# MPEG Standards

- **MPEG**
  - **M**oving **P**icture **E**xperts **G**roup
- Standards
  - MPEG-1
  - MPEG-2
  - MPEG-4
  - MPEG-7
  - MPEG-21

# What is MPEG

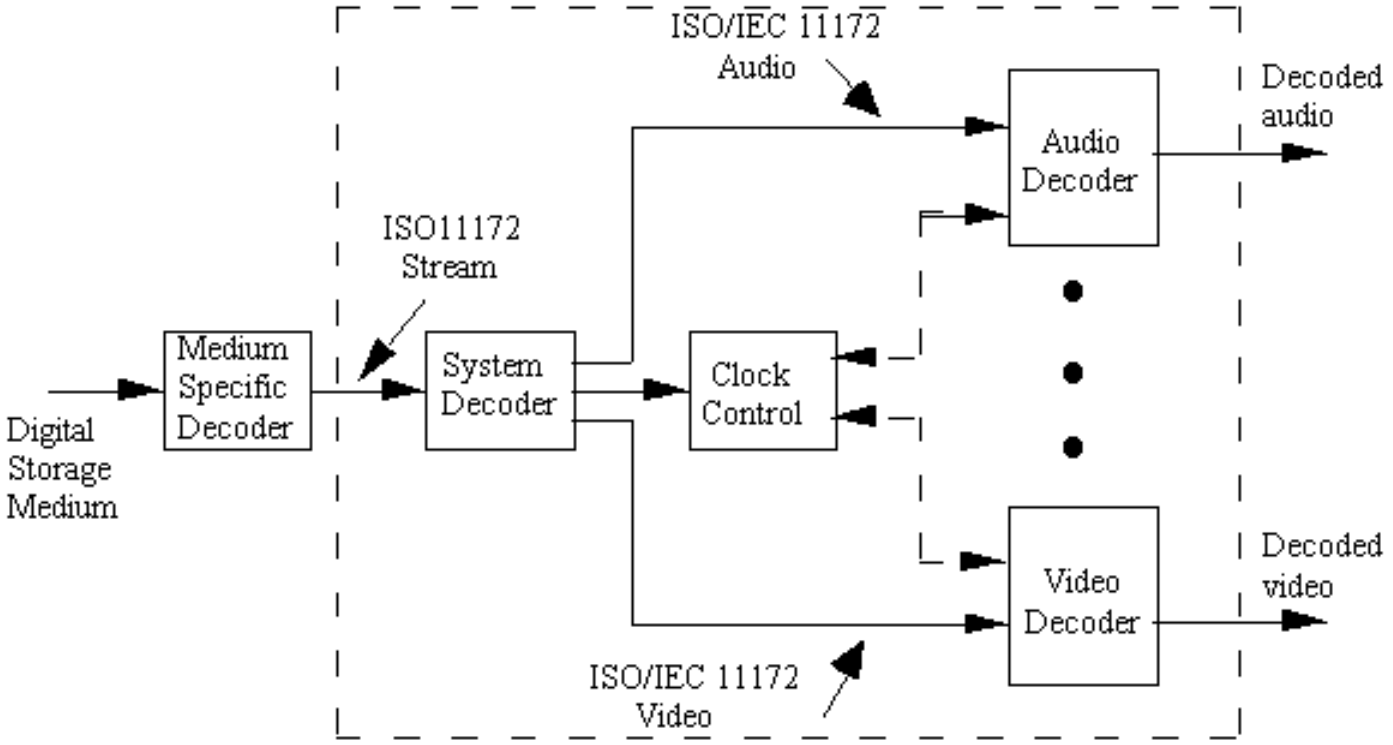
- MPEG: Moving Picture Experts Group
  - established in 1988
- ISO/IEC JTC 1 /SC 29 / WG 11
  - Int. Standards Org. / Int. Electro-technical Commission
  - Joint Technical Committee Number 1
  - Subcommittee 29, Working Group 11
- Develop standards for the coded representation of
  - moving picture and associated audio
- Sometimes collaborating with other standard organization
  - VCEG (ITU-T Video Coding Experts Group)
  - W3C (World Wide Web Consortium)
  - Web3D (Web3D Consortium, precious VRML)

# Overview of MPEG Standards

- MPEG-1 (1992)
  - Coding of video and audio for storage media (CD-ROM, 1.5Mbps)
  - VCD, MP3
- MPEG-2 (1994)
  - Coding of video and audio for transport and storage (4~80Mbps)
  - Digital TV (HDTV) and DVD
- MPEG-4 (v1:1999, v2: 2000, v3: 2001)
  - Coding of natural and synthetic media objects
  - Web and mobile applications
- MPEG-7 (2001~)
  - Multimedia content description for AV materials
  - Media searching and filtering
- MPEG-21 (2001~)
  - Multimedia framework for integration of multimedia technologies
  - Transparent and augmented use of multimedia resources

# MPEG-1 System

- Standard had three parts/layers: Video, Audio, and System (control interleaving of streams)
- combines one or more data streams from the video and audio parts with timing information to form a single stream suited to digital storage or transmission

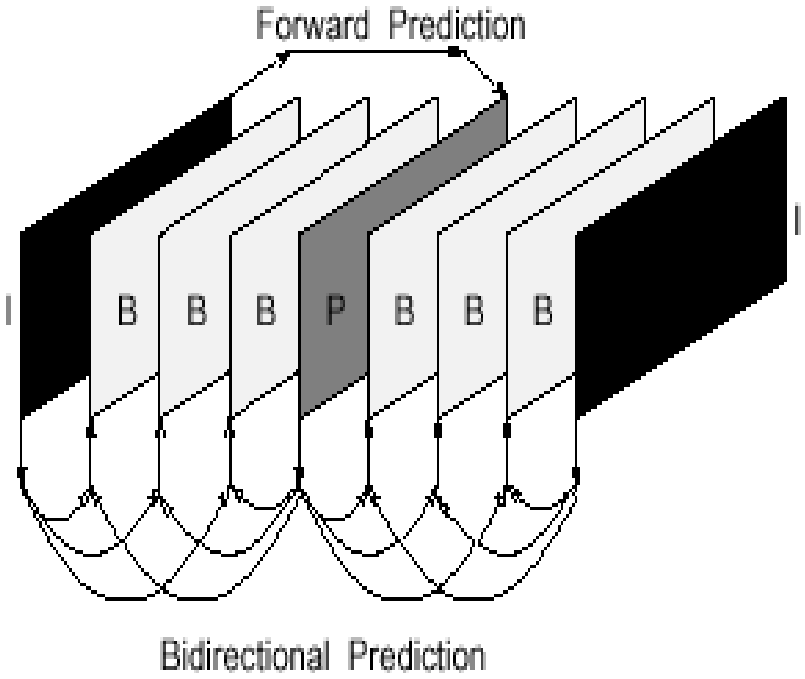


# MPEG-1 Video Layer

- For compressing video (NTSC 625-line and 525-lines)
- CIF/SIF (352x288/240)
- YCrCb: **4:2:0** sub-sampling
- Storage media at continuous rate of about **1.5 Mbps**
- **Intra-frame encoding**: DCT-based compression for the reduction of spatial redundancy (similar to JPEG)
- **Inter-frame encoding**: block-based *bidirectional* motion compensation for the reduction of temporal redundancy
- The difference signal, the prediction error, is further compressed using the discrete cosine transform (DCT) to remove spatial correlation and is then quantized.
- Finally, the motion vectors are combined with the DCT information, and coded using variable length codes

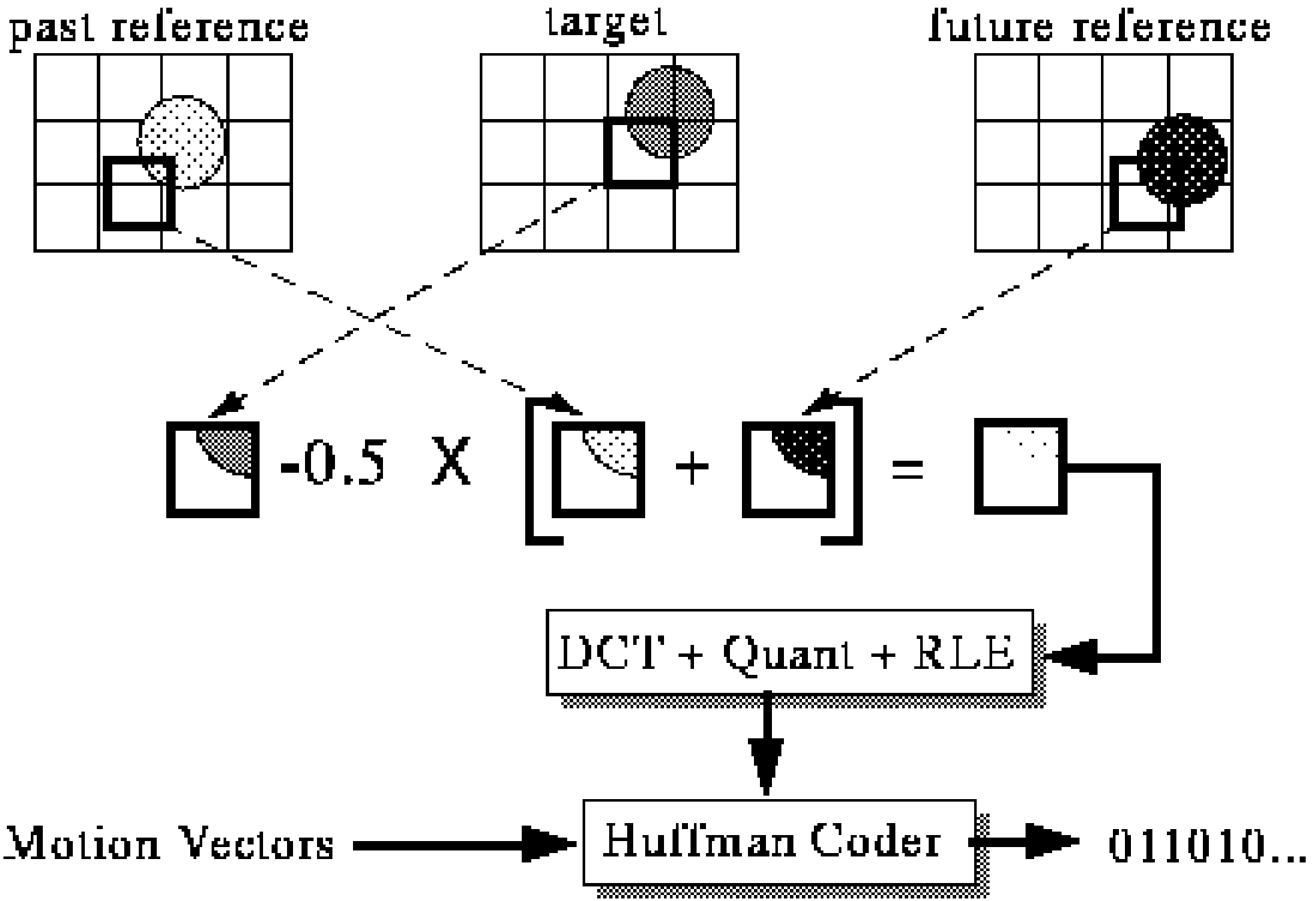
# Frame Sequence of MPEG-1

- **I-frames**
  - Intra-coded frames providing access points for random access
  - Moderate compression
- **P-frames**
  - Predicted frames with reference to a previous **I** or **P** frame
- **B-frames**
  - Bidirectional frames encoded using the previous and the next **I/P** frames
  - Maximum compression

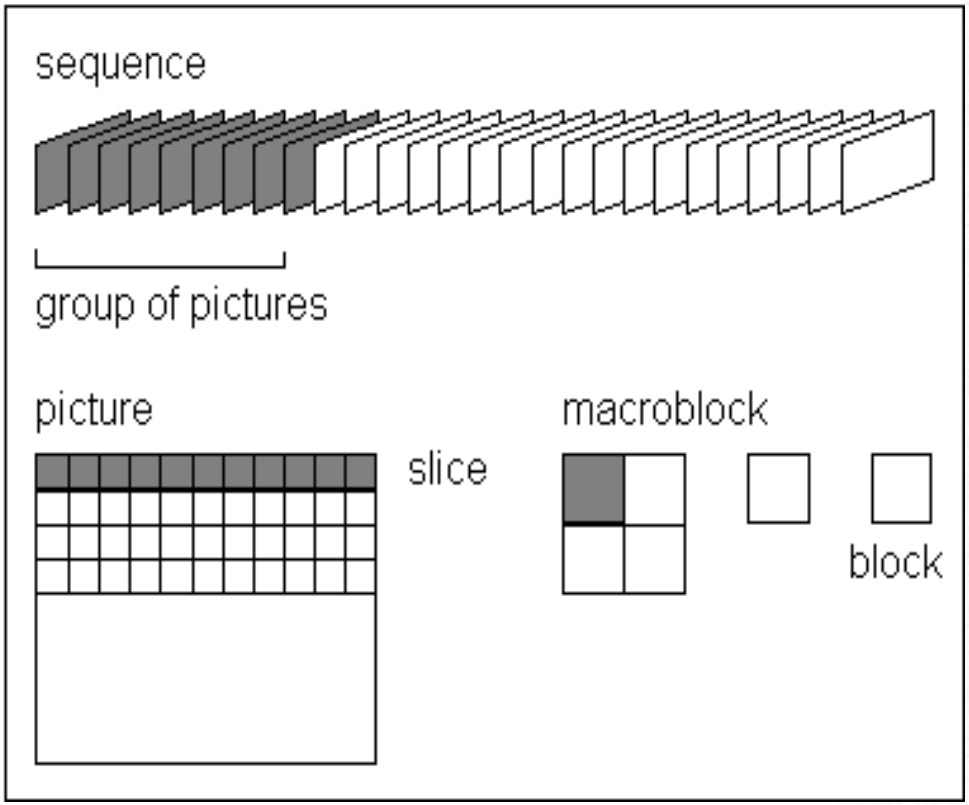


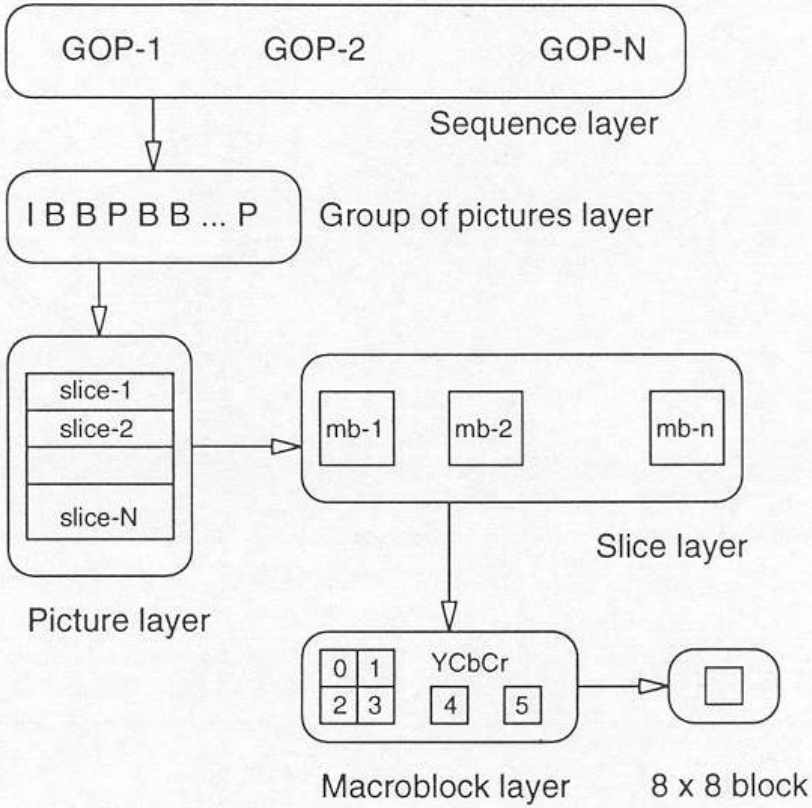

| Fr. Type | Size | Compr Ratio |
|---|---|---|
| I | 18 KB | 7:1 |
| P | 6 KB | 20:1 |
| B | 2.5 KB | 50:1 |
| Average | 4.8 KB | 27:1 |

# Bidirectional Motion Compensation

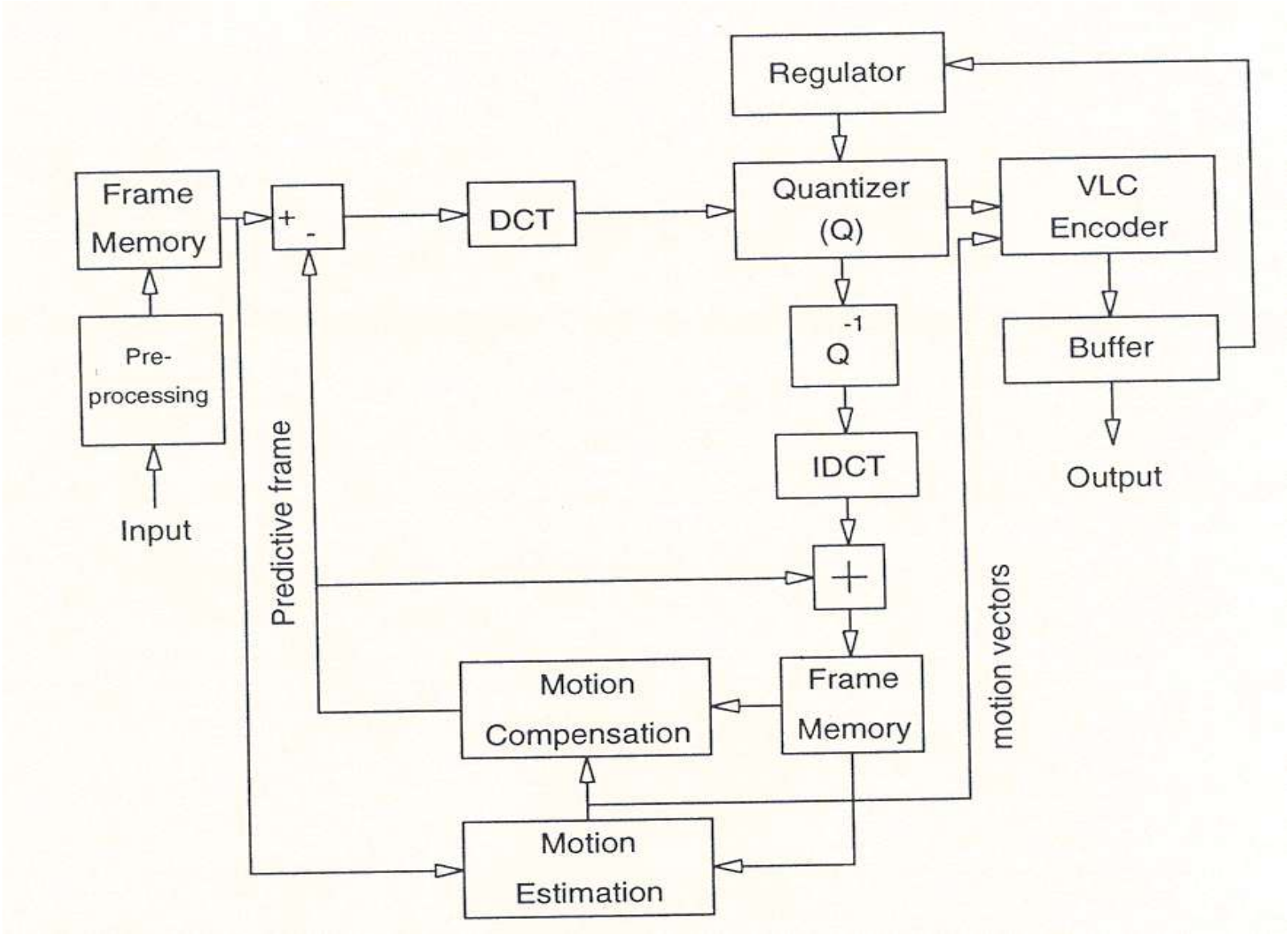


# Syntax Layers in MPEG-1

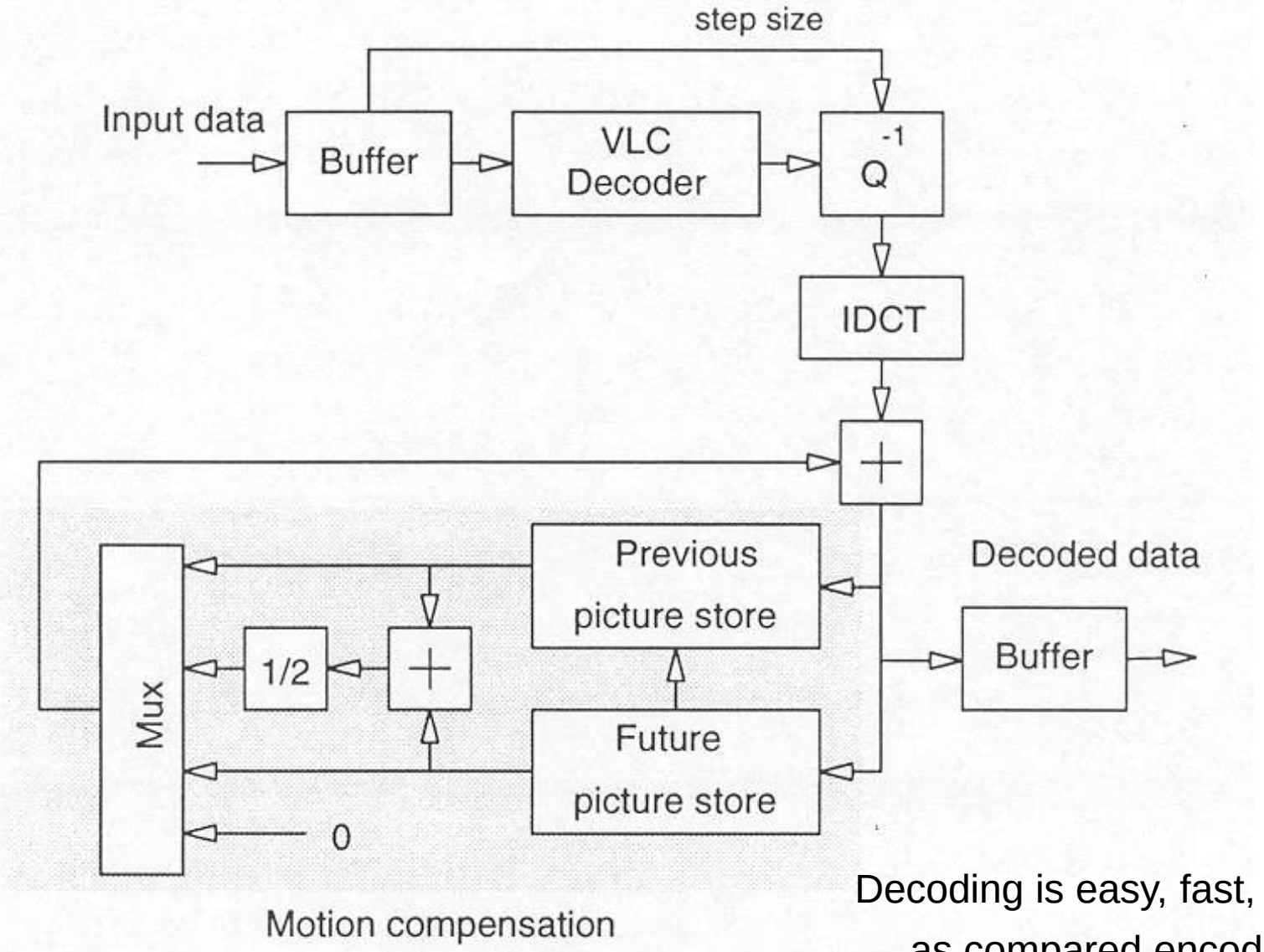


# MPEG-1 Encoder



# MPEG-1 Decoder




Decoding is easy, fast, cheap as compared encoding

# Differences from H.261

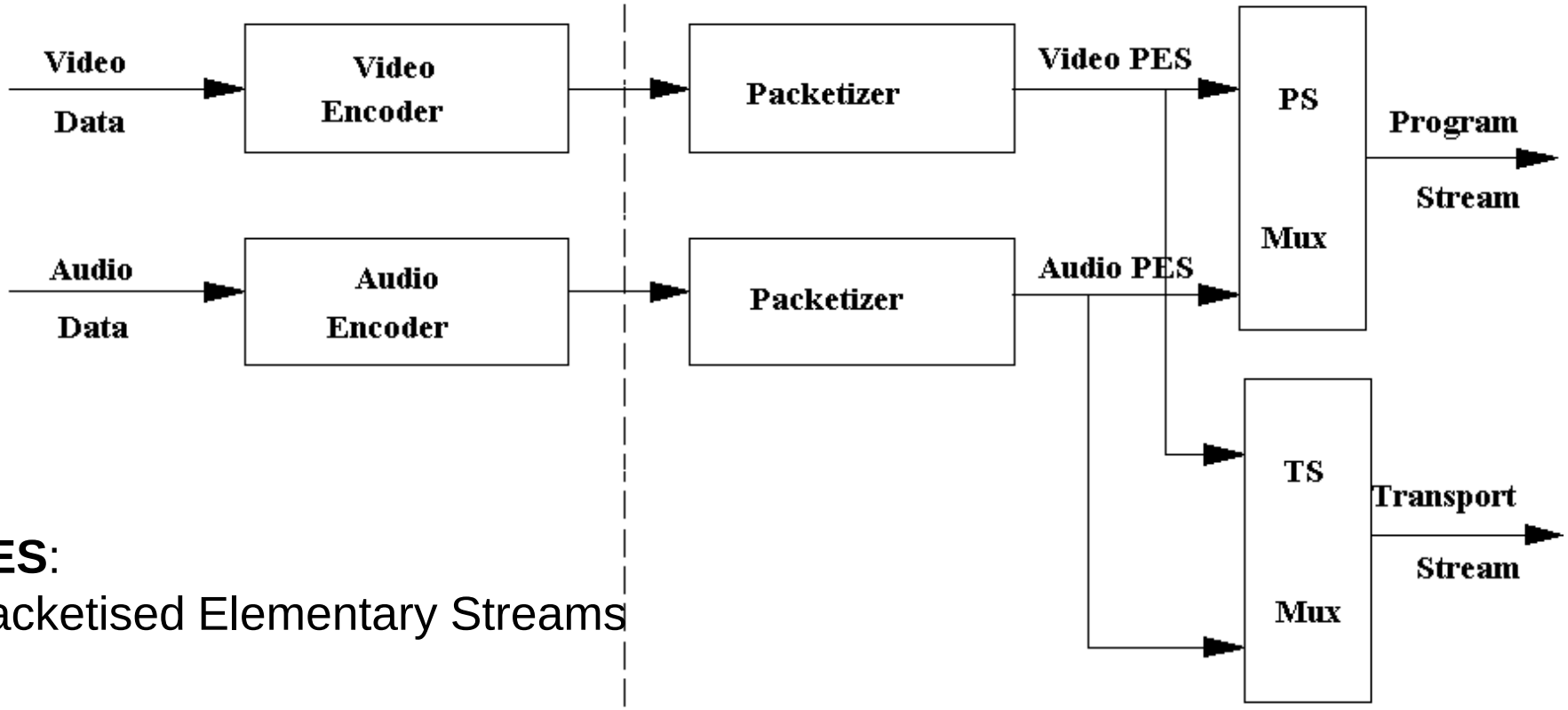
- Larger gaps between I and P frames, so need to expand motion vector search range.
- To get better encoding, allow motion vectors to be specified to fraction of a pixel (1/2 pixel).
- Bitstream syntax must allow random access, forward/backward play, etc.
- Added notion of *slice* for synchronization after loss/corrupt data.
- B frame macroblocks can specify *two* motion vectors (one to past and one to future), indicating result is to be averaged.

# MPEG-2

- Unlike MPEG-1 which is basically a standard for storing and playing video on a single computer
- MPEG-2 is a standard for digital TV (HDTV and DVD)

| Level | Size | Pixels/sec | Bit-rate (Mb/s) | Application |
|---|---|---|---|---|
| Low | 352 x 288 x 30 | 3 M | 4 | VHS, TV |
| Main | 720 x 576 x 30 | 12 M | 15 | Studio TV |
| High 1440 | 1440 x 1152 x 60 | 96 M | 60 | Consumer HDTV |
| High | 1920 x 1152 x 60 | 128 M | 80 | HDTV, Film |

# MPEG-2 System

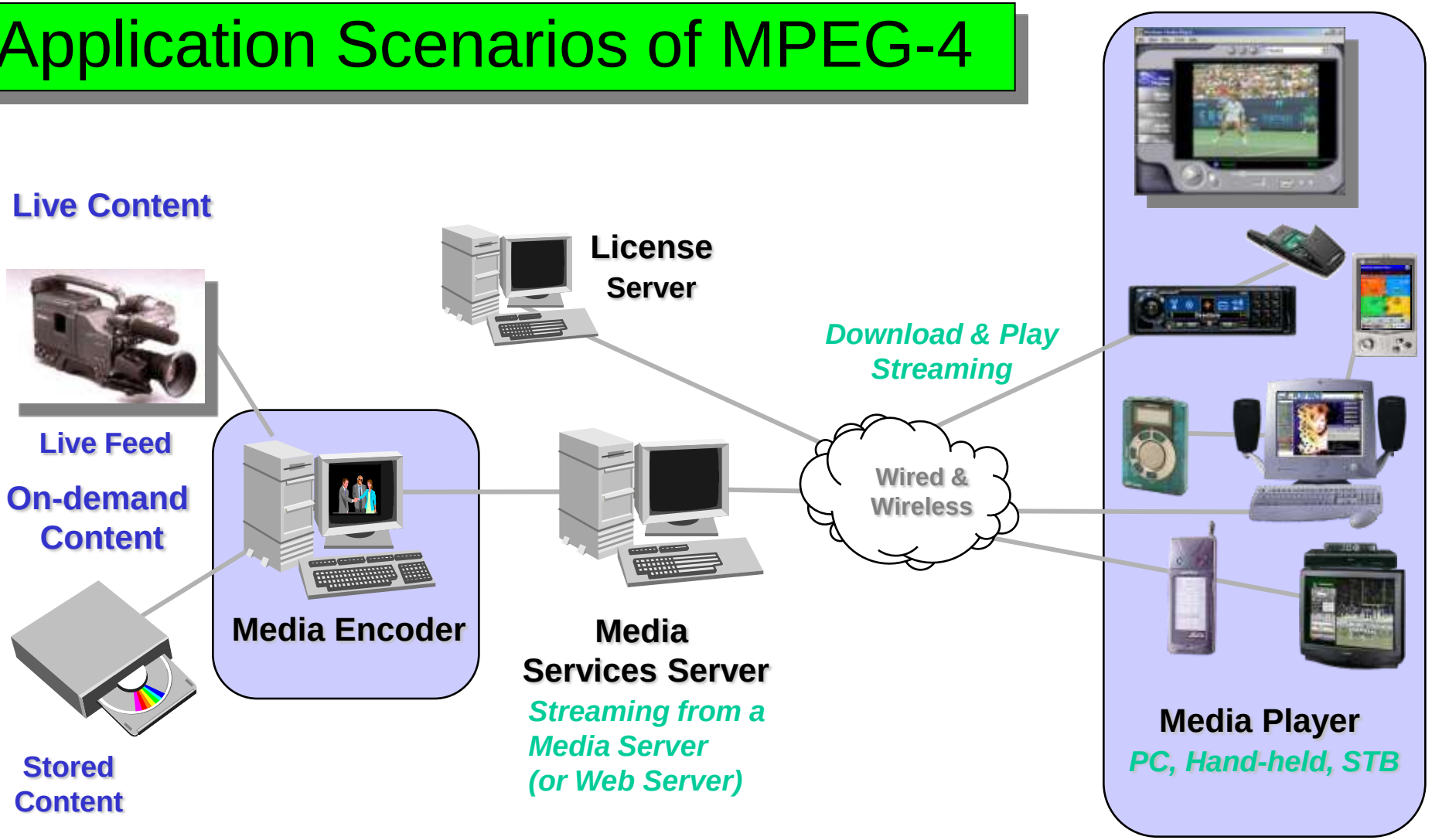



**PES**:
Packetised Elementary Streams

# New Features in MPEG-2

- Support both field prediction and frame prediction.
- Besides 4:2:0, also allow 4:2:2 and 4:4:4 subsampling
- Scalable Coding
  - SNR Scalability -- similar to JPEG Progressive mode, adjusting the quantization steps of the DCT coefficients (image quality)
  - Spatial Scalability -- similar to hierarchical JPEG, multiple spatial resolutions (image size: CIF, SDT to HDTV).
  - Temporal Scalability -- different frame rates (5~60f/s)
- Many minor fixes

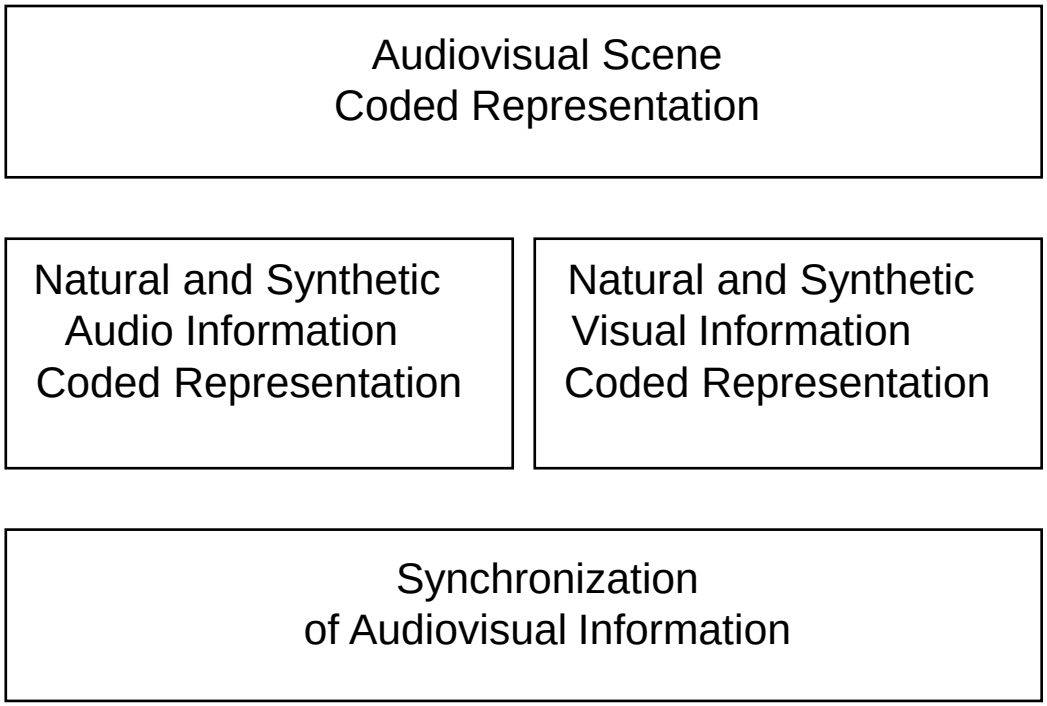
# Application Scenarios of MPEG-4



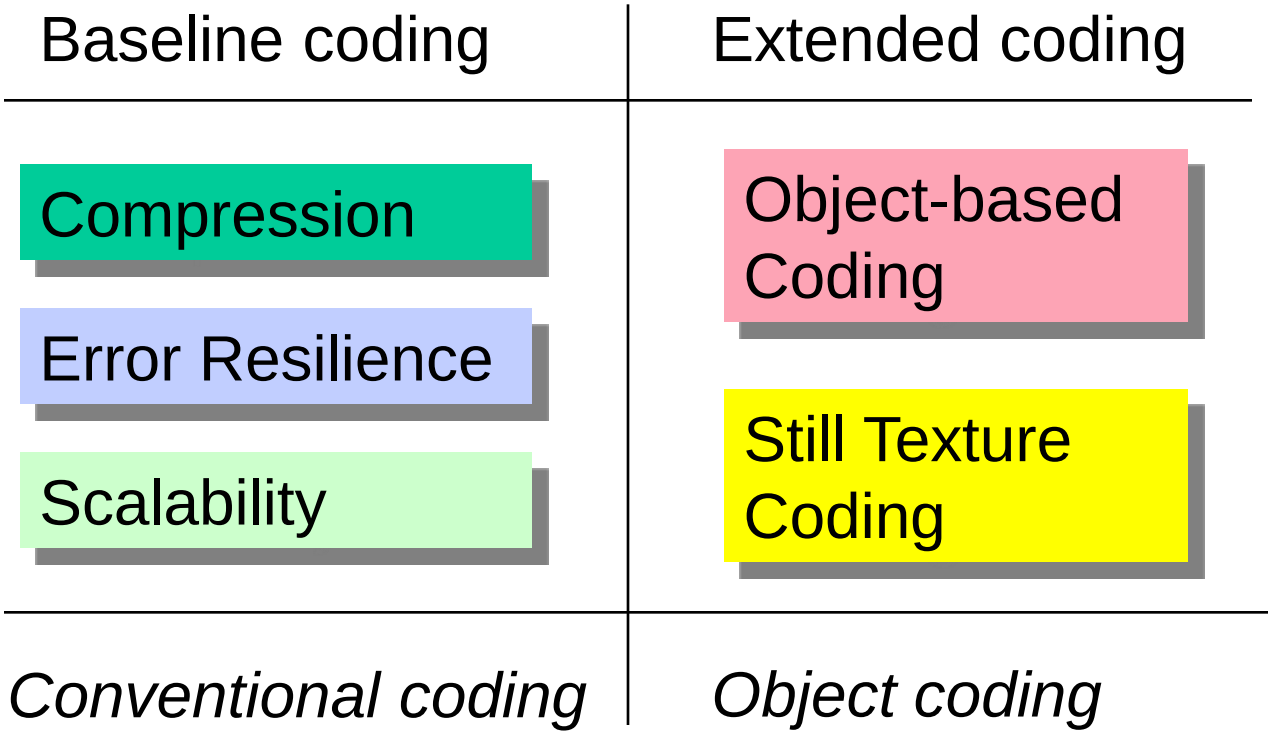

Compression → Access → Interaction

# Overview of MPEG-4

- The **coded representation** of the combination of streamed elementary audiovisual information
- 1) **Compression**, 2) **Content-based interactivity**, 3) **Universal access**
- To provide a bridge between the Web and conventional AV media
- To delivery streaming AV media on the Internet and wireless networks



# MPEG-4 Video Coding




Natural visual coding for captured pictures
Synthetic visual coding for graphic/animation pictures
Synthetic/Natural Hybrid Coding (SNHC) for the mixed two

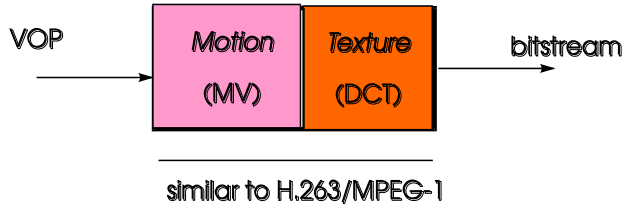
# Integration of Natural and Synthetic Contents

Augmented/Mixed Reality

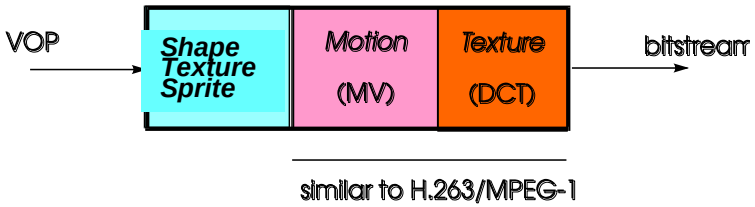
# Baseline and Extended Coding

VOP: Visual Object Plane (MPEG-4 term for a frame)

MPEG-4 Core Coder




Extended MPEG-4 Core Coder



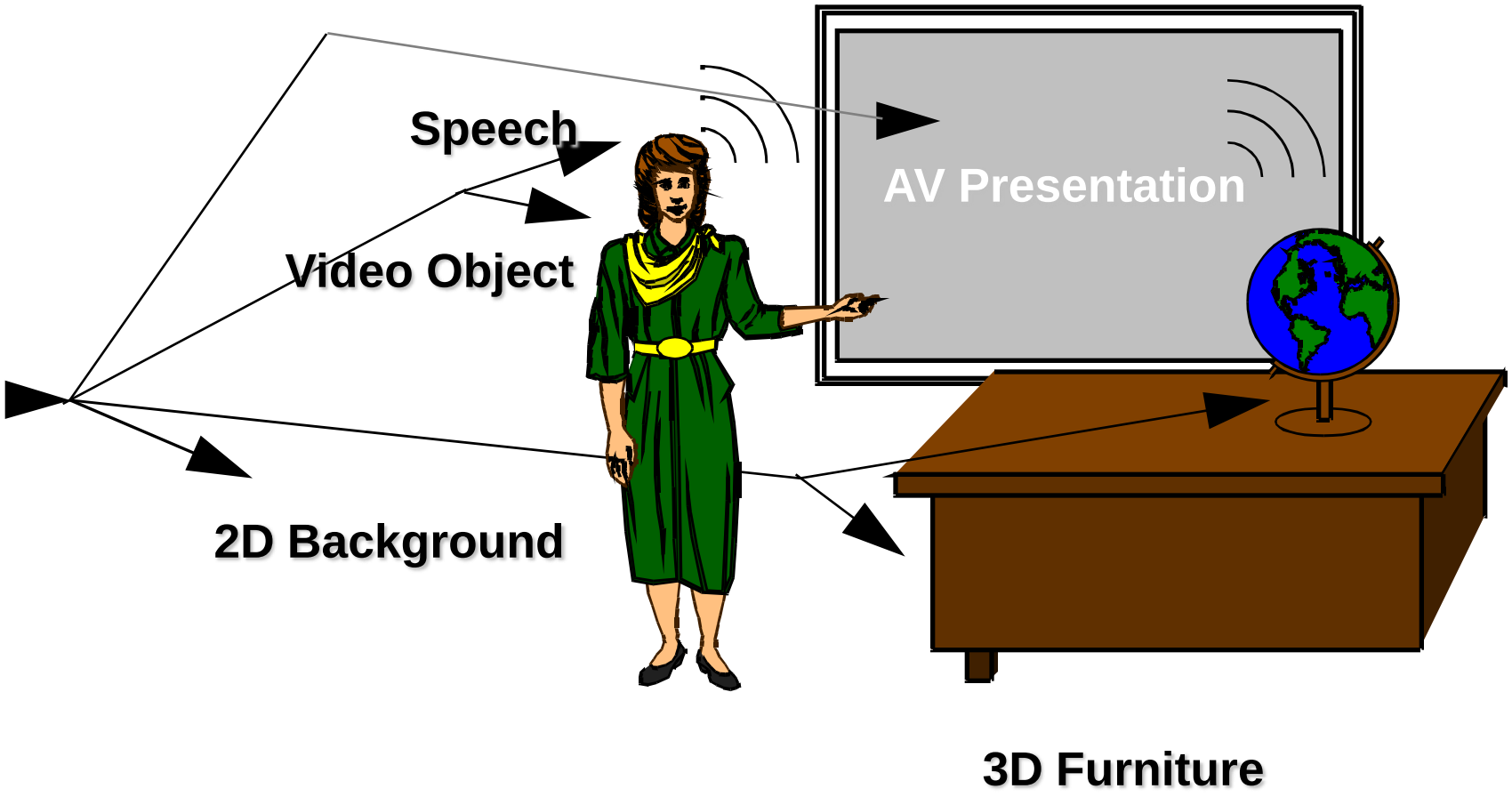
# MPEG-4 Baseline Coding

- Support both progressive and interlaced scanning
- Arbitary size from 8x8 to 2048x2048
- YCrCb: 4:0:0, 4:2:0, 4:2:2 and 4:4:4
- Continuously various frame rate
- Bit rates: 5Kbps ~ 1Gbps from very small TV to Studio TV
  - low (<64Kbps), intermediate (64~484kbps)
    high (384K~4Mbps) and very high (>4Mbps)
- MPEG-4 Video is Compatible to Baseline H.263
- And Almost Compatible to MPEG-1
- And almost compatible to MPEG-2
- Better coding efficiency than MPEG-1/2 and H.263

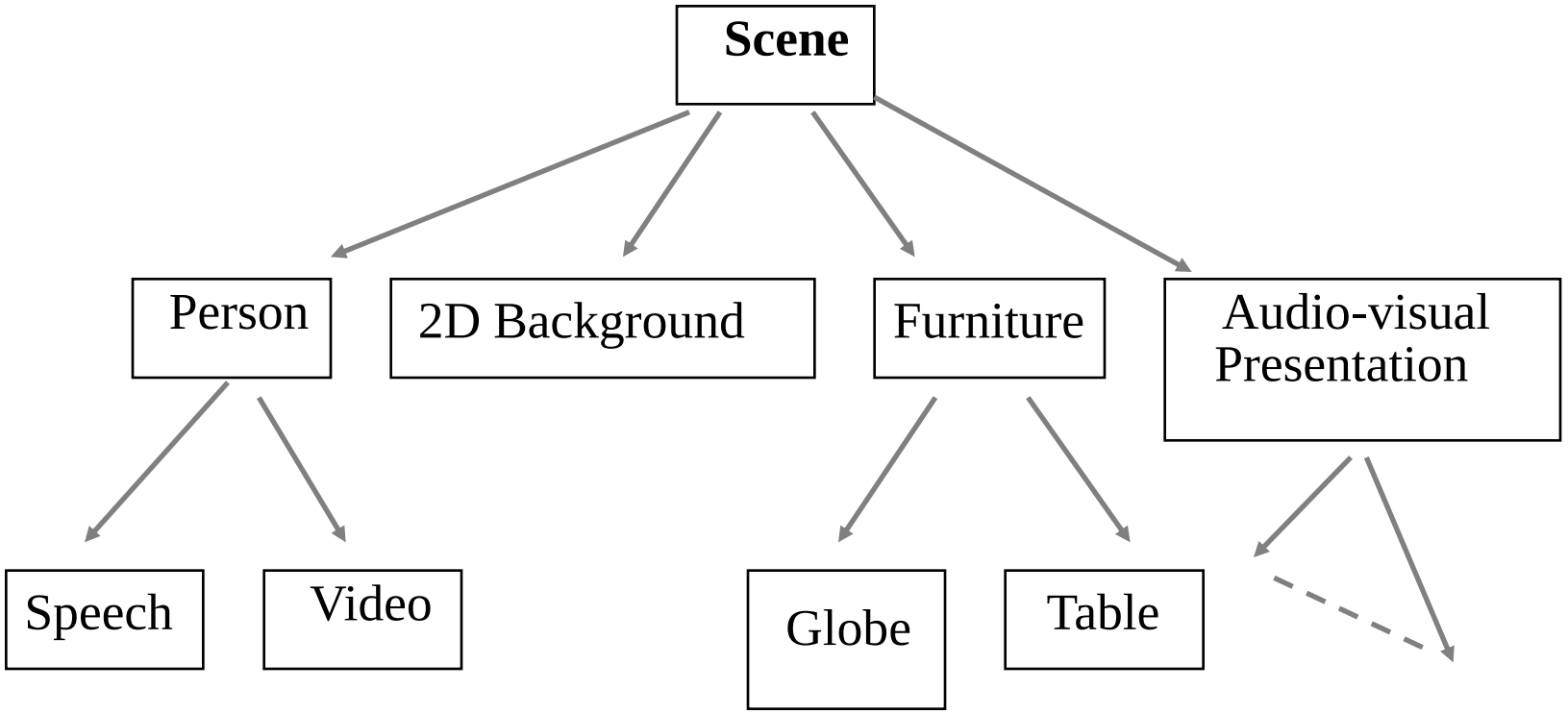
# - Extended Functionalities - *Object-Based Coding of Video*

- Object-Based Coding = Content-Based Coding
- Object-based coding increases compression efficiency
- Object-based coding allows the user to access arbitrarily-shaped objects in a coded scene
- Object-based coding enables high interaction with scene content
- Manipulation of scene content on bitstream level

# Objects in Audio-Visual Scene

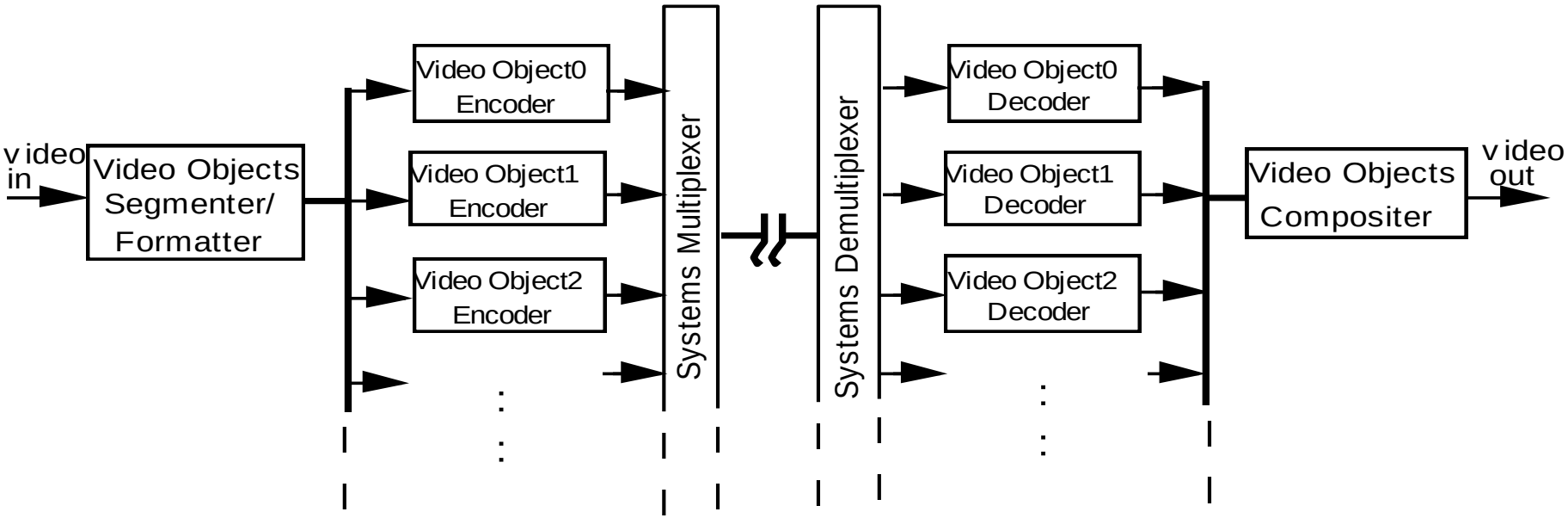


# BIFS – BInary Format for Scene

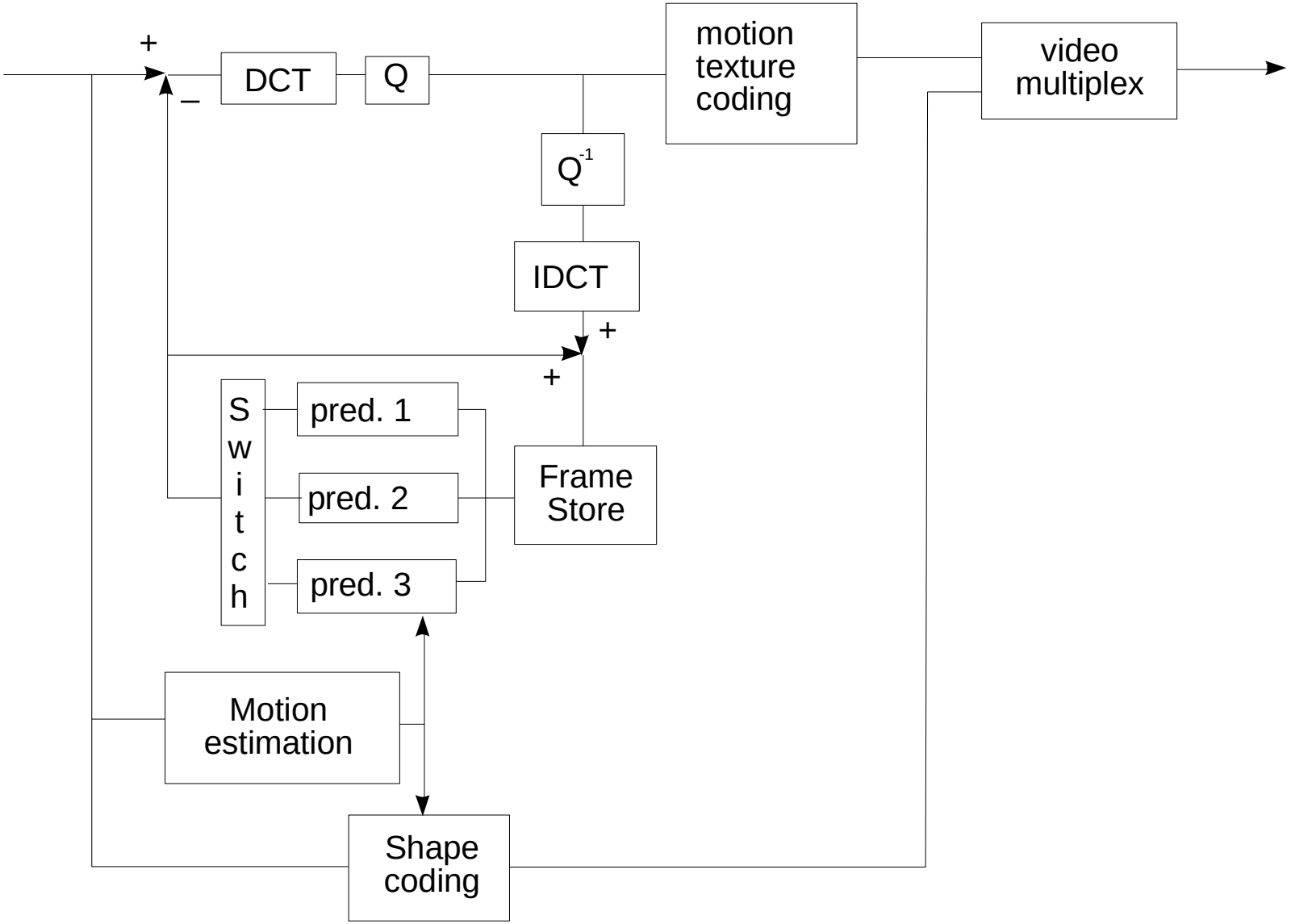


# Object-Based Coding

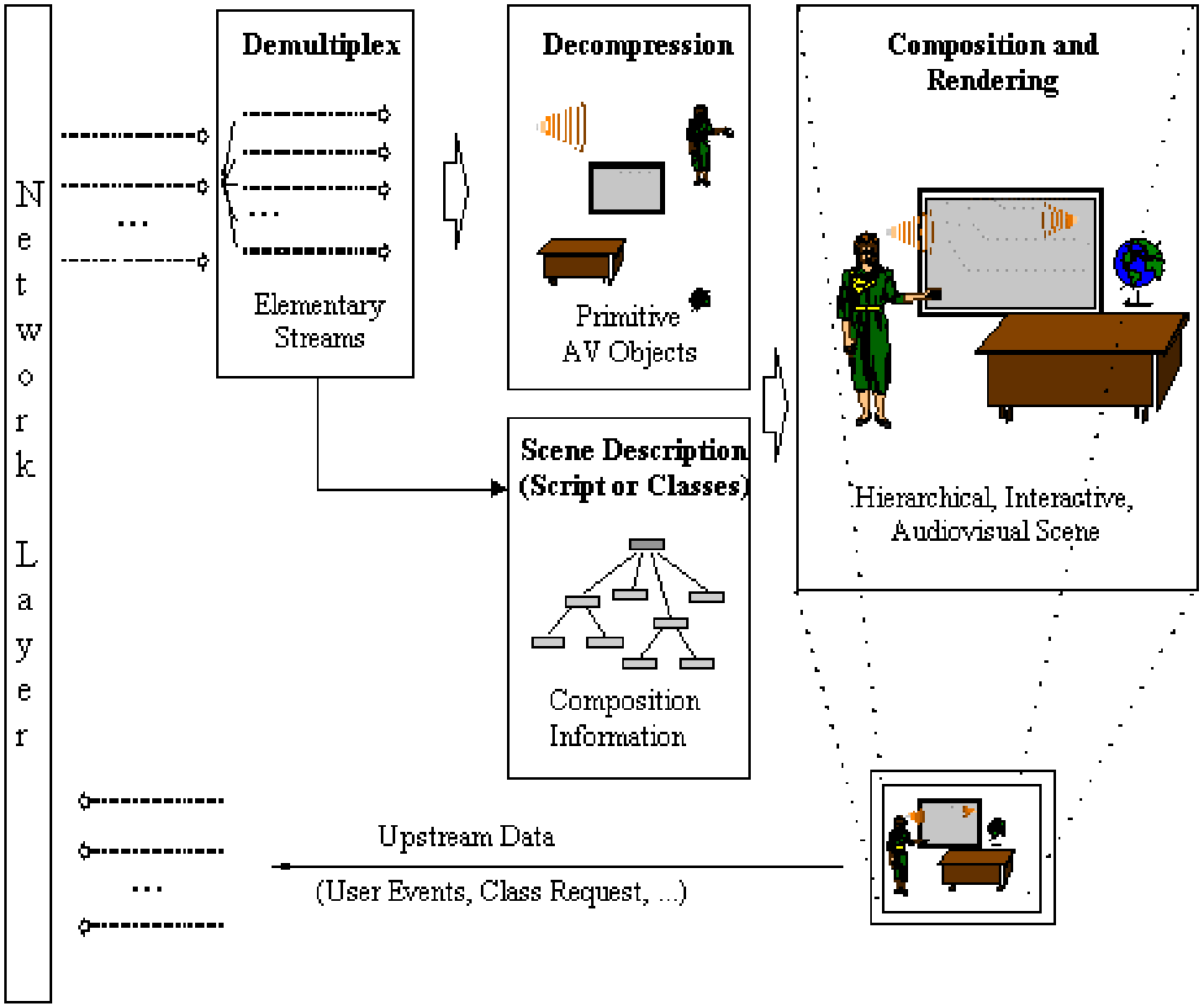
- Each video object in a scene is coded and transmitted separately



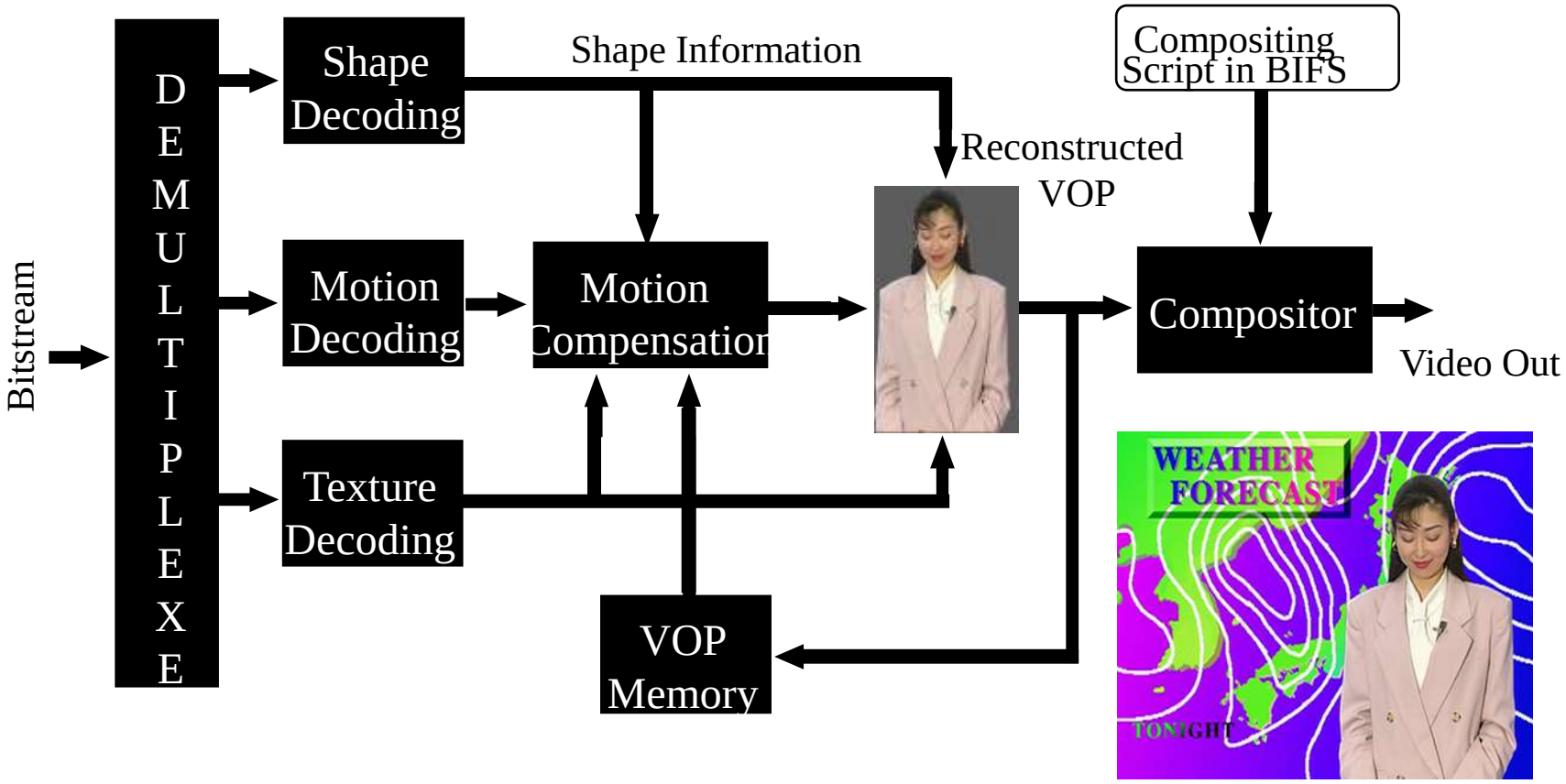
# Object-Based Encoding



# Scene Reconstruction



# Example of Video Decoding



# Sprite Coding

- Original in computer graphics
- Long term background objects
- Real time rotation, translation, zooming

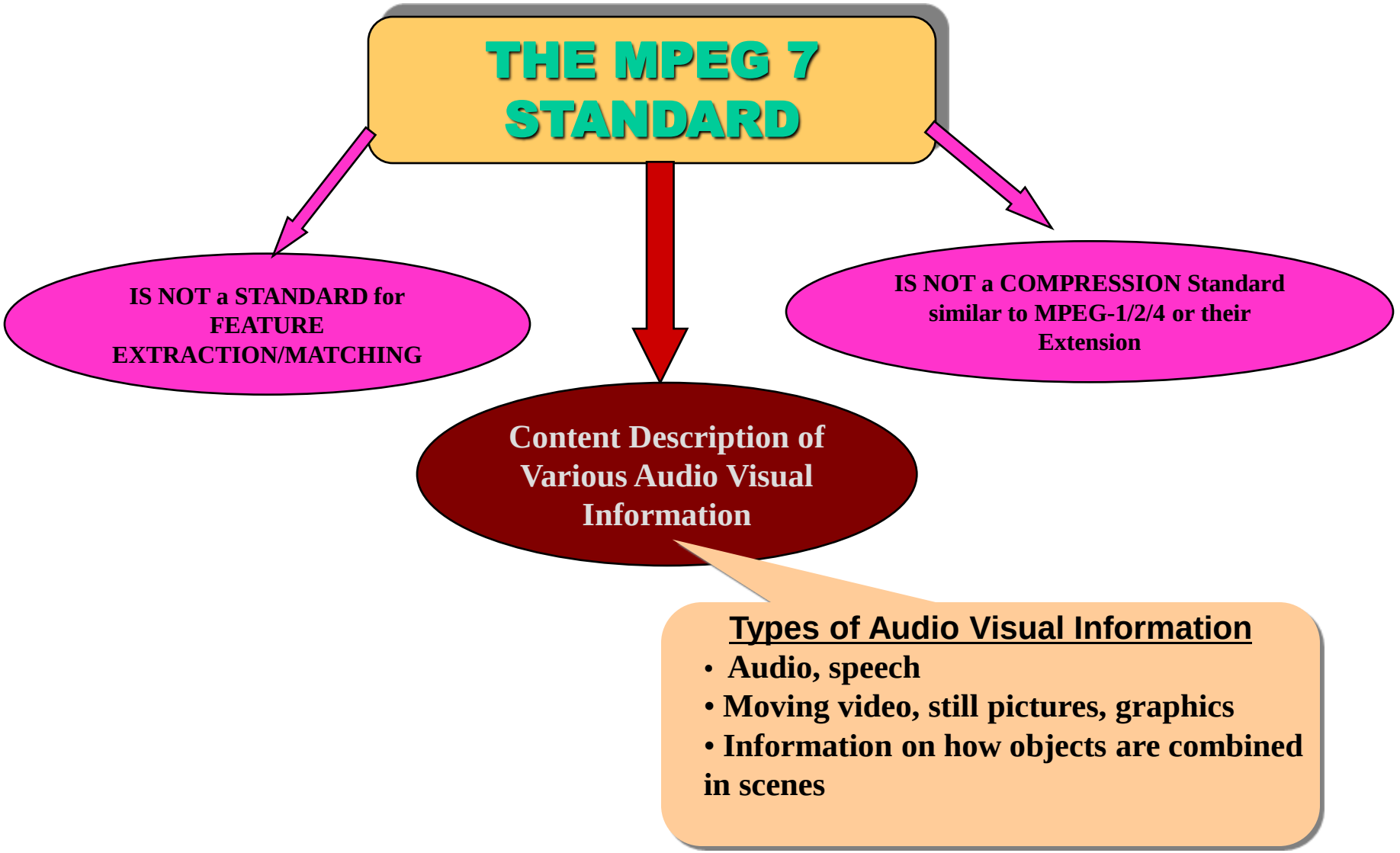

sprite

player

# Various Applications of MPEG-4

- **IVS Internet Video Streaming**
- **VA Video Archive**
- **VCD Video Content Distribution**
- **IMM Internet Multimedia**
- **IVG Interactive Video Games**
- **IPC Interpersonal Communications (videoconferencing, videophone, etc.)**
- **ISM Interactive Storage Media (optical disks, etc.)**
- **MMM Multimedia Mailing**
- **NDB Networked Database Services (via ATM, etc.)**
- **WMM Wireless Multimedia**

# MPEG-7: What Is It ?



# Why do we need MPEG-7 ?

*Support for Advanced Query*

**N e e d**

- *Fast & Accurate Access*
- *Personalized Content Production and Consumption*
- *Content Management*
- *Automation*

+

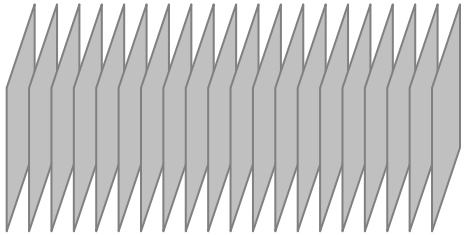
- **Visual**
- **Audio**
- **Sketch**

# Main Elements of MPEG-7

- Descriptors (D)
  - syntax and semantics of each feature representation
- Description Schemes (DS)
  - structure and semantics of the relationships between components
- Description Definition Language (DDL)
  - creation of new DS's
  - modification/extension  of existing DS's

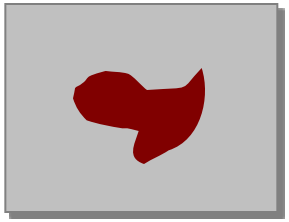
# Low level Audio and Visual descriptors

## Video segments



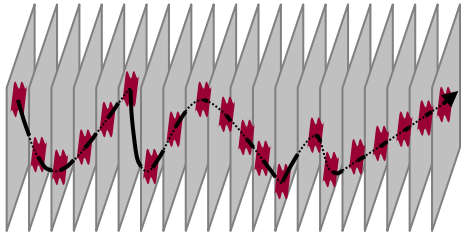
Contents
- Color
- Camera motion
- Motion activity
- Mosaic

## Still regions



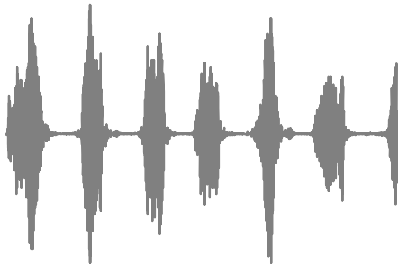
Contents
- Color
- Shape
- Position
- Texture

## Moving regions



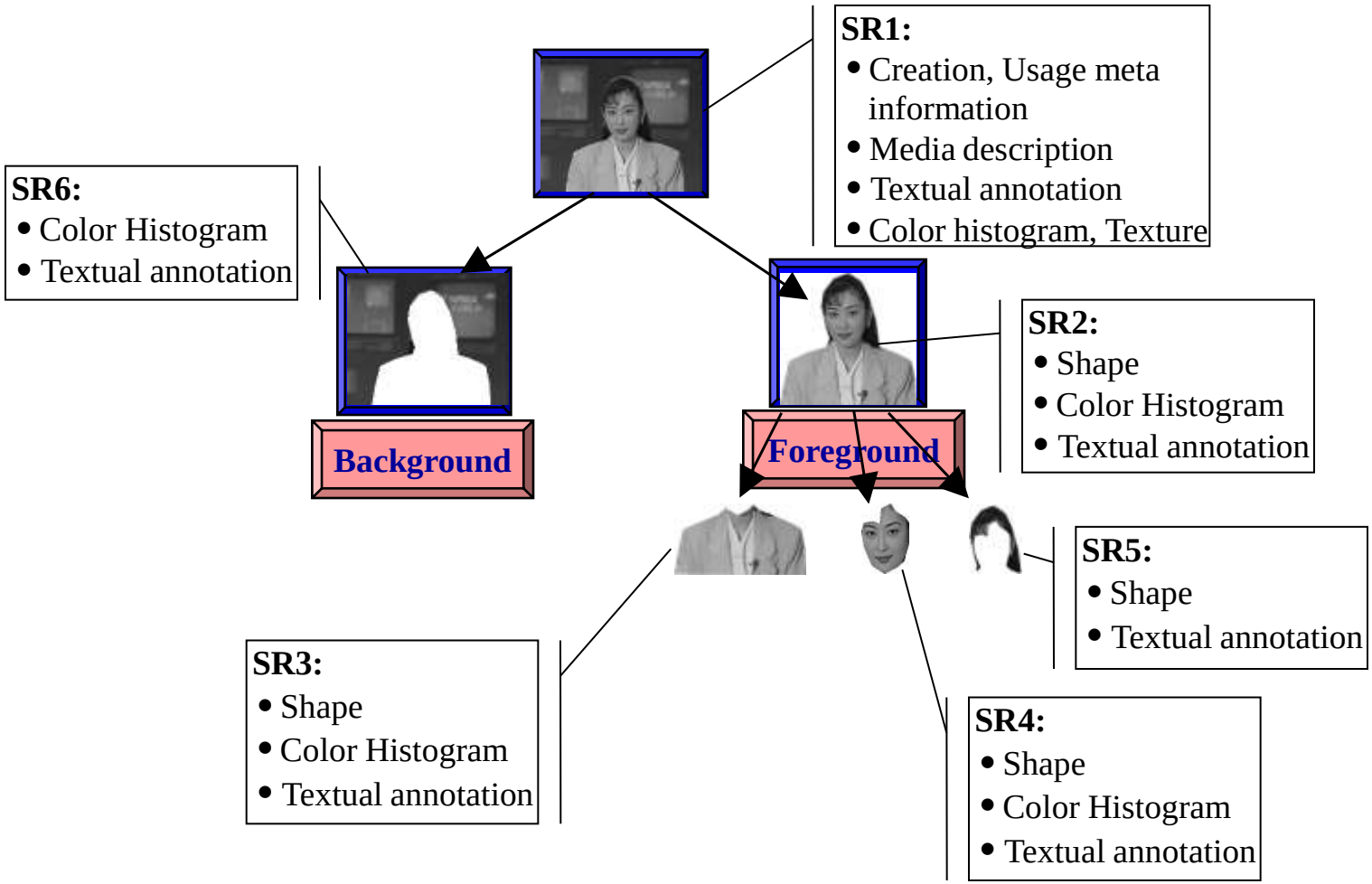
Contents
- Color
- Motion trajectory
- Parametric motion
- Spatio-temporal shape

## Audio segments

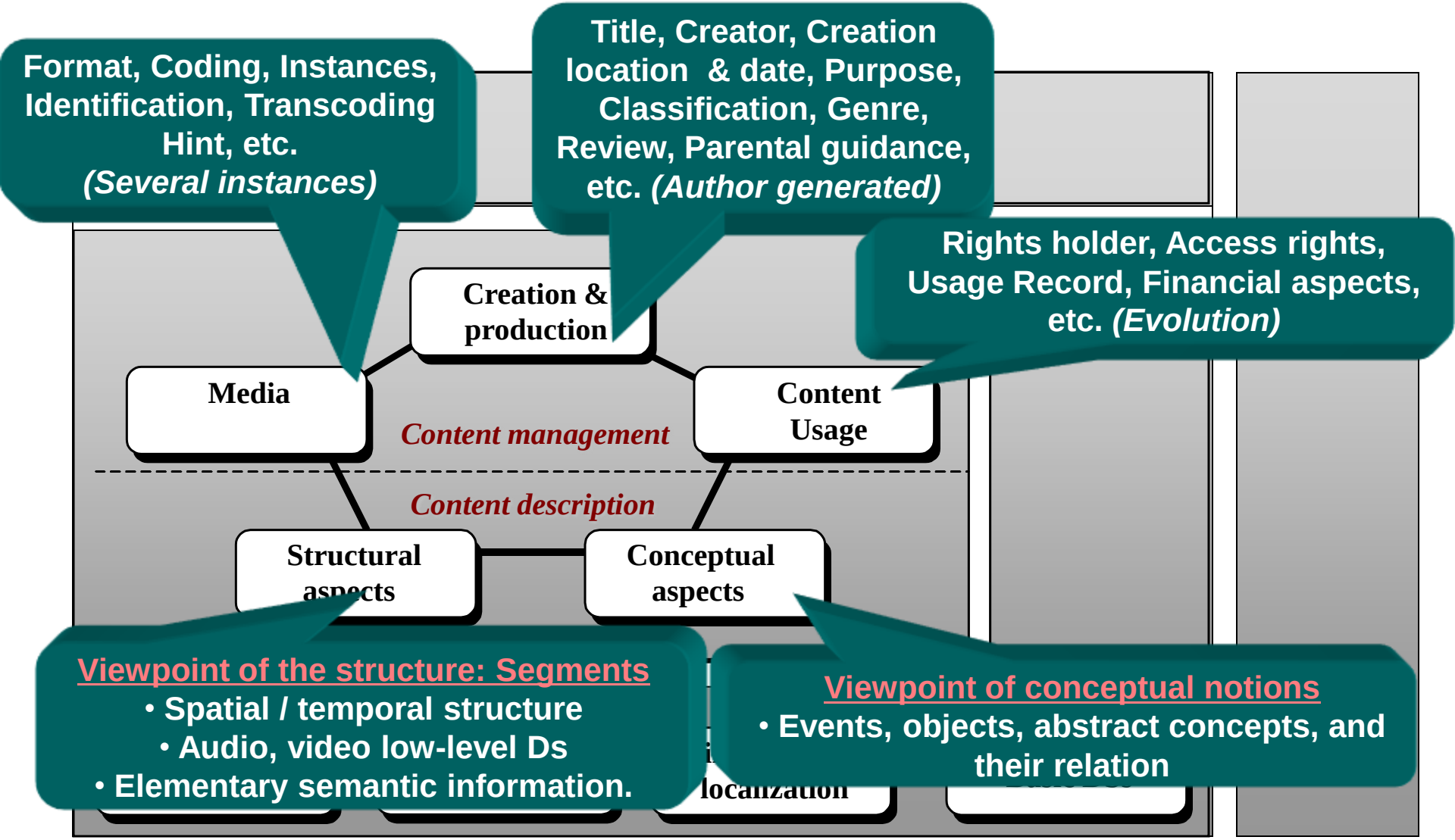


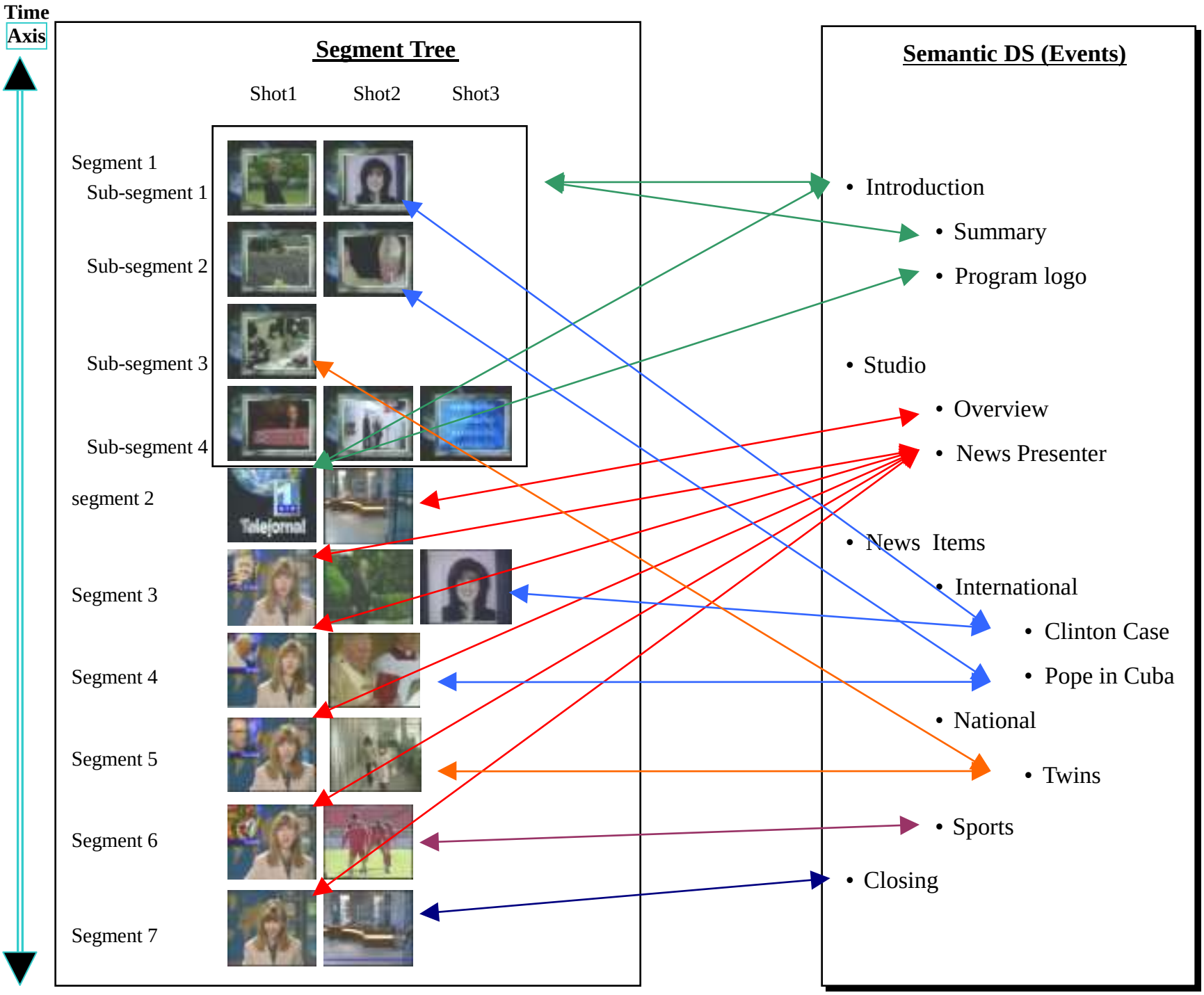
Contents
- Spoken content
- Spectral characterization
- Music: timbre, melody

# Low Level Descriptors and Segment Trees



# Content Management and Description




Time Axis
Segment Tree
Shot1
Shot2
Shot3
Segment 1
Sub-segment 1
Sub-segment 2
Sub-segment 3
Sub-segment 4
segment 2
Segment 3
Segment 4
Segment 5
Segment 6
Segment 7
Semantic DS (Events)
• Introduction
• Summary
• Program logo
• Studio
• Overview
• News Presenter
• News Items
• International
• Clinton Case
• Pope in Cuba
• National
• Twins
• Sports
• Closing

# MPEG-21

- Seeks to describe a multimedia framework and set out a vision for the future of an environment that is capable of supporting the delivery and use of all content types by different categories of users in multiple application domains
- Financial, content, consumer, technology, delivery applications
- MPEG-21 digital item – A structured digital object with a standard representation, identification and metadata with this framework. This entity is also the fundamental unit for distribution and transaction within this framework.
  - Digital Item Declaration
  - Digital Item Representation
  - Digital Item Identification and Description
  - Digital Item Management and Usage
  - Intellectual Property Management and Protection
  - Terminals and Networks
  - Event Reporting

# Demos of Video Coding