The First International Conference on Advances in Multimedia
MMEDIA 2009
July 20-25, 2009 - Colmar, France

MULTIMEDIA STANDARDS

Peter L. Stanchev

Kettering University, Flint, Michigan 48504, USA

Institute of Mathematics and Informatics, BAS, Bulgaria

pstanche@kettering.edu

Kettering University is nationally ranked in the Top Five in two categories and in the Top Twenty overall in the United States.

INSTITUTE OF MATHEMATICS AND INFORMATICS

INFORMATION SYSTEMS DEPARTMENT

Research:

To advance peoples knowledge about the efficient and effective utilization of information technology through the way of individuals, organizations, society, and nations for the improvement of economic and social welfare.

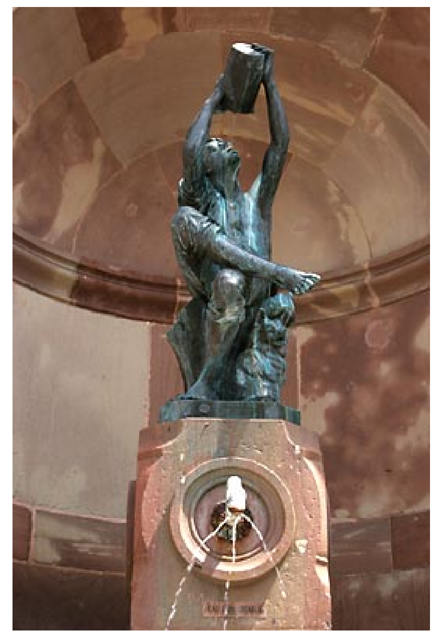
Education:

To develop information system leaders for the 21st century who will not only acquire the skills and knowledge necessary for their next job but also the tools for building a life-long successful career.

Citizenship:

To contribute significantly to developing and sustaining democratic and knowledgeable schools, communities, and societies.

http://www.youtube.com/watch?v=ENmjlpuzX4Y



THE TUTORIAL OBJECTIVES

- Introduction
- MPEG (Moving Picture Experts Group)
  - MPEG-1, MPEG-2, MPEG-4
  - MPEG-7 Multimedia Content Description Interface
  - MPEG-21 Multimedia Framework
- Standards for Scene Description
- Open Standards for interactive TV (iTV)
- Standards in the Audiovisual Media
- Web 2.0 Tools
- Semantic Web Technology
- Conclusions

MULTIMEDIA

Colmar was founded in the 9th century. This was the location where Charles the Fat held a diet in 884. Colmar was granted the status of a free imperial city of the Holy Roman Empire in 1226. During the Thirty Years' War, the city was taken by the armies of Sweden in 1632, who held it for two years. The city was conquered by France under Louis XIV in 1697.
In 1679 (Treaties of Nijmegen) Colmar was ceded to France. With the rest of Alsace, Colmar was annexed by the newly formed German Empire in 1871 as a result of the Franco-Prussian War. It returned to France after World War I, was annexed by Nazi Germany in 1940, and then reverted to French control after the battle of the "Colmar Pocket" in 1945.
The Colmar Treasure, hidden during the Black Death, was discovered here in 1863.

STANDARDS

- **JPEG (**Joint Photographic Experts Group)
- **MPEG (**Moving Pictures Expert Group)
- **MHEG (**Multimedia and Hypermedia Information Coding Experts Group)
- **SGML** (Standard Generalized Markup Language)
- **DC** (Dublin Core)
- **RDF** (Resource Definition Framework)

MPEG

In 1988, the International Standards Organization (ISO) formed the Motion Picture Experts Group (MPEG) with the primary goal of developing coding techniques to achieve good quality audio and video with the compact disc as the target application. The objective was an open, timely and interoperable coding standard which at a low cost implemented leading technology and had the potential for performance improvement even after the standard became official.

MPEG

- Through the years, the scope of the work evolved to include a standard for high-quality, efficient coded digital TV (MPEG-2), a standard for multimedia applications (MPEG-4) and a standard for a multimedia content description interface (MPEG-7), multimedia framework (MPEG-21)
- In 1996, MPEG was presented with the Emmy award for the MPEG-1 and MPEG-2 standards.
- The standards only standardize the format for representing data input to decoders and a set of rules for how to interpret this data. The specific encoding methods are not standardized.

MPEG-1

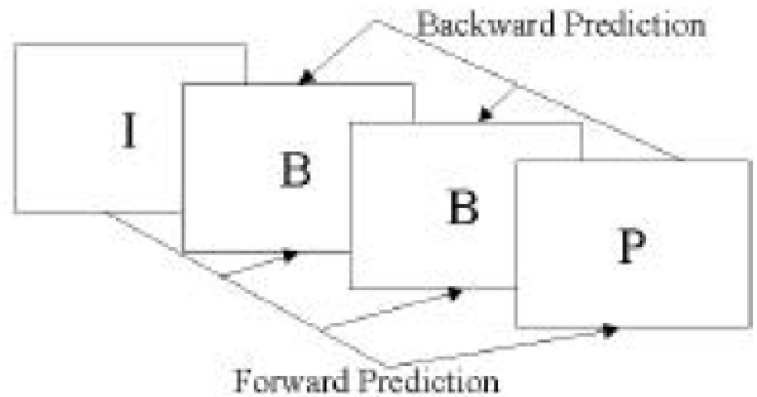
The original target for the MPEG-1 standard was good-quality audio and video at a total bit rate of about 1.4 Mbit/sec. Requirements derived from the target application were: coding of good-quality video from 1 to 1.5 Mbit/sec, random access to frames in limited time (i.e. frequent access points), fast forward and reverse capabilities, synchronized playback of audiovisual data and the standard to be possible to implement in real-time decoders at a reasonable cost.

MPEG-1

- The MPEG-1 Systems standard defines how to multiplex audio and video packets into a single, synchronized stream. The Systems Time Clock (STC) operates at 90 kHz and is the reference time base.
- The MPEG-1 Video standard specifies the elementary video bit stream syntax and the video decoding process. Originally, MPEG-1 Video supported bit rates of about 1.2 Mbit/sec. However, to avoid the restricting of future applications the syntax was made flexible to support image sizes up to 4096 x 4096, many frame rates (23.97, 24, 25, 29.97, 50, 59.94, and 60 frames/sec) and higher bit rates.

MPEG-1

There are three kinds of frame or picture coding supported in the standard. Intra (I) pictures are coded separately by themselves, predictive (P) pictures are coded with respect to the previous I or P-picture, and bidirectionally predictive (B) pictures are coded with respect to the previous and next I or P-picture. B-pictures are the least expensive in respect of compression efficiency and I-pictures the most expensive.



MPEG-1

MPEG-1 uses both spatial and temporal redundancies when compressing the video signal. Coding each frame separately using block Discrete Cosine Transforms (DCT), quantizing the resulting 8 x 8 blocks of DCT coefficients, zigzag scanning and variable length coding utilizes the spatial redundancy. Frames representing cuts in the scene are typically coded as I-pictures. The MPEG-1 Audio standard specifies the elementary audio bitstream syntax and the audio decoding process.

MPEG-1

- The standard consists of three layers (I, II and III) that represent increasing complexity and coding efficiency. The popular MP3 music format is actually not MPEG-3 (which does not exist) but MPEG-1 layer III. Available sampling rates are 32, 44.1 or 48 kHz. In addition, layer III supports variable bit rate coding.
- MPEG-1 Audio can be either mono, stereo, dual mono with two separate channels or joint stereo using interchannel redundancies.

MPEG-1

In November 1992 MPEG-1 became an international standard for video and audio compression. It produces video recorder-quality output that can be transmitted over twisted pair transmission lines for modest distances. MPEG-1 is also used for storing movies on CD-ROMs.

MPEG-2

- MPEG-2 is targeted at broadcasting and CD-ROM applications and supports high resolution and good quality such as high-definition TV (HDTV). Like MPEG-1, the standard is aimed at compressing data. Requirements derived from the target applications included very good quality, high-resolution video at 4 to 15 Mbit/sec, multi-channel audio with high quality, and backward compatibility with MPEG-1.
- It has improved error resilience compared to that of MPEG-1, which was aimed at audiovisual applications with very low error rates.

MPEG-2

- MPEG-2 includes two types of streams: the program stream, which corresponds to a modified MPEG-1 Systems stream, and the transport stream, which offers robustness to errors occurring in noisy environments and the ability to include multiple programs in a single stream.
- MPEG-2 can thus be said to contain two different standards in one with different algorithms for encoding, packaging, multiplexing and decoding the data.

MPEG-2

The MPEG-2 Video standard can be viewed as a superset of the MPEG-1 Video standard. It supports the same frame rates as MPEG-1 Video but picture sizes up to 16,384. The codpictures containing I, P and B-pictures. The pictures are then divided into slices, which in turn are divided into macro blocks. The macro blocks are divided into luminance and chrominance blocks. To provide both forward and backward compatibility with MPEG-1 is done much like with MPEG-1 video but it also supports interlaced video. It is still based on motion compensated block DCT coding.

MPEG-2

- MPEG-2 Audio output consists of five channels: Left (L), Center (C), Right (R), Left surround (Ls) and Right surround (Rs). Ls and Rs loudspeakers are typically placed in the back left and right, respectively.
- In addition, multiple channels can be used to provide multilingual systems, the director's comments etc.
- MPEG-2 became an international standard in 1996

MPEG-4

- MPEG-4 was originally aimed at medium-resolution video conferencing over low bandwidths at low frame rates but was later modified to define coding of audiovisual objects in applications such as Internet video, virtual reality games, interactive shopping and media databases. The main functionalities addressed are the ability to interact with objects in scenes, universal accessibility in terms of storage and transmission media, and improved compression compared to earlier standards.

MPEG-4

MPEG-4 has three abstraction layers.

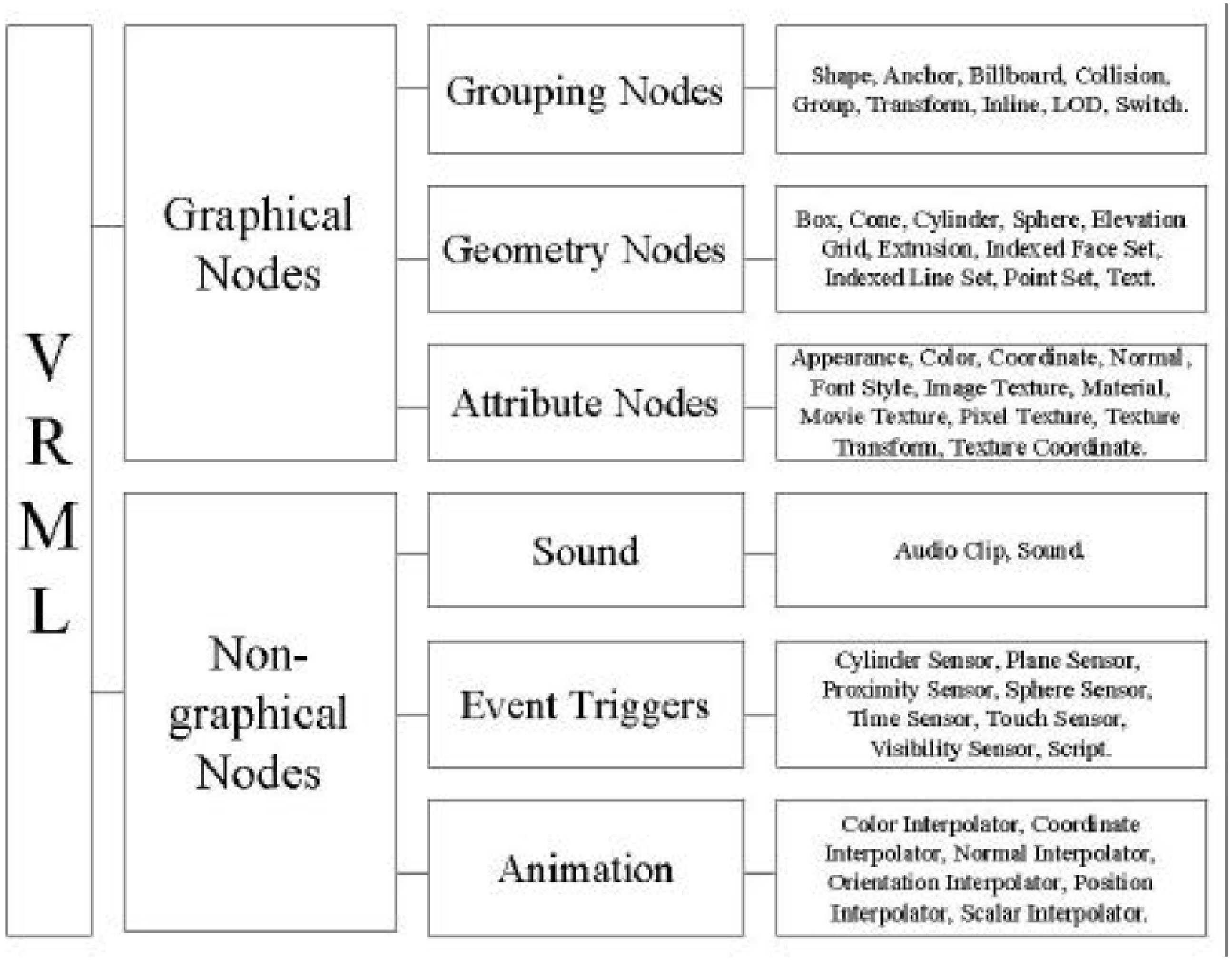
The delivery layer allows transparent access to content independently on how it is stored or delivered;

the sync layer manages the elementary streams;

the compression layer handles the media encoding and decoding.

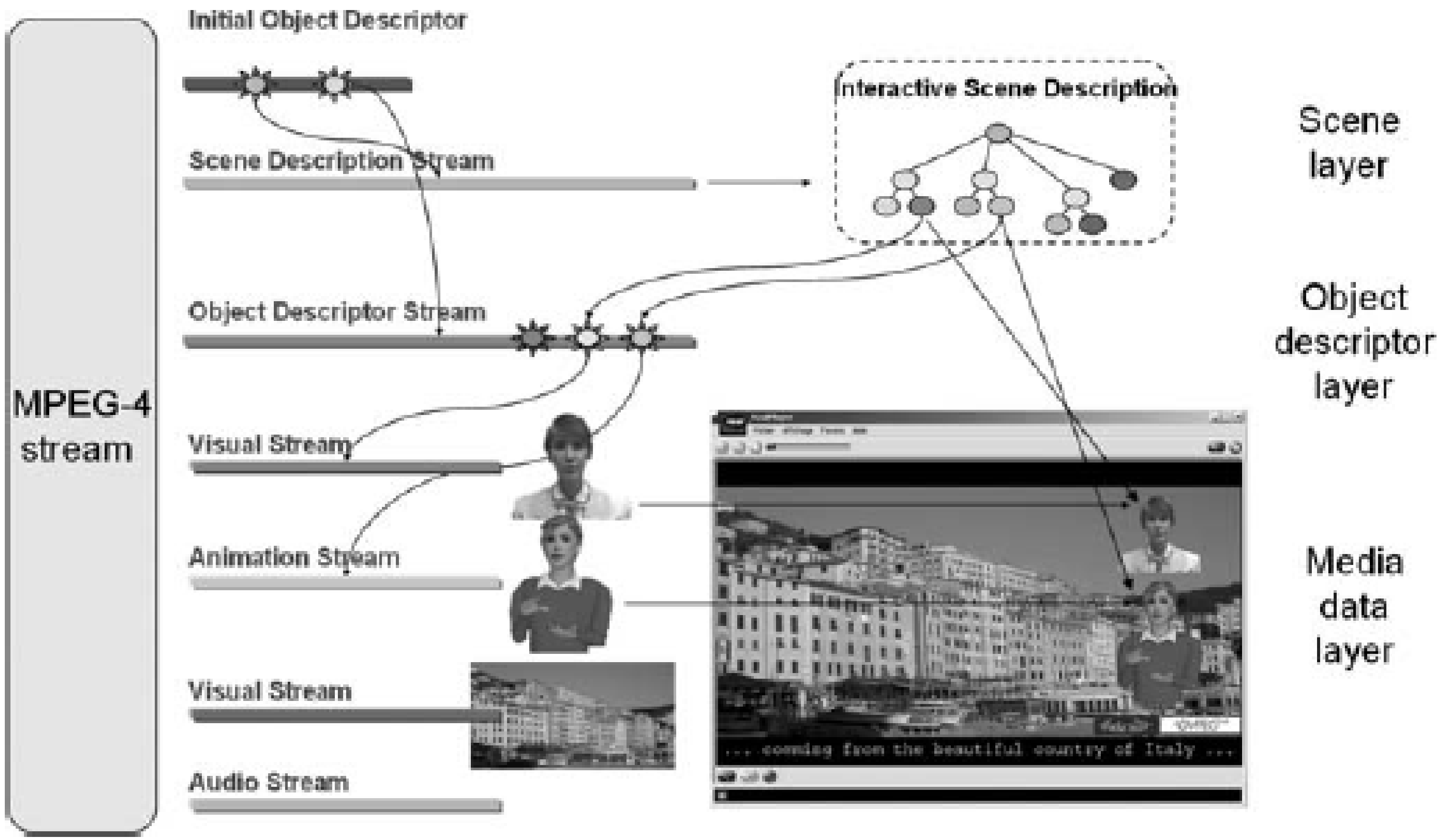
Compression Layer	MPEG-4 Audio MPEG-4 Visual	Media aware Delivery unaware
Sync Layer	MPEG-4 Systems	Media unaware Delivery unaware
Delivery Layer	MPEG-4 DMIF	Media unaware Delivery Aware

MPEG-4



VRML
Graphical Nodes
Grouping Nodes
Shape, Anchor, Billboard, Collision, Group, Transform, Inline, LOD, Switch.
Geometry Nodes
Box, Cone, Cylinder, Sphere, Elevation Grid, Extrusion, Indexed Face Set, Indexed Line Set, Point Set, Text.
Attribute Nodes
Appearance, Color, Coordinate, Normal, Font Style, Image Texture, Material, Movie Texture, Pixel Texture, Texture Transform, Texture Coordinate.
Non-graphical Nodes
Sound
Audio Clip, Sound.
Event Triggers
Cylinder Sensor, Plane Sensor, Proximity Sensor, Sphere Sensor, Time Sensor, Touch Sensor, Visibility Sensor, Script.
Animation
Color Interpolator, Coordinate Interpolator, Normal Interpolator, Orientation Interpolator, Position Interpolator, Scalar Interpolator.

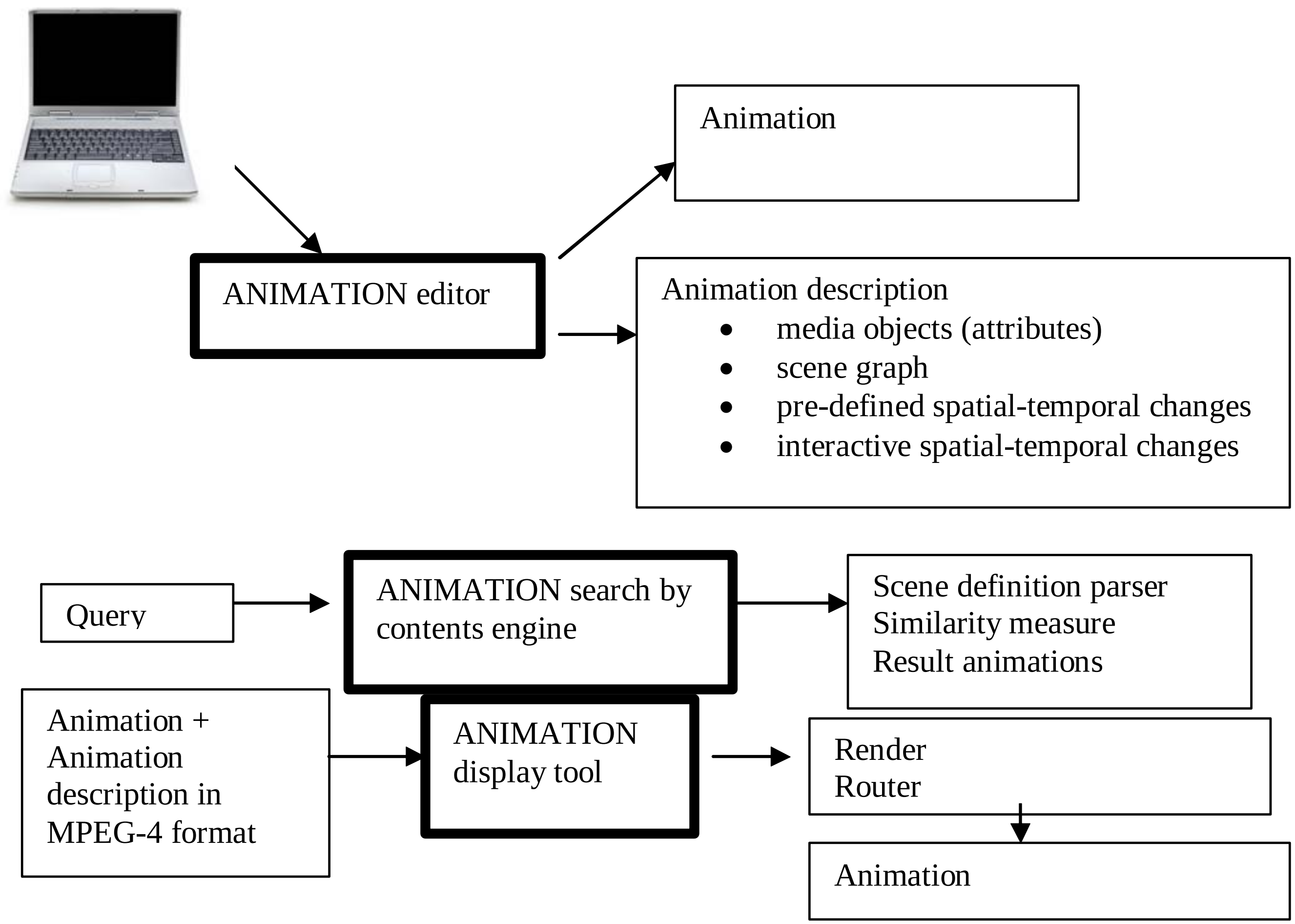
MPEG-4 TERMINAL DATA FLOW



MPEG-4

- MPEG-4 Audio consists of audio streams related to each other by mixing, effect processing, switching, and setting of their 3-D spatial locations. The MPEG-4 Audio standard includes low bit rate audio coding tools for speech, high-quality audio coding tools for music, and synthetic audio coding tools for text-to-speech functions and Structured Audio
- The MPEG-4 content can be delivered across channels from low bit rate wireless to high-speed ATM or from DVDs.
- MPEG-4 became an international standard in May 1999.

MPEG-4 EXAMPLE.
THE ANIMATION SYSTEM



MPEG-4 EXAMPLE.
DESCRIPTION 2D SCENE NODES

Appearance {exposedFieldSFNodematerialNULL exposedFieldSFNodetextureNULL exposedFieldSFNodetextureTransformNULL}

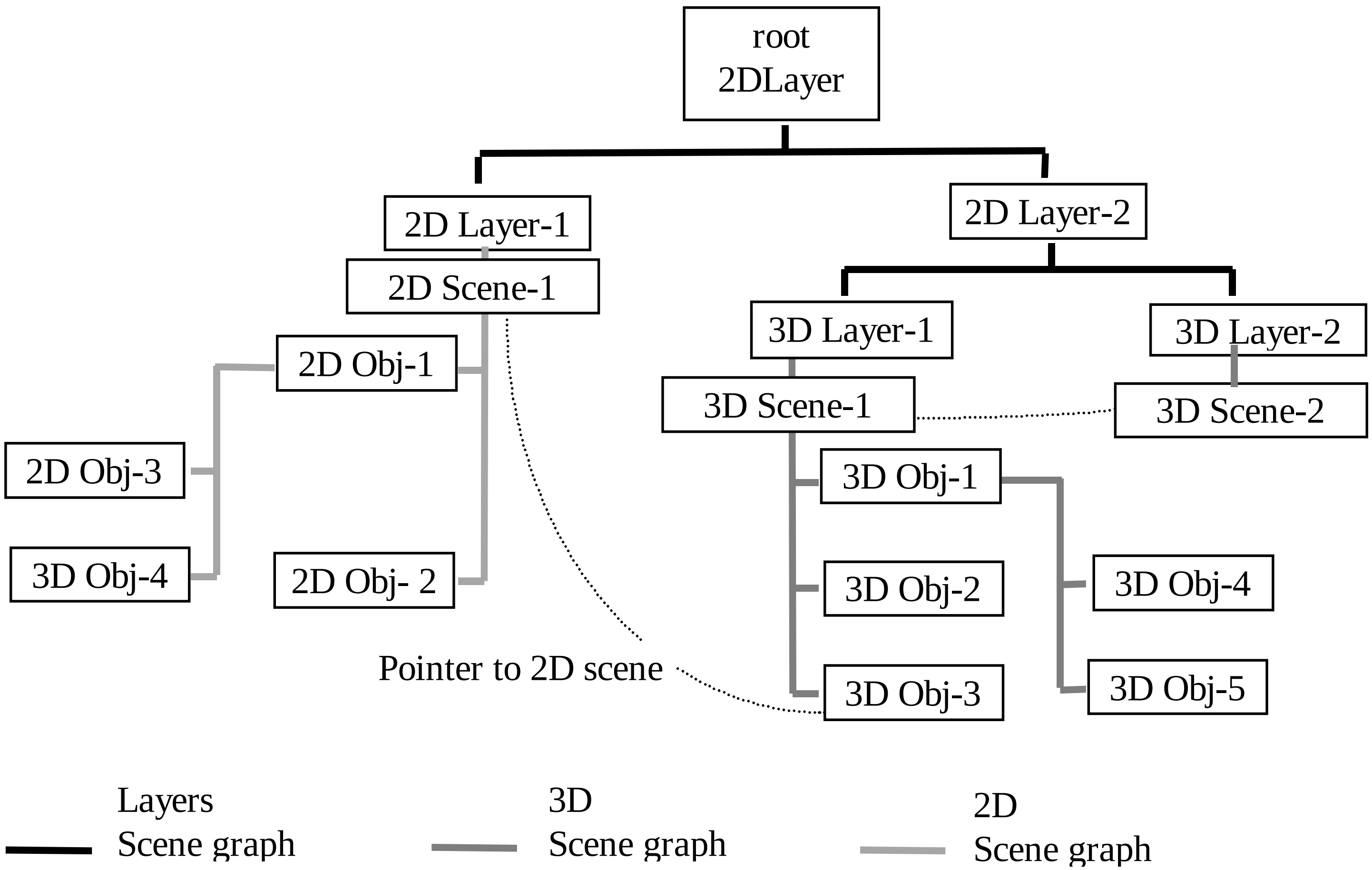
Background2D {EventInSFBoolset_bindNULL ExposedFieldMFStringurlNULL EventOutSFBoolisBound}

Color {ExposedFieldMFColorcolorNULL}

Coordinate2D {ExposedFieldMFVec2fPointNULL}

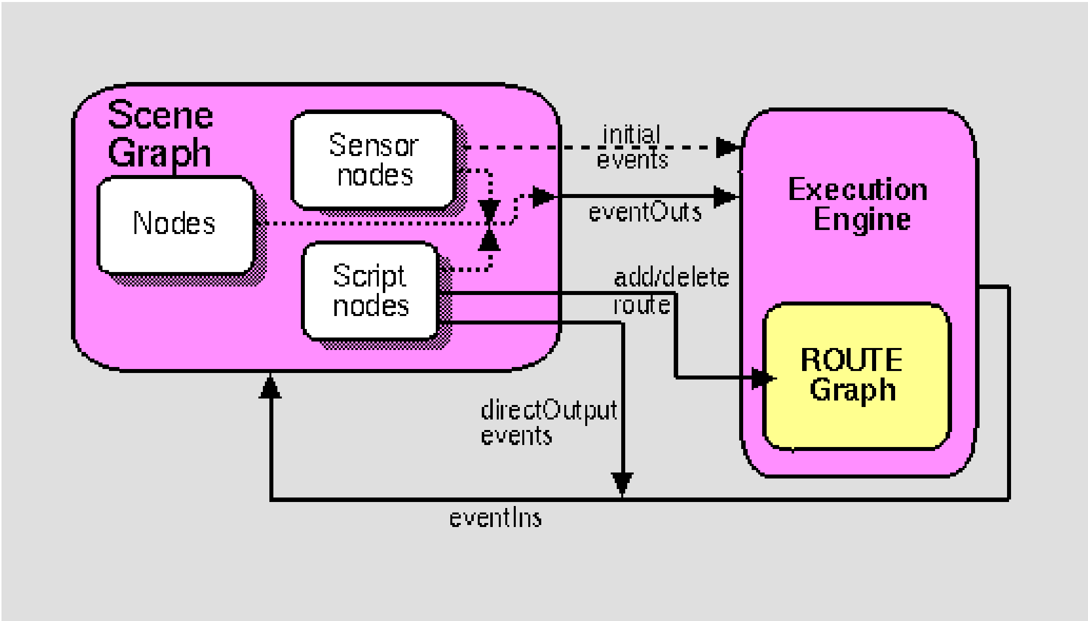
LineProperties {exposedFieldSFColorlineColor0, 0,0 exposedFieldSFInt32lineStyle0 exposedFieldSFFloat width1 }

MPEG-4 EXAMPLE

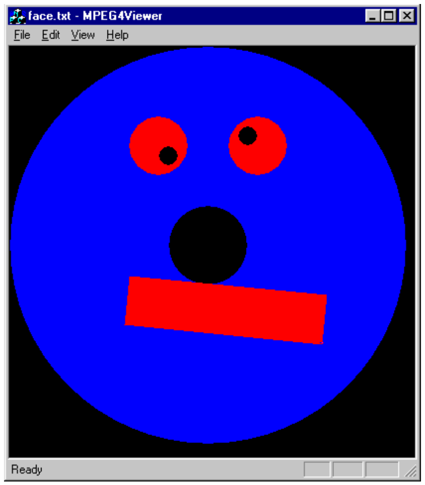


MPEG-4 EXAMPLE.
THE ANIMATION SYSTEM

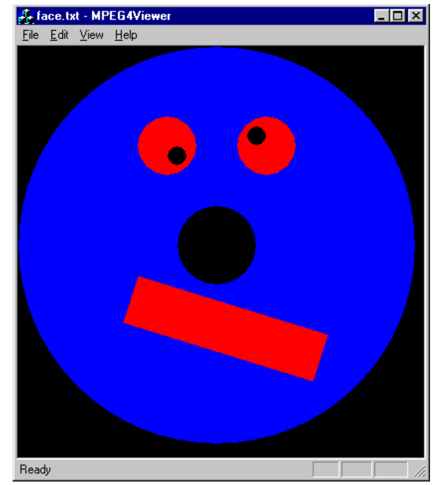
- Render
- Router
- Execution model

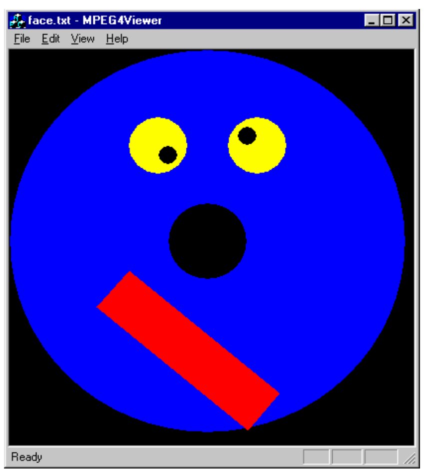


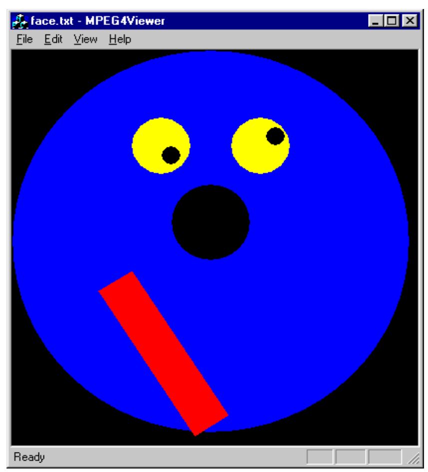
MPEG-4 EXAMPLE.
THE ANIMATION SYSTEM

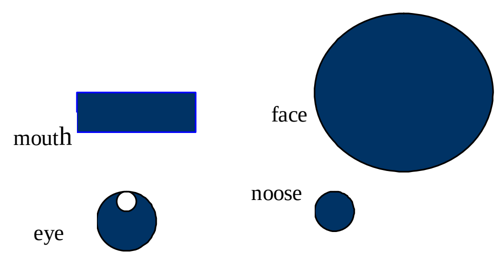


[viewer]









MPEG-7. DRIVING PRINCIPLES

- Wide application base
- Relation with content
- Wide array of data types
- Media independence
- Object-based
- Format independent
- Abstraction level
- Extensibility

MPEG-7 MAIN ELEMENTS

Description Tools:

- Descriptors (D), that define the syntax and the semantics of each feature (metadata element); and
- Description Schemes (DS), that specify the structure and semantics of the relationships between their components, that may be both Descriptors and Description Schemes

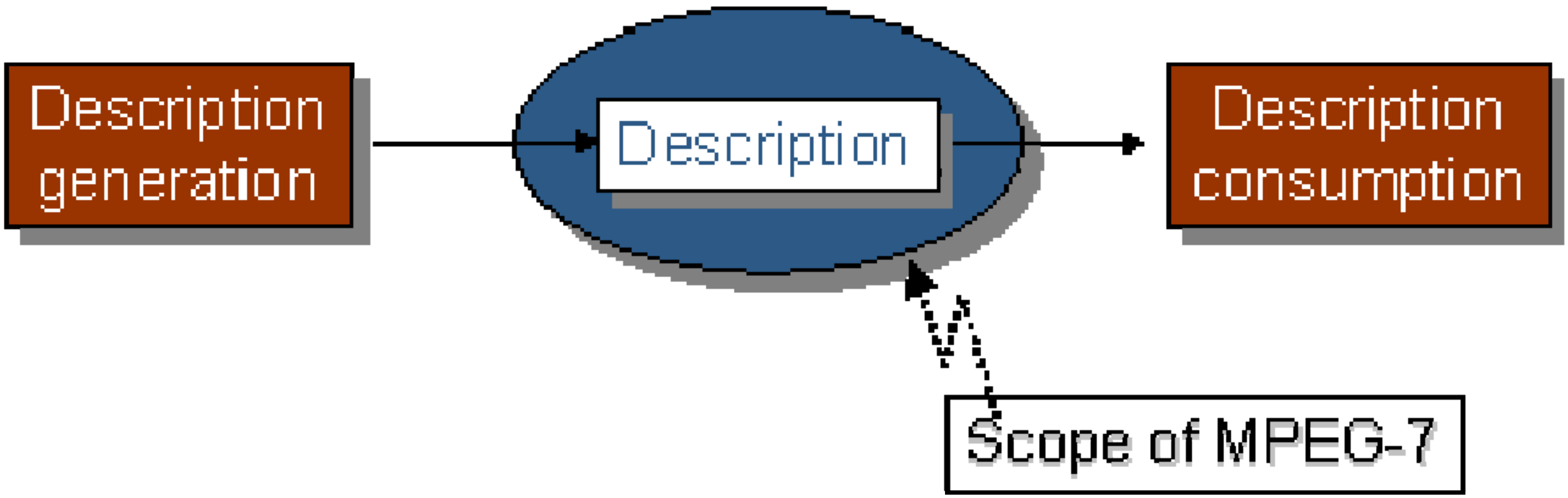
THE MPEG-7 DESCRIPTIONS OF CONTENT

- Structural information on spatial, temporal or spatial-temporal components of the content (scene cuts, segmentation in regions, region motion tracking).
- Information about low level features in the content (colors, textures, sound timbres, melody description).
- Conceptual information of the reality captured by the content (objects and events, interactions among objects).

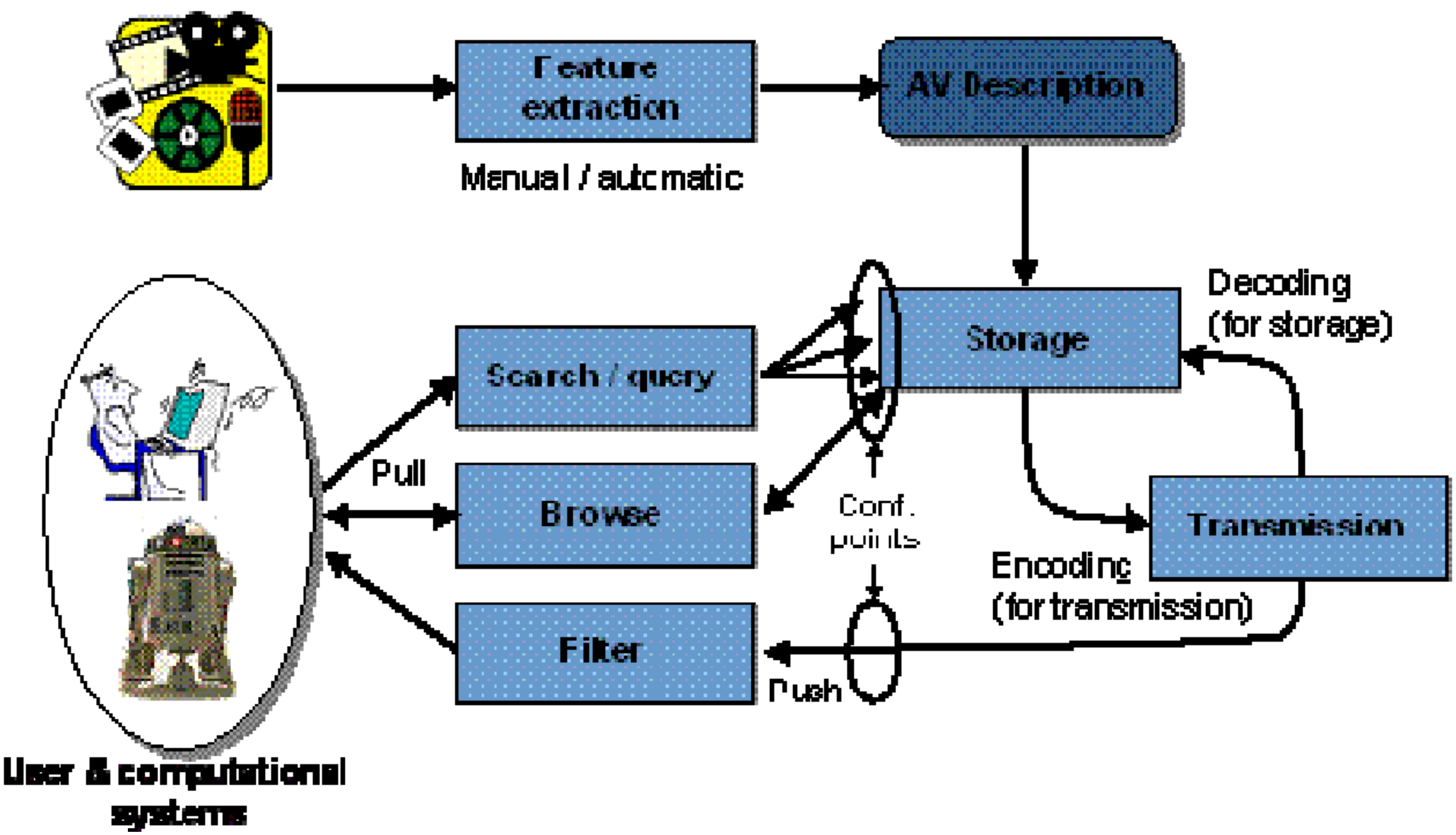
THE MPEG-7 DESCRIPTIONS OF CONTENT

- Information about how to browse the content in an efficient way (summaries, variations, spatial and frequency subbands).
- Information about collections of objects.
- Information about the interaction of the user with the content (user preferences, usage history).

SCOPE OF MPEG-7



ABSTRACT REPRESENTATION OF POSSIBLE APPLICATIONS USING MPEG-7



Schema:

```
<complexType name=”VideoDoc">
   <element name=“Title”…./>
   <element name=“Producer”…./>
   <element name=“Date”…./>
</complexType>
<complexType name=”NewsDoc" base=”VideoDoc" derivedBy="extension">
   <element name=“Broadcaster”…. maxOccurs=“0”/>
   <element name=“Time”….   maxOccurs=“0”/>
<element name=”VideoCatalogue">
   <complexType>
      <element name="CatalogueEntry" minOccurs="0" maxOccurs="*" type=”VideoDoc"/>
   </complexType>
```

This permits VideoDoc elements, as well as types derived from VideoDoc to be used as a child of VideoCatalogue, e.g., NewsDoc

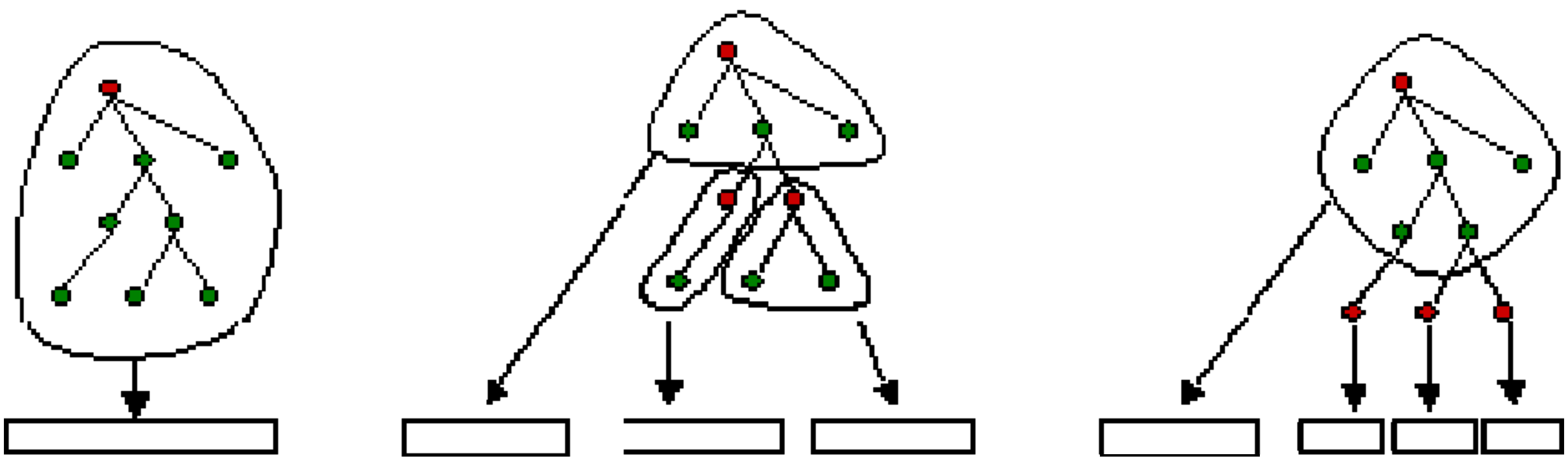
Instance doc:

```
<CatalogueEntry xsi:type=”NewsDoc">
            <Title>”CNN 6 oclock News” </Title>
            <Producerr>David James</Author>
            <Date>2004</Date>
            <Broadcaster>CNN</Channel>
</CatalogueEntry>
```

BINARY FORMAT FOR MPEG-7

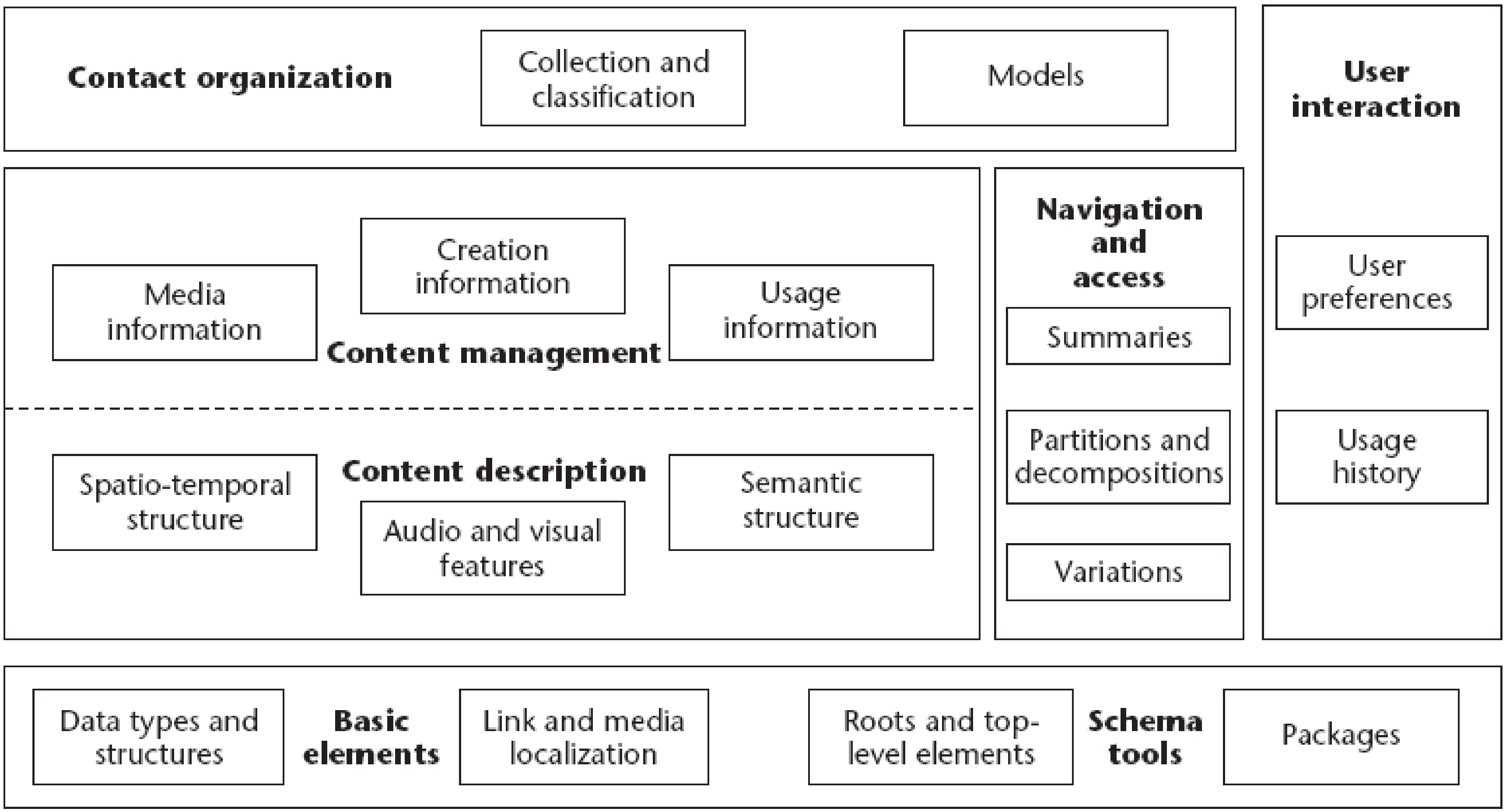
ShemaID
Context mode
Context path

Length
Substitution Code
Type Code
Attributes
Content

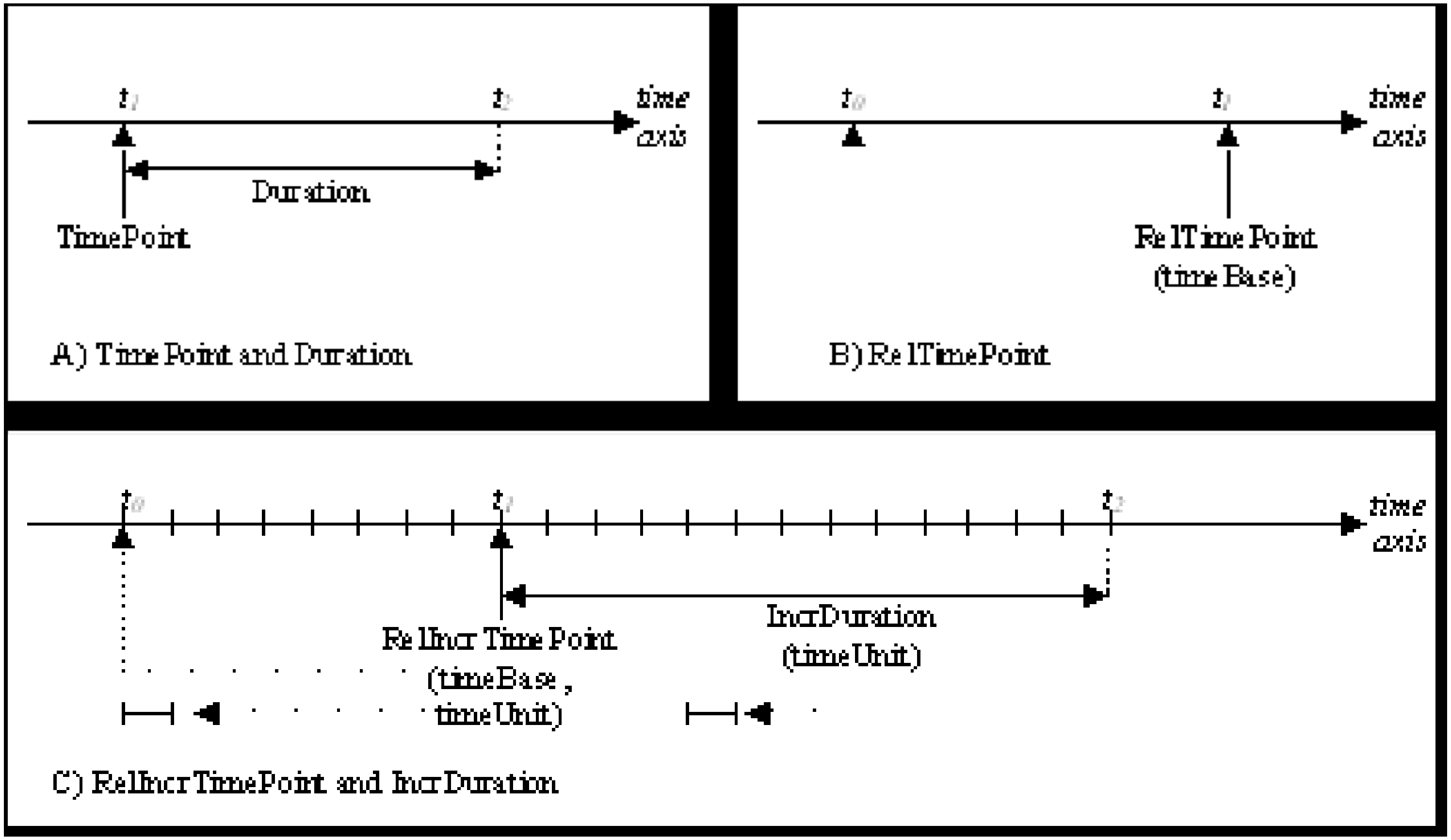


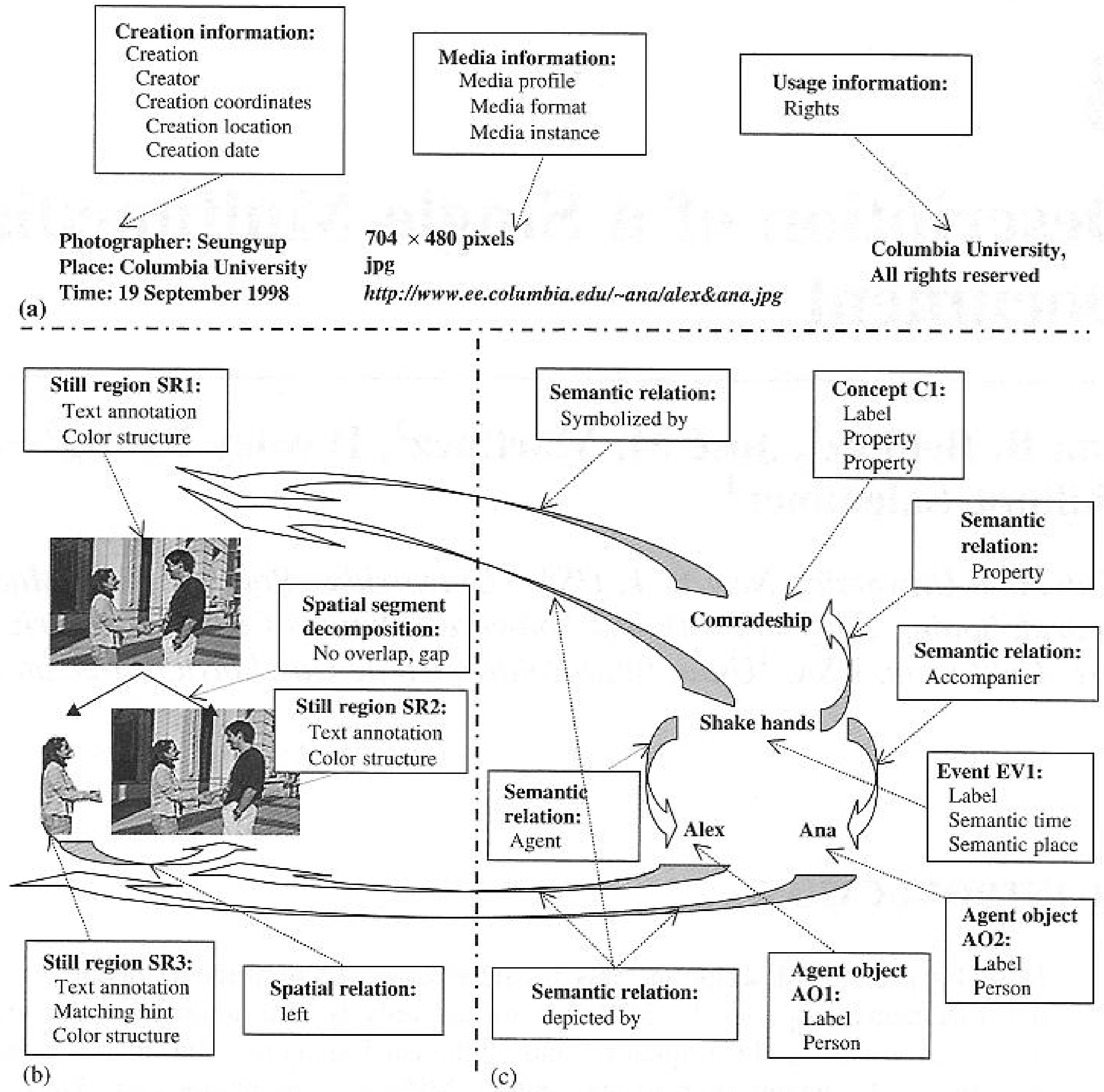
Different streaming strategies of the same XML file

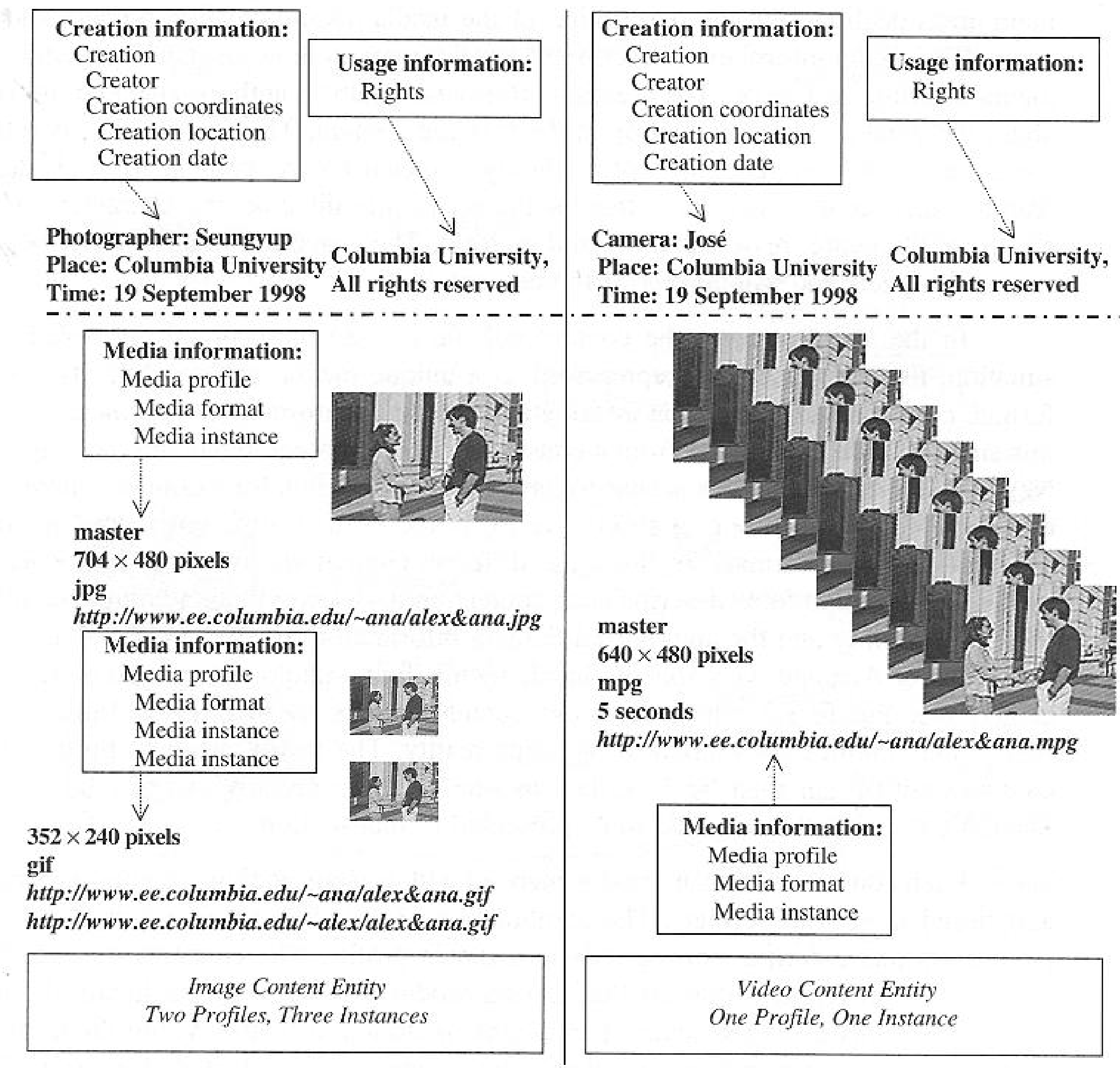
MULTIMEDIA DESCRIPTION SCHEMES



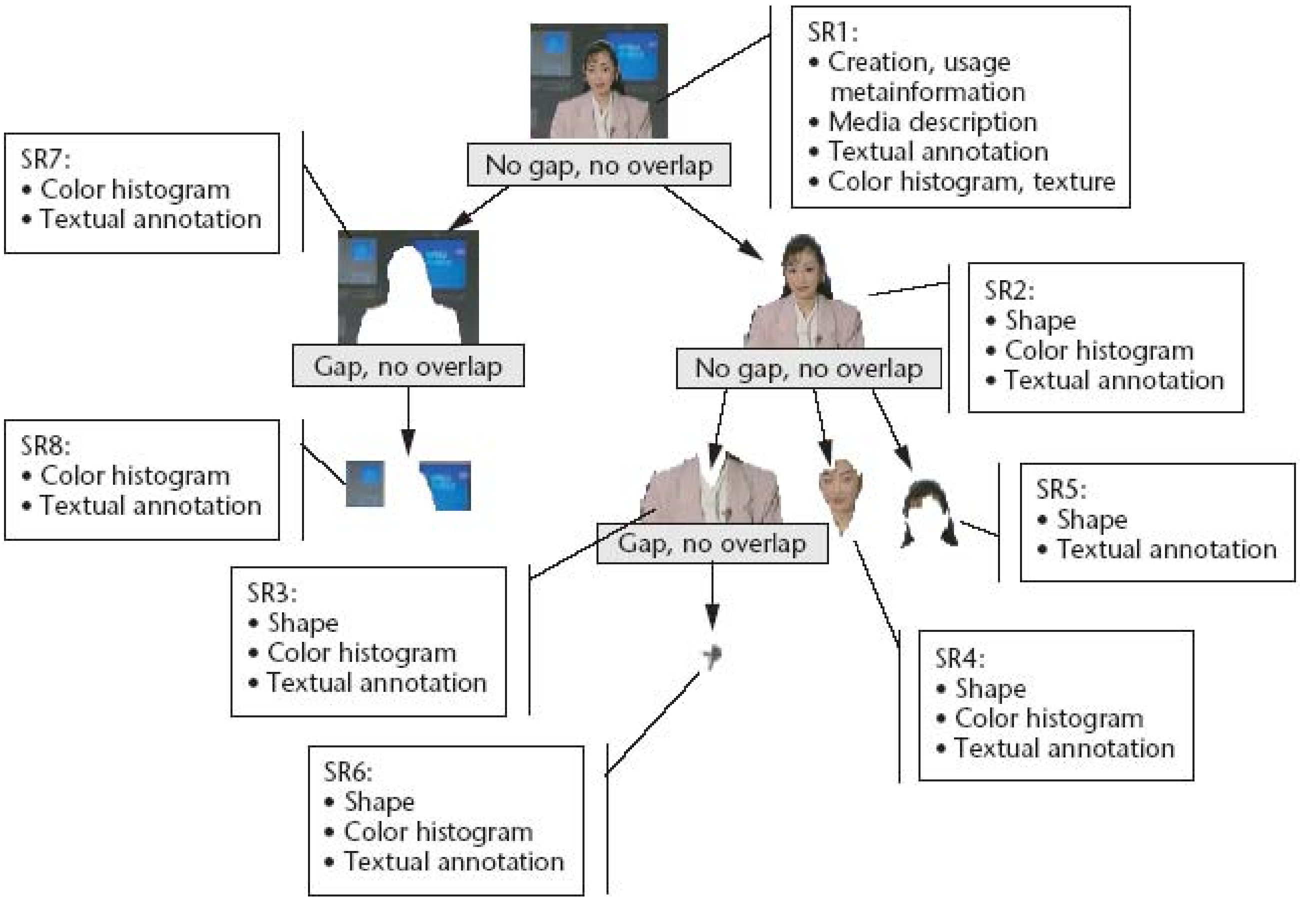
MULTIMEDIA DESCRIPTION SCHEMES. OVERVIEW OF THE TIME DSs



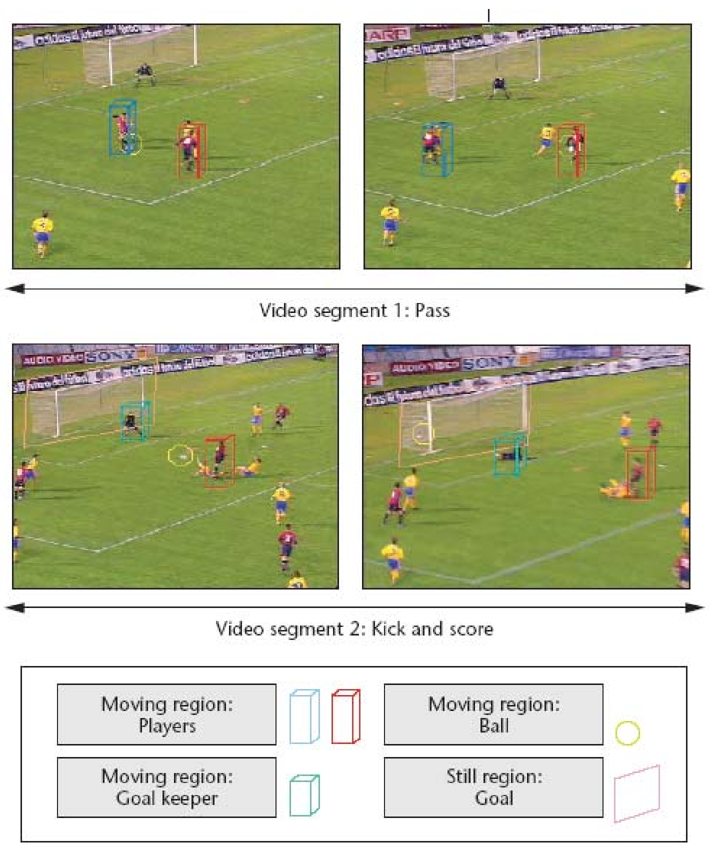

Creation information:
Creation
Creator
Creation coordinates
Creation location
Creation date
Media information:
Media profile
Media format
Media instance
Usage information:
Rights
Photographer: Seungyup
Place: Columbia University
Time: 19 September 1998
704 × 480 pixels
jpg
http://www.ee.columbia.edu/~ana/alex&ana.jpg
Columbia University,
All rights reserved
(a)
Still region SR1:
Text annotation
Color structure
Semantic relation:
Symbolized by
Concept C1:
Label
Property
Property
Semantic relation:
Property
Comradeship
Spatial segment decomposition:
No overlap, gap
Semantic relation:
Accompanier
Still region SR2:
Text annotation
Color structure
Shake hands
Event EV1:
Label
Semantic time
Semantic place
Semantic relation:
Agent
Alex
Ana
Agent object AO2:
Label
Person
Still region SR3:
Text annotation
Matching hint
Color structure
Spatial relation:
left
Semantic relation:
depicted by
Agent object AO1:
Label
Person
(b)
(c)

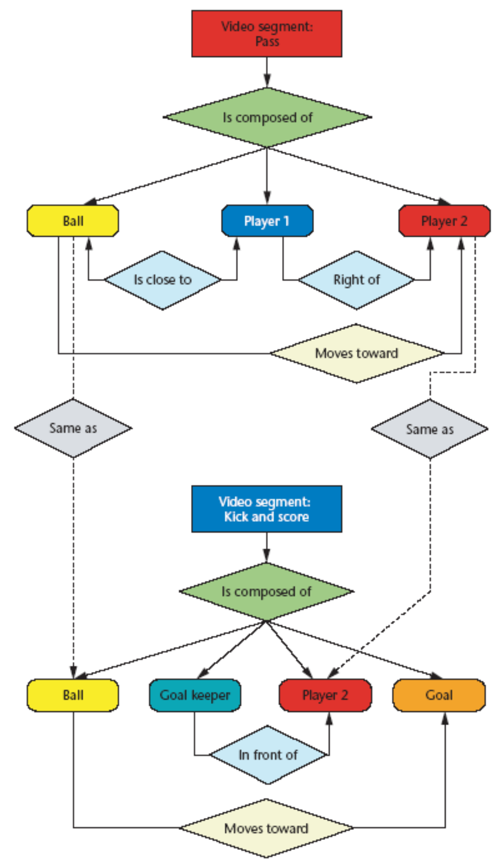

Creation information:
Creation
Creator
Creation coordinates
Creation location
Creation date
Usage information:
Rights
Photographer: Seungyup
Place: Columbia University
Time: 19 September 1998
Columbia University,
All rights reserved
Creation information:
Creation
Creator
Creation coordinates
Creation location
Creation date
Usage information:
Rights
Camera: José
Place: Columbia University
Time: 19 September 1998
Columbia University,
All rights reserved
Media information:
Media profile
Media format
Media instance
master
704 × 480 pixels
jpg
http://www.ee.columbia.edu/~ana/alex&ana.jpg
Media information:
Media profile
Media format
Media instance
Media instance
352 × 240 pixels
gif
http://www.ee.columbia.edu/~ana/alex&ana.gif
http://www.ee.columbia.edu/~alex/alex&ana.gif
master
640 × 480 pixels
mpg
5 seconds
http://www.ee.columbia.edu/~ana/alex&ana.mpg
Media information:
Media profile
Media format
Media instance
Image Content Entity
Two Profiles, Three Instances
Video Content Entity
One Profile, One Instance

MULTIMEDIA DESCRIPTION SCHEMES. CONTENT MANAGEMENT



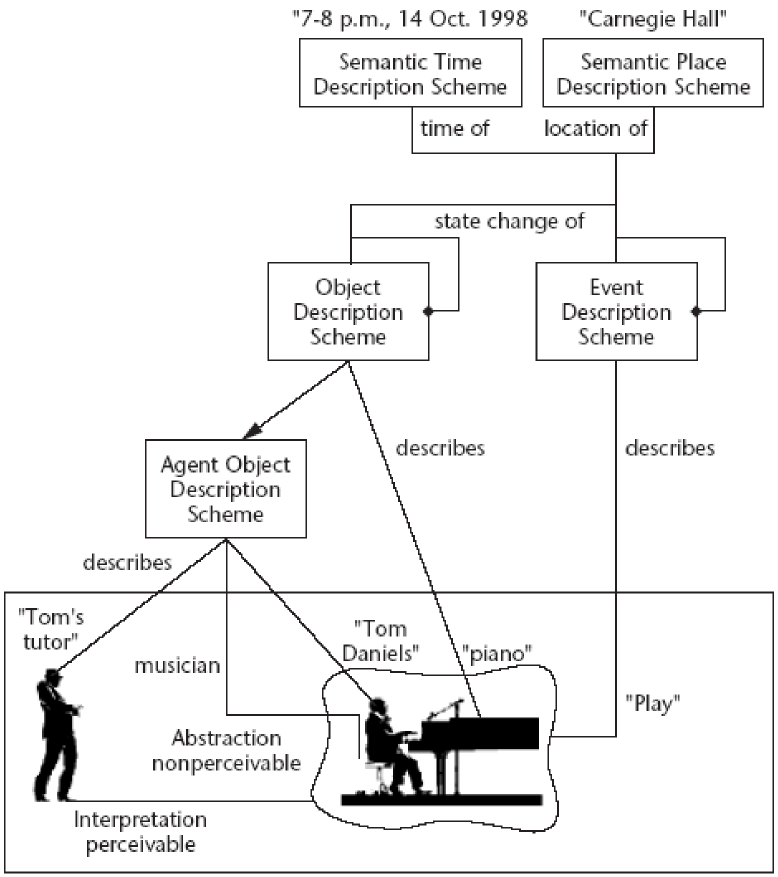
MULTIMEDIA DESCRIPTION SCHEMES. CONTENT MANAGEMENT

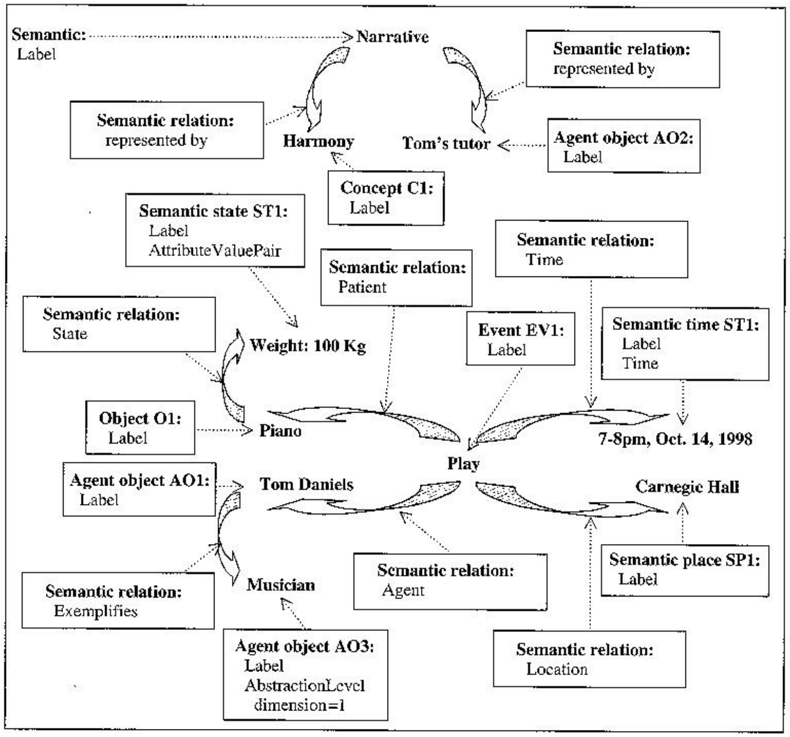




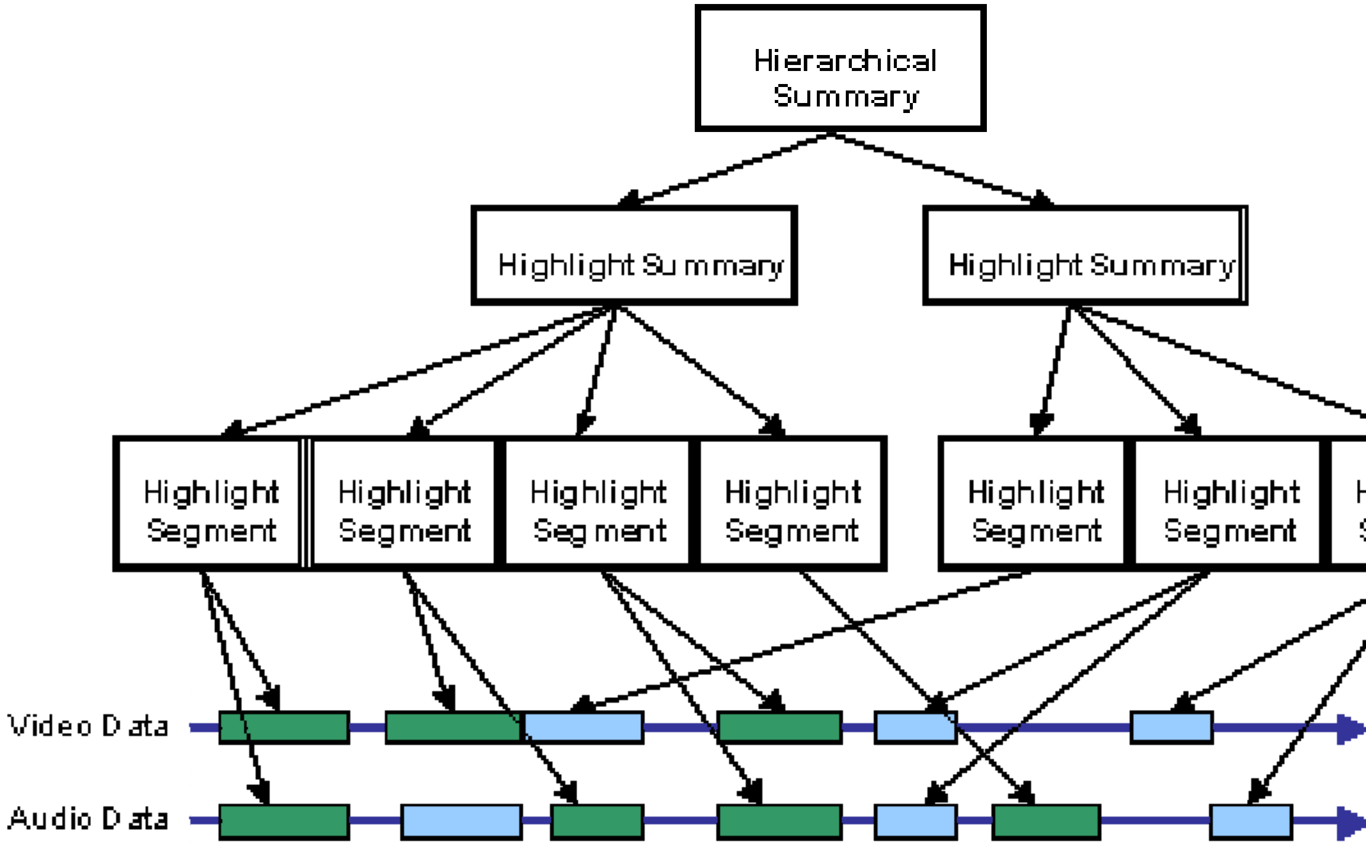
Example of Video Segments and Regions for the Segment Relationship Graph

MULTIMEDIA DESCRIPTION SCHEMES. CONCEPTUAL ASPECTS

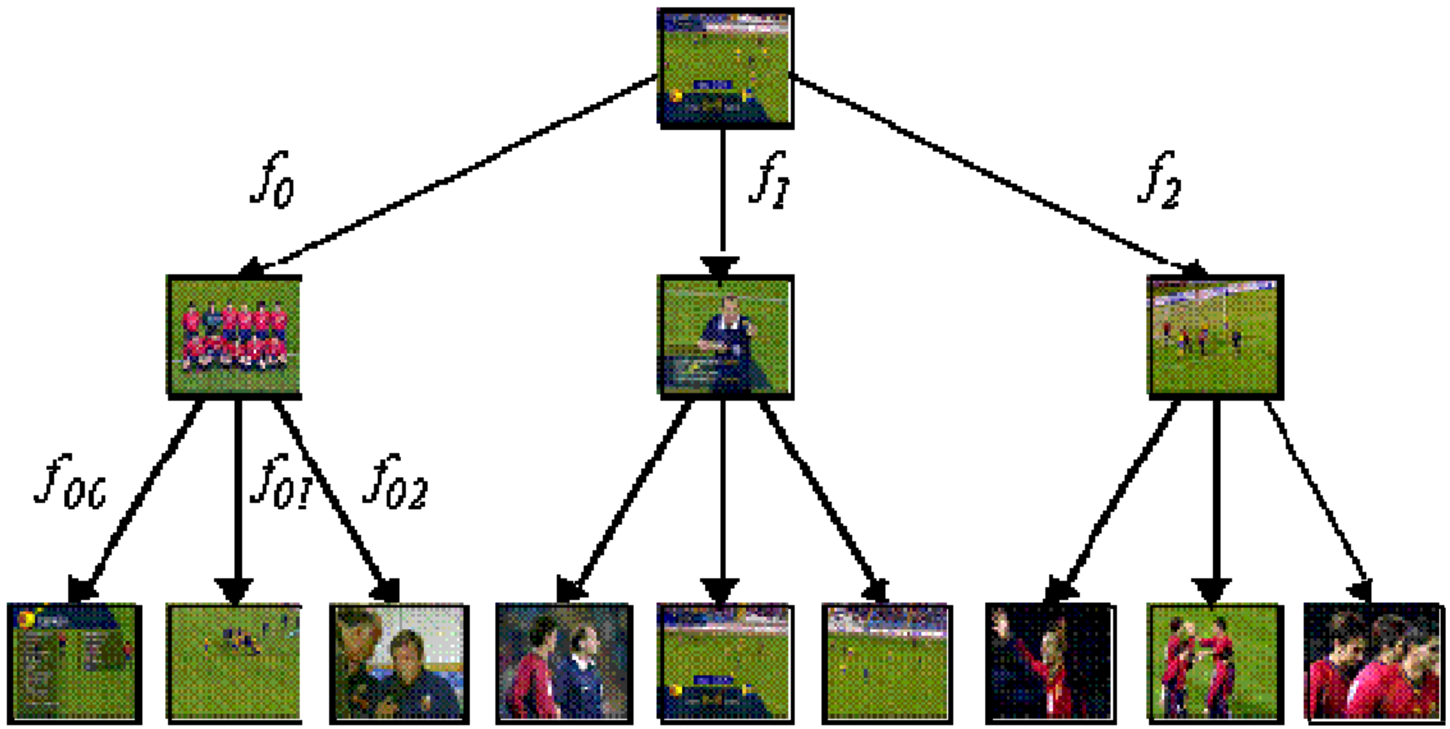




MULTIMEDIA DESCRIPTION SCHEMES. HIERARCHICAL SUMMARY DS

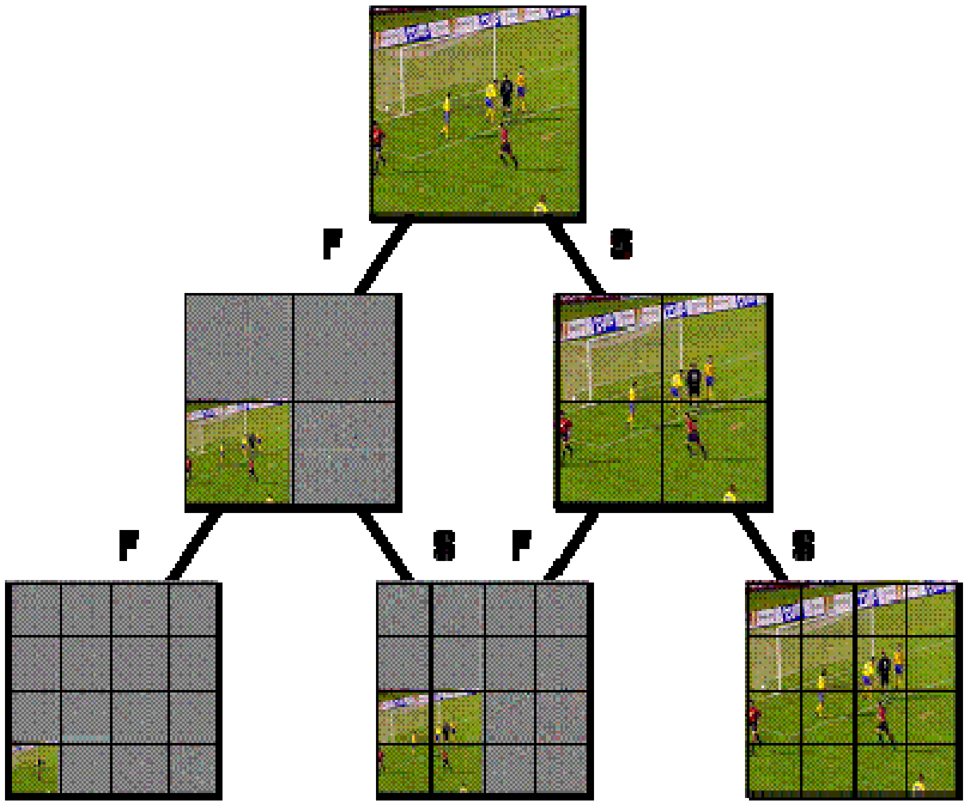


MULTIMEDIA DESCRIPTION SCHEMES. HIERARCHICAL SUMMARY DS



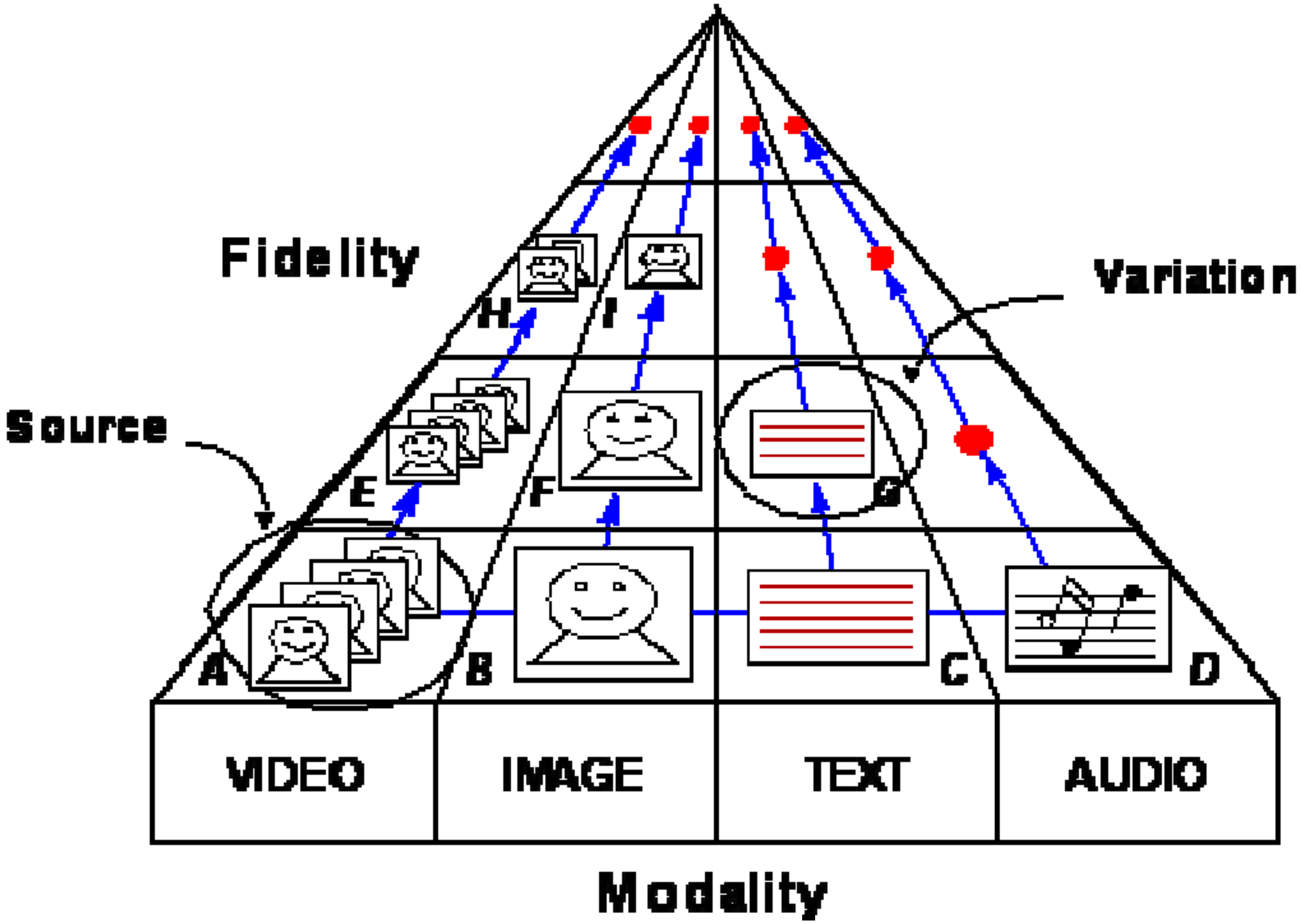
Example of a Hierarchical Summary of a video of a soccer game providing a multiple level key-frame hierarchy. The Hierarchical Summary denotes the fidelity (i.e., f0, f1) of each key-frame with respect to the video segment referred to by the key-frames at the next lower level.

MULTIMEDIA DESCRIPTION SCHEMES. PATRICIAN AND DECOMPOSITION

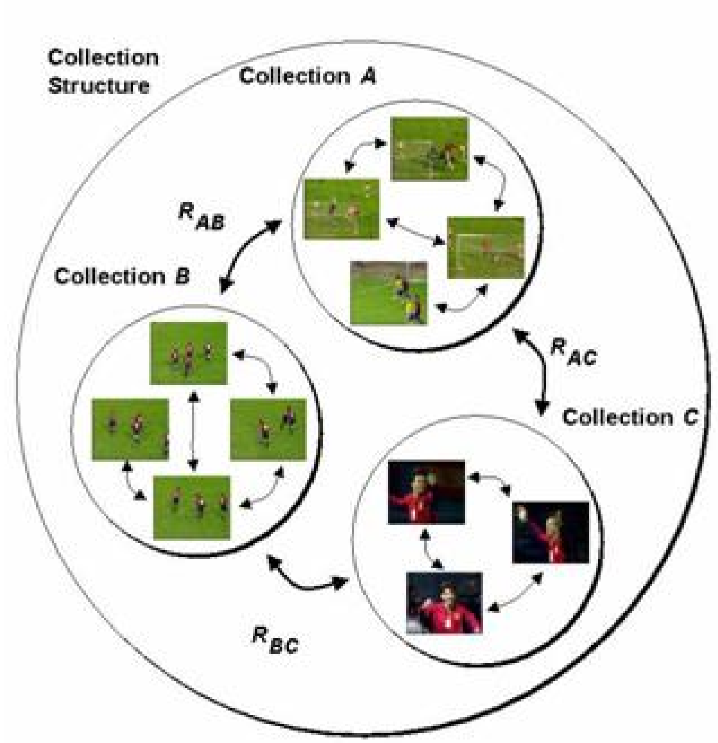


The Space and Frequency Graph describes the decomposition of an audio or visual signal in space (time) and frequency. S - spatial; F - frequency

MULTIMEDIA DESCRIPTION SCHEMES. AV PROGRAM



MULTIMEDIA DESCRIPTION SCHEMES. CONTENT ORGANIZATION



Probability model
Analytic model
Cluster model
Classification model

The Collection Structure DS describes collections of audio-visual content including the relationships (i.e., R_{AB}, R_{BC}, R_{AC}) within and across Collection Clusters.

VISUAL DESCRIPTORS. TOOLS

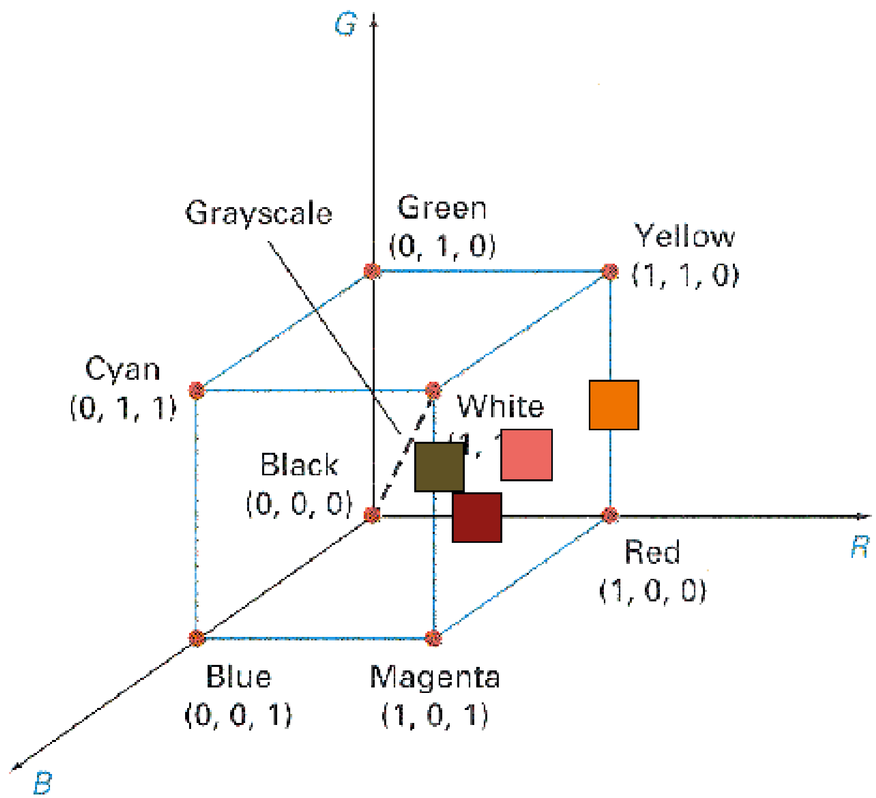
- Basic Elements (used by the other Visual descriptors): grid layout, time series, 2D–3D multiple view, spatial 2D coordinates, and temporal interpolation.
- Color Descriptors: Color Space, Color Quantization, Scalable Color, Dominant Color, Color Layout, Color Structure, and Group-of-Frames/Group-of-Pictures Color.
- Texture Descriptors: Homogeneous Texture, Non-Homogeneous Texture (Edge histogram), and Texture Browsing.
- Shape Descriptors: Region-Based, Contour-Based, and 3D Shape.
- Motion Descriptors (for video): Motion Activity, Camera Motion, Parametric Motion, and Motion Trajectory.
- Location Descriptors: Region Locator and Spatio–Temporal Locator.

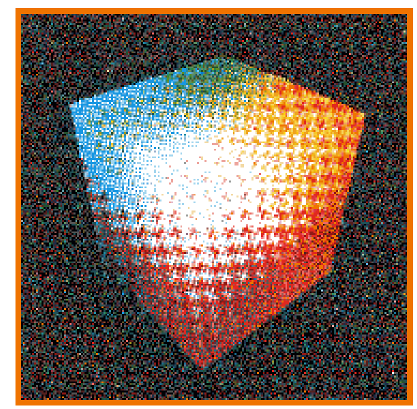
COLOR MODELS

	Monochrome	RBG	YC_rC_b	HS V	HMMD	Linear Matrix
MPEG -1,-2,-4	✍	✍	✍			
MPEG-7	✍	✍	✍	✍	✍	✍

RGB MODEL

- RGB Color cube





YCBCR MODEL

- describing the luminance and two color difference signals.
- Y = 0.299*R + 0.587*G + 0.114*B
- Cb = - 0.169*R - 0.331*G + 0.500*B
- Cr = 0.500*R – 0.419*G – 0.081*B

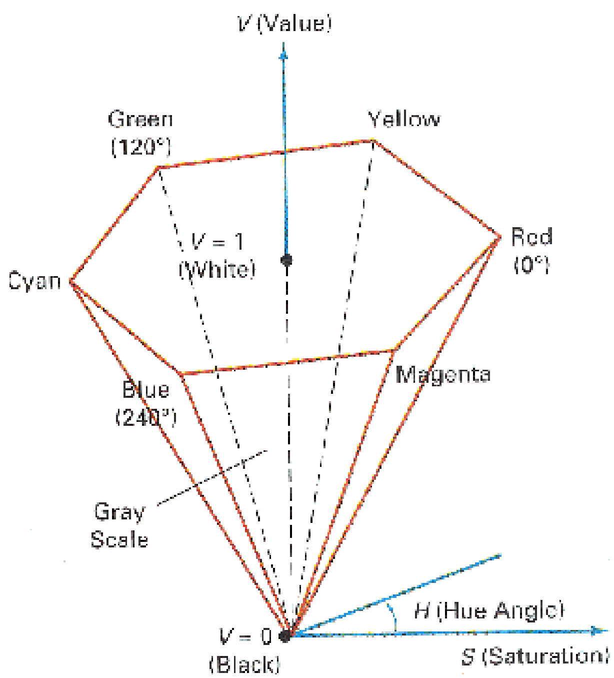
HSV MODEL

HSV model (hue, saturation, value): used by artists

- Hue: the name that we give to a color: red, yellow, cyan, etc.
- Saturation: “how pure the color is”, that is how far is a hue from the same hue mixed with white so that the color has more pastel shade /=1 pure red, yellow, blue, etc./
- Value/Light - how bright the color appear

HSV MODEL

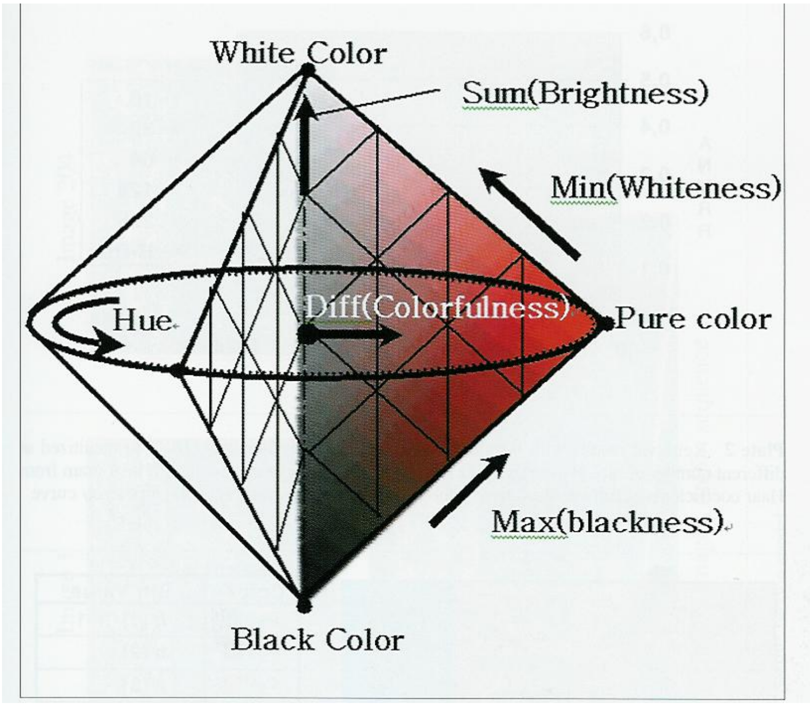
HSV model (hue, saturation, value)



H	S	V	Color
0	1.0	1.0	Red
120	1.0	1.0	Green
240	1.0	1.0	Blue
*	0.0	1.0	White
*	0.0	0.5	Gray
*	*	0.0	Black
60	1.0	1.0	
270	0.5	1.0	
270	0.0	0.7	

HMMD MODEL

- Hue, Max, Min, Diff and Sum values. The max value represents the blackness and the min value represents the whiteness of a color. The diff value represents the colorfulness by indicating how much gray the color contains. The sum value represents the brightness.



COLOR DESCRIPTORS. DOMINANT COLOR DESCRIPTOR

Dominant color descriptor is most suitable for representing local (object or image region) features where a small number of colors are enough to characterize the color information in the region of interest. Whole images are also applicable, for example, flag images or color trademark images. Color quantization is used to extract a small number of representing colors in each region/image. The percentage of each quantized color in the region is calculated correspondingly. A spatial coherency on the entire descriptor is also defined, and is used in similarity retrieval.

COLOR DESCRIPTORS. DOMINANT COLOR DESCRIPTOR

The result of the method is a vector from integer numbers, presented as $F = \{(c_i, p_i, v_i), s\}$, (*i*=*1*,*2*, …, *N*), where *N* is the number of dominant colors. The vector components are: the dominant color value c_i (RGB color space vector); p_i - normalized fraction of pixels corresponding to color c_i; optimal color variance v_i, describes the variance of the color values of the pixels in a cluster around the corresponding color; coherency s represents the overall spatial homogeneity of the dominant colors.

$$F_1 = \{(c_1^{'}, p_1^{'}, v_1^{'}), s_1^{'}\} \qquad (i = 1,2,\dots,N_1)$$ and

$$F_2 = \{(c_2^{'}, p_2^{'}, v_2^{'}), s_2^{'}\} \qquad (i = 1,2,\dots,N_2)$$

where $p \in [0,31]$, $c_i = rgb2luv(c_i^{'})$, $v_i = \begin{cases} 60.0 & v_i^{'} = 0 \\ 90.0 & v_i^{'} = 1 \end{cases}$,

$$p_i = \frac{(p_i^{'} + 0.5)/31.9999}{\sum_i p_i}$$, and if

$$f_{x_i y_j} = \frac{1}{2\pi\sqrt{2\pi \times (v_{x_i}^{(l)} + v_{y_j}^{(l)}) \times (v_{x_i}^{(u)} + v_{y_j}^{(u)}) \times (v_{x_i}^{(v)} + v_{y_j}^{(v)})}} \times \exp\left\{-\frac{1}{2}\left[\frac{c_{x_i}^{(l)} - c_{y_j}^{(l)}}{v_{x_i}^{(l)} + v_{y_j}^{(l)}} + \frac{c_{x_i}^{(u)} - c_{y_j}^{(u)}}{v_{x_i}^{(u)} + v_{y_j}^{(u)}} + \frac{c_{x_i}^{(v)} - c_{y_j}^{(v)}}{v_{x_i}^{(v)} + v_{y_j}^{(v)}}\right]\right\}$$

and $D_v = \sqrt{\sum_{i=1}^{N_1}\sum_{j=1}^{N_1} p_{1_i} p_{1_j} f_{1_i 1_j} + \sum_{i=1}^{N_2}\sum_{j=1}^{N_2} p_{2_i} p_{2_j} f_{2_i 2_j} - 2\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} p_{1_i} p_{2_j} f_{1_i 2_j}}$ then

the distance is calculated as: $D = [0.3 \times abs(s_1 - s_2) + 0.7] \times D_v$.

DOMINANT COLOR DESCRIPTOR

```xml
<VisualDescriptor xsi:type="DominantColorType">
        <SpatialCoherency>19</SpatialCoherency>
        <Value>
            <Percentage>10</Percentage>
            <Index>71 55 41</Index>
            <ColorVariance>1 0 0</ColorVariance>
        </Value>
        <Value>
            <Percentage>19</Percentage>
            <Index>104 98 69</Index>
            <ColorVariance>0 0 0</ColorVariance>
        </Value>
        <Value>
            <Percentage>2</Percentage>
            <Index>184 184 156</Index>
            <ColorVariance>1 0 0</ColorVariance>
        </Value>
    </VisualDescriptor>
```

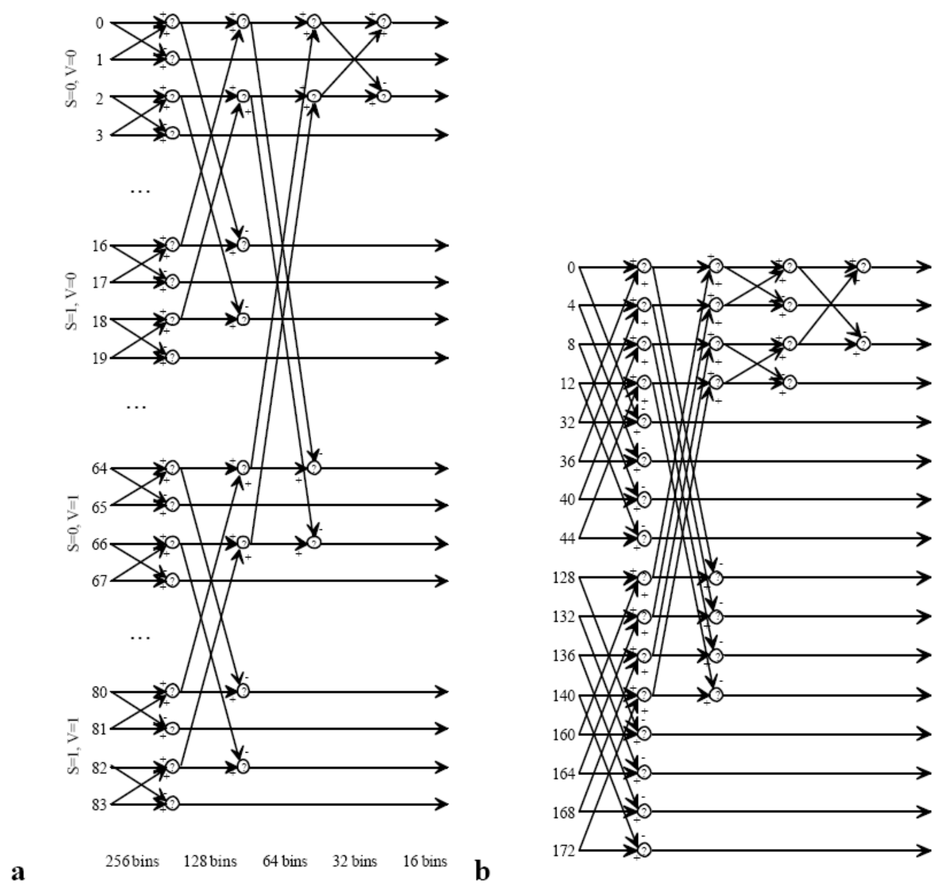
COLOR DESCRIPTORS.
DOMINANT COLOR DESCRIPTOR

443.523

COLOR DESCRIPTORS. SCALABLE COLOR DESCRIPTOR

The Scalable Color Descriptor is a Color Histogram in HSV Color Space, which is encoded by a Haar transform. Its binary representation is scalable in terms of bin numbers and bit representation accuracy over a broad range of data rates. The Scalable Color Descriptor is useful for image-to-image matching and retrieval based on color feature. Retrieval accuracy increases with the number of bits used in the representation.

COLOR DESCRIPTORS.
SCALABLE COLOR DESCRIPTOR



(a) Basic unit of Haar transform (b) schematic diagram of SCD generation

COLOR DESCRIPTORS. SCALABLE COLOR DESCRIPTOR

The distance matching can be done either in the Haar coefficient domain or in the histogram domain. In the case where only the coefficient signs are retained, the matching can be done efficiently in the Haar coefficient domain by calculating the Hamming distance as the number of bit positions at which the binary bits are different using an XOR operation on the two descriptors to be compared. This induces only a marginal loss in similarity matching precision compared to reconstructing the color histogram and performing histogram matching, while the computational cost is considerably lower.

COLOR DESCRIPTORS. SCALABLE COLOR DESCRIPTOR

<VisualDescriptor xsi:type="ScalableColorType" numOfBitplanesDiscarded="0" numOfCoeff="64">
<Coeff>-74 -57 -86 69 -34 14 22 28 -31 13 11 7 -13 14 19 22
-5 1 0 8 -15 5 0 5 9 2 2 0 -6 5 1 -4 3 3 0 1 0 0 0 -3 5 2 1 3 1 2 3 -1 -4 -3 2 2 2 3
3 0 -3 0 0 -2 1 0 -3 -3</Coeff>
</VisualDescriptor>

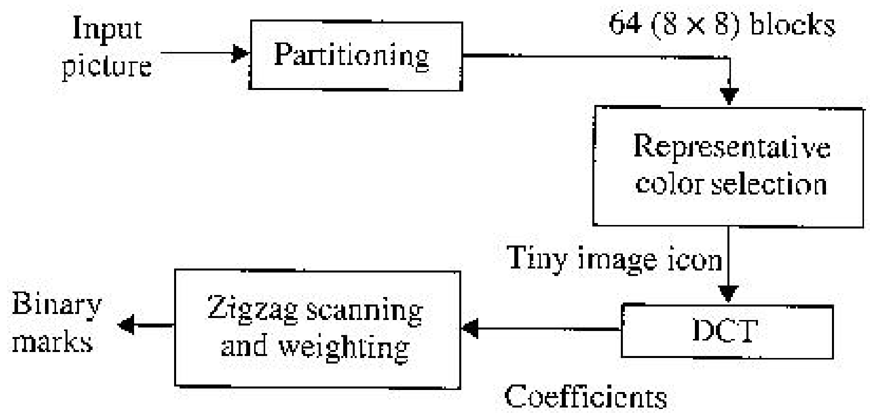
COLOR DESCRIPTORS. SCALABLE COLOR DESCRIPTOR

336.00

COLOR DESCRIPTORS.
COLOR LAYOUT DESCRIPTOR

Color layout descriptor effectively represents the spatial distribution of color of visual signals in a very compact form. This compactness allows visual signal matching functionality with high retrieval efficiency at very small computational costs. It provides image-to-image matching as well as ultra high-speed sequence-to-sequence matching, which requires so many repetitions of similarity calculations.

COLOR DESCRIPTORS.
COLOR LAYOUT DESCRIPTOR



The method output is a vector from integer numbers, describes *{DY, DCr, DCb}* coefficients, where *Y* is the coefficient value for luminance, *Cr, Cb* coefficient values for chrominance.

For matching two descriptions *{DY, DCr, DCb}* and *{DY', DCr', DCb'}* the following formula:

$$D = \sqrt{\sum_i w_{yi}(DY_i - DY_i')^2} + \sqrt{\sum_i w_{bi}(DCb_i - DCb_i')^2} + \sqrt{\sum_i w_{ri}(DCr_i - DCr_i')^2}$$ is used, where

i represents the zigzag- scanning order of the coefficients.

COLOR DESCRIPTORS. COLOR LAYOUT DESCRIPTOR

<VisualDescriptor xsi:type="ColorLayoutType">
<YDCCoeff>15</YDCCoeff>
<CbDCCoeff>19</CbDCCoeff>
<CrDCCoeff>37</CrDCCoeff>
<YACCoeff5>14 23 15 14 14 </YACCoeff5>
<CbACCoeff2>16 16 </CbACCoeff2>
<CrACCoeff2>17 14 </CrACCoeff2>
</VisualDescriptor>

COLOR DESCRIPTORS. COLOR LAYOUT DESCRIPTOR

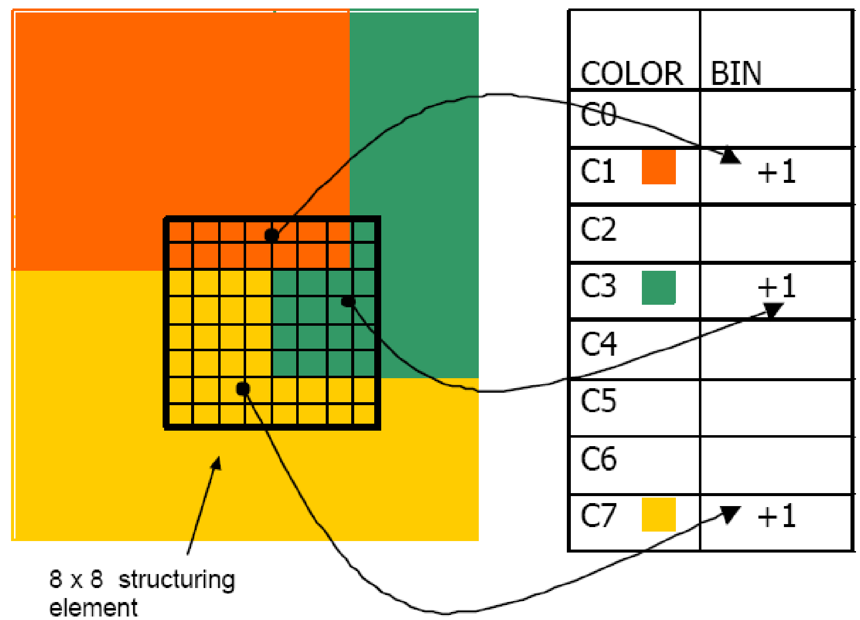
48.9038

COLOR DESCRIPTORS.
COLOR STRUCTURE DESCRIPTOR

The Color structure descriptor is a color feature descriptor that captures both color content (similar to a color histogram) and information about the structure of this content. The extraction method embeds color structure information into the descriptor by taking into account all colors in a structuring element of 8x8 pixels that slides over the image, instead of considering each pixel separately. Unlike the color histogram, this descriptor can distinguish between two images in which a given color is present in identical amounts but where the structure of the groups of pixels having that color is different in the two images. Color values are represented in the double-coned HMMD color space, which is quantized non-uniformly into 32, 64, 128 or 256 bins. Each bin amplitude value is represented by an 8-bit code.

COLOR DESCRIPTORS. COLOR STRUCTURE DESCRIPTOR

- The output of the method is a vector from integer numbers, presented by a 256 bin histogram.
- The matching is done by minimizing the distance calculated as the sum of the differences between the corresponding bins in any two color-structure histograms.



COLOR DESCRIPTORS. COLOR STRUCTURE DESCRIPTOR

<VisualDescriptor xsi:type="ColorStructureType" colorQuant="2">
<Values>0 0 0 0 0 0 0 0 42 8 1 0 0 0 0 0 0 0 0 0
0 0 0 0 235 255 75 41 0 0 0 0 0 0 0 0 5 1 0 0 91 76 12 2 0 0 0 0 0
0 0 0 26 3 0 0 4 6 0 0 0 0 0 0</Values>
</VisualDescriptor>

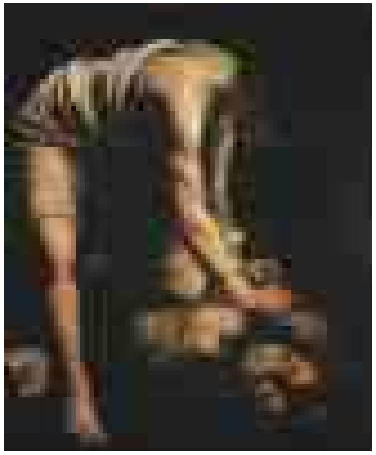
COLOR DESCRIPTORS. COLOR STRUCTURE DESCRIPTOR

5.42353

MPEG 7 system

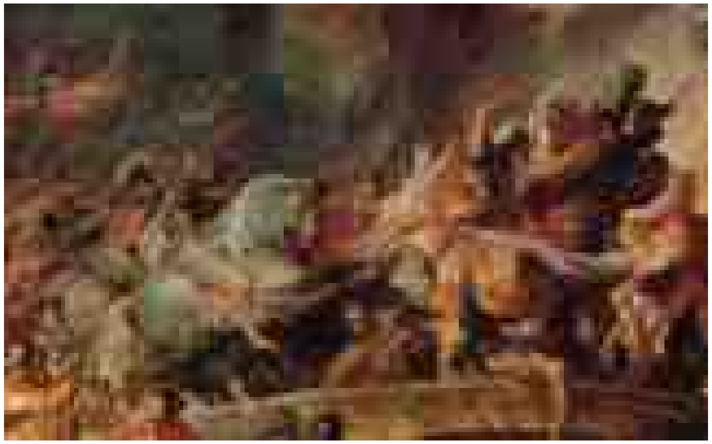
HIGH LEVEL COLOR CHARACTERISTICS. BAROQUE

Baroque period, era in the history of the Western arts roughly coinciding with the 17th century. The work that distinguishes the Baroque period is stylistically complex, even contradictory. In general, however, the desire to evoke emotional states by appealing to the senses, often in dramatic ways, underlies its manifestations. Some of the qualities most frequently associated with the Baroque are **grandeur, sensuous richness, drama, vitality, movement, tension, emotional exuberance**, and a tendency to blur distinctions between the various arts.



Rubens, Peter Paul
Battle of the Amazons

David and Goliath
Caravaggio,
Michelangelo



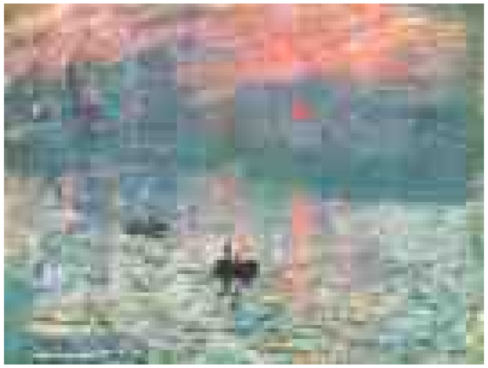
HIGH LEVEL COLOR CHARACTERISTICS. CUBISM

Cubism, highly influential visual arts style of the 20th century that was created principally by the painters Pablo Picasso and Georges Braque in Paris between 1907 and 1914. The Cubist style emphasized the flat, two-dimensional surface of the picture plane, rejecting the traditional techniques of perspective, foreshortening, modeling, and refuting time-honored theories of art as the imitation of nature. Cubist painters were not bound to copying form, texture, color, and space; instead, they presented a new reality in paintings that depicted radically fragmented objects, whose several sides were seen simultaneously.

Portrait of Picasso

HIGH LEVEL COLOR CHARACTERISTICS. EXPRESSIONISM

Expressionism, artistic style in which the artist seeks to depict not objective reality but rather the subjective emotions and responses that objects and events arouse in him. He accomplishes his aim through distortion, exaggeration, primitivism, and fantasy and through the vivid, jarring, violent, or dynamic application of formal elements.

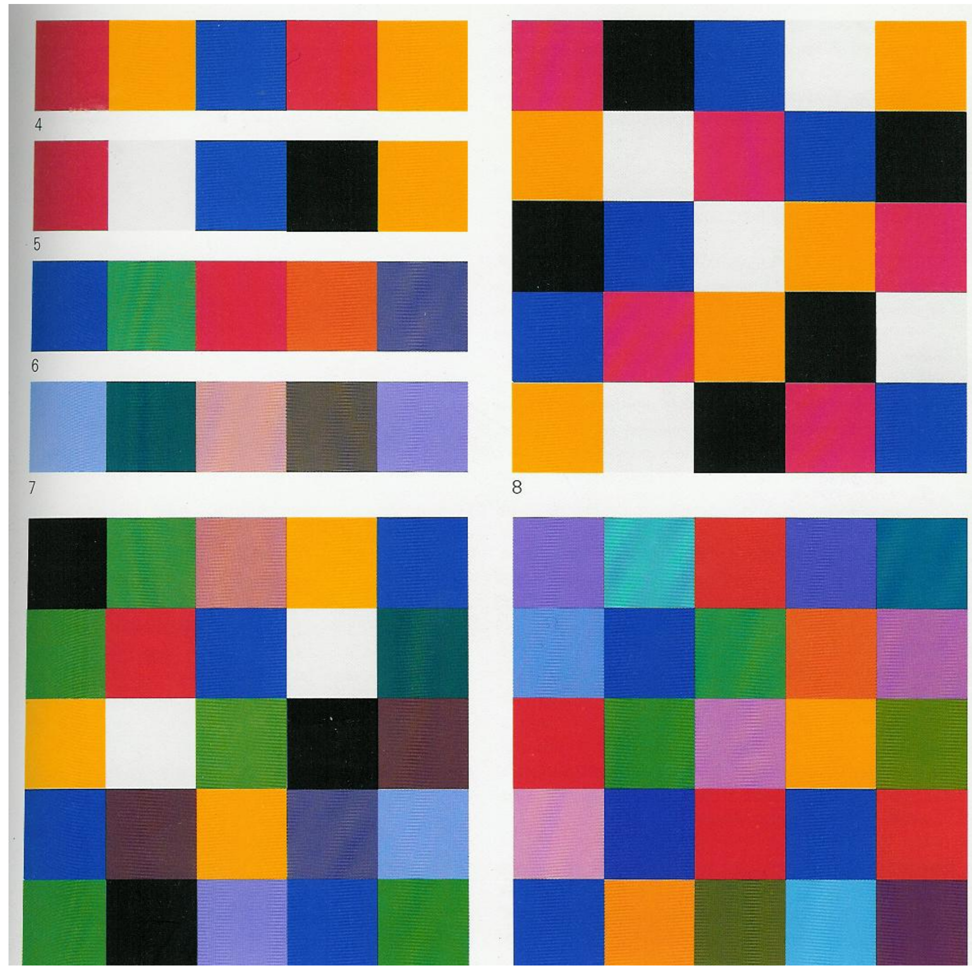


Soleil Levant
Claude Monet

HIGH LEVEL COLOR CHARACTERISTICS. SYMBOLISM

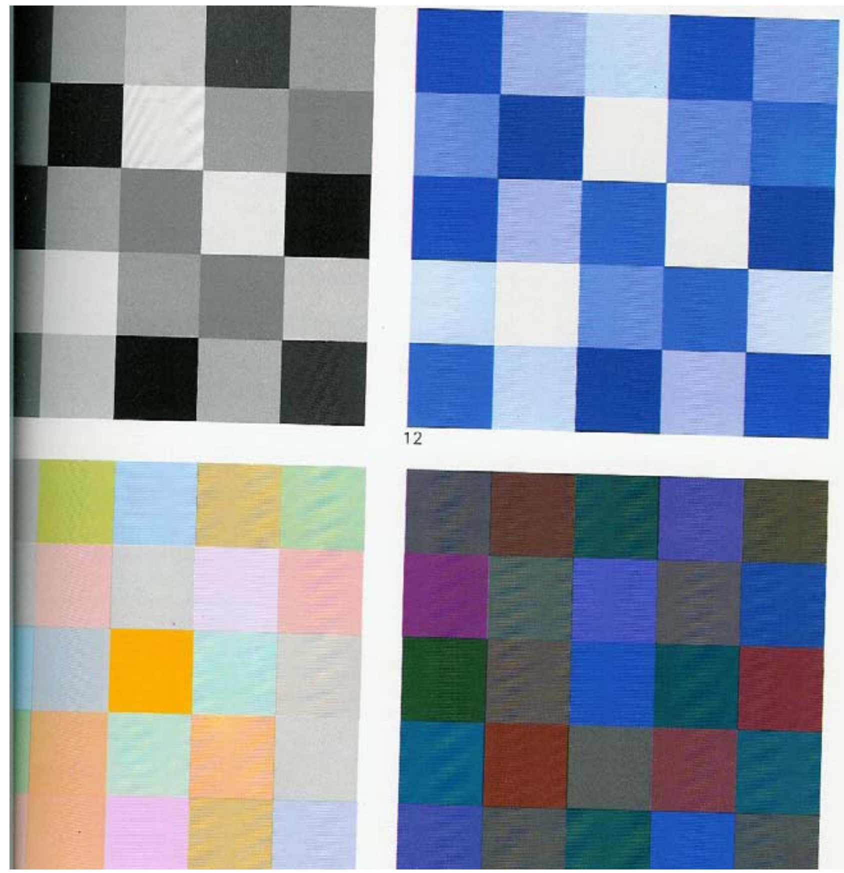
CONTRAST OF HUE

It presents undiluted colors in their most intense luminosity. Some color combinations are: yellow/red/blue, red/blue/green, blue/yellow/violet, yellow/green/violet/red/, violet/green/blue/orange /black.



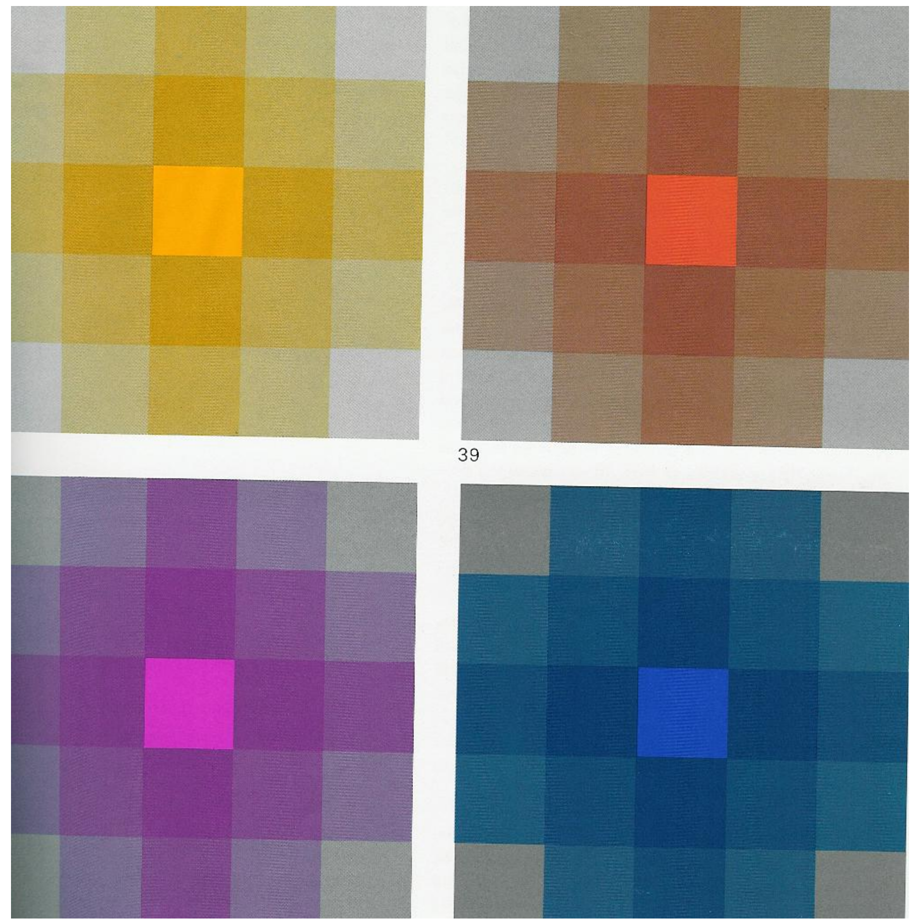
COMPLEMENTARY CONTRACT

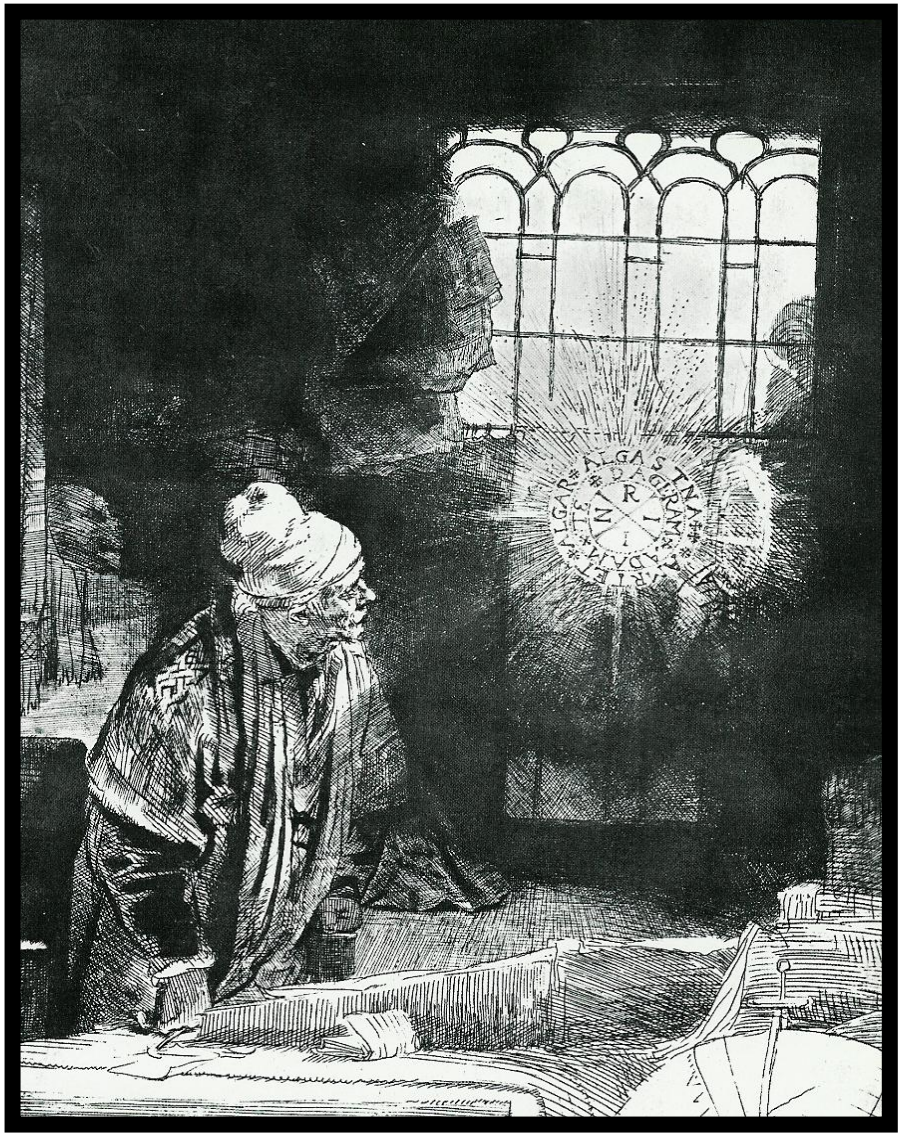
Two colors are called complementary if their pigments mixed together yield in neutral grey. Examples are: yellow-violet, blue-orange, red-green. This contrast gives the effect of a stability fixed image.



SATURATION CONTRACT

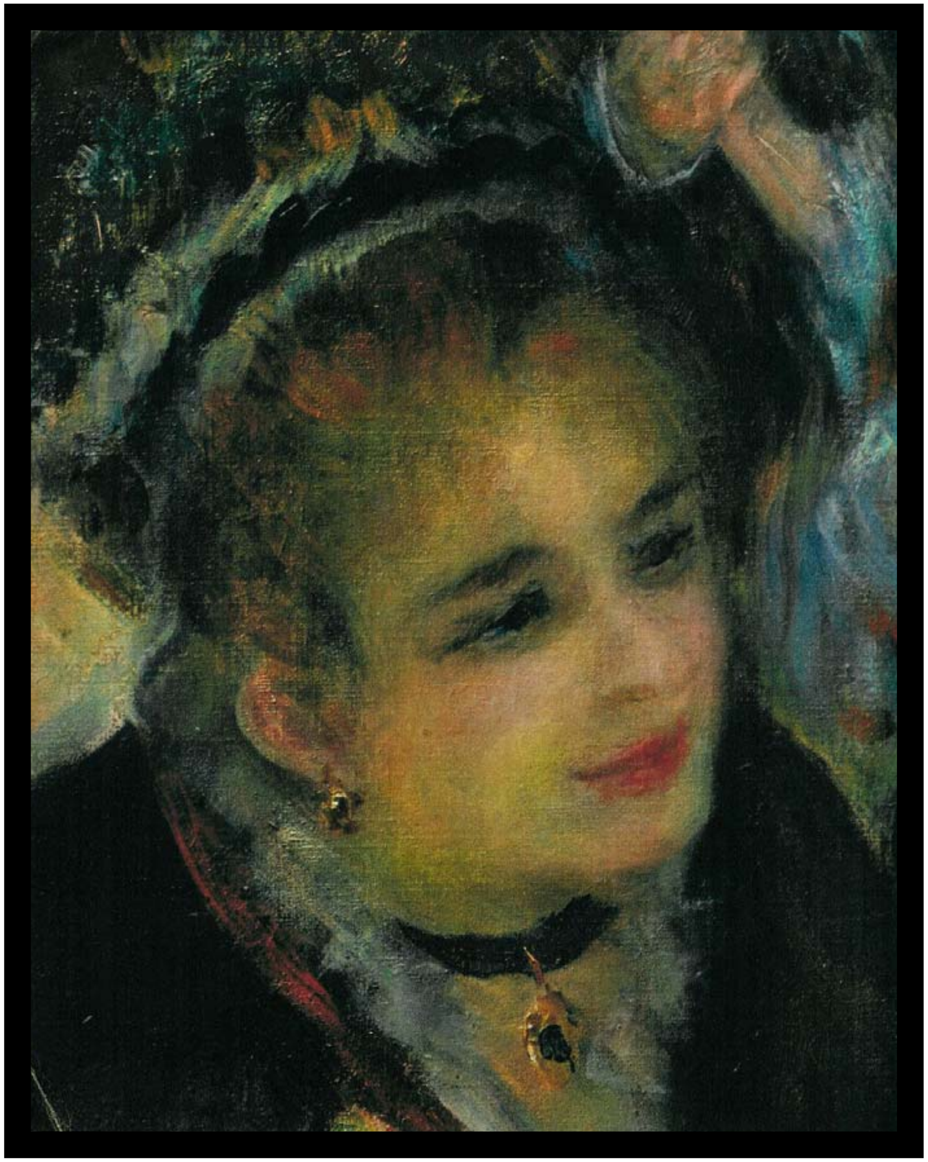
Saturation relates to the degree of purity of the color.

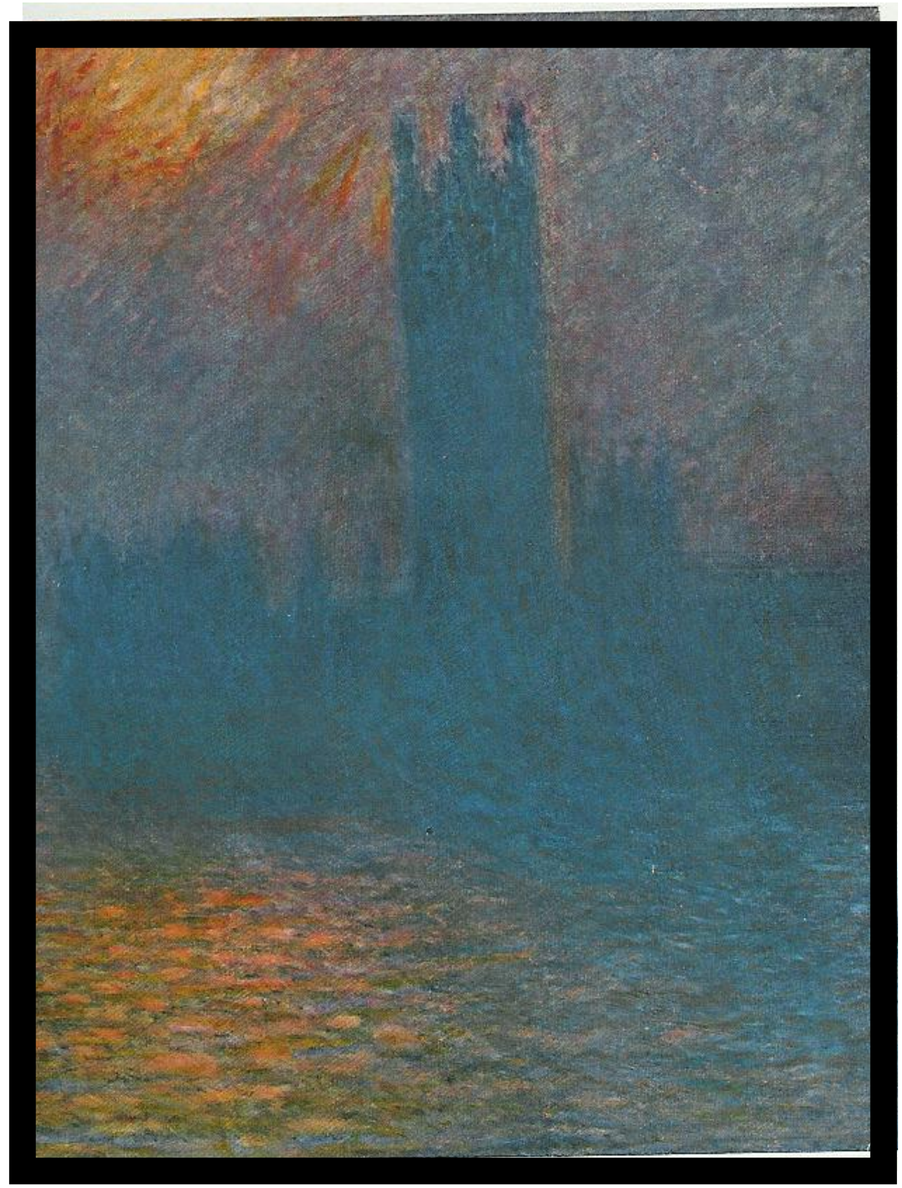




Light-dark contrast:

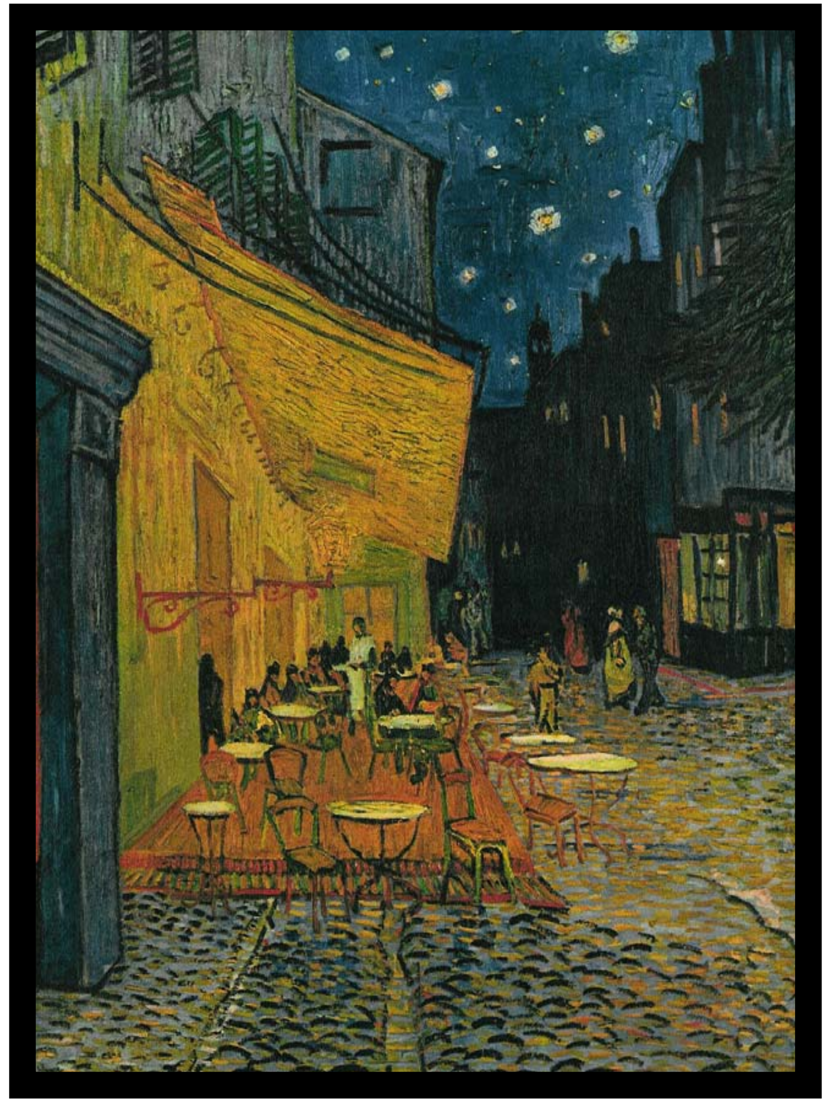
“Dr. Faustus in his study room” sketching by Rembrandt
and “Guitar on Mantelpiece” by Paulo Picasso





Cold-warm-contrast:

“Le Moulin de la Galette” by Auguste Renoir,
and “Houses of Parliament” by Clode Monet



Simultaneous contrast:

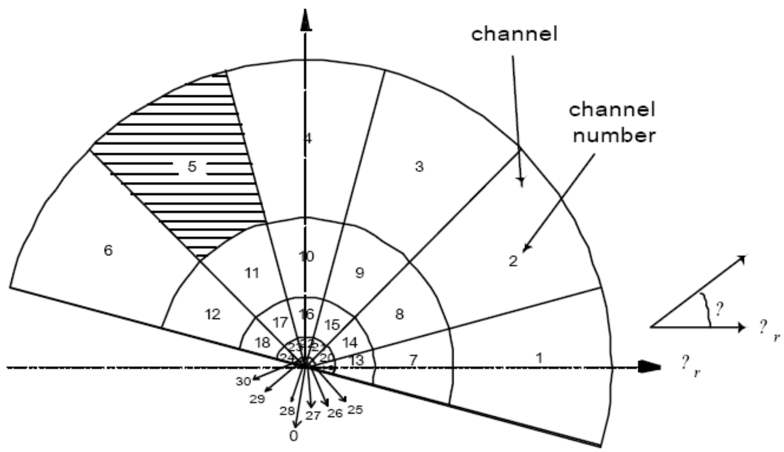
"Stripping of Christ" by El Greco, and
"Café at Evening" by Vincent van Gogh

TEXTURE DESCRIPTORS

TEXTURE DESCRIPTORS. HOMOGENOUS TEXTURE DESCRIPTOR

Homogeneous texture has emerged as an important visual primitive for searching and browsing through large collections of similar looking patterns. An image can be considered as a mosaic of homogeneous textures so that these texture features associated with the regions can be used to index the image data. Examples of queries that could be supported in this context could include “Retrieve all Land-Satellite images of Santa Barbara which have less than 20% cloud cover” or “Find a vegetation patch that looks like this region”.

TEXTURE DESCRIPTORS. HOMOGENOUS TEXTURE DESCRIPTOR



- Frequency layout for feature extraction using 6 orientations and 5scales.
- The matching is done by summing the normalized weighted absolute difference between two sets of feature vectors.

HOMOGENOUS TEXTURE DESCRIPTOR

/VisualDescriptor>
<VisualDescriptor xsi:type="HomogeneousTextureType">
<Average>83</Average>
<StandardDeviation>79</StandardDeviation>
<Energy>185 151 148 152 172 180 194 174 169
156 155 182 170 170 141 162 149 171 145 132 132 147 141 126
125 93 83 112 89 89</Energy>
<EnergyDeviation>185 150 145 149 173 176 182
169 170 153 141 177 155 165 124 150 134 167 139 109 130 128 136
123 117 75 64 102 65 75</EnergyDeviation>
</VisualDescriptor>

HOMOGENOUS TEXTURE DESCRIPTOR

3.74064

TEXTURE DESCRIPTORS.
TEXTURE BROWSING DESCRIPTOR

The Texture Browsing Descriptor is useful for representing homogeneous texture for browsing type applications, and requires only 12 bits (maximum). It provides a perceptual characterization of texture, similar to a human characterization, in terms of regularity, coarseness and directionality.

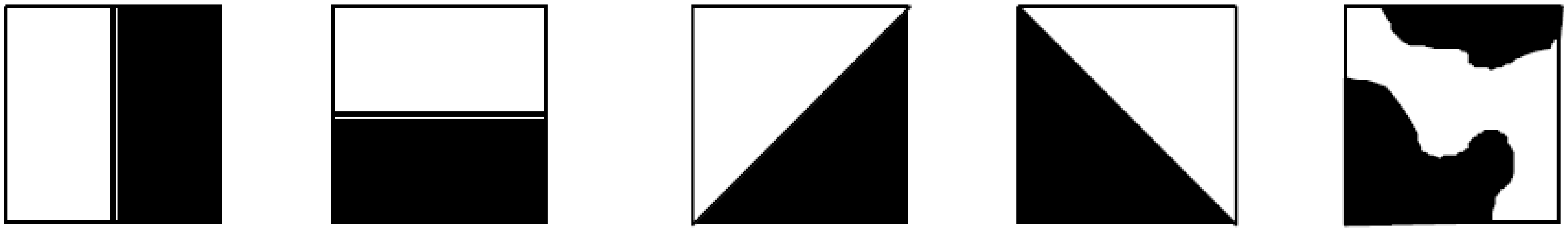
TEXTURE DESCRIPTORS. TEXTURE BROWSING DESCRIPTOR

First, the image is filtered with a bank of orientation and scale tuned filters (modeled using Gabor functions); from the filtered outputs, two dominant texture orientations are identified. Three bits are used to represent each of the dominant orientations. This is followed by analyzing the filtered image projections along the dominant orientations to determine the regularity (quantified to 2 bits) and coarseness (2 bits x 2).

TEXTURE DESCRIPTORS.
TEXTURE BROWSING DESCRIPTOR

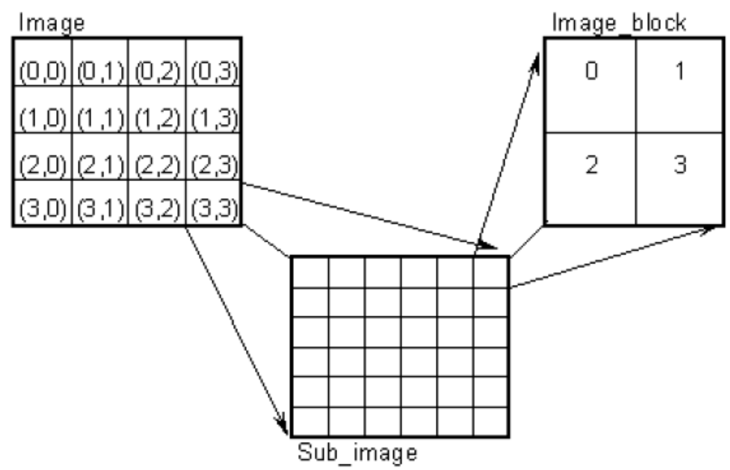
```
<VisualDescriptor xsi:type="TextureBrowsingType">
    <Regularity>irregular</Regularity>
    <Direction>0Degree</Direction>
    <Scale>fine</Scale>
    <Direction>90Degree</Direction>
    <Scale>fine</Scale>
  </VisualDescriptor>
```

TEXTURE DESCRIPTORS.
EDGE HISTOGRAM DESCRIPTOR



a) vertical edge | b) horizontal edge | c) 45 degree edge | d) 135 degree edge | e)non-directional edge

H_E	Semantics
h(0)	Relative population of vertical edges in subimage at (0,0)
h(1)	Relative population of horizontal edges in subimage at (0,0)
h(2)	Relative population of 45° edges in subimage at (0,0)
h(3)	Relative population of 135° edge in subimage at (0,0)
h(4)	Relative population of nondirectional edges in subimage at (0,0)
⋮	⋮
h(75)	Relative population of vertical edges in subimage at (3,3)
h(76)	Relative population of horizontal edges in subimage at (3,3)
h(77)	Relative population of 45° edges in subimage at (3,3)
h(78)	Relative population of 135° edges in subimage at (3,3)
h(79)	Relative population of nondirectional edges in subimage at (3,3)

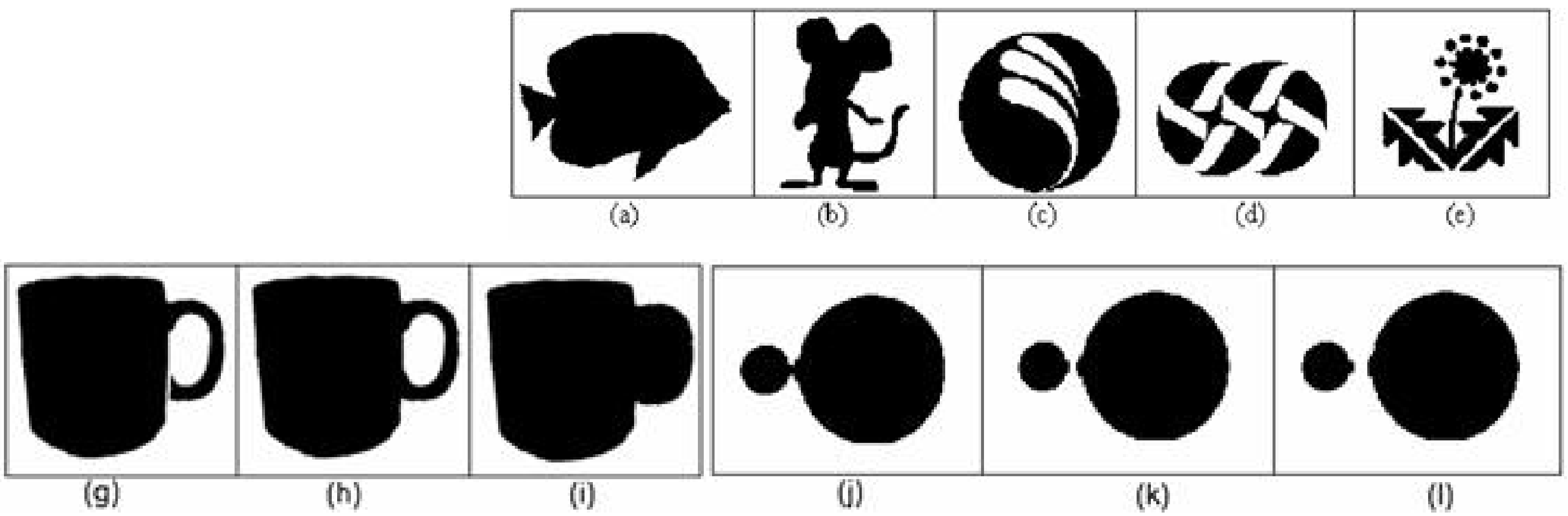


TEXTURE DESCRIPTORS.
EDGE HISTOGRAM DESCRIPTOR

<VisualDescriptor xsi:type="EdgeHistogramType">
<BinCounts>2 4 4 2 5 1 5 2 4 2 0
6 1 1 0 1 6 1 1 1 3 5 3 3 3 3 4 5 3 5 0 6 4 5 3 1 4 5 2
4 1 2 2 3 2 2 2 2 3 2 5 1 2 1 5 0 3 2 4 3 2 4 4 5 6 3 2
4 3 6 4 3 6 4 6 3 4 4 6 6</BinCounts>
</VisualDescriptor>

SHAPE DESCRIPTORS. REGION SHAPE

The feature extraction is based on a set of Angular Radial Transform (ART) coefficients. ART is a complex 2-D transform defined on a unit disc with polar coordinates.

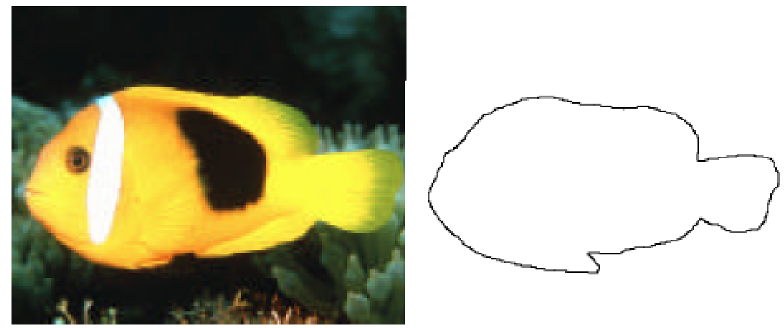


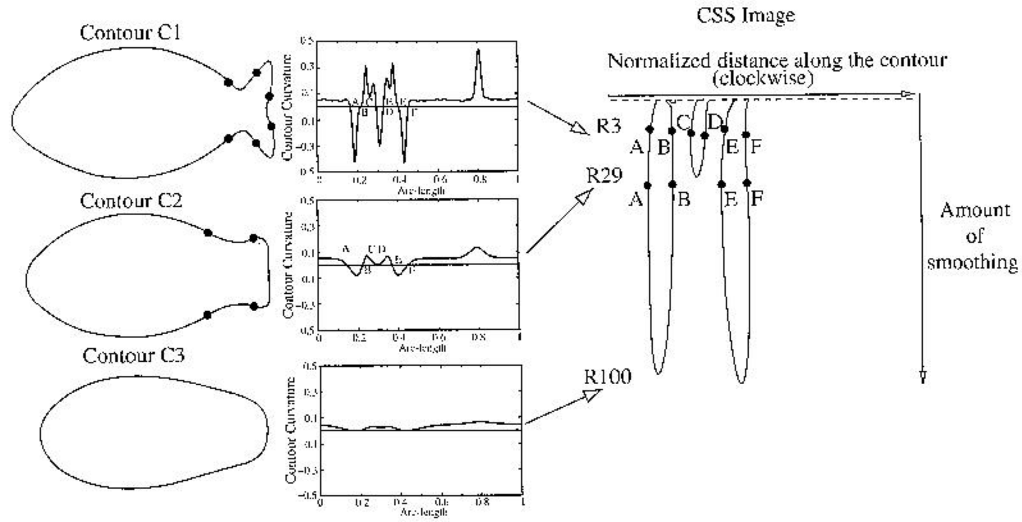
SHAPE DESCRIPTORS. REGION SHAPE

```
<VisualDescriptor xsi:type="RegionShapeType">
            <MagnitudeOfART>
15 8 13 14 13 15 14 15 15 14 14 15 9 15 15 13 15 15 6 1
5 15 13 15 15 5 14 15 13 15 15 9 13 15 12
15</MagnitudeOfART>
         </VisualDescriptor>
      </StillRegion>
      <StillRegion>  … </StillRegion>
      <StillRegion>  … </StillRegion>
      <StillRegion>  … </StillRegion>
      </SpatialDecomposition>
```

SHAPE DESCRIPTORS. CONTOUR SHAPE

The Contour Shape descriptor captures characteristic shape features of an object or region based on its contour. It uses so-called Curvature Scale-Space representation, which captures perceptually meaningful features of the shape.





AUDIO DESCRIPTORS

Audio Framework

Silence Description

Timbral Temporal

Log Attack Time Description
Temporal Centroid Description

Basic Spectral

Audio Spectrum Envelope Description
Audio Spectrum Centroid Description
Audio Spectrum Spread Description
Audio Spectrum Flatness Description

Basic

Audio Waveform Description
Audio Power Description

Timbral Spectral

Harmonic Spectral Centroid Description
Harmonic Spectral Deviation Description
Harmonic Spectral Spread Description
Harmonic Spectral Variation Description
Spectral Centroid Description

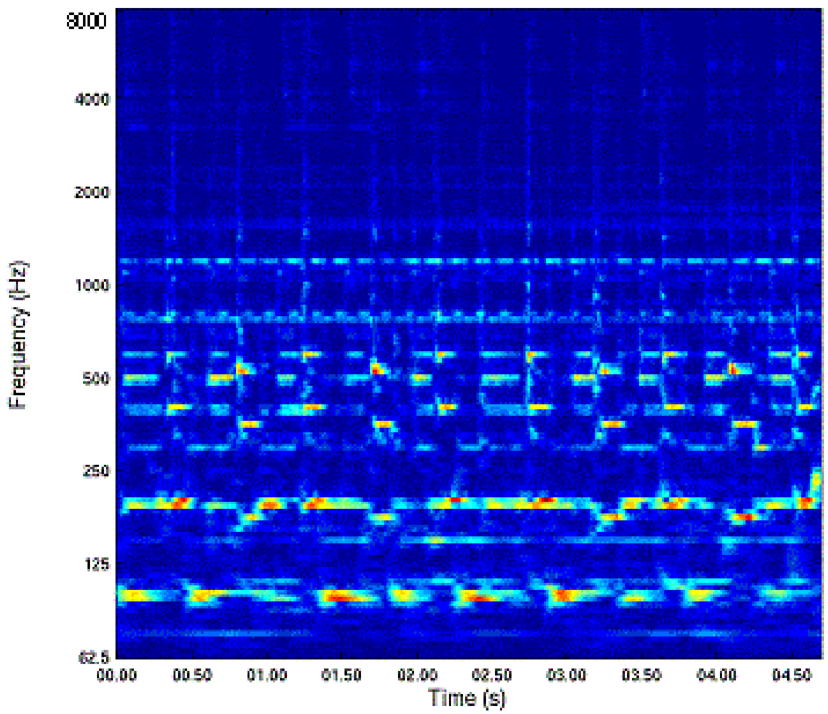
Spectral Basic

Audio Spectrum Basis Description
Audio Spectrum Projection Description

Signal Parameters

Audio Harmonicity Description
Audio Fundamental Frequency Description

AUDIO DESCRIPTORS. AUDIOSPECTRUMENVELOPE

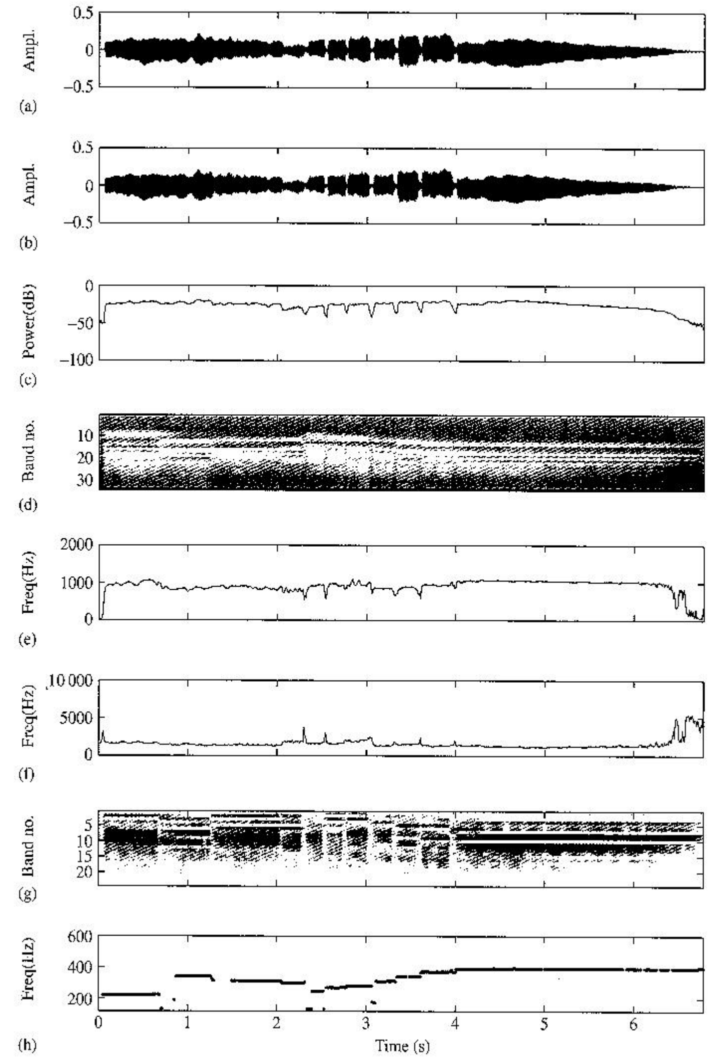


AudioSpectrumEnvelope description of a pop song. The required data storage is NM values where

N is the number of spectrum bins and

M is the number of time points

AUDIO DESCRIPTORS



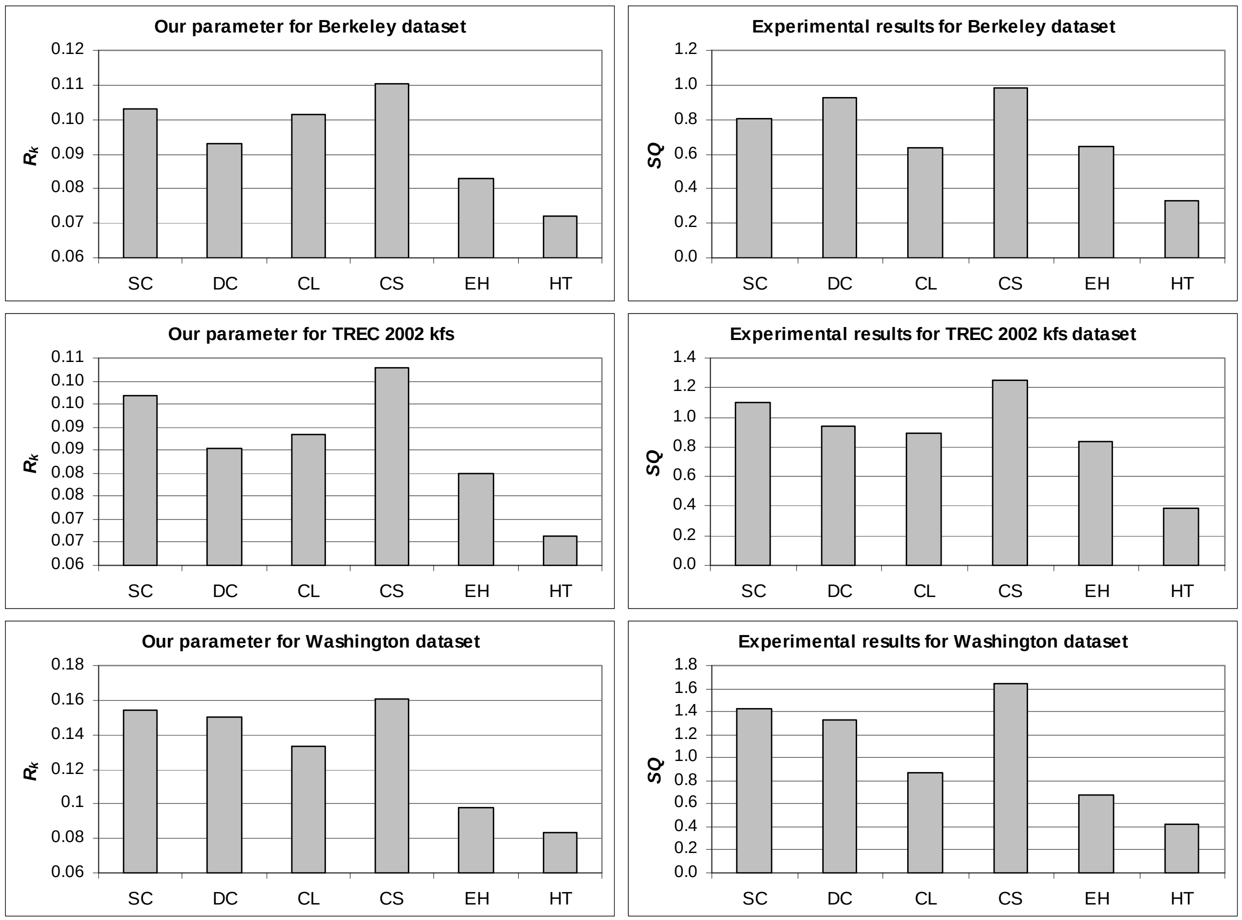
Cor anglais playing a solo excerpt from the overture *Le Carnaval romain* by Berlioz: (a) Original 44.1 kHz sampled waveform; (b) AudioWaveformType Min-Max representation; (c) AudioPowerType; (d) AudioSpectrumEnvelopeType '1/4 Octave' resolution. (e) AudioSpectrumCentroidType; (f) AudioSpectrumSpreadType; (g) AudioSpectrumFlatnessType 4-band resolution and (h) AudioFundmental FrequencyType

AUDIO DESCRIPTORS.
MUSIC STYLES “PROGRAMME”

Music of a narrative or descriptive kind. The term was introduced by Liszt, who defined a programme as “a preface added to a piece of instrumental music . . . to direct (the listener's) attention to the poetical idea of the whole or to a particular part of it”.

Programme music, which has been contrasted with Absolute Music, is distinguished by its attempt to depict objects and events. The concept is much older than Liszt. Kuhnau's six Bible sonatas (1700) are each preceded by a summary of what the music is meant to convey, and the 'programmes' of Vivaldi's 'Four Seasons' concertos are contained in sonnets appended to the music.

EVALUATING THE EFFECTIVENESS OF MPEG-7 DESCRIPTORS



Rk and *SQ* (for the compound experiments) for each dataset

MPEG 21 (ISO/IEC 21000-2)

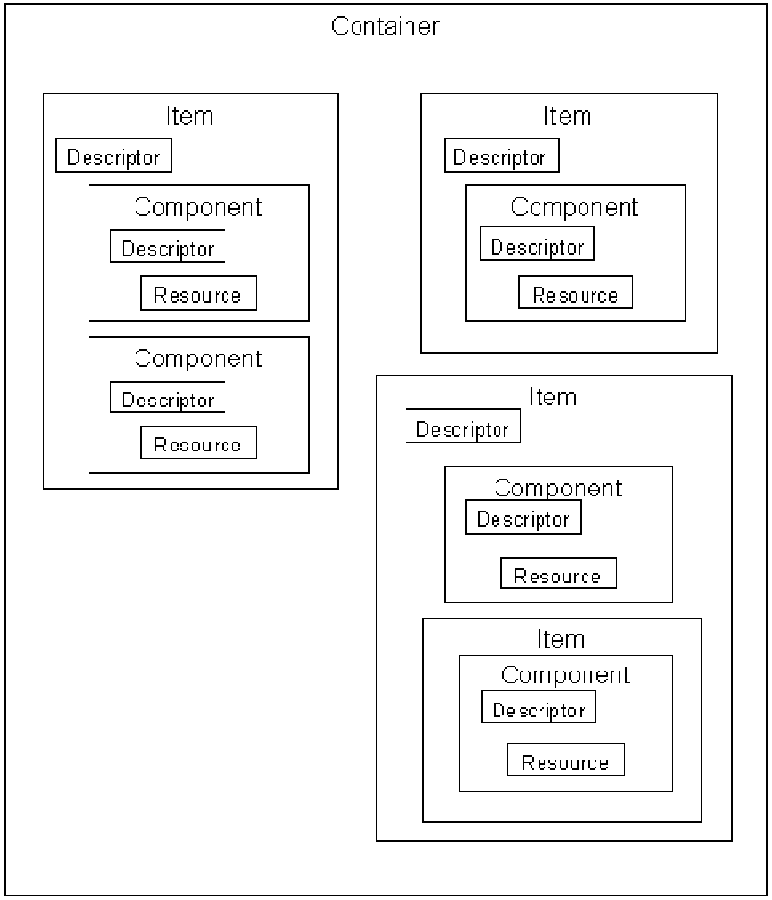
A multimedia framework “to enable transparent and augmented use of multimedia resources…”

- images (JPEG2000, GIF, PNG, ...)
- video (MPEG-4, QuickTime, ASF, ...)
- audio (WAV, MP3, ...)
- text (txt, doc, ...)

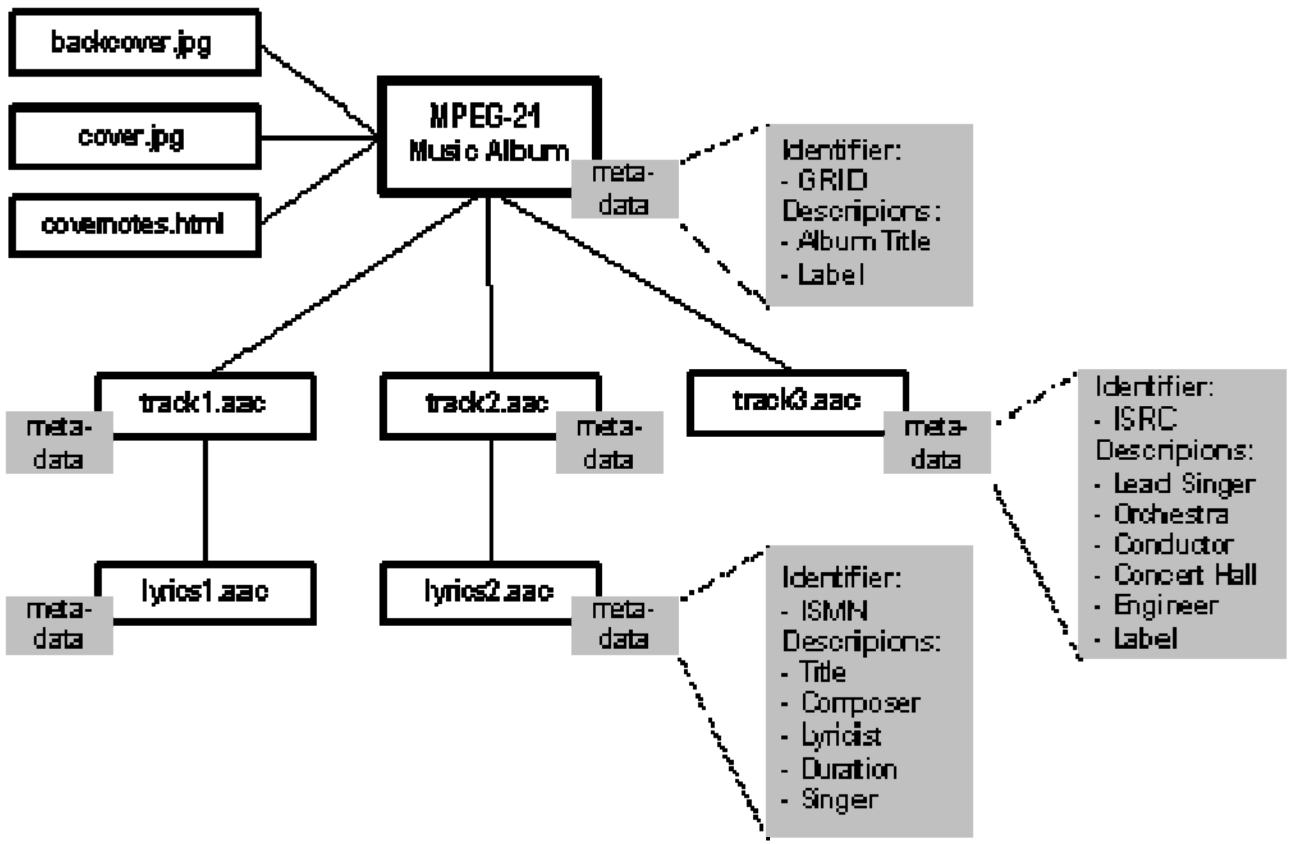
DIGITAL ITEM

- A structured digital object with a standard representation, identification and metadata
- The fundamental unit of distribution & transaction in the MPEG-21 frame
- Digital Items can be expressed using Digital Item Declaration Language (XML schema)
- Digital Item = resources + metadata + structure
  - Resource: individual asset(s)
  - Metadata: information about or pertaining to the Item
  - Structure: relationships between the parts of the Item

MPEG-21 DIGITAL IDENTIFICATION DECLARATION MODEL



MPEG-21 IDENTIFYING DIGITAL ITEMS



MPEG-21 BUILDING BLOCKS

DID technology is described in three sections

- model
  - describes set of abstract terms and concepts
  - Digital Item is the digital representation of "a work”
  - Digital Item is the thing that is acted upon within the model
  - Digital Items are managed, handled, processed, described, exchanged, collected
- representation
  - DID elements are represented in XML
  - normative description of their syntax and semantics
- schema
  - normative XML schema
  - comprising entire grammar of the DID

MPEG-21 building blocks

- Container - structure allowing items and/or containers to be grouped
- Contain Descriptor - allows for "labeling" of containers
- Item - grouping of sub-items and/or components that are bound to relevant descriptors
- Item Descriptors contain information about the item may contain choices allow items to be customized or configured may be conditional

MPEG-21 BUILDING BLOCKS

- Component - binding of a resource to its descriptors. Descriptors typically contain control/structural information for the resource
- Resource - individually identifiable asset a video or audio clip, image, textual asset must be locatable via an unambiguous address
- Anchor - binds descriptors to a fragment
- Fragment - corresponds to a specific location or range within a resource
- Statement - literal textual value contains information, but not an asset
- Choice - describes set of related selections
- Selection - describes specific decision that will affect one or more conditions somewhere within an item if chosen
- Condition - describes the enclosing element as being optional
- Predicate - can have 3 obvious possible values -true, false, undecided
- Assertion - defines a full or partially configured state of a choice

MPEG BASED SYSTEMS

- MILOS
- IBM Analysis and Retrieval Systems
- IBM with post office movie
- MyMultimediaWorld
- MPEG7

STANDARDS FOR SCENE DESCRIPTION

- VRML (Virtual Reality Modeling Language)
- X3D (eXtensible 3D)
- SMIL (Synchronized Multimedia Integration Language)
- SVG (Scalable Vector Graphics)
- MPEG-4
- Adobe Flash, Autodesk 3ds.

	VRML	X3D	SMIL	SVG	MPEG-4
Standard name	Virtual Reality Modelling Language	eXtensible 3D	Synchronised Multimedia Integration Language	Scalable Vector Graphics	Coding of audio-visual objects
Organisation	ISO	ISO	W3C	W3C	ISO
Publication Year	1997	2005	1998-2007	2003	1998-2007
Media types	3D graphics and ref. to video / audio	2D/3D graphics and ref. to video / audio	Ref. to 2D graphics, video / audio	2D graphics, text and image	Still image, video, audio, 2D/3D graphics
Animation	Linear Interpolation	Linear Interpolation	Local animation	Not relevant	Linear and higher order interpolation
Compression	No	Yes	No	No	Yes
Streaming	Only for video / audio	Only for video / audio	Only for video / audio	Not supported	Video, audio, scene graph, graphics animation
Synchronisation	Only between animation tracks	Only between animation tracks	Not supported per frame	Not relevant	Full support
Interactivity	Yes	Yes	Yes	Yes	Yes

SVG

- **Scalable.** In terms of graphics, scalable means not being limited to a single, fixed, pixel size. The same SVG graphic can be placed at different sizes on the same Web page, and re-used at different sizes on different pages. SVG graphics can be magnified to see fine detail, or to aid those with low vision
- **Vector.** Vector graphics contain geometric objects such as lines and curves. SVG gives control over the rasterisation process. SVG also provides client-side raster filter effects.
- **Graphics.** SVG fills a gap in the market by providing a rich, structured description of vector and mixed vector/raster graphics; it can be used standalone, or as an XML namespace with other grammars.

example1, example2, example3

OPEN STANDARDS FOR INTERACTIVE TV (iTV)

- Participation in a quiz show
- Additional information on news topics
- Directly buy a product presented in a commercial
- Electronic Program Guides (EPGs)
- Video on Demand (VoD)
- YouTube, MyVideo

BASIC TECHNOLOGIES FOR iTV

- Set-top Boxes. A set-top box (STB) is a device that forms a link between a TV-set and an external source such as satellite dish, antenna.
- Video Encoding and Transport. Audio and video encoding standards play a major role in the world of digital TV (DTV). MPEG-2, MPEG-4 Part 10 Advanced Video Codec (MPEG-4 AVC or H.264) is used in several recent DTV standards

DIGITAL iTV STANDARDS AND ORGANIZATIONS

- Digital Video Broadcasting Project - http://www.dvb.org/
- Advanced Television Systems Committee (ATSC) - http://www.atsc.org/
- CableLabs - http://www.cablelabs.com/
- The Association of Radio Industries and Businesses (ARIB) - http://www.arib.or.jp/english/

MHEG

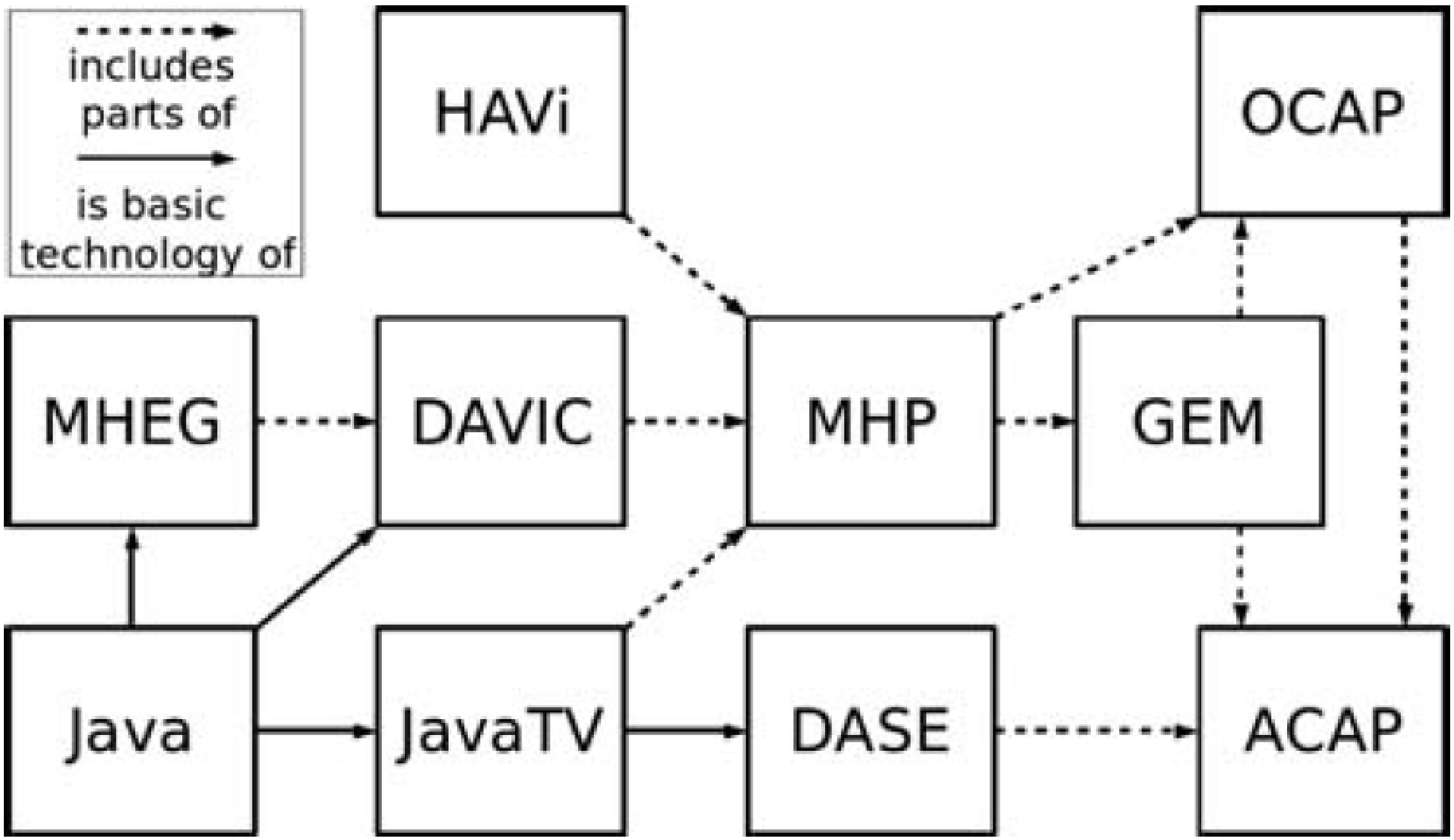
The Multimedia and Hypermedia Information Coding Expert Group (MHEG), a subgroup of the International Organization for Standardization (ISO), published the MHEG standard in 1997.

- MHEG-1: MHEG Object Representation Base Notation (ASN.1).
- MHEG-2: Should provide an encoding based on SGML instead of ASN.1 but was never finished.
- MHEG-3: MHEG Script Interchange Representation.
- MHEG-4: MHEG Registration Procedure.
- MHEG-5: Support for Base Level Interactive Applications.
- MHEG-6: Support for Enhanced Interactive Applications.
- MHEG-7: Interoperability and Conformance Testing for MHEG-5.
- MHEG-8: XML Notation for MHEG-5.

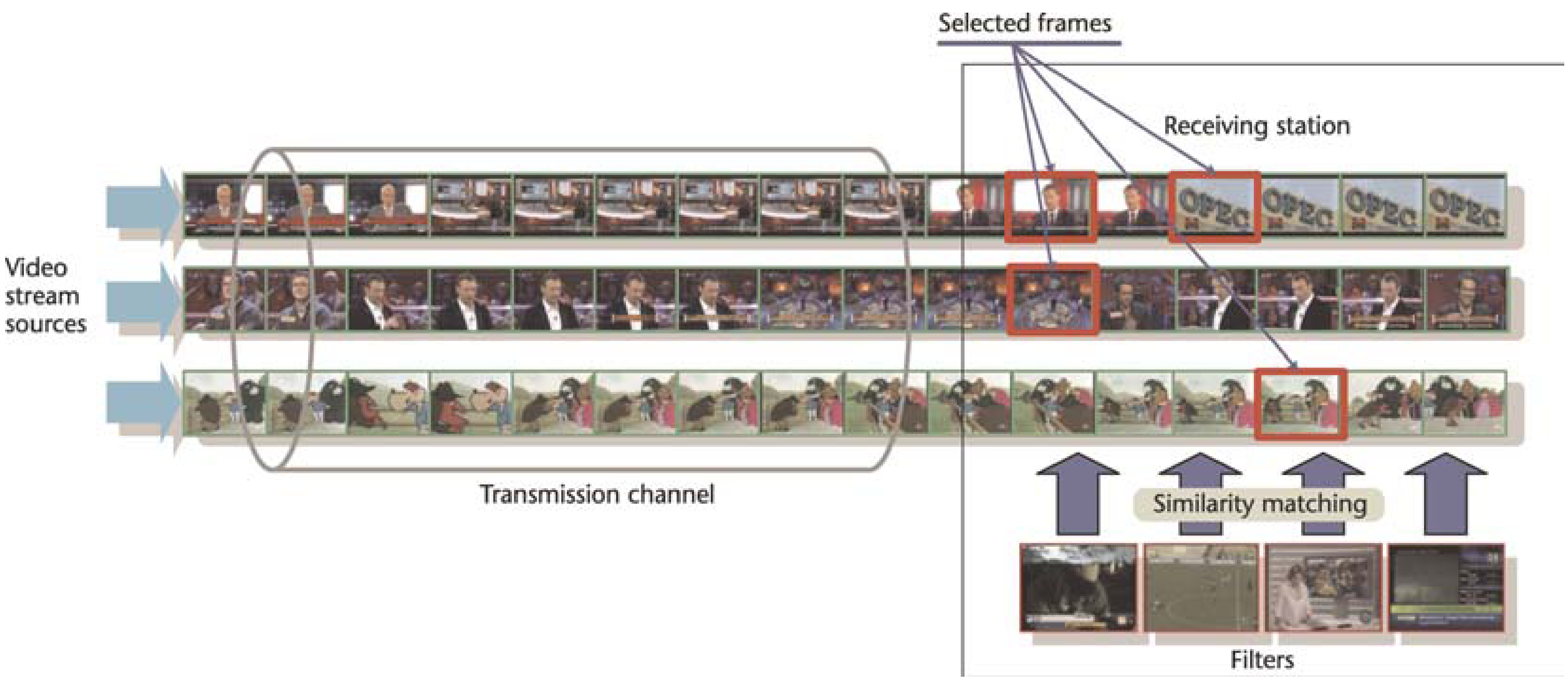
iTV STANDARTDS

- DAVIC **-** The Digital Audio-Video Council (DAVIC) was founded in 1994 and completed its work in 1999. Not all of the DAVIC specifications are used, but major parts are referenced in many other standards.
- Java TV - The Java TV API is an extension of the Java platform. The Java Media Framework (JMF), though not part of the Java TV API provides the foundation for management and control of time-based media, such as video, audio and others. Broadcast Data API provides access to different kinds of broadcast data. Java applications for digital receivers that use the Java TV API are called Xlets.
- The Multimedia Home Platform was specified by the MHP group.
  - Enhanced Broadcast Profile (Profile 1): The simplest version of an MHP environment supports the Enhanced Broadcast Profile.
  - Interactive Broadcast Profile (Profile 2): This profile includes support for a standardized return channel
  - Internet Access Profile (Profile 3): It is extended with support for Internet applications.
  - IPTV Profile (Profile 4): The most enhanced profile is the IPTV Profile. This profile integrates support for DVB-IPTV into the MHP platform.

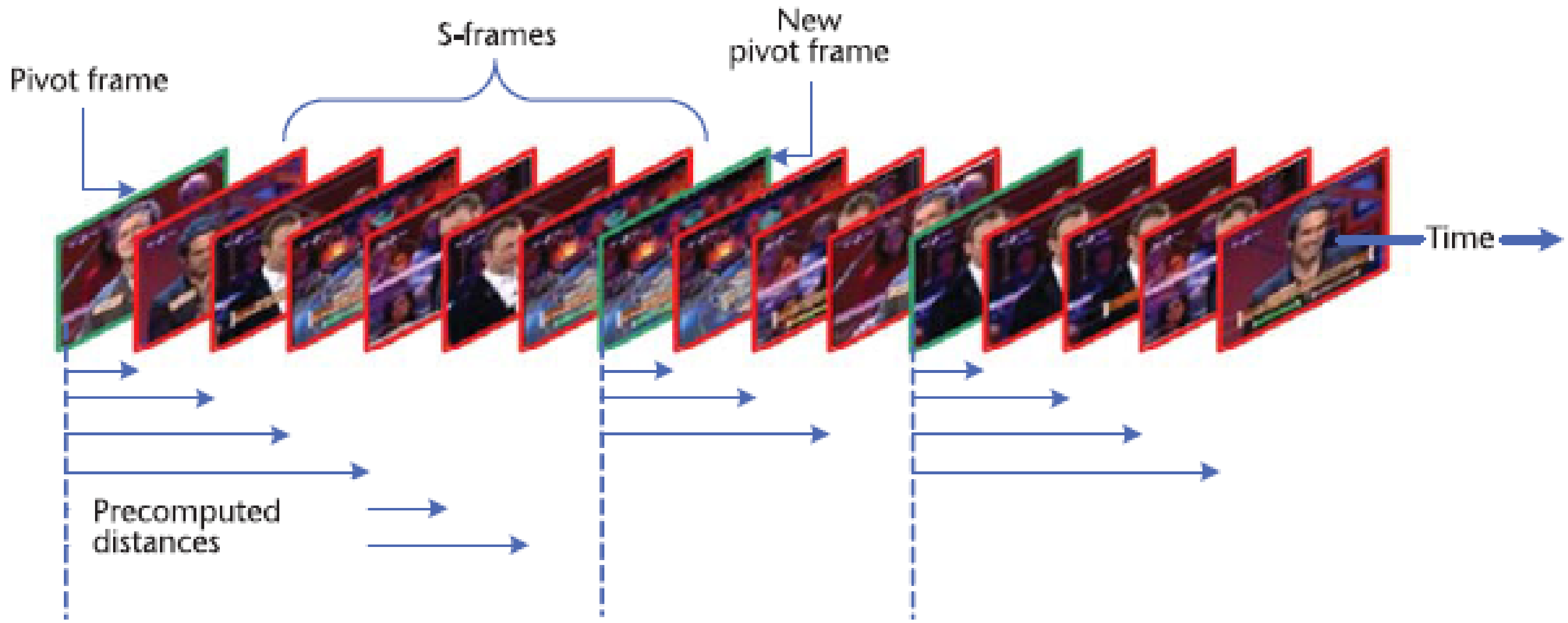
DIGITAL iTV STANDARDS



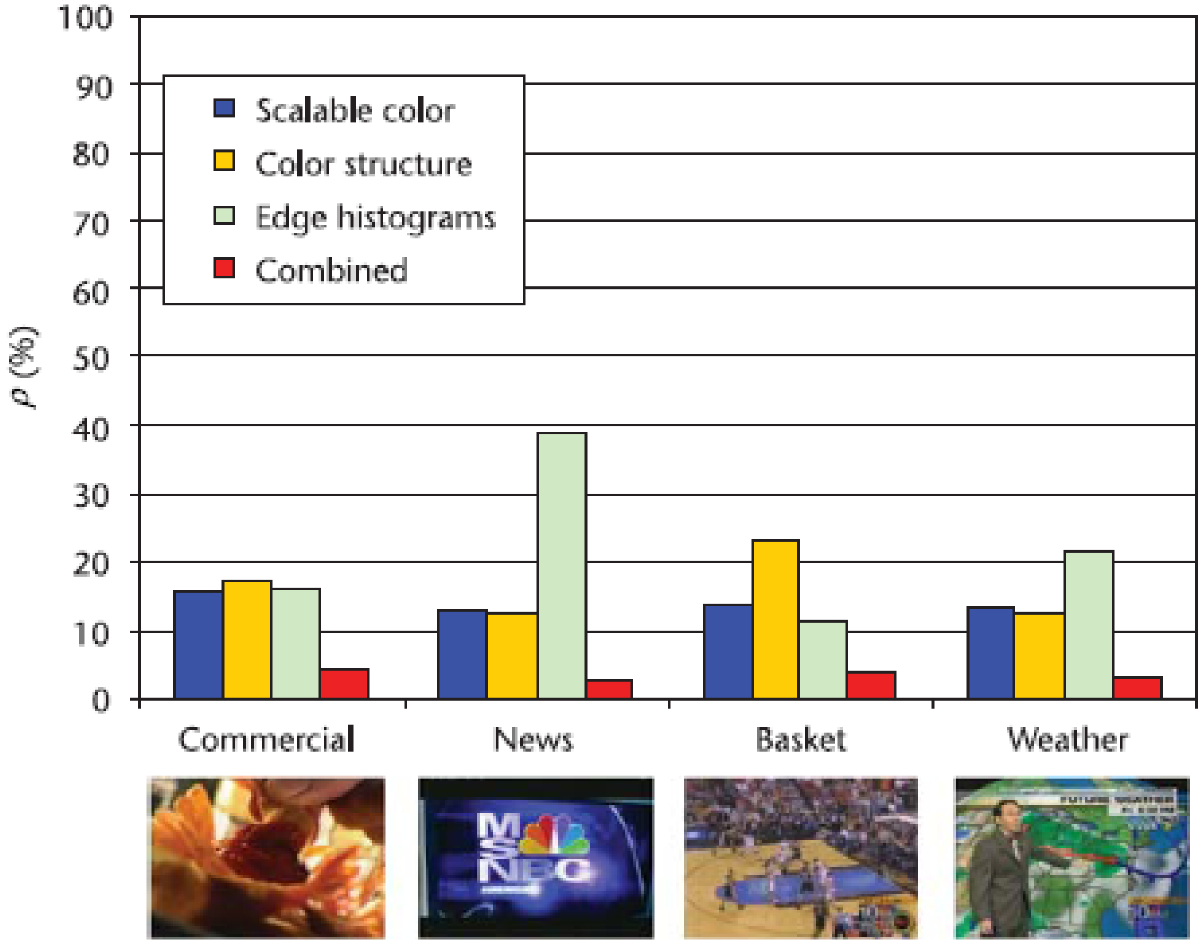
EFFICIENT VIDEO-STREAM FILTERING



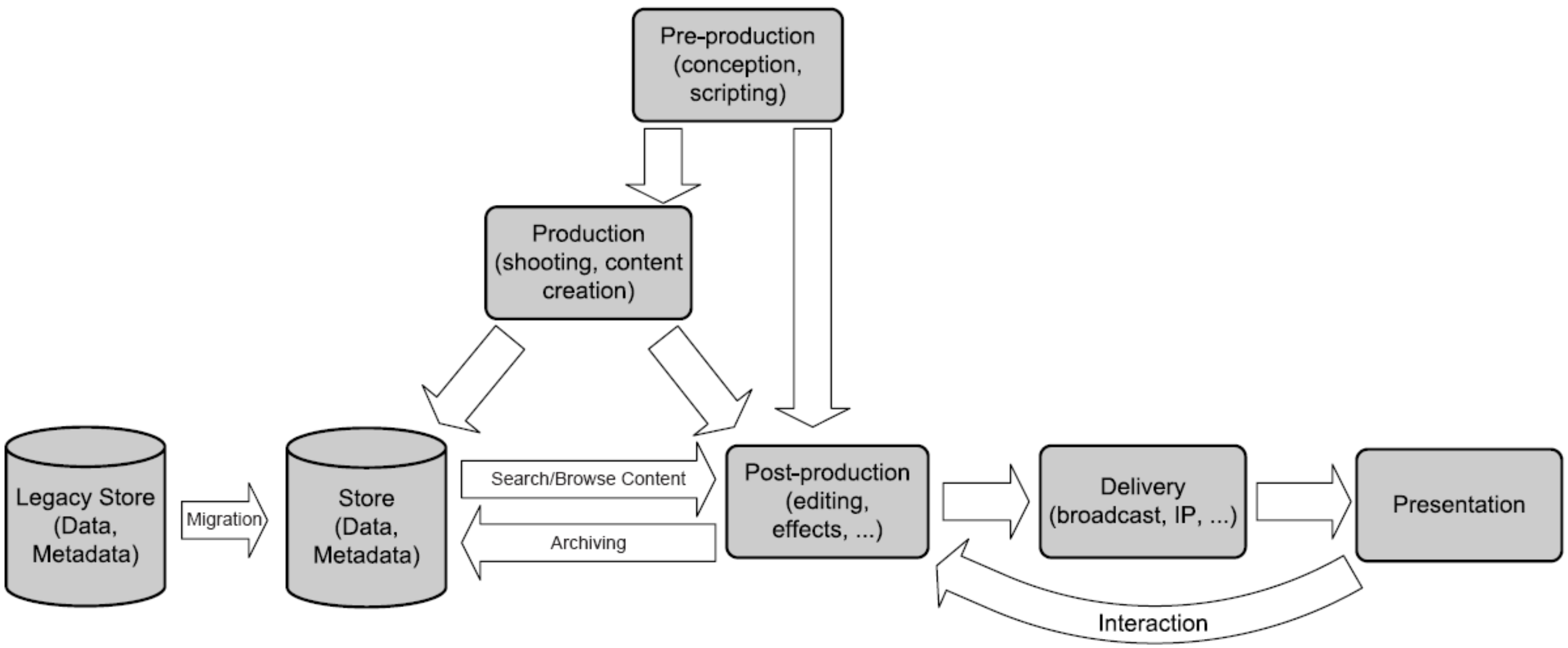
EFFICIENT VIDEO-STREAM FILTERING



EFFICIENT VIDEO-STREAM FILTERING



STANDARDS IN THE AUDIOVISUAL MEDIA



DUBLIN CORE

Used to describe digital materials such as video, sound, text or images

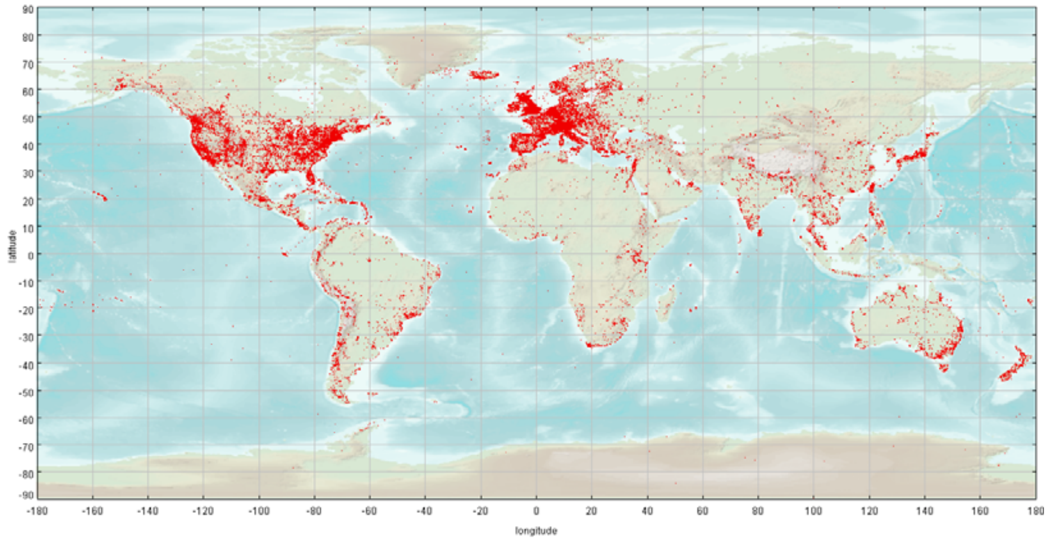
Title
Language
Publisher
Description
Identifier
Rights
Contributor
Relation
Coverage
Subject
Source
Creator
Format
Date
Type

STANDARDS IN THE AUDIOVISUAL MEDIA

- EBU P Meta. The European Broadcasting Union (EBU) has defined P Meta as a metadata vocabulary for programme exchange in the professional broadcast industry.
- Material Exchange Format. The Material Exchange Format (MXF) is a standard issued by Society of Motion Picture and Television Engineers (SMPTE), defining the specification of a file format for the wrapping and transport of essence and metadata in a single container.
- SMPTE Descriptive Metadata Scheme 1 (DMS-1), formerly know as Geneva Scheme uses metadata sets defined in the SMPTE Metadata Dictionary
- SMPTE Metadata Dictionary (is a large thematically structured list of narrowly defined metadata elements, defined by a key, the size of the value and its semantics.
- Standard Media Exchange Framework (SMEF) is a data model defined by the BBC to describe the metadata related to media items (media objects) and programmes and parts thereof (editorial objects), down to the shot level.
- Controlled Vocabulary and Ontologies. Audiovisual content descriptions often contain references to semantic entities such as objects, events, states, places, and times.

WEB 2.0

- Web 1.0 focused on a relatively small number of companies and advertisers producing content = “brochure web”
- Web 2.0 involves the user. Web 1.0 is as a lecture and Web 2.0 is a conversation.
- Many Web 2.0 companies are built almost entirely on user-generated content and harnessing collective intelligence.
- Google, MySpace, Flickr, YouTube and Wikipedia, users create the content, while the sites provide the platforms.


latitude
longitude

WEB 2.0

- Key tools for building RIAs = Adobe's Flex, Microsoft's Silverlight, ASP.NET, Ajax and Sun's JavaServer Faces
- Other Web development tools and technologies = Adobe's Dreamweaver, JSON, the web servers IIS and Apache, MySQL, PHP and ASP.NET
- Web services allow you to incorporate functionality from existing applications into your own applications quickly and easily.
  - Amazon Web Services
  - Google Maps web services
  - eBay web services

XML BASICS

- XML documents are readable by both humans and machines
- XML permits document authors to create custom markup for any type of information
  - Can create entirely new markup languages that describe specific types of data, including mathematical formulas, chemical molecular structures, music and recipes
- An XML parser is responsible for identifying components of XML documents (typically files with the .xml extension) and then storing those components in a data structure for manipulation

XML BASICS

- An XML document can optionally reference a Document Type Definition (DTD) or schema that defines the XML document’s structure
- An XML document that conforms to a DTD/schema (i.e., has the appropriate structure) is valid
- If an XML parser (validating or nonvalidating) can process an XML document successfully, that XML document is well-formed

DOCUMENT TYPE DEFINITIONS (DTDS)

- DTDs and schemas specify documents' element types and attributes, and their relationships to one another
- DTDs and schemas enable an XML parser to verify whether an XML document is valid (i.e., its elements contain the proper attributes and appear in the proper sequence)
- A DTD expresses the set of rules for document structure using an EBNF (Extended Backus-Naur Form) grammar
- In a DTD, an ELEMENT element type declaration defines the rules for an element. An ATTLIST attribute-list declaration defines attributes for a particular element

XML VOCABULARIES

- Some XML vocabularies
  - MathML (Mathematical Markup Language)
  - Scalable Vector Graphics (SVG)
  - Wireless Markup Language (WML)
  - Extensible Business Reporting Language (XBRL)
  - Extensible User Interface Language (XUL)
  - Product Data Markup Language (PDML)
  - W3C XML Schema
  - Extensible Stylesheet Language (XSL)

RSS

- RSS stands
  - RDF (Resource Description Framework) Site Summary
  - Also known as Rich Site Summary and Really Simple Syndication
- An XML format used to syndicate simple website content
  - news articles, blog entries, product reviews, podcasts, vodcasts and more
- RSS feed contains
  - RSS root element with a version attribute
  - channel child element with item subelements
  - Depending on the RSS version, the channel and item elements have certain required and optional child elements

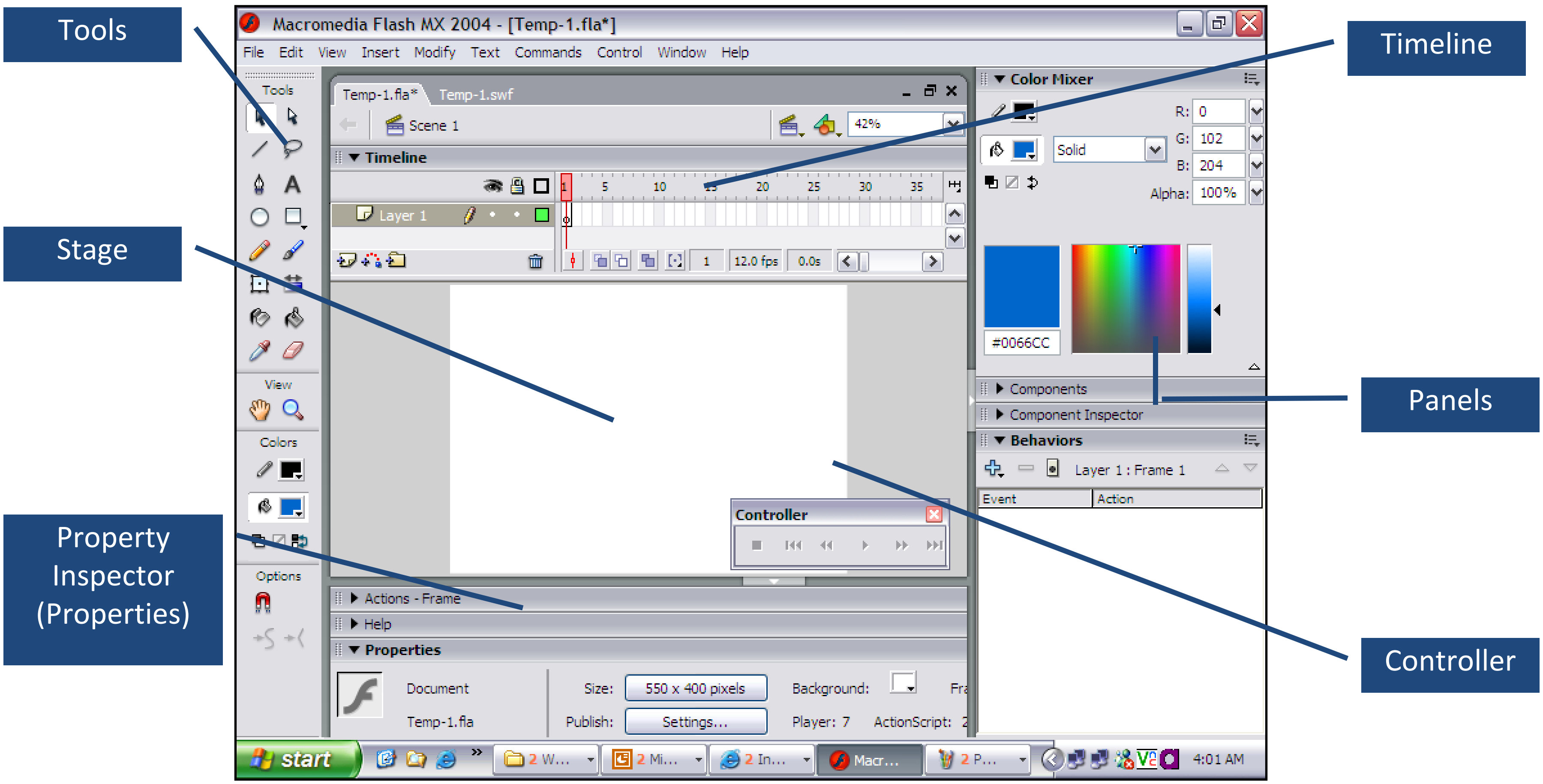
RSS

- item elements
  - provide the feed subscriber with a link to a web page or file, a title and description of the page or file
- Enables website developers to draw more traffic.
- Required child elements of channel in RSS 2.0
  - Description
  - Link
  - Title
- Required child elements of item in RSS 2.0
  - Title or description

RSS

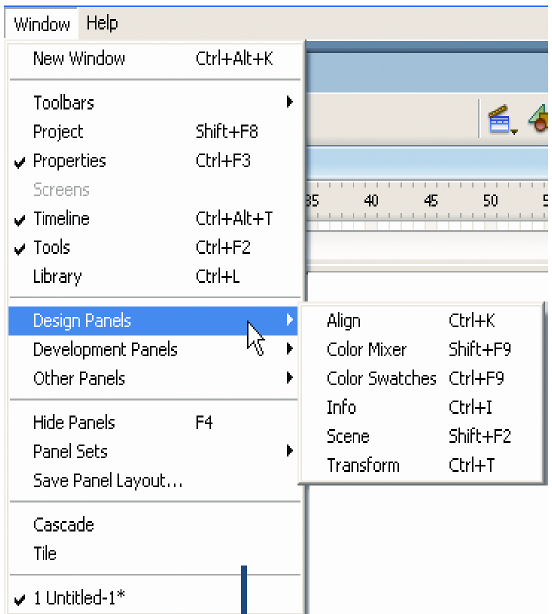
- RSS aggregator
  - keeps tracks of many RSS feeds
  - brings together information from the separate feeds
- Many sites provide RSS feed validators
  - validator.w3.org/feed
  - feedvalidator.org
  - www.validome.org/rss-atom/
- The DOM and XSL can be used to create RSS aggregators

FLASH WORK SPACE



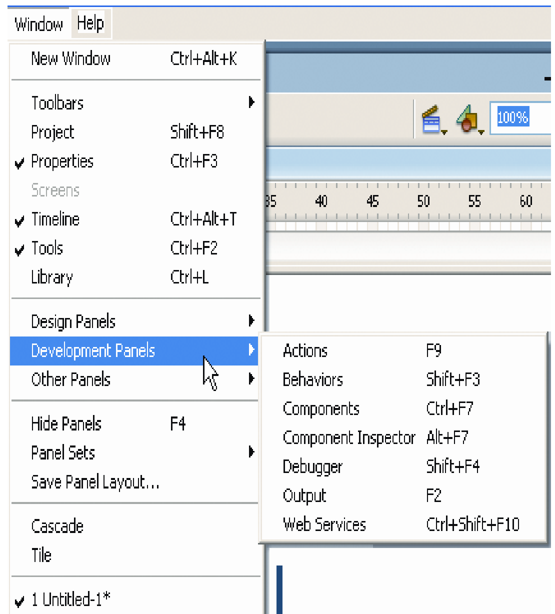
FLASH. EXPLORING PANEL SETS

- Design



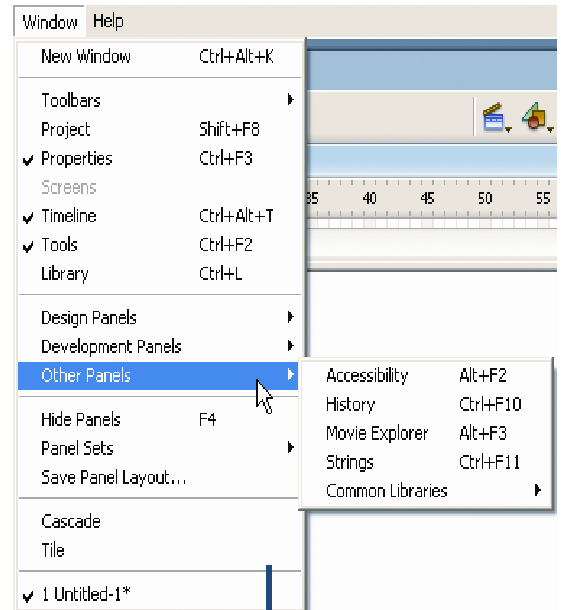
Arrange Scenes, Color and Transform Objects

- Develop



ActionScript, behavior control, & build components

- Other



ActionScript, behavior control, and build components

FUTURE OF THE WEB

- Resource Description Framework (RDF)
  - Based on XML
  - Used to describe content in a way that is understood by computers
  - Connects isolated databases across the web with consistent semantics

SEMANTIC WEB TECHNOLOGY
ONTOLOGY

- Before Aristotle: study of the nature of being
- Since Aristotle: study of knowledge representation and reasoning
- Ontologies
  – Ways of organizing and describing related items, and are used to represent semantics.
  – Another way of cataloging the Internet

CONCLUSIONS.
THE WORLD OF MASHUPS

- Mashups can search a database for items that have known addresses
- http://tutorlinker.com/
- http://findnearby.net/
- http://www.opensecrets.org/travel/index.php
- http://www.cogandsprocket.com/adgen/adgen.html

CONCLUSIONS. FUTURE OF THE WEB

- Tagging and Folksonomies
  - Early hints a “web of meaning”
  - “loose” classification system
- Semantic Web
  - Next generation in web development,
  - “web of meaning”
  - Depends heavily on XML and XML-based technologies
- Microformats
  - Standard formats for representing information aggregates that can be understood by computers, enabling better search results and new types of applications

Multimedia, Quo vadis?

Univ. Prof. DI Dr. Laszlo Böszörmenyi
Klagenfurt University, Austria

Acknowledgment

This work is partially financed by Bulgarian National Science Fund under the project D002-308/19.12.2008 "Automated Metadata Generating for e-Documents Specifications and Standards".

REFERENCES

Tutorials

- Stanchev P., Image Data Bases, ACM Summer School on Multimedia, Albena, Bulgaria, 22-28 Sept 1997.
- Stanchev P., Image Data Models, Third International Conference, Visual Information and Information Systems, Amsterdam, The Netherlands, June 1999.
- Stanchev P., Introduction to MPEG 7: Multimedia Content Description Interface, 7-th IASTED International Conference on Computer Graphics and Imaging, CGIM 2004, Kauai, Hawaii, 2004
- Stanchev P., Introduction to MPEG 7: Multimedia Content Description Interface, International Conference on "Visualization, Imaging and Image Processing (VIIP 2004), September 6-8, 2004, Marbella, Spain
- Stanchev P., Semantic Video and Image Retrieval - State of the Art and Challenges, ACMSE 2007, The 45th ACM Southeast Conference, Winston-Salem, North Carolina, USA, March 24, 2007, 512-513

REFERENCES

- Multimedia Semantics - The Role of Metadata. Series: Studies in Computational Intelligence , Vol. 101, Granitzer, Michael; Lux, Mathias; Spaniol, Marc (Eds.) 2008, XII, 262 p. 73 illus., Hardcover ISBN: 978-3-540-77472-3
- (Harvey & Paul) Deitel & Associates, Internet & World Wide Web: How to Program, 4/E, Prentice Hall, 2008, ISBN-10: 0131752421, ISBN-13: 9780131752429
- Steffen Staab: Ontology Learning, Inaugural workshop of the language, interaction and computation lab, May 29, 2007, Rovereto, Italy
- http://www.chiariglione.org/mpeg/tutorials/seminars/mp21-2005/02_MPEG-21%20DID%20tutorial.pdf
- http://www.mpeg.org/
- http://www.mheg.org/users/mheg/index.php
- http://www.go2web20.net/
- http://www.w3.org/

REFERENCES

- F. Falchi, C. Gennaro, P. Savino, P. Stanchev, "Efficient Video Stream Filtering", IEEE Multimedia, January – March 2008, 52-61
- P. Stanchev, D. Green Jr., B. Dimitrov, Semantic Notation and Retrieval in Art and Architecture Image Collections, Journal of Digital Information Management, Vol. 3, No. 4, December 2005, 218-221
- K. Ivanova, P. Stanchev, B. Dimitrov, Analysis of the Distributions of Color Characteristics in Art Painting Images, Serdica J. Computing 2 (2008), 111-136
- Stanchev P., Amato G., Falchi F., Gennaro C., Rabitti F., Savino P., "Selection of MPEG-7 Image Features for Improving Image Similarity Search on Specific Data Sets", 7-th IASTED International Conference on Computer Graphics and Imaging, CGIM 2004, Kauai, Hawaii, 2004, 395-400.

The First International Conference on Advances in Multimedia
MMEDIA 2009
July 20-25, 2009 - Colmar, France

MULTIMEDIA STANDARDS

Peter L. Stanchev

Kettering University, Flint, Michigan 48504, USA

Institute of Mathematics and Informatics, BAS, Bulgaria

pstanche@kettering.edu