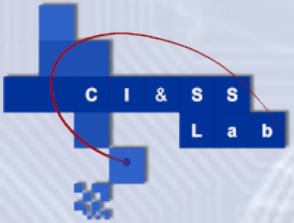


UNIVERSITÀ PARTHENOPE
DiST
DIPARTIMENTO DI SCIENZE E TECNOLOGIE

Intelligent Signal Processing

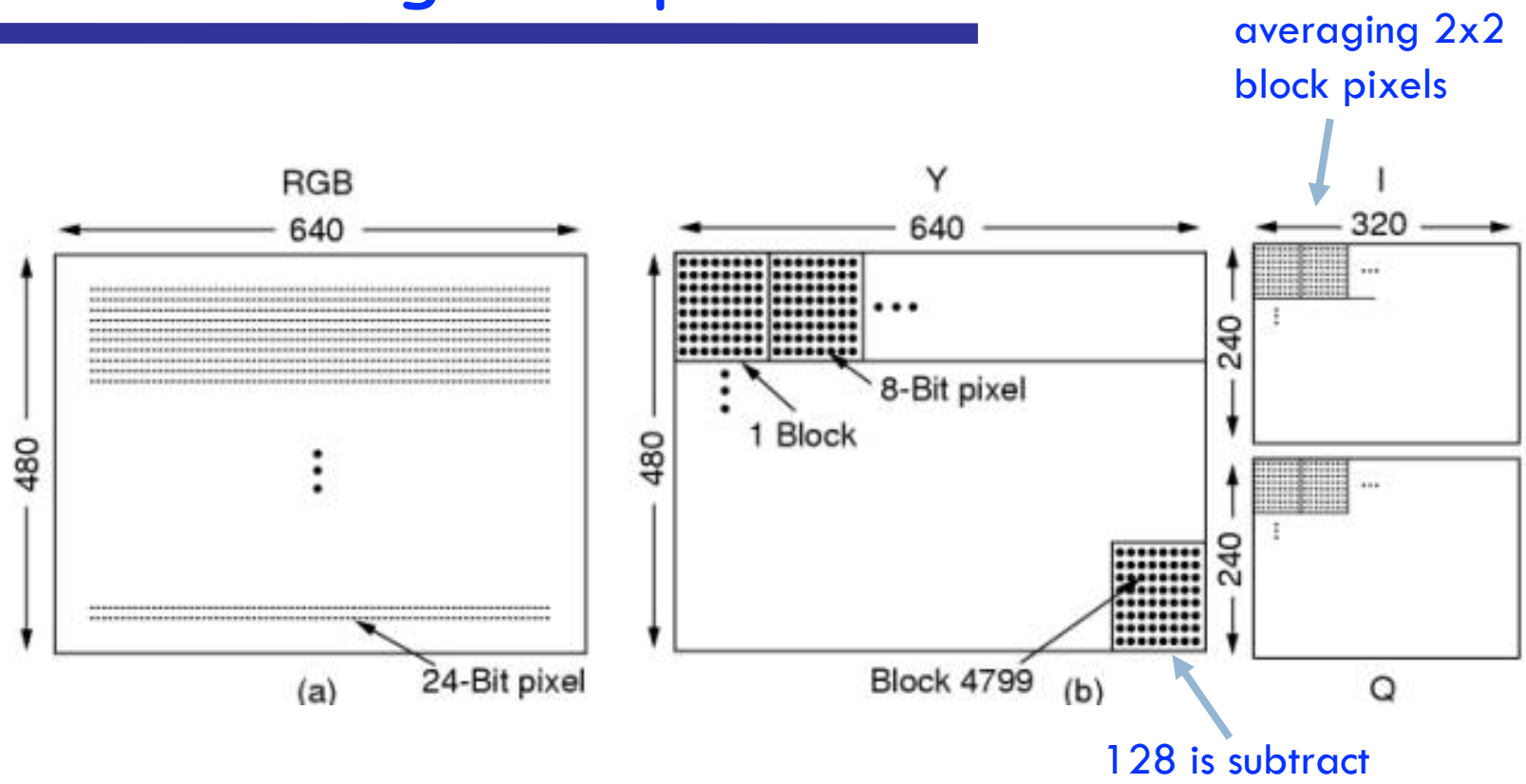
Video Compression

Angelo Ciaramella

JPEG standard

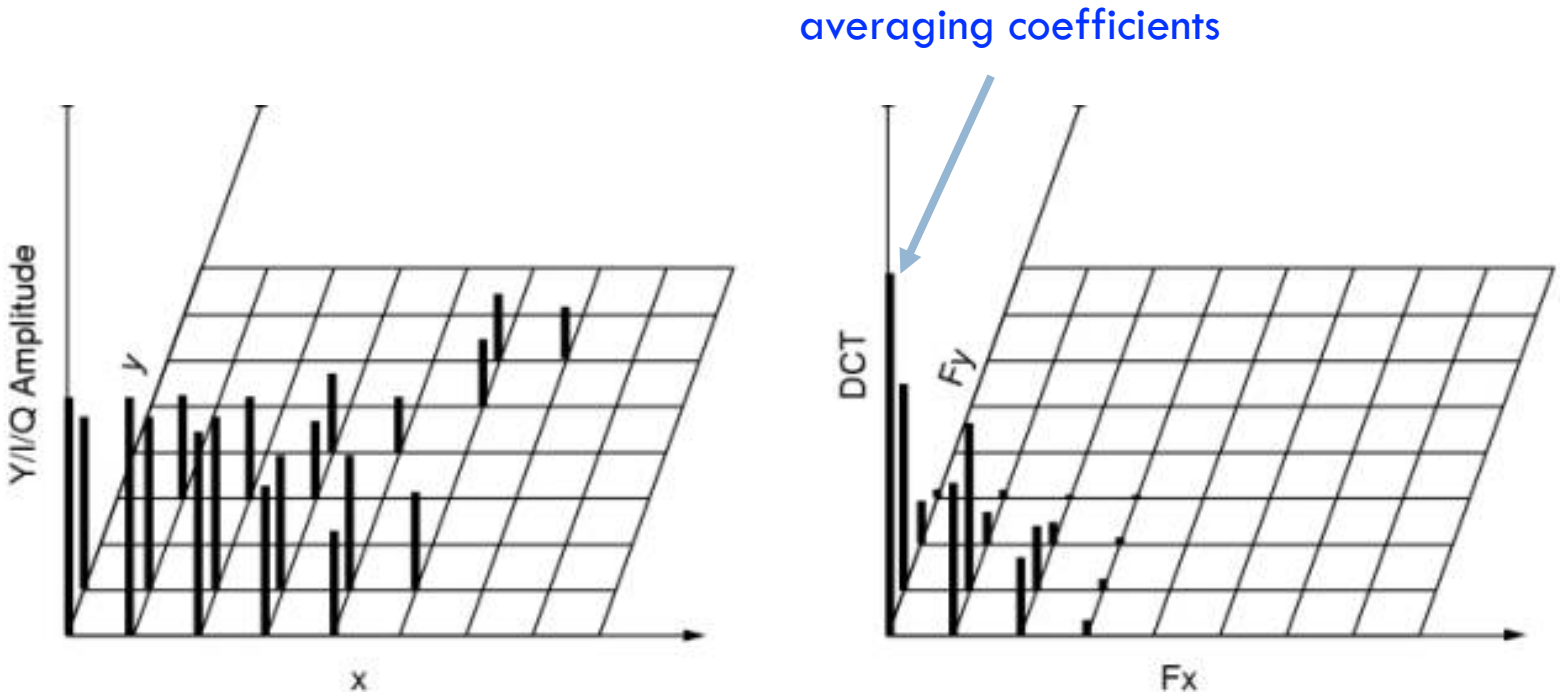
- JPEG (Joint Photographic Expert Group)
  - developed by experts on behalf of the ISO-IEC
  - International Standard 10918
  - lossy compression for digital images
    - images produced by digital photography
  - degree of compression can be adjusted
    - tradeoff between storage size and image quality.
    - typically achieves 10:1 compression with little perceptible loss in image quality

JPEG encoding – step 1



JPEG encoding – YIQ Block preparation

JPEG encoding – step 2



JPEG encoding – DCT coefficients

JPEG encoding – step 3

DCT Coefficients

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantized coefficients

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantization table

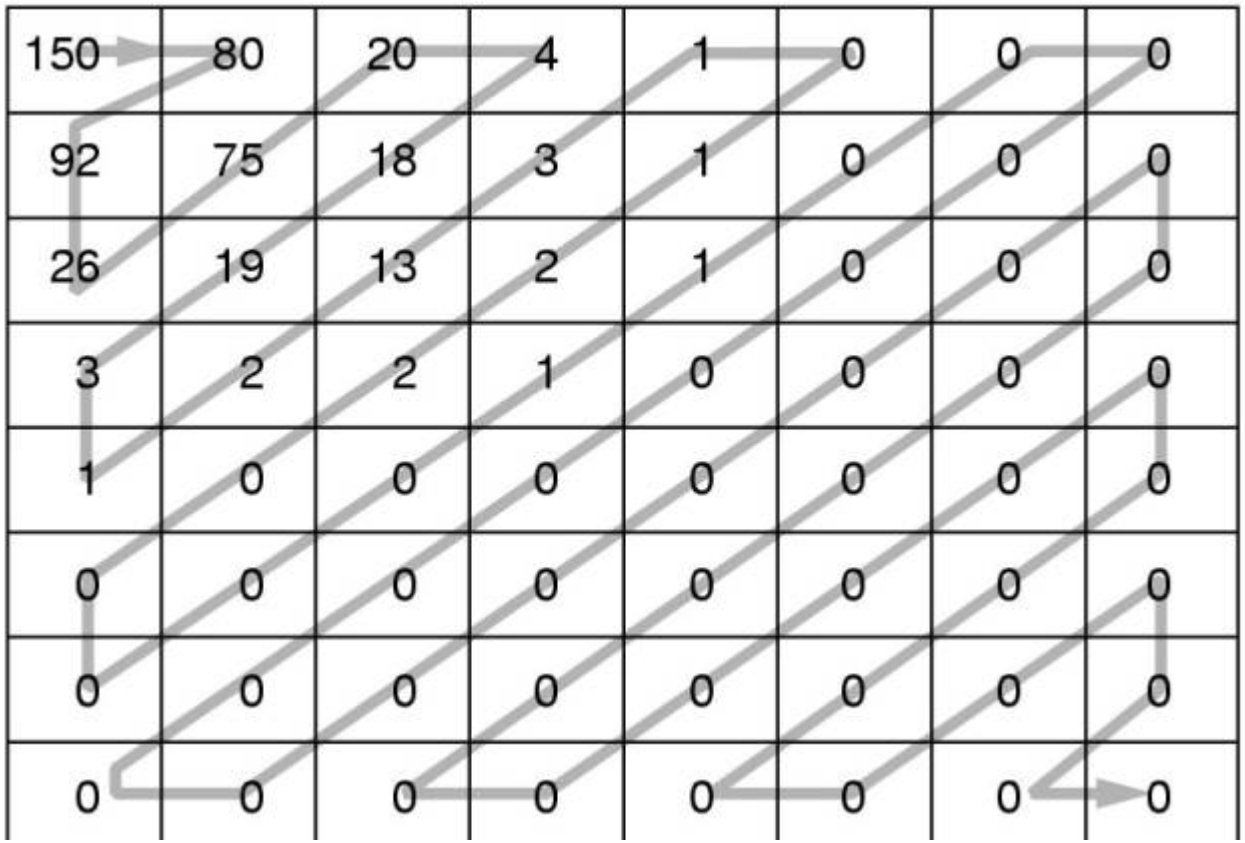
1	1	2	4	8	16	32	64
1	1	2	4	8	16	32	64
2	2	2	4	8	16	32	64
4	4	4	4	8	16	32	64
8	8	8	8	8	16	32	64
16	16	16	16	16	16	32	64
32	32	32	32	32	32	32	64
64	64	64	64	64	64	64	64

JPEG encoding – quantization

JPEG encoding - step 4

- The coefficient (0,0) is substituted by the difference with the same coefficient of the adjacency matrix
  - a low value since the coefficients are similar

JPEG encoding – step 5



JPEG encoding – matrix linearization

RLE is used

JPEG encoding - step 6

- A Huffman encoding scheme is used
- Decoding is obtained by inverting the steps

DV standard

- DV standard
  - each frame is encoded with JPEG
  - *high compression rate*

Source	Mbps	GB/ora
MPEG-2 (640x480)	4	1.76
DV (720x480)	25	11

MPEG

- Moving Picture Experts Group (MPEG)
  - working group of authorities that was formed by ISO and IEC
  - standards for audio and video compression and transmission
  - established in 1988 by the initiative of
    - Hiroshi Yasuda (Nippon Telegraph and Telephone)
    - Leonardo Chiariglione
  - The first meeting was in May 1988 in Ottawa, Canada

MPEG standards

- MPEG-1 (1993)
  - *Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s*
  - ISO/IEC 11172
- MPEG-2 (1995)
  - *Generic coding of moving pictures and associated audio information*
  - ISO/IEC 13818
- MPEG-3
  - *standardizing scalable and multi-resolution compression*
  - intended for HDTV compression
  - was merged with MPEG-2

MPEG standards

- MPEG-4 (1998)
  - *Coding of audio-visual objects*
  - ISO/IEC 14496
- MPEG-7 (2002)
  - *Multimedia content description interface*
  - ISO/IEC 15938
- MPEG-21 (2001)
  - *Multimedia framework*
  - ISO/IEC 21000

MPEG standards

- MPEG-A (2007)
  - *Multimedia application format*
  - ISO/IEC 23000
  - e.g., MPEG music player application format, MPEG photo player application

- MPEG-B (2006)
  - *MPEG systems technologies*
  - ISO/IEC 23001
  - e.g., Binary MPEG format for XML, Fragment Request Units, Bitstream Syntax Description Language (BSDL)

MPEG standards

- MPEG-C (2006)
  - *MPEG video technologies*
  - ISO/IEC 23002
  - e.g., accuracy requirements for implementation of integer-output 8x8 inverse discrete cosine transform

- MPEG-D (2007)
  - *MPEG audio technologies*
  - ISO/IEC 23003
  - e.g., MPEG Surround, SAOC-Spatial Audio Object Coding and USAC-Unified Speech and Audio Coding

MPEG standards

- MPEG-E (2007)
  - *Multimedia Middleware*
  - ISO/IEC 23004
  - e.g., Architecture, Multimedia application programming interface (API), Component model

- MPEG-V (2011)
  - *Media context and control*
  - ISO/IEC 23005
  - e.g., Avatar characteristics, Sensor information, Architecture

MPEG standards

- MPEG-M (2010)
  - *MPEG eXtensible Middleware* (MXM)
  - ISO/IEC 23006
  - e.g., MXM architecture and technologies, API, MPEG extensible middleware (MXM) protocols
- MPEG-U (2010)
  - *Rich media user interfaces*
  - ISO/IEC 23007
  - e.g., Widgets

MPEG standards

- MPEG-H (2013)
  - *High Efficiency Coding and Media Delivery in Heterogeneous Environments*
  - ISO/IEC 23008
  - Part 1 – MPEG media transport; Part 2 – High Efficiency Video Coding; Part 3 – 3D Audio
- MPEG-DASH (2012)
  - *Information technology – Dynamic adaptive streaming over HTTP* (DASH)
  - ISO/IEC 23009
  - Media presentation description and segment formats

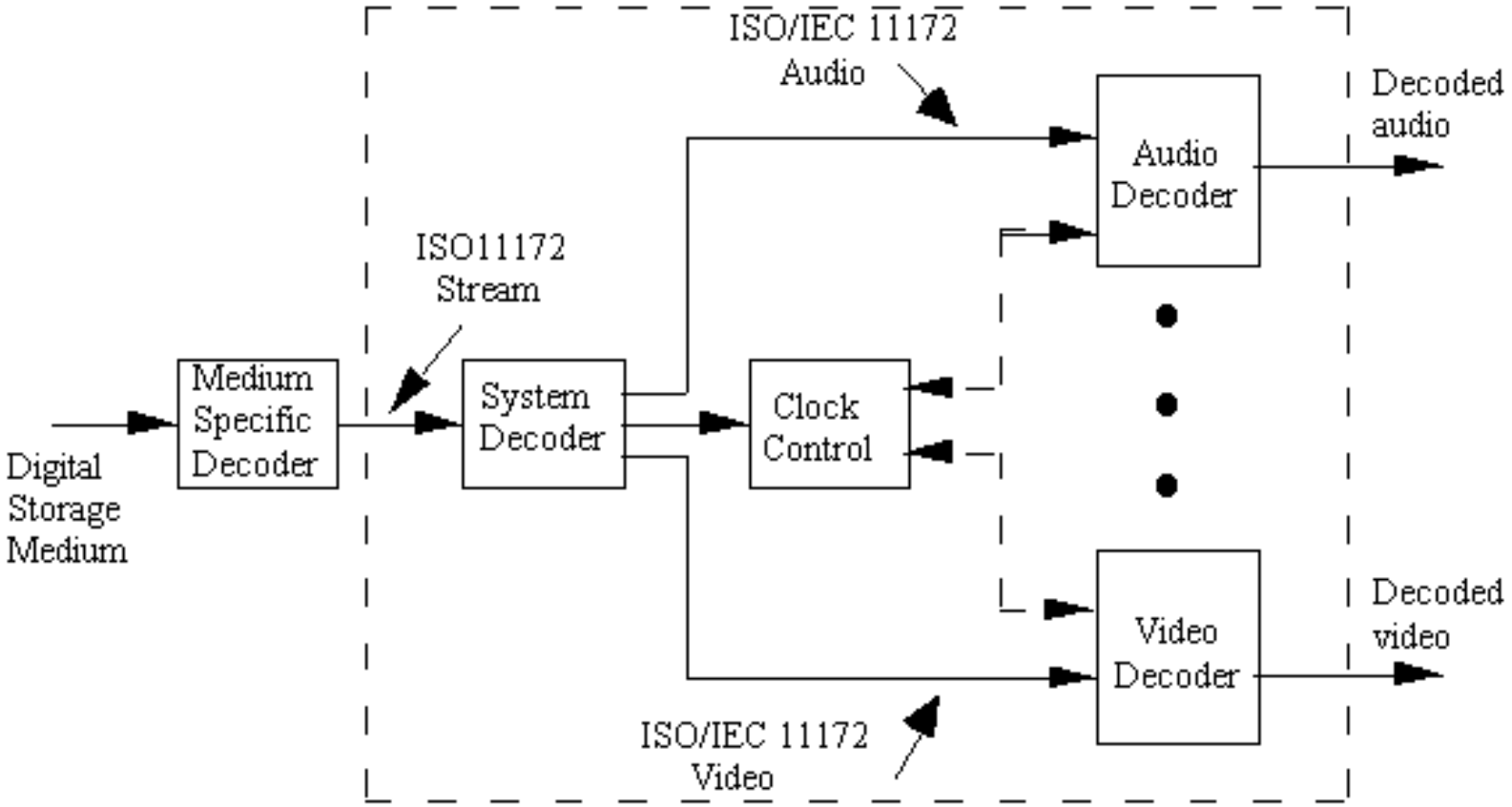
MPEG-1

- MPEG-1
  - standard for lossy compression of video and audio
  - designed to compress VHS-quality raw digital video and CD audio down to 1.5 Mbit/s (26:1 and 6:1 compression ratios respectively)
  - without excessive quality loss
    - video CDs
    - digital cable/satellite TV
    - digital audio broadcasting (DAB)

MPEG-1

- The standard consists of five Parts
  - ISO/IEC 11172-1 (1993)
    - System
  - ISO/IEC 11172-2 (1993)
    - Video
  - ISO/IEC 11172-3 (1993)
    - Audio
  - ISO/IEC 11172-4 (1995)
    - Compliance Testing
  - ISO/IEC TR 11172-5 (1998)
    - Software simulation

MPEG-1 - System

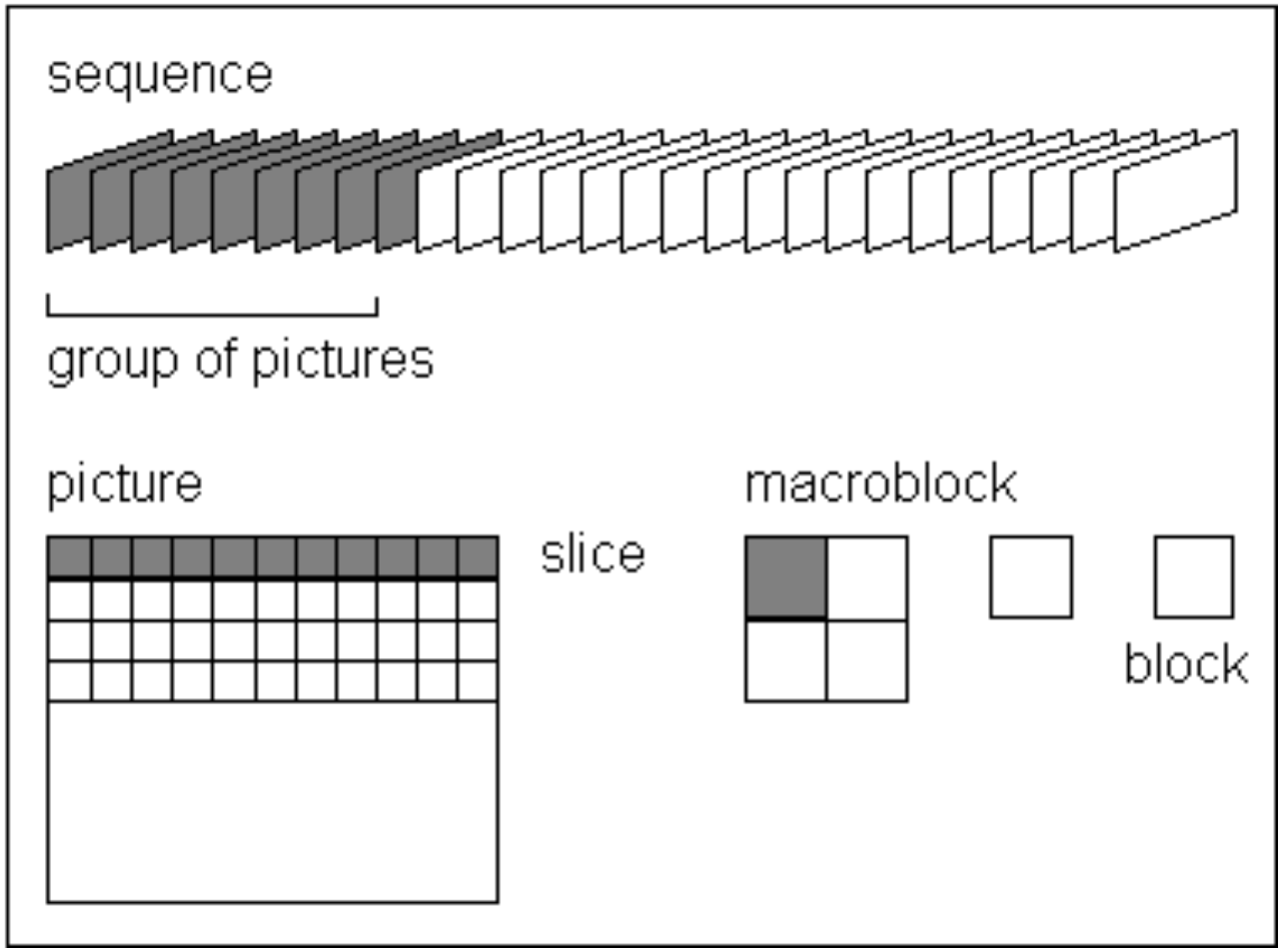


ISO/IEC 11172-1: System

MPEG-1 - Video

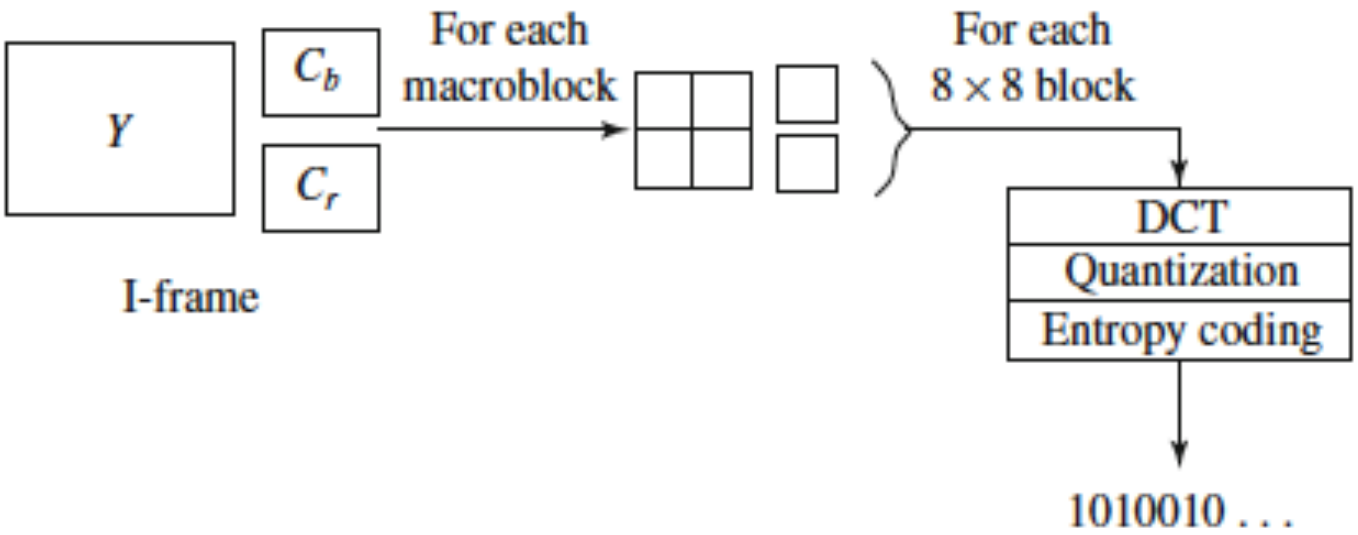
- MPEG-1 has several frame/picture types
  - I-frame (Intra-frame)
    - decoded independently of any other frames
    - can be considered effectively identical to baseline JPEG images
    - also in H.261 encoding standard
  - P-frame (Predicted-frame)
    - also be called forward-predicted frames
    - improve compression by exploiting the temporal redundancy in a video
    - store only the *difference* in image from the frame (either an I-frame or P-frame) immediately preceding it (*anchor frame*)
    - the difference between a P-frame and its anchor frame is calculated using *motion vectors* on each *macroblock* of the frame
    - Motion vector data will be embedded in the P-frame for use by the decoder
    - also in H.261 encoding standard

MPEG-1 - Video



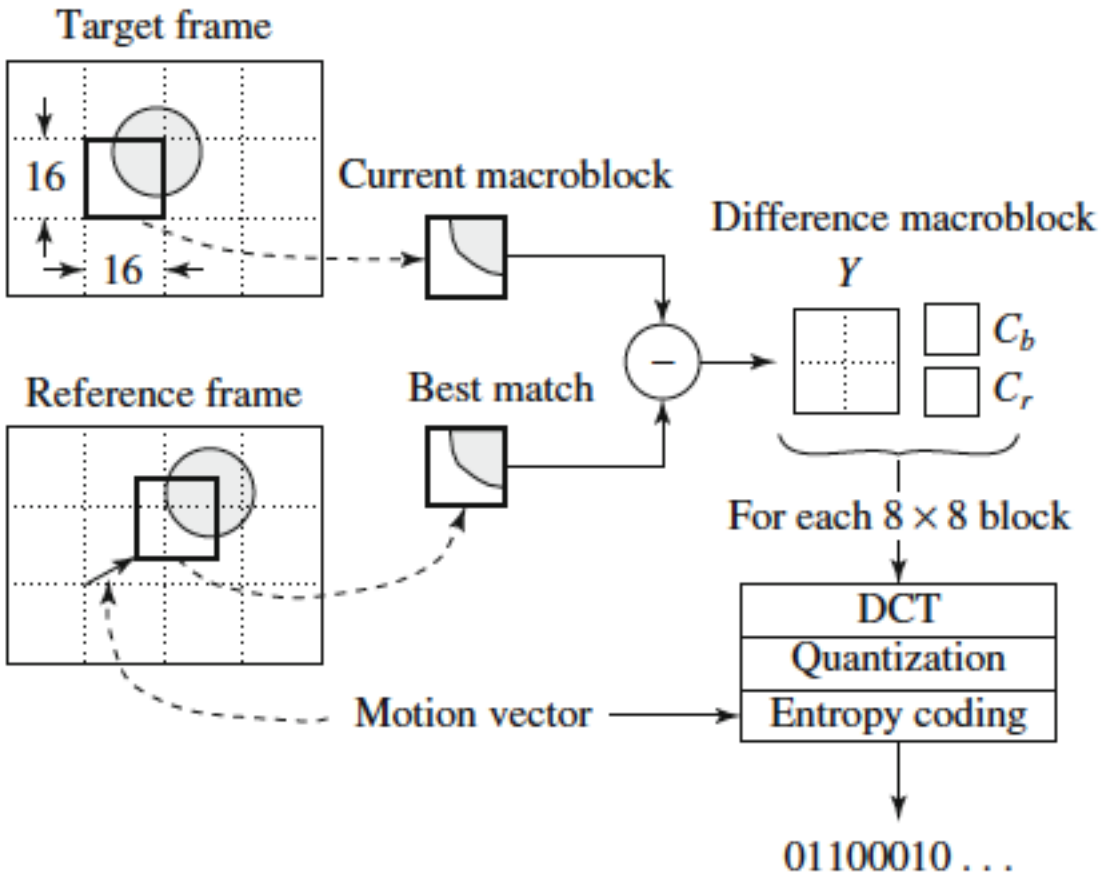
Sequence of pictures and macroblocks

MPEG-1 - Video



I-frame coding

MPEG-1 - Video

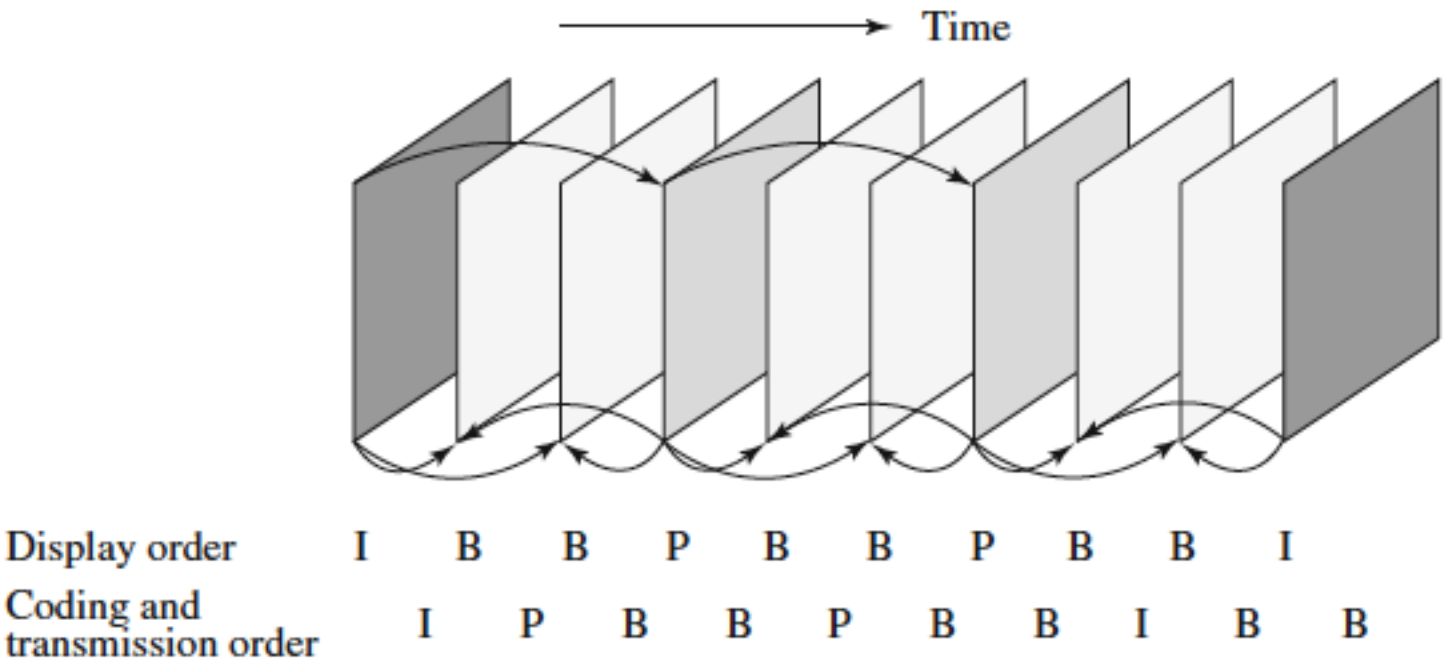


P-frame coding based on motion compensation

MPEG-1 - Video

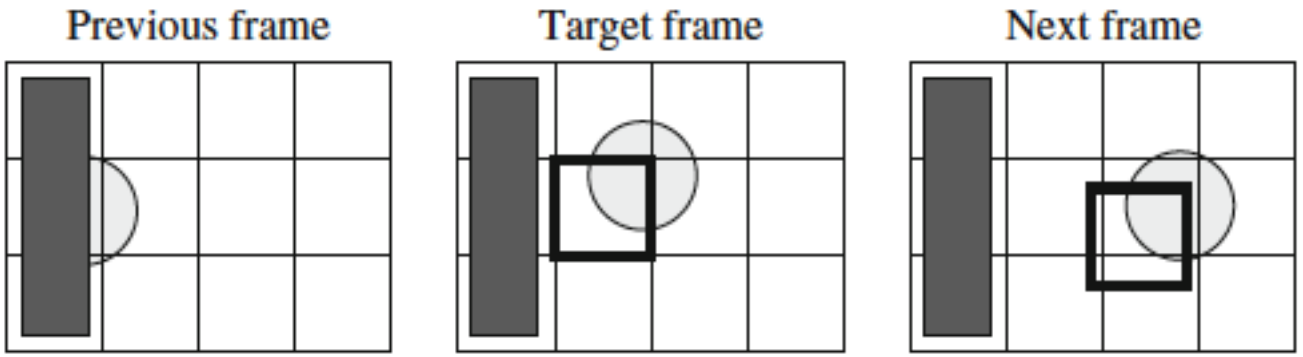
- MPEG-1 has several frame/picture types
  - B-frame (Bidirectional-frame)
    - make predictions using both the previous and future frames (i.e. two anchor frames)
    - requires larger data buffers and causes an increased delay on both decoding and during encoding
  - D-frame
    - independent images (intra-frames) that have been encoded using DC transform coefficients only
    - very low quality
    - are only used for fast previews of video, for instance when seeking through a video at high speed

MPEG-1 - Video



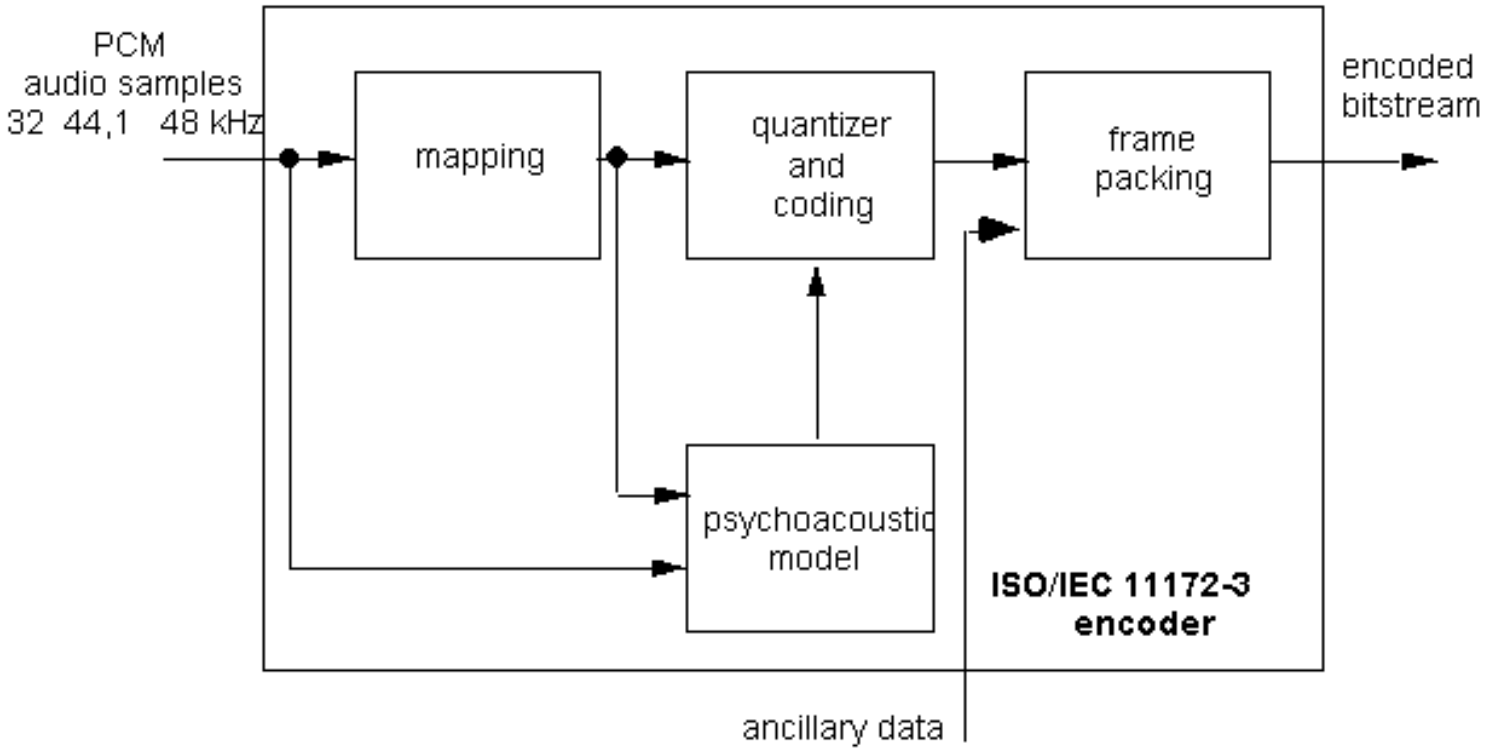
MPEG frame sequence

MPEG-1 - Video



The need for bidirectional frame

MPEG-1 - Audio



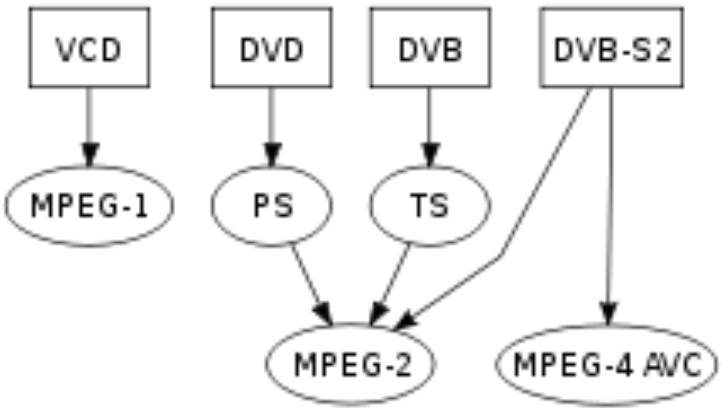
Audio encoding

MPEG-1

- Part 4
  - procedures for testing conformance
  - provides two sets of guidelines and reference bitstreams for testing the conformance of MPEG-1 audio and video decoders, as well as the bitstreams produced by an encoder
- Part 5
  - Reference software
  - C reference code for encoding and decoding of audio and video, as well as multiplexing and demultiplexing

MPEG-2

- MPEG-2
  - generic coding of moving pictures and associated audio information
  - combination of lossy video compression and lossy audio data compression methods
  - storage and transmission of movies using currently available storage media and transmission bandwidth



MPEG-2 is used in Digital Video Broadcast and DVDs.
The MPEG transport stream, TS, and MPEG program stream, PS, are container formats

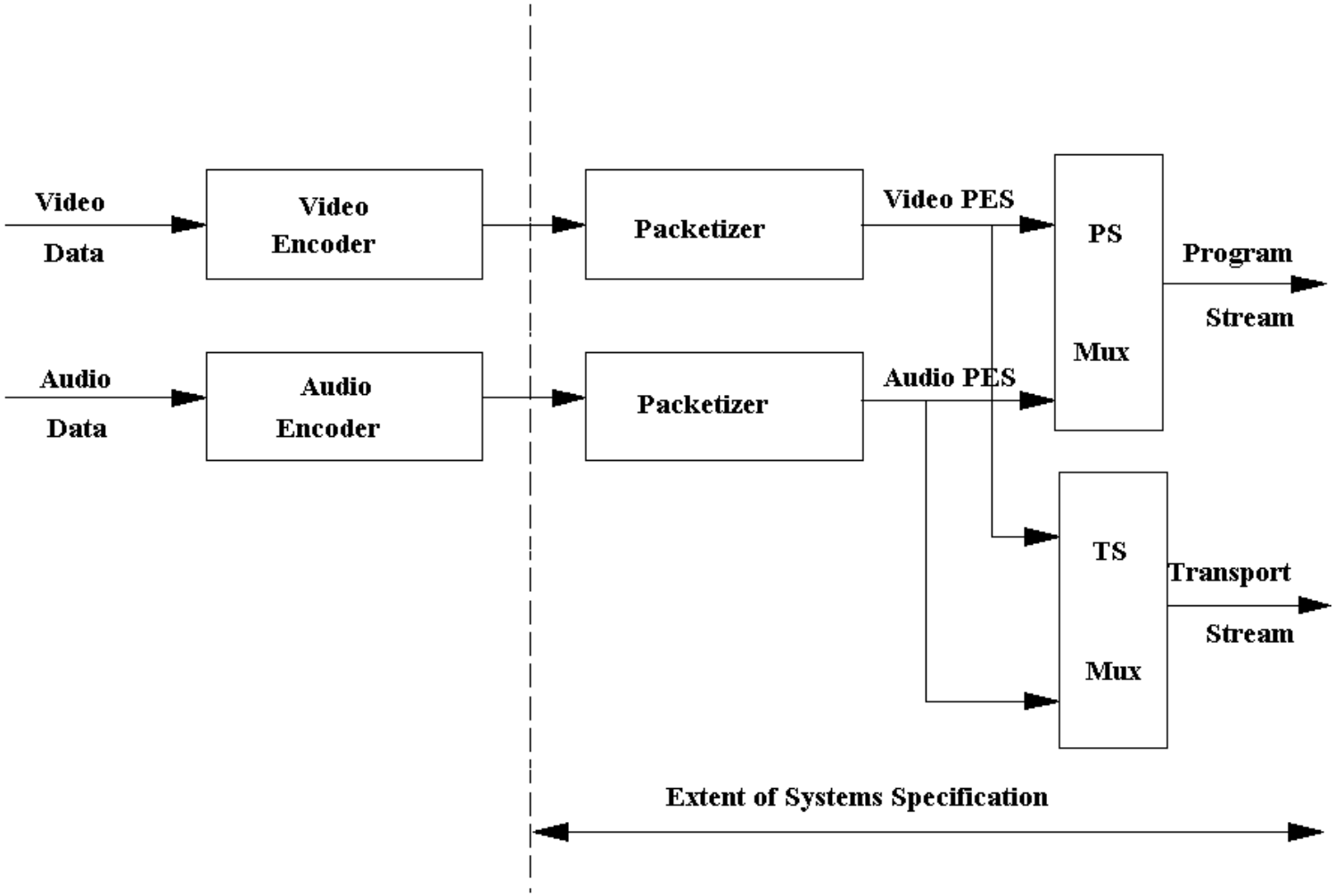
MPEG-2

- The standard consists of 9 Parts
  - ISO/IEC 13818-1 (2000)
    - Systems
  - ISO/IEC 13818-2 (2000)
    - Video
  - ISO/IEC 13818-3 (1998)
    - Audio
  - ISO/IEC 13818-4 (1998)
    - Conformance Testing
  - ISO/IEC 13818-1 (1997)
    - Software simulation

MPEG-2

- The standard consists of 9 Parts
  - ISO/IEC 13818-6 (1998)
    - Extensions for DSM-CC
  - ISO/IEC 13818-7 (1997)
    - Advanced Audio Coding (AAC
  - ISO/IEC 13818-8 (1996)
    - Extension for real time interface for systems decoders
  - ISO/IEC 13818-9 (1999)
    - Conformance extensions for Digital Storage Media Command and Control (DSM-CC)

MPEG-2 - System



ISO/IEC 13818 - System

MPEG-2 - Video

- Video encoding
  - similar to the previous MPEG-1 standard
  - provides support for interlaced video, the format used by analog broadcast TV systems
  - MPEG-2 Video and Systems are also used in some HDTV transmission systems

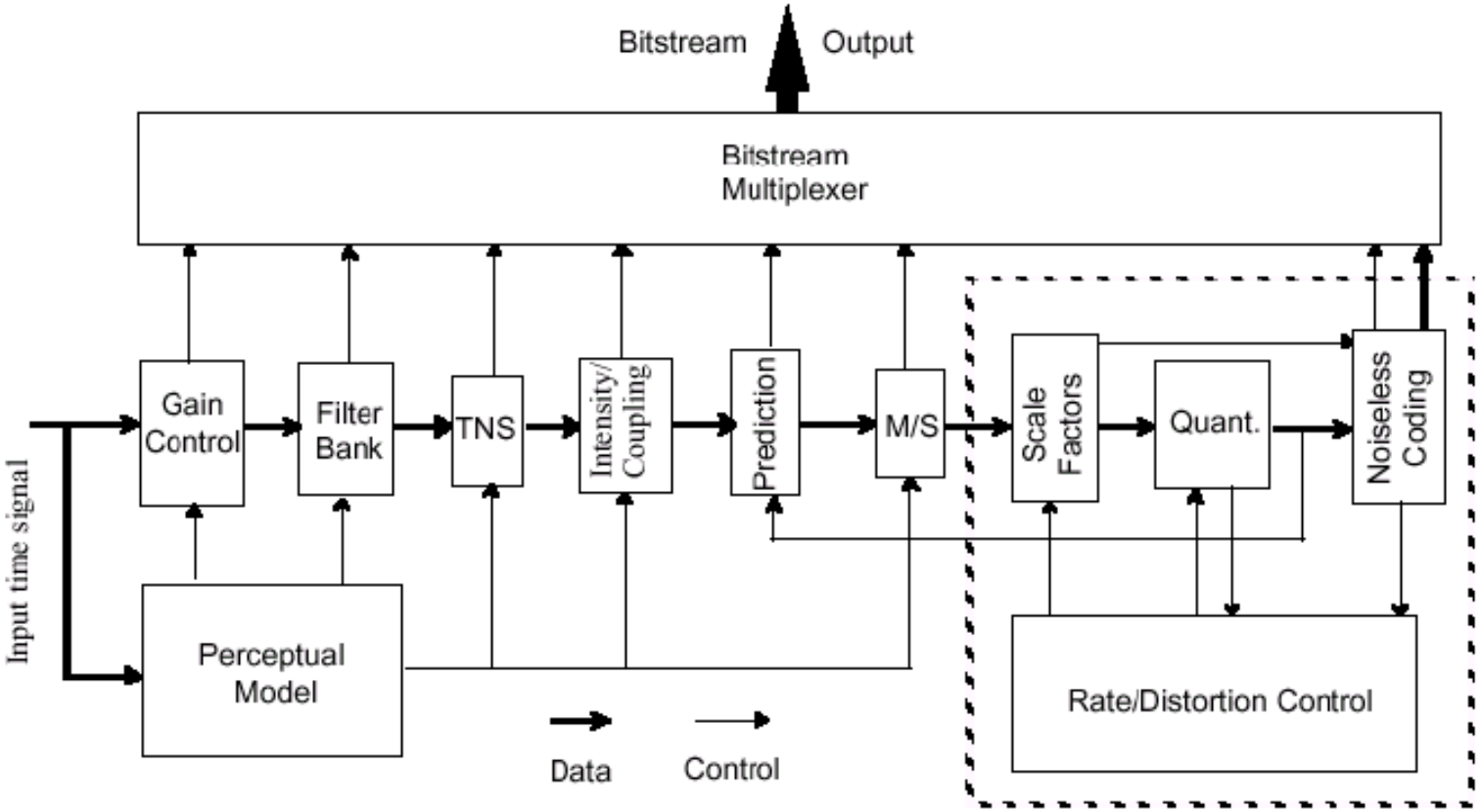
MPEG-2 - Audio

- Audio encoding
  - MPEG-2 introduces new audio encoding methods compared to MPEG-1
    - MPEG-2 Part 3
      - enhances MPEG-1's audio by allowing the coding of audio programs with more than two channels, up to 5.1 multichannel
    - MPEG-2 Part 7
      - specifies a rather different, non-backwards-compatible audio format
      - is referred to as MPEG-2 AAC (Advanced Audio Coding)
      - AAC is more efficient

MPEG 2 - AAC

- Adanced Audio Coding (AAC)
  - improvement for multichannel encoding
  - 48 channels
  - samplig frequency from 8 to 96 KHz for each channel

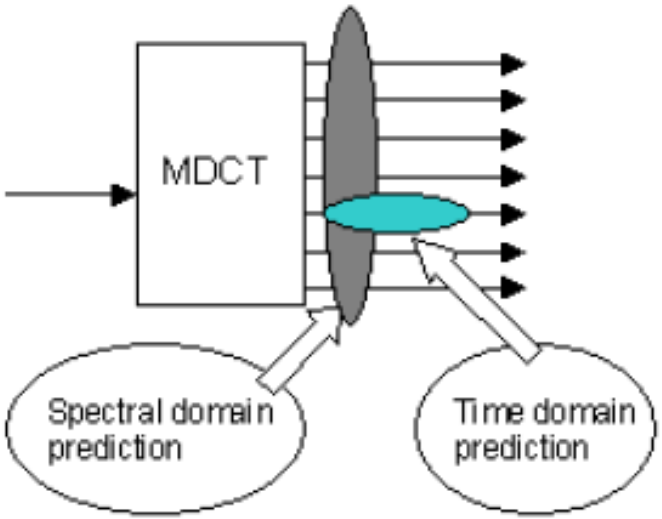
MPEG 2 - AAC



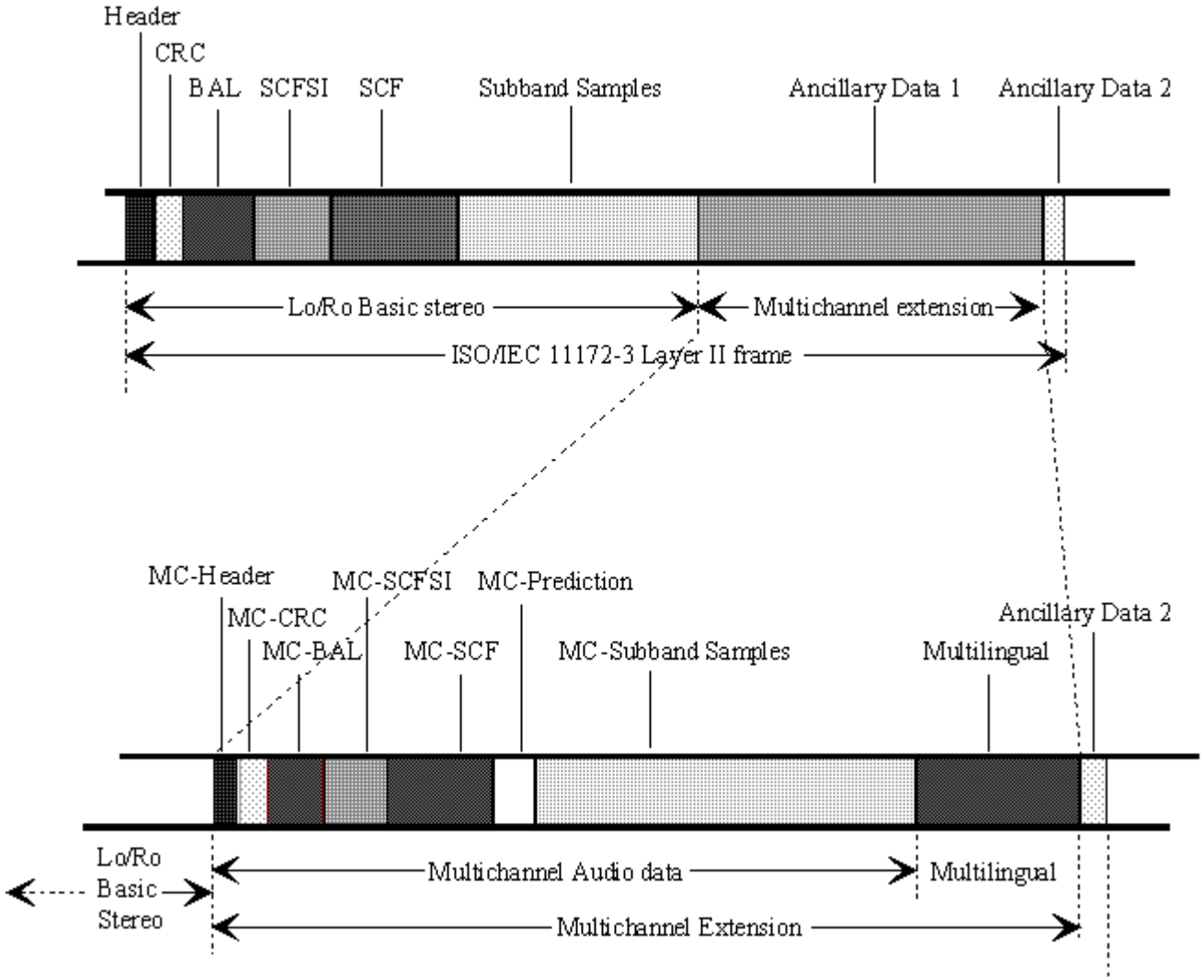
AAC encoding scheme

MPEG 2 - AAC

- Main concept – prediction
  - Prediction
  - Temporal Noise Shaping (TNS)



MPEG-2 - Audio

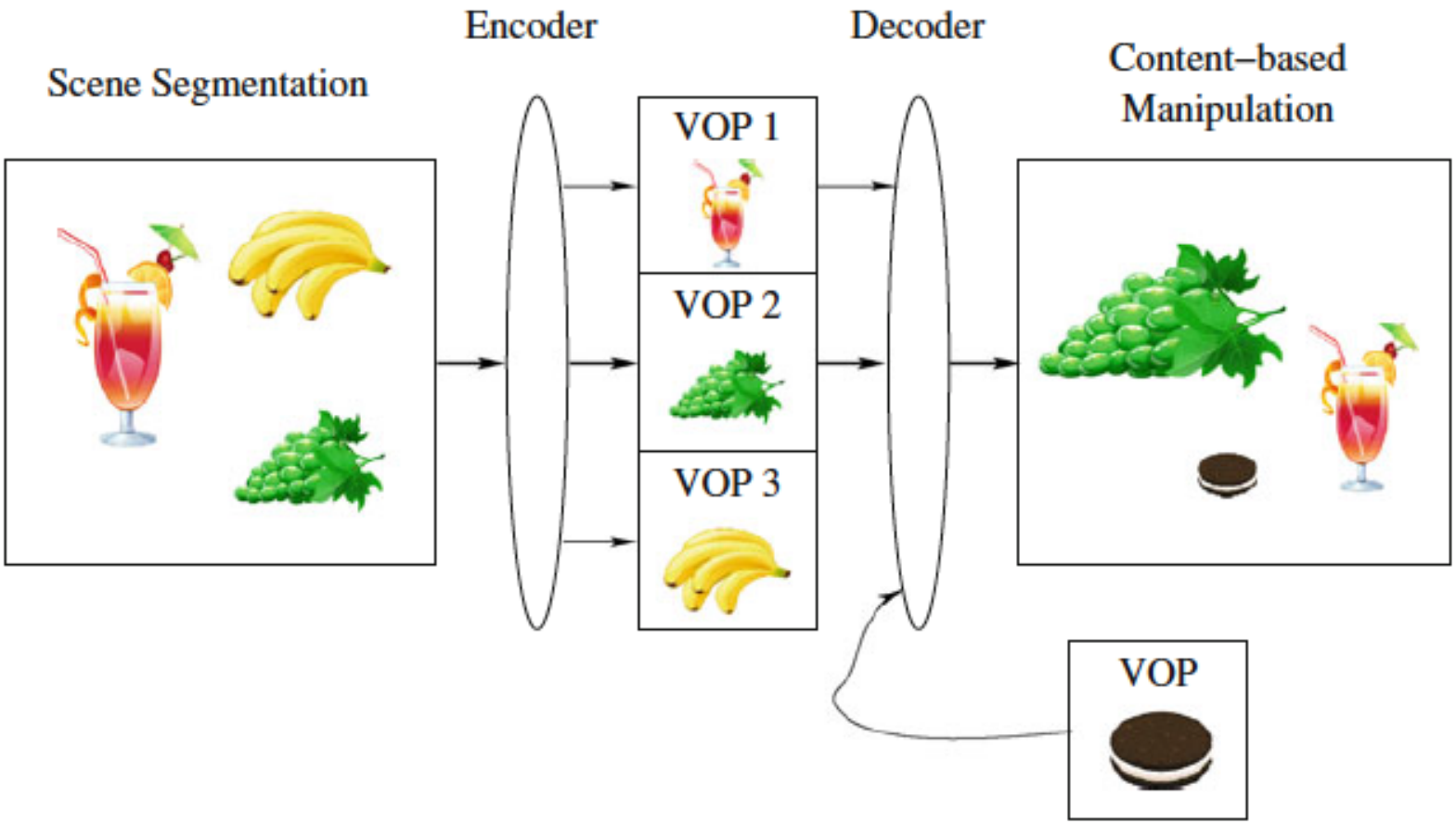


Multichannel Audio information

MPEG-4

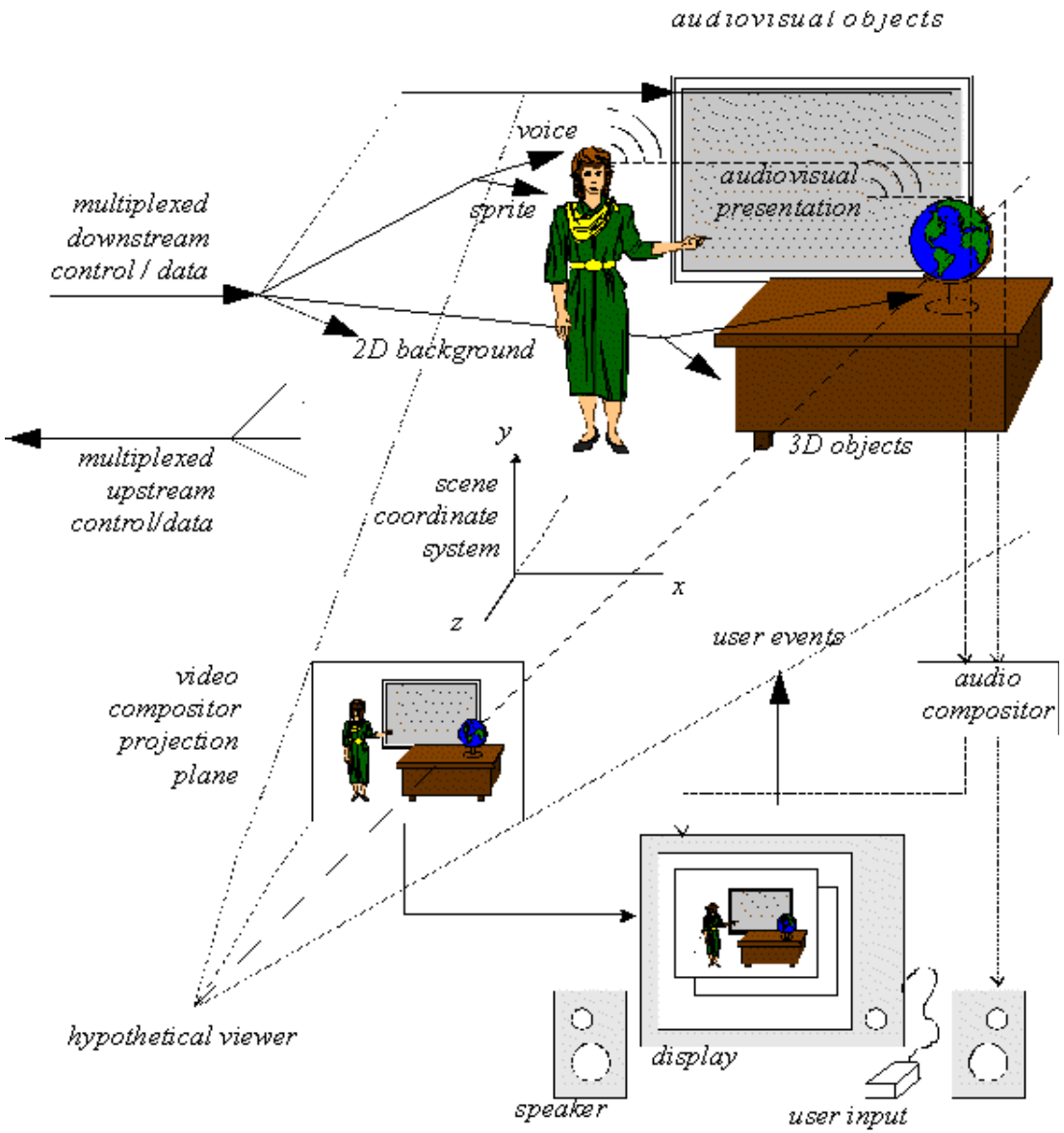
- Besides compression, it pays great attention to user interactivity
  - allows a larger number of users to create and communicate their multimedia presentations and applications on new infrastructures
    - Internet, mobile/wireless networks, …
  - adopt a new object-based coding approach
    - *media objects* are entities
    - media objects (audio and visual objects) can be either natural or synthetic
- bitrate covers a large range, between 5 kbps and 10 Mbps

MPEG-4



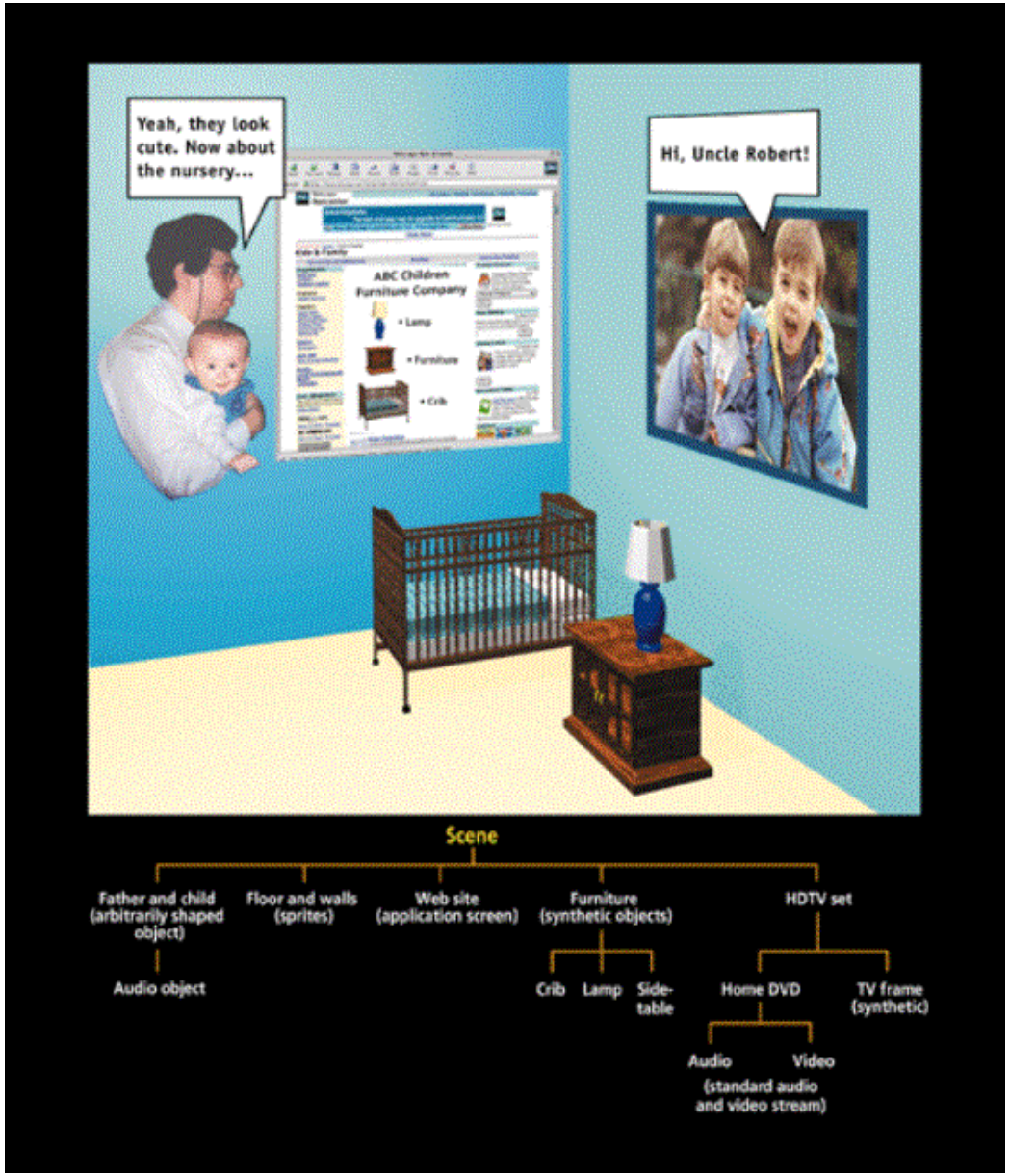
Composition and manipulation of MPEG-4 videos (VOP = Video object plane)

MPEG-4



Example of a MPEG-4 scene

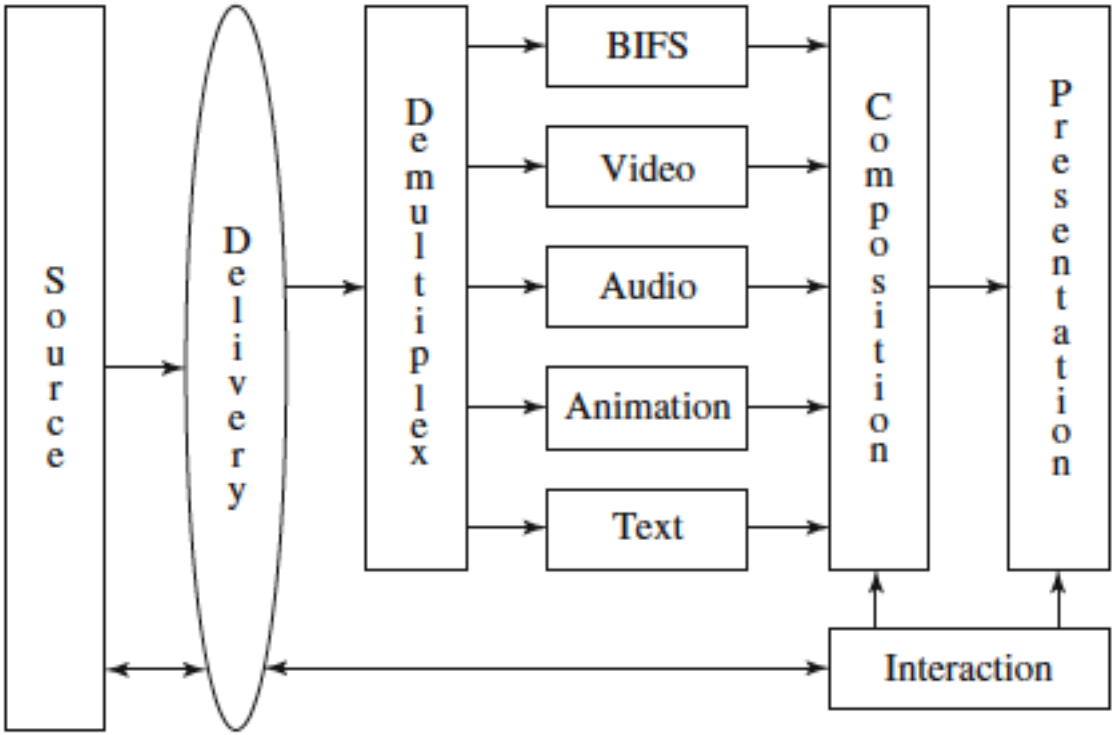
MPEG-4



Example of a MPEG-4 scene

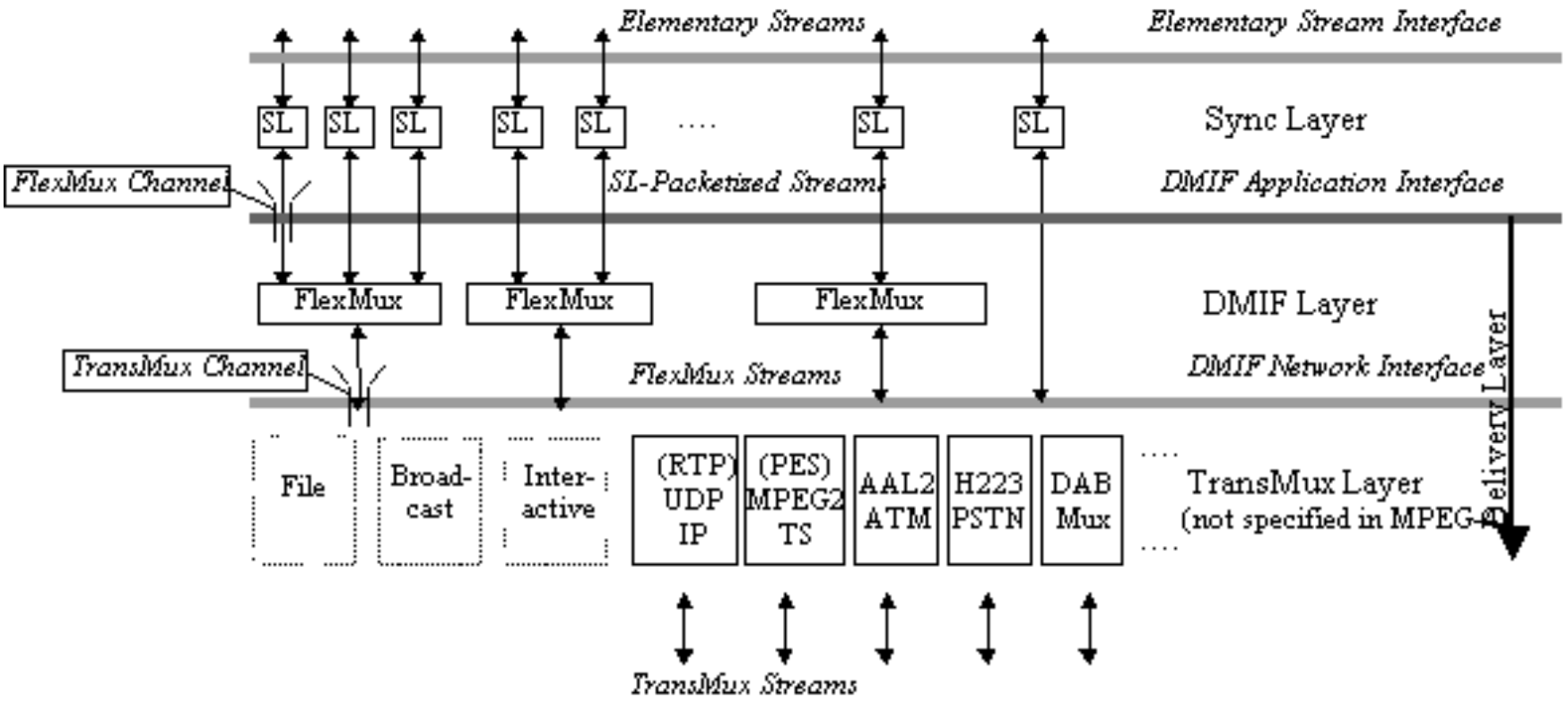
Hierachical scene composition

MPEG-4



MPEG-4 reference model

MPEG-4

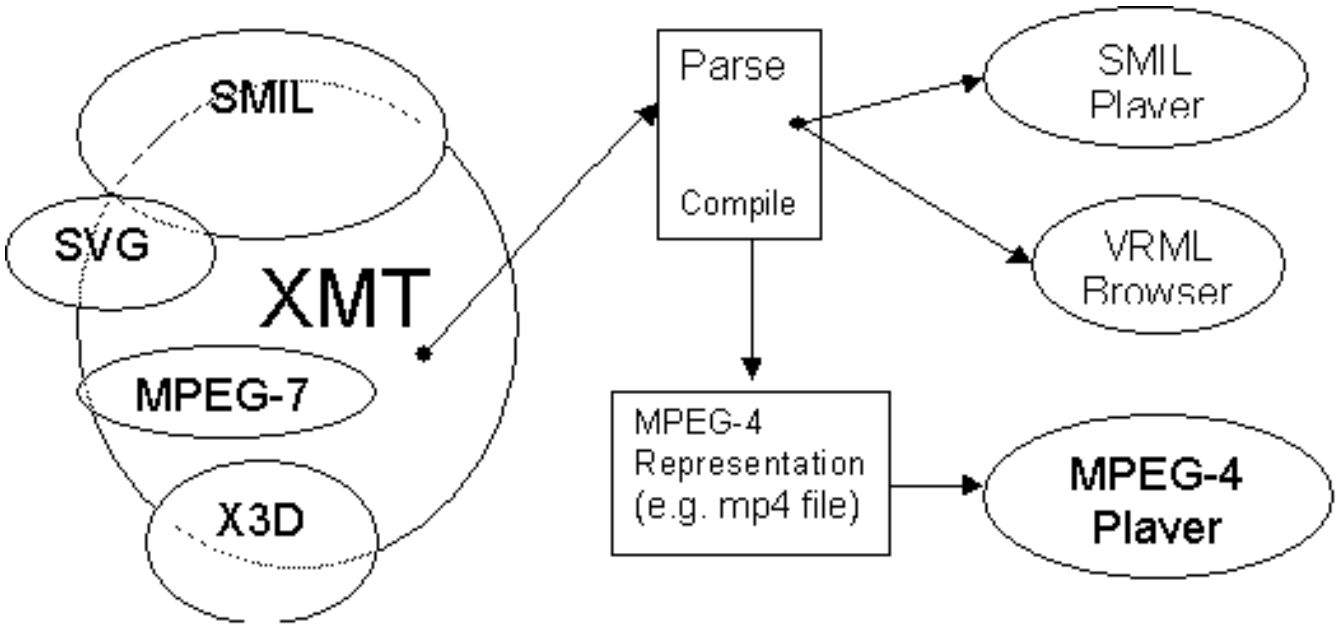


Layers of the system

BIFS

- BIFS - Binary Format for Scenes
  - facilitates the composition of media object in the scene
  - scene graph
    - nodes – audiovisual primitives and attributes
    - graph structure – spatial and temporal relationship of objects in the scene
  - enhancement of Virtual Reality Modeling Language (VRML)

BIFS

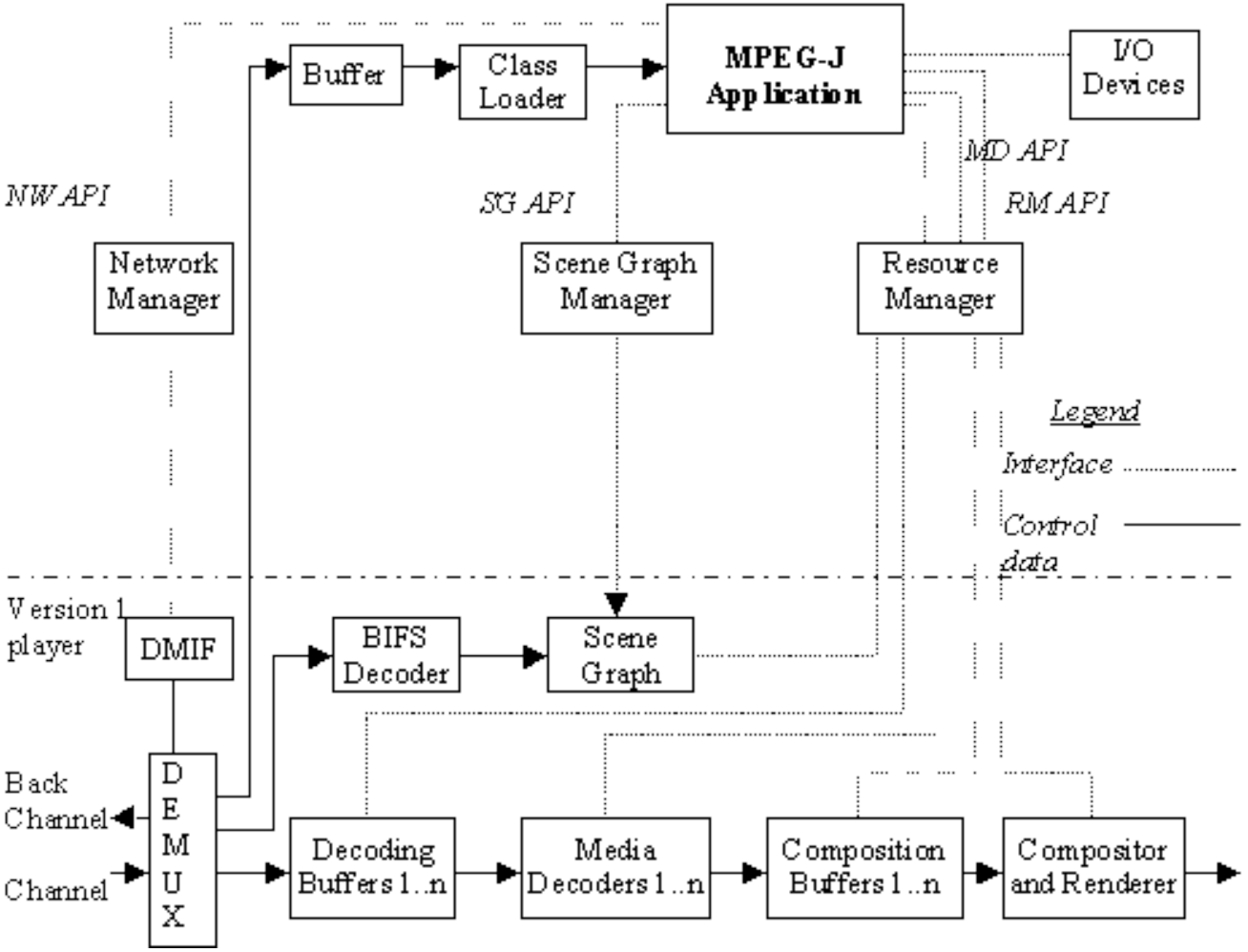


BIFS interfaces

MPEG-J

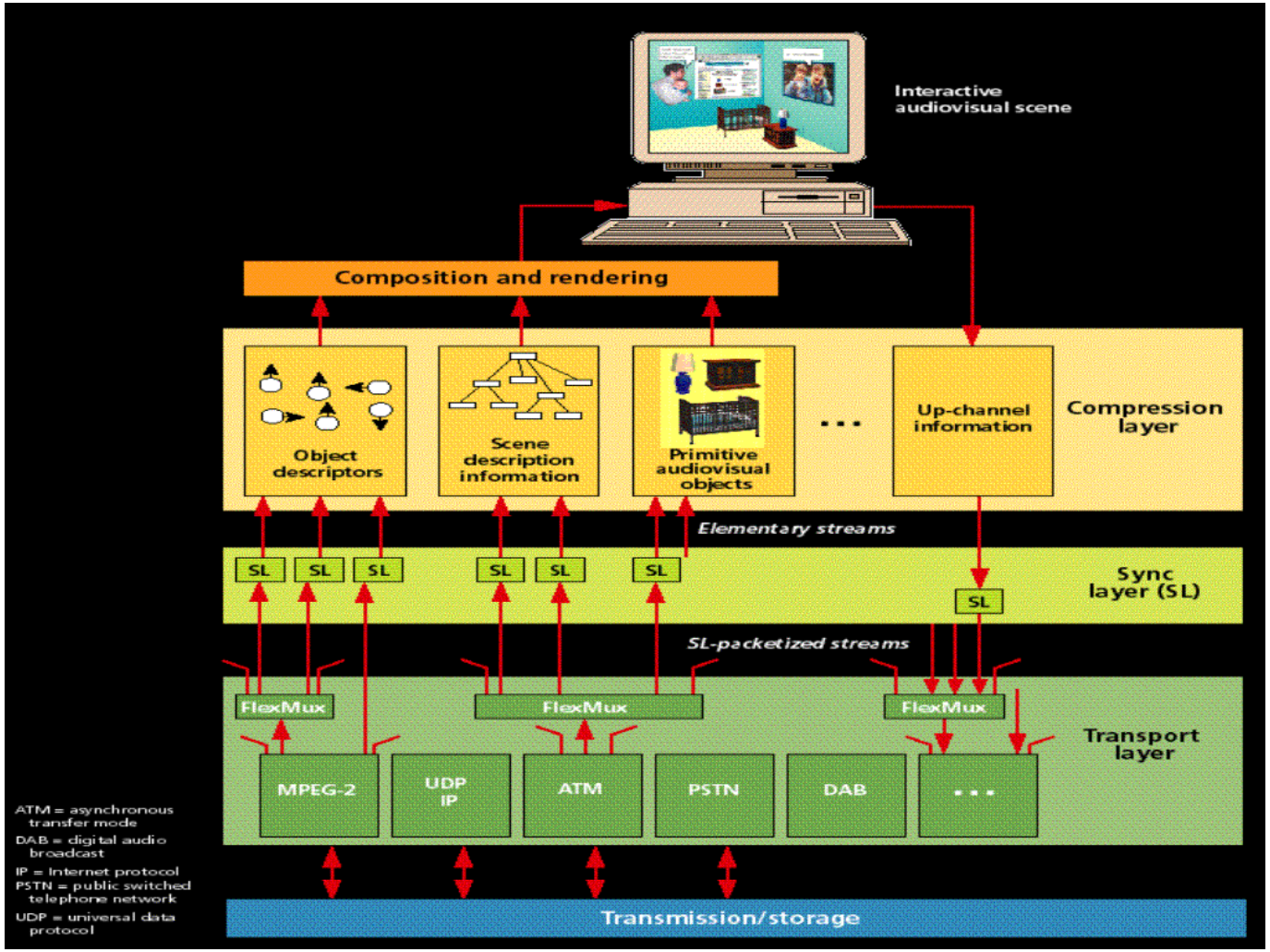
- MPEG-J
  - programming environment
  - Java applications can access Java packages and APIs and enhance users' interactivity

MPEG-J



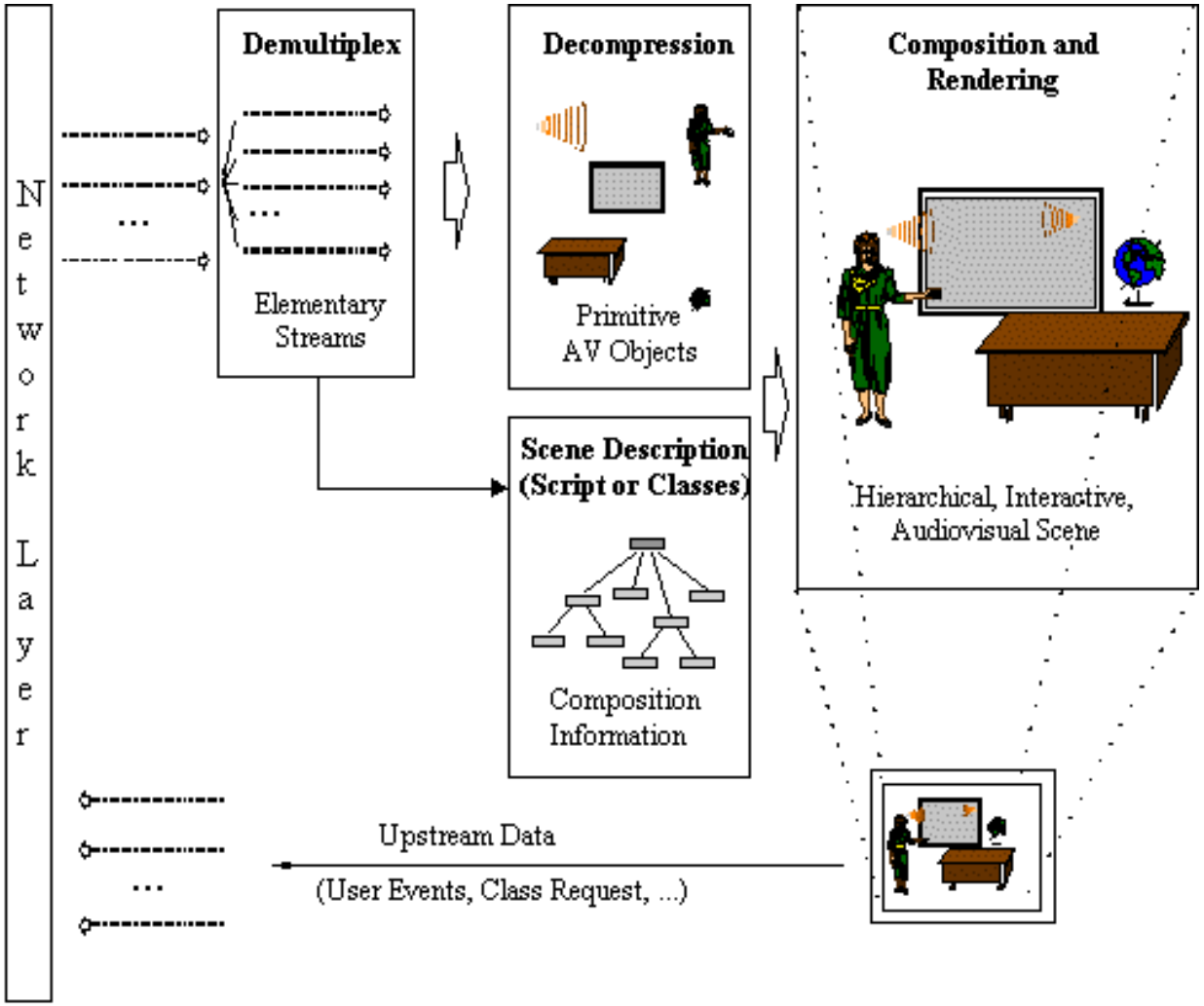
MPEG-J interfaces

MPEG-4



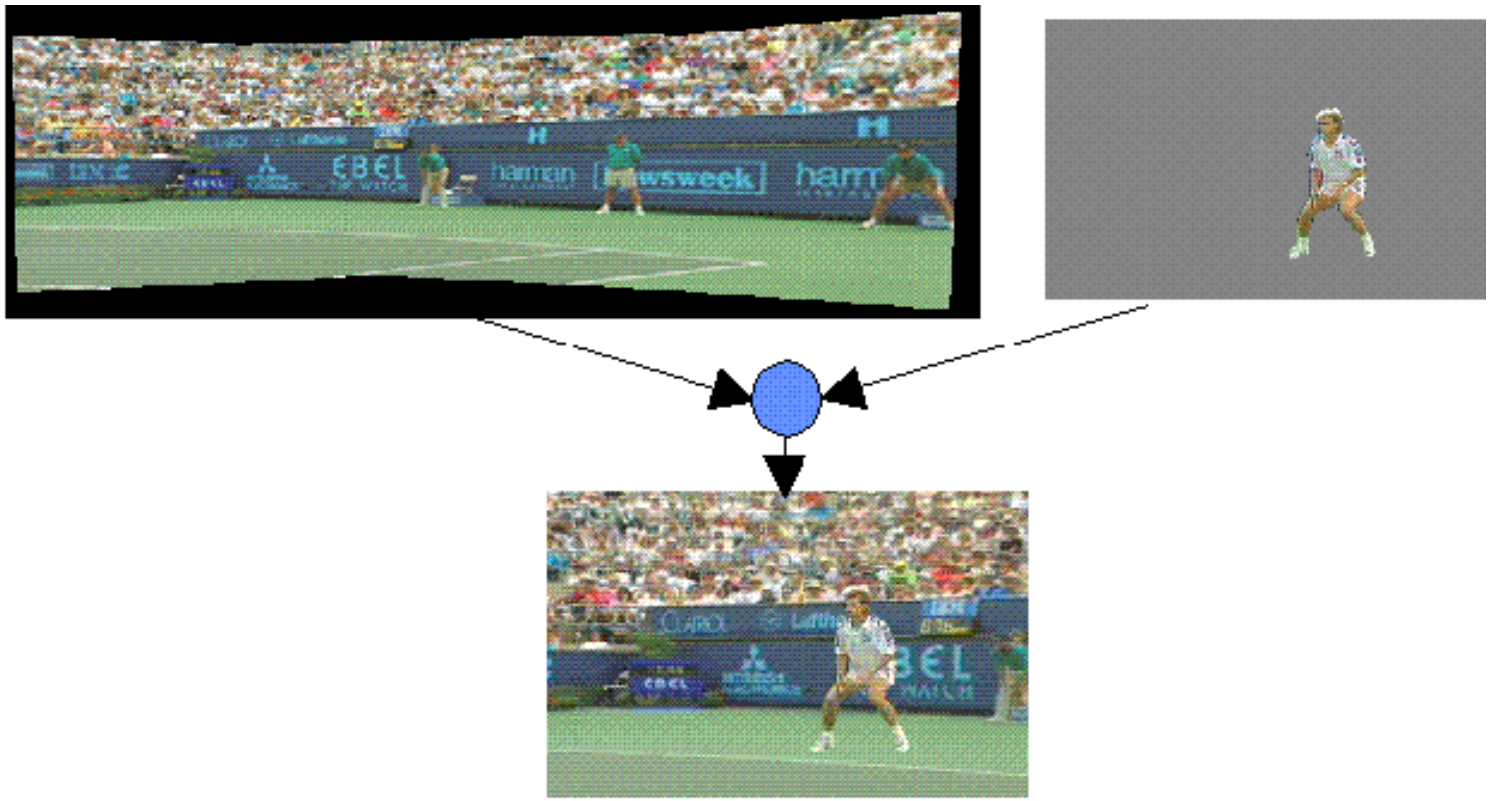
MPEG-4 components and layers

MPEG-4



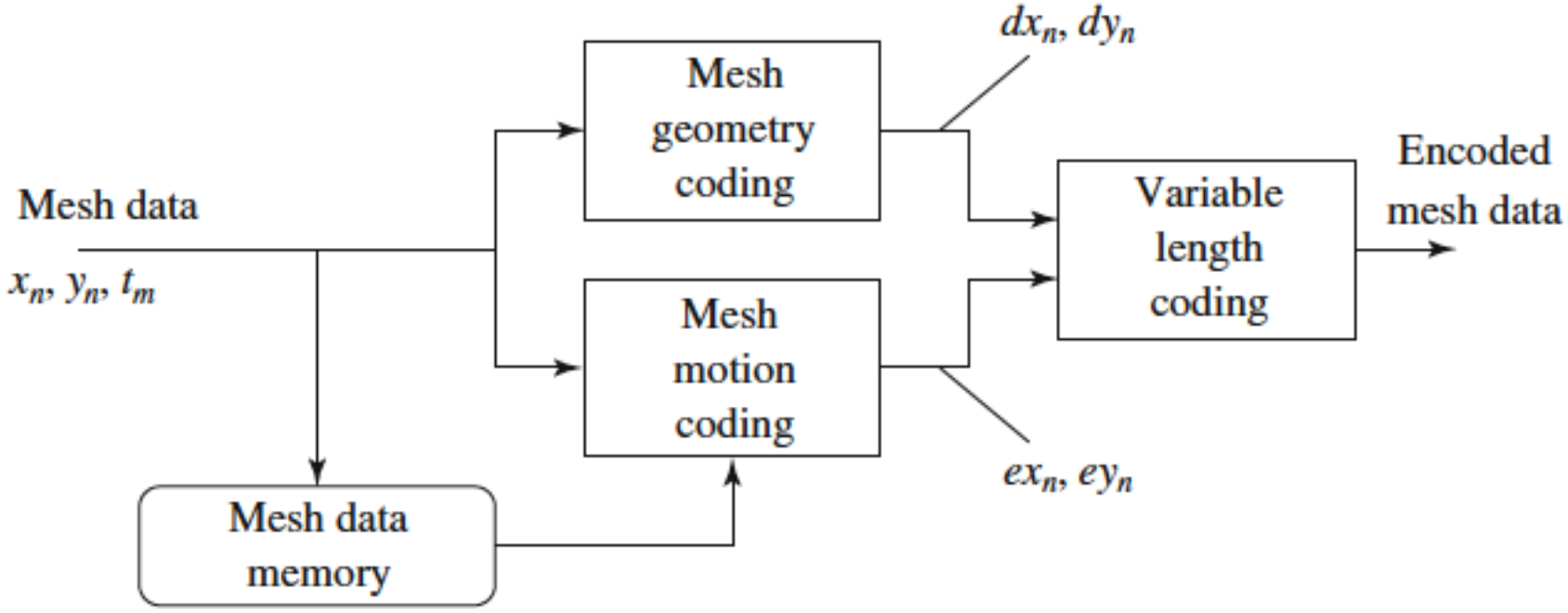
Decoding, composition and rendering

Sprite coding



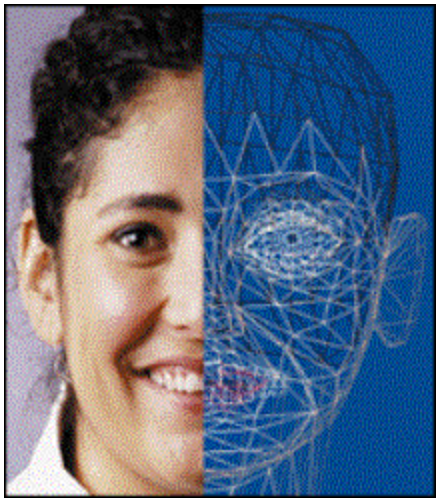
Example of sprite coding to compose an image

Synthetic objects



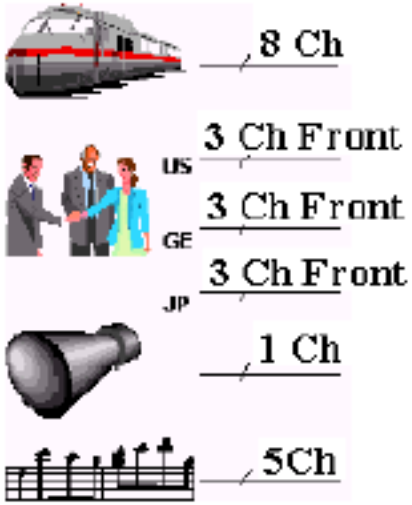
2D Mesh object plane encoding process

Synthetic objects



Examples of mashes applications in MPEG-4 (avatar)

MPEG 4



MPEG 4 considers each audio as an independent object

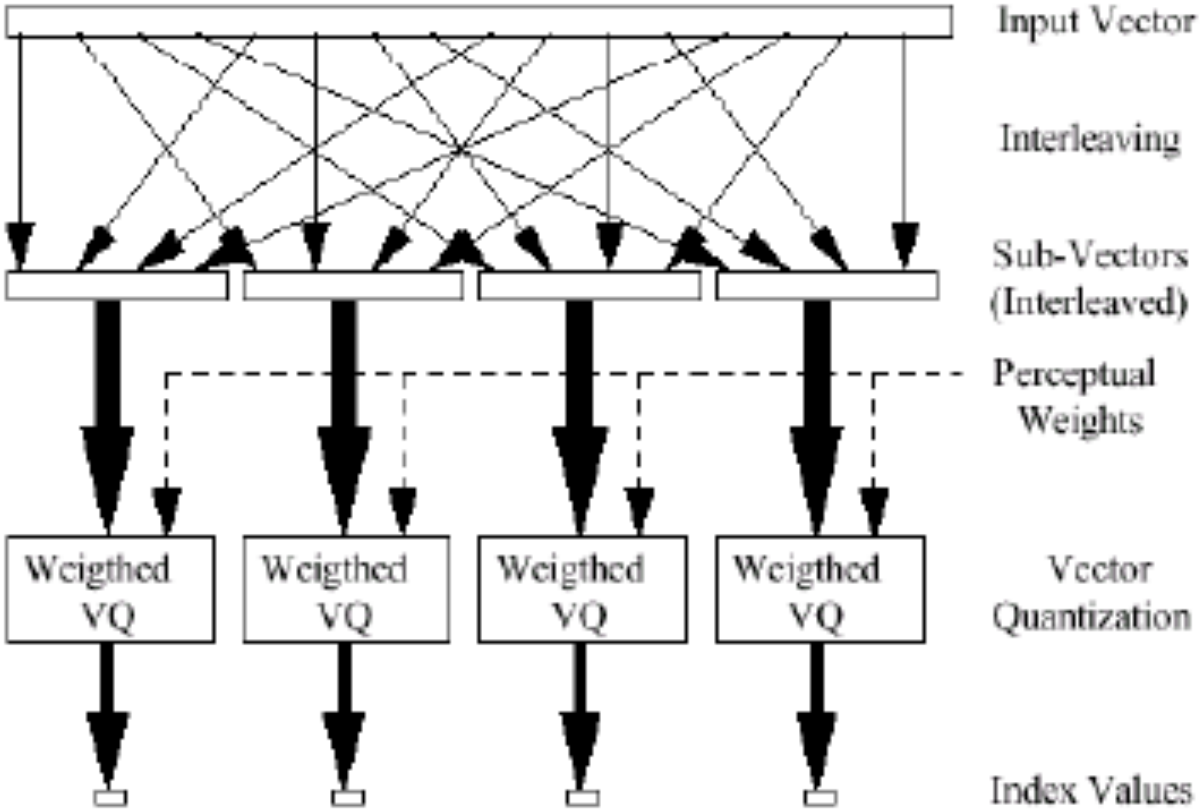
MPEG 4 - Speech Signal

- Speech signal
  - Synthesis Decoding Code Excited Linear Predictive (CELP)
    - Bitrate from 4 to 24 Kbit/s
  - Harmonic vector eXcitation Coding (HVXC)
    - Bitrate from 2 to 4 Kbit/s

General Audio

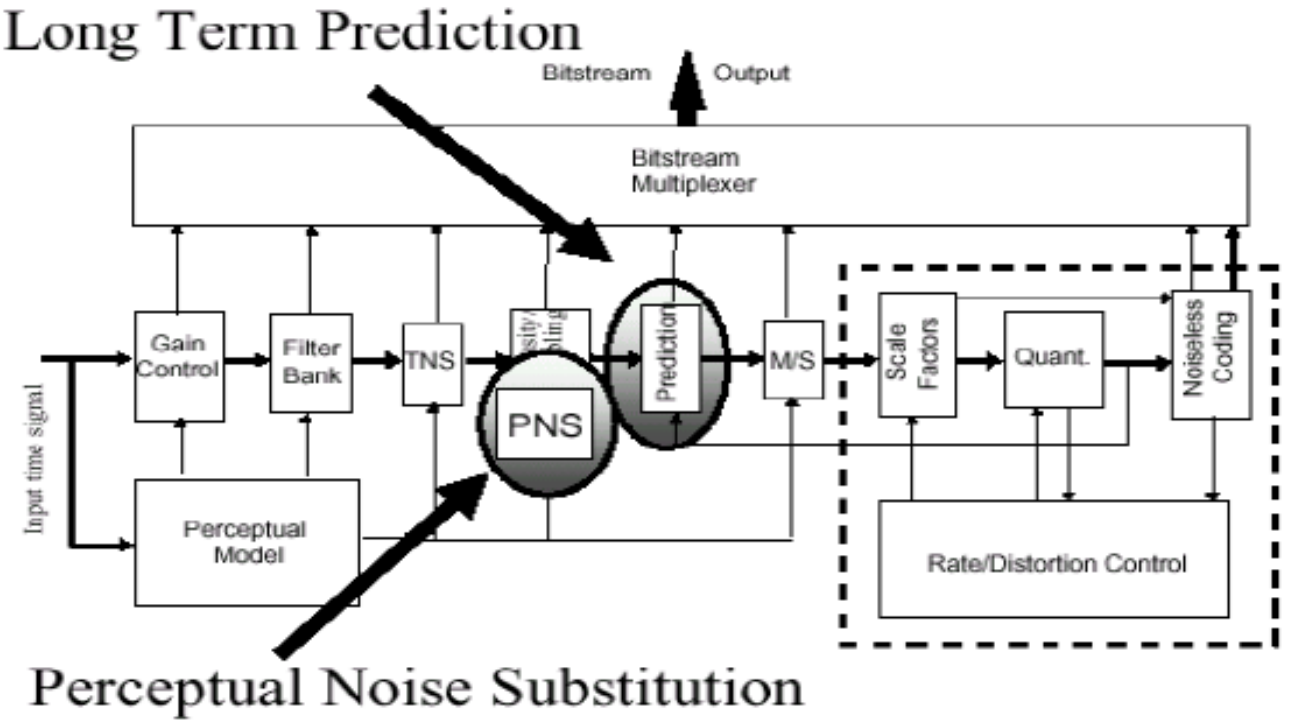
- General Audio
  - Transform-domain Weighted Interleaved Vector Quanrization (TwinVQ)
    - less than 16 Kbit/s
  - AAC for greater bitrates

General Audio



TwinVQ scheme

AAC



TwinVQ scheme

Synthesized Speech

- Text to Speech
  - production of a sound voice from a text
  - Interface for compressed data

Synthesized Audio

- Structured Audio Orchestra Language (SAOL)
  - Set of musical instruments for reproducing
- Structured Audio Score Language (SASL)
  - what to produce

MPEG-7

- The standard consists of 12 Parts
  - ISO/IEC 15938-1 (2002)
    - Systems
  - ISO/IEC 15938-2 (2002)
    - Description definition language
  - ISO/IEC 15938-3 (2002)
    - Visual
  - ISO/IEC 15938-4 (2002)
    - Audio
  - ISO/IEC 15938-5 (2003)
    - Multimedia description schemes

MPEG-7

- The standard consists of 12 Parts
  - ISO/IEC 15938-6 (2003)
    - Reference software
  - ISO/IEC 15938-7 (2003)
    - Conformance testing
  - ISO/IEC 15938-8 (2002)
    - Extraction and use of MPEG-7 descriptions
  - ISO/IEC 15938-9 (2005)
    - Profiles and levels
  - ISO/IEC 15938-10 (2005)
    - Schema definition

MPEG-7

- The standard consists of 12 Parts
  - ISO/IEC 15938-11 (2005)
    - MPEG-7 profile schemas
  - ISO/IEC 15938- (2008)
    - Query format

MPEG-7

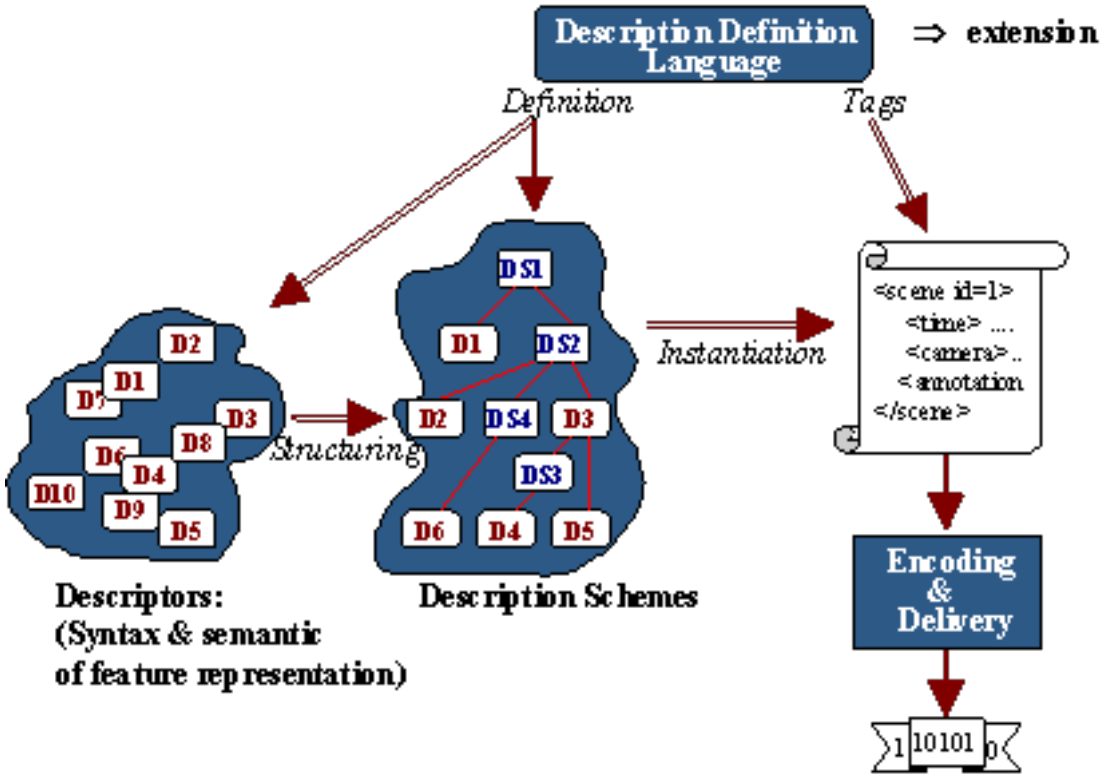
- Main objective
  - audiovisual content-based retrieval (or audiovisual object retrieval)
  - applications such as digital libraries
  - supports a variety of multimedia applications
    - pictures, graphics, 3D models, audio, speech, video, composition information, …

MPEG-7

- Descriptor (D)
  - low-level features
    - color,texture, shape, and motion
  - high-level features of semantic objects
    - events and abstract concepts
- Description Scheme (DS)
  - Specification of the structure and relationship between Ds and DSs
- Description Definition Language (DDL)
  - Syntactic rules to express and combine DSs and Ds

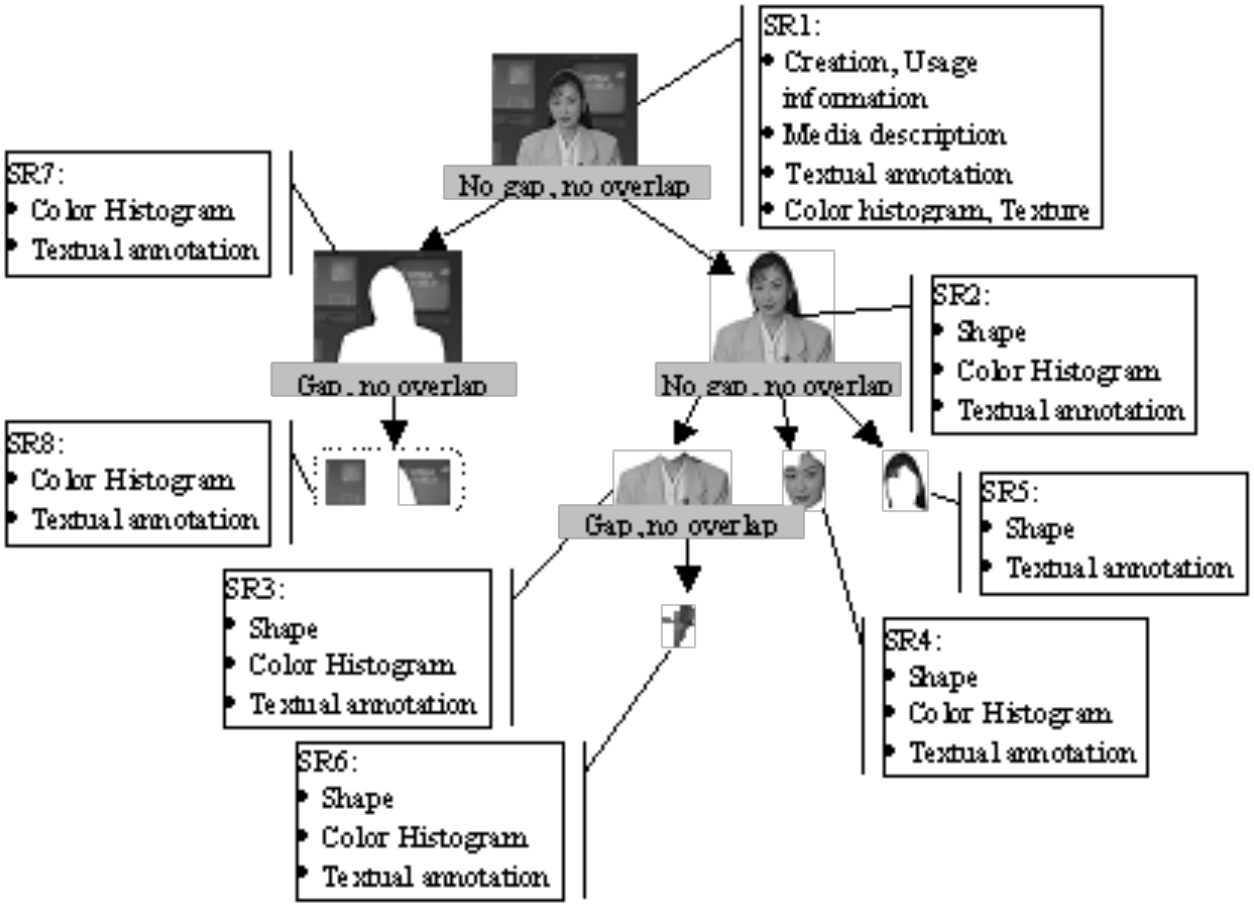
MPEG-7

Based on XML



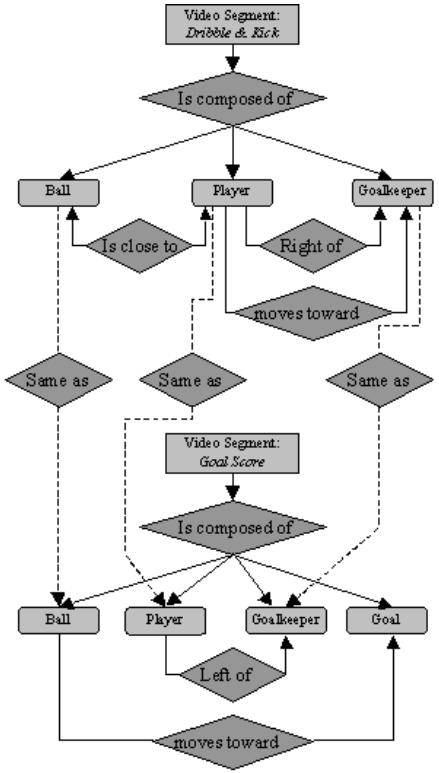
Main parts of MPEG-7 standard

MPEG-7

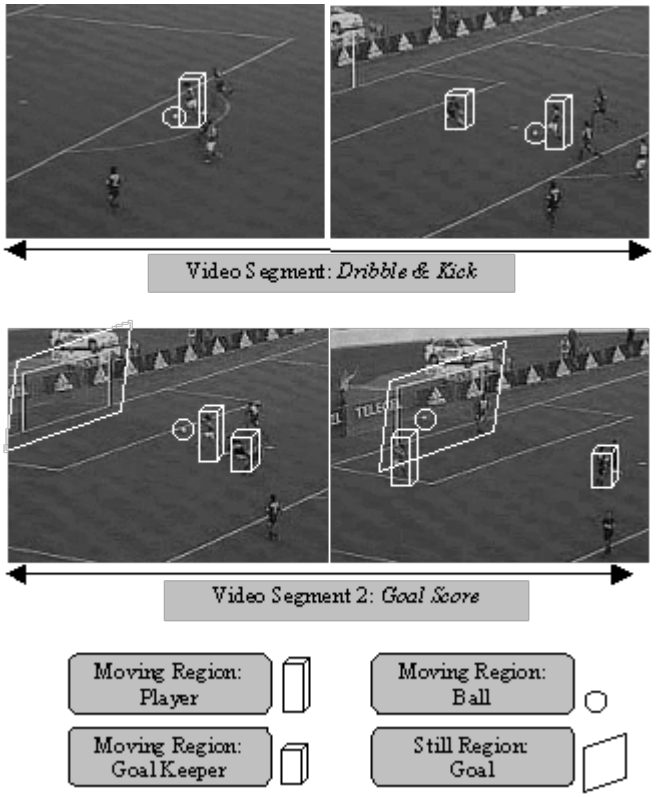


Example of descriptors

MPEG-7

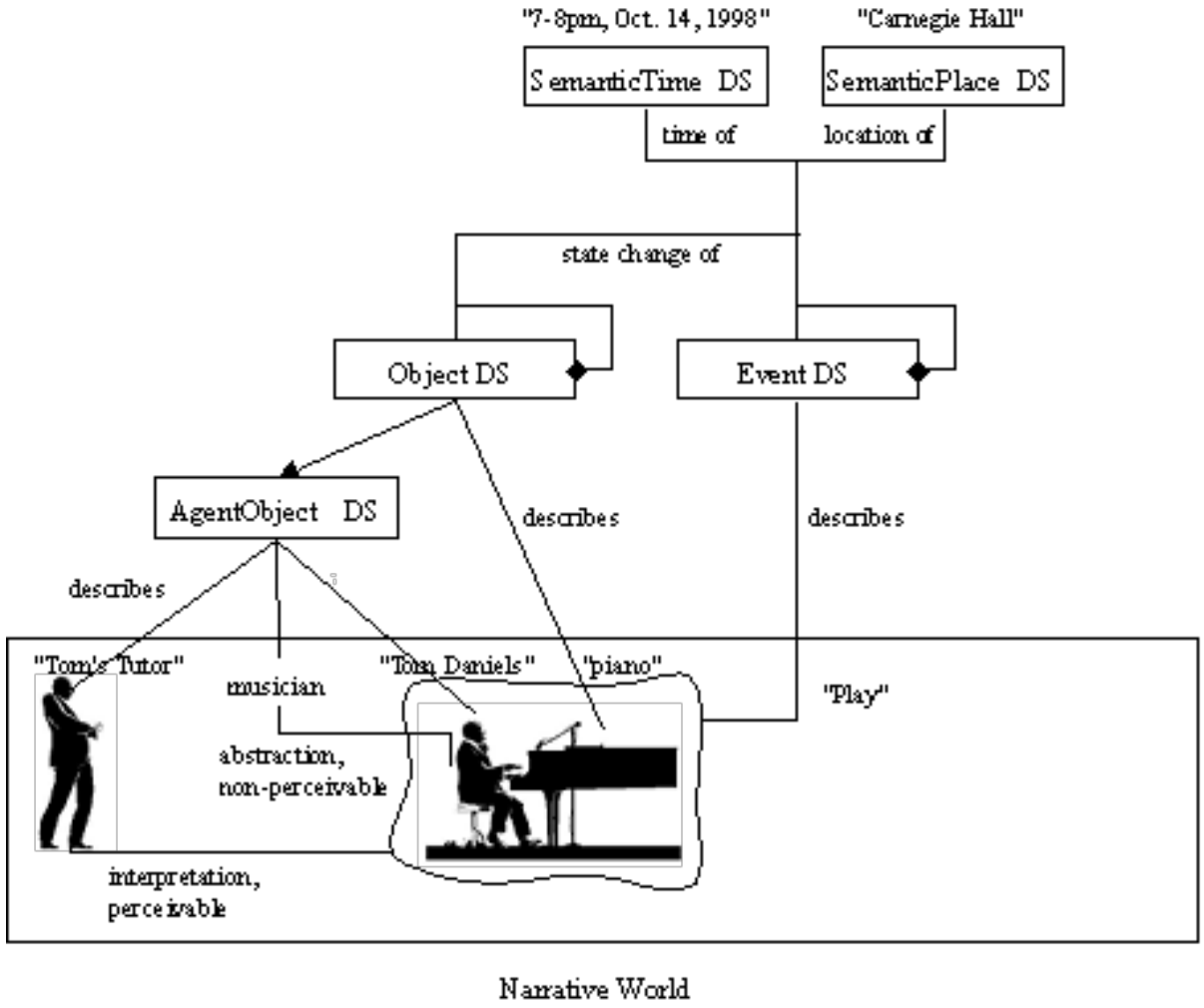


Still region DS



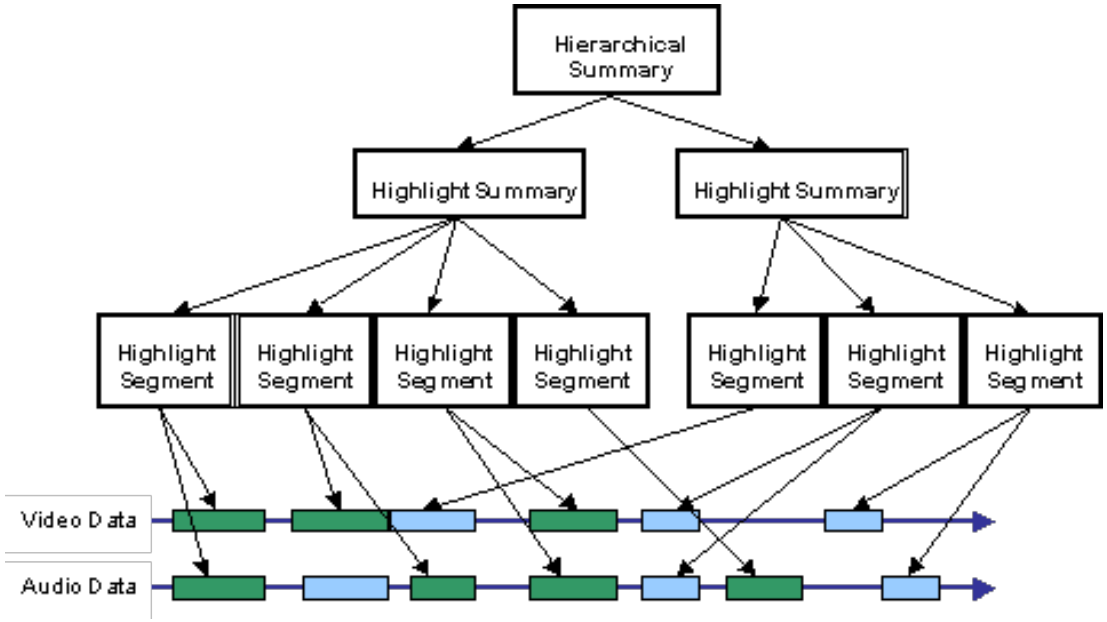
Example of content description in DS

MPEG-7

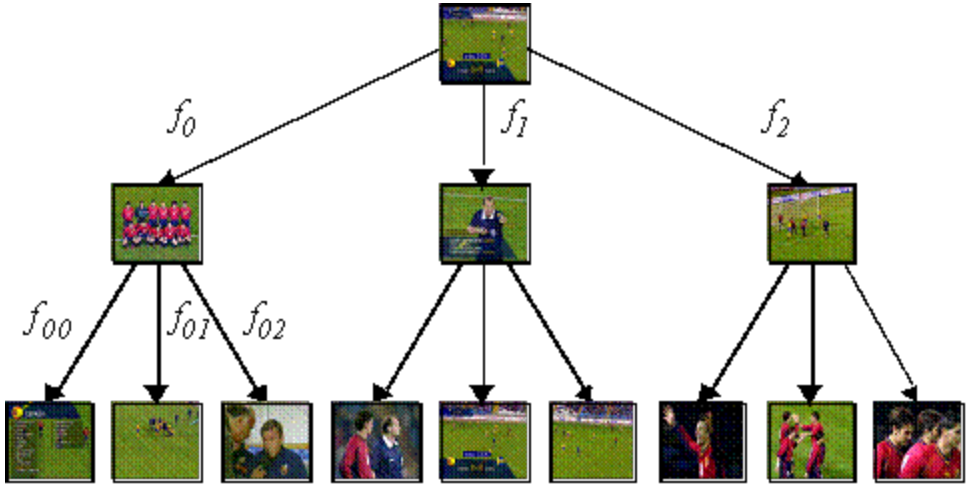


Conceptual description DS

MPEG-7

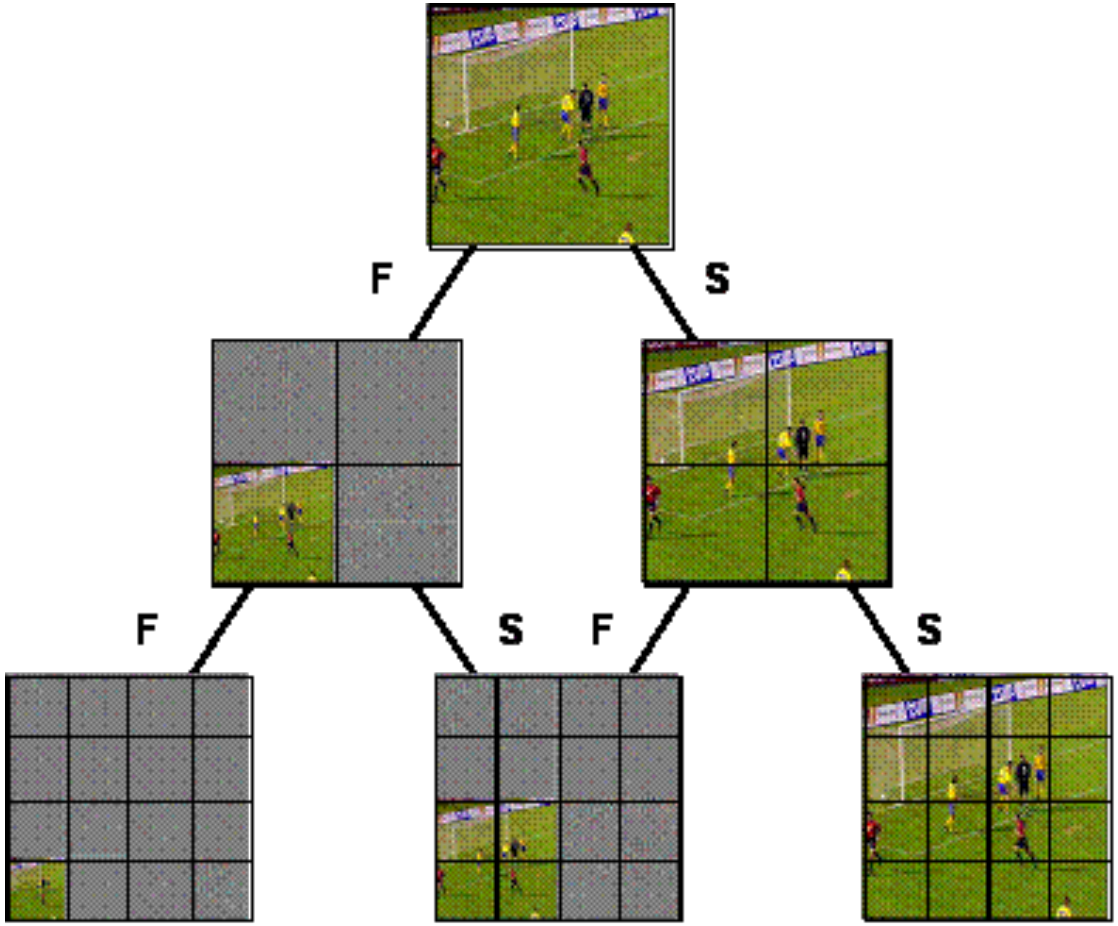


Summaries



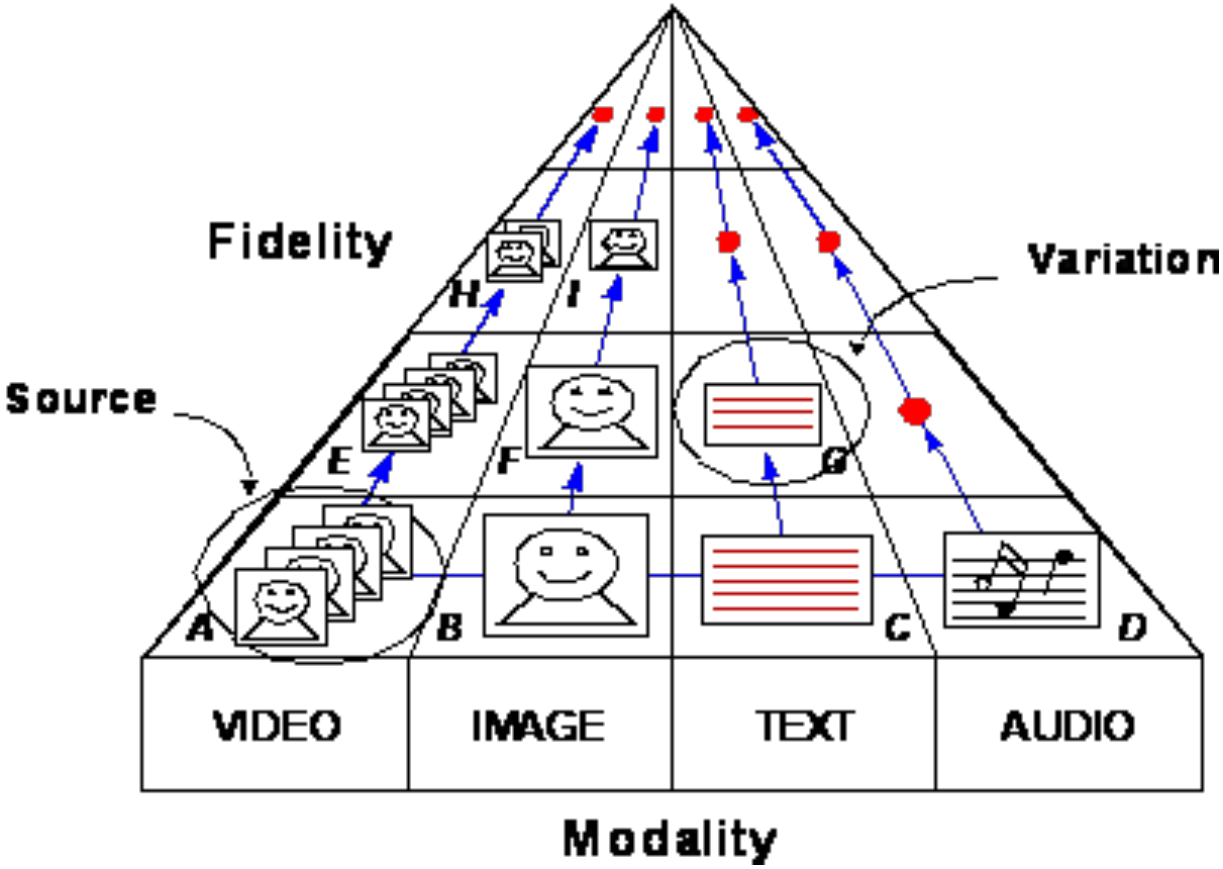
Navigation and access in DS

MPEG-7



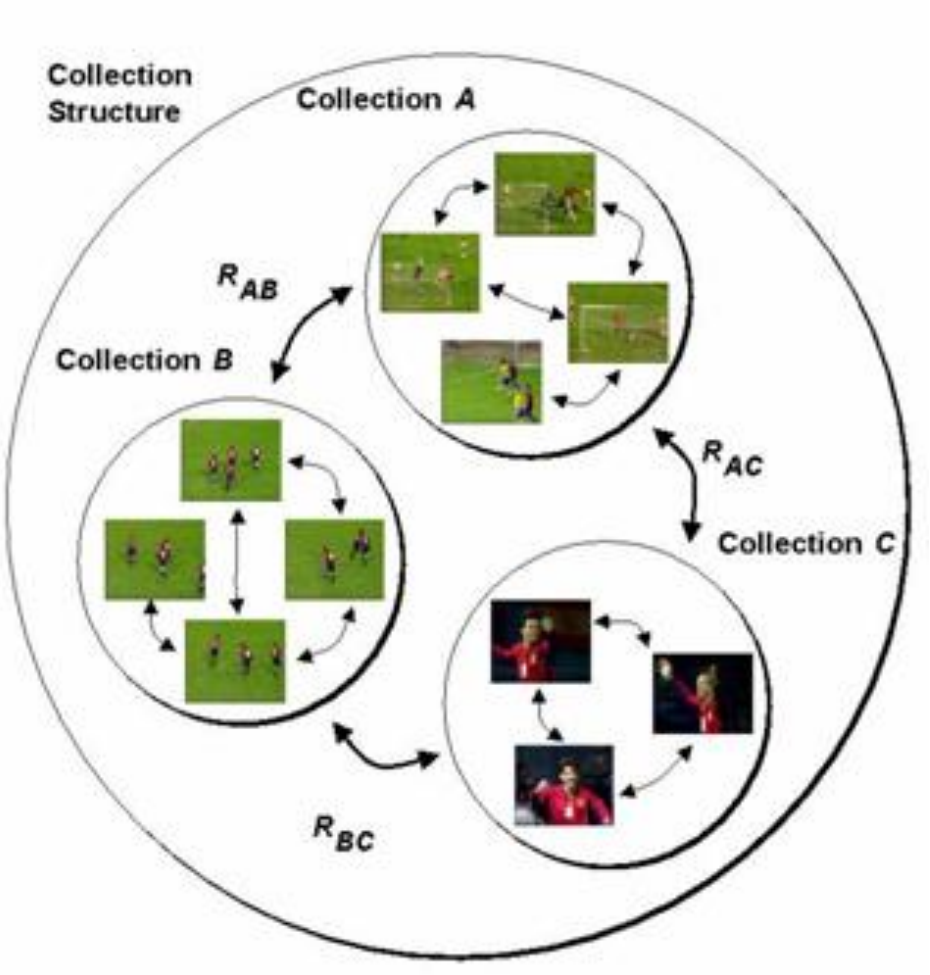
View decomposition in DS

MPEG-7



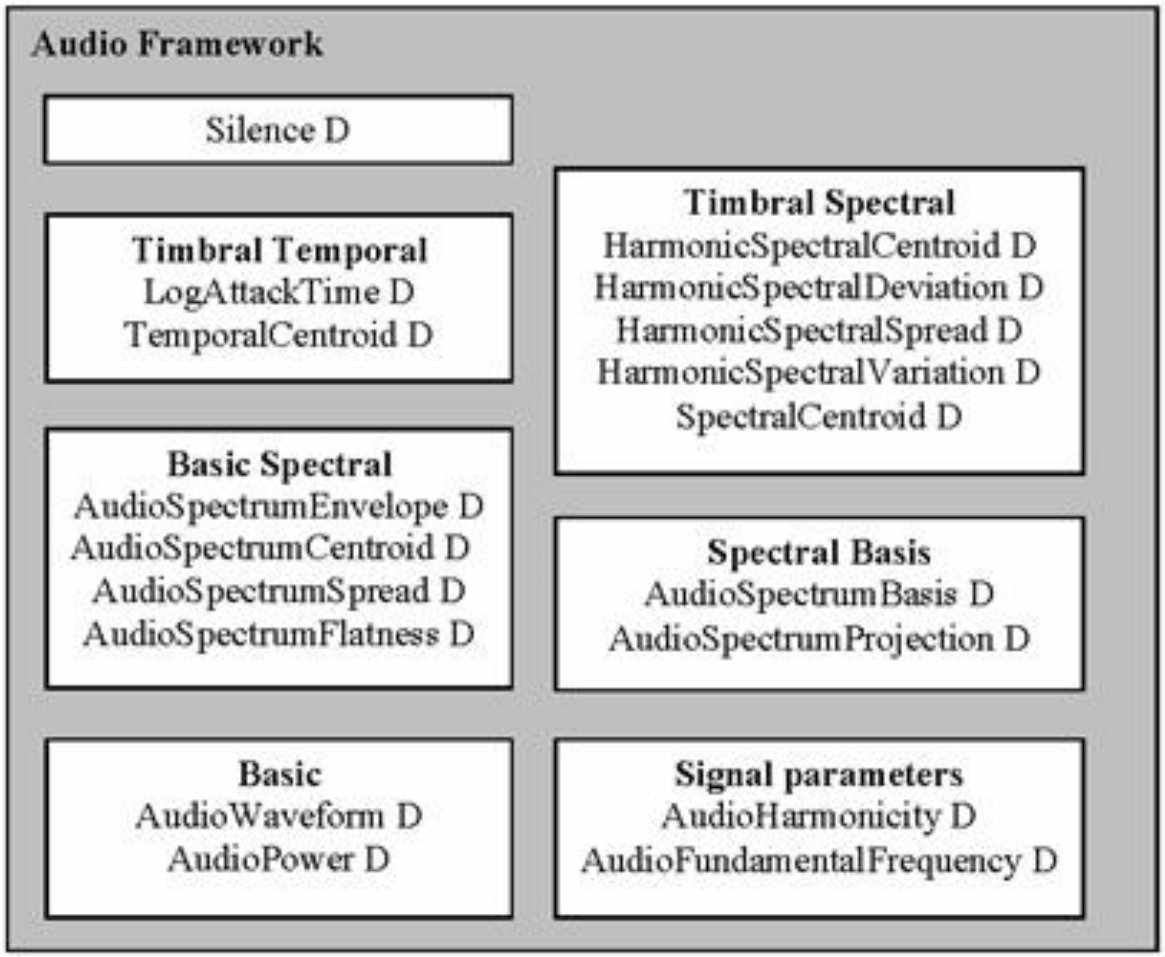
Variation DS

MPEG-7



Collection structure DS

MPEG-7

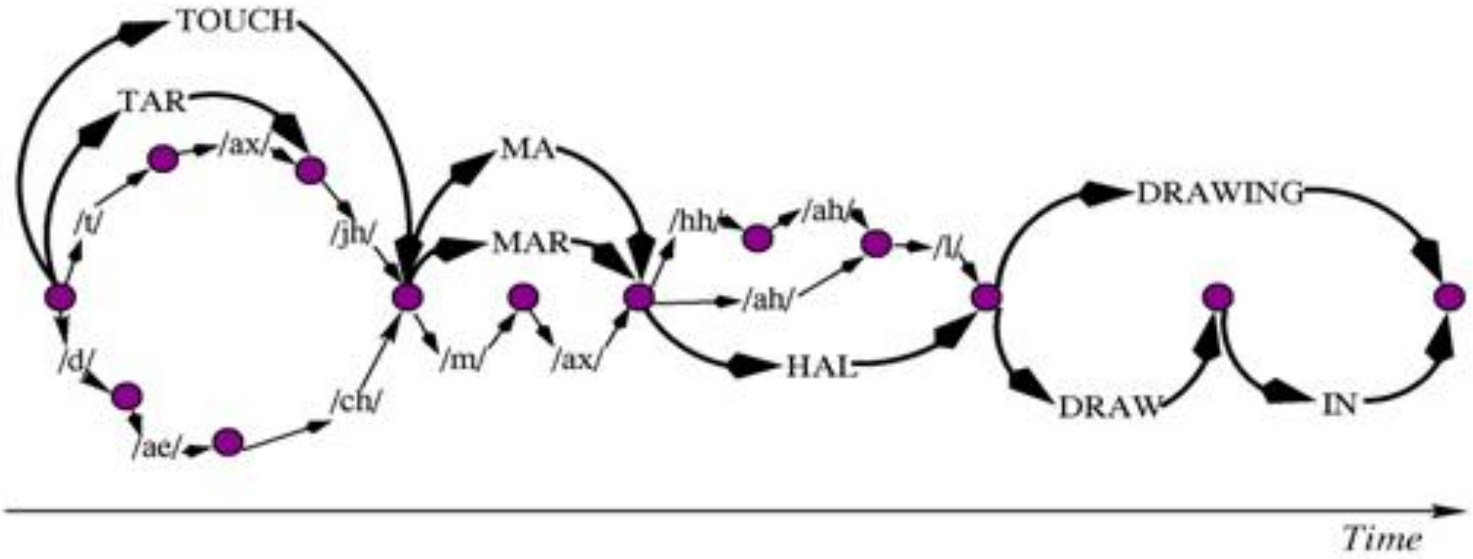


Audio DS

MPEG-7

- High level Audio DS
  - Audio Signature Description Scheme
  - Musical Instrument Timbre Description Tools
  - Melody Description Tools
  - General Sound Recognition and Indexing Description Tools
  - Spoken Content Description Tools

MPEG-7



Speech and DS

Dolby Digital

- Cinematographic encoding
  - Batman in 1992
- Home Theather (e.g., 5.1 Dolby System)
  - 5 channels
    - left, right, middle, left surround, right surround
  - Low Frequency Effects (LFE)
    - 1/10 sampling of the other channels

AC-3 Dolby Digital

- AC-3
  - perceptual digital audio coding technique that reduces the amount of data needed to produce high-quality sound
  - coding system designed specifically for multichannel digital audio