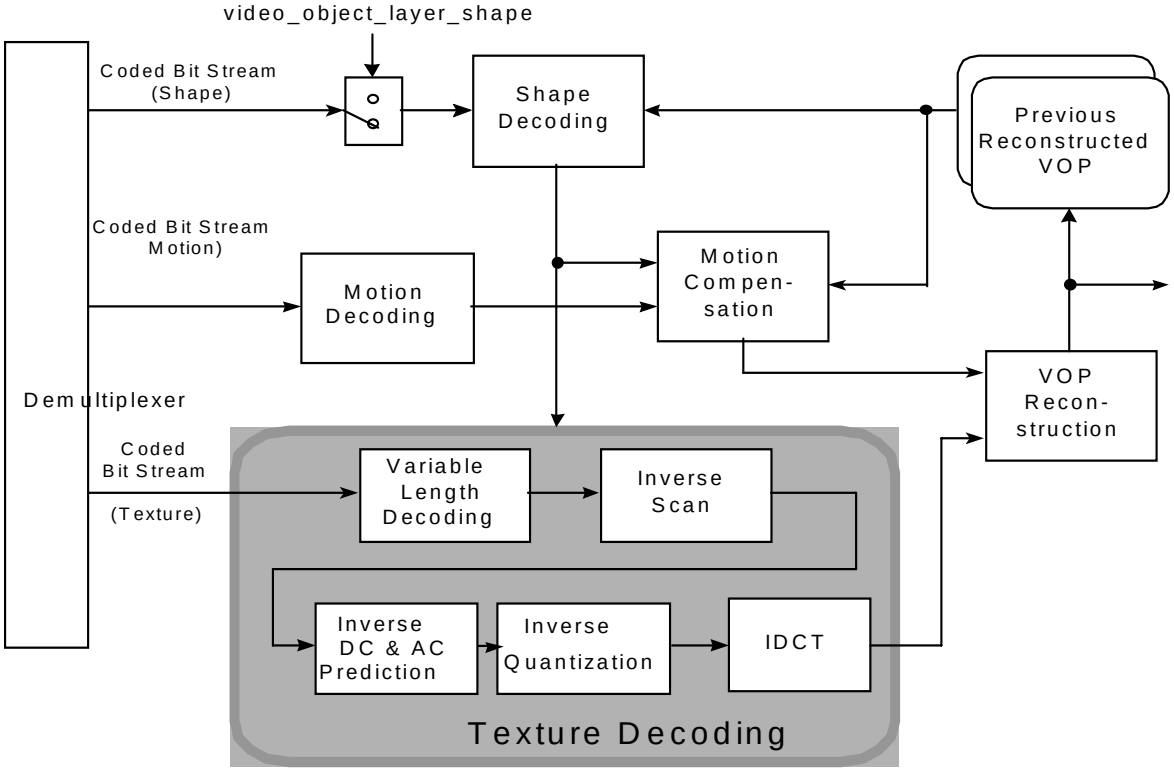
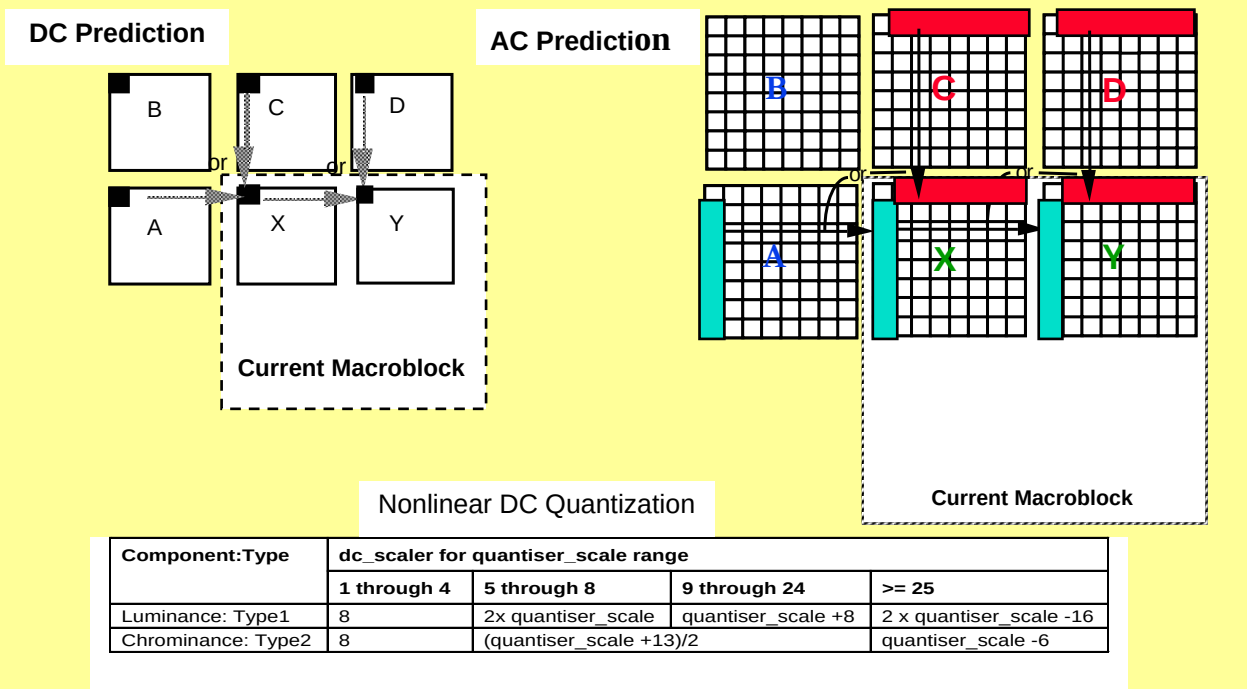
MPEG-4 Video Decoder:



Coding of Intra VOPs: Improved adaptive intra macro-block coding as compared to MPEG-1/2



Component:Type	dc_scaler for quantiser_scale range			
	1 through 4	5 through 8	9 through 24	>= 25
Luminance: Type1	8	2x quantiser_scale	quantiser_scale +8	2 x quantiser_scale -16
Chrominance: Type2	8	(quantiser_scale +13)/2		quantiser_scale -6

H.263 vs MPEG Quantization:

1.H.263 Quantization

– Intra DC: LEVEL = COF// dc_scaler

dc_scaler uses two Qp dependent tables one for Luminance other for Chrominance

– Intra AC Forward Quant: *LEVEL = |COF|/(2xQp); clip to [-127:127]*

– Inter Forward Quant: *LEVEL = (|COF|-Qp/2)/(2xQp); clip to [-127:127]*

2.MPEG Quantization

– Intra DC: *LEVEL = COF// dc_scaler*

dc_scaler uses two Qp dependent tables one for Luminance other for Chrominance

– Intra AC Forward Quant: *ac~[i][j] = (16*ac[i][j])//WI[i][j]; clip to [-2048, 2047]*

LEVEL = (ac~[i][j] + sign(ac~[i][j])((a*Qp) // b)) / (2*Qp); clip to [-127:127]*

– Inter Forward Quant: ac~[i][j] = (16*ac[i][j])//WN[i][j]; clip to [-2048, 2047]

*LEVEL = ac~[i][j] / (2*Qp); clip to [-127:127]*

DCT Coefficient Quantization and Scanning:

Two methods:

- Method 1 uses quantizer of H.263.
- Method 2 uses quantizer of MPEG-1/2.

Intra uses:

(i) Nonlinear quantizer for DC coefficients.

(ii) Quantizer matrix (MPEG-1/2) as well as partial quantizer matrix loading.

(iii) Quantization of DCT coefficients adaptively scanned; 1 new scan compared to MPEG-2.

Alternate-Horizontal scan

0	1	2	3	10	11	12	13
4	5	8	9	17	16	15	14
6	7	19	18	26	27	28	29
20	21	24	25	30	31	32	33
22	23	34	35	42	43	44	45
36	37	40	41	46	47	48	49
38	39	50	51	56	57	58	59
52	53	54	55	60	61	62	63

Alternate-Vertical scan

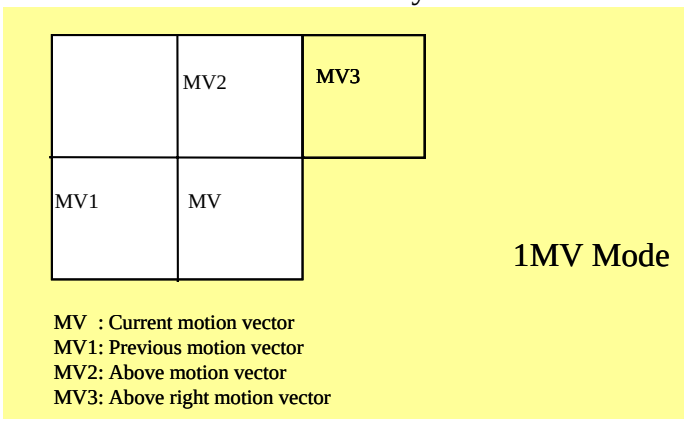
0	4	6	20	22	36	38	52
1	5	7	21	23	37	39	53
2	8	19	24	34	40	50	54
3	9	18	25	35	41	51	55
10	17	26	30	42	46	56	60
11	16	27	31	43	47	57	61
12	15	28	32	44	48	58	62
13	14	29	33	45	49	59	63

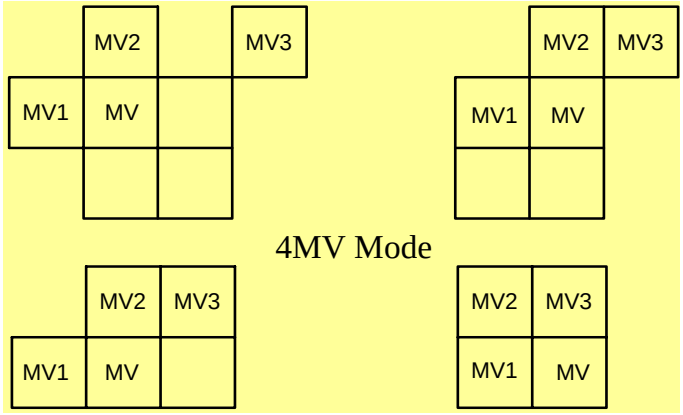
Zig-zag scan

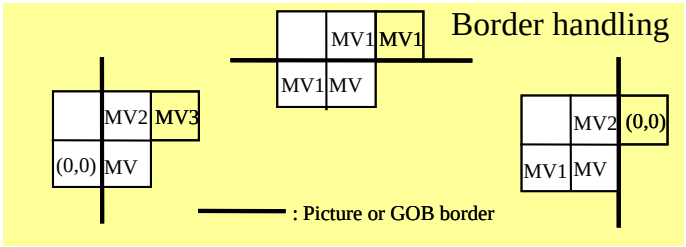
0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Predictive Motion Vector Coding:

$$MV_x = \text{Median}\{MV1_x, MV2_x, MV3_x\}$$
$$MV_y = \text{Median}\{MV1_y, MV2_y, MV3_y\}$$



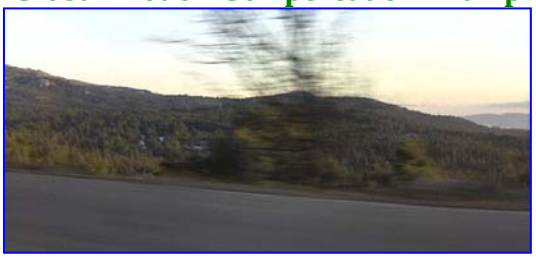




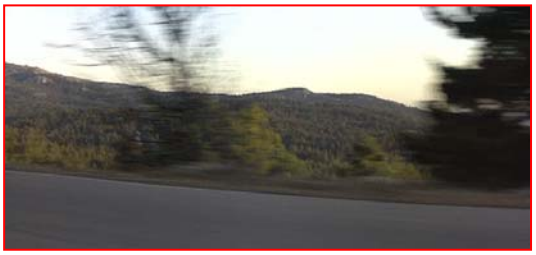
-If one or more of MV1, MV2, and MV3 is outside the picture border or the current video packet, then it is classified as *invalid*.

1) If only one MV is invalid, it is set to 0
2) If two MVs are invalid, they are set to the third MV
3) If all three MVs are invalid, they are all set to 0.

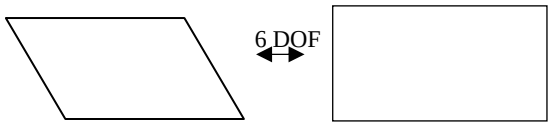
Global Motion Compensation Example:



Frame 3



Frame 6



Affine : $(a_1, a_2, a_3, a_4, a_5, a_6)$

$x' = a_1 + a_3x + a_4y$

$y' = a_2 + a_5x + a_6y$

$a_1 = 77.06$, $a_3 = 1.046$, $a_4 = -0.104$

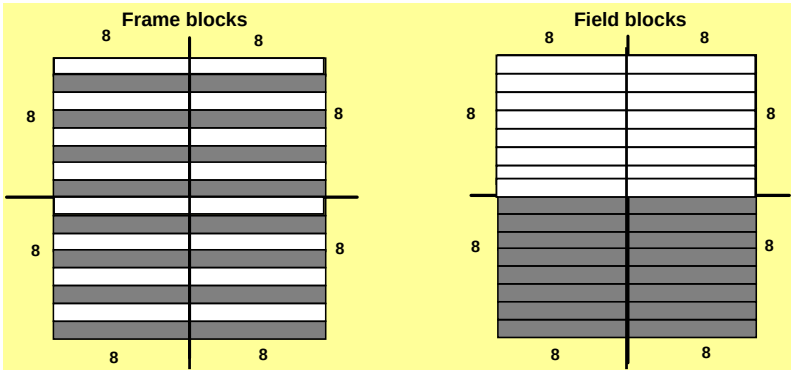
$a_2 = 9.38$, $a_5 = -2e\text{-}4$, $a_6 = 0.98$

Reduced Resolution Coding:

- Encodes the I and P-VOP in reduced spatial resolution or normal resolution(normal VM scheme) adaptively.
- **Resolution Decision**: The decided resolution is signaled to decoder through the "vop_reduced_resolution" flag in VOP header. If the reduced resolution is selected, the VOP is encoded as the following:
    - **Motion Estimation**: Motion estimation is done in 32 x 32 macroblocks with expanded search area and with restricted search positions in half pel search.
    - **Motion Compensation**: Motion compensation is done in 32 x 32 macroblocks, and the prediction error is generated in each 32 x 32 macroblock.
    - **Scaling down of the Motion Vectors**: The scaled down motion vectors are encoded as described in the VM.
    - **Down Sampling of the Prediction Error :** The prediction error is down-sampled to the half resolution. After the down sampling, the size of macroblock is 16x16 and the size of block is 8x8.
    - **Texture Coding of the Prediction Error** : The down sampled prediction error is encoded exactly same as described in the current VM(DCT, Q, VLC)

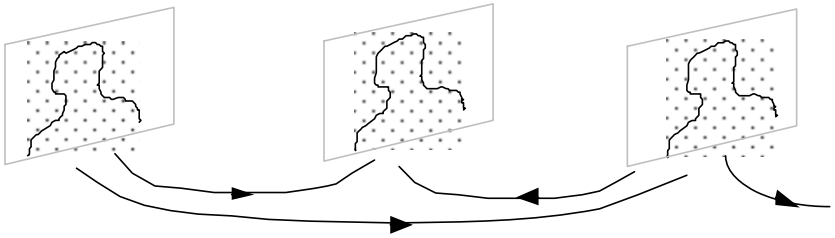
Interlaced Video Coding: Allows interlaced I-, P-, B- VOP coding

- MC for field or frame pictures similar to MPEG-2
- Modified AC/DC prediction
- Frame/Field DCT similar to MPEG-2
- Modified rules for scanning of coefficients

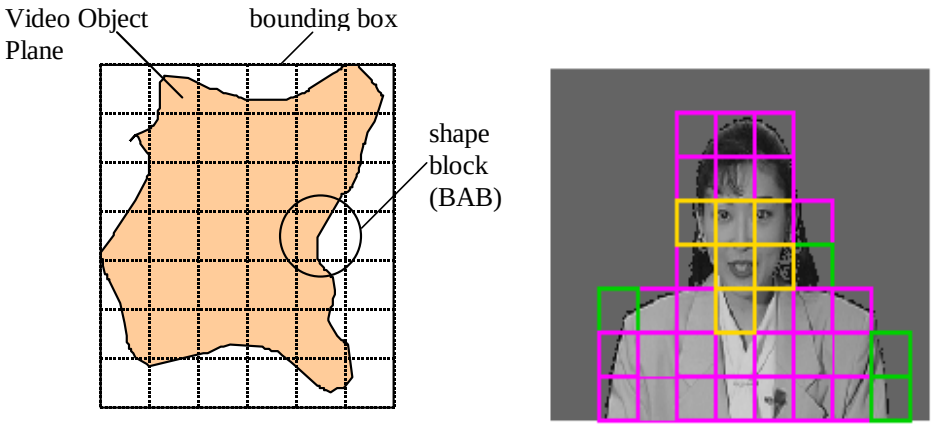


VOP Coding Types:

- VOP coding involves shape, motion and texture coding
- I, P, and B VOPs similar to I, P and B pictures
- I-VOPs have only intra macroblocks; optimized coding of intra
- P or B VOPs has inter or intra macroblocks; improved MC tools allow optimized coding
- B-VOPs further optimized by adding a "direct" prediction mode



Shape Coding:



1. Binary Shape coding
   - Lossless coding (JBIG-like context-based arithmetic coding)
   - Lossy coding
   - Possible feathering around borders in the binary mode
2. Gray-level Shape Coding

Motion Compensation of VOPs:

- 16x16 block MC same as in MPEG-1/2
- Half or Quarter pel motion compensation
- Global MC using 8 motion parameters wrt the previous VOP