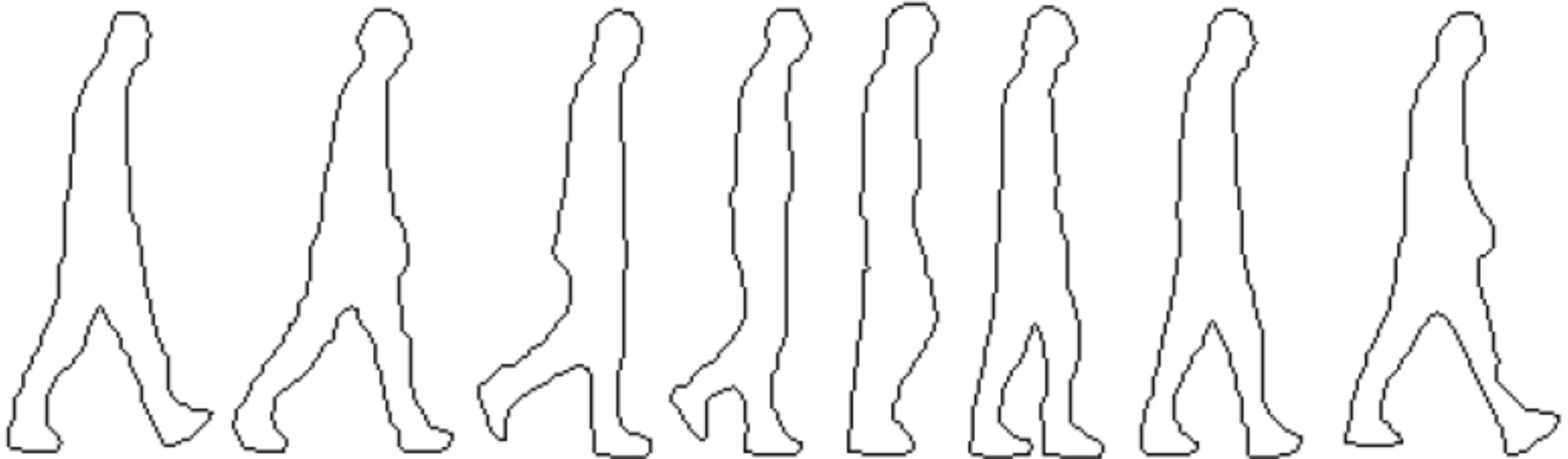


## Homogeneous Texture Descriptor

- Man kann ein Bild auch als Mosaik aus homogenen Texturen betrachten.
- Blickwinkel und Entfernung spielen keine Rolle.



Objekterkennung ist ein wichtiger Bestandteil der Videoanalyse.



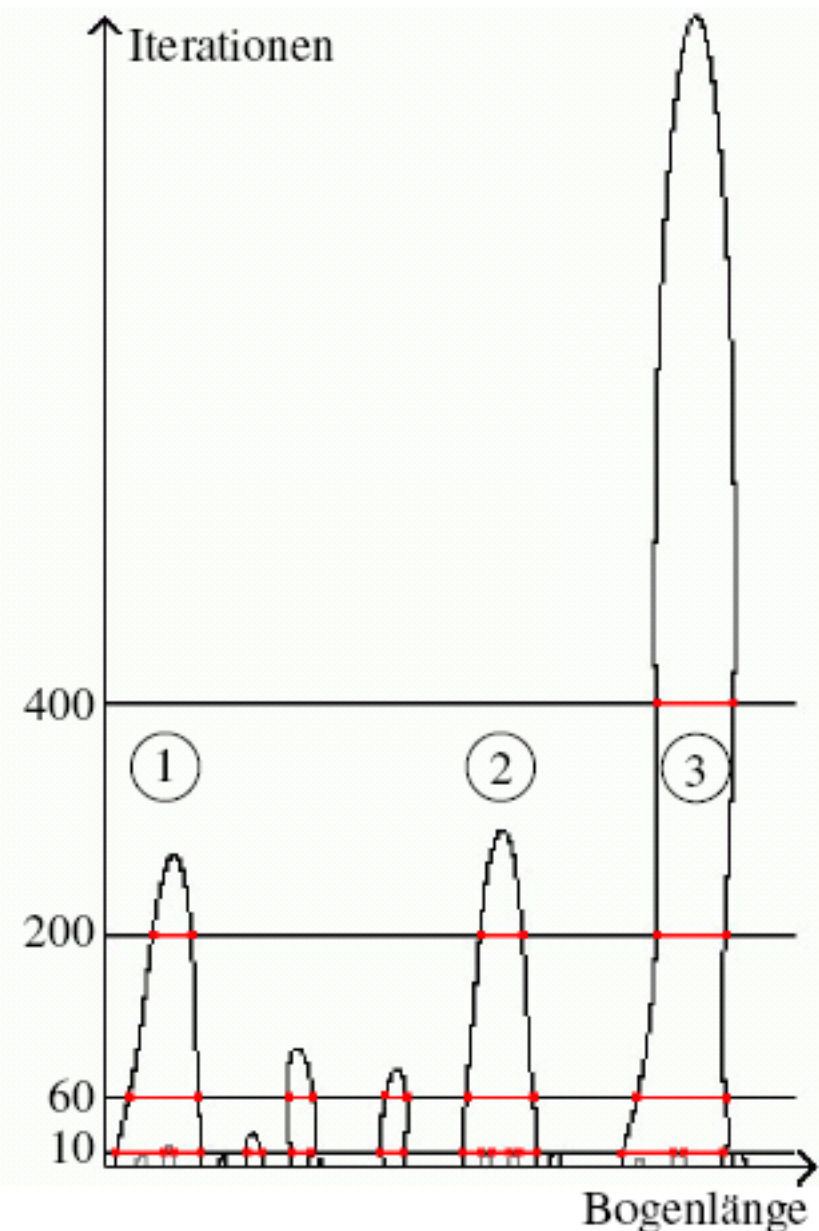
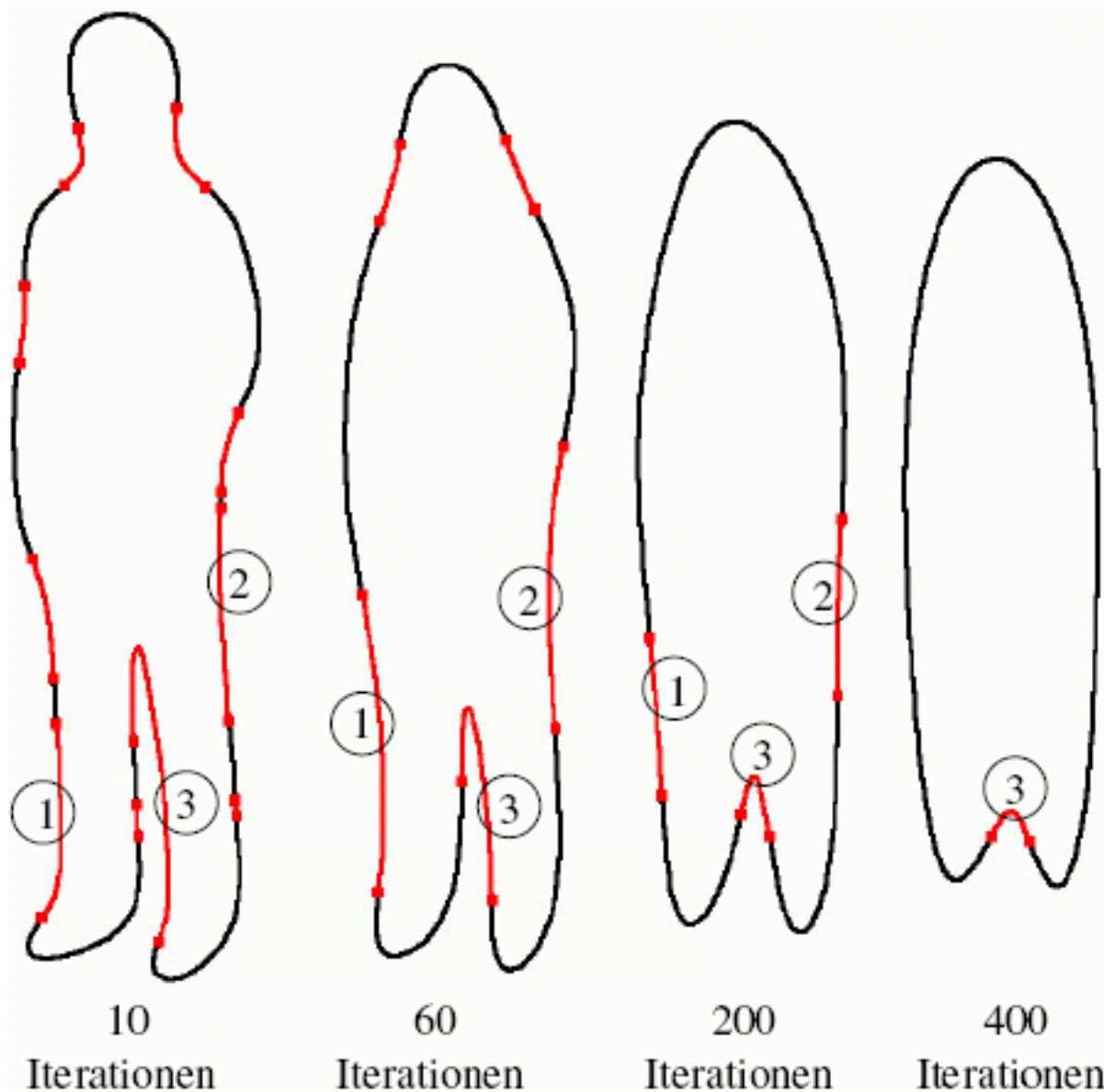
- Das menschliche Gehirn erkennt anhand der Kontur(änderung) dreidimensionale Objekte.



**MPEG-7 Contour Shape Descriptor**

(Quelle: Kopf, 2006)

# Contour Shape Descriptor



(Quelle: Kopf, 2006)



## Distanzmaß

Suchanwendungen benötigen ein **Maß**, mit dem Eigenschaften verglichen werden können

Da viele Deskriptoren als Vektoren(paare) dargestellt werden können, ist die Verwendung der **Euklidischen Distanz** sehr verbreitet.

### Euklidische Distanz:

$$d(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$

- Multimodale Suche: Skizze oder Sprache anstatt eines Begriffs
- Video-on-Demand: Untertitel, Sprache oder Musik nach persönlichen Präferenzen
- Videoüberwachung: Erkennung von gefährlichen Objekten
- Inhaltsadaption: Speech-to-Text, Picture-to-Text, ...



### **Vorteile von MPEG-7:**

- universell einsetzbar
- durch DDL erweiterbar
- Binärformat für Streaming

### **Nachteile von MPEG-7:**

- sehr komplex
- wird bisher nur vereinzelt eingesetzt

## Offene Fragen ?

