An Overview of H.263 and H.263+

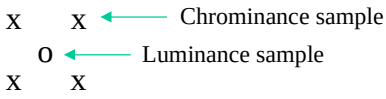
Thanks that Some slides come from Sharp Labs of America, Dr. Shawmin Lei
January 1999

ITU-T Video Coding Standards

- H.261: for ISDN
- H.263: for PSTN (very low bit rate video)
  - Four optional modes
- H.263 Version 2 (H.263+): an extension of H.263
  - 12 more negotiable modes
  - Scalable bit streams
  - Enhance performance over packet-switched networks
  - Support custom picture size and clock frequency
  - Provide supplemental display and external usage capabilities
- H.263++: under development, backward compatible to H.263+
- H.26L: under development, not backward compatible to H.263+

H.263 - Source Format

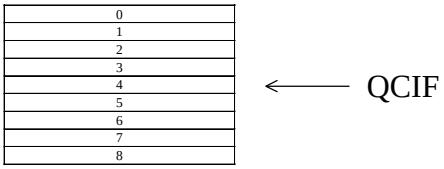
- 5 Formats: sub-QCIF (128x96), QCIF (176x144), CIF (352x704), 4CIF (704x576), and 16CIF (1408x1152)
- Color Format: 4:1:1



- Aspect Ratio:
  - Pixel Aspect Ratio: (4/3)*(288/352) = 1.09 for all 5 formats
  - Picture Aspect Ratio: 4:3 except for the sub-QCIF format

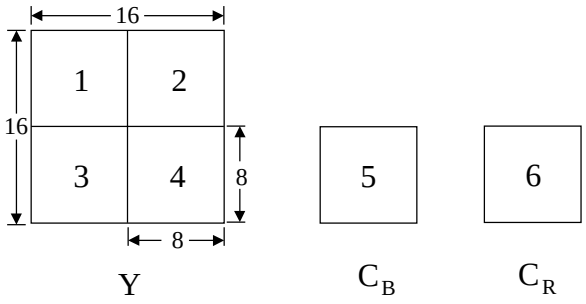
H.263: Syntax Structure

- Picture Layer
- Group-of-Block (GOB) Layer
  - A GOB comprises k*16 lines (k=1 for sub-QCIF, QCIF, and CIF; k=2 for 4CIF; k=4 for 16CIF)



- Macroblock Layer:
  - A macroblock covers 16x16 luminance pixels area.
  - Usually contains 6 blocks except for PB-frame mode (12 blocks instead).
- Block Layer: Each block contains 8x8 pixels.

Macroblock Structure

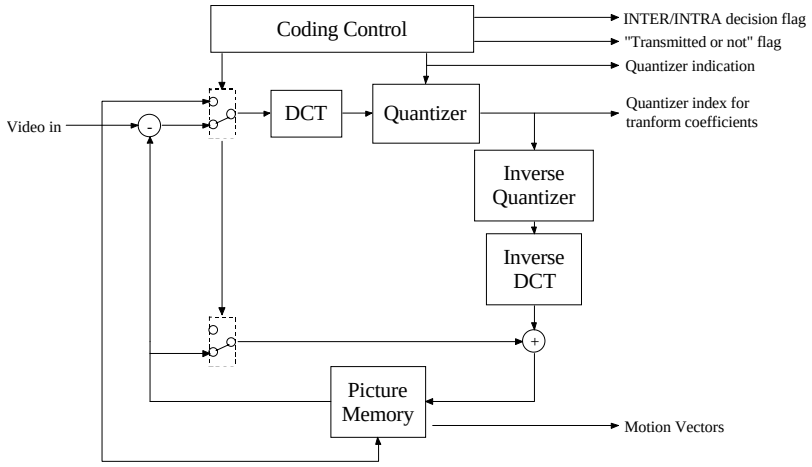


Note: In PB-frame mode, 6 P-blocks are followed by 6 B-blocks.

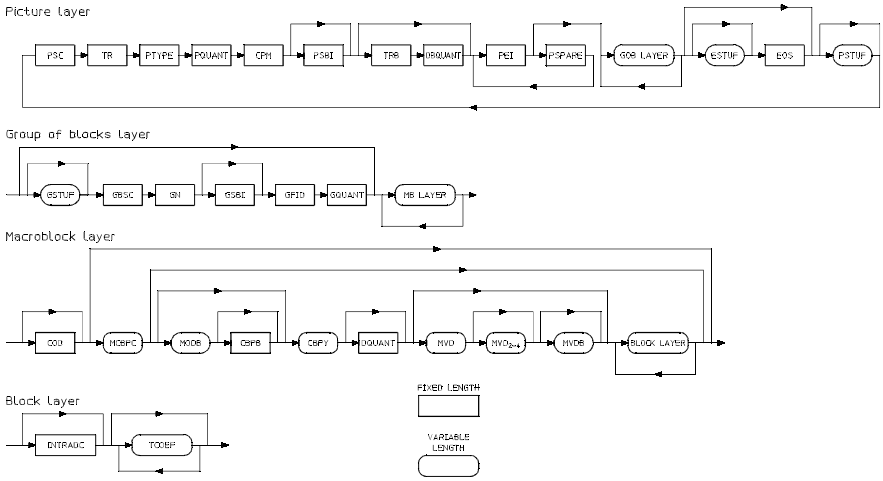
H.263 Baseline - Coding Techniques

- Motion-compensated prediction
  - Block-based, translation model, one motion vector per 16x16 macro-block
  - Half-pixel accuracy, [-16, 15,5] range
- Transform: 8x8 DCT
- Quantization: uniform with a central dead zone around zero
  - Same step size within a macro-block (i.e., uniform quantization matrix)
  - Even quantization levels from 2 to 62 are allowed.
  - DC coefficients of an intra block are always quantized by a step size of 8.
- Entropy Coding:
  - Zig-zag scan
  - 3-dimensional run-length VLC: (LAST, RUN, LEVEL)

H.263 Video Encoder



H.263 Video Bitstream Syntax



H.263 Syntax Elements

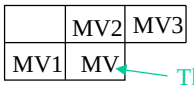
- Picture Layer
  - –PSC (22): Picture Start Code
  - –TR (8): Temporal Reference
  - –PTYPE (13): Picture Type Info.
  - –PQUANT (5): Picture Quantizer
  - –CPM (1): Continuous Presence Multipoint
  - –PSBI (2): Picture Sub-Bitstream Indicator
  - –TRB (3): Temp. Ref. for B-picture
  - –DBQUANT (2): DQUANT for B-pic.
  - –PEI (1): Extra Insertion Information
  - –PSPARE (8): Spare Information
  - –ESTUF (V<8): Stuffing
  - –EOS (22): End of Sequence
  - –PSTUF (V<8): Stuffing
- GOB Layer
  - –GSTUF (V<8): Stuffing
  - –GBSC (17): GOB Start Code
  - –GN (5): Group Number
  - –GSBI (2): GOB Sub-Bitstream Indica.
  - –GFID (2): GOB Frame ID
  - –GQUANT (5): GOB QUANT Info.

H.263 Syntax Elements (Continued)

- Macroblock Layer
  - –COD (1): Coded Macroblock Indic.
  - –MCBPC (V): Macroblock type & Coded Block Pattern for Chrominance
  - –MODB (V): MB Mode for B-blocks
  - –CBPB (6): Coded Block Pattern for B-blocks
  - –CBPY (V): Code Block Pattern for Y
  - –DQUANT (2): Differential QUANT
  - –MVD (V): Motion Vector Difference
  - –MVD2-4 (V): MVDs in Adv. Pred.
  - –MVDB (V): MVD for B-blocks
- Block Layer
  - –INTRADC (8): DC coefficient for INTRA blocks
  - –TCOEF (V): Transform Coefficients

Major Differences between H.261 and H.263 Baseline

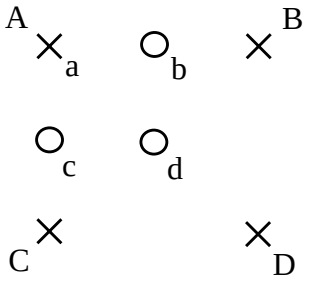
- **Source Formats**: H.263 supports 5 while H.261 supports 2.
- **PSC Byte Alignment:** Yes for H.263 but no for H.261
- **PQUANT:** Added in H.263 Picture Layer.
- **GOP Structure**
- **MB coded or not:** Identified by COD in H.263, instead of MBA.
- **Motion Compensation Accuracy:** Half-pixel accuracy for H.263
- **Loop Filter:** None in H.263 while optional in H.261
- **Motion Vector Predictor:**
  - H.263: Median value of the three candidate motion vectors (MV1-3)
  - H.261: Motion vector of the preceding macroblock (MV1).



H.261 and H.263 Baseline (Continued)

- **MB Quant:**
  - DQUANT (2 bits) specifies a change up to 2 steps in H.263.
  - MQUANT (5 bits) specifies absolute value of the new QUANT in H.261.
- **Entropy Coding of DCT Coefficients:**
  - H.263: (LAST, RUN, LEVEL)
  - H.261: (RUN, LEVEL) and EOB

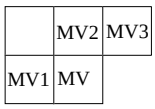
Bilinear Interpolation for Half Pixels



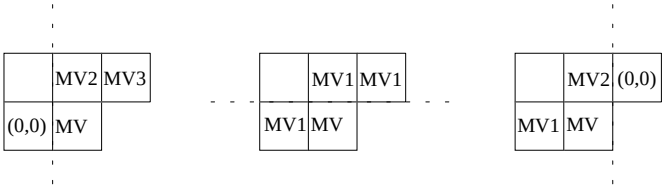
× Integer pixel position
○ Half pixel position

a = A
b = (A + B + 1) / 2
c = (A + C + 1) / 2
d = (A + B + C + D + 2) / 4

Motion Vector Prediction



MV : Current motion vector
MV1: Previous motion vector
MV2: Above motion vector
MV3: Above right motion vector



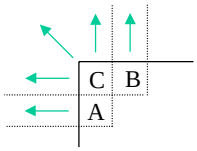
- - - - - : Picture or GOB border

H.263 Optional Modes

- Unrestricted Motion Vector Mode (Annex D)
  - MVs are allowed to point outside (outside pixels obtained from boundary repetition extension)
  - Larger ranges: [-31.5, 31.5] instead of [-16, 15.5]
- Syntax-Based Arithmetic Coding Mode (Annex E)
  - Provide about 5% bit rate reduction and rarely used
- Advanced Prediction Mode (Annex F)
  - Allow 4 motion vectors per MB, one for each 8x8 block
  - Overlapped block motion compensation for luminance
  - Allow MVs point outside of picture.
  - Reduce blocking artifacts and increase subjective picture quality.
- PB-Frames Mode (Annex G)
  - Double the frame rate without significant increase in bit rate

Unrestricted Motion Vector Mode

- Motion vectors are allowed to point outside the picture.
  - Outside referenced pixels are extended from boundary pixels.



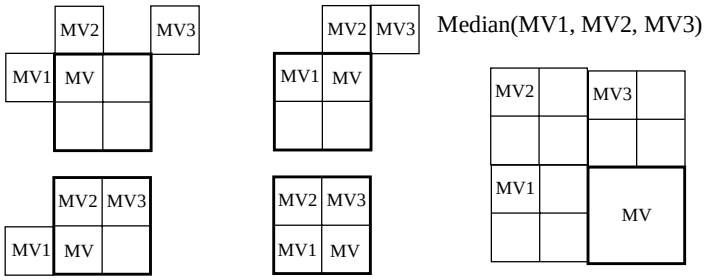
- Extended motion vector range from [-16, 15.5] to [-31.5, 31.5], with the following restrictions, depending on its predictor (P):
  - If 31.5 >= P >= 16.5, 31.5 >= MV >= 0
  - If 16 >= P >= -15.5, P + 15.5 >= MV >= -16 + P
  - If -16 >= P >= -31.5, 0 >= MV >= -31.5

Syntax-based Arithmetic Coding Mode

- Encode **macroblock layer** and **block layer** by **arithmetic codes**.
- Different coding models (just as different Huffman table) for different syntax elements.
- The coding models are **fixed** and defined in Annex E.
- More fine coding models:
  - Separate CBPB models for Y and UV
  - Different models for the first three TCOEFs
  - Different sets of coefficient models for inter and intra
- Allow fractional number of bits.
- The encoder needs to be flushed before sending PSC or GBSC, or at the end of sequence.
- Use bit stuffing (insert “1” after each successive 14 “0”s) to avoid start code emulation.

Advanced Prediction Mode

- Allow 4 motion vectors per macroblock:
  - One/four vectors decision
  - Motion vectors are differentially coded with a predictor as



- The chrominance MV is the sum of 4 MVs divided by 8.

Advanced Prediction Mode (Continued)

- Overlapped Motion Compensation for Luminance
  - Each pixel in a block is a **weighted average** of **3** prediction values.
    - The value (P0) predicted from the MV of the current block
    - The value (P1) predicted from the MV of the closer above or below block
    - The value (P2) predicted from the MV of the closer left or right block
  - If the closer macroblock was not coded, its MV is set to zero.
  - If the closer macroblock is outside picture, its MV is set to the MV of the current block.
  - If the closer macroblock was INTRA coded, its MV is set to the MV of the current block except in PB-frames mode the MV for the B-block is used.
  - If the current block is at the bottom of the MB (i.e., block 3 & 4), the MV of the below block is set to the MV of the current block.
  - If the PB-frames mode is also used, the overlapped motion compensation is only used for prediction of the P-pictures, not for the B-pictures.

Weighting Factors

4	5	5	5	5	5	5	4
5	5	5	5	5	5	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	4

Weighting Factor (W0) for P0

2	2	2	2	2	2	2	2
1	1	2	2	2	2	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	2	2	2	2	1	1
2	2	2	2	2	2	2	2

Weighting Factor (W1) for P1

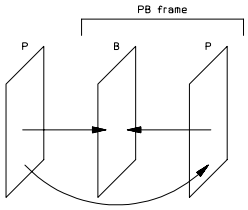
2	1	1	1	1	1	1	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	1	1	1	1	1	1	2

Weighting Factor (W2) for P2

The current pixel = (W0*P0 + W1*P1 + W2*P2 + 4)/8

PB-Frames Mode

- A PB-frame consists of 2 pictures being coded as one unit.



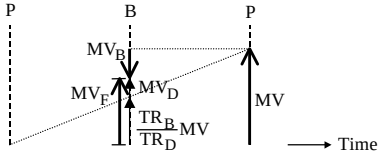
- In a PB-frame, a MB consists of 6 P-blocks and 6 B-blocks.
- B-blocks are always INTER coded (if coded), even in an INTRA macroblock. Thus, MVD is also included for INTRA macroblocks in PB-frames and is used by B-blocks.

PB-Frames Mode (continued)

- Calculation of Vectors for the luminance B-block in a PB-frame:

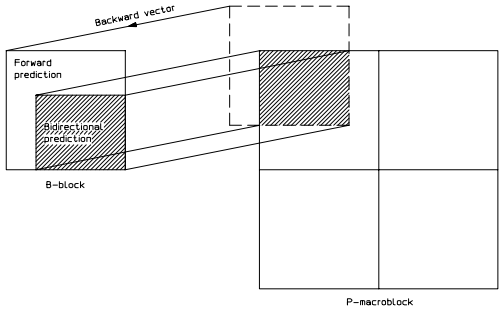
$MV_F = (TR_B \times MV) / TR_D + MV_D$
$MV_B = ((TR_B - TR_D) \times MV) / TR_D$ if MV_D is equal to 0
$MV_B = MV_F - MV$ if MV_D is unequal to 0

Where MV: the motion vector for the P-block
MV_D: the delta motion vector given by MVDB
MV_F: forward motion vector (from previous P-picture)
MV_B: backward motion vector (from current P-picture)
TR_D: Temporal Reference for the P-picture
TR_B: Temporal Reference for the B-picture



PB-Frames Mode (continued)

- P-macroblock (Y and Cr, Cb) is first reconstructed.
- Prediction of a B-block in a PB-frame:
  - For pixels where MVB points inside the P-macroblock, use bi-directional prediction (average of forward and backward predictions).
  - For all other pixels, use only forward prediction..



H.263+ Standard

- Official name should be **H.263 Version 2** approved in Jan. 1998.
- Backward Compatible with H.263 Version 1: H.263 is one of many modes in H.263+.
- Objectives:
  - Broaden the range of applications
  - Improve compression efficiency
- Custom Source Format (picture size, aspect ratio, clock frequency)
- Scalability
- Modified Unrestricted Motion Vector Mode
- 12 new optional modes

H.263+ Unrestricted Motion Vector Mode

- Motion vectors over picture boundaries (same)
- New restrictions for motion vector values:
  - No elements more than 15 pixels out horizontally or vertically
- Extension of the motion vector range (not depend on MV predictor)
  - When UUI = “1”:

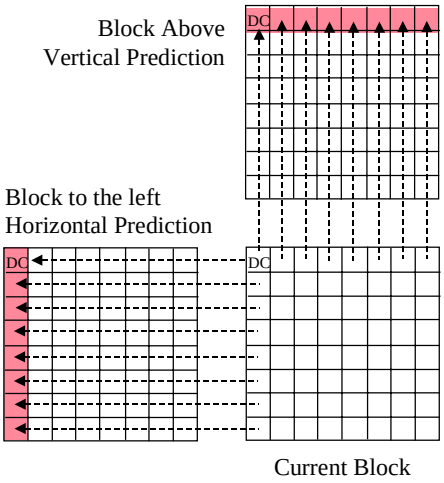
Picture width	Horizontal motion vector range	Picture height	Vertical motion vector range
4, …, 352	[-32, 31.5]	4, …, 288	[-32, 31.5]
356, …, 704	[-64, 63.5]	292, …, 576	[-64, 63.5]
708, …, 1408	[-128, 127.5]	580, …, 1152	[-128, 127.5]
1412,…, 2048	[-256, 255.5]		

  - When UUI = “01”: not limited except by its picture size.
- New Reversible VLC’s (RVLC’s) are used for encoding MVD.
  - These codes are single valued for easier implementation.

Advanced INTRA Coding Mode

- Use **inter block prediction** to improve compression efficiency of INTRA macroblocks.
- 3 INTRA_MODE Prediction with different scan patterns:
  - DC Only Prediction: The DC is predicted by the average of DCs of above and left blocks. (use zig-zag scan)
  - Vertical DC & AC Prediction: The first row of coefficients are predicted by those of the above block. (use alternate-horizontal scan)
  - Horizontal DC & AC Prediction: The first column of coefficients are predicted by those of the left block. (use alternate-vertical scan)
  - ✎ If the predictor (both predictors for DC only mode) is (are) not available, 1024 is used as the predictor.
- INTRA residuals are quantized using a variable step size without a “dead-zone”.
- A separate VLC table for intra coded coefficients

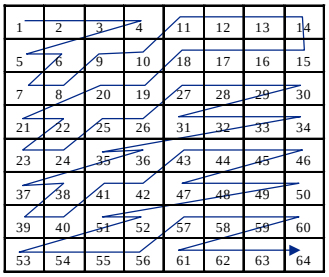
Inter Block Prediction in Advanced INTRA Coding Mode



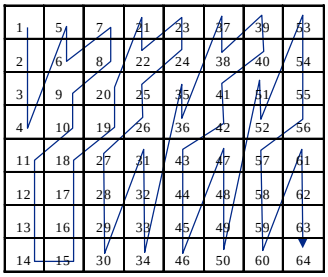
Zig-Zag Scan of DCT Coefficients

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

Alternate Scan Patterns



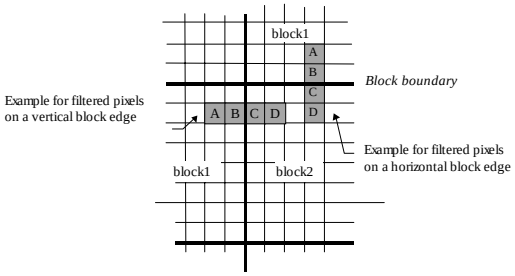
(a) Alternate-Horizontal scan



(b) Alternate-Vertical scan (as in Recommendation H.262)

Deblocking Filter Mode

- An optional deblocking filter within the coding loop
  - Filtered Pixels



  - The weight of the filter's coefficients depends on QUANT
  - The filtering is also tapped off when it appears to be a true edge.
- 4 motion vectors per macroblock as Advanced Prediction Mode
- Motion vectors over picture boundaries as Unrestricted MV Mode

Slice Structured Mode

- Slice Structure, instead of GOP structure
- Allow subdivision of a picture into segments containing variable numbers of macroblocks
- Two additional submodes for different order of transmission:
  - Rectangular Slice Submode
  - Arbrary Slice Ordering Submode
- No data dependence cross the slice boundaries within the current picture
- Flexible structure useful for error resilience or region of interest

Supplemental Enhancement Information Mode

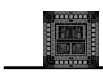
- Supplemental information can be included to support decoder features and functionalities within the video bit stream.
- Supplemental information includes support for
  - **Picture freeze**: can be partial (rectangular) area freeze
  - **Picture snapshot**: allow part of or the full picture to be used as a still image
  - **Video segmentation**: can be used by an external application
  - **Progressive refinement**: quality refinement instead of pictures at different times
  - **Chroma keying**: to represent transparent or semi-transparent pixels which can be blended with background picture

Improved PB-frames Mode

- B-part of an improved PB frames can be predicted by
  - Bidirectional prediction: same as PB-frames mode, but no delta vector
  - Forward prediction: a separate MV, not delta, is transmitted
  - Backward prediction: no MV is transmitted
- Improve quality of B pictures, especially around scene changes

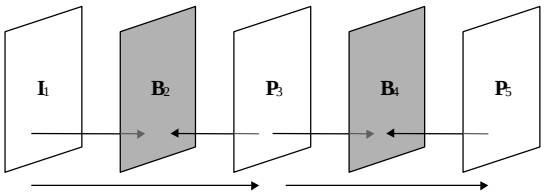
Reference Picture Selection Mode

- Can select the reference picture to suppress temporal error propagation.
- Additional picture memory is required.
- Can use backward channel to indicate which part of which pictures have been correctly decoded.
- Two back-channel mode switches: ACK and NACK



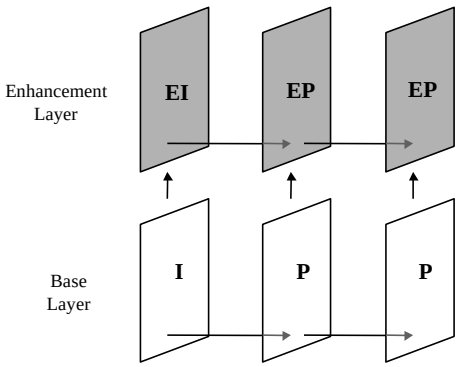
Temporal, SNR, and Spatial Scalability Mode

- Useful for decoders with different capability in multipoint and broadcast video applications.
- Temporal Scalability
  - B-pictures allows prediction from either or both forward and backward
  - B-pictures are separate entities, different from B-part of PB frames
  - Multiple B-pictures can be inserted between pairs of reference pictures.



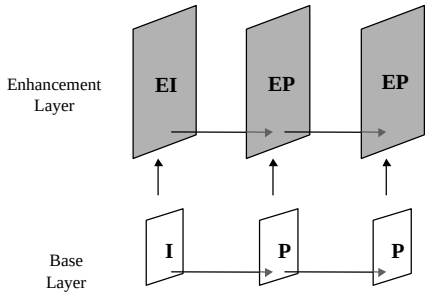
SNR Scalability

- Use finer quantizer to encode the difference picture in an enhancement layer.

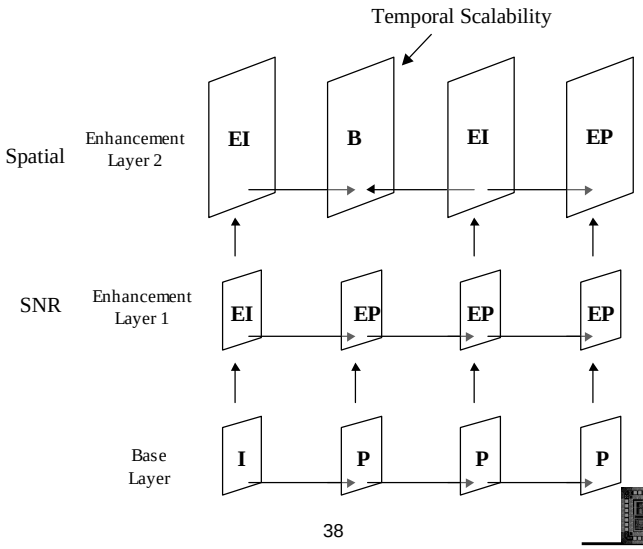


Spatial Scalability

- Allow multiresolution bit stream.
- Very similar to SNR Scalability, except the base layer contains lower resolution pictures and interpolation is used to predict the enhancement layer.
- The interpolation filters are normative part of the standard.



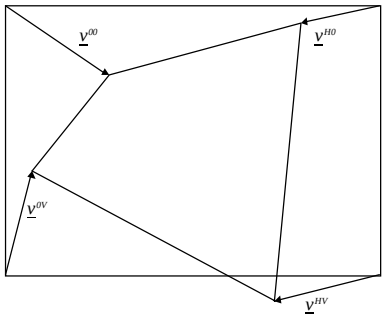
Multilayer Scalability



Reference Picture Resampling Mode

- Describe an algorithm to warp the reference picture prior to its use for prediction.
- Useful for resampling a reference picture having a different source format.
- Can also specify a “global motion” warping alteration of shape, size, location, and rotation.
- The simplest form: implicit factor of 4 resampling
  - Only FIR filters needed for upsampling and downsampling
- The reference picture resampling is defined in terms of the displacement of four corner of the current picture area.

Warping in Reference Picture Resampling Mode

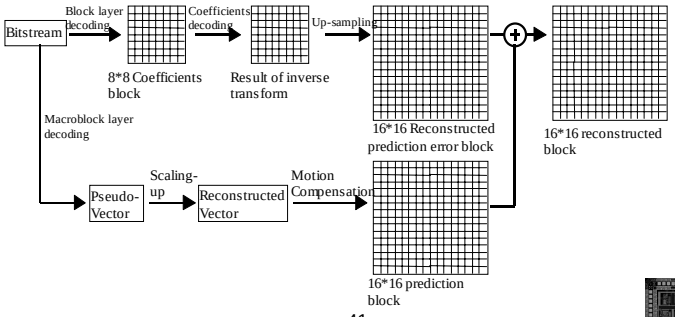


Bilinear Interpolation: $\bar{v}(x,y) = \left(1-\frac{y}{V}\right)\left[\left(1-\frac{x}{H}\right)\underline{v}^{00} + \left(\frac{x}{H}\right)\underline{v}^{H0}\right] + \left(\frac{y}{V}\right)\left[\left(1-\frac{x}{H}\right)\underline{v}^{0V} + \left(\frac{x}{H}\right)\underline{v}^{HV}\right].$

H: current picture width
V: current picture height

Reduced Resolution Update Mode

- Allow updating picture at lower resolution
- Most useful for highly active scenes with detailed backgrounds.
- Filtering is performed along the edges of the 16x16 reconstructed blocks at the encoder and decoder.
- Block decoding in Reduced Resolution Update Mode:



Independent Segment Decoding Mode

- No data dependencies across the segment boundaries are allowed.
- Segment boundaries are treated as picture boundaries.
- A segment is
  - A GOB with non-empty header plus consecutive GOBs with empty headers, if Slice Structured Mode is not used;
  - A Slice, if Slice Structured Mode is used.
- Limit the propagation of errors.
- Enhance error resilience and recovery capabilities.
- Better used with slice structures.
- Constraints:
  - Only rectangular slices are allowed, if Slice Structured Mode is used.
  - Segmentation of current frame shall be the same as that of its reference frames.

Alternative INTER VLC Mode

- Allow VLC tables originally designed for INTRA Coding to be used for some INTER coding coefficients and CBPY data.
- Better efficiency when small quantizer step sizes are used or when significant changes are evident in the picture.
- Note that VLC tables for INTER and INTRA contain the same codewords, but different interpretation of LEVEL and RUN.
- The INTRA VLC table can be used only when the decoder can detect its use by decoding with INTER VLC table first and resulting more than 64 coefficients.
- INTRA CBPY table is used for encoding INTER CBPY when both C_B and C_R blocks have at least one non-zero coefficient.

Modified Quantization Mode

- Allow modification of the quantizer at macroblock layer to any value, not limited to +1, -1, +2, and -2.
  - DQUANT uses 2 bits (started with “1”) to specify small changes.

	Change of QUANT	
Prior QUANT	DQUANT = 10	DQUANT = 11
1	+2	+1
2–10	–1	+1
11–20	–2	+2
21–28	–3	+3
29	–3	+2
30	–3	+1
31	–3	–5

  - It uses 6 bits (started with “0”) to specify other changes.
    Codeword: 0xxxxx
    where the last 5 bits specify the new QUANT value.

Modified Quantization Mode (Continued)

- Enhance chrominance quality by a finer quantizer.

Range of QUANT	Value of QUANT_C
1–6	QUANT_C = QUANT
7–9	QUANT_C = QUANT – 1
10–11	9
12–13	10
14–15	11
16–18	12
19–21	13
22–26	14
27–31	15

- Improve picture quality by extending the range of representable quantized DCT coefficients, not limited by [-127, +127].

Recommended Optional Enhancement

- For backward compatibility to H.263:
  - Advanced Prediction Mode (Annex F)
- Level 1 Preferred Modes:
  - Advanced INTRA Coding Mode (Annex I)
  - Deblocking Filter Mode (Annex J)
  - Supplemental Enhancement Information (Full-Frame Freeze Only) (Annex L Section L.4)
  - Modified Quantization Mode (Annex T)
- Level 2 Preferred Modes:
  - Unrestricted Motion Vectors Mode (AnnexD)
  - Slice Structured Mode (Annex K)
  - Reference Picture Resampling (Implicit Factor-of-4 Mode Only) (Annex P)

Recommended Optional Enhancement (Continued)

- Level 3 Preferred Modes:
  - Advanced Prediction Mode (Annex F)
  - Improved PB-frames (Annex M)
  - Independent Segment Decoding (Annex R)
  - Alternate INTER VLC (Annex S)