

MPEG Standards Enabling Universal Multimedia Access

MPEG-7 Universal Multimedia Access

Hermann Hellwagner · Christian Timmerer

Dept. of Information Technology, Klagenfurt Univ.,
Austria

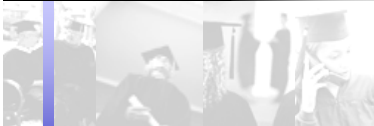
1st Int'l. Conf. on

Automated Production of Cross Media Content for Multi-channel Distribution
~AXMEDIS 2005~

December 1, 2005

Acknowledgement:

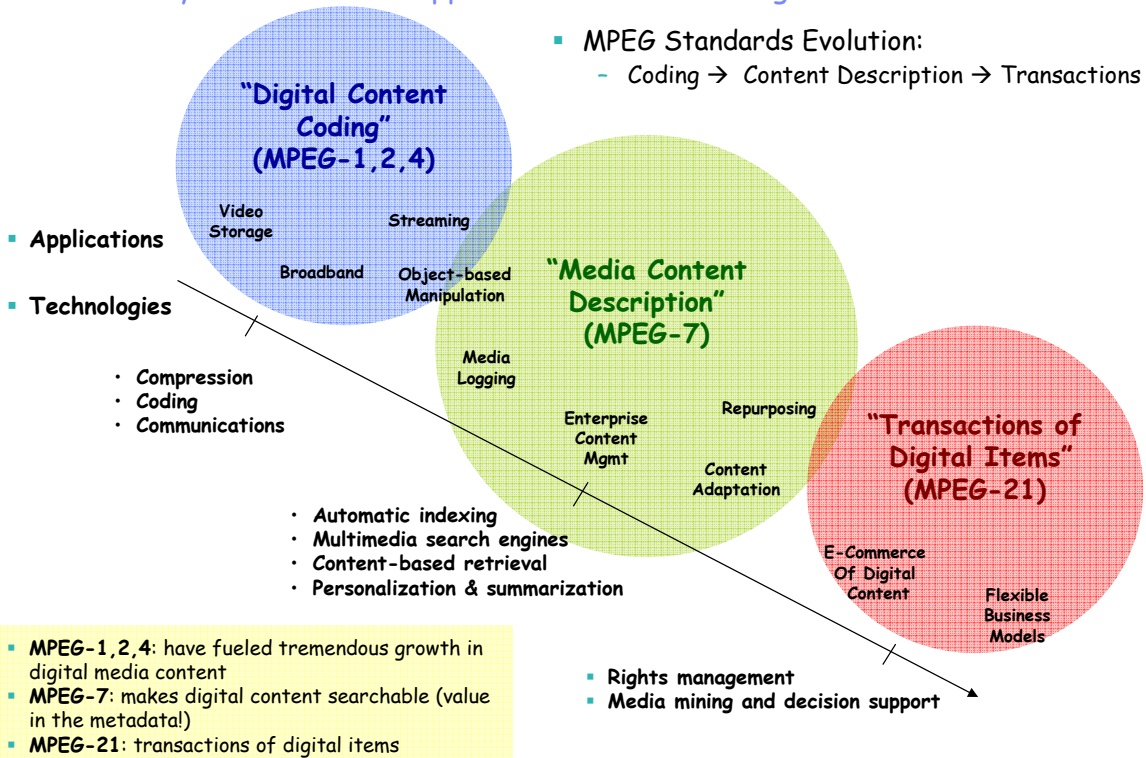
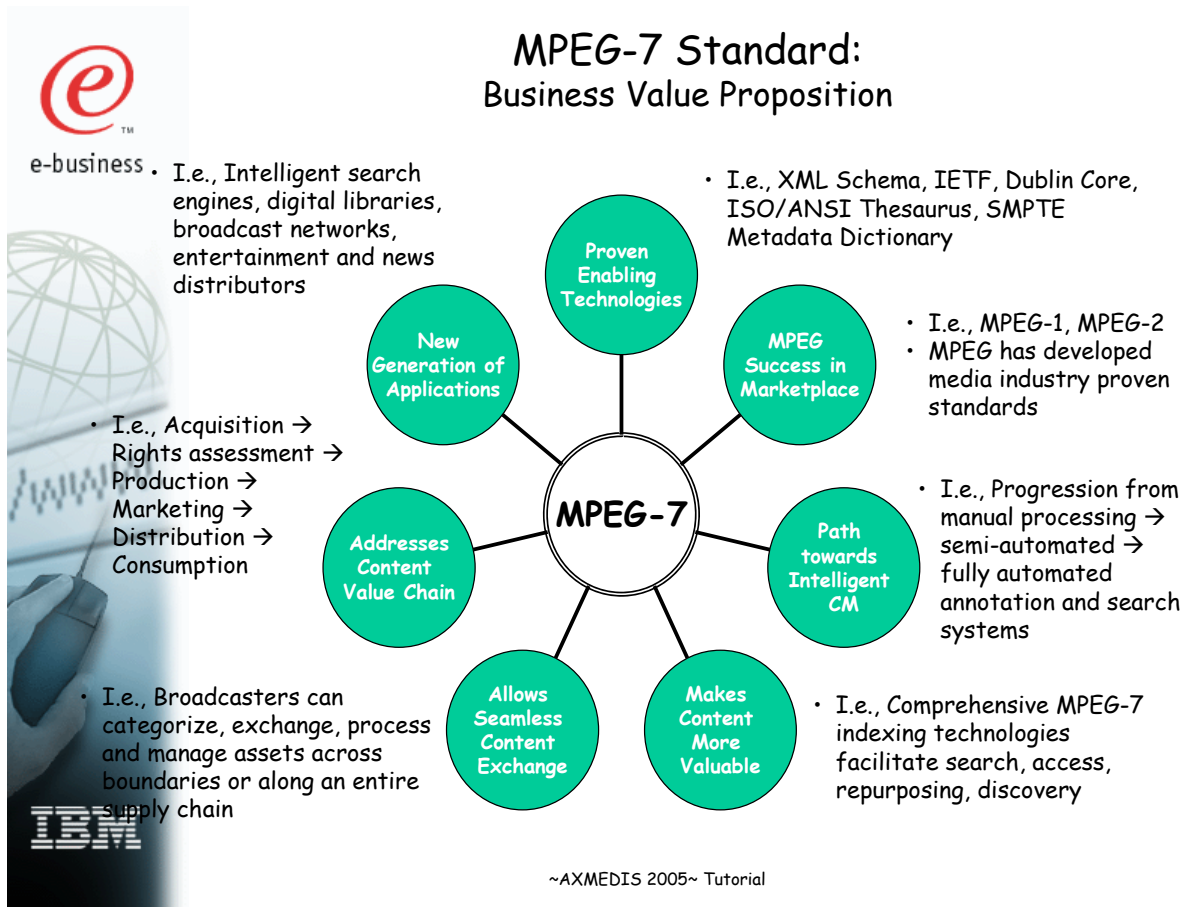
John R. Smith, IBM T. J. Watson Research Center
<http://www.research.ibm.com/people/j/jsmith/>



MPEG-7 Universal Multimedia Access

- MPEG-7 as Part of the MPEG Family of Standards
- Brief Overview of MPEG-7
- MPEG-7 UMA Tools
 - Variations
 - Media Information/Profile
 - Space and Frequency Views
 - Summaries
 - Semantic Descriptions
 - Transcoding Hints

MPEG Family of Standards: Applications and Technologies

MPEG-7 Standard:
Business Value Proposition



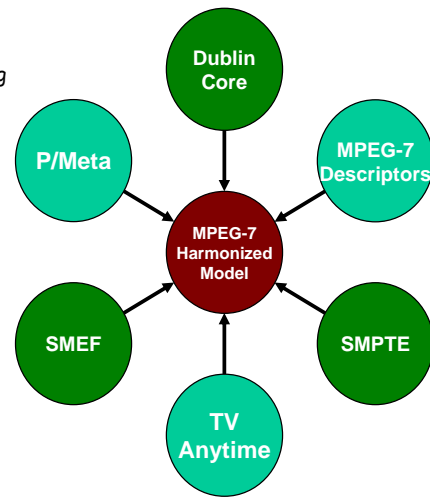
e-business



Harmonization of Multimedia Metadata Standards

MPEG-7, SMEF, P/Meta, SMPTE, Dublin Core, TV-Anytime, Indecs

- **MPEG-7: Moving Picture Experts Group**
 - Infrastructure standard for Multimedia Metadata
 - Supports interpretation of the information's meaning
 - Supports broad range of applications
- **SMEF - Standard Media Exchange Framework:**
 - BBC developed data models for information involved in the Production, Development, Use, and Management of media assets
- **P/Meta - EBU P/Meta Project:**
 - Exchange of program content between high-level business functions of EBU members: Production, Delivery/Broadcast, & Archive
- **SMPTE - Metadata dictionary & MXF:**
 - Addresses Program Interchange independent of format
- **Dublin Core Metadata Initiative:**
 - Interoperable online metadata standards supporting broad range of purposes and business models.
- **TV-Anytime - TV-Anytime Metadata:**
 - Attractors/descriptors used e.g. in Electronic Program Guides (EPG), or in Web pages to describe content.
- **Indecs - Indecs Metadata Framework**
 - An international initiative of rights owners creating metadata standards for e-commerce.



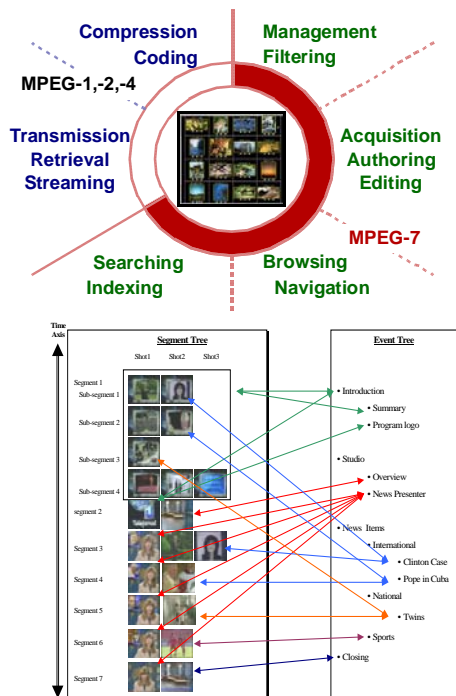
- **MPEG-7 Harmonized Model:**
 - Harmonized elements with other standards and existing practices
 - Extensible framework
 - Registration authority for classification schemes, controlled terms, ontologies

~AXMEDIS 2005~ Tutorial

IBM T. J. Watson Research Center

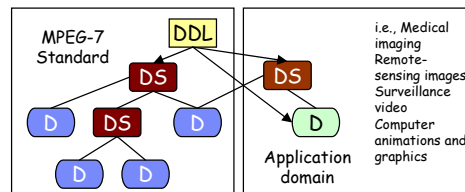


MPEG-7 Overview (XML Metadata for Multimedia Content Description)



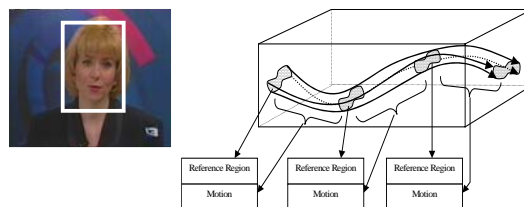
■ MPEG-7 Normative elements:

- Descriptors (D) and Description Schemes (DS)
- Description Definition Language (DDL)
- Extensible for application domains

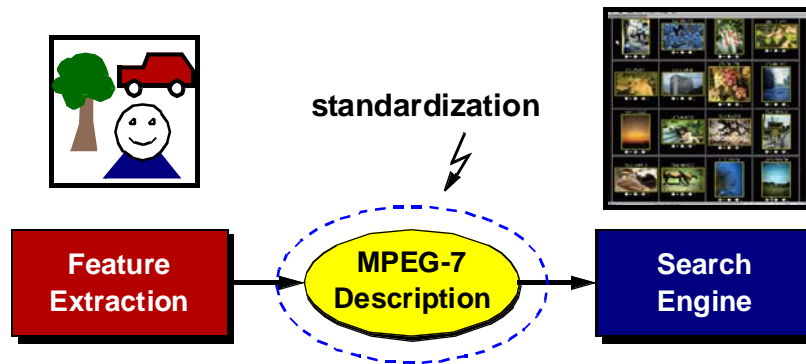


■ Rich, highly granular multimedia content description:

- Video segments, moving regions, shots, frames, ...
- Audio-visual features: color, texture, shape, ...
- Semantics: people, events, objects, scenes, ...



MPEG-7 Standard Scope

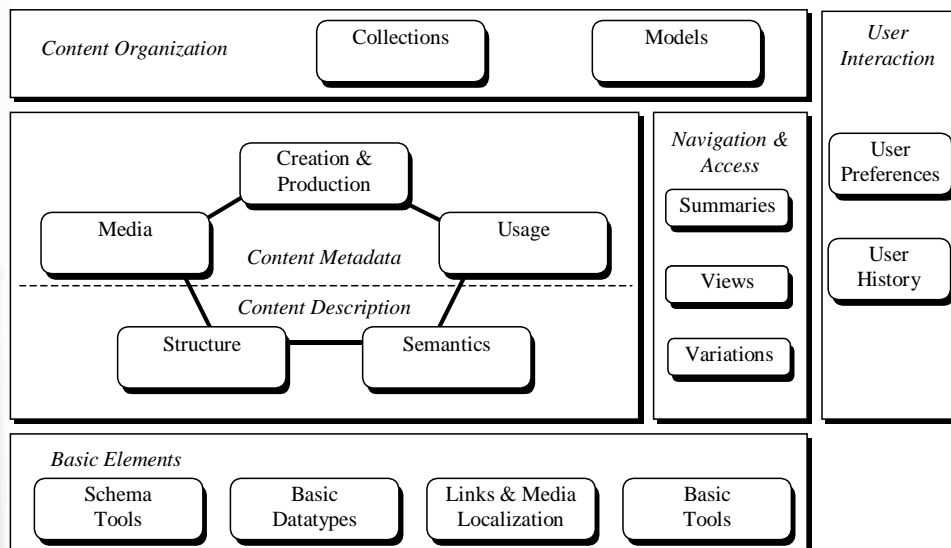


Industry Competition

Feature Extraction: Content analysis (D, DS) Feature extraction (D, DS) Annotation tools (DS) Authoring (DS)	MPEG-7 Scope: Description Schemes (DSs) Descriptors (Ds) Language (DDL) Coding Schemes (CS) MPEG-7 Concepts	Search Engine: Searching & filtering Classification Complex querying Indexing Personalization
---	---	---

~AXMEDIS 2005~ Tutorial

MPEG-7 Multimedia Description Schemes

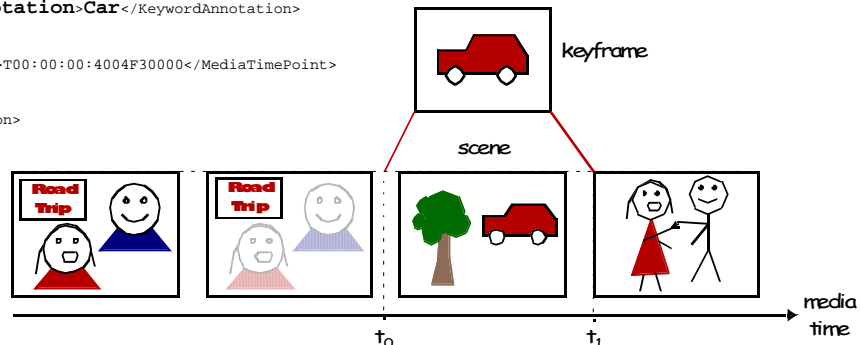


~AXMEDIS 2005~ Tutorial

Example MPEG-7 XML Video Description

- Example of MPEG-7 description of video scene & keyframe:

```
<Mpeg7>
  <Description xsi:type="ContentEntityType">
    <MultimediaContent xsi:type="VideoType">
      <Video>
        <MediaLocator><MediaUri>video.mpg</MediaUri></MediaLocator>
        <TemporalDecomposition>
          <VideoSegment>
            <TextAnnotation type="scene">
              <FreeTextAnnotation>Drive to the country</FreeTextAnnotation>
            </TextAnnotation>
            <MediaTime>
              <MediaTimePoint>T00:00:00:0F30000</MediaTimePoint>
              <MediaIncrDuration mediaTimeUnit="PT1001N30000F">12</MediaIncrDuration>
            </MediaTime>
            <TemporalDecomposition>
              <VideoSegment>
                <TextAnnotation type="keyframe">
                  <KeywordAnnotation>Car</KeywordAnnotation>
                </TextAnnotation>
                <MediaTime>
                  <MediaTimePoint>T00:00:00:4004F30000</MediaTimePoint>
                </MediaTime>
              </VideoSegment>
            </TemporalDecomposition>
          </VideoSegment>
        </TemporalDecomposition>
      </Video>
    </MultimediaContent>
  </Description>
</Mpeg7>
```



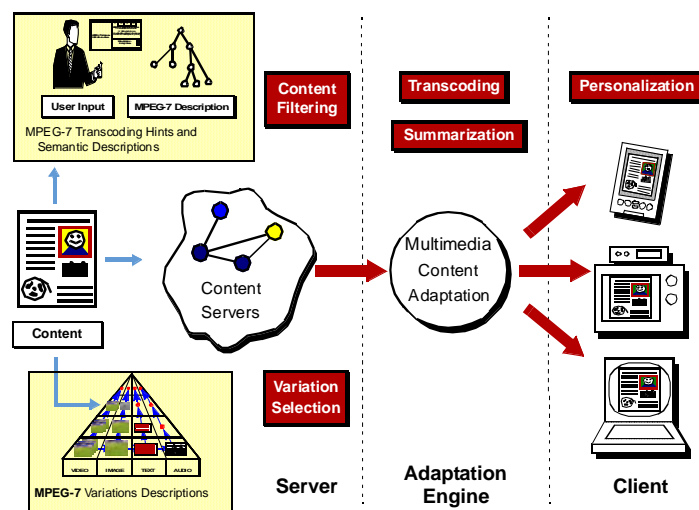
Universal Multimedia Access (UMA)

- 1st generation UMA:

- Enable access of multimedia content anytime, anywhere, and without concern for particular formats, devices, networks, etc.
- Transcoding proxy architectures (eg., IBM Websphere Transcoding Publisher, Telestream Flipfactory)

- 2nd generation UMA:

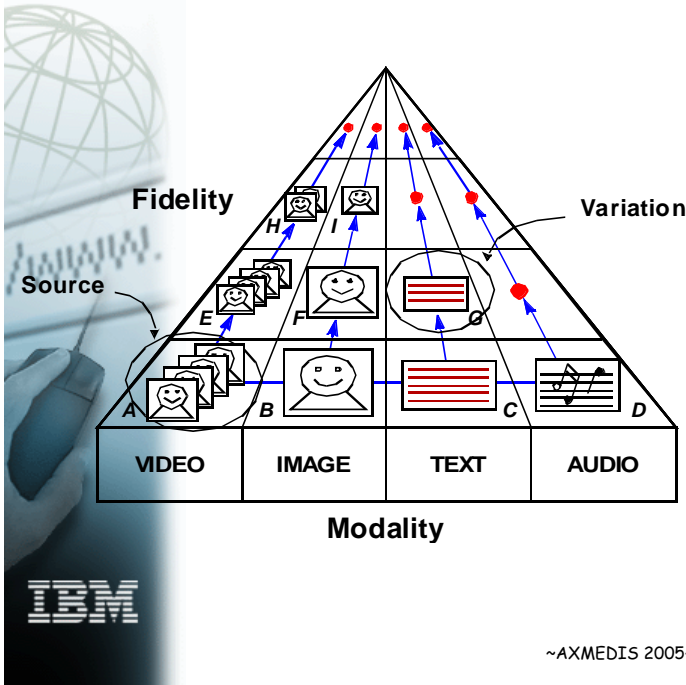
- Customizes content at a high-semantic level through summarization and personalization according to user preferences and usage context



- UMA standards support:

- MPEG-7 UMA Application (Transcoding Hints DS, Variations DS, Summary DS, Semantic DS, Views DS)
- MPEG-21 Digital Item Adaptation (Usage Environment Description, Adaptation QoS, generic Bitstream Syntax Description, etc.)

Example: MPEG-7 MDS Variation Description



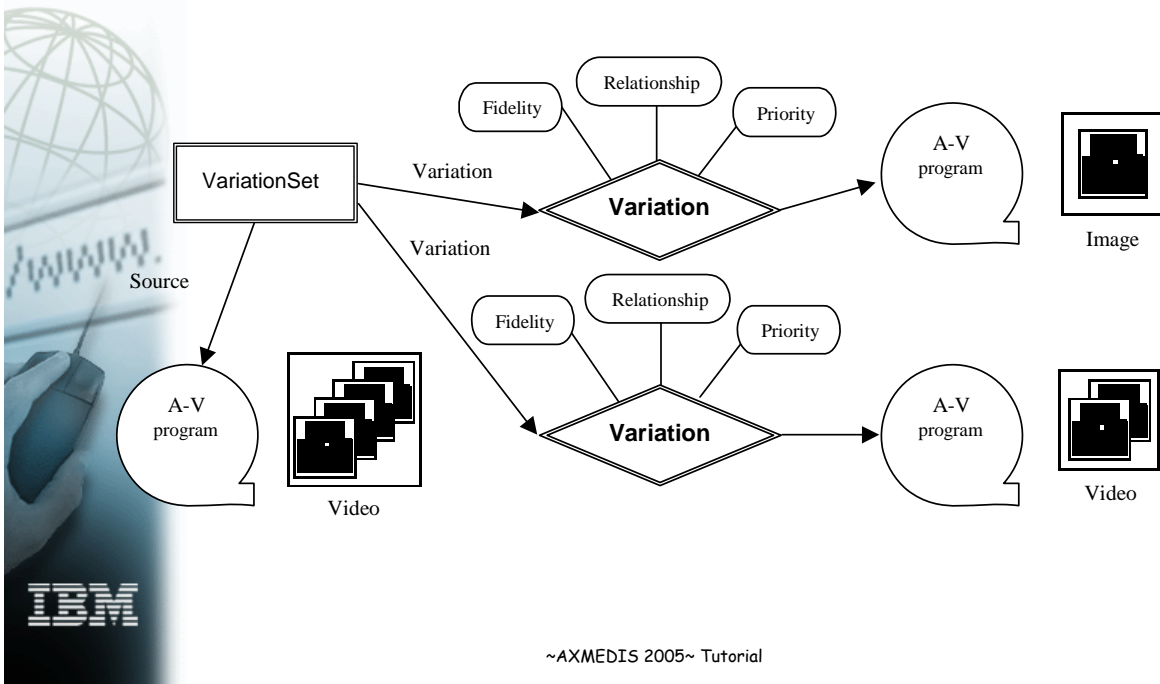
```
<VariationSet>
  <Source><Video>
    <MediaInformation><MediaInstance>
      <MediaLocator><MediaURI> A
    </MediaURI></MediaLocator>
    </MediaInstance></MediaInformation>
    <CreationInformation><Creation>
      <Title><TitleText> Soccer video
    </TitleText></Title>
    </Creation></CreationInformation>
  </Video></Source>

  <Variation fidelity="0.85" priority="1">
    <VariationProgram>
      <MediaLocator><MediaURI> A
    </MediaURI></MediaLocator>
    </VariationProgram>
    <VariationRelationship> extract
  </VariationRelationship>
</Variation>

  <Variation fidelity="0.75" priority="3">
    <VariationProgram>
      <MediaLocator><MediaURI> C
    </MediaURI></MediaLocator>
    </VariationProgram>
    <VariationRelationship> extract
  </VariationRelationship>
    <VariationRelationship> languageTranslation
  </VariationRelationship>
</Variation>
</VariationSet>
```

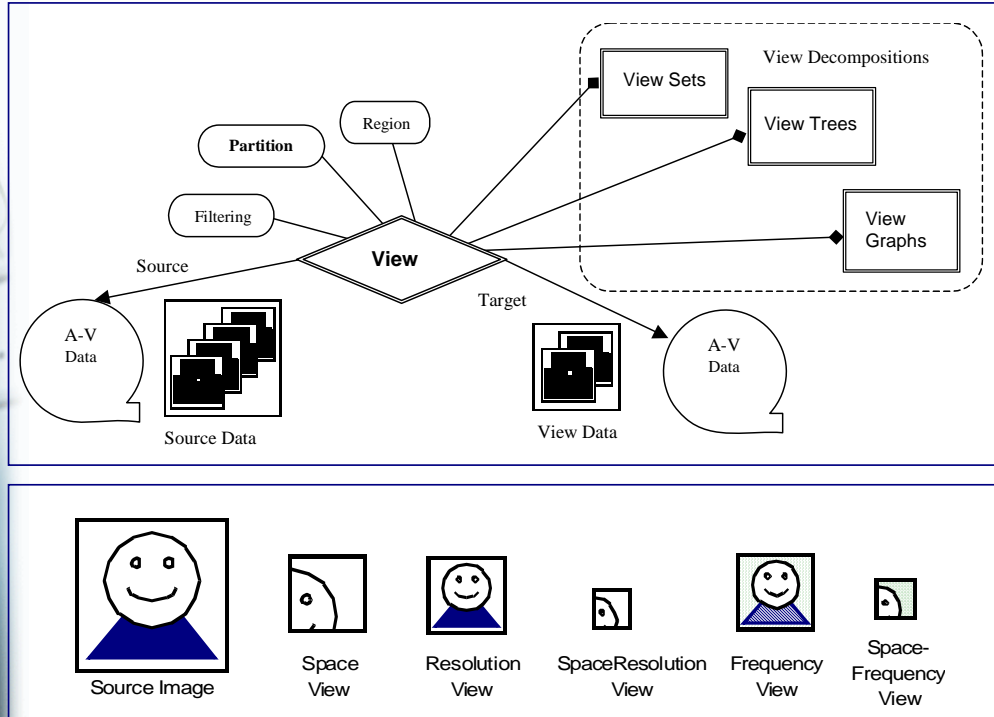
~AXMEDIS 2005~ Tutorial

MPEG-7 MDS: Variations DS



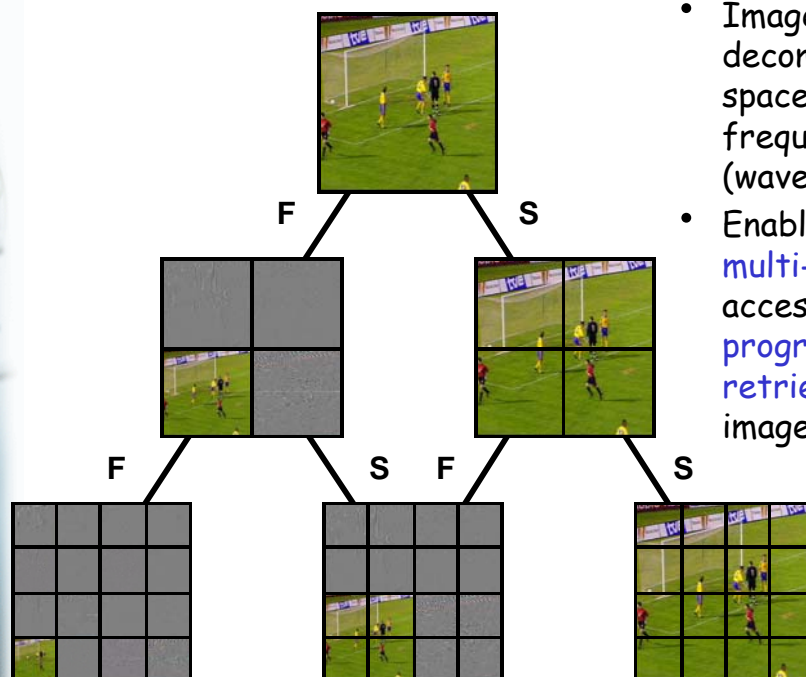
~AXMEDIS 2005~ Tutorial

MPEG-7 MDS: Space and Frequency Views



~AXMEDIS 2005~ Tutorial

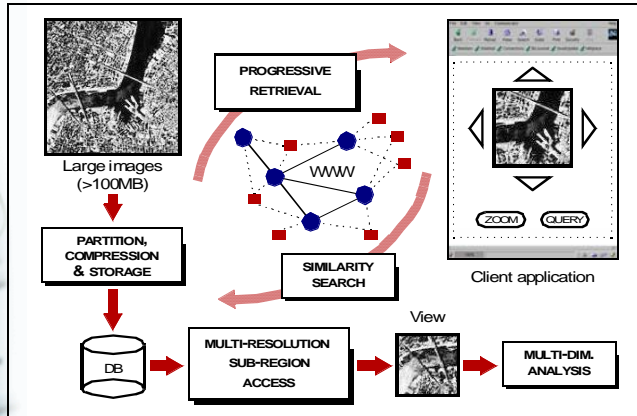
MPEG-7 MDS: Space and Frequency Graph



- Image view decomposition in space and frequency (wavelets)
- Enables efficient multi-resolution access and progressive retrieval of the image data

~AXMEDIS 2005~ Tutorial

MPEG-7 Space and Frequency Graph



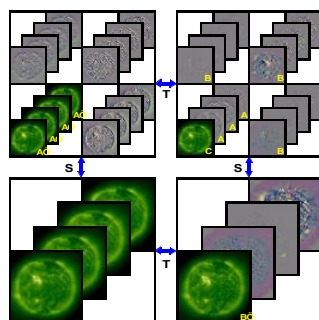
- **SFGraph** - space and frequency wavelet decomposition
- Scalable image compression
- Progressive image retrieval

- **Image Zooming** - allows progressive retrieval of large images and maps



~AXMEDIS 2005~ Tutorial

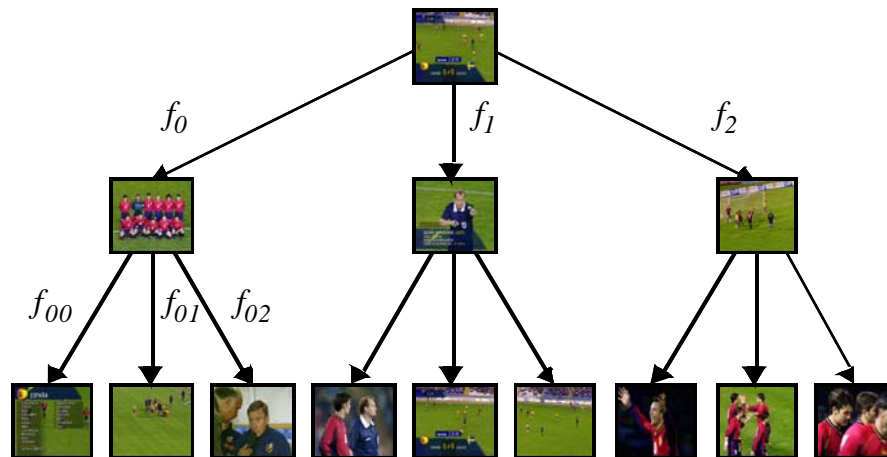
MPEG-7 Video View Graph Decomposition



- **VideoZoom** - 3-D wavelet decomposition
- Progressive video retrieval
- Scalable video delivery
- Secure video transcoding

~AXMEDIS 2005~ Tutorial

MPEG-7 Interactive Video Summary Example

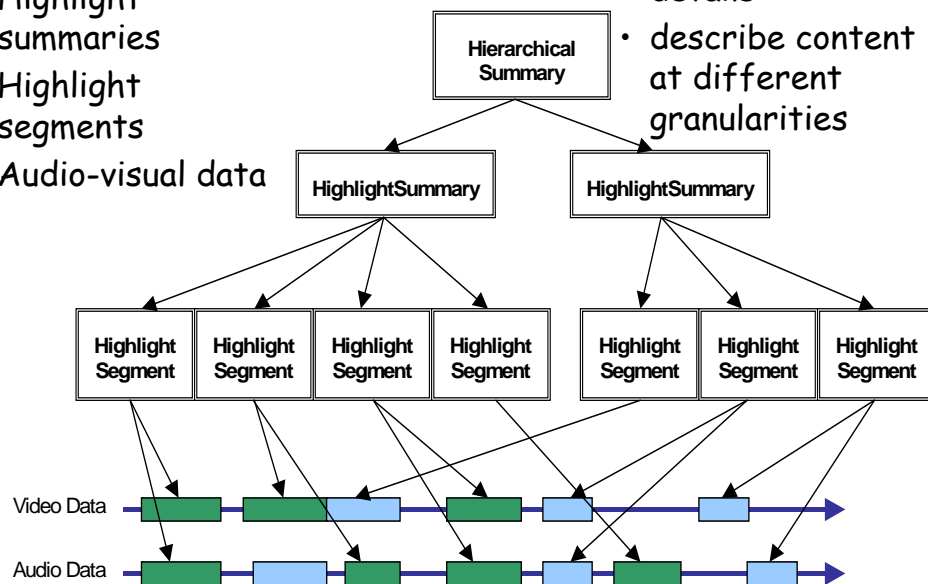


- **Hierarchical Video Summary:** fidelity (i.e., f_0, f_1, \dots) of the key-frames with respect to the video segments referred to by the key-frames at the next lower level

~AXMEDIS 2005~ Tutorial

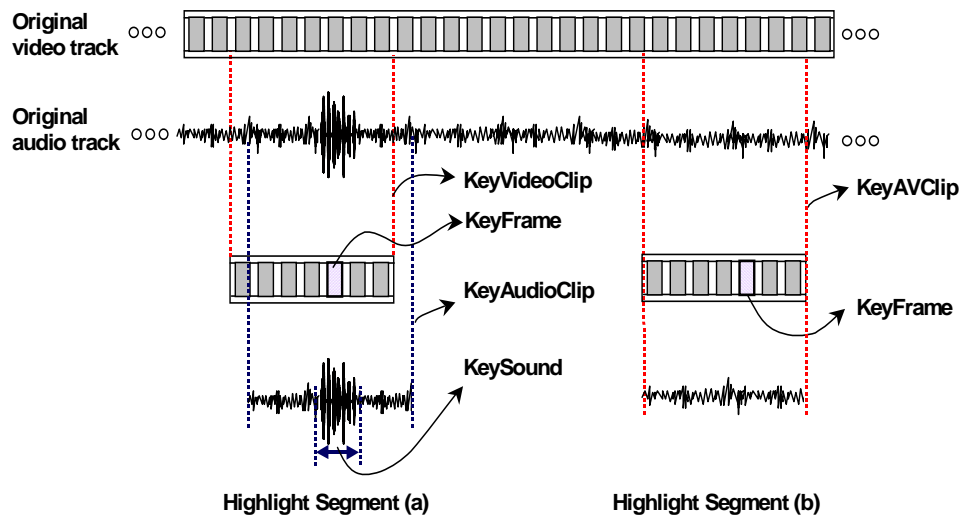
MPEG-7 MDS: Hierarchical Summary DS

- Hierarchical summary
- Highlight summaries
- Highlight segments
- Audio-visual data
- describe different levels of temporal details
- describe content at different granularities



~AXMEDIS 2005~ Tutorial

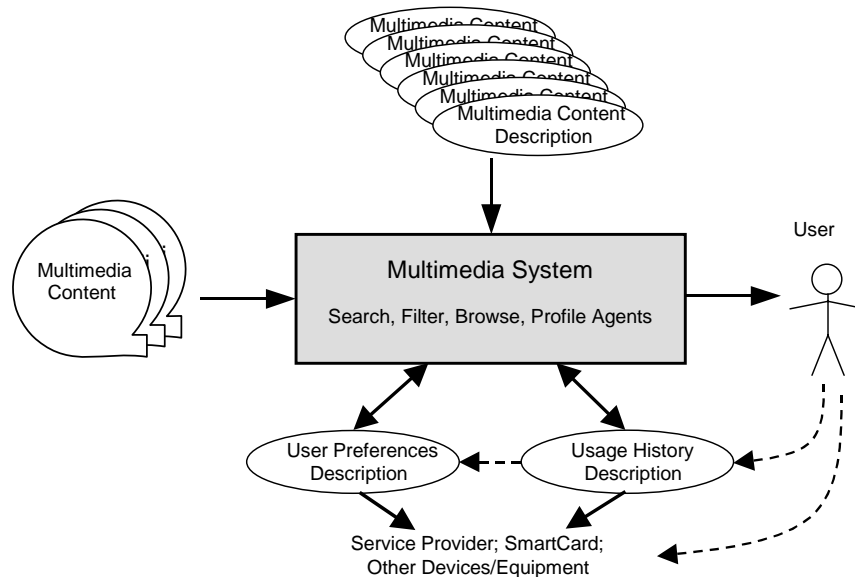
MPEG-7 MDS: Sequential Summary DS



- **Sequential summary:** contains highlight segments
- **Highlight segments:** contains key audio & video clips, images, and sounds

~AXMEDIS 2005~ Tutorial

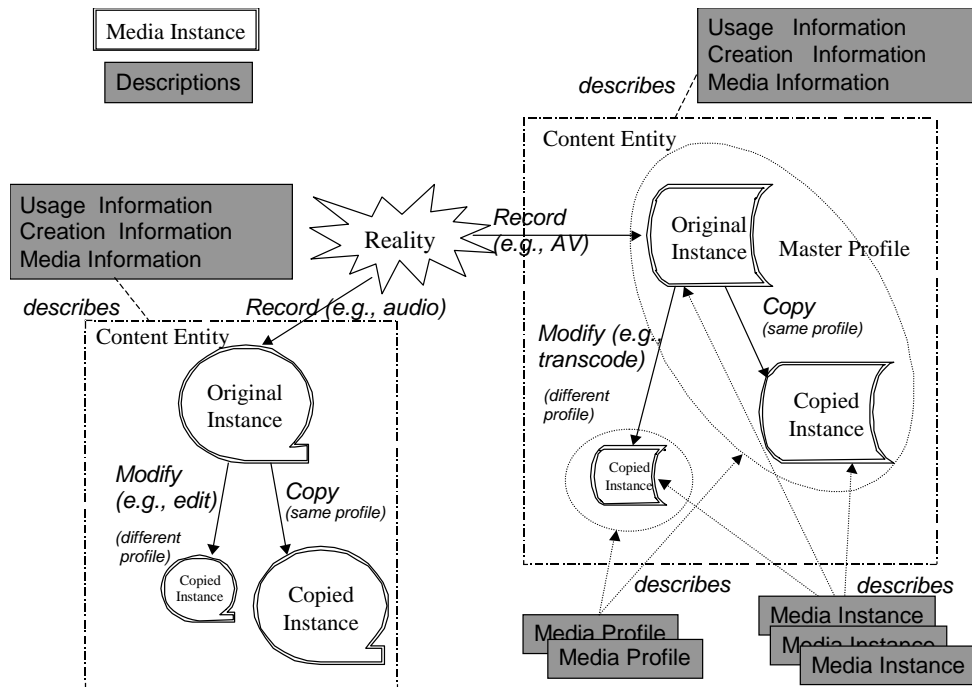
MPEG-7 Personalized filtering, search and browsing of multimedia content



- ➔ need to know more information about the actual media and its semantics
- ➔ in order to meet the user preference and usage history

~AXMEDIS 2005~ Tutorial

Media Instance	Descriptions
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1





 e-business

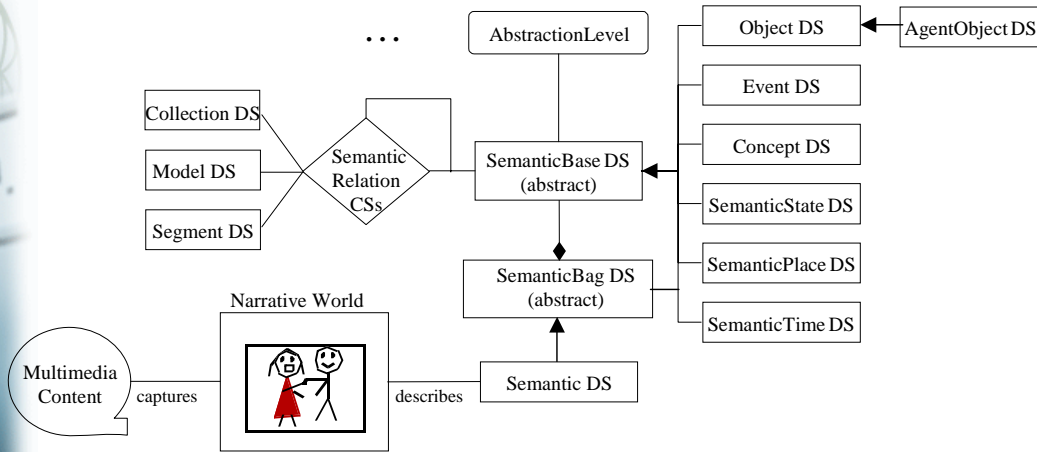




```
<mpeg7>
<Description xai:type="ContentEntityType">
<MultimediaContent xai:type="VideoType">
<Video>
-MediaInformation id="news1_media">
<MediaIdentification>
<EntityIdentifier organization="MPEG" type="MPEGContentSetId">
mpeg7_content:news1
</EntityIdentifier>
<VideoDomain href="urn:mpeg:mpeg7:cs:VideoDomainCS:2001:1.2">
<Name xml:lang="en">Natural</Name>
</VideoDomain>
</MediaIdentification>
-MediaProfile
-MediaFormat
<Content href="MPEG7ContentCS">
<Name>audiovisual</Name>
</Content>
-Medium href="urn:mpeg:mpeg7:cs:MediumCS:2001:1.1">
<Name xml:lang="en">CD</Name>
</Medium>
-FileFormat href="urn:mpeg:mpeg7:cs:FileFormatCS:2001:3">
<Name xml:lang="en">mpeg</Name>
</FileFormat>
<FileSize>666478608</FileSize>
-VisualCoding
<Format
href="urn:mpeg:mpeg7:cs:VisualCodingFormatCS:2001:1"
colorDomain="color">
<Name xml:lang="en">MPEG-1 Video</Name>
</Format>
<PixelFormat>aspectRatio="0.75" bitsPer="8"/>
<Frame height="288" width="352" rate="25"/>
</VisualCoding>
</MediaFormat>
-MediaInstance id="onlinempeg7cs17news1">
<InstanceIdentifier organization="MPEG"
type="MPEG7ContentSetOnlineId">mpeg7/content17/news1
</InstanceIdentifier>
-MediaLocator
<MediaUri>http://www.mpeg.org/mpeg7/contentset/17/news1.mpg</MediaUri>
</MediaLocator>
</MediaInstance>
</MediaProfile>
</MediaInformation>
</Video>
</MultimediaContent>
</Description>
</mpeg7>
```

~AXMEDIS 2005~ Tutorial

MPEG-7 Semantics of Multimedia Content



~AXMEDIS 2005~ Tutorial

MPEG-7 Semantic Type

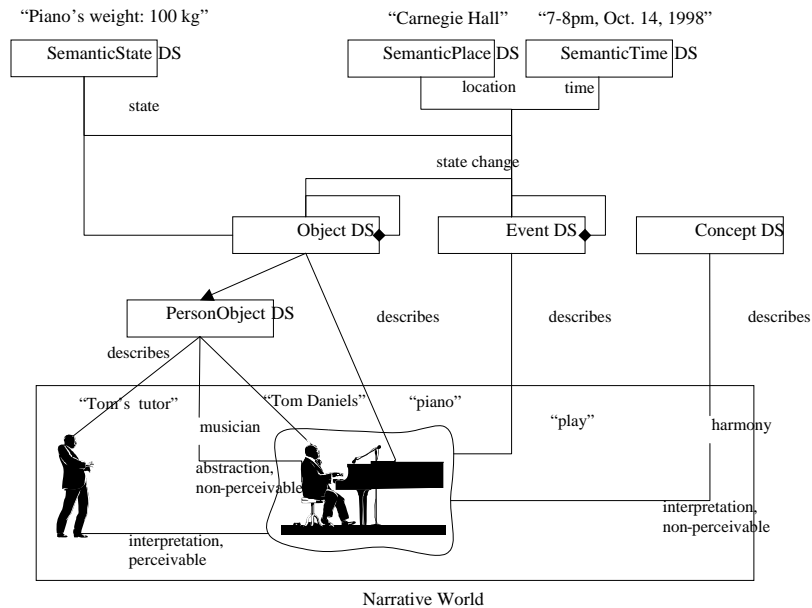
```
<Mpeg7>
  <Description xsi:type="SemanticDescriptionType">
    <Semantics>
      <Label>
        <Name> Concept of freedom </Name>
      </Label>
      <SemanticBase xsi:type="ConceptType">
        <Label>
          <Name> Freedom </Name>
        </Label>
        <Property>
          <Name> Open </Name>
        </Property>
        <Property>
          <Name> Outspoken </Name>
        </Property>
        <Property>
          <Name> Frank </Name>
        </Property>
        <MediaOccurrence type="symbol">
          <MediaLocator>
            <MediaUri> liberty.gif
          </MediaUri>
          </MediaLocator>
        </MediaOccurrence>
      </SemanticBase>
    </Semantics>
  </Description>
</Mpeg7>
```

~AXMEDIS 2005~ Tutorial



e-business

MPEG-7 MDS: Example Semantics Description (Narrative World)



Object: *piano*

People:
musician,
tutor

Event: *playing*
piano

Place:
Carnegie Hall

Time: *1998*

State: *tuning*

Concept:
Harmony

~AXMEDIS 2005~ Tutorial

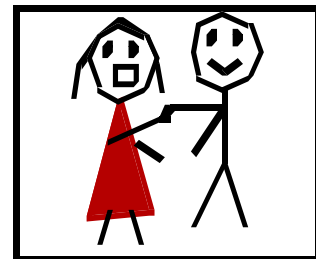


e-business

MPEG-7 MDS: Event DS Example

- The following example gives an MPEG-7 description of the event of handshake between people:






```
<Mpeg7>
  <Description xsi:type="SemanticDescriptionType">
    <Semantics>
      <Label>
        <Name> Shake hands </Name>
      </Label>
      <SemanticBase xsi:type="AgentObjectType" id="A">
        <Label href="urn:example:acs">
          <Name> Person A </Name>
        </Label>
      </SemanticBase>
      <SemanticBase xsi:type="AgentObjectType" id="B">
        <Label href="urn:example:acs">
          <Name> Person B </Name>
        </Label>
      </SemanticBase>
      <SemanticBase xsi:type="EventType">
        <Label href="urn:example:handshake"><Name> Handshake </Name></Label>
        <Definition>
          <FreeTextAnnotation> Clasping of right hands by two people </FreeTextAnnotation>
        </Definition>
        <Relation type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2001:agent" target="#A"/>
        <Relation type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2001:accompanier" target="#B"/>
      </SemanticBase>
    </Semantics>
  </Description>
</Mpeg7>
```



~AXMEDIS 2005~ Tutorial

MPEG-7 Media Resource Transcoding: Signal Level

- **Transcoding Objectives:**
 - Maximize Content Value (eg., quality, SNR), Minimize Transcoding Complexity (eg., delay); meet usage context requirements (eg., terminal and network characteristics)
- **Transcoding Architectures and Algorithms:**
 - **Image Transcoding:** scalable image transcoding (eg., wavelet, DCT), cropping, rescaling
 - **Video Transcoding:** bit-rate, spatio-temporal resolution, quality
 - **Modality Conversion:** Speech \leftrightarrow Text; Video \rightarrow Image; Multimedia \rightarrow Metadata text

Workstation	Color PC	TV browser	HHC	PDA	Smart Phone
					"bridge"
compress: 38 KB	23 KB	8 KB	4 KB	0.6 KB	100 B
size: 256 x 256	192 x 192	128 x 128	96 x 96	64 x 64	-
color: 24 bit RGB	24 bit RGB	256 colors	4 bit gray	B/W	-

MPEG-7 Transcoding Hints

- **Importance** - specifies the relative importance of segments, regions, objects, or audio-visual programs
- **Spatial resolution hint** - specifies the maximum allowable spatial resolution reduction factor for perceptibility
- **Shape hint** - specifies the amount of shape change in the media.
- **Difficulty hint** - specifies the transcoding difficulty of the media.
- **Motion hints** - specifies motion uncompensability and motion intensity information
 - Motion uncompensability specifies the amount of new content in a segment or region
 - Motion intensity specifies the motion intensity of a segment or region

MPEG-7 Transcoding Hints



```
<StillRegion>
<SpatialGeometry>
<BoundingBox>
<ValueVector> 76 </ValueVector>
<ValueVector> 220 </ValueVector>
<ValueVector> 114 </ValueVector>
<ValueVector> 40 </ValueVector>
</BoundingBox>
</SpatialGeometry>
<SegmentHint importance="1.0"/>
</StillRegion>
```

Transcoding Hints:

- Importance
- Resolution
- MPEG coding parameters: motion, shape

Related information:

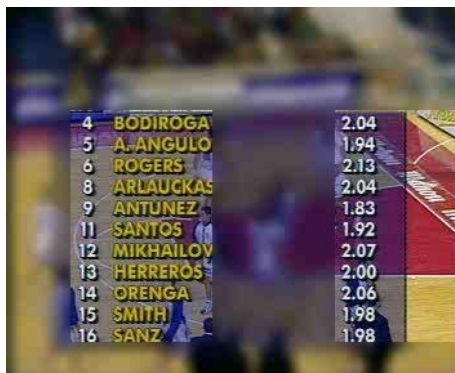
- Image type
- Regions
- Objects
- Semantics



Image/region	Transcoding Hint (dB)	Without (dB)
Face	36.74	22.33
Text	32.02	22.69
Image	17.62 (8 kB)	28.48 (8 kB)

MPEG-7 Transcoding Hints - Transcoding Example

Transcoding using hints (regional)



Transcoding w/o hints (global)



Image/region	Segment Hint (dB)	Without (dB)
Text 1	32.54	23.76
Text 2	34.93	24.92
Image	15.11 (12 kB)	22.92 (11 kB)

MPEG-7 Transcoding Hints - Transcoding Example

Transcoding using hints (regional)



Transcoding w/o hints (global)



Image/region	Segment Hint (dB)	Without (dB)
Face	38.40	31.21
Text	35.52	26.04
Image	17.51 (8 kB)	28.94 (7 kB)





MPEG-7 Image Transcoding as Optimization Problem



- **Constraint optimization:**
 - Constraint on image size from device
 - Benefit gained from content value of transcoded image
- **Problem statement:**
 - Maximize total content value given constraints

- **Region metadata annotation (R_i):**
 - **Importance Hint (I_i):** indicates relative importance of region ($0 < I_i < 1$)
 - **Spatial Resolution Hint (S_i):** indicates minimum resolution of region for preserving details ($0 < S_i < 1$)
 - **Content Value (V_i):** indicates value of the transcoded region $V_i = f(I_i, S_i)$





Example: Image Transcoding (w/o Transcoding Hints)

			
1280 x 960	640 x 480	320 x 240	160 x 120
PC (SVGA)	TV Browser	Handheld	PDA/Phone

Global Image Adaptation:

- Global image size reduction adapts to device screen size
- Loses details of important regions as all regions are reduced equally
- Results in lower content value for given data size

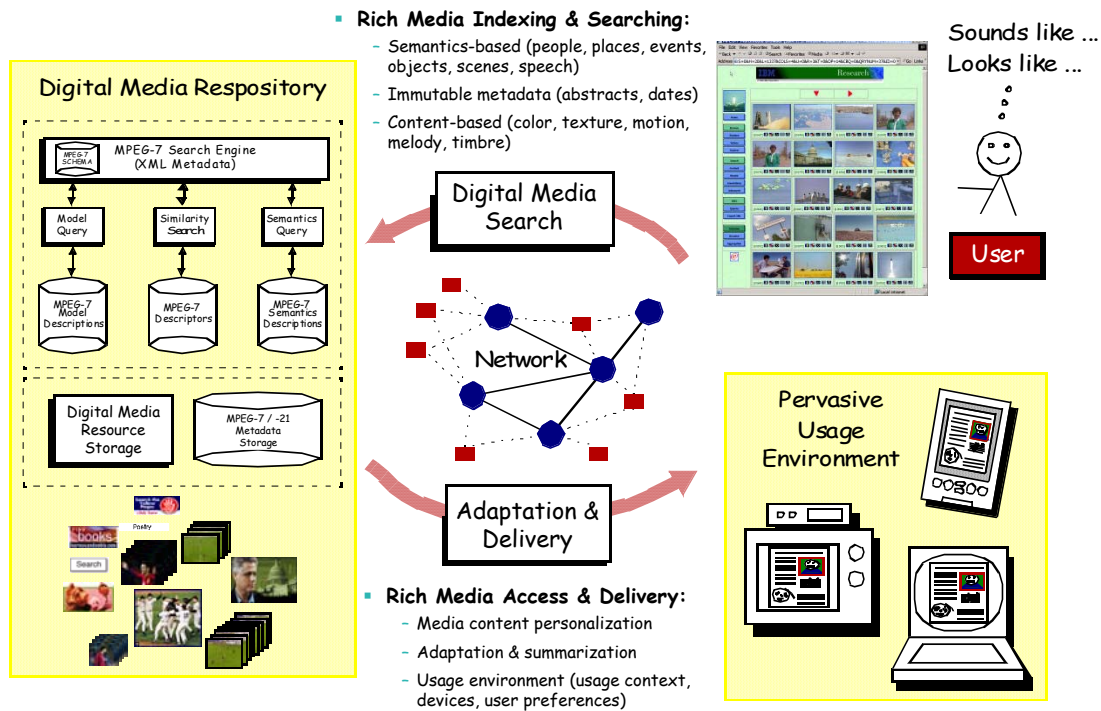
Example: Image Transcoding (with Transcoding Hints)

			
1280 x 960	640 x 480	320 x 240	160 x 120
PC (SVGA)	TV Browser	Handheld	PDA/Phone

MPEG-7 Transcoding Hint-based Adaptation:

- Combination of cropping and scale reduction to adapt to device screen size
- Preserves details of important regions (cost-based optimization)
- Results in higher content value for given data size

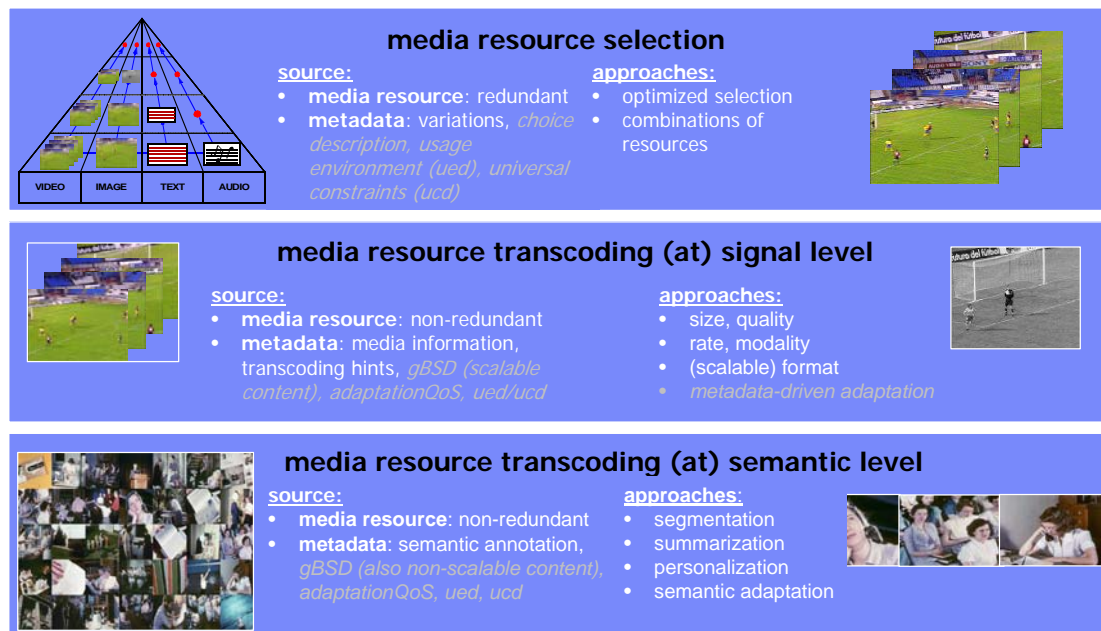
Digital Media Content Access and Distribution



35

UMA, MPEG-7, MPEG-21 | 1st Int'l. Conf. on Automated Production of Cross Media Content for Multi-channel Distribution ~AXMEDIS 2005~ | December 1, 2005 © 2004 IBM Corporation

multimedia content adaptation: approaches



36

UMA, MPEG-7, MPEG-21 | 1st Int'l. Conf. on Automated Production of Cross Media Content for Multi-channel Distribution ~AXMEDIS 2005~ | December 1, 2005 © 2004 IBM Corporation

Research Topic: Multi-Transcoding

