New Video Coding standards MPEG-4, HEVC

Georgios Tziritas

Computer Science Department

http://www.csd.uoc.gr/~tziritas

MPEG-4 : introduction

Motion Picture Expert Group
Publication 1998 (Intern. Standardization Organization)
Multimedia Communications, Internet

- **Improved coding efficiency over MPEG-2**
- **Ability to interact with the audio-visual scene at the receiver**
- **Integration of natural and synthetic objects**
- **Temporal, spatial and quality scalability**
- **Error resilience to enable robust transmission**

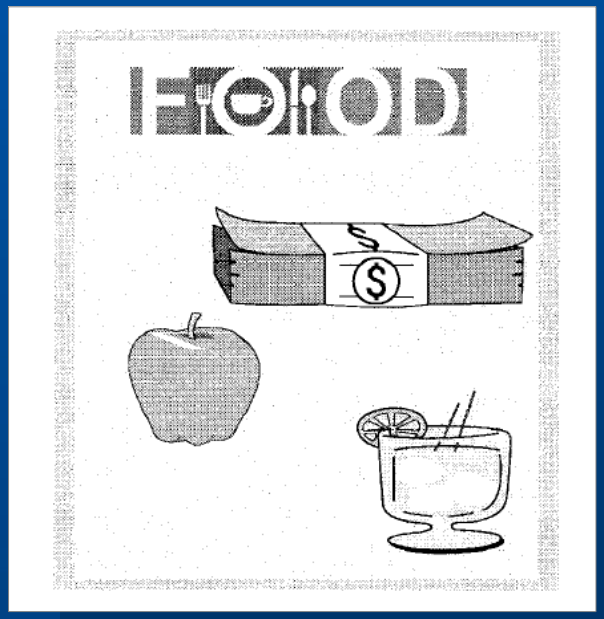
Covers from 5 kbits/s to 10 Mbits/s

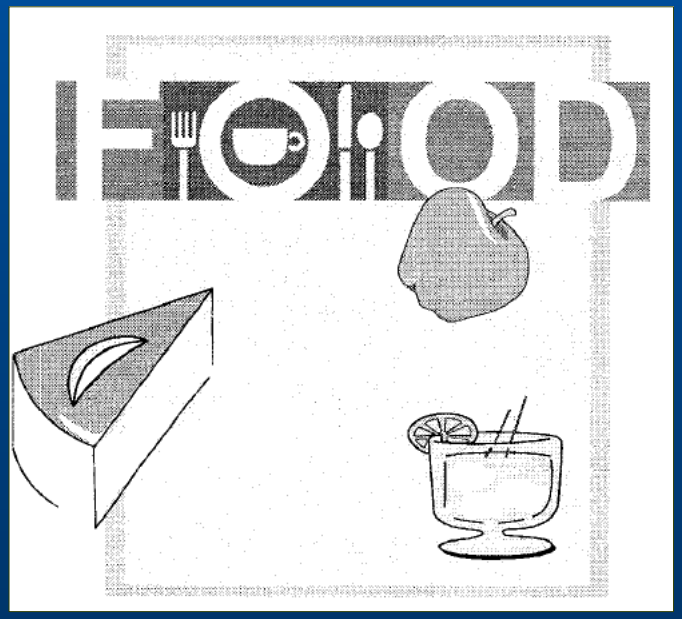
5-64 kbits/s for mobile communications
2 Mbits/s for TV broadcasting

Object processing

Object manipulation :

cut, paste, move, geometric transform, appearance change





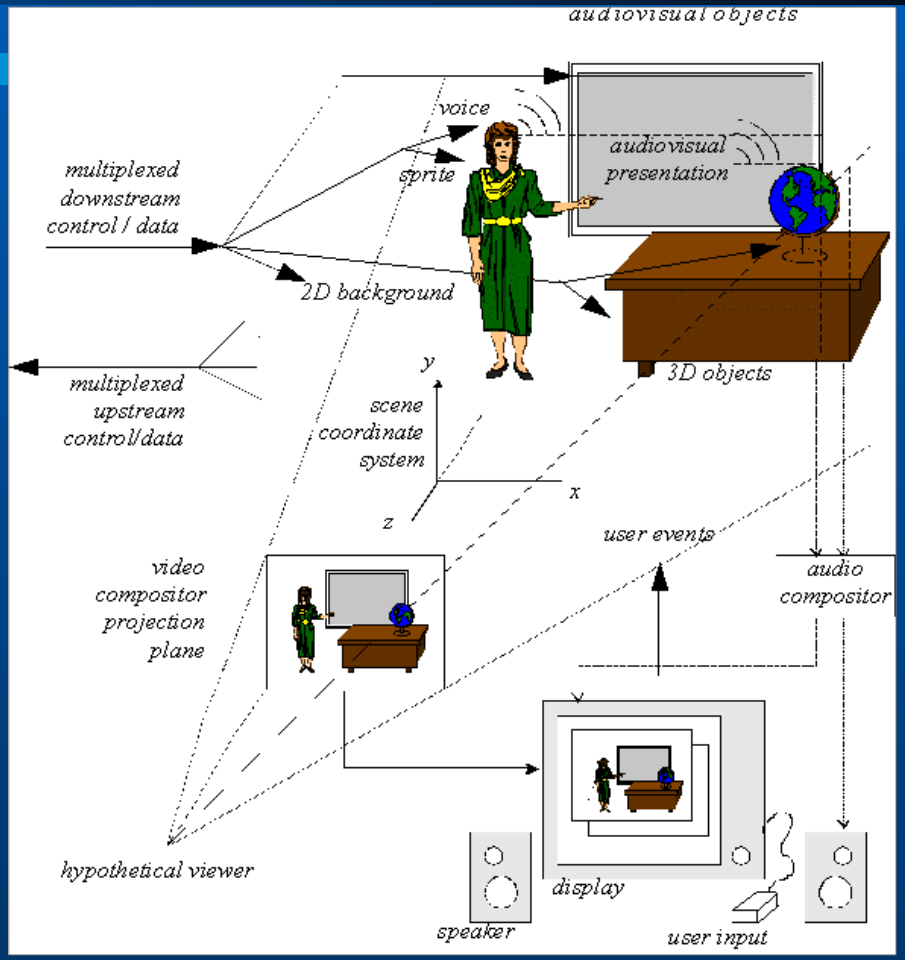
Audiovisual scenes rendering

Audiovisual object AVO :
Audio, visual, audiovisual

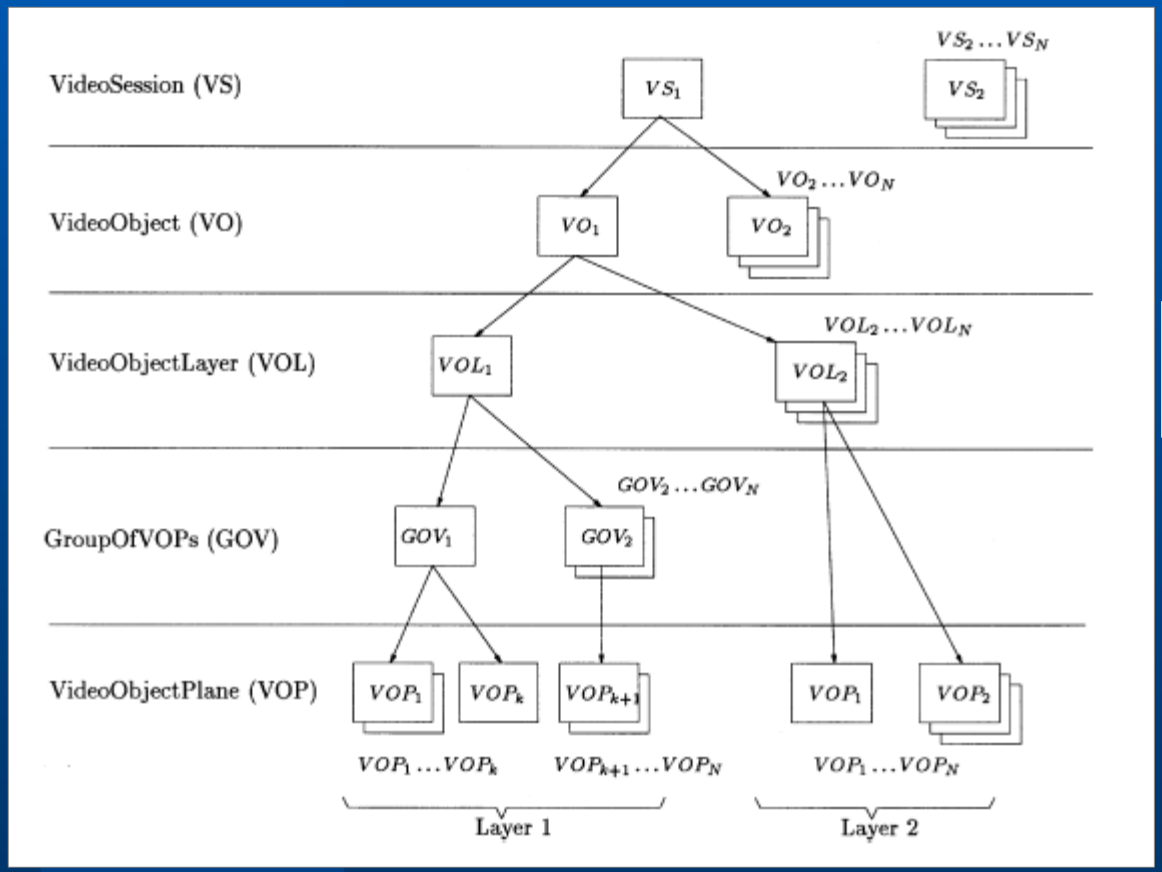
Description of scene and AVO

Hierarchical structure

Specific objects :
text, graphics,
speaking face,
animated human body



Hierarchical structure



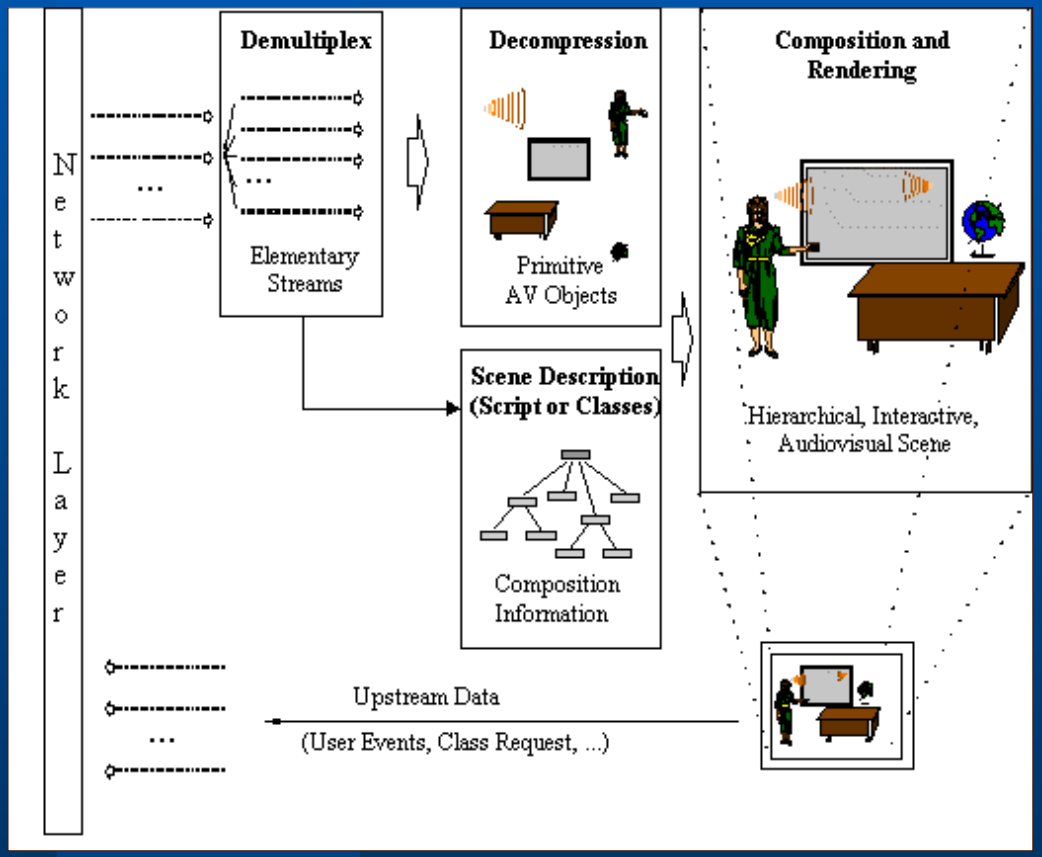
delivers the complete MPEG-4 visual scene

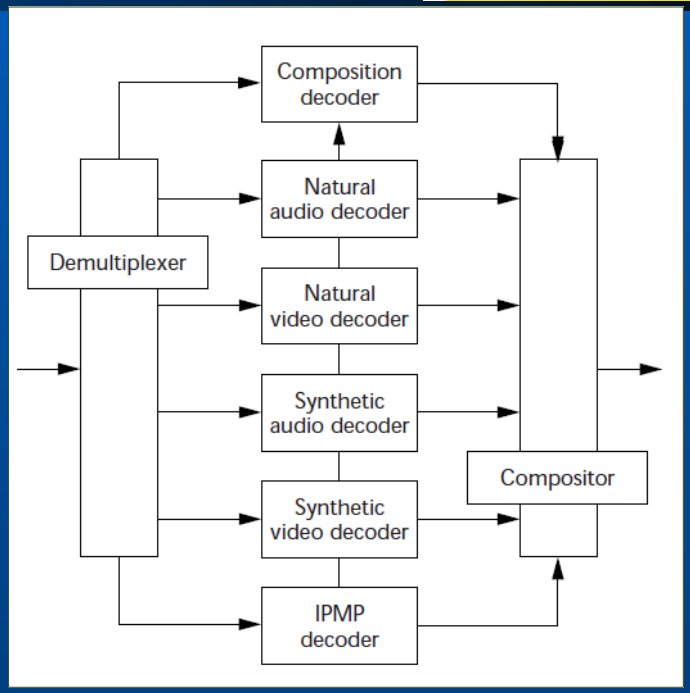
a particular object in the scene of arbitrary shape

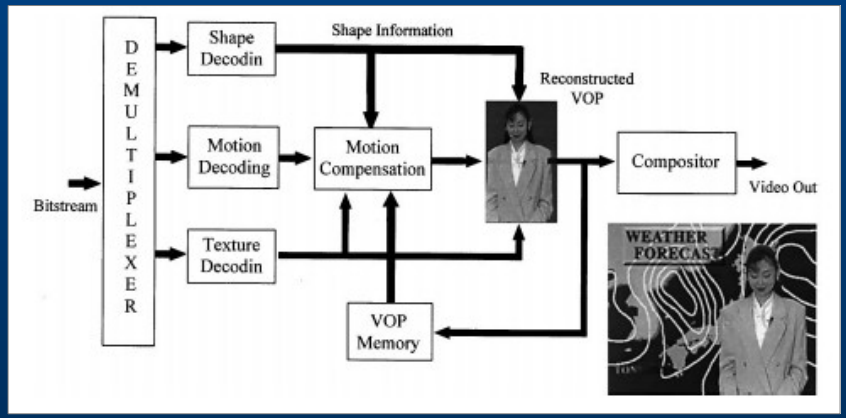
a way to support (multi-layered) scalable coding

a snapshot of a VO at a particular moment

Receiver







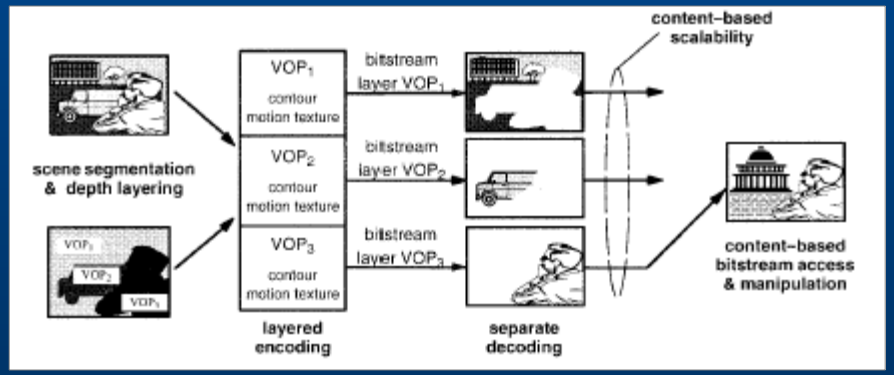
Spring 2018

Coding modes

- **Video object encoding**
  - **Similar to MPEG-1 (MPEG-2), in addition shape coding**
  - **Sprite panorama coding**
- **Mesh coding**
  - **Triangular mesh for articulated motion and temporal interpolation**
- **Model-based coding**
  - **Face and human body**
- **Static texture coding**
  - **Wavelet transform**

Video object coding

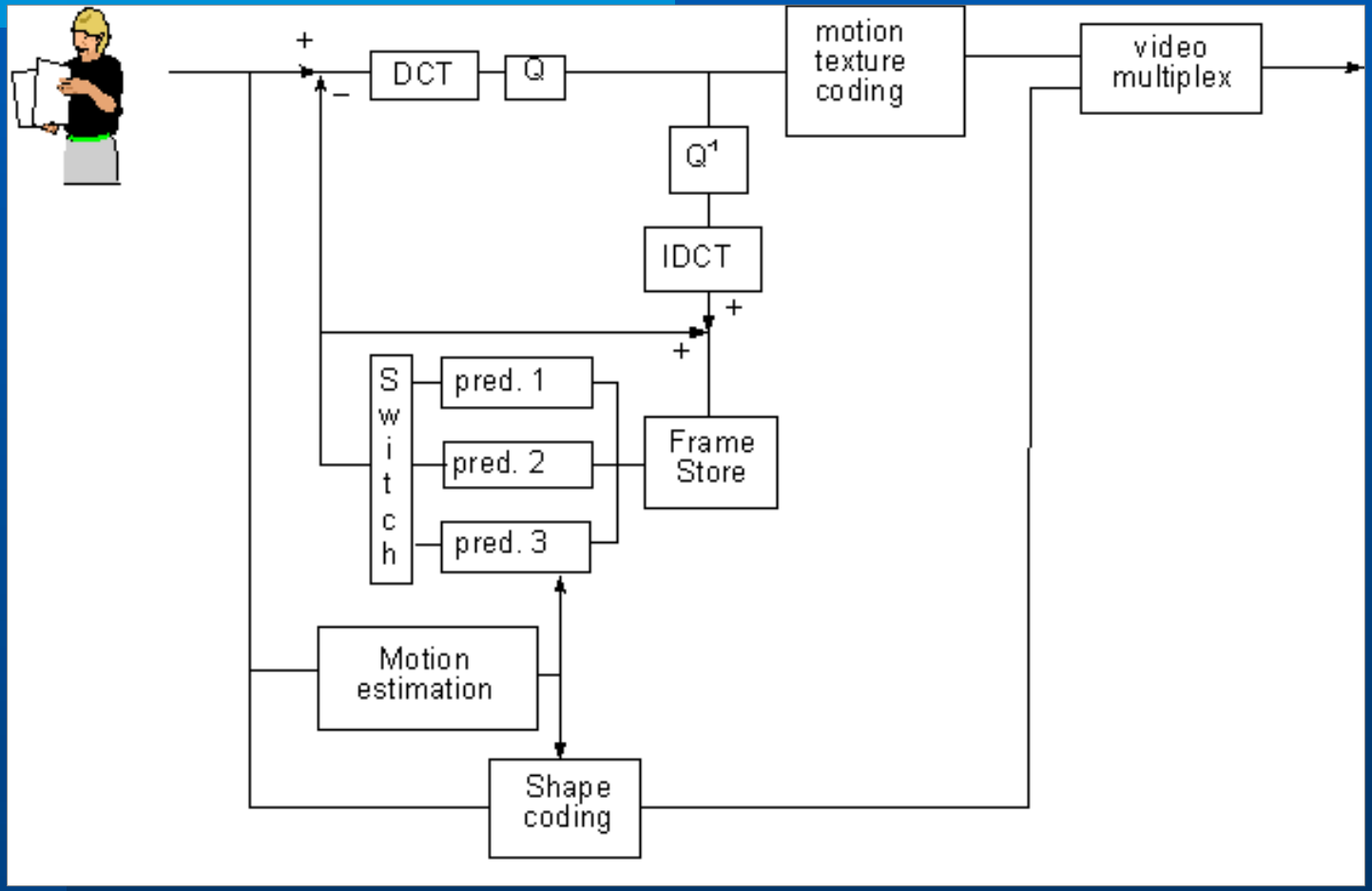
- image and video
- texture for meshes
- 2-D grid mesh
- temporal mesh variation
- random access to objects
- image and video manipulation
- content-based scaling
- spatial, temporal and quality scaling



Video object coding

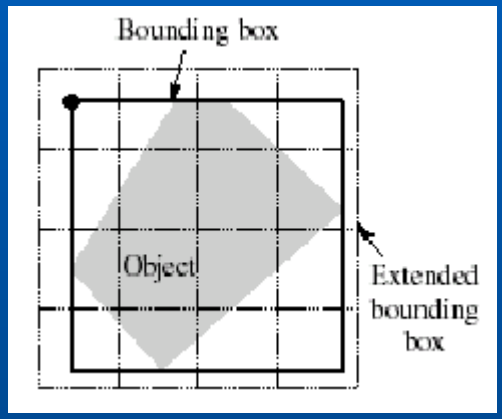
Video object, region of arbitrary shape (VO)
Snapshot : video object plane (VOP)
VOP definition : shape and texture
In addition motion for video objects
Rectangular objects : compatibility with
MPEG-1, MPEG-2, H.263

Video object encoder

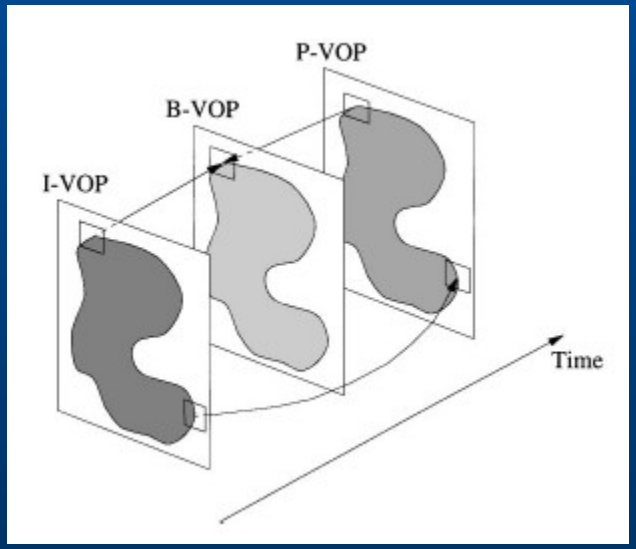


Motion and texture coding

Motion vector in macroblocks (MB)
- **1/4 pixel accuracy**
- **maybe overlapped MB, or splitted to 4 blocks**
- **predictive coding / motion compensation**
- **bounding box / MB**



Spatial compression using DCT
(similar to JPEG)
Shape-adaptive DCT coding
Uniform quantizer / visibility



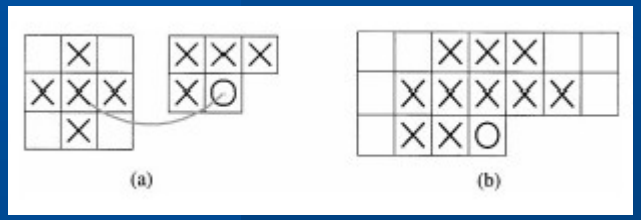
Shape coding

Transparency array α (binary or grayscale 8 bits)

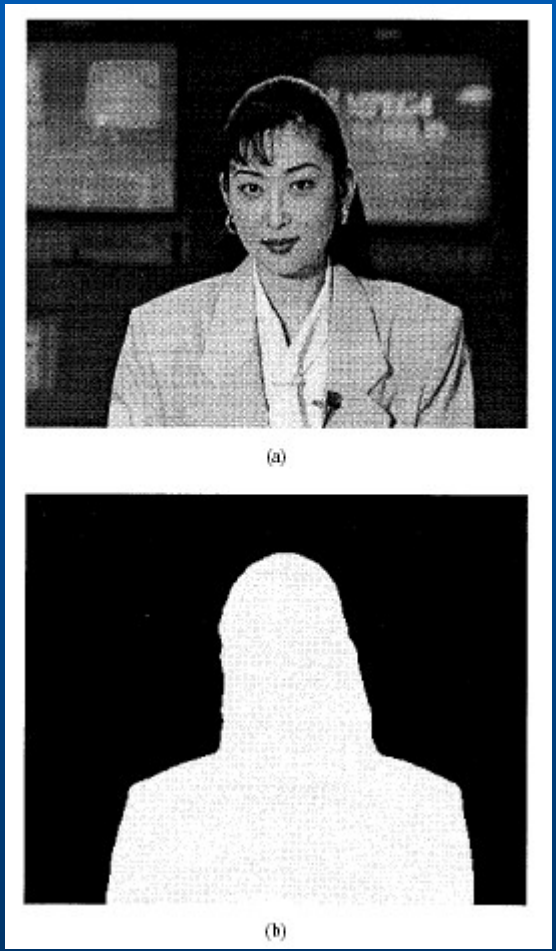
Spatial or temporal coding, lossy or lossless

Block 16 x 16 : transparent, opaque, boundary

Context-based arithmetic encoder



DCT-based coding for α array (grayscale 8 bits)

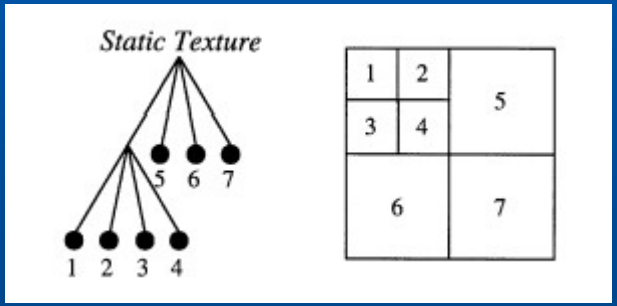


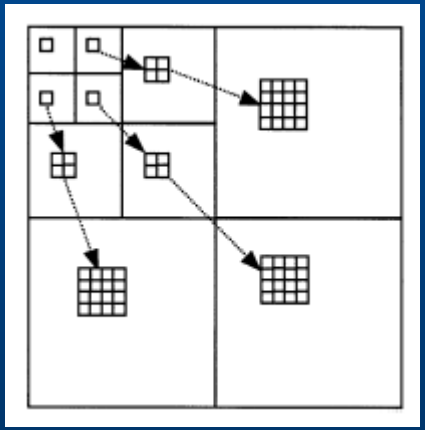
Static texture coding

2-D wavalet transform using Daubechies biorthogonal filters

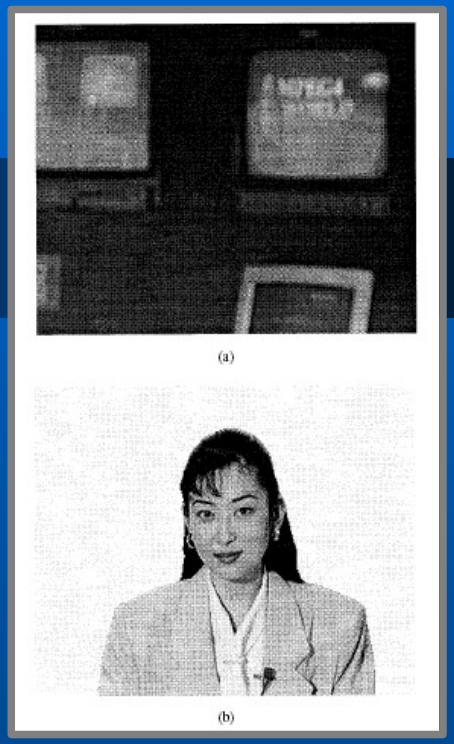
Spatial prediction for lowest frequency subband

Multi-scale zero-tree to better track locations of all coefficients





Sprite panorama coding

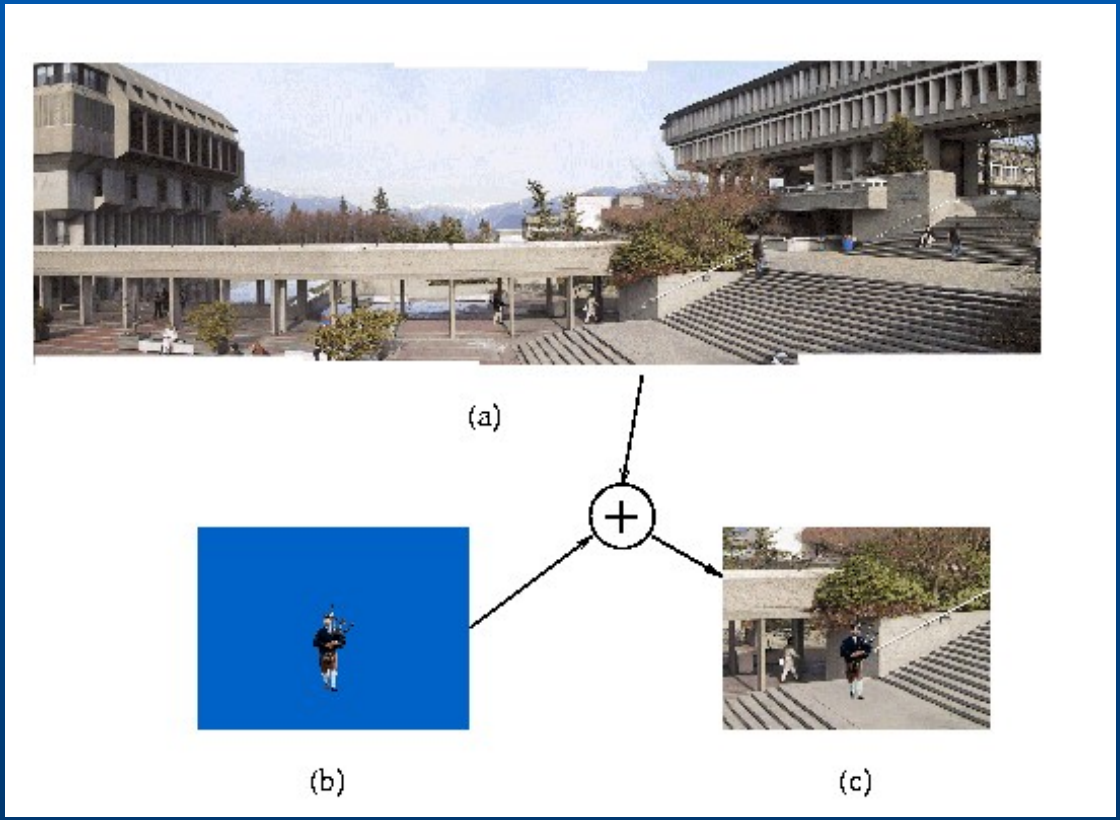


A still image that describes the static background over a sequence of video frames

The large sprite panoramic image can be encoded and sent to the decoder only once at the beginning of the video sequence

Sprite panorama coding

Scene composition



Camera motion compensation

Global motion parametric model

$$x_i' = \frac{a_0 + a_1 x_i + a_2 y_i}{a_6 x_i + a_7 y_i + 1}$$

$$y_i' = \frac{a_3 + a_4 x_i + a_5 y_i}{a_6 x_i + a_7 y_i + 1}$$

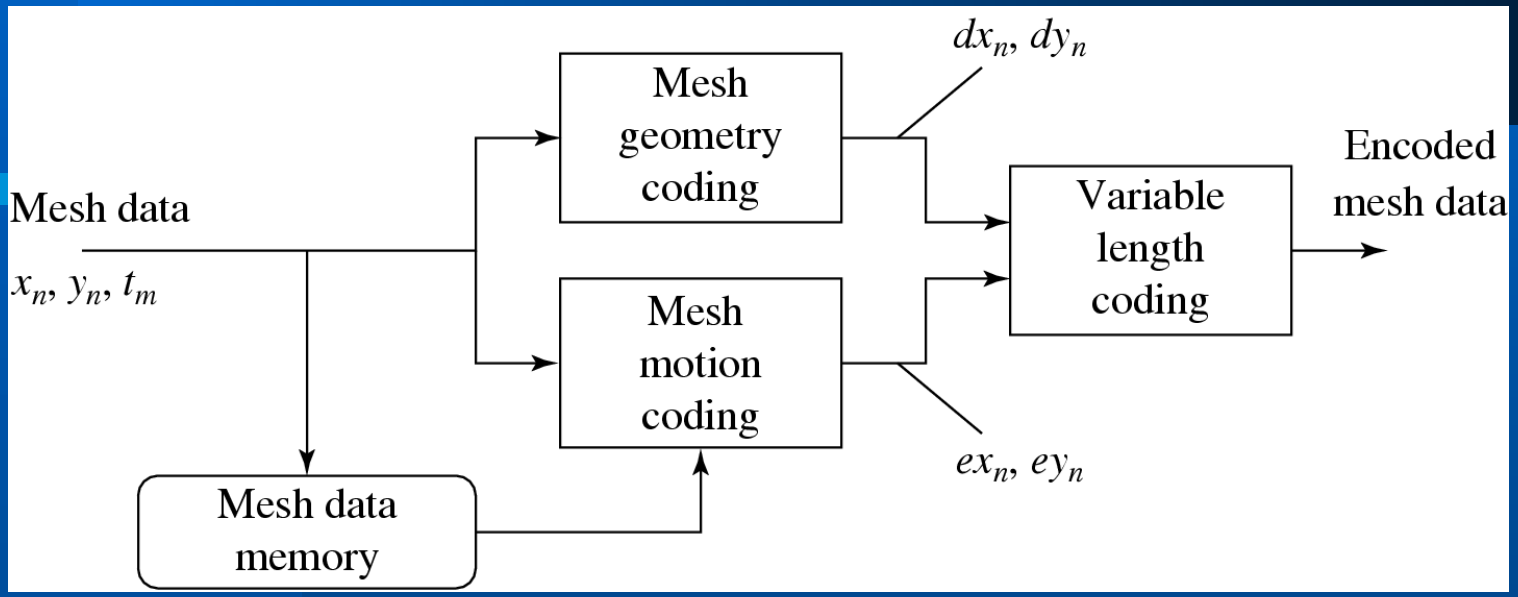
2-D mesh object coding

Computer graphics

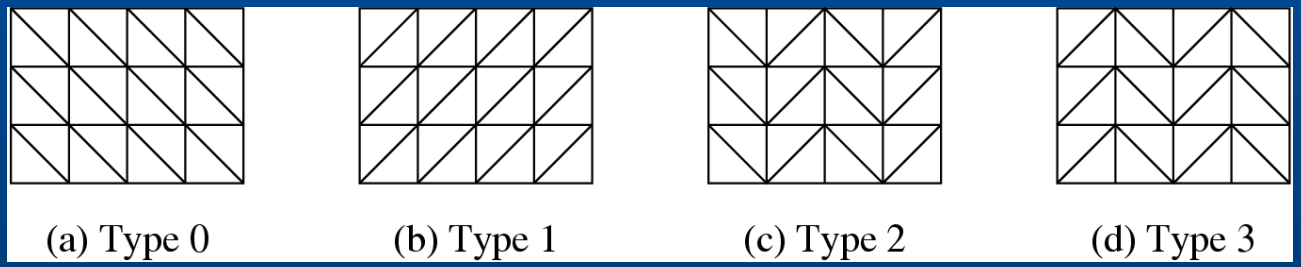
2-D mesh : a tessellation of a 2-D planar region using polygonal patches
Triangular mesh

Parametric motion in each triangle

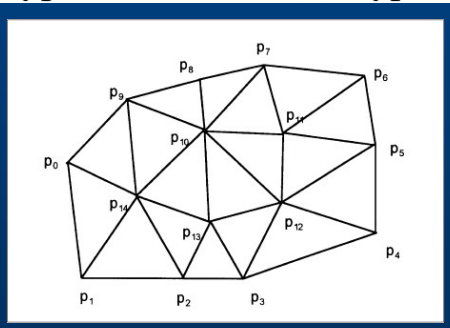
Mesh grid coding



Uniform mesh

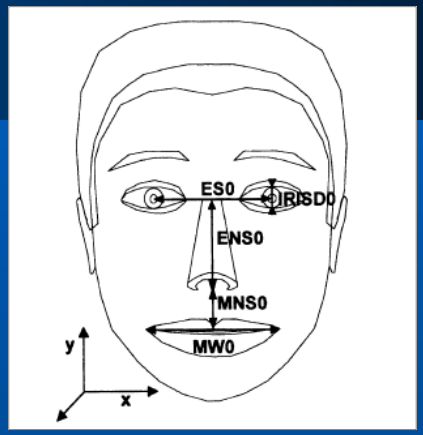


Data adaptive mesh : Delaunay
Differential node coding
Node motion vectors



Texture rendering

Animated face coding

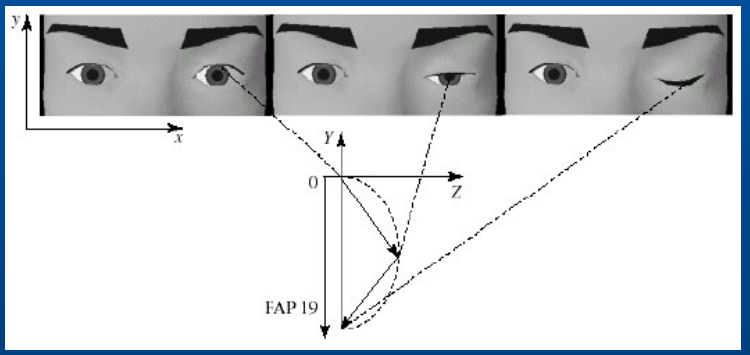


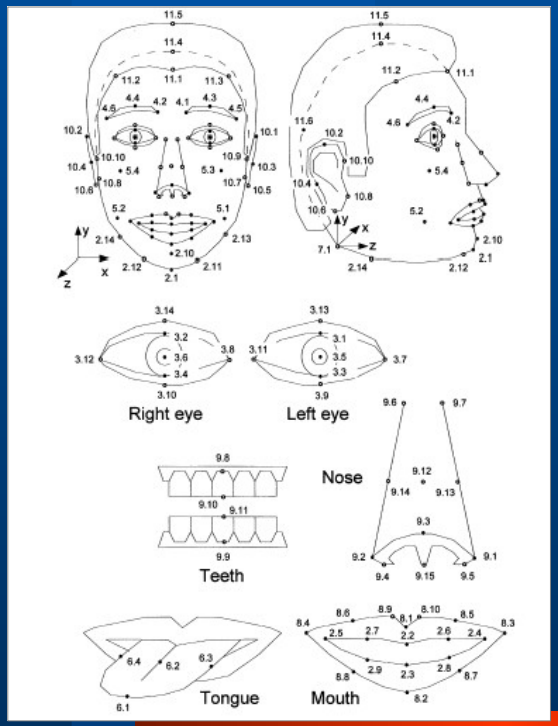
Face

- face definition parameters (FDP)
- face animation parameters (FAP)

Definition :

- 84 FDP
- 68 FAP





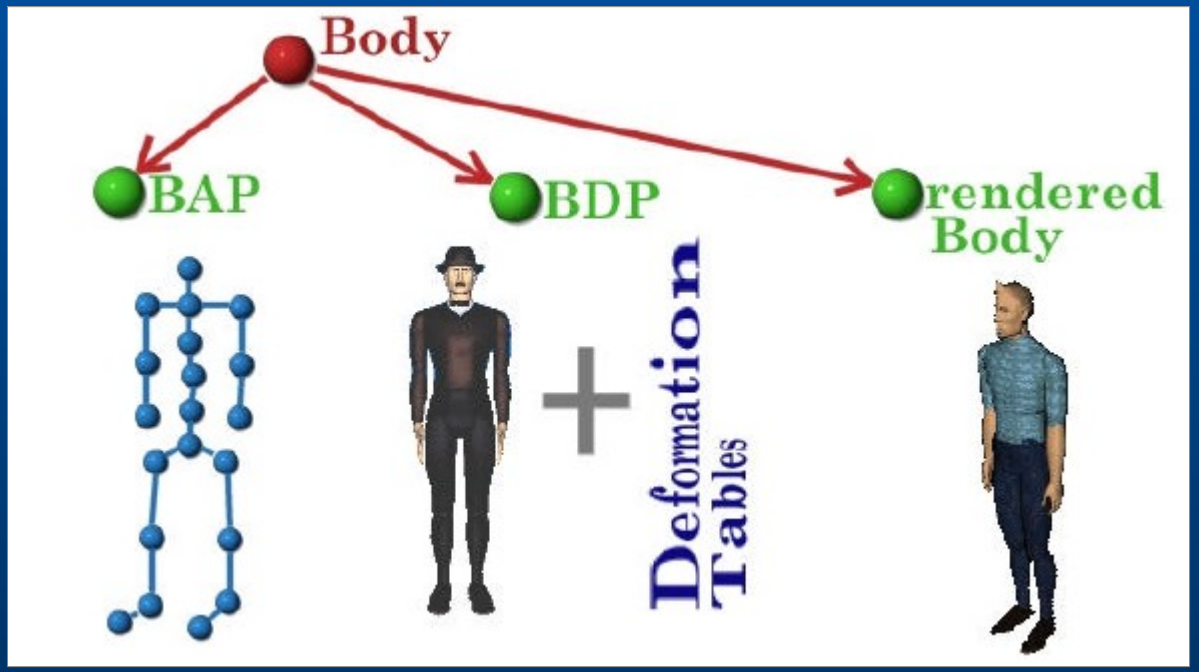
VRML

Speech synthesis combined to mouth animation

Animated body coding

- 3-D body definition parameters
- Body animation parameters

VRML / H-anim
Humanoid animation



MPEG-4 AVC

Advanced Video Coding : higher video coding efficiency

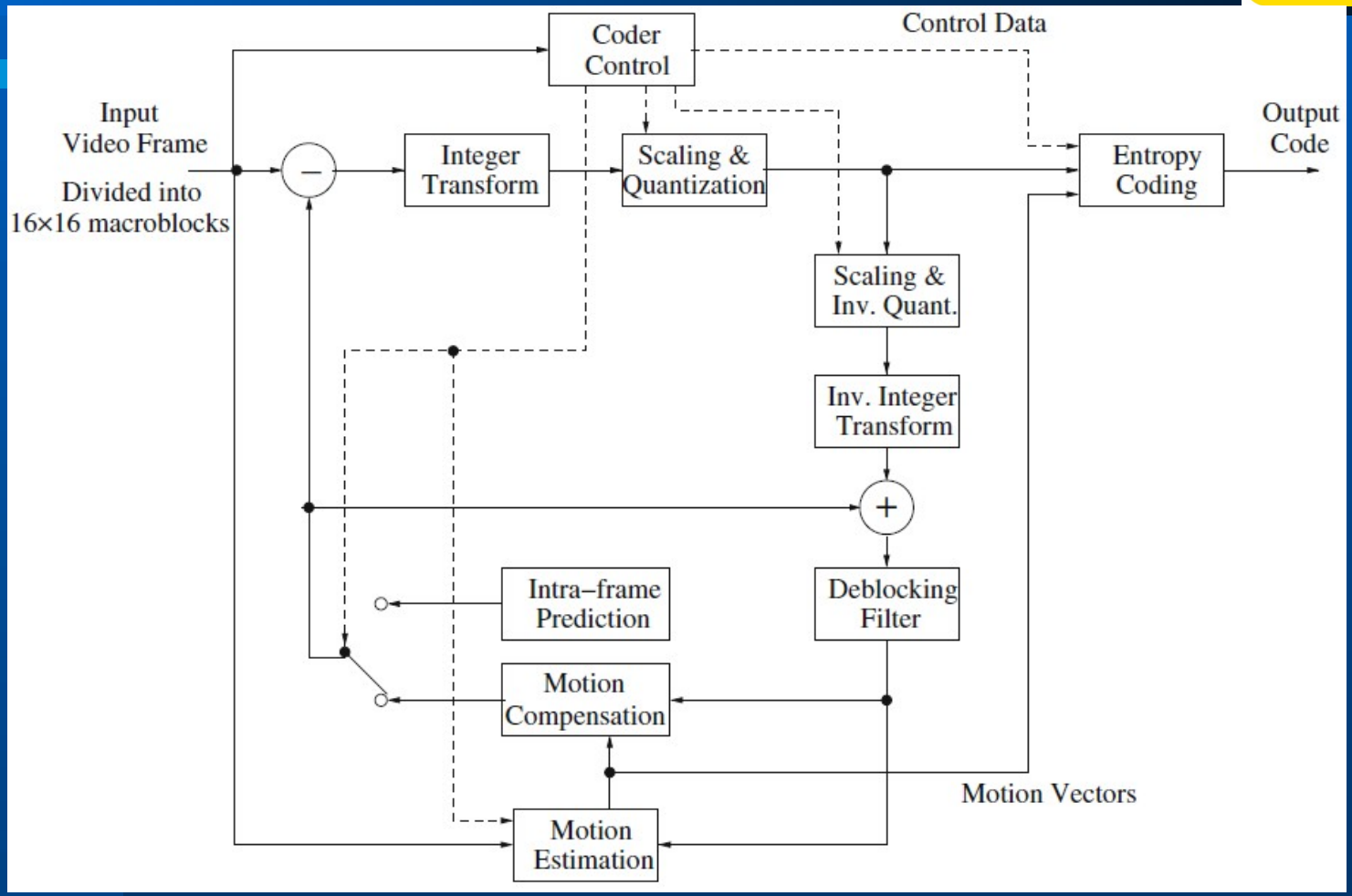
First release 2003 (ISO, ITU)

Scalable encoding (2007)

Multiple view encoding (2009)

Satelite TV 1.5 Mbits/s
Terrestrial and cable digital TV
Blu-ray Disc, HD DVD
Network communications (Ethernet, DSL, mobile, ...)
Flash, YouTube

MPEG-4 AVC hybrid coder



MPEG-4 AVC main features

- Variable block size motion compensation
- Multiple reference picture motion compensation
- Integer transform in 4 x 4 blocks (more efficient)
- Directional spatial prediction
- Exp-Golomb encoding
- Robust to transmission errors

HEVC / H.265

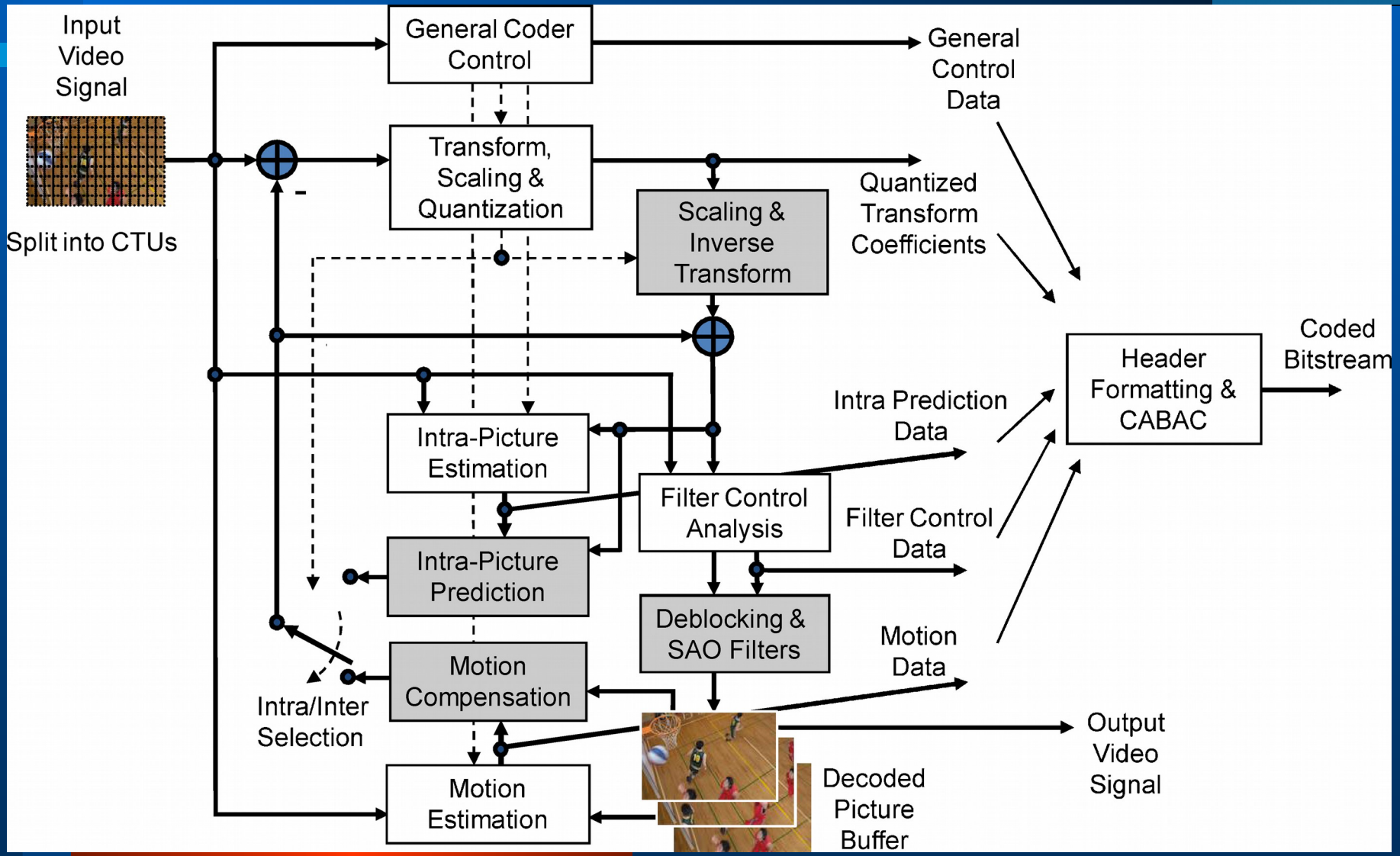
High-efficiency Video Coding
High-resolution

First release (2013)

Scalable coding (2014)

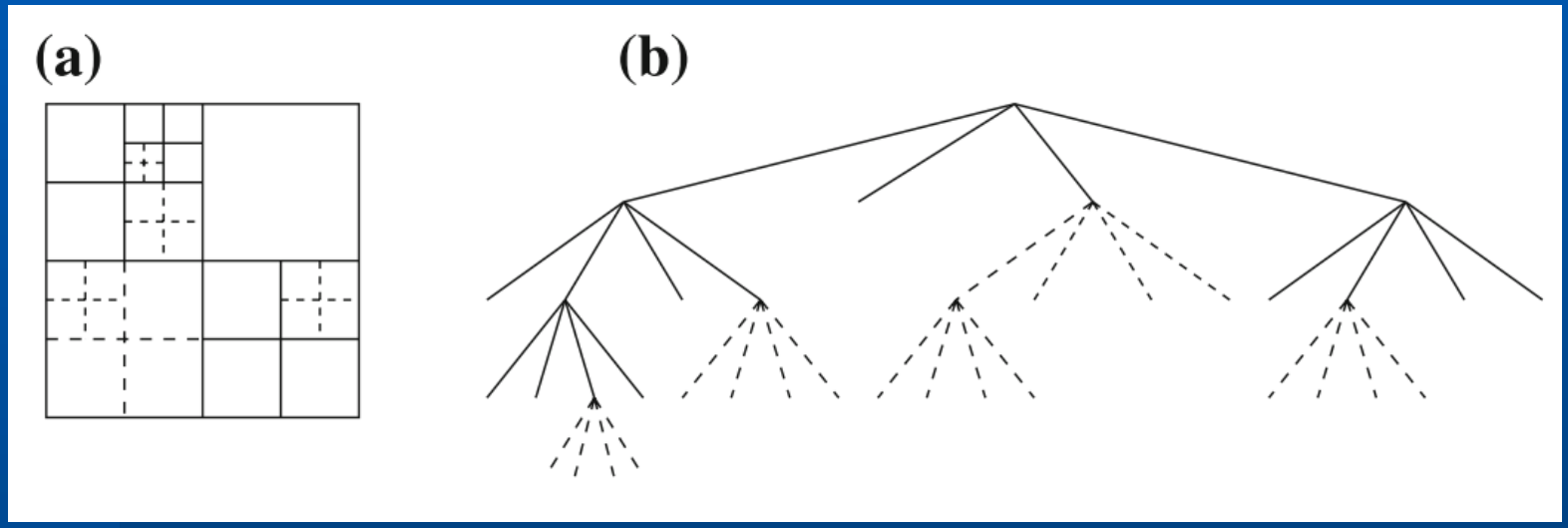
Multi-view coding (2014-15)

HEVC / H.265 hybrid coder



Motion compensation

H.265 HEVC
High Effeciency Video Coding



Hierarchical quad-tree decomposition
Initial block 64 x 64

1/4 pixel accuracy

Comparison

H.265 HEVC
High Effeciency Video Coding

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR INTERACTIVE APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	H.263 CHC	MPEG-4 ASP	MPEG-2 H.262 MP
HEVC MP	40.3%	67.9%	72.3%	80.1%
H.264/MPEG-4 AVC HP	–	46.8%	54.1%	67.0%
H.263 CHC	–	–	13.2%	37.4%
MPEG-4 ASP	–	–	–	27.8%

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR ENTERTAINMENT APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/ H.262 MP
HEVC MP	35.4%	63.7%	65.1%	70.8%
H.264/MPEG-4 AVC HP	–	44.5%	46.6%	55.4%
MPEG-4 ASP	–	–	3.9%	19.7%
H.263 HLP	–	–	–	16.2%