VIDEO COMPRESSION STANDARDS

MPEG 1 STANDARD

BASIC OBJECTIVES

- Targeted for multimedia CD-ROM applications at a bit rate of1.5 Mbits/sec
- Specifies a syntax for representation of encoded bitstream and a method of decoding
- MPEG-1 does not stipulate use of specific algorithms for bitstream generation and allows substantial flexibility
- syntax supports operations such as motion estimation; motion compensated prediction; Discrete Cosine transforms (DCT); quantization and variable length coding
- The standard supports a number of parameters that can be specified in the bit-stream itself and a variety of picture sizes, aspect ratios etc. are permissible.

BASIC OBJECTIVES

In addition, MPEG-1 standard supports the following application specified features:

- Frame–based random access of video: This is achieved by allowing independent access-points (I-frames) to the bit-stream.
- Fast-forward and fast reverse (FF/FR) searches: This refers to the scanning of the compressed bitstream to search for the desired portions of the video stream.
- Reverse playback of video
- Edit ability of the compressed bit stream
- Reasonable coding / decoding delay of about 1 sec to give the impression of interactivity.

CONSTRAINED PARAMETERS IN MPEG-1

The following set of constrained parameters are specified to aid hardware implementations

- Maximum number of pixels / line : 720
- Maximum number of lines/picture: 576
- Maximum number of picture/sec : 30
- Maximum number of macroblock/ pictures: 396
- Maximum number of macroblocks/sec: 9900
- Maximum bit-rate; 1.86 Mbit/sec
- Maximum decoder buffer size: 376,832 bits

PICTURE TYPES IN MPEG-1

- The MPEG-1 standard supports the following three picture types:
- I- picture
- P-picture
- B-picture

1. INTRAFRAME CODED PICTURES(I-PICTURES)

- These pictures are coded without reference to other pictures in the video sequence. I-pictures therefore do not use any motion estimation and motion compensation and the frames are treated just like still images
- The pixel intensity values are DCT encoded in a manner similar to JPEG and compression is achieved by a combination of quantization and run length coding of zero coefficient
- The first frame of every video sequence must necessarily be an I-picture, since it does not have any past reference
- In MPEG-1 standard, frames are encoded as I-pictures at regular intervals to enforce updating with the current content. This is done at the beginning of every group of pictures (GOP)

1. INTRAFRAME CODED PICTURES(I-PICTURES)

Advantages:

- I-pictures have better reconstruction quality
- I-pictures allow random access and fast forward / fast rewind (FF/FR) functionalities in the bitstream

Disadvantages:

- Since there is no temporal prediction, I pictures achieve very poor compression performance and require significant bits for encoding

2. INTERFRAME PREDICTED PICTURES(P-PICTURES)

- These pictures are coded with reference to the nearest (in temporal order) coded I-picture or P-picture, using motion compensation for prediction

Advantages:

- Since these pictures use the temporal redundancy for encoding, they achieve better compression performance as compared to I-pictures

Disadvantages:

- these pictures do not allow random access and FF/FR functionalities in the bitstream
- temporal prediction does not work where there are scene changes

3. BI-DIRECTIONALLY PREDICTED PICTURES(B-PICTURES)

- Have best compression performance
- use bi-directional motion estimation with reference to the nearest coded I-picture and / or P-pictures on either side of the B-picture in temporal order
- To achieve high compression ratio in the encoded bit stream, most of the frames in a video sequence are encoded as B-pictures
- For P-pictures, as well as B-pictures, the error in prediction through motion compensation is DCT-encoded and the compression is achieved through quantization and run-length encoding for zero coefficients. For both these pictures, entropy-coded motion vectors form part of the bitstream.

3. BI-DIRECTIONALLY PREDICTED PICTURES(B-PICTURES)

Advantages:

- Have the best compression performance

Disadvantages:

- Like the P-pictures, B-pictures also do not allow random access and FF/FR functionalities in the bit-stream

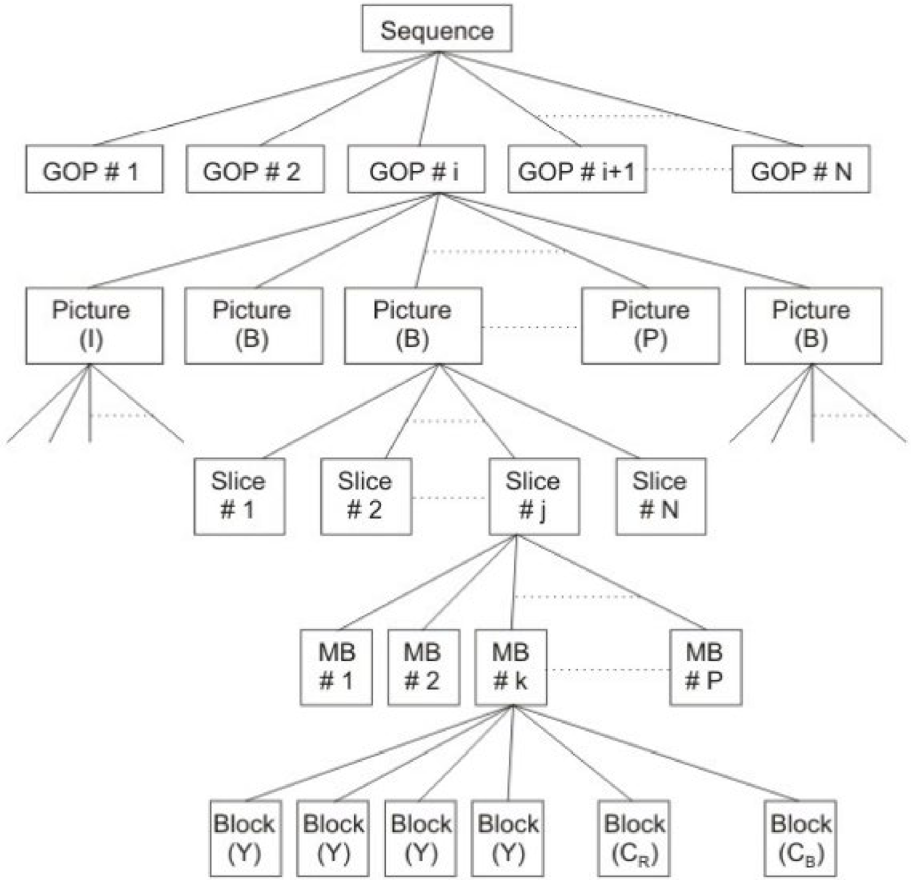


Fig. 23.1: Hierarchical data structure in MPEG-1.

HIERARCHICAL DATA STRUCTURE IN MPEG-1

- MPEG-1 data stream follows a 6-layer hierarchical structure
- Top-most level – we have video sequence
- Consists of several groups of pictures(GOP) in the next level. Each GOP begins with an I-picture and taking this as the reference, P-picture is encoded N frames later than the I-picture in temporal order. The (N-1) frames in between are encoded as B-pictures. The P-pictures in turn predict the next P-picture that occurs N-frames later

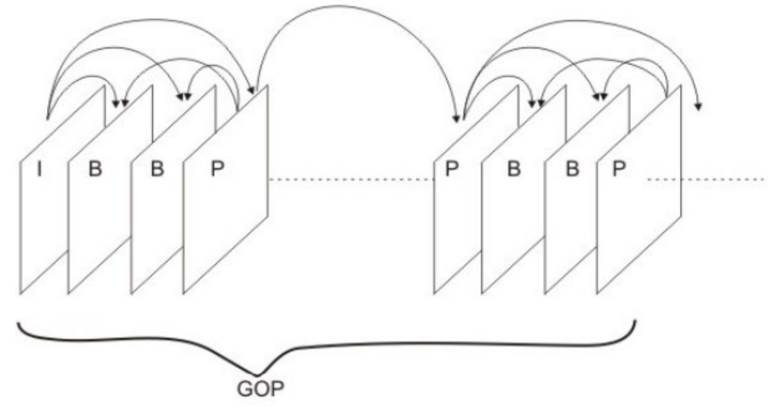


Fig 23.2 Compression of a GOP

HIERARCHICAL DATA STRUCTURE IN MPEG-1

- At the next level of hierarchy, pictures are composed of slices, which are essentially sequence of macroblocks in raster scan order and are designed for error recovery.
- Below the slice layer are the macroblocks, whose composition is illustrated

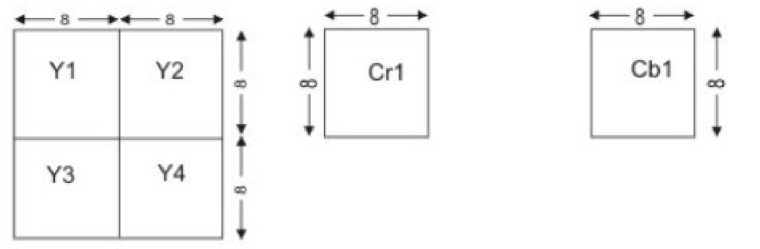


Fig 23.3 Composition of a macro block

HIERARCHICAL DATA STRUCTURE IN MPEG-1

- MPEG-1 standard works on colour images in Y-Cr-Cb format and every R-G-B sequence must be converted to Y-Cr-Cb format before encoding through MPEG-1.
- Each macroblock is composed of 16 x 16 pixels of luminance (Y) channel and one block of 8 x 8 pixels from each of the Cr and Cb channels.
- The groups of 16x16 luminance pixels are further subdivided into four blocks of 8x8 pixels. In MPEG-1 standard, motion compensation is applied on 16x16 pixels and the DCT is applied on 8 x 8 pixels.

MACROBLOCK TYPES SUPPORTED BY MPEG-1 STANDARD

The MPEG-1 standard supports the following macroblock types, depending upon the picture-types, i.e., I, P- and B.

- Macroblock types for I-Pictures
- Macroblock types for P-Pictures
- Macroblock types for B-Pictures

1. MACROBLOCK TYPES FOR I-PICTURES

There are two types of macroblocks (MB) in the I-picture

- “Intra” MBs are coded with the current quantization matrix x
- “Intra-A’ (Intra-Adaptive) MBs are coded with the current quantization matrix elements divided by a quantization scale parameter MQUANT, which is transmitted as a part of MB header. In Intra-A, each macroblock can have different quantization step-size, unlike “Intra” MBs. This mode is preferred for images, whose level of details varies significantly from one region of the image to the other. Due to the provision of “Intra-A” mode in MPEG, the encoding of I-pictures in MPEG is 30% more efficient than JPEG.

2. MACROBLOCK TYPES FOR P-PICTURES

The allowable MB types for P-pictures are as follows :

- Intra
- Intra-A
- Intra-D
- Inter-DA
- Inter-F
- Inter-FD
- Inter-FDA
- Skipped

2. MACROBLOCK TYPES FOR P-PICTURES

- The “Intra” and “Intra-A” modes are the same as those under I-pictures
- There are some macroblocks in a P-picture where no reliable temporal reference exists due to change in scene content, occlusion/ uncovering of objects, objects moving out of frame, new objects moving into the frame etc. In such cases, those macroblocks may be coded as “Intra” or “Intra-A”, although it may still be a P-picture.
- The “skipped” modes is used for stationary macroblocks which neither undergo any noticeable displacements nor have any noticeable changes in pixel intensity values.

2. MACROBLOCK TYPES FOR P-PICTURES

MBs classified as "Inter" are interframe coded and the temporal prediction may use motion compensation and / or adaptive quantization. The letters that follows "Inter" are defined as

- D : DCT of the prediction error will be coded.
- F : Forward motion compensation is on.
- A : Adaptive quantization is on.

3. MACROBLOCK TYPES FOR B-PICTURES

B-pictures allow al the MB types supported by P-pictures. In addition to forward motion compensation ("F"-type), B-pictures allow backward motion compensation (B-type) and bi-directional (interpolative) motion compensation (I-type). Hence, following MB-types are supported

- Intra
- Intra-A
- Inter-F
- Inter-FD
- Inter-FDA
- Inter-B

3. MACROBLOCK TYPES FOR B-PICTURES

- Inter-BD
- Inter-BDA
- Inter-I
- Inter-ID
- Inter-IDA
- Skipped

MPEG-2 STANDARD

BASIC OBJECTIVES

- Compression, coding and transmission of high quality multi-channel, multimedia signals for terrestrial broadcast, digital, cable TV distribution, broadband networks etc.
- Defining profiles and levels as the subset of syntax to suit wide range of applications.
- Scalable bit stream
- Error-correction capabilities
- Backward compatibility with MPEG-1, so that every MPEG-2 compatible decoder can decode a valid MPEG-1 bit stream

PROFILES AND LEVELS OF MPEG-2

- Since MPEG-2 standard encompasses diverse applications requirements, a single syntax was defined by integrating many video coding algorithms
- implementation of the full syntax was not very practical and some subsets of the syntax were defined with some profiles and levels
- Accordingly a decoder’s capabilities to decode a particular bit stream get defined.

PROFILES AND LEVELS OF MPEG-2

MPEG-2 supports following five profiles in decreasing order of hierarchy.

- High
- Spatial scalable
- SNR scalable
- Main
- Simple

PROFILES AND LEVELS OF MPEG-2

Each profile adds a new set of algorithms and acts as a superset of the algorithms supported in the profile below. A level specifies the range of parameters that are supported by the implementation, i.e., image size, frame rates and bit-rates. MPEG-2 supports following four levels

- High
- High-1440
- Main
- Low

Table 24.1 Algorithms and functionalities under each profile

Profile	Algorithms and functionalities
High	All functionalities provided by spatial scalable profile plus • 3-layers of SNR and spatial scalable coding • 4 :2 :2 YUV representation
Spatial scalable	All functionalities provided by SNR scalable profile plus • 2- layers of spatial scalable coding • 4 : 1 :1 YUV representation
SNR scalable	All functionalities provided by Main profile plus • 2-layers of SNR scalable coding • 4 :2:0 YUV representation
Main	All functionalities provided by simple profile plus : • Coding interlaced video • Random access • B-picture prediction modes • 4 :2 :0 YUV representation
Simple	Does not support B-picture prediction

Table 24.2 Upper bound of parameters at each level

Level	Parameters constraints
High	1920 pixels/line, 1152 lines/frame, 60 frame/sec
High-1440	1440 pixels/line, 1152 lines/frame, 60 frames/sec
Main	720 pixels/line, 576 lines/frame, 30 frames/sec
Low	352 pixels/line, 288 lines/frame, 30 frames/sec.

INTERLACED VIDEO: FRAME PICTURE AND FIELD PICTURE

- Interlaced scanning in which a frame is partitioned into a set of odd-numbered scan lines (referred to as odd field) and a set of even numbered scan lines (referred to as even field)

MPEG-2 supports two new picture formats – frame pictures, and field pictures.

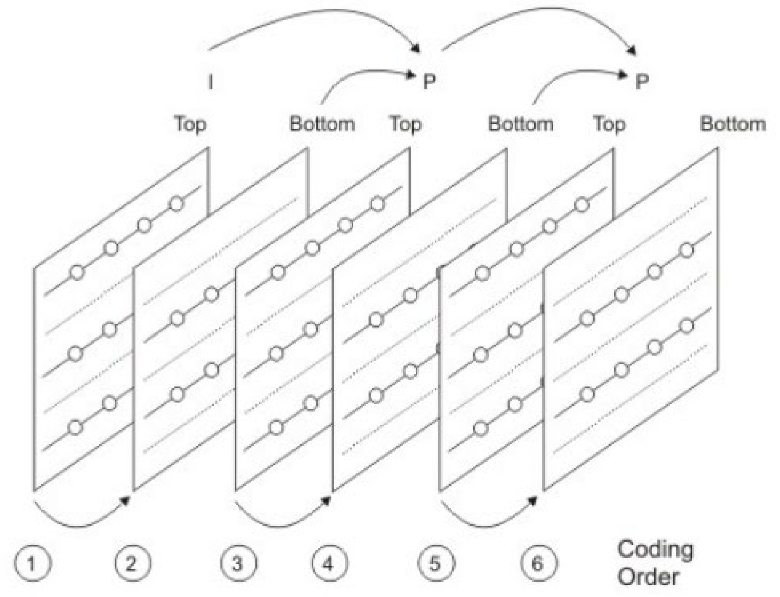
- In field picture, every field is coded separately. Every field is separated into non-overlapping macroblock and DCT is applied on a field basis
- In frame pictures, the two fields are coded together as a frame, similar to the conventional coding of progressive video sequence.

INTERLACED VIDEO: FRAME PICTURE AND FIELD PICTURE

- Frame pictures are preferred for relatively still images
- Field pictures give better results in presence of significant motion
- Each frame picture or a field picture may be I-type, P-type or B-type

FIELD AND FRAME PREDICTION

- It is possible to predict a field picture from previously decoded field pictures.
- Each odd field (top field) is coded using motion compensated inter-field prediction based on the previously coded even field (bottom field).
- Each even field may either be predicted through motion compensation on a previously coded even field or from previously coded odd field belonging to the same picture
- Within a field picture, all predictions are field predictions



24.1 Field prediction in MPEG-2

FIELD AND FRAME PREDICTION

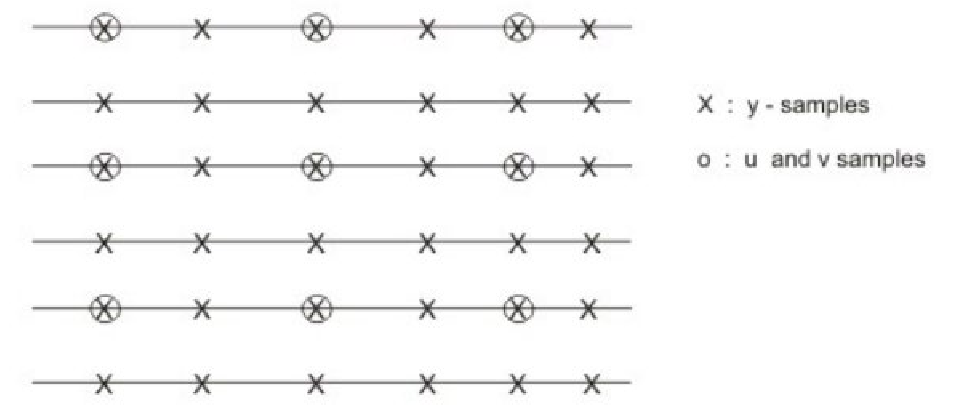
- Frame pictures can either have a frame prediction or field prediction and the prediction mode may be selected on a macroblock to macroblock basis
- MPEG-2 also supports a dual prime prediction in which two independent predictions are made - one for the 8-lines which correspond to the odd (top)field, another for the 8 even (bottom)field lines

CHROMINANCE FORMAT FOR MPEG-2

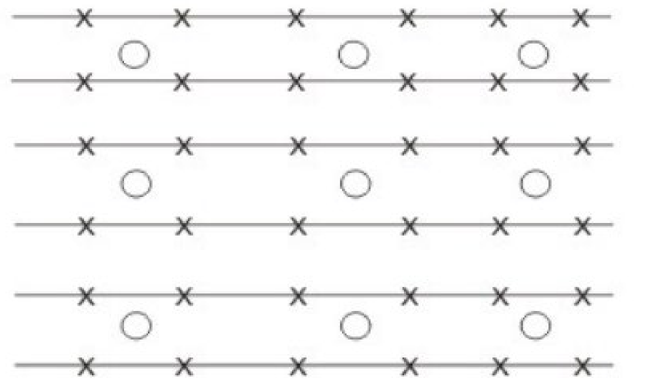
- Chrominance format describes the ratio between the horizontal spatial sampling frequencies of the luminance and chrominance components.
- The chrominance format is expressed as three numbers - the first represents the luminance (Y) sampling frequency, the second and the third represent chrominance U and V sampling frequencies respectively. By convention, the first number is always taken as 4
- In MPEG-1, both U and V are sampled at half the sampling rate of Y in both horizontal and vertical directions. It should have been called as 4:1:1, but is referred to as 4:2:0 since the relative positions of luminance and chrominance in these two formats differ.

CHROMINANCE FORMAT FOR MPEG-2

- In 4:2:0, the chrominance samples are located in between the grids for luminance samples, whereas in 4:1:1 format, the U and V samples have same spatial locations as that of Y.



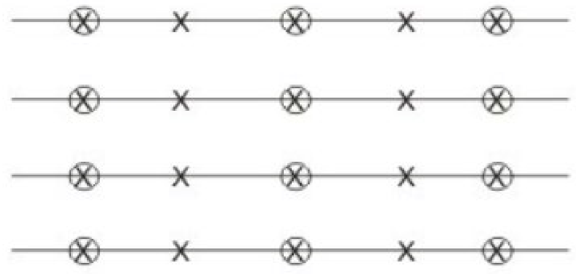
(a)



(b)

CHROMINANCE FORMAT FOR MPEG-2

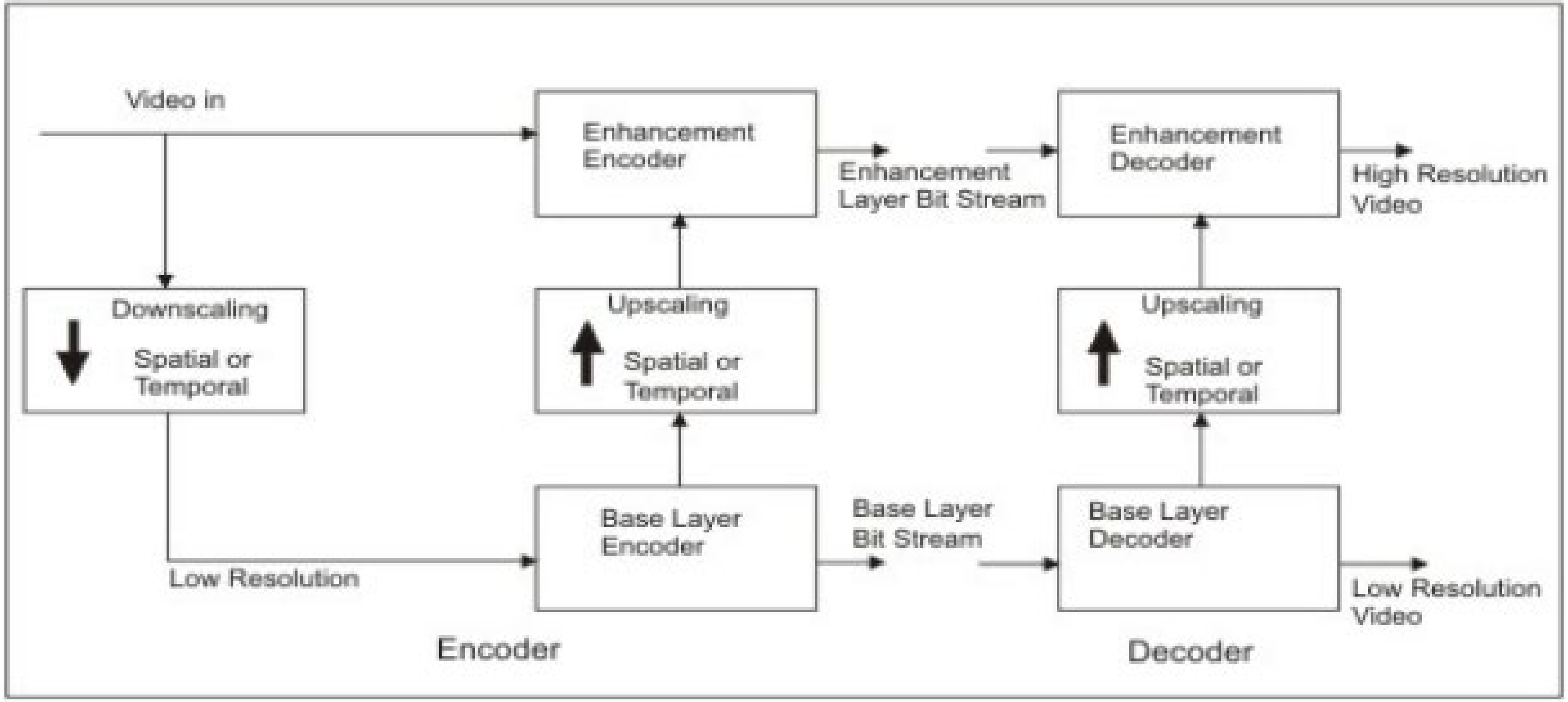
- MPEG-2 not only supports the 4:2:0 format, but also the 4:2:2 format in which case, the chrominance sub-sampling is done in only one direction (horizontal), but in the vertical, the same sampling frequency as that of luminance is maintained.



(c)

SCALABILITY SUPPORT OF MPEG-2

- MPEG-2 standard supports scalability to provide interoperability between different services and to support receivers with different display capabilities. Receivers not having the capability to reconstruct full resolution video can decode only a subset of the layered bitstream to reconstruct a reduced resolution video.
- The bit-stream is organized into layers having two or three hierarchies. The bottom of the hierarchy contains base layer, which every receivers and every application must make use of. Above the base layer, enhancement layers exist, which will be used by high-end applications.
- The scalability support is of particular interest for SDTV (Standard Definition Television) and HDTV applications. Instead of providing separate bitstreams for SDTV and HDTV, one common scalable bitstream is provided



24.3 Multi-scale video encoding and decoding

SCALABILITY SUPPORT OF MPEG-2

- a downscaled version is encoded into a base-layer bitstream with reduced bit-rate
- The reconstructed base-layer video is up-scaled spatially or temporally to predict the original input video
- The prediction error is encoded into an enhancement layer bitstream.
- The scalable coding can be used to encode video with a suitable bit-rate allocated to each layer in order to meet the specific bandwidth requirement of the transmission channels or the storage media

SCALABLE CODING SCHEMES

MPEG-2 has standardized three scalable coding schemes

- signal-to-noise ratio (SNR) scalability
- spatial scalability
- temporal scalability

1. SNR SCALABILITY

- intended for use in video applications involving telecommunications, video services with multiple qualities
- The SNR scalable algorithms use a frequency (DCT-domain) scalability technique in which both base-layer and the enhancement layers are encoded at the same spatial scale but using different quantization for DCT coefficients
- At the base-layer, the DCT coefficients are coarsely quantized to achieve moderate image quality at reduced bit rate. The enhancement layer encodes the difference between the nonquanitized DCT coefficients and the coarsely quantized coefficients from the base-layer with fine quantization step-sizes
- The SNR scalability is obtained as a straight forward extension to the main profile and obtains good coding efficiency.

2. SPATIAL SCALABILITY

- designed to support displays having different spatial resolution using one common layered bit-stream
- This scheme best suits SDTV/HDTV applications
- The base-layer encodes a spatially down-sampled video sequence and the enhancement layer encodes the extra information that would be necessary to support higher spatial resolution displays.
- The spatial scalability algorithm is based on the classical pyramidal approach for progressive image coding.

3. TEMPORAL SCALABILITY

- intended for use in systems where a migration into the higher temporal resolution from a lower one may be necessary
- Temporal scalability is achieved by skipping certain fields/ frames at the baselayer. The skipped frames are then encoded at the enhancement layer.
- The enhancement layer forms its predictions from either the decoded picture at the base layer or from previous temporal prediction at the enhancement layer
- Temporal scalability can be used to accommodate both interlaced and progressive video. The base layer can be interlaced and the enhancement layer can be a progressive HDTV video sequence

DATA PARTITIONING IN MPEG-2 BIT-STREAM

- MPEG-2 bit-stream has a provision for data partitioning according to the priorities to support error concealment in presence of transmission or channel errors.
- Similar to the SNR scalability, the algorithm is based upon the separation of DCT coefficients in two layers with different error likelihood.
- This scheme is implemented with a very low complexity as compared to the scalable coding schemes.

MPEG-4 STANDARD

BASIC OBJECTIVES

- Support for content-based manipulation and bit stream editing
- Ability to combine synthetic scenes or objects with natural scenes and objects
- Provisions for efficient random access of video frames or objects
- Better visual quality at comparable bit rates, as compared to its earlier standards.
- Ability to encode multiple views, for example stereoscopic video
- Provisions for error robustness to allow access to a variety of wireless and wired networks, storage media
- Scalability with fine granularity in content, quality and complexity

CONTENT BASED INTERACTIVITY

- In MPEG-4, audio and video data are content based, which allow independent access and manipulation of audio-visual objects in the compressed domain. Transformation of existing objects (re-positioning, scaling, and rotations), addition of new objects, removal of existing objects etc. are all within the scope of manipulation.
- The object manipulations are possible through simple operations performed on the bit stream

TOOLBOX APPROACH OF MPEG-4

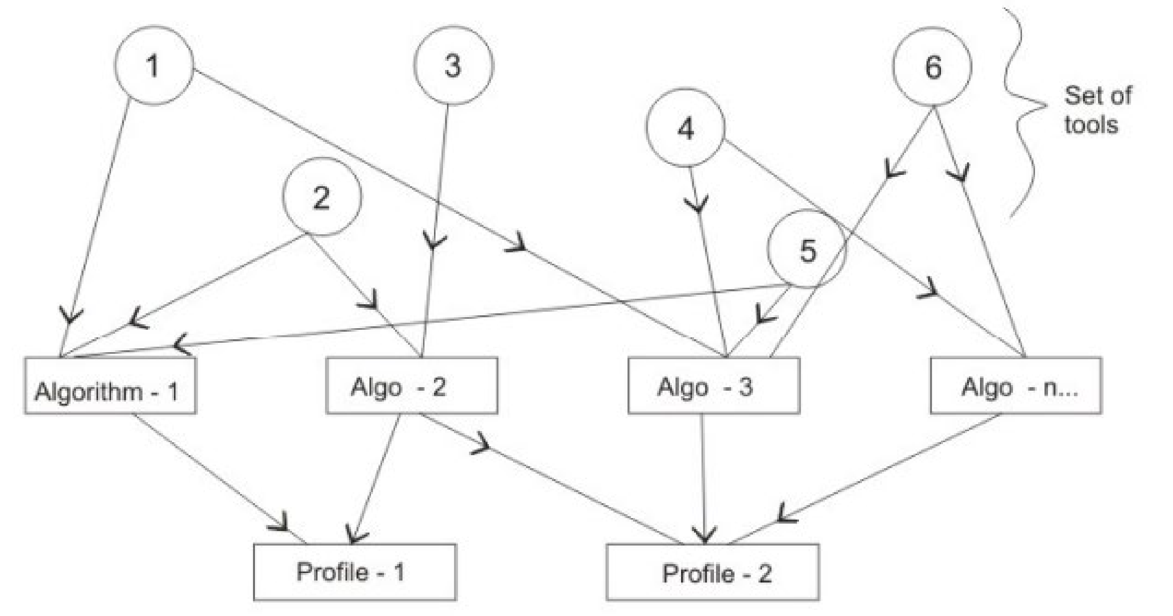


Figure 25.2 MPEG-4 "Toolbox" approch

TOOLBOX APPROACH OF MPEG-4

- In the two earlier standards, complete algorithms for audio, video and system aspects were standardized.
- MPEG-4 in contrast, follows a toolbox approach in which tools are standardized.
- Video tools include a complete algorithm, or individual modules such as shape coding, motion compensation, texture coding etc.
- These independent coding tools can be bound together using the MPEG-4 Systems Description Language (MSDL).
- offers flexibility to address variety of requirements

VIDEO OBJECT REPRESENTATION AND ENCODING LAYERS

- To achieve content-based interactivity, MPEG-4 has standardized on the video object representation.
- A sequence is composed of one or more audio visual objects (AVO). AVOs can be either an audio object resulting out of speech, music, sound effects, etc or a video object (VO) representing a specific content
- A video object may be present over a large collection of frames. A snapshot of a video object in one frame is defined as the video object plane (VOP) and is the most elementary form of content representation

VIDEO OBJECT REPRESENTATION AND ENCODING LAYERS

- For content representation using VOPs, an input video sequence is segmented into a number of arbitrarily shaped regions (VOPs). Each of the regions may possibly cover particular image or video content of interest. The shape and the location of the region can vary from frame to frame
- The shape, motion and texture information of the VOPs belonging to the same VO is encoded and transmitted into a Video Object Layer (VOL). Since typically there are several video objects, the bit stream should also include information on how to combine the different VOLS to reconstruct the video.

VIDEO OBJECT REPRESENTATION AND ENCODING LAYERS

Fig. 25.3 Snapshot of a Video Sequence

Fig 25.4 Binary alpha-plane

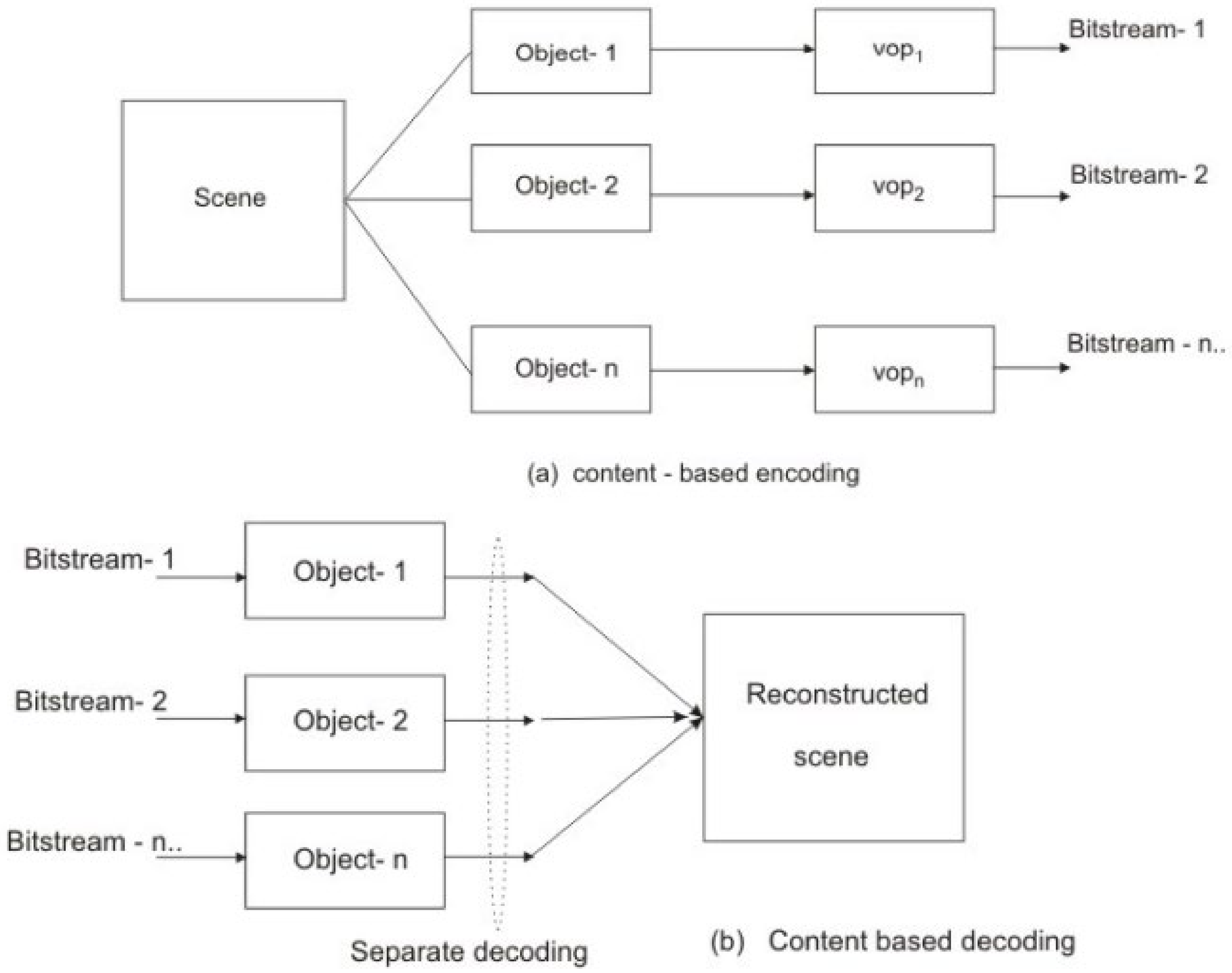


Figure 25.5 Content based encoding and decoding in MPEG - 4

VIDEO OBJECT REPRESENTATION AND ENCODING LAYERS

- The scene is first segmented into a number of VOPs, each of which specifies particular image sequence content and is coded into a separate VOL. It is possible to reconstruct the original video if all the VOLs are considered.
- However, contents can be decoded by considering only a subset of all VOLs and this allows content based interactivity.

Each VOL encoding has three components

- Shape (contour) coding
- Motion estimation and compensation
- Texture coding

VOP ENCODING AND SHAPE ADAPTIVE MACROBLOCK GRID

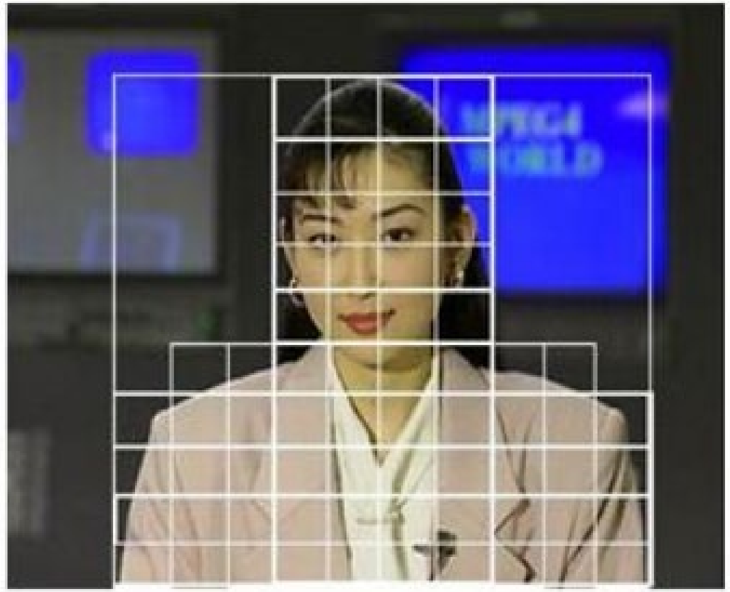


Fig 25.6 VOP Window and VOP macroblocks

VOP ENCODING AND SHAPE ADAPTIVE MACROBLOCK GRID

- The VOP image window is a rectangular window having size in multiples of 16 pixels in each image direction that surrounds the foreground VOP. This window is adjusted to collocate with the top-most and leftmost border of the VOP.
- The position of the VOP image window is defined with respect to a reference window of constant size by specifying a shift parameter.

VOP ENCODING AND SHAPE ADAPTIVE MACROBLOCK GRID

The VOP image window is composed of macro block of size 16 x 16 pixels, which are of three types:

- Macroblocks which do not belong to the VOP at all. These are inactive macroblock with respect to the VOP and are not encoded in the VOL.
- Macroblocks which partly belong to the VOP. These are the boundary macroblocks for the VOP and require some special consideration during its encoding.
- Macroblocks which fully belong to the VOP. These are the standard macroblock for the VOP.

ENCODING OF VOP'S

The VOLs compose the bit-streams for the VOPs and their encoding have three major components

- Shape
- Motion
- Texture.

1. SHAPE CODING IN MPEG-4

- Since video objects in MPEG-4 are of arbitrary shape, encoding of shapes form an essential part of encoding. Whether a pixel belongs to the VOP or not is specified by a binary map known as alpha plane which has an entry of "1" if the pixel belongs to VOP and is "0" otherwise

Shape coding techniques may be broadly classified as

- contour based
- bit-map based

1. SHAPE CODING IN MPEG-4

- The contour based techniques extract and encodes a description of the closed contour enclosing the shape
- It is the vertex-based coding that approximates the shape using a polygonal approximation.
- First, the longest axis of the shape is found and its two end points are used as the initial polygon. For each polygon line, it is checked if the approximation lies within the tolerance.
- If not, a new vertex is inserted at the point of largest prediction error. Each new polygon side is checked again for approximation and the process is iteratively repeated.

1. SHAPE CODING IN MPEG-4

- The bit-map based techniques are applied directly to the binary alpha-plane, within the conventional block-based framework

Bit map based shape coding techniques may be broadly categorized as:

- Modified Read (MR) approach, used in fax
- Context based arithmetic encoding (CAE), which has been adopted in JBIG standard.

2. MOTION ESTIMATION IN VOP'S

- The shape-adaptive macroblock grid is used for motion estimation in VOPs.
- The standard macroblock within the grid are motion compensated, following the approaches adopted in the earlier two MPEG standards.
- In contour macroblocks, an image padding method is employed in the reference for these macroblocks, which can be seen as an extrapolation of pixel values outside the VOP based on the values inside the VOPs.
- After padding the reference VOP, a polygon matching technique is employed for motion estimation and compensation. "Polygon" refers to the part of the contour macroblock which belongs to the active area inside the VOP frame to be coded and excludes the pixels outside this area.

2. MOTION ESTIMATION IN VOP'S

Based on the motion estimation and motion compensation philosophy, three types of VOPs can be defined :

- I-VOP : These are the intra-coded VOPs, similar to the intra coded frames (I-picture) where no motion estimation is employed and only texture coding is done.
- P-VOP : These VOPs use forward prediction for motion compensation, very similar to the P-picture.
- B-VOP: These VOPs are bi-directionally predicted, very similar to the Bpicture.

3. TEXTURE CODING

- Texture coding is to be performed on the I-VOP or the residual errors after the motion compensation in the P-VOPs and B-VOPs.
- For texture coding too, the shape adaptive macroblock grid is used. For each macroblock, a maximum of four 8 x 8 luminance blocks and two 8 x 8 chrominance block are employed.
- Special adaptation is required for the contour blocks, where image padding technique is used to fill the macro block content outside a VOP before applying the DCT in intra-VOP.

4. MULTIPLEXING OF SHAPE, MOTION AND TEXTURE INFORMATION

The video object layer (VOL) is formed by multiplexing the encoded VOP information in the following order:

- Shape encoding
- Motion vector encoding
- Texture coding

SPATIAL AND TEMPORAL SCALABILITY OF MPEG-4

- each VOP can be encoded to multiple number of VOLs of which only one forms the base layer and the remaining ones compose the enhancement layers. The layered bitstream has a major advantage in terms of prioritized transmission and error resiliency.

Two types of scalabilities are supported in VOP encoding process

- Spatial Scalability
- Temporal Scalability

1. SPATIAL SCALABILITY

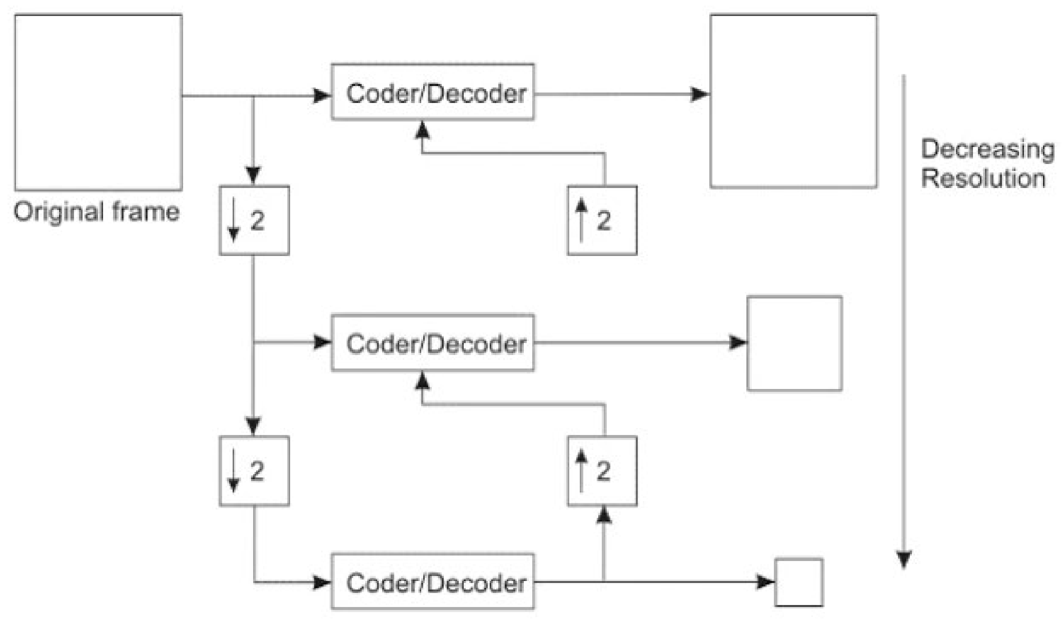


FIGURE 25.7 Spatial scalability of VOPs in MPEG - 4

1. SPATIAL SCALABILITY

- This is very similar to the spatial scalability support in MPEG-2. Here, multi resolution representations of the VOPs are formed by spatially downsampling the input video signal into a number of levels.
- The lowest resolution level supports the base-layer bit stream and the subsequent upper resolutions are predicted by upsampling from the lower resolution of the VOP.

2. TEMPORAL SCALABILITY

- Like the spatial scalability temporal scalability too generates a layered bit stream in which the base-layer is formed by temporally subsampling the video objects and the enhancement layers are obtained by temporal prediction from the lower layers
- Using the MPEG-4 VOP temporal scalability approach, it is possible to have different frame rates for different video objects.

SPRITE CODING IN MPEG-4

- The object based coding in MPEG-4 essentially requires video segmentation algorithm to extract the foreground from the background. This idea is extended to sprite coding, in which the background is reconstructed and transmitted separately from the foreground, using a very sophisticated motion analysis and prediction strategies.
- the large, static panorama picture is first transmitted to the receiver. For each frame, camera parameters are transmitted separately, which facilitates extraction of frame backgrounds from the panorama. The foreground is encoded separately and the receiver composes the scene from the separately transmitted foreground and the background.
- Since sprite-coding requires only one time transmission of the background, substantial coding gain is usually achieved as compared to the usual block based encoding of the entire scene.

FACIAL FEATURE ANIMATION CAPABILITIES OF MPEG-4

- The sprite coding concepts can be extended to the model based video coding for head-and shoulder video sequences. Such model based coding techniques use a 3-D wire mesh model of a human head and shoulders. A sprite image of a person is mapped on to the 3-D surface to represent the texture details of the person.
- Both model and human –face sprites are required to be sent by the transmitter to the receiver in the beginning and subsequently, for each frame, only a few parameters, that represent the motion of the person are to be transmitted. Transmission of 2 to 6 motion parameters per frame is sufficient for excellent predictions of the face region.

REFERENCES

- “Multimedia Systems Design” Prabhat K. Andleigh, Kiran Thakrar
- https://nptel.ac.in/courses/117105083