

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 Feature Extraction for Images and Video

8.3 Feature Extraction for Audio

8.4 Selected Metadata Standards (including MPEG-7)

8.5 Semantic Web Technologies for Multimedia

Literature:

Rosenblatt/Trippe/Mooney, Digital Rights Management, Chapter 6

Troncy/Huet/Schenk, Multimedia Semantics - Metadata, Analysis and Interaction, Wiley 2011

Unlabelled Stuff

- The Unlabeled Video Tape Problem
 - Even worse with digital media:
Various formats, variants
- Digital media production:
 - Labeling of parts to be composed
 - » Date, time, format, ...
 - Representing the composition
- Digital media on the Internet
 - Identifying digital media
 - » Title, author, genre, ...
 - Searching for specific media, e.g. audio, video content
 - Fine-grained search within media
 - » e.g. person search within video content
 - Bringing together related media (e.g. text news and photos)
 - » (Automated) syndication



Content, Essence, Metadata

- Content
 - consists of *essence* data and *metadata*
- Essence
 - parts of content that directly represent program material such as audio, video, graphic, still-image, text, or sensor-data
- Metadata
 - parts of content that contain data used
 - » to *describe* essence or
 - » to provide information on its *use*
 - metadata objects sometimes called “mobs”
- Metadata may be
 - stored separately from the essence data
 - combined with the essence data (“embedded metadata”)

Source: AAF Developer Overview

Types of Multimedia Metadata

- Technical Metadata:
 - Form (data format, representation parameters like resolution, color depth...)
 - For live captured material: Time, date, location of original occurrence
 - Technical parameters of capture (e.g. aperture, exposure etc. for images)
- Content Description Metadata:
 - High level, structured:
 - » Title, author, composer, artist, cast,
 - High level, unstructured:
 - » Summary, textual description, thumbnail, ...
 - Low level:
 - » Objects and time positions
 - » Audio and video features: Key, mood, tempo ...
- Additional information:
 - Digital rights, classification, context, further links, ...

Types of Origin for Metadata

- Automatic creation or derivation:
 - All technical metadata
 - Extracted data features
= mainly low level metadata (e.g. average brightness, musical tempo)
- Retrieval from external databases:
 - High-level metadata
 - Retrieval may be based on identifier or analysis of media content
 - Example: GraceNote database for music
- Manual addition
 - Archival, indexing, annotation, ...

Metadata Problems

- Creation metadata
 - During the creation of media essence, metadata is created but often ignored
 - Example: EXIF data in JPEG
- Manually added metadata
 - Users notoriously ignore the administration of metadata
- Metadata incompatibility
 - Metadata exists in various formats specific for media types, applications, product vendors, ...
 - Exchange of metadata is difficult
- Broad range of metadata
 - Metadata exists on various levels, covering all is expensive
- Metadata economy
 - How much of the metadata will be used?
 - When to create metadata?

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 Feature Extraction for Images and Video

8.3 Feature Extraction for Audio

8.4 Selected Metadata Standards (including MPEG-7)

8.5 Semantic Web Technologies for Multimedia

Literature:

Troncy/Huet/Schenk, Multimedia Semantics -

Metadata, Analysis and Interaction, Wiley 2011

Chapter 4 of this book: R. Benmokhtar, B. Huet, G. Richard, S. Essid:

Feature Extraction for Multimedia Analysis

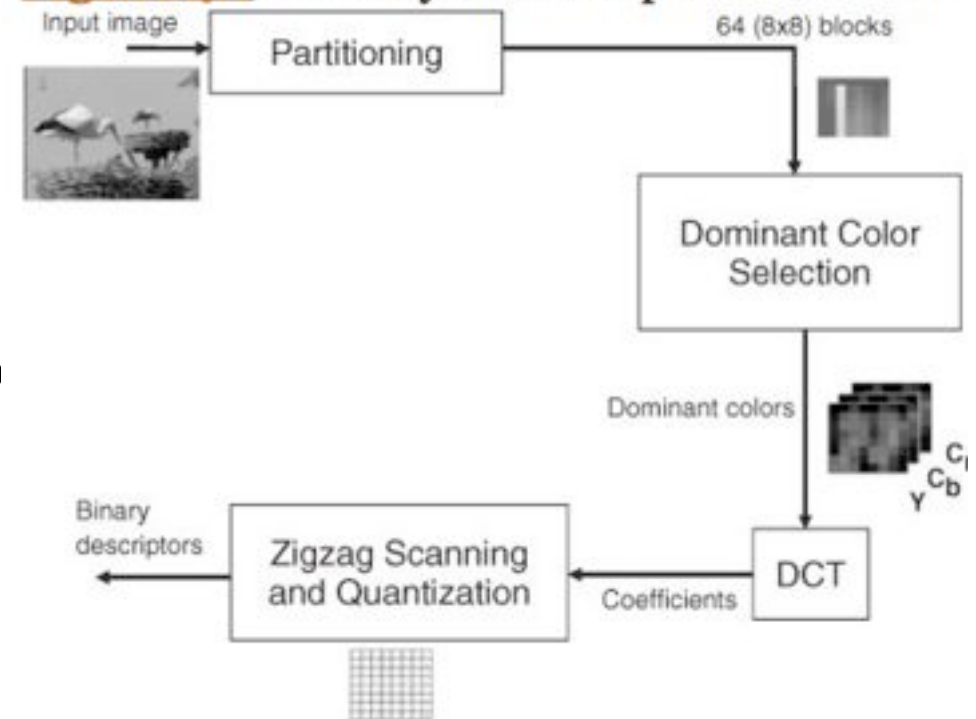
Features of Multimedia Signals

- Feature: Condensed information from multimedia signal
 - Image, video, audio, text
 - Purpose: Description of content with *high variance* and *high discriminance*
 - Identification of similar content with respect to certain aspects
- Low-level features
 - Derived from signal processing algorithms
 - Selection of low-level features defined in MPEG-7 standard (see below)
- Multi-feature and multi-modal fusion:
 - Combined use of multiple characteristics
- Case study:
 - Automatic identification of violence in video material
 - Specifically scenes containing *punching*

Low-Level Visual Descriptors (Examples) (1)

- Color Descriptors
 - Dominant Color Descriptor (DCD, Cieplinski 2000):
 - » Small set (four to six) of dominant colors, histogram
 - » Taking into account variance, spatial coherence (3x3 masking window)
 - Color Layout Descriptor (CLD, Kasutani and Yamada 2001):
 - » Representing spatial distribution of colors
 - Scalable Color Descriptor (MPEG-7 2001)
 - » HSV color space with fixed quantisation
 - » Using Haar transform to create small histogram

Figure 4.1 Color layout descriptor extraction.



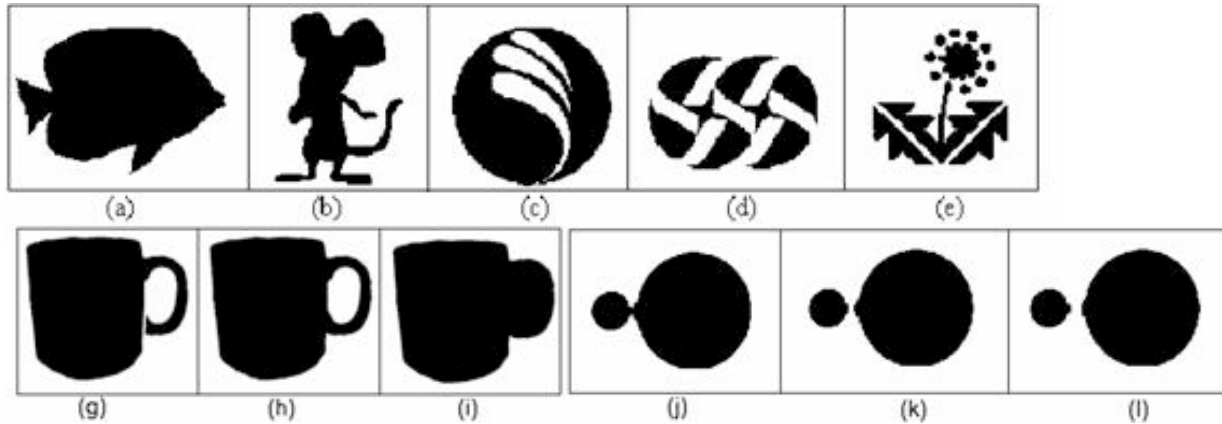
Source: Benmokhtar et al.

Low-Level Visual Descriptors (Examples) (2)

- Texture Descriptors
 - Homogeneous Texture Descripture (HTD, Manjunath et al. 2002)
 - » Statistical analysis on local spatial frequencies
 - » 30 “channels” from 6 frequency times and 5 orientation channels
 - » Energy and energy deviation for each of the channels
 - Edge Histogram Descriptor (EHD, Park et al. 2000)
 - » For 16 subimages (4x4), detects five types of edges
 - » 80-dimensional vector
- Shape Descriptors
 - Region-based Shape Descriptor (R-SD, Manjunath et al. 2002)
 - » Using *Angular-Radial Transformation* (see next slides)
 - Contour-based Shape Descriptor (C-SD, Zhang and Lu 2003)
 - » Using *Curvature Scale Space* (CSS) representation (see next slides)

Examples for Shape Descriptors

Region shapes:

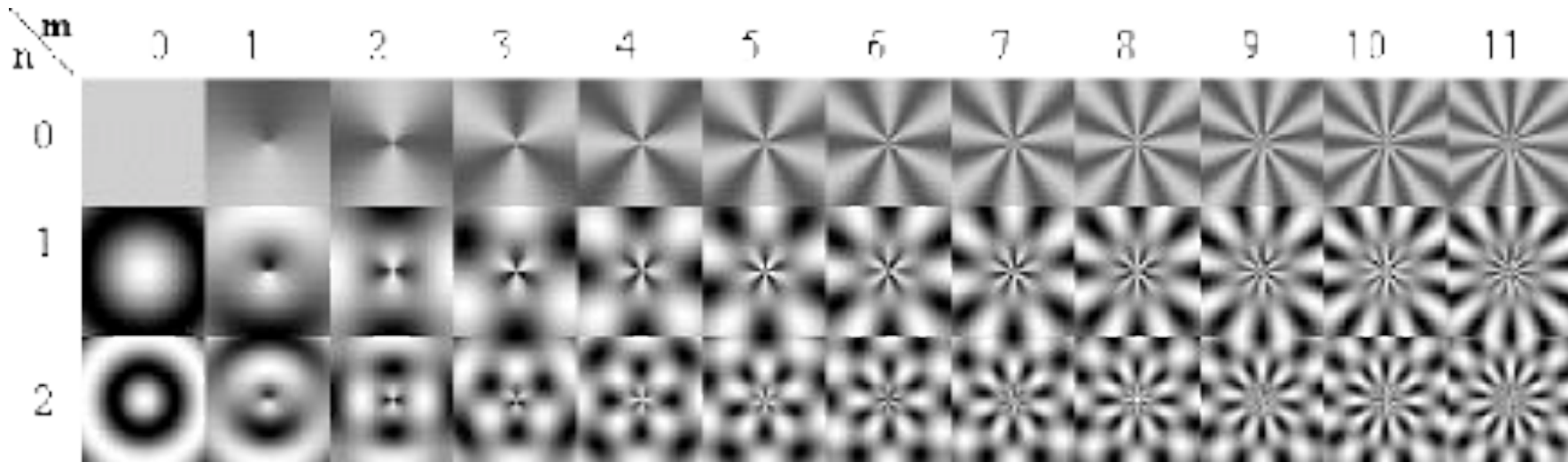


Contour shapes:



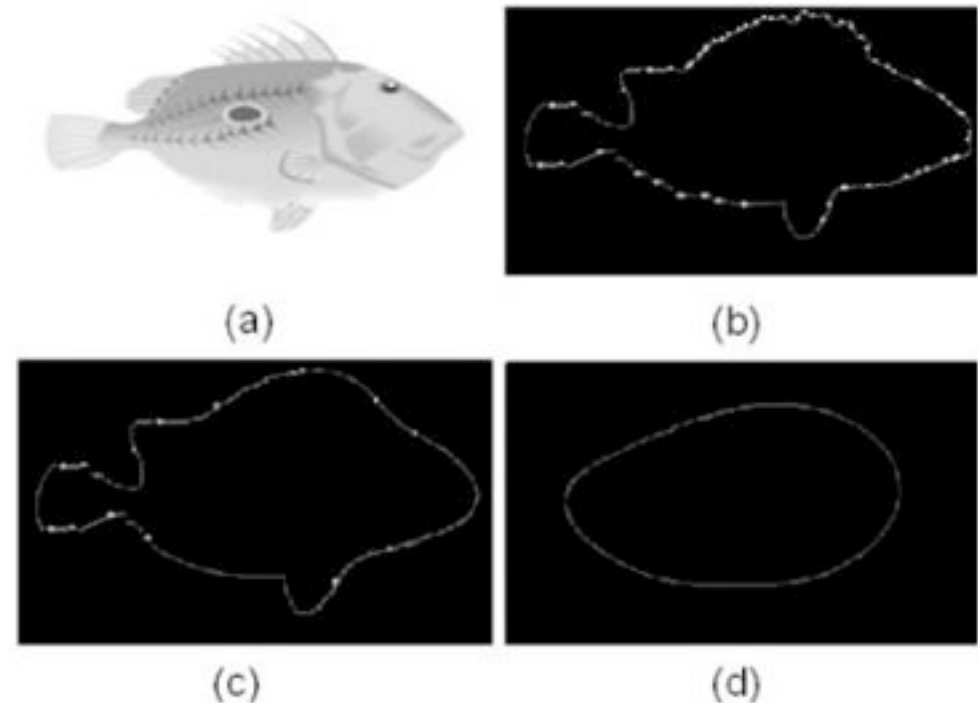
Angular Radial Transformation (ART)

- Convert image information into angular and radial parts
- Represent image as coefficients of basis functions
- First 36 basis functions:



Curvature-Scale Space Computation

- Curvature is a local measure of how fast a curvature is turning
 - Contour is sampled with increasing precision and smoothed stepwise until a convex shape is obtained
 - During the step-wise smoothing, specific points are saved into the descriptor:
 - » Points separating convex and concave parts of contour
 - » Peaks of contour map
 - Additional global values:
 - » Eccentricity
 - » Circularity
 - » Number of CSS peaks



Source: Benmokhtar et al.

Motion Descriptors for Video (Examples)

- Camera Motion Descriptor (CMD, Manjunath et al. 2002)
 - Camera operations:
panning, tracking, tilting, booming, zooming, dollying, rolling
- Motion Activity Descriptor (MAD, Sun et al. 2002)
 - Statistical analysis of motion vectors from differential frames
 - *Intensity* of motion (motion vector magnitude)
 - *Direction* of motion (dominant direction of vectors)
 - *Spatial distribution* of activity
 - *Temporal distribution* of activity
- More complex descriptors describing long-term trajectories of objects

Case Study: Violence Detection in Video

- Only few descriptors are suitable
- Only combination of multiple descriptors can achieve results
- Possible solution:
 - Detection of human bodies using contour and region shapes
 - Detection of rapid and significant movements of persons and objects
 - Detection of bleeding using color descriptors
 - Combination with audio analysis (punch sound, screams)
- Combination of automatic pre-filtering of material with human analysis
- Important information extremely difficult to obtain through analysis
 - e.g. presence of (possibly concealed) weapons

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 Feature Extraction for Images and Video

8.3 Feature Extraction for Audio

8.4 Selected Metadata Standards (including MPEG-7)

8.5 Semantic Web Technologies for Multimedia

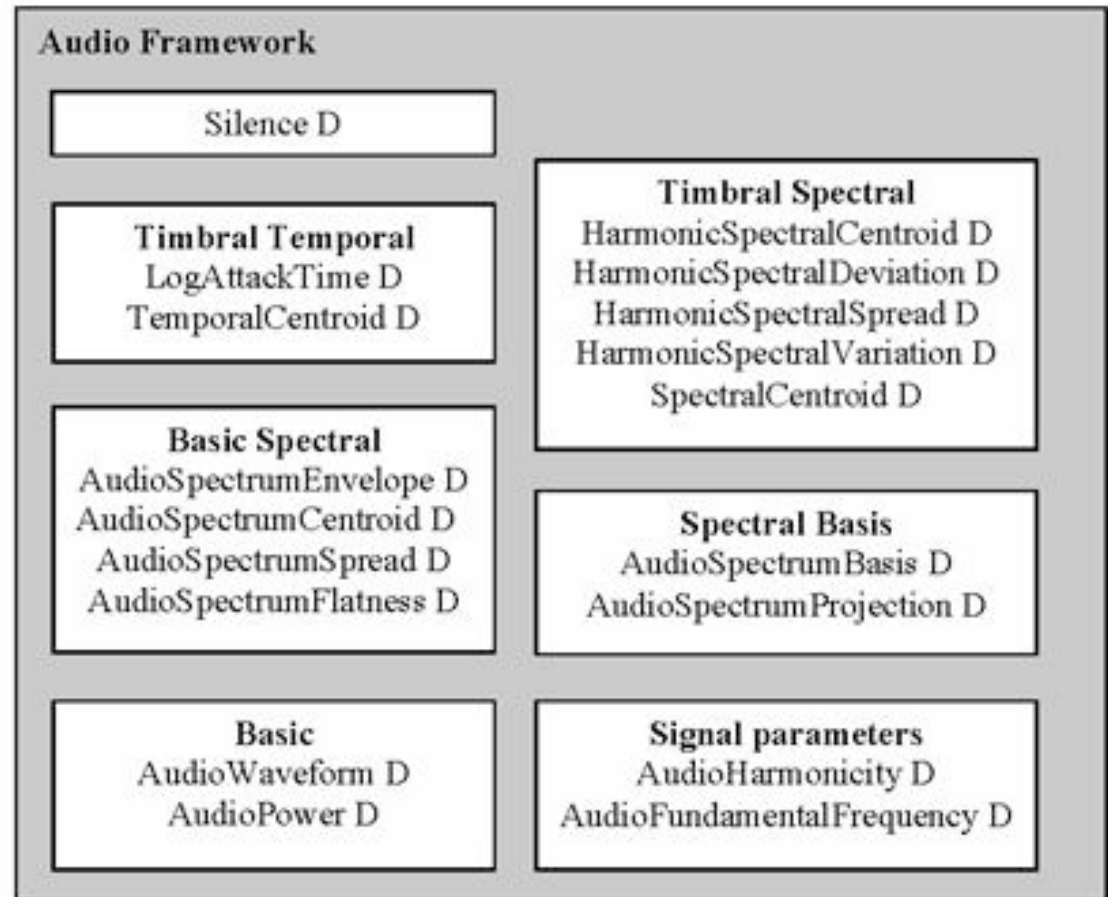
Literature:

Communications of the ACM 49(8), August 2006,
Special section on Music Information Retrieval, pp. 28-60

S. Dasiopoulou, E. Giannakidou, G. Litos, P. Malasioti, and Y. Kompatsiaris.
2011. A survey of semantic image and video annotation tools. In *Knowledge-driven multimedia information extraction and ontology evolution*, G. Paliouras, C. D. Spyropoulos, and G. Tsatsaronis (Eds.). Springer, 196-239.

MPEG-7 Audio Low-Level Descriptors

- Structures:
 - Single scalar value
 - Series of sampled values
- Features:
 - See figure
- MPEG-7 descriptions may contain features described using different (external) methods and algorithms



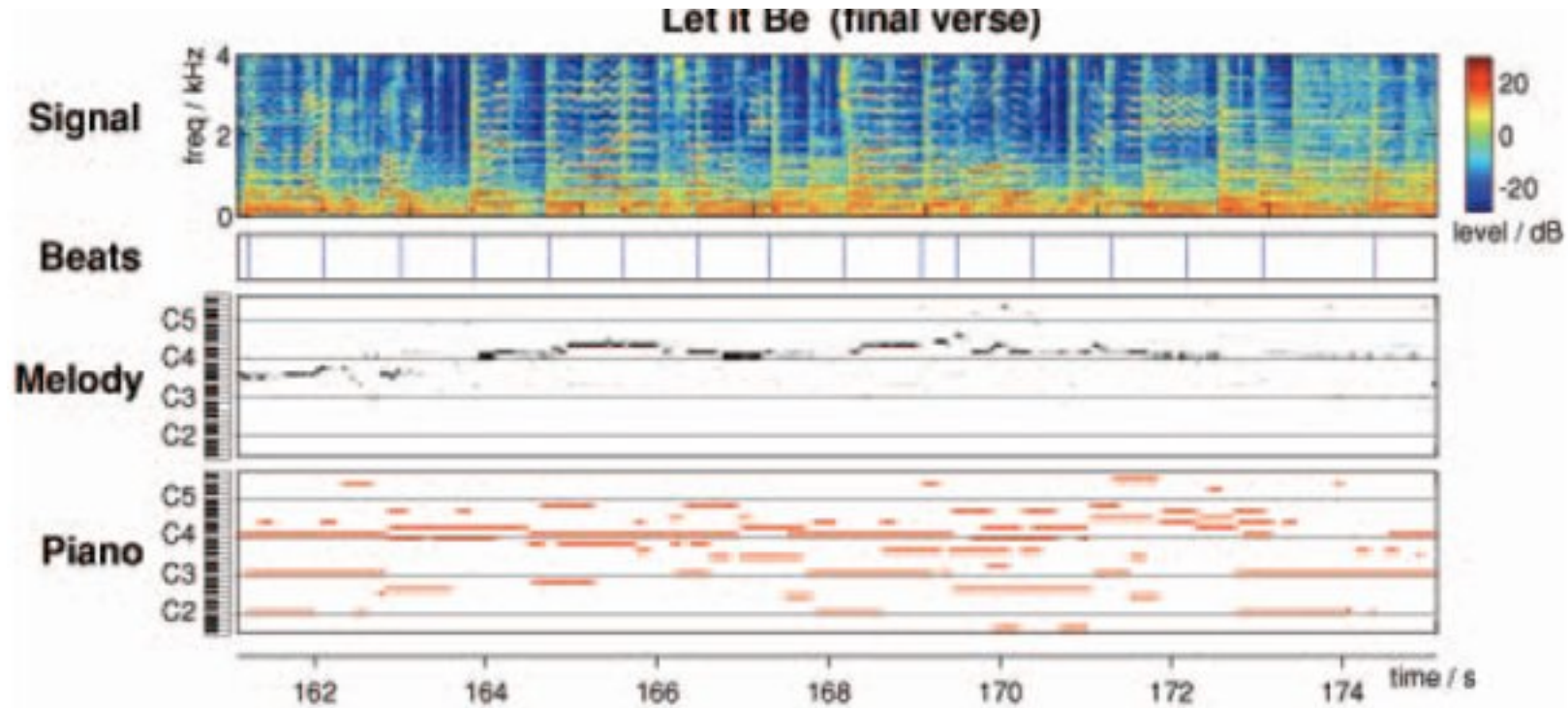
MPEG-7 Audio High-Level Descriptors

- Audio signature
 - Statistical summary of spectral flatness descriptor
 - Fingerprinting, identification of audio content
- Musical instrument timbre
- Melody description
 - MelodyContour (terse, efficient)
 - MelodySequence
 - » Query by Humming
- General sound recognition and indexing
 - Probabilistic classifiers for sound classes
- Spoken content
 - Output and intermediate results of Automatic Speech Recognition (ASR)

Timescales of Musical Information

- Individual music note events
 - Extraction of the music score
 - Identification of instrument playing
- Chords (simultaneous notes)
 - Identification of chords
- Phrase level
 - Tempo extraction
 - Identification of phrases (based on repetition/alternation of segments)
e.g. identification of chorus
- Piece level
 - Genre identification (“rock”, “jazz”, “classical”)

Automatic Score Transcription

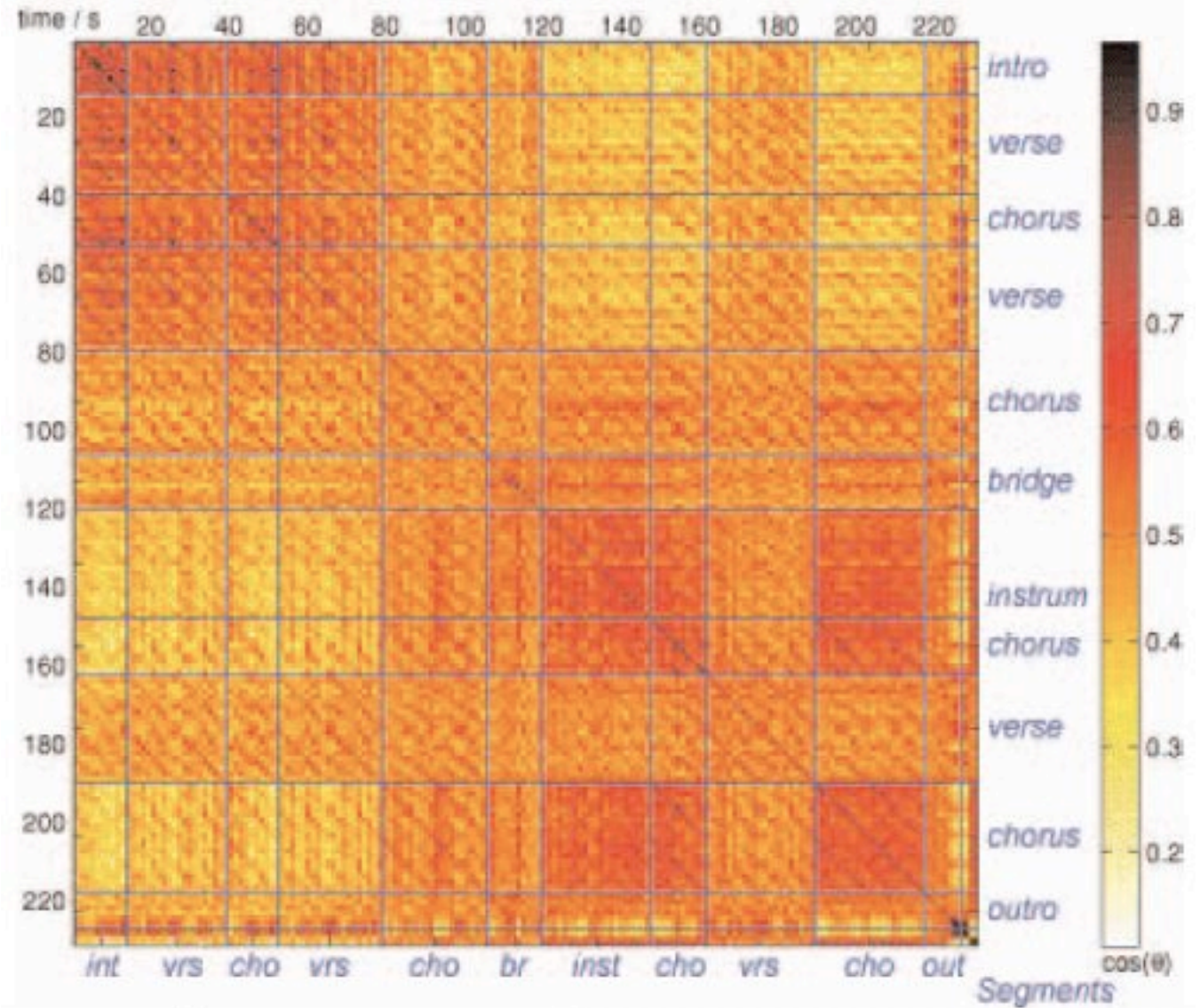


- Beats determined by tempo-smoothed event detector
- Melody recognized by general-purpose support-vector classifier
 - Trained to recognize spectral slices to be labelled with pitch values

Automatic Phrase Detection

- Self-similarity matrix
 - Values represent acoustic similarity
 - Looking for diagonal ridges off the main diagonal
 - Blue lines are manually inserted for comparison

See also:
<http://www.fxpal.com/publications/FXPAL-PR-99-093.pdf>



Example: Shazam Music Tagging (1)



- Commercial service for mobile phones:
Identify music from a short audio sample (*query by example*)
 - See <http://www.shazam.com> (London, founded 2000)
 - A. Wang: The Shazam Music Recognition Service, *Comm. ACM* Aug. 2006
- Challenges:
 - Distinguishing music from noise
 - Dealing with distortions
 - Keeping fingerprints small (in order to deal with millions of songs)
- Basic idea:
 - Spectrogram peaks (energy distribution in time and frequency)¹
 - Few “anchor” peaks are combined with peaks in a certain surrounding zone (time and frequency offsets)
 - » Combinatorial hashing creates 32b fingerprint hash token

¹An overlapping Short-Time Fourier Transform is calculated at regular intervals on the audio data, and a power level is calculated for each resulting time-frequency bin. A bin is a peak if its power level is greater than all the other bins in a bounded region around the bin.

Example: Shazam Music Tagging (2)

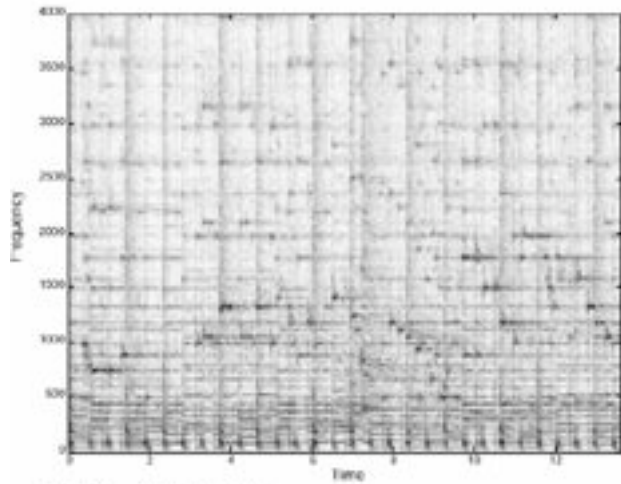


Fig. 1A - Spectrogram

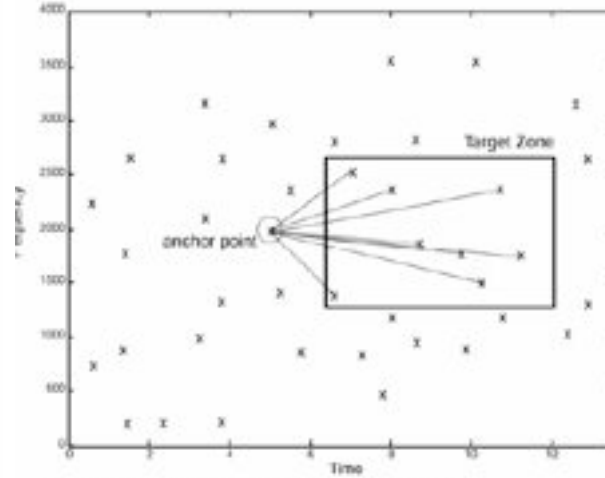


Fig. 1C - Combinatorial Hash Generation

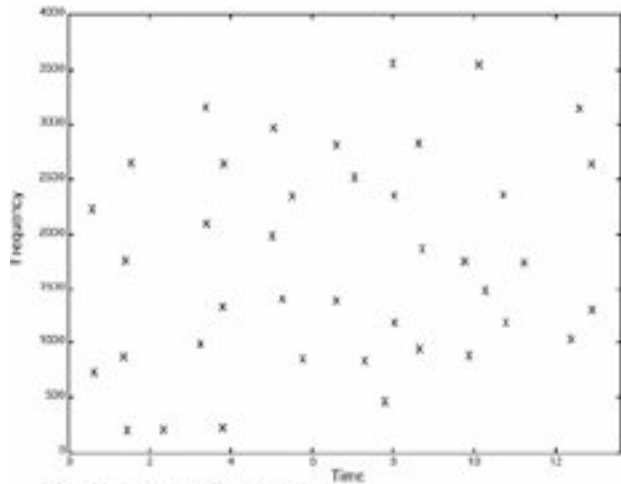


Fig. 1B - Constellation Map

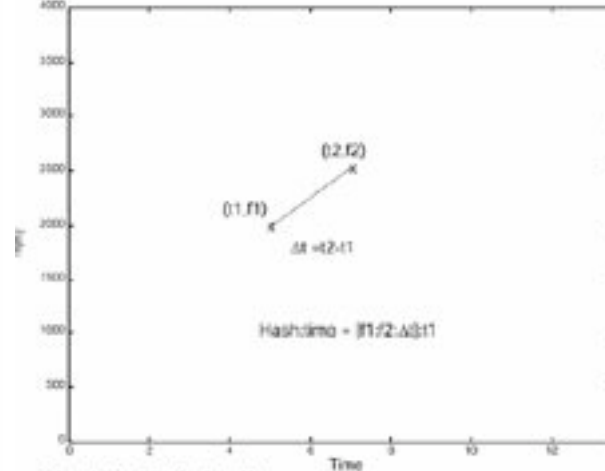


Fig. 1D - Hash details

<http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>

Fingerprint Complexity Tradeoff

- Computing a more complex fingerprint:
 - Increases search time (more tokens to inspect)
 - Improves entropy
 - » Better descriptiveness distinguishes more clearly between items
- Shazam example:
 - Combinatorial expansion increases token number by factor 10 (roughly)
 - Combinatorial expansion accelerates index search by a factor of more than a million!

Example: Shazam Music Tagging (3)

- Comparing tokens from sample and database:
 - Only tokens having peaks from target signal are relevant
 - Even presence of a few well matching tokens is significant
- Temporal alignment of fingerprint features:
 - Matching set of features must have identical relative positions in time
 - Find linear time correspondence
 - » By searching a histogram of relative time differences for peaks

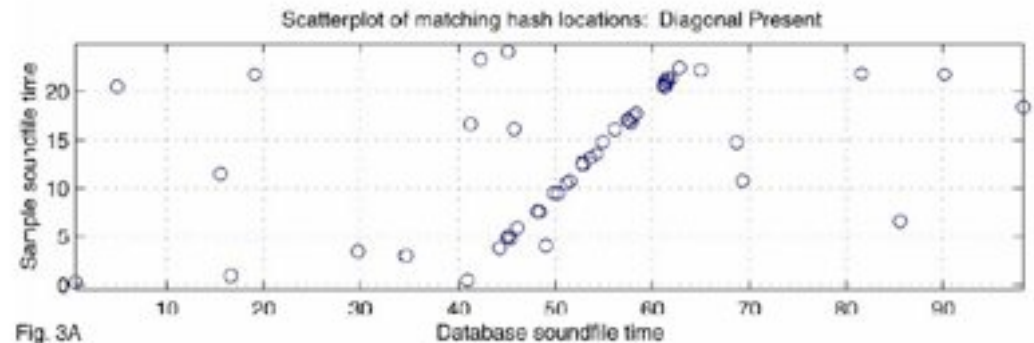


Fig. 3A

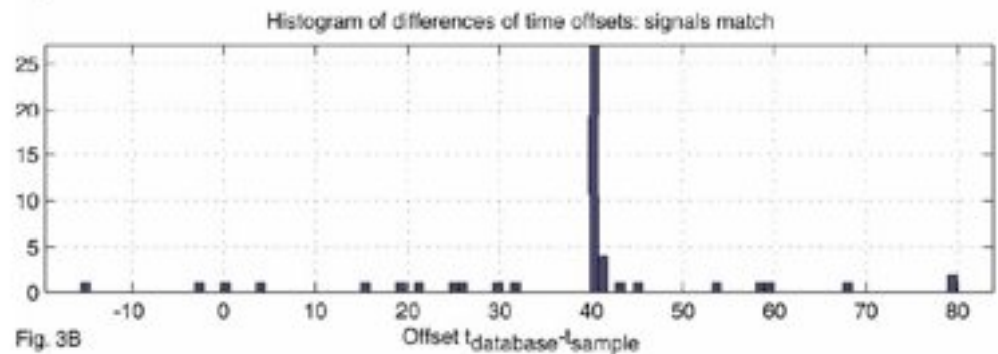


Fig. 3B

Example: Shazam Music Tagging (4)

- Commercial situation:
 - 2012: Shazam app used 10 million times per day
 - More than 11 million tracks in database
- Without Internet connectivity (1999/2000):
 - Query via speech channel, result via text message
- Smartphone apps (Shazam/Encore)
 - Require Internet connectivity
 - Query and result via Internet
- Steady changes in business model:
 - Secondary content for TV
 - Music retail
 - Social music network



“2580” service



shazam.org

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 Feature Extraction for Images and Video

8.3 Feature Extraction for Audio

8.4 Selected Metadata Standards (including MPEG-7)

8.5 Semantic Web Technologies for Multimedia

Literature:

Troncy/Huet/Schenk, Multimedia Semantics - Metadata, Analysis and Interaction, Wiley 2011

B. S. Manjunath, Philippe Salembier, Thomas Sikora:
Introduction to MPEG-7, Wiley 2002

www.chiariglione.org, mpeg-7.joanneum.at,
www.multimedia-metadata.info

MPEG-7

- Moving Picture Experts Group (MPEG)
 - = ISO/IEC JTC1/SC29/WG11 “Moving Pictures and Audio”
 - Main Web presence now: www.chiariglione.org/mpeg
- MPEG-7 “Multimedia Content Description Interface” (since 1996)
 - “ ... a standard for describing the multimedia content data that supports some degree of interpretation of the information’s meaning, which can be passed onto, or accessed by, a device or a computer code. MPEG-7 is not aimed at any one application in particular; rather, the elements that MPEG-7 standardizes support as broad a range of applications as possible.”
- ISO/IEC 15938 standard since 2002, parts still being added
 - MPEG 7 Audiovisual Description Profile (AVDP): 2012
- Industrial uptake very slow
 - Ambitious standard
- Some research and open source prototypes available
 - See e.g.
<http://mpeg7.joanneum.at>,
<http://www.multimedia-metadata.info>

Parts of the MPEG-7 Standard

- MPEG-7 Systems
- MPEG-7 Description Definition Language (DDL)
 - Descriptors (D) and description schemes (DS) specify the syntax and semantics of each *feature* (metadata element)
 - DDL allows the creation of Ds and DSs
 - » XML-based language with some small extensions to XML Schema
- MPEG-7 Visual
- MPEG-7 Audio
- MPEG-7 Multimedia Description Schemes
- MPEG-7 Reference Software
 - eXperimentation Model XM
- MPEG-7 Conformance (rules for conformance checking)
- Extraction and use of MPEG-7 descriptors
- MPEG-7 Profiles and Levels (Profile Schemas, Schema Definition)
- MPEG-7 Query Format

MPEG-7 Profiles

Part 9 of MPEG-7 (2005):

- Simple Metadata Profile (SMP)
 - Single document or simple collection, similar to EXIF or ID3
- User Description Profile
 - Tools for describing personal preferences and usage patterns
 - Adopted by TV-Anytime standard
- Core Description Profile
 - Collections of multimedia content, description of relationships

Later profiles:

- Audiovisual Description Profile (AVDP)
 - Targeted at requirements of audiovisual media production
 - Mainly driven by European Broadcasting Union (EBU)

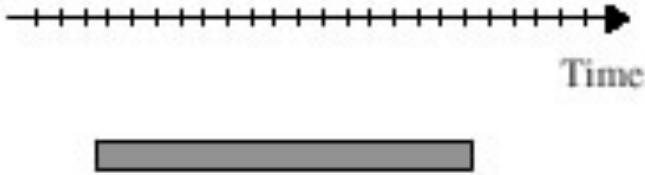

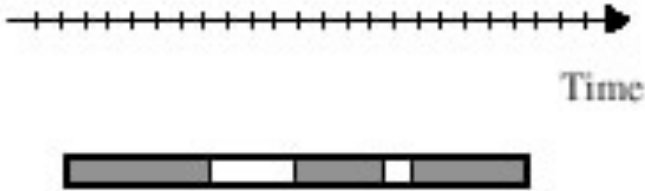

Application Areas of MPEG-7

- Architecture, real estate, and interior design (e.g., searching for ideas).
- Broadcast media selection (e.g., radio channel, TV channel).
- Cultural services (e.g., virtual museums).
- Digital libraries (e.g., image catalogue, musical dictionary).
- Education (e.g., repositories of multimedia courses).
- Home Entertainment (e.g., home video management).
- Investigation services (e.g., human characteristics recognition, forensics).
- Journalism (e.g. searching for video footage of political event).
- Multimedia directory services (e.g. yellow pages, tourist information).
- Multimedia editing (e.g., personalized electronic news service, media authoring).
- Remote sensing (e.g., cartography, ecology, natural resources management).
- Shopping (e.g., searching for clothes that you like).
- Surveillance (e.g., traffic control, surface transportation).
- ...

Structural Content Description: Segments

- A segment represents a section of an audio-visual content item.
- The Segment Description Scheme (DS) is an abstract class (in the sense of object-oriented programming).
- It has nine major subclasses:
 - Still Region DS (spatial)
 - Video Segment DS (temporal)
 - Moving Region DS (spatiotemporal)
 - Audio Segment DS (temporal)
 - AudioVisual Segment DS (temporal)
 - AudioVisual Region DS (spatiotemporal)
 - Still Region 3D DS (3D spatial)
 - Ink Segment DS (electronic ink from pen, smartboard etc.)
 - Multimedia Segment DS (composite of segments)

Examples of Segments

Temporal segment (Video, audio, audio-visual and ink segment)	Spatial segment (Still region)
 <p data-bbox="450 826 495 866">(a)</p> <p data-bbox="819 740 1133 866">Segment composed of one connected component</p>	 <p data-bbox="1200 826 1245 866">(b)</p> <p data-bbox="1563 740 1877 866">Segment composed of one connected component</p>
 <p data-bbox="450 1326 495 1366">(c)</p> <p data-bbox="819 1240 1133 1366">Segment composed of three connected components</p>	 <p data-bbox="1200 1326 1245 1366">(d)</p> <p data-bbox="1563 1240 1877 1366">Segment composed of three connected components</p>

Structural Relations of Segments

- Content structure:
 - Either hierarchical segment decomposition
 - Or general segment relationship graph
- Predefined structural relations in MPEG-7 (can be extended):
 - Generic:
 - » Identical, union, disjoint
 - Spatial:
 - » South, north, west, east, northwest, northeast, southwest, southeast, left, right, below, above, over, under
 - Temporal:
 - » Precedes, follows, meets, metBy, overlaps, overlappedBy, contains, during, strictContains, strictDuring, starts, startedBy, finishes, finishedBy, coOccurs, contiguous, sequential, coBegin, coEnd, parallel, overlapping
- For each relation, the inverse relation is implicitly defined.

Semantic Segmentation/Annotation of Images

```

<Descriptor xsi:type = "ScalableColorType"
  NumberOfCoefficients = "4"
  NumberOfBitplanesDiscarded = "0">
  <Coefficients>
    -217 71 57 8 29 ... 6 0 1
  </Coefficients>
</Descriptor>

```

Sky
Light blue,
Uniform texture



```

<Descriptor xsi:type = "RegionShapeType">
  <MagnitudeOfART>
    15 73 ... 9 12 12 11
  </MagnitudeOfART>
</Descriptor>

```

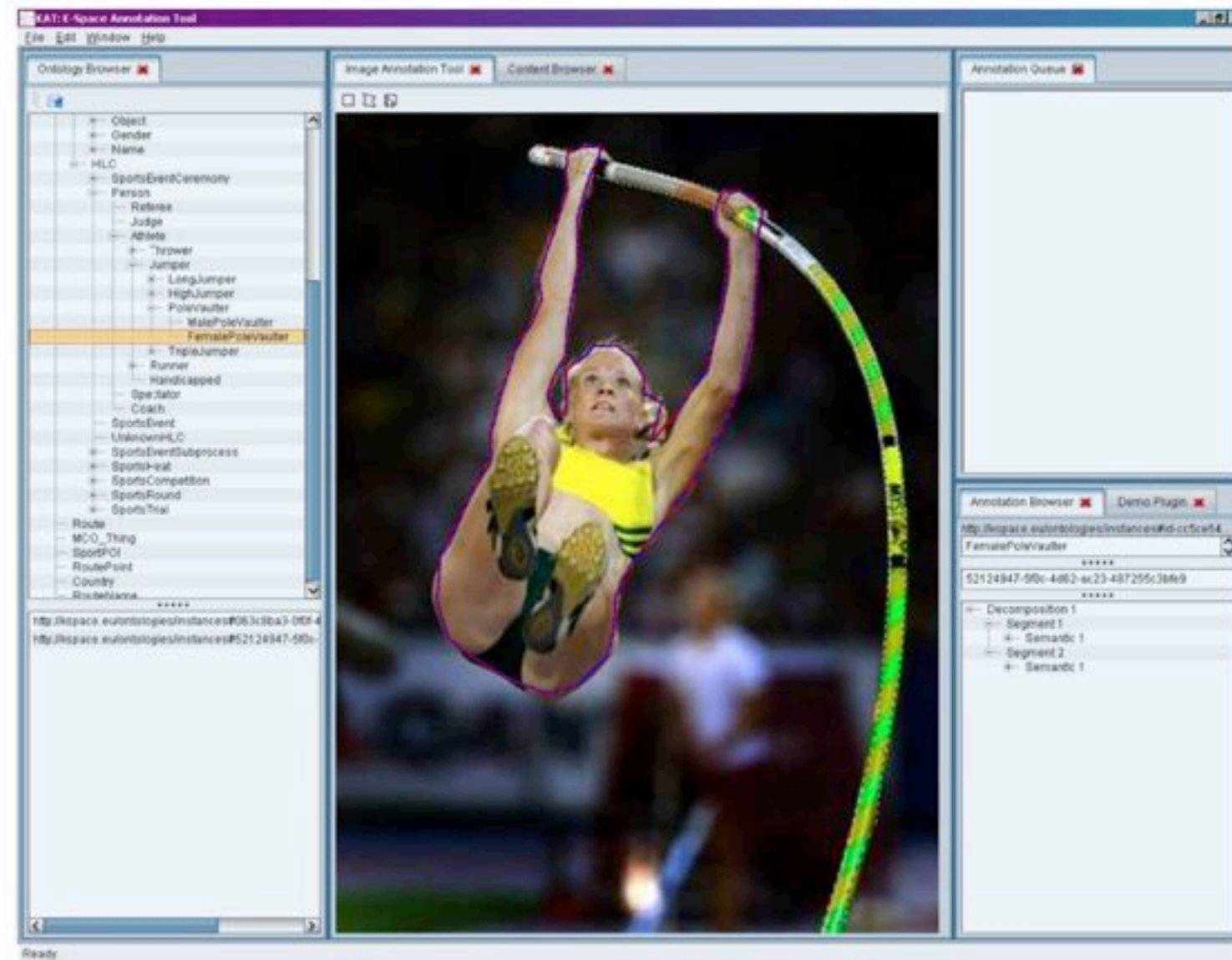
Athlete,
Pole Vaulter

adjacent

Horizontal
Bar

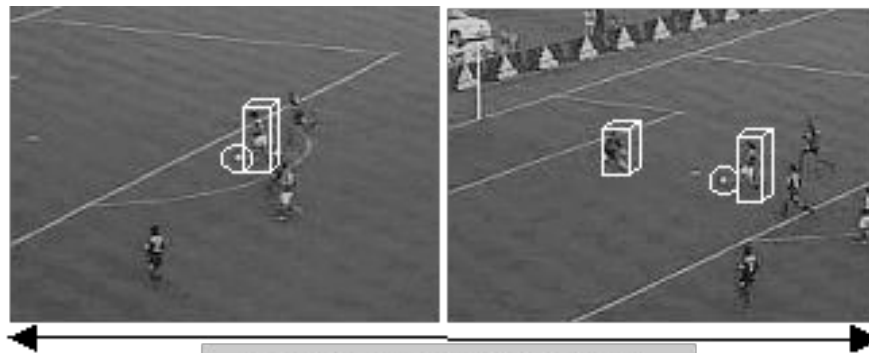
Dasiopoulou et al.

Example of Visual Annotation Tool

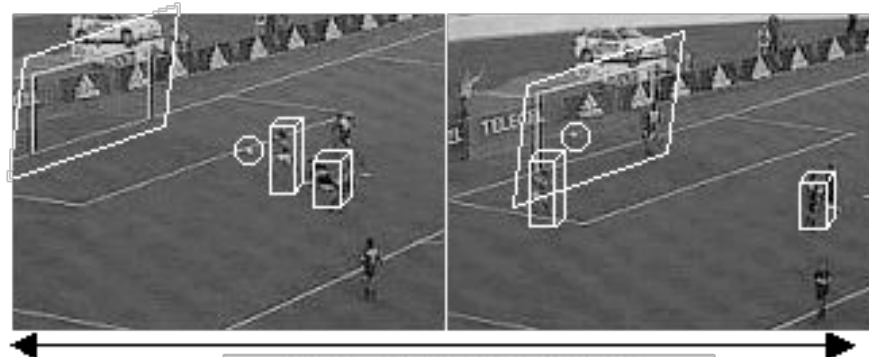


KAT tool
K-Space EU project

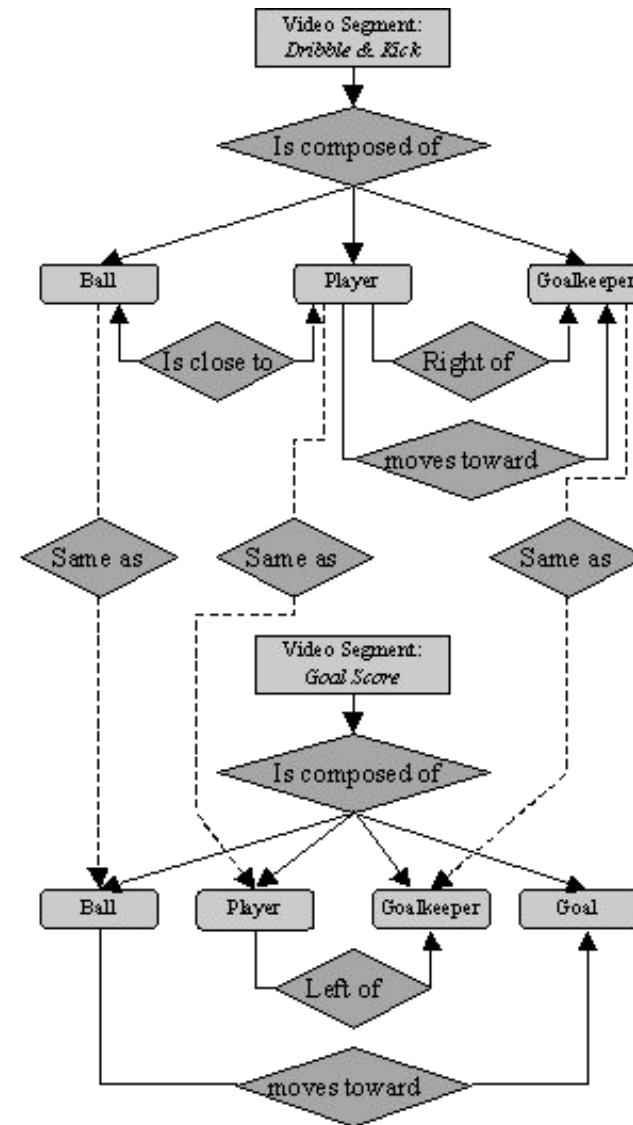
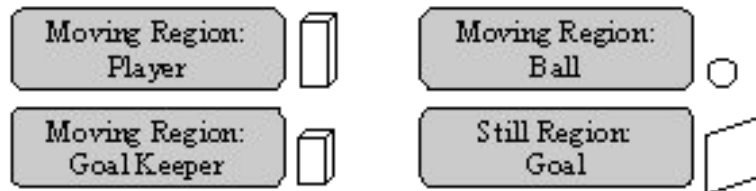
Video Segmentation with Moving Regions



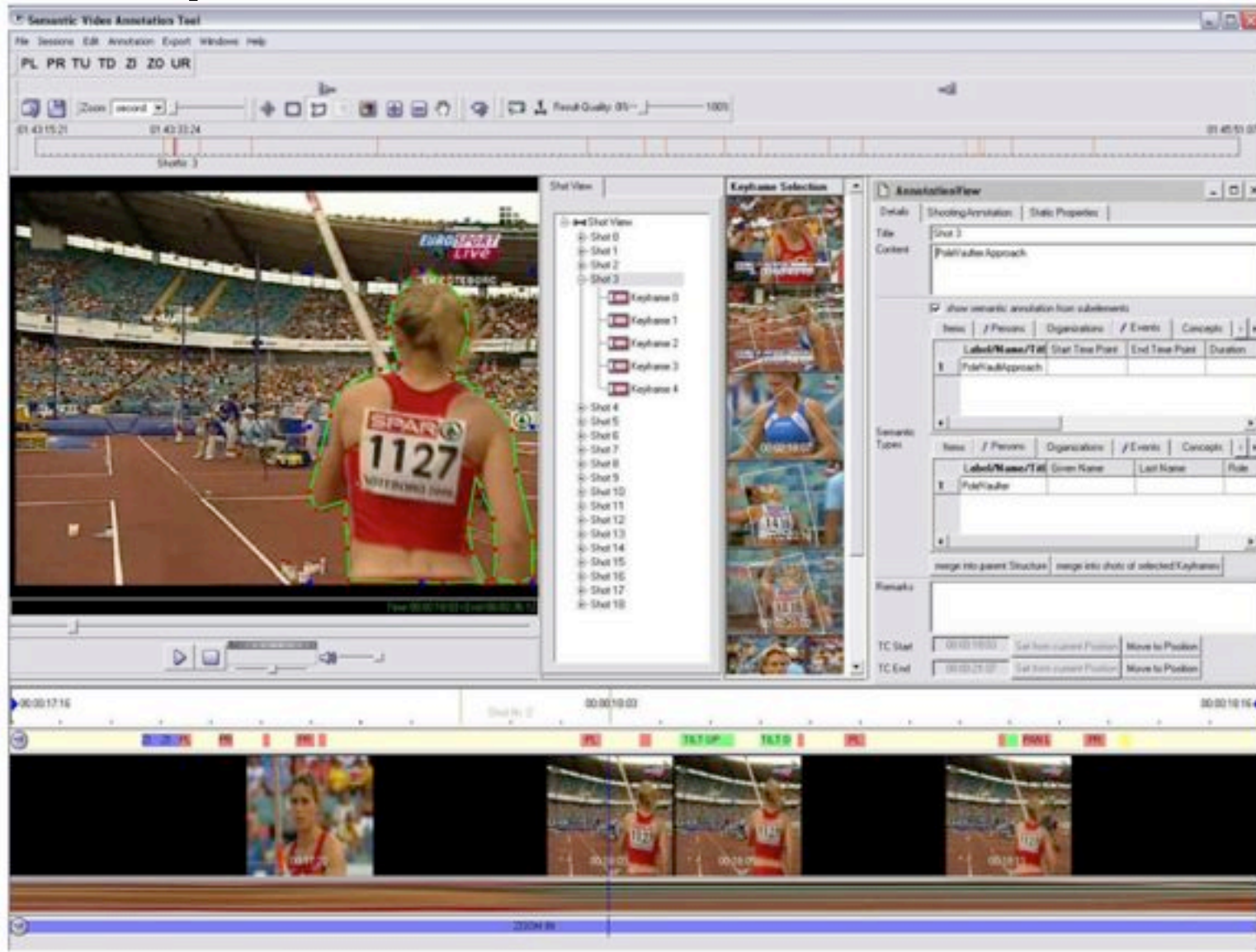
Video Segment: *Dribble & Kick*



Video Segment 2: *Goal Score*

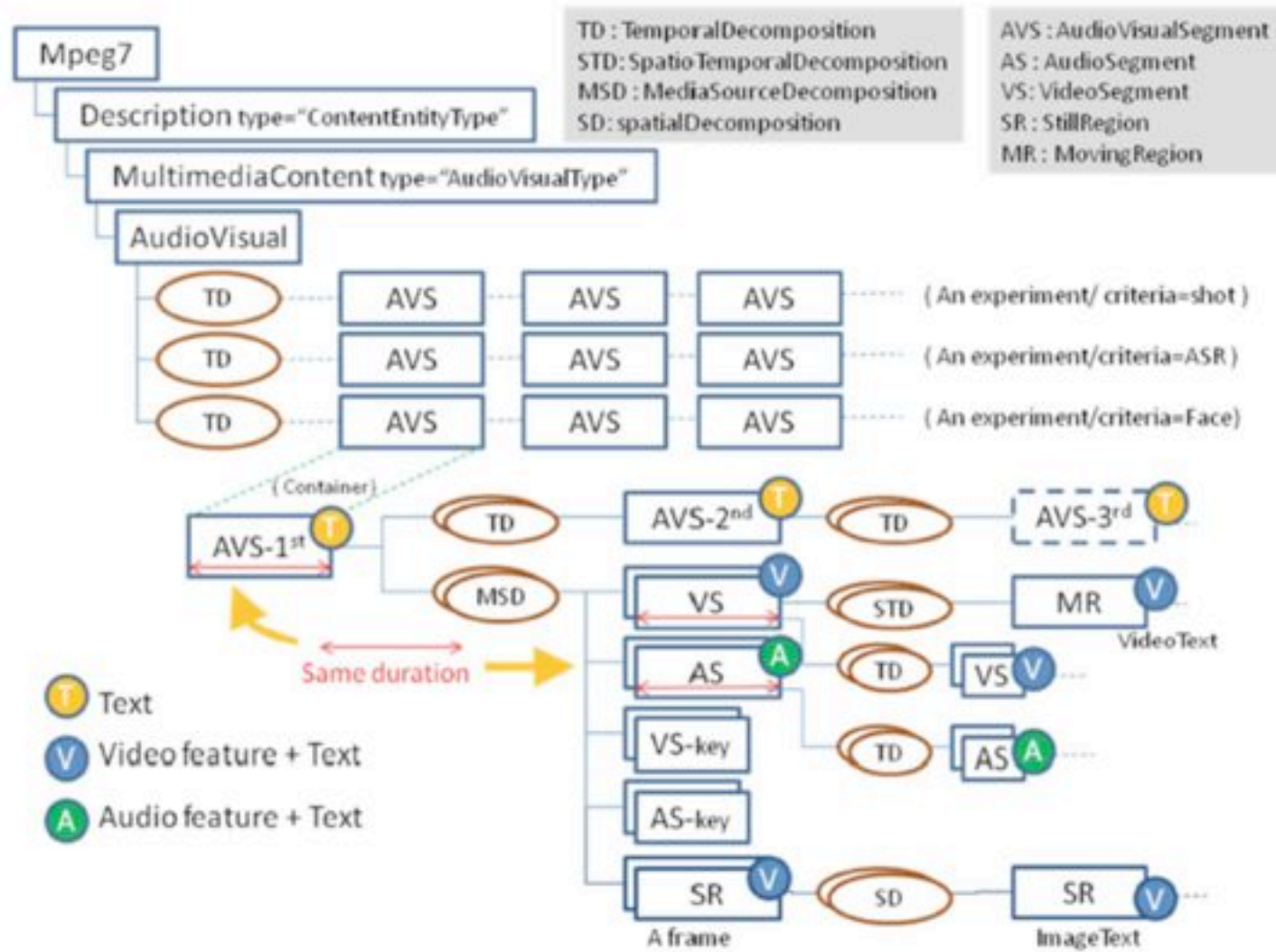


Example of Video Annotation Tool



Semantic Video Annotation Suite (joanneum.at)

Audiovisual Description Profile Structure



joanneum.at

Metadata in Classic Multimedia Formats

- EXIF (Exchangeable Image File Format)
 - EXIF header for captured image or sound
 - Massively used in still-picture cameras
- ID3 for MP3
 - ID3 tag: association of information frames (each for specific metadata)
 - Predefined frames: identification, technical metadata, rights, lyrics, ...
 - Extensible by new frames
- News/G2
 - Developed by IPTC (International Press Telecommunications Council)
 - XML-based specific languages: NewsML-G2, SportsML-G2, EventsML-G2, ProgramGuideML-G2, WeatherML
 - Controlled vocabularies, e.g. IPTC News Codes

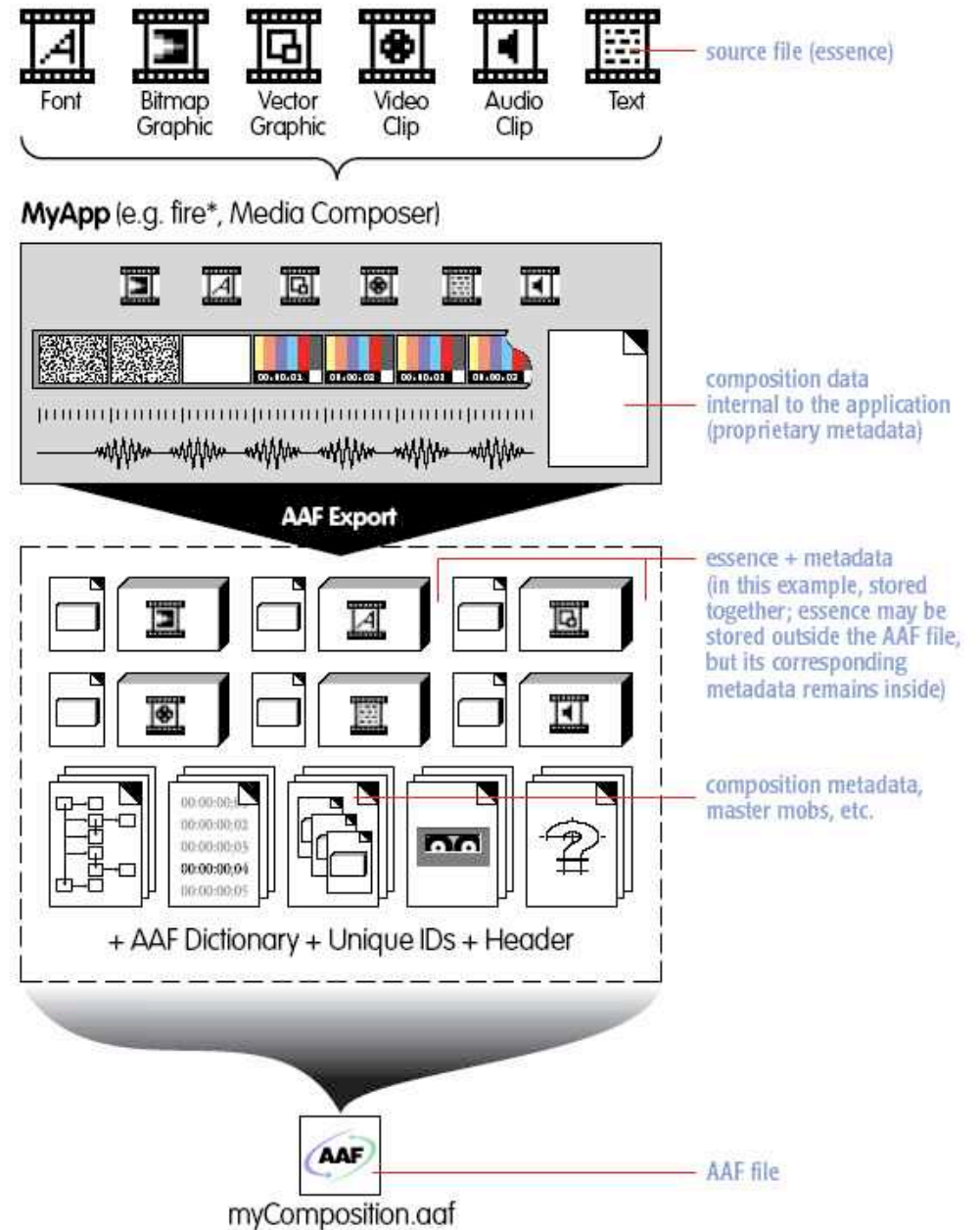
Selected Media Metadata Standards

- Dublin Core Metadata Initiative (DCMI) & PRISM (Publishing Requirements for Industry Standard Metadata)
 - Oriented towards books, magazines, journals etc.
 - Uses XML, RDF, Dublin Core
 - dublincore.org, www.prismstandard.org
- TV Anytime (www.tv-anytime.org)
 - Devoted to audio-visual services making use of local mass-storage
 - Focus on Electronic Program Guide and user profiles
- EBU P/Meta
 - Devoted to material exchange between broadcasting stations
 - Vocabulary for program structure and metadata
- SMPTE Metadata Dictionary
 - Structured list of 1500 metadata elements, used e.e. in MXF format
- Commercial solutions e.g. by Rovi (www.rovicorp.com), ex Macrovision

Integration of Digital Media in Video Production

- Example: Putting together all audio elements for a film soundtrack
 - Music tracks, ambient sound tracks, performer's synchronized sound, ...
 - Metadata related to creation process need homogeneous treatment
- Standards in the broadcasting industry
 - SMPTE (Society of Motion Picture and Television Engineers)
 - EBU (European Broadcasting Union)
 - Working on hardware-based standards for a long time
- EBU/SMPTE Task Force for Harmonized Standards for the Exchange of Program Material as Bit Streams (1996-1999)
 - Results further developed into Advanced Authoring Format (AAF)
 - AAF: Industry-driven, cross-platform, multimedia file format
 - "Advanced Media Workflow Association" (AMWA)
 - » see <http://www.amwa.tv/>

Interchanging Compositions with AAF



Adobe XMP

- Defined by Adobe 2001, since 2007 under BSD license
- Embedding of metadata into distributed files
 - In particular into PDF
- Data model and XML-Based storage model
 - Following the RDF description principle
- Formal schema definitions for metadata properties
- Application:
 - Adobe products (e.g. Photoshop, In-Design)
 - International Press and Telecommunications Council (IPTC) has integrated XMP into its Image Metadata specifications

8 Multimedia Content Description

8.1 Metadata: Concepts and Overview

8.2 Feature Extraction for Images and Video

8.3 Feature Extraction for Audio

8.4 Selected Metadata Standards (including MPEG-7)

8.5 Semantic Web Technologies for Multimedia

Literature:

Troncy/Huet/Schenk, Multimedia Semantics - Metadata, Analysis and Interaction, Wiley 2011

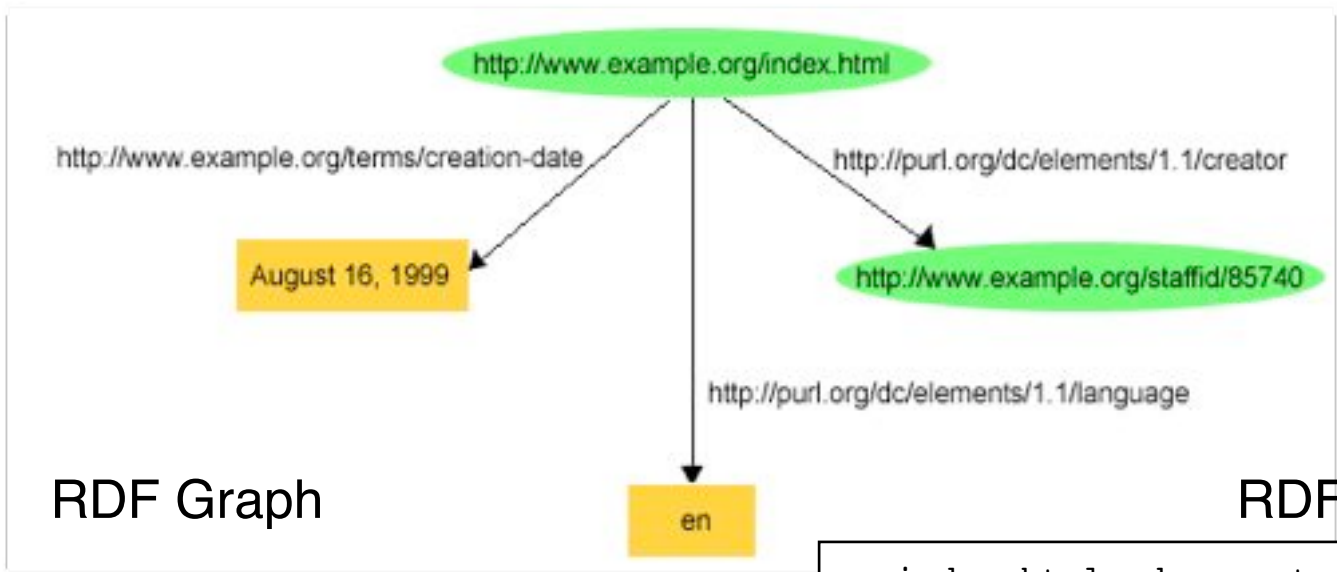
W3C: RDF Primer, <http://www.w3.org/TR/rdf-primer/>

Resource Description Framework RDF

- Language for representing information about resources in the WWW
 - W3C's Semantic Web activity
- *Resource*: Anything that can be identified by a URI (e.g. all Web pages)
- *Property*: An attribute of a described thing which can take on specific values
- *Statement*: A triple consisting of
 - *Subject*: Some resource to be described
 - *Predicate*: A property of the subject
 - *Object*: A specified value
- Properties, values and statements are resources themselves,
 - i.e. can be identified by a URI
 - i.e. can be subject to further description
- RDF documents are collections of (triple) statements
 - written either in XML or in specialized notations (e.g. "Turtle")

RDF Example

- `http://www.example.org/index.html` has a **creator** whose value is John Smith
- `http://www.example.org/index.html` has a **creation-date** whose value is August 16, 1999
- `http://www.example.org/index.html` has a **language** whose value is English



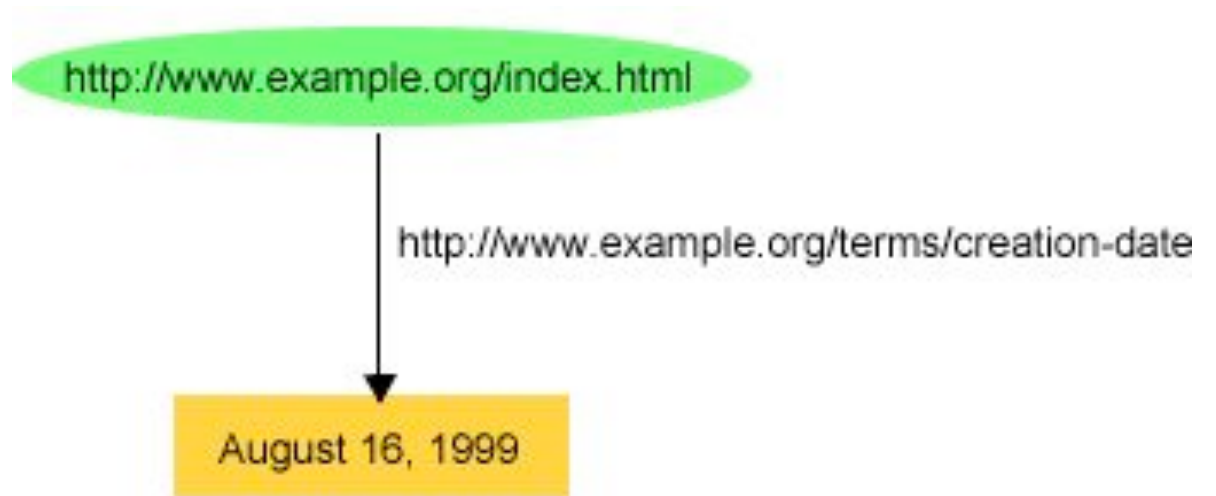
RDF Graph

RDF Triples (using namespaces)

```
ex:index.html dc:creator exstaff:85740 .
ex:index.html exterms:creation-date "August 16, 1999" .
ex:index.html dc:language "en" .
```

RDF/XML: XML Notation for RDF

```
<?xml version="1.0"?>  
<rdf:RDF  
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"   
  xmlns:exterms="http://www.example.org/terms/">  
  
  <rdf:Description  
    rdf:about="http://www.example.org/index.html">  
    <exterms:creation-date>August 16, 1999  
    </exterms:creation-date>  
  </rdf:Description>  
  
</rdf:RDF>
```



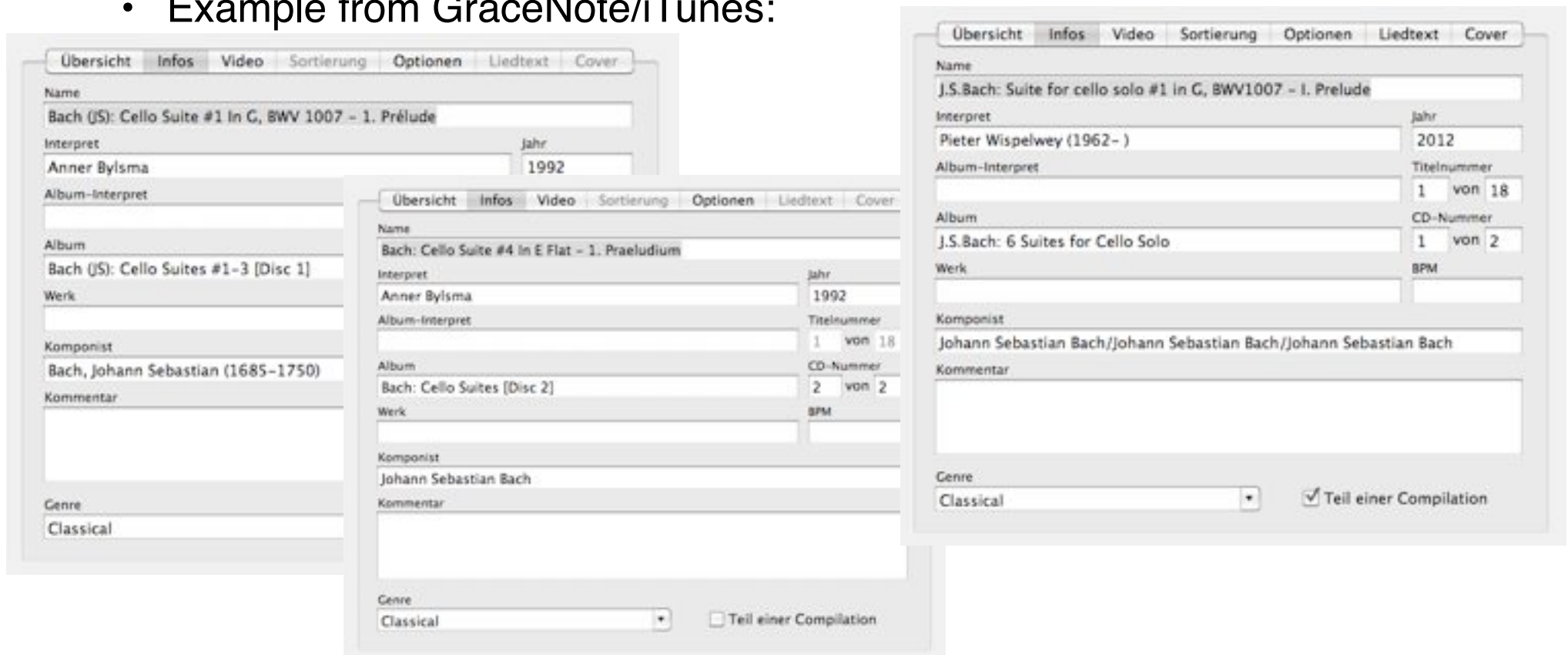
Example: Audio Metadata in DC-based RDF/XML

Multiple statements within one RDF description element

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
    rdf:about="http://www.medien.ifi.lmu.de/team/
    heinrich.hussmann/files/mmn7.mov">
    <dc:creator>Heinrich Hussmann</dc:creator>
    <dc:title>Multimedia Content Description I</dc:title>
    <dc:description>Discusses multimedia metadata
    standards.</dc:description>
    <dc:date>2012-12-20</dc:date>
    <dc:format>audio/mp4</dc:format>
  </rdf:Description>
</rdf:RDF>
```

The Need for Ontologies

- Fixed schemata for text-based annotation are insufficient for practical purposes (or are often misused).
- Fine-grained and flexible semantic schemata are needed.
- Example from GraceNote/iTunes:



Defining Ontologies

- *Ontology*: Controlled vocabulary to express semantic information
 - Knowledge representation through concepts in terms of: types, properties, interrelationships
- RDF Schema
 - Simple set-theoretic ontologies, defines vocabularies for RDF
- OWL, Web Ontology Language by W3C
 - Extension of *RDF Schema*
 - Based on *Description Logics*, powerful mathematical semantics
 - Ontologies can be denoted in RDF syntax itself
- *W3C Ontology for Media Resources*
 - <http://www.w3.org/TR/mediaont-10> (Recommendation Feb 2012)
 - Defines a standard terminology for multimedia segmenting and annotation
 - Gives detailed mappings for most commonly used metadata standards (incl. DC, EXIF, ID3, MPEG-7, QuickTime, YouTube, ...)
 - Fully defined in OWL/RDF

Embedding Semantic Metadata into the Web

- RDFa: “RDF in Attributes”
 - Structured Data Markup for Web Documents
 - W3C WG Note August 2013, see <http://www.w3.org/TR/xhtml1-rdfa-primer/>
 - Idea: Add information to Web documents on the meaning of its contents in machine-readable form
 - Applicable to all XML languages and HTML5

- Example:

```
<html>
  <head>...</head>
  <body> ...
    <h2 property="http://purl.org/dc/terms/title">
      The Trouble with Bob</h2>
    <p>Date:
      <span property="http://purl.org/dc/terms/created">
        2011-09-10</span>
    </p>...
  </body>
```

- Can be used, for instance, to add license conditions to links!