mi
PROFESSUR
MEDIENINFORMATIK

03 Meta Data

Lecture Media Retrieval

Maximilian Eibl, Medieninformatik, TU Chemnitz

Meta Data

Def.: Structured data about digital objects describing data, context, content, structure and life cycle.
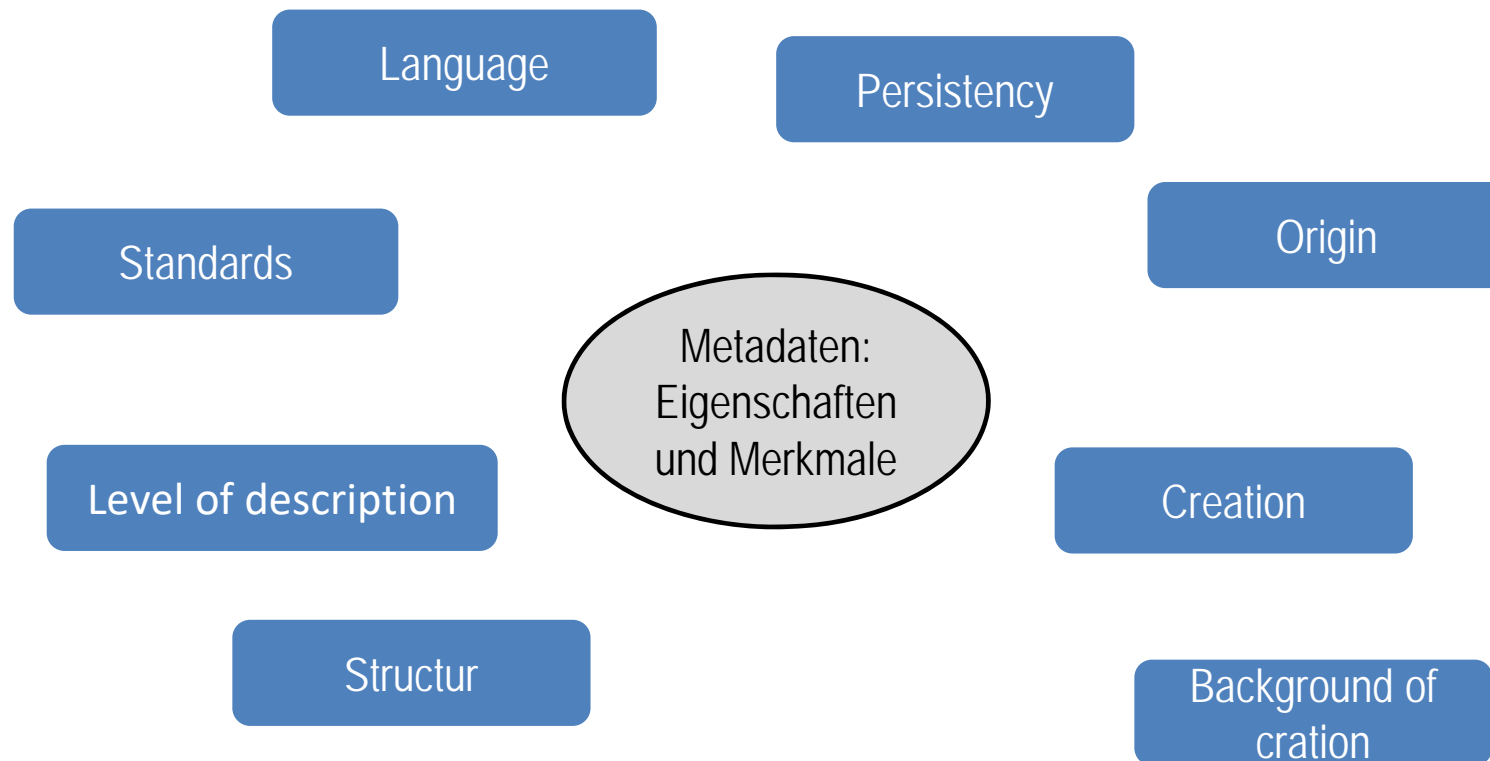
Meta Data

- Classic tools: Thesaurus, Classification
  - SWD
  - Dublin-Core
  - …
- New tools:
  - Ontology
  - Folksonomy
  - Heterogenity

Kinds of Meta Data

- Describing MD
  - Content MD
  - Filmographic MD
- Administrative MD
- Structural MD
- Archival MD
- Legal MD
- Technical MD
- Indentificational MD
- Process oriented MD

Metadaten: Eigenschaften

```
        Language              Persistency

   Standards                                    Origin

                      Metadaten:
                     Eigenschaften
   Level of description  und Merkmale           Creation

        Structur                          Background of
                                             cration
```

Meta Data in Video

text detection

text transcription

face / person detection

speaker

face / person identification

language

object detection
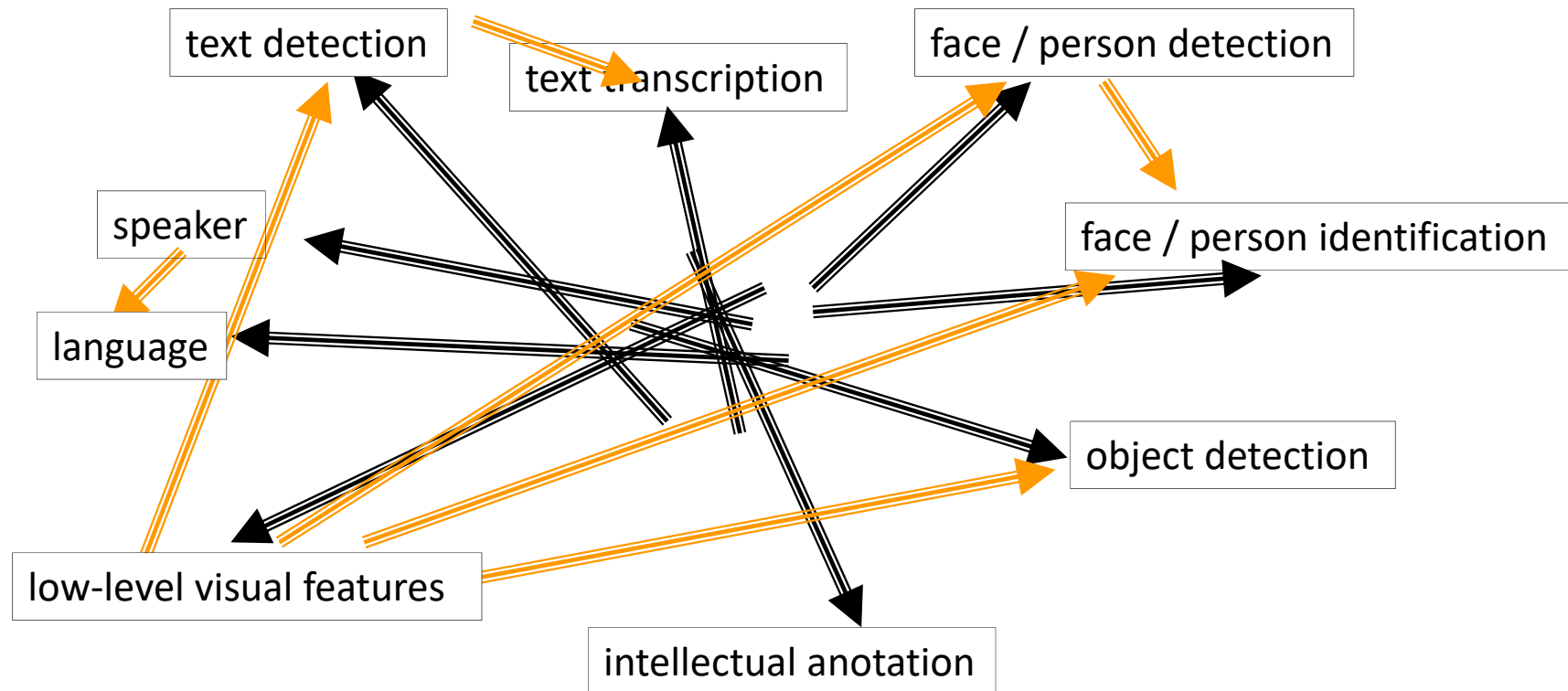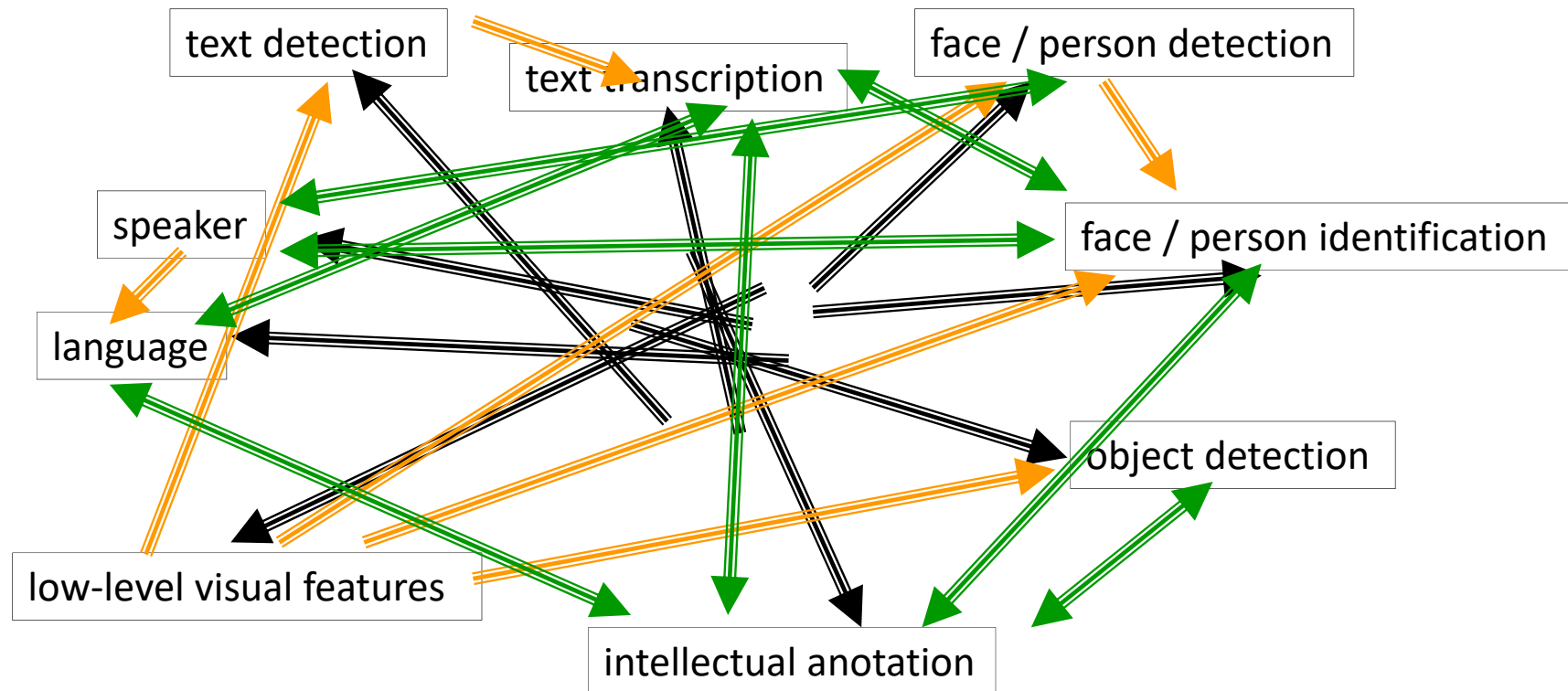
low-level visual features

intellectual anotation

Jens Nestler
Pressesprecher Feuerwehr Chemnitz

Meta Data: Extraction

Metadata: Conditions

text detection

text transcription

face / person detection

face / person identification

speaker

language

object detection

low-level visual features

intellectual anotation
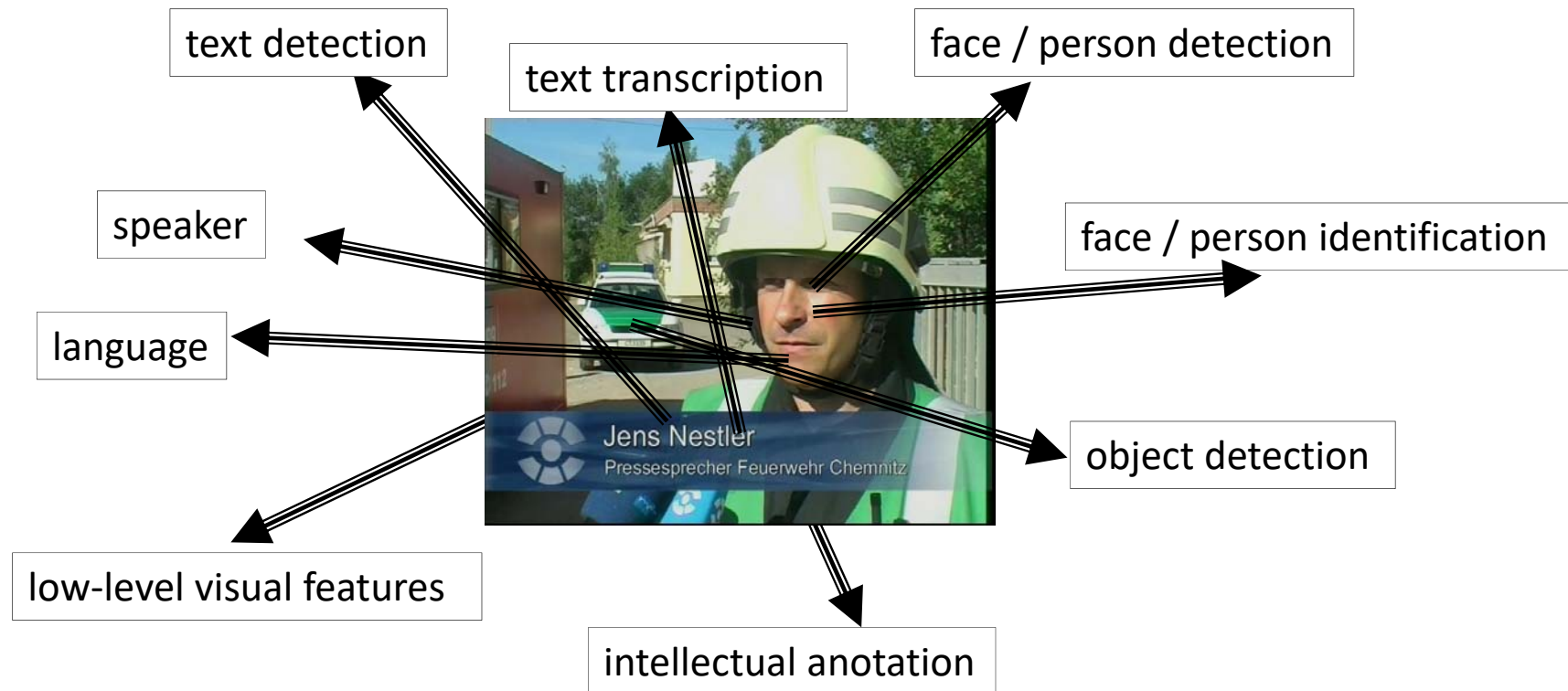
Metadaten: Validation

Metadaten

author - editor - rapporteur - chief of service - cutter - origin - identification of the organization - sponsor - scientific advisor - draftsman / graphic artist - working title - contribution - subtitles - entry title - file name - recording title - original title - production number - program identifier – broadcast title or main title - time slot - title - series or series subtitles - series or series titles - subtitles - number of rights - copyright notice - license period start - license period end - source - restriction on use expiration date - restriction of use explanation - limitation of use date of disposition - room layout - noise reduction - image proportion - duration of broadcasting - soundtrack - additional services - notes video content - recording / production location - accompanying material HTML - file - accompanying material manuscript - accompanying material press text - accompanying material broadcasting schedule - accompanying material other - accompanying material narration text - remarks - contents - color – film sound - episode - photo - GEO-descriptor - music - original soundtrack - production stage - production parts - realization level - editorial information - property descriptor - descriptive text - first broadcast date – rebroadcast date - technical acceptance date - timecode beginning - timecode end - soundtrack content - transfer date - time start of broadcast - expected broadcast date - expected rebroadcast date

Metadaten

Card Sorting

Videoreporter

Cutter

Redakteure

Kontakter

MPEG-7: Motivation & Definition (Day 2001: Introduction)

"How many times have you seen science fiction movies such as 2001: A Space Odyssey and thought, "Wow, we're so far away from having any of the fancy gadgets depicted in these movies!" In 2001, Hal, the talking computer intelligently navigates and retrieves information or runs complex operations instigated by spoken input. Or how about using an image-based query, say an image of the motorbike used by Arnold Schwartzenegger in the movie T2, to find images of similarly looking motorbikes. Dreams or reality? […] As more and more audiovisual information becomes available from many sources around the world, many people would like to use this information for various purpose"

"MPEG-7 is a multimedia content description standard, […], that addresses how humans expect to interact with computer systems, since it develops rich descriptions that reflect those expectations."

"MPEG-1, -2, and -4 make content available. MPEG-7 lets you to find the content you need."

MPEG-7 - Motivation

- Increasing relevance of non-textual media
- Classical annotation systems are primarily text oriented
- Increasing need for support in descriptions for multi media
- Different to MPEG-1, -2, -4: MPEG-7 does not code content but descriptions
- Aim: make media usable in different applcations

Status of MPEG-7

- Development 1996 – 2001
- Great bandwidth of abstraction levels and descriptions
    - „low-level"
    - „high level"
    - Flexible use
- Description schemes for different media
- Extendable schemes
- Support

Parts of MPEG-7

- MPEG-7 Systems

- MPEG-7 Description Definition Language

- MPEG-7 Visual

- MPEG-7 Audio

- MPEG-7 Multimedia Description Schemes MDS

- MPEG-7 Reference Software

- MPEG-7 Conformance

- MPEG-7 Extraction and Use of Descriptions

Typical Areas von MPEG-7-Descriptions

- Classial Meta Data

- Perceptional features

- Content usage and cross references

Basic Components of MPEG-7

- Description tools → descriptors *(Beschreibungswerkzeuge)*
- Description schemes *(Beschreibungsschemata)*
- Description definition language *(Beschreibungssprache)*
- System tools

Bestandteile einer MPEG-7-Beschreibung II: Zusammenspiel

Description Definition Language

Definition

Tags

structuring

instantiation

DS1

D1 D2 DS2 DS3

D3 D4 D5 DS4 D6

D7 D8 D9 D10

D4
D1
D3 D6
D2
D5
D7 D10
D9
D8

```
<scene id=1>
 <timeStart>…</timeStart>
 <timeEnd>…</timeEnd>
 <mainColor>
  <red>…</red>
  <green>…</green>
  <blue>…</blue>
 </mainColor>
 ….
```

Encoding
&
Delivery

Descriptors

Description Schemes
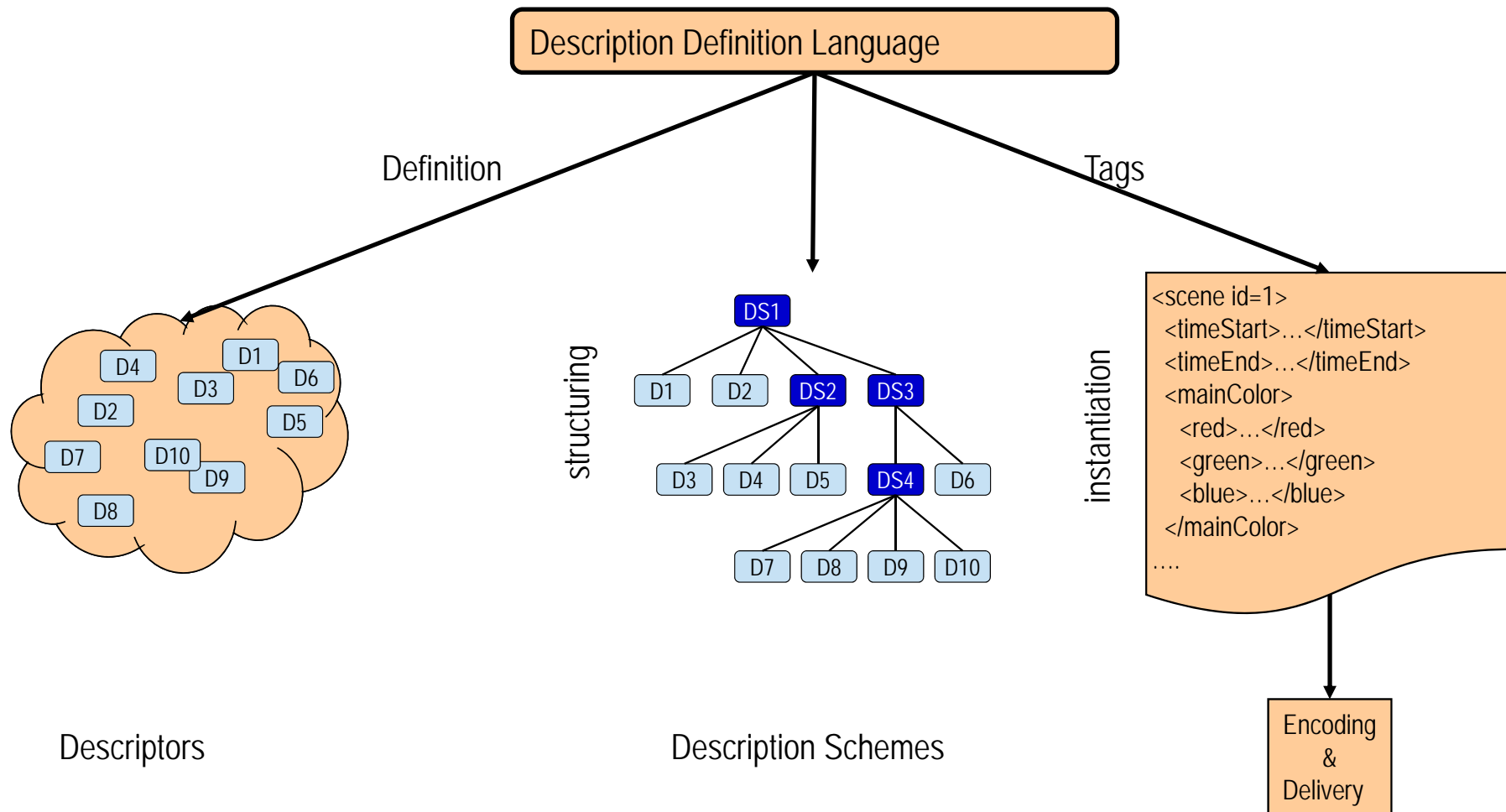
Anwendungskette von MPEG-7

Standardisierung

| Generierung von Beschreibungen | → | Beschreibung | → | Nutzung von Beschreibungen |

Merkmalsanalyse:
Inhaltsanalyse (D, DS)
Merkmalsextraktion (D, DS)
Annotationswerkzeuge (DS)
Authoring (DS)
automatisch (feature extraction)
semi-automatisch (werkzeug-unterstützt)
manuell
...

Bereich von MPEG-7:
Description Schemes (DSs)
Descriptors (Ds)
Language (DDL)

Anwendung:
Suchen, Filtern,
Klassifikation,
Manipulation,
Zusammenfassen,
Indexieren

Die MPEG-7 DDL - description definition language

- Syntax basiert auf XML und XML Schema

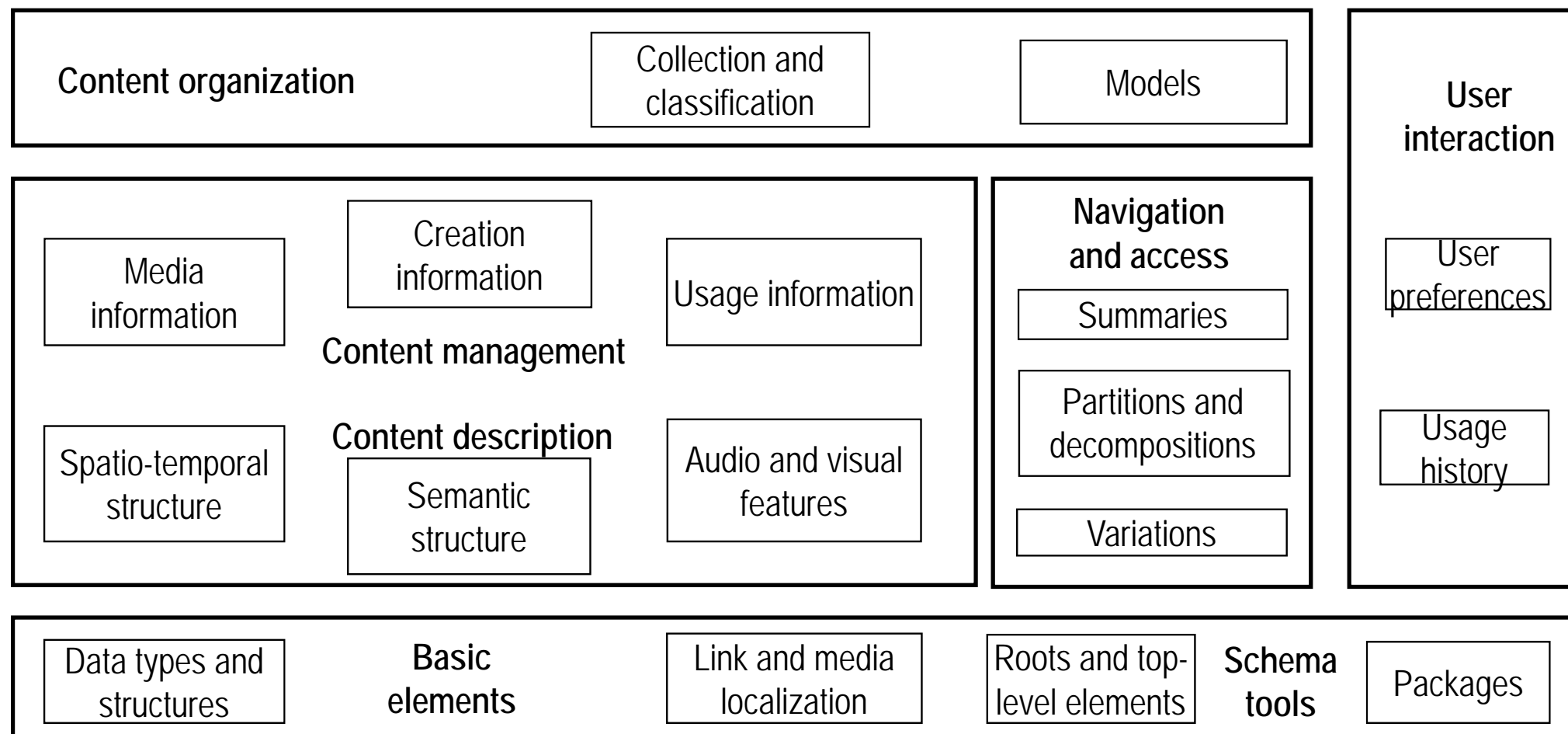- Teile:

  - XML Schema structural language components (d. h. Teil des W3C XML Schema-Standards)

  - XML Schema datatype language components (d. h. Teil des W3C XML Schema-Standards)

  - Erweiterungen von XML Schema, die MPEG-7-spezifisch sind, z. B. Einführung zusätzlicher Datentypen für die Beschreibung multimedialer Daten (Vektoren, Matrizen, getypte Referenzen

- Ergebnis: MPEG-7-Beschreibungen sind als XML repräsentierbar und (damit) interoperabel (ungeachtet der Möglichkeit der Binärkodierung)
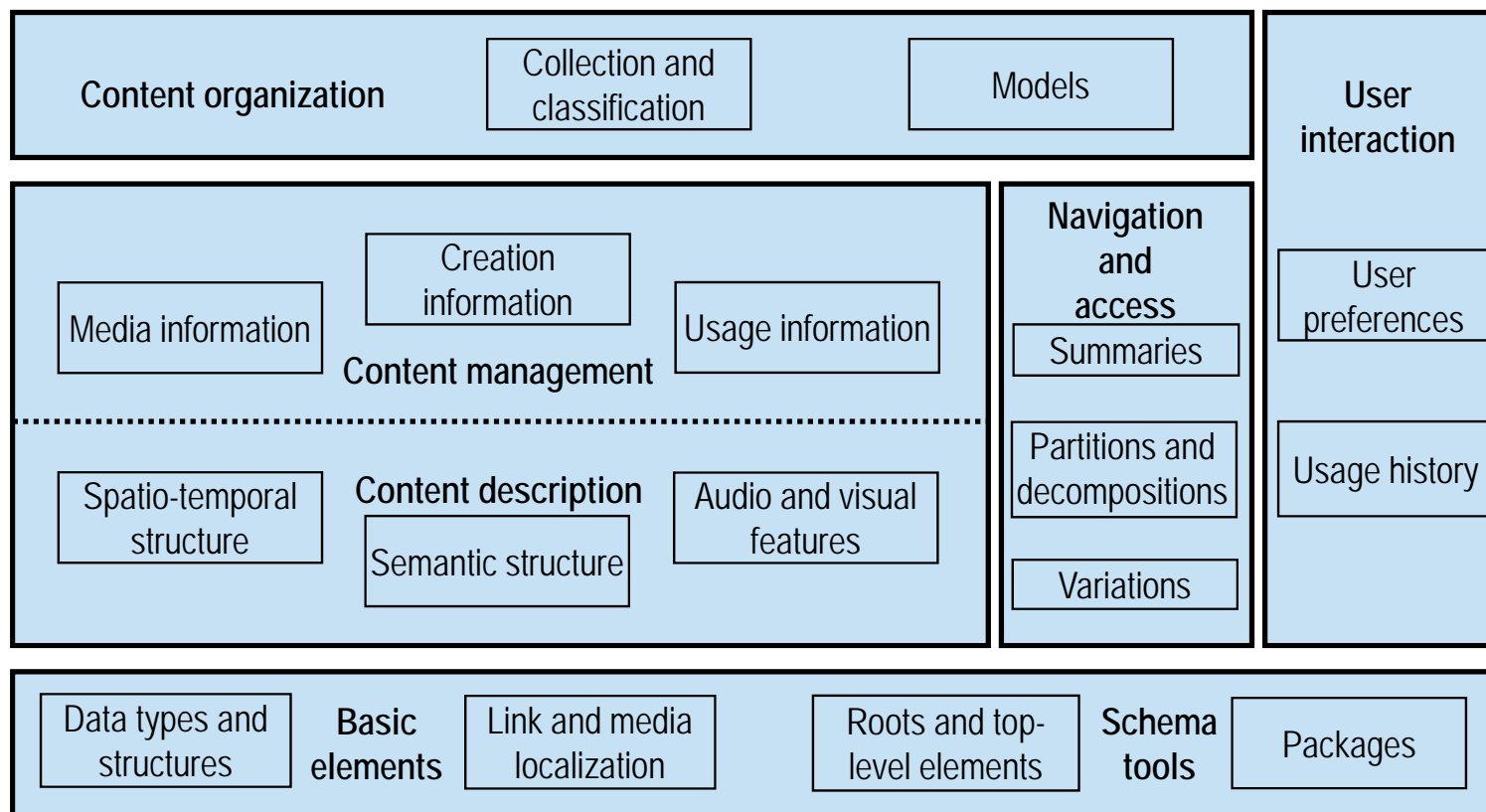
MPEG-7 Multimedia Description Schemes (DS)

mi

Content organization

Collection and classification	Models

User interaction

User preferences

Usage history

Content management

Media information	Creation information	Usage information

Content description

Spatio-temporal structure	Semantic structure	Audio and visual features

Navigation and access

Summaries

Partitions and decompositions

Variations

Schema tools

Data types and structures	**Basic elements**	Link and media localization	Roots and top-level elements	Packages

MPEG-7 Multimedia Description Schemes (DS)

Content organization				User interaction
	Collection and classification		Models	

					Navigation and access	User preferences
Media information	Creation information		Usage information		Summaries	
	Content management					
Spatio-temporal structure	Content description		Audio and visual features		Partitions and decompositions	Usage history
	Semantic structure				Variations	

Data types and structures	Basic elements	Link and media localization		Roots and top-level elements	Schema tools	Packages

Overview Visual Description Tools I
(ISO/IEC 15938-3)

Basic Structures

Descriptor Containers
- GridLayout
- TimeSeries
- MultipleView

BasicSupportingTools
- TemporalInterpolation
- Spatial2DcoordinateSystem

Visual Features

Color

ColorFeatureDescriptors
- DominantColor
- ScalableColor
- ColorLayout
- ColorStructure
- GofGopColor

ColorSupportingTools
- ColorSpace
- ColorQuantization

Texture
- HomogeneousTexture
- TextureBrowsing
- EdgeHistogram

Shape
- RegionShape
- ContourShape
- Shape3D

Motion
- CameraMotion
- MotionTrajectory
- ParametricMotion
- MotionActivity
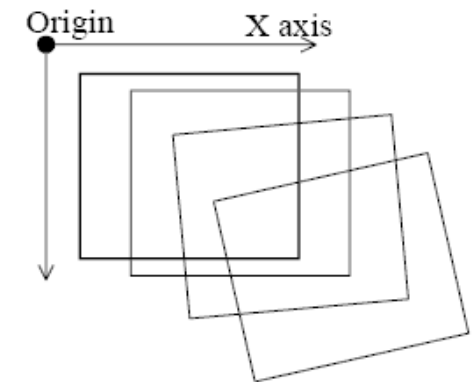
Localization
- RegionLocator
- SpatiaoTemporalLocator

Other
- FaceRecognition
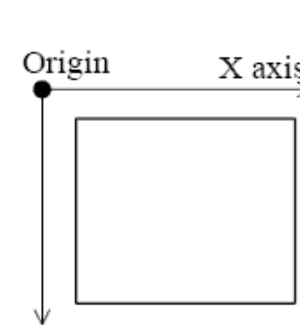
a) "Local" coordinates   b) "Integrated" coordinates

Overview Visual Description Tools II
(ISO/IEC 15938-3)

Basic Structures

Descriptor Containers
GridLayout
TimeSeries
MultipleView

BasicSupportingTools
TemporalInterpolation
Spatial2DcoordinateSystem

Visual Features

Color

ColorFeatureDescriptors
DominantColor
ScalableColor
ColorLayout
ColorStructure
GofGopColor

ColorSupportingTools
ColorSpace
ColorQuantization

Texture
HomogeneousTexture
TextureBrowsing
EdgeHistogram

Shape
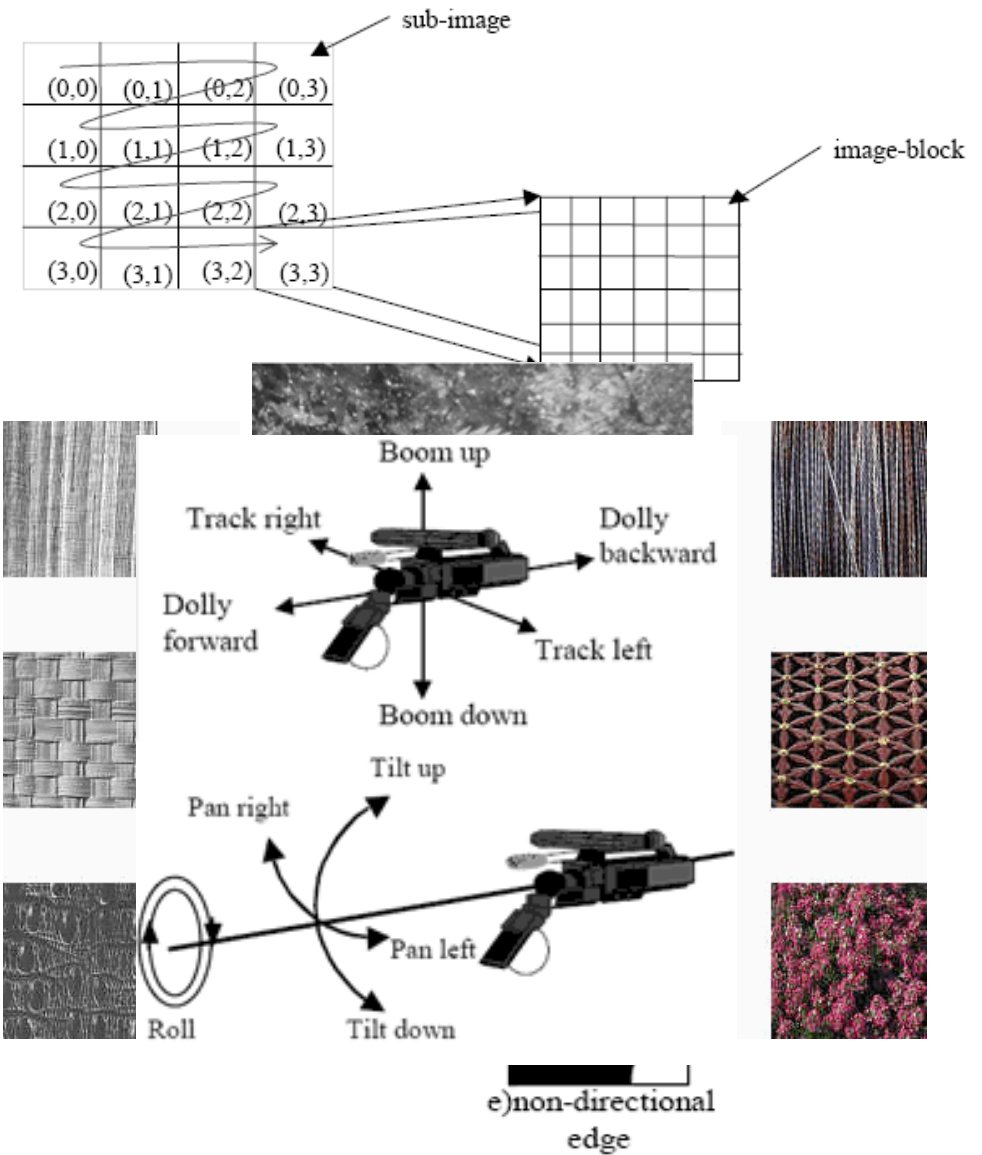RegionShape
ContourShape
Shape3D

Motion
CameraMotion
MotionTrajectory
ParametricMotion
MotionActivity

Localization
RegionLocator
SpatiaoTemporalLocator

Other
FaceRecognition

sub-image

image-block

(0,0)	(0,1)	(0,2)	(0,3)
(1,0)	(1,1)	(1,2)	(1,3)
(2,0)	(2,1)	(2,2)	(2,3)
(3,0)	(3,1)	(3,2)	(3,3)

Boom up

Track right

Dolly backward

Dolly forward

Track left

Boom down

Tilt up

Pan right

Pan left

Roll

Tilt down

e) non-directional edge

Audio Descriptors

- Describe audio content
- Application areas
  - robust audio matching
  - timbre matching
  - Melody search (Melody Contour Description Scheme bzw. Melody Sequence Description Scheme)
  - sound recognition / indexing
  - Spoken language description

Audio Framework (ISO/IEC 15938-4)

Silence D

Timbral Temporal
  LogAttackTime D
  TemporalCentroid D

Timbral Spectral
  HarmonicSpectralCentroid D
  HarmonicSpectralDeviation D
  HarmonicSpectralSpread D
  HarmonicSpectralVariation D
  SpectralCentroid D

Basic Spectral
  AudioSpectrumEnvelope D
  AudioSpectrumCentroid D
  AudioSpectrumSpread D
  AudioSpectrumFlatness D

Spectral Basis
  AudioSpectrumBasis D
  AudioSpectrumProjection D

Basic
  AudioWaveform D
  AudioPower D

Signal Parameters
  AudioHarmonicity D
  AudiFundamentalFrequency D

High Level Audio Description Tools

- Audio Signature Description Scheme
- Musical Intrument Timbre Description Tools
- Melody Description Tools
- General Sound Recognition and Indexing Description Tools
- Spoken Content Description Tools

Example for a MPEG-7-description I

```
<Mpeg7 xmlns="http://www.mpeg7.org/2001/MPEG-7_Schema" xml:lang="en"
   type="complete">
   <ContentManagement xsi:type="CreationInformationType">
      <Creation>
         <Title xml:lang="de">
            MPEG-7: The Multimedia Content Description Interface
         </Title>
         <Abstract>
            <FreeTextAnnotation>
               Overview on MPEG-7 in the couse of the lecture Media Retrieval
            </FreeTextAnnotation>
            <StructuredAnnotation>
               <What>
                  <Name>
                     Overview on MPEG-7
                  </Name>
               </What>
            </StructuredAnnotation>
         </Abstract>
...
```

Example for a MPEG-7-description II

```
...
        <Creator>
          <Role href="urn:mpeg:mpeg7:cs:RoleCS:AUTHOR">
            <Agent xsi:type="PersonType">
              <Name>
                <GivenName>Maximilian</GivenName>
                <FamilyName>Eibl</FamilyName>
              </Name>
            </Agent>
        </Creator>
        <CreationCoordinates>
          <CreationDate>
            <TimePoint>2017-12-17</TimePoint>
          </CreationDate>
        </CreationCoordinates>
      </Creation>
...
```

Example for a MPEG-7-description III

```
...
    <Classification>
      <Genre href="urn:mpeg:ContentCS:1">
        <Name xml:lang="de">
          Information
        </Name>
      </Genre>
      <Language type="main">de</Language>
      <Release date="20012-06-19"/>
    </Classification>
    <RelatedMaterial>
      <MediaLocator>
        <MediaUri>
          http://www.tu-chemnitz.de/cs/mi/
        </MediaUri>
      </MediaLocator>
    </RelatedMaterial>
  </ContentManagement>
</Mpeg7>
```

Visual Descriptors / Example of a Video Description

```
<Mpeg7>
    <Description xsi:type="ContentEntity">
        <MultimediaContent xsi:type="VideoType">
            <Video id="video_example">
                <MediaInformation>...</MediaInformation>
                <TemporalDecomposition gap="false" overlap="false">
                    <VideoSegment id="VS1">
                        <MediaTime>
                            <MediaTimePoint>T00:00:00</MediaTimePoint>
                            <MediaDuration>PT2M</MediaDuration>
                        </MediaTime>
                        <VisualDescriptor  xsi:type="GoFGoPColorType" aggregation="average">
                            <ScalableColor  numOfCoef="8" numOfBitplanesDicarded="0">
                                <Coeff>1 2 3 4 5 6 7 8</Coeff>
                            </ScalableColor>
                        </VisualDescriptor>
                    </VideoSegment>
                </TemporalDecompostion>
            </Video>
        </MultimediaContent>
    </Description>
</Mpeg7>
```
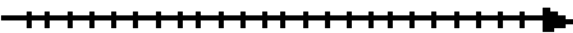
(see Martínez 2002B: 85, Fig. 3)

Content Description

- Segment DS
    - Multimedia Segment DS
    - AudioVisual Region DS
    - AudioVisual Segment DS
    - Audio Segment DS
    - Still Region DS
    - Still Region 3D DS
    - Moving Region DS
    - Video Segment DS
    - Ink Segment DS
- Graph DS

Segment DS, Examples for Segments

| Temporal segment (VideoSegment, AudioSegment)	Spatial segment (StillRegion)
a) Time — Segment composed of one connected component	b) Segment composed of one connected component
c) Time — Segment composed of three connected components	d) Segment composed of three connected components

Special Properties of Segment Description

Property	Video Segment	Still Region	Moving Region	Audio-segment
Time	X		X	X
Shape		X	X	
Color	X	X	X	
Texture		X		
Movement	X		X	
Camera Movement	X			
Audio			X	X

Example of a MPEG-7-Description of Still Regions)
(Martínez 2002B: 91, Fig. 6)
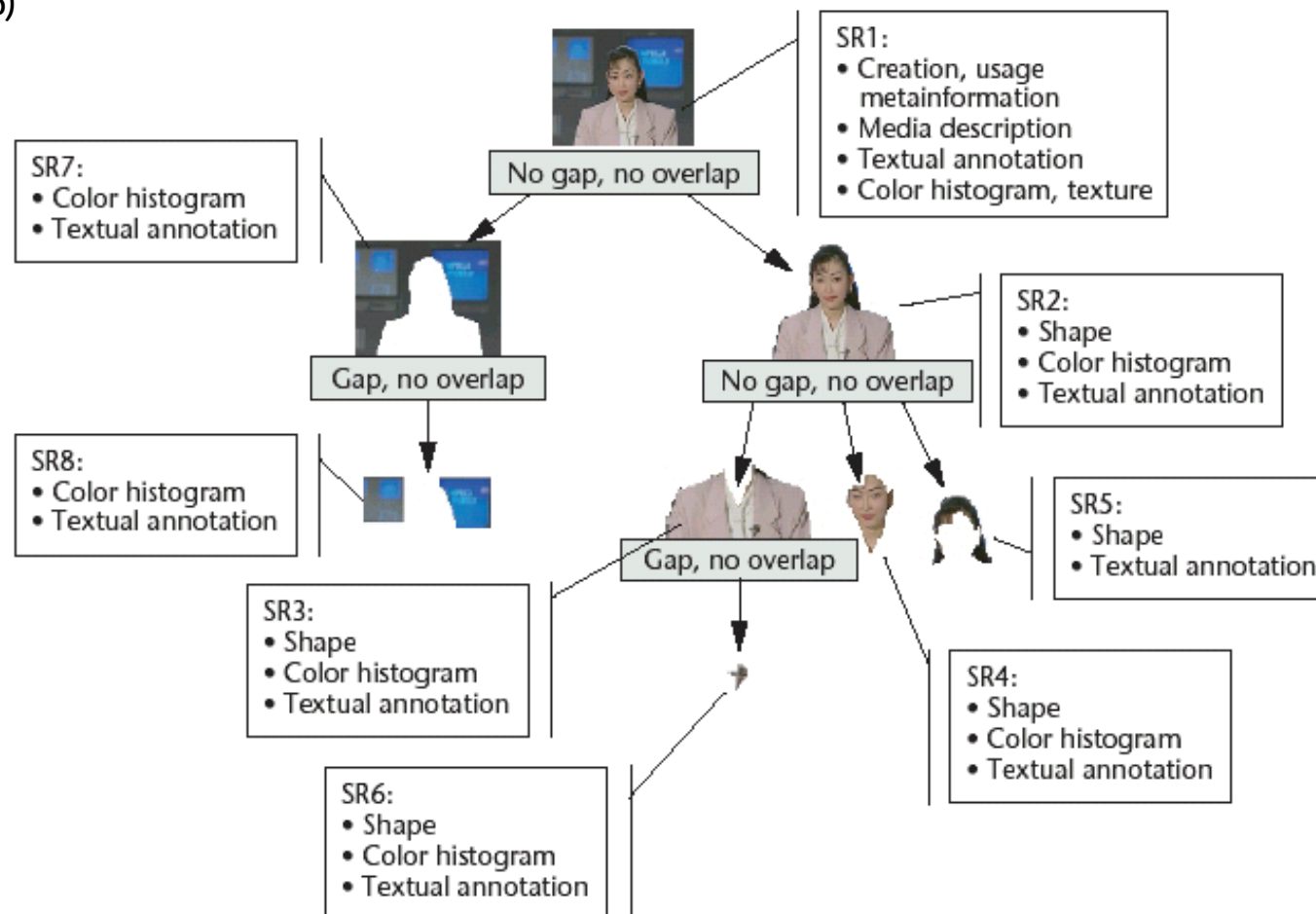
Beschreibung typischer Anwendungsszenarien im Rahmen des Standards

- Applikationstypen als Teil des sog. MPEG-7 eXperimental Model, das während der Entwicklung des Standards als gemeinsame Entwicklungsbasis diente

- Anwendungsklassen (Haupteinteilung: feature extraction / client applications (s. o.))
  - extraction application: Zu gegebenen Multimediadaten werden Beschreibungselemente nach MPEG-7 generiert
  - search & retrieval: Mit Hilfe einer Beschreibungsdatenbank können zu einer Anfrage geeignete Media-assets selektriert werden
  - media transcoding: Mit Hilfe vorliegender Beschreibungen und ggf. Anfragen werden Medien umkodiert
  - description filtering: Filtern und selektieren von Beschreibungen durch Anfragen

- Weitere Einteilungsmöglichkeiten
  - nach Inhaltsart
  - nach push- oder pull-Anwendungen

MPEG-7: Beispiele für Anwendungen nach „Branchen"

- Digitale Bibliotheken, Medienarchive
- Media Authoring (inkl. Personalisierung)
- Kulturdienste (Museen etc.)
- Multimedia-Verzeichnisdienste
- Streaming Media („Rundfunk")
- E-Commerce (personalisierte Anwendungen)
- Überwachung (Verkehr, Logistik)
- Biometrik
- Entertainment
- Soziale Dienste
- .....