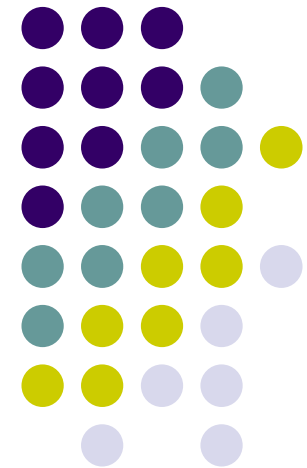


Video Compression Standards (II)

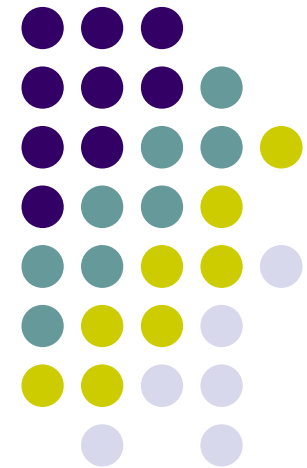
A/Prof. Jian Zhang

NICTA & CSE UNSW
COMP9519 Multimedia Systems
S2 2009

jzhang@cse.unsw.edu.au



Tutorial 2 : Image/video Coding Techniques





Basic Transform coding *Tutorial 2*

- Discrete Cosine Transform
 - For a 2-D input block U , the transform coefficients can be found as $Y = CUC^T$
 - The inverse transform can be found as $Y = CUC^T$
 - The $N \times N$ discrete cosine transform matrix $C=c(k,n)$ is defined as:

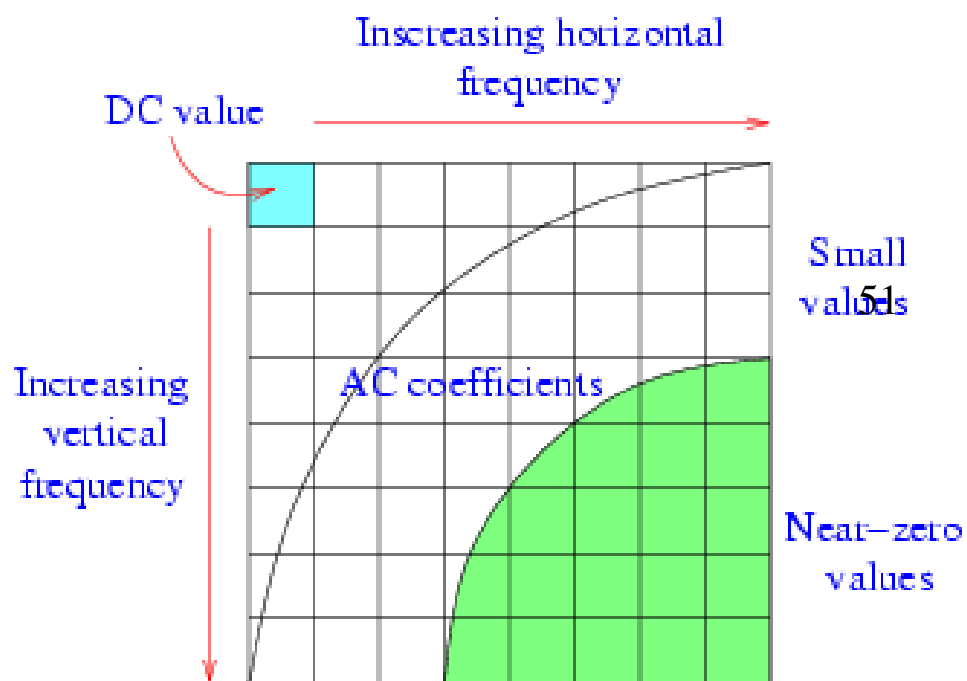
$$c(k, n) = \begin{cases} \frac{1}{\sqrt{N}} & \text{for } k = 0 \text{ and } 0 \leq n \leq N - 1, \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N} & \text{for } 1 \leq k \leq N - 1 \text{ and } 0 \leq n \leq N - 1. \end{cases}$$



Basic Transform coding *Tutorial 2*

- The distribution of 2-D DCT Coefficients

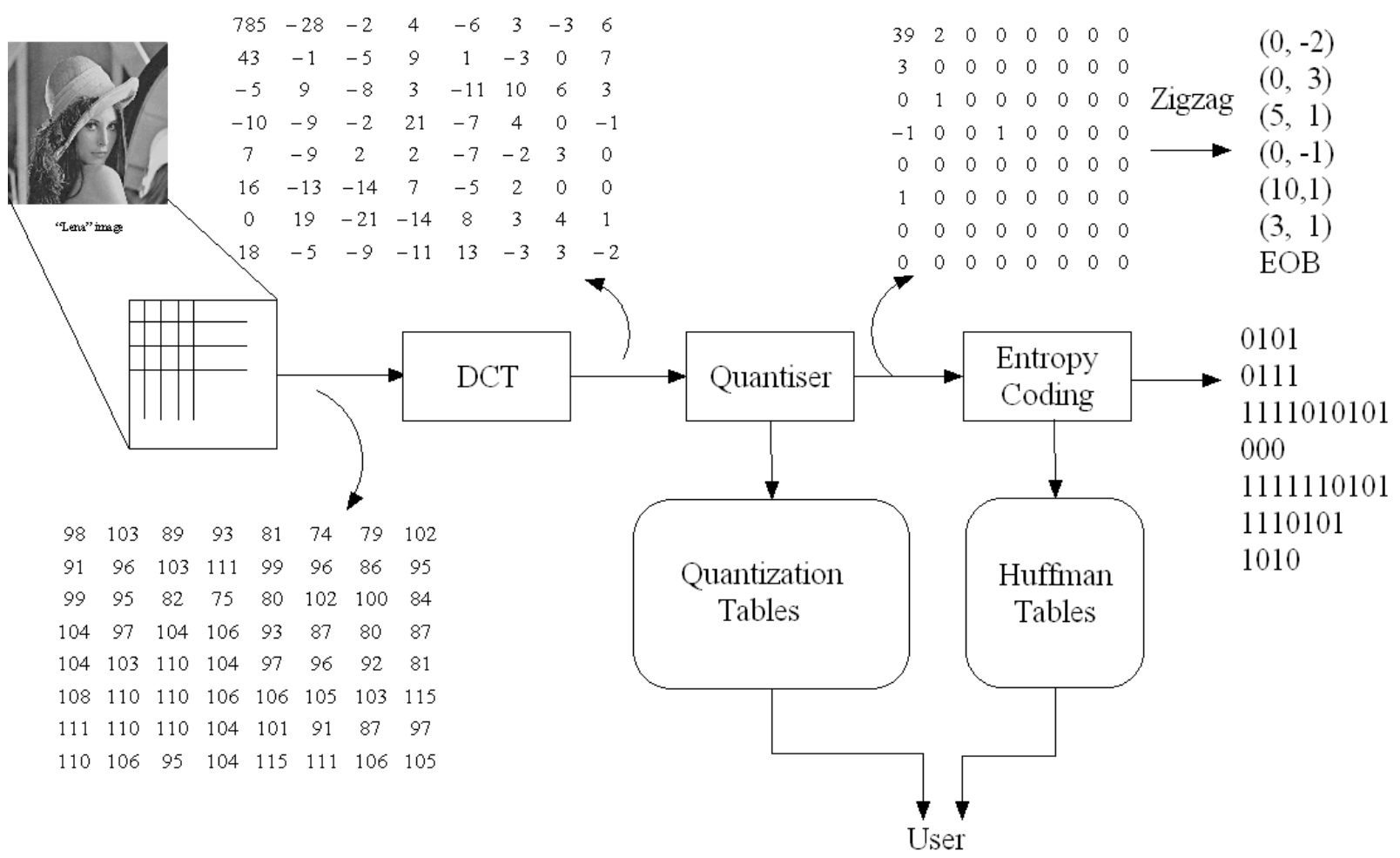
Ref: H. Wu



68	3	5	-2	0	0	-2	0
-10	0	-4	3	0	0	0	0
9	3	0	0	0	-2	0	0
3	2	0	3	0	2	-2	0
0	0	2	-2	0	0	0	0
0	2	-2	2	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0



JPEG DCT-Based Encoding *Tutorial 2*

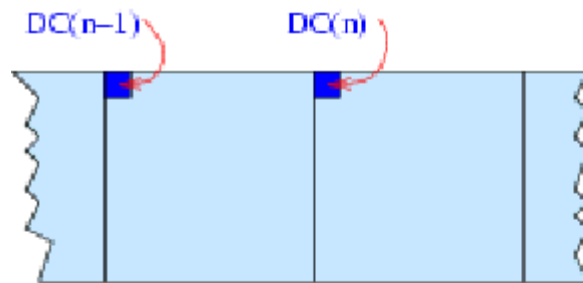


Coding of DCT Coefficients (DC)

Tutorial 2



- DC coefficient is coded differentially as (size, amplitude). There are 12 size categories



$$\Delta DC_i = DC_i - DC_{i-1}$$

-3 36 39

Size: 2

-3 00

-2 01

2 10

3 11

[Coeff]	Size	[Code]	Length
0	0	00	2+0
1	1	010	3+1
2..3	2	011	3+2
4..7	3	100	3+3
8..15	4	101	3+4
16..31	5	110	3+5
32..63	6	1110	4+6
64..127	7	11110	5+7
128..255	8	111110	6+8
256..511	9	1111110	7+9
512..1023	10	11111110	8+10
1024..2047	11	111111110	9+11

Final code:01100

Coding of DCT Coefficients (AC)

Tutorial 2



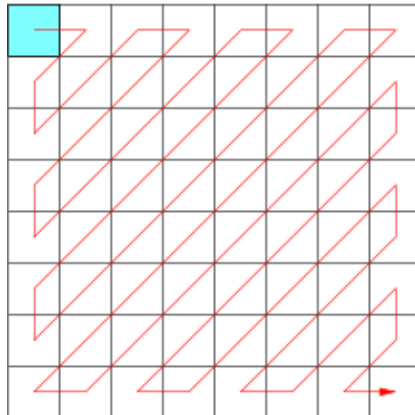
- AC coefficients are re-arranged to a sequence of (run, level) pairs through a zigzag scanning process
 - Level is further divided into (Size Categories, Amplitude).
 - Run and size are then combined and coded as a single event (2D VLC)
 - An 8-bit code 'RRRRSSSS' is used to represent the nonzero coefficients
 - The SSSS is defined as size categories from 1 to 11
 - The RRRR is defined as run-length of zeros in the zig-zag scan or number of zeros before a nonzero coefficient
 - The composite value of RRRRSSSS is then Huffman coded
- Ex: 1) RRRRSSSS=11110000 represents 15 run '0' coef. and followed by a '0' coef.
- 2) Multiple symbols used for run-length of '0' coef. exceeds 15
- 3) RRRRSSSS=00000000 represents end-of-block (EOB)

Coding of DCT Coefficients (AC)

Tutorial 2



Zig-Zag scan



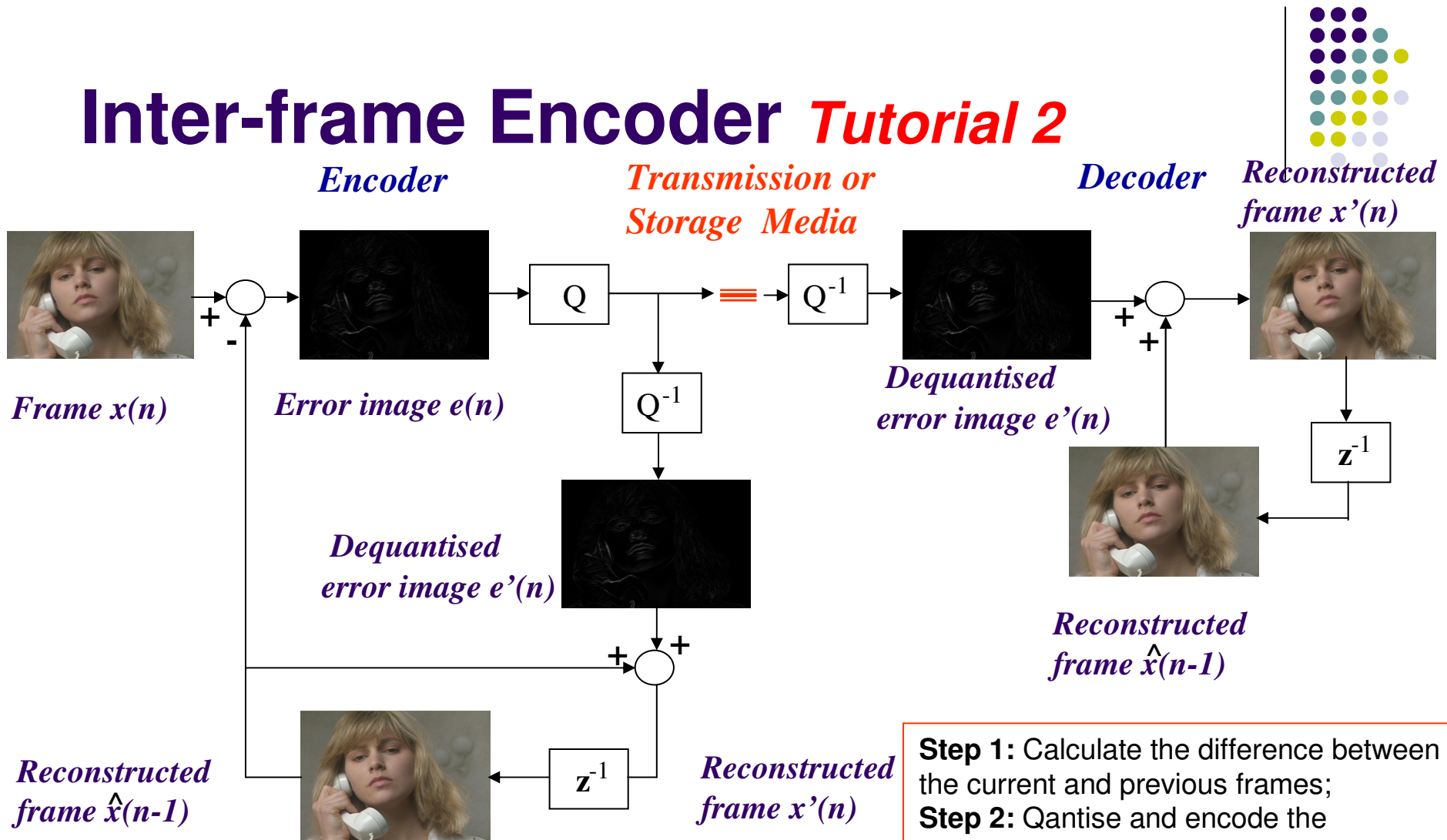
SSSS

RRRR

	0	1	2	9	10	11
0	EOB	Composite values				
.	N/A					
.	N/A					
15	ZRL					

Zero Run	Category	Code length	Code word
0	1	2	00
0	2	2	01
0	3	3	100
0	4	4	1011
0	5	5	11010
0	6	6	111000
0	7	7	1111000
.	.	.	.
1	1	4	1100
1	2	6	111000
1	3	7	1111000
1	4	9	11110110
.	.	.	.
2	1	5	10011
2	2	8	1111000
.	.	.	.
3	1	6	111010
3	2	9	11110111
.	.	.	.
4	1	6	111011
5	1	7	1111010
6	1	7	1111011
7	1	8	11111000
8	1	8	11111010
9	1	9	111111000
10	1	9	111111001
11	1	9	111111010
.	.	.	.
.	.	.	.
EOB	.	4	1010

Inter-frame Encoder *Tutorial 2*



Step 1: Calculate the difference between the current and previous frames;
Step 2: Quantise and encode the difference image.
Step 3: Add the dequantised (residual) image to the previous frame to reconstruct the current frame of image.

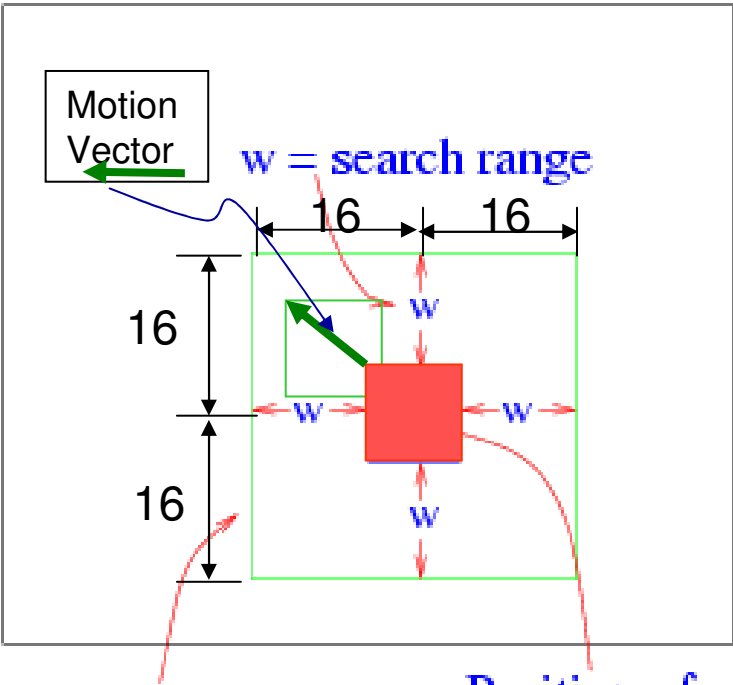
Block Based Motion Estimation

Tutorial 2



- Block base search

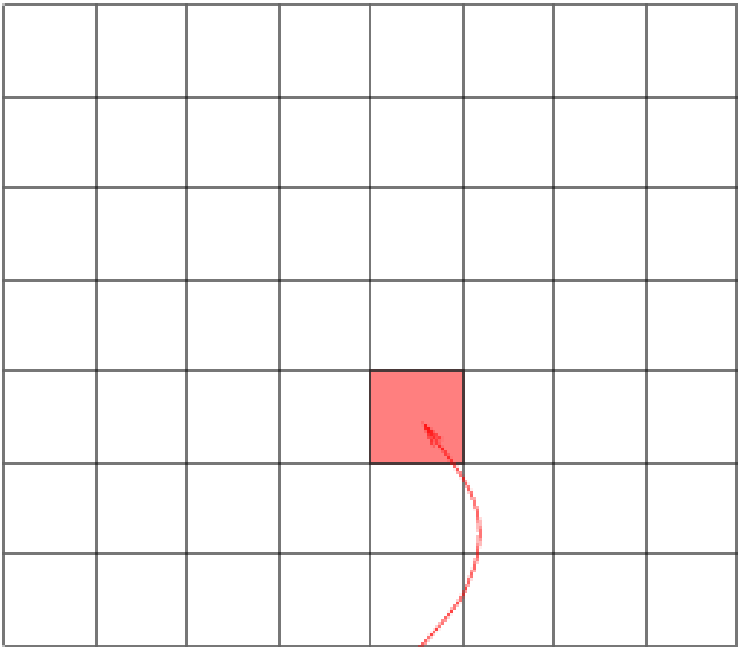
Reference Frame



Search Window

Position of Current Block

Current Frame



Current Block

16x16 -- Macroblock

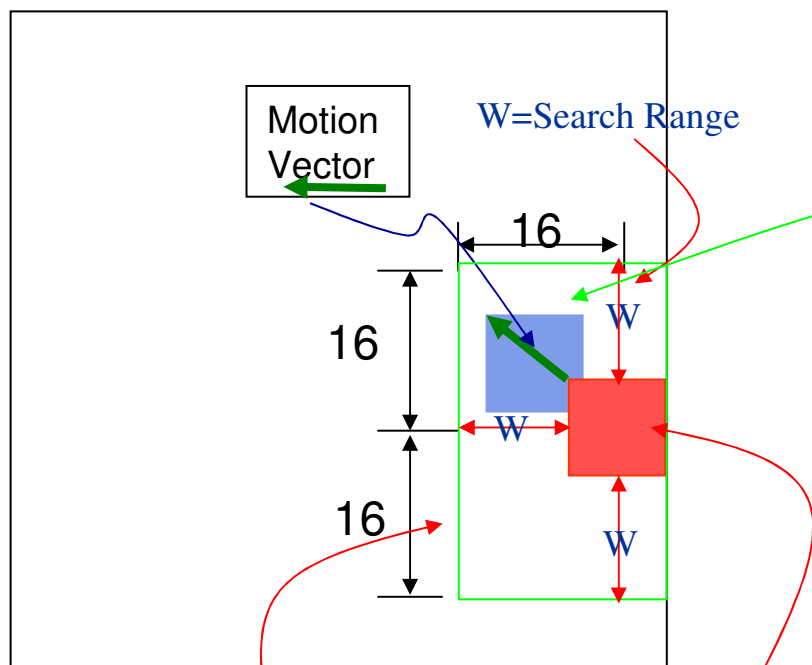
Block Based Motion Estimation

Tutorial 2



- Block base search

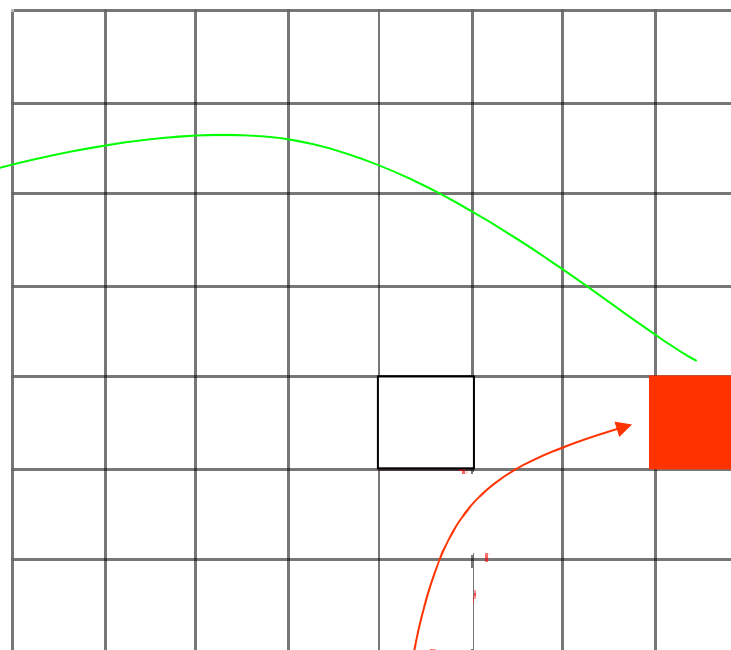
Reconstructed Frame



Search Window

Position of
Current Block

Current Frame



Current Block

16x16 -- Macroblock

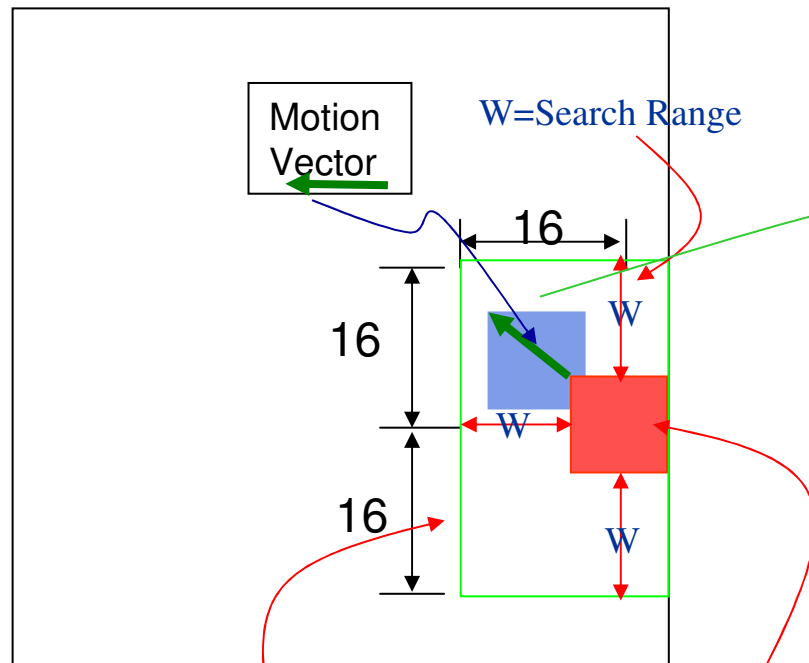
Block Based Motion Estimation

Tutorial 2



- Block base search

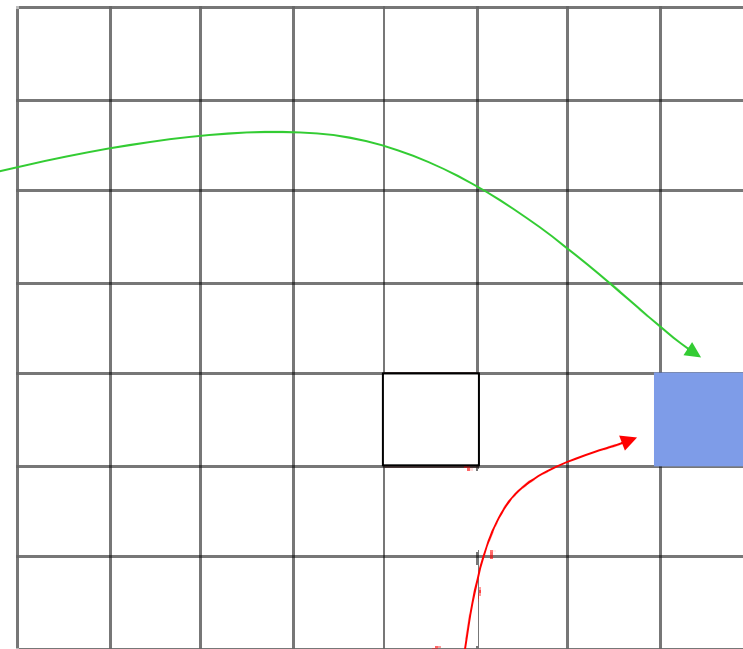
Reconstructed Frame



Search Window

Position of
Current Block

Motion Compensated Frame

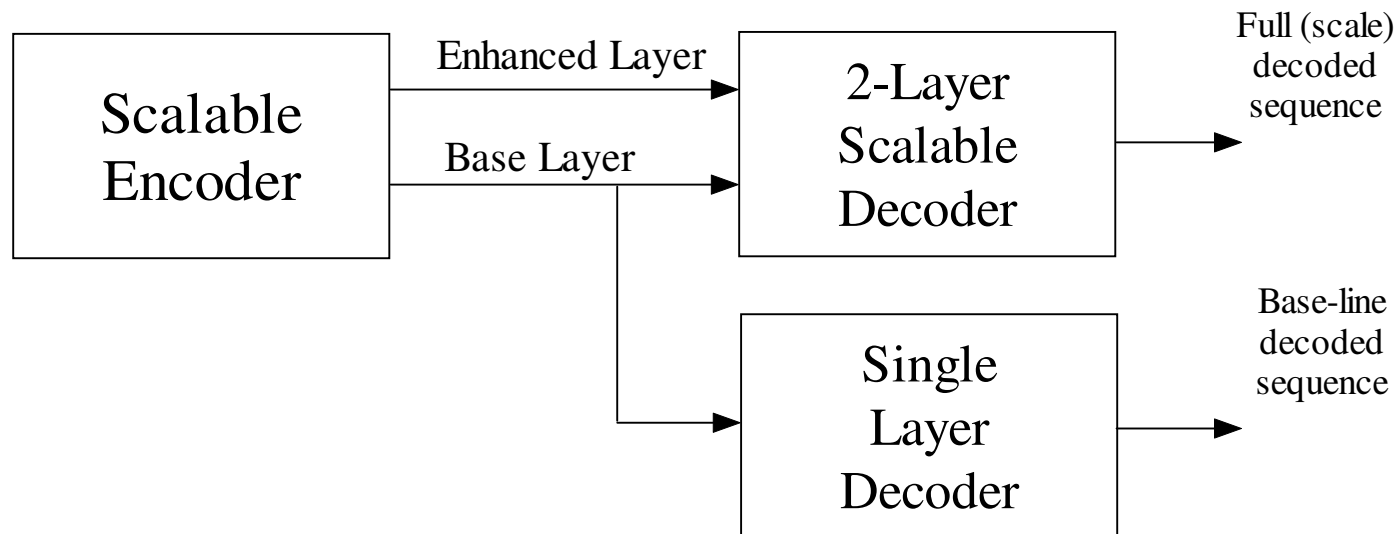


Motion Compensated MB
16x16 -- Macroblock

4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



- Scalable video coding means the ability to achieve more than one video resolution or quality simultaneously.



4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability

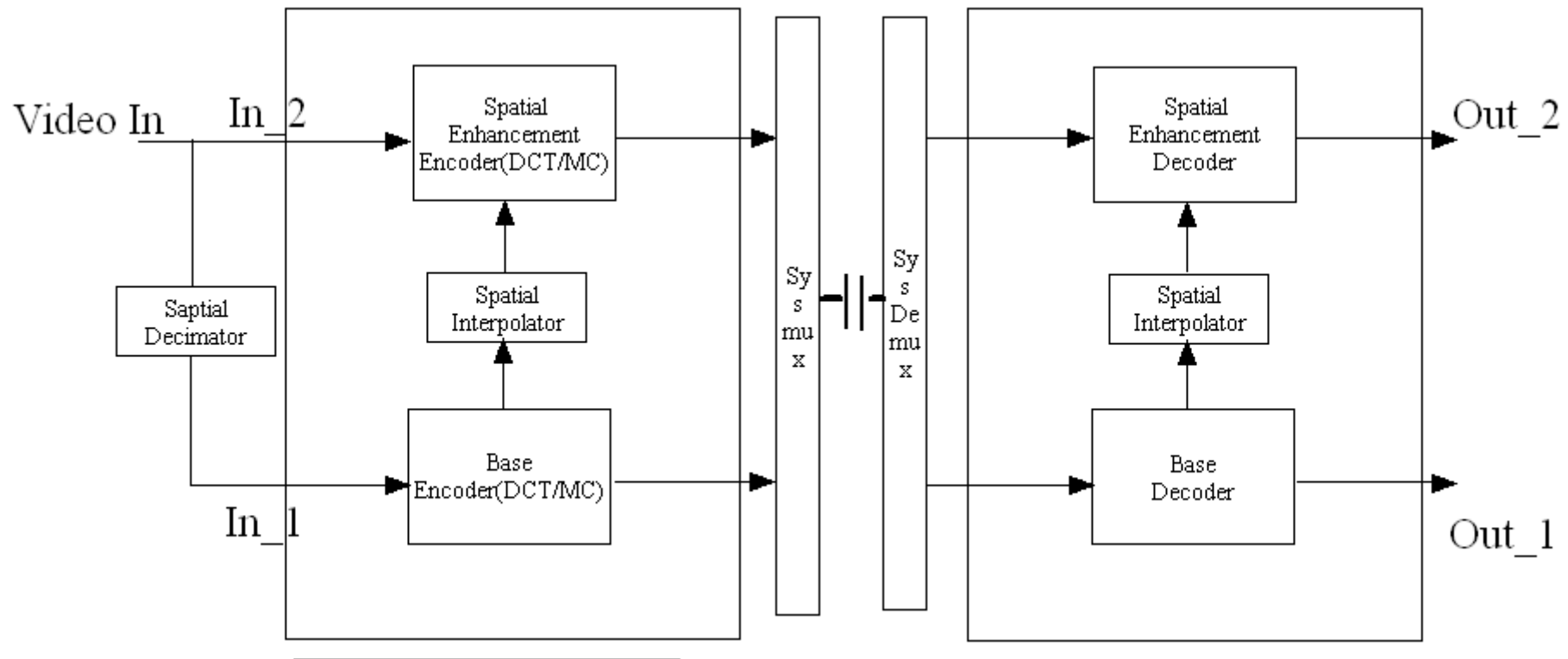


- Spatial Scalability
 - A spatially scalable coder operates by filtering and decimating a video sequence to a smaller size prior to coding.
 - An up-sampled version of this coded base layer representation is then available as a predictor for the enhanced layer
 - As prediction is performed in the spatial domain, the coding at the base layer can take any other standards including (MPEG-1 or H.261).
 - This is an important feature to address compatibility in layered codec

4.1 Digital Video Coding (DVC) Standards – MPEG-2 Scalability



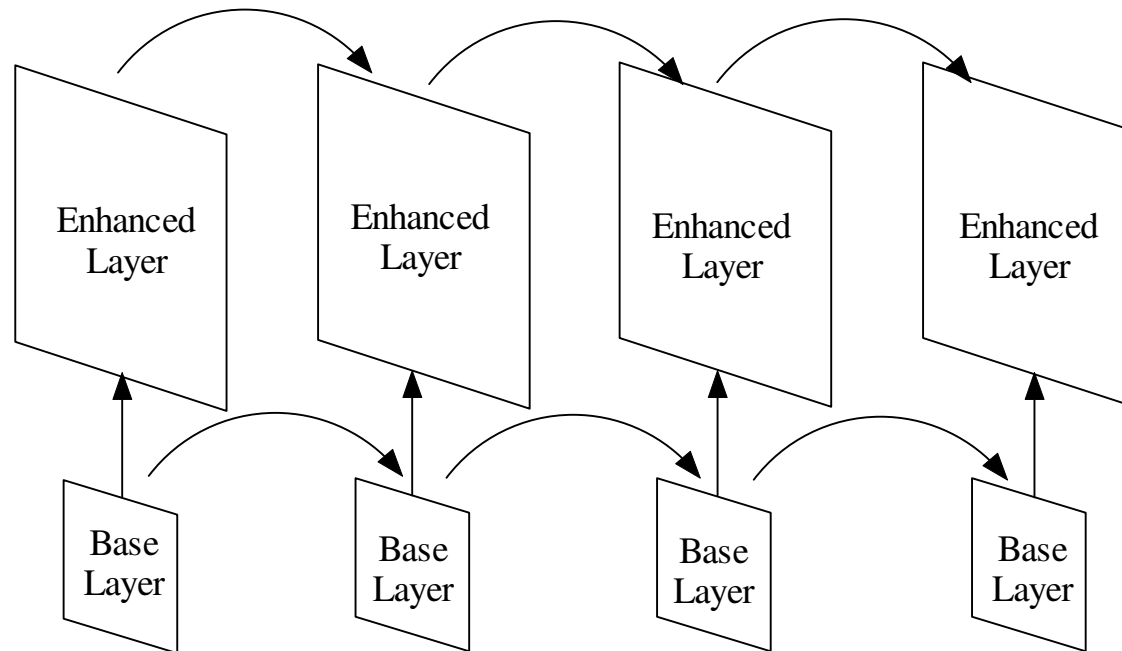
- Spatial Scalability – Spatial Scalability Codec



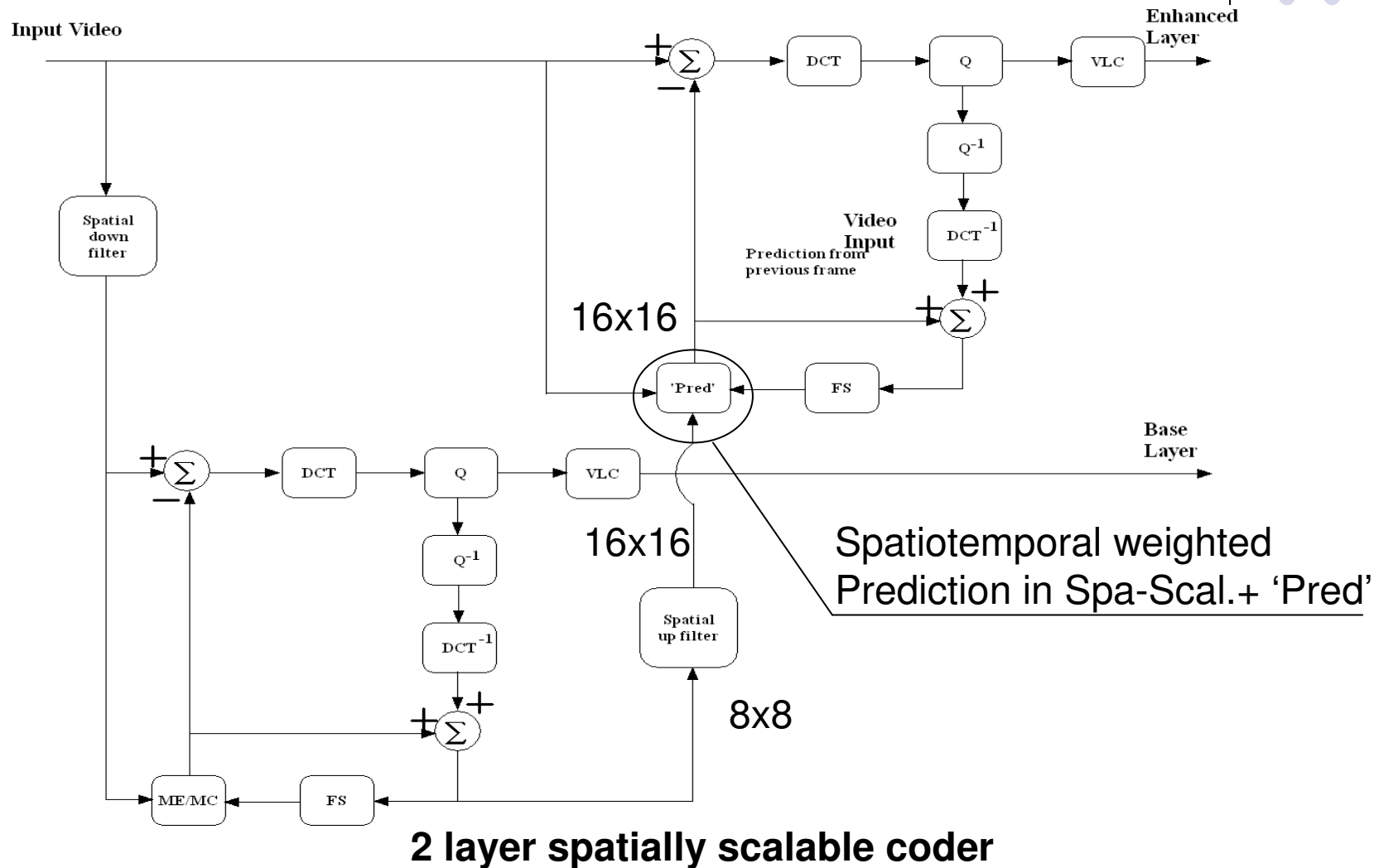
4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



- Spatial Scalability Types
 - Progress to progress
 - Progress to interlaced
 - Interlaced to progress
 - Interlaced to interlaced



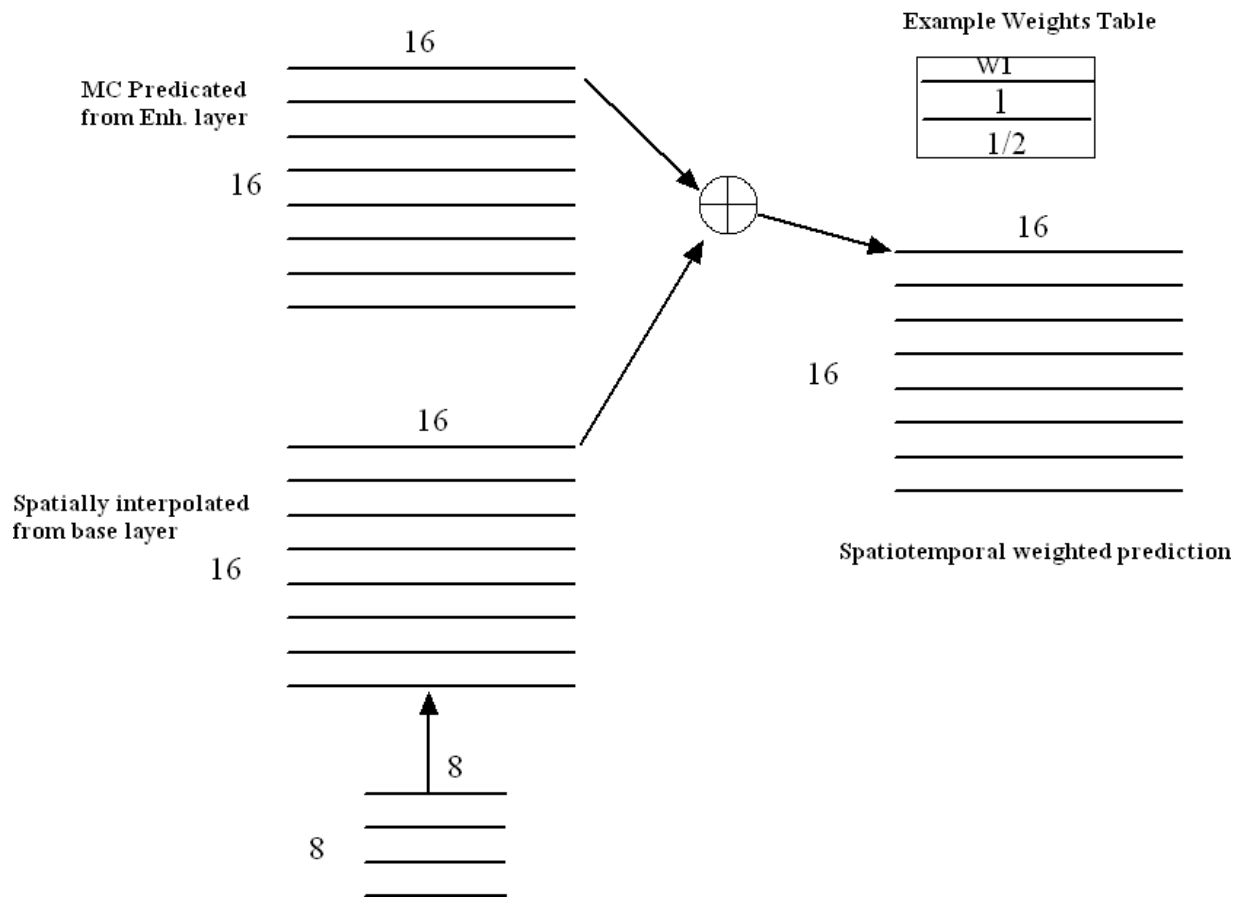
4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



- Spatiotemporal weighted Prediction



4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability

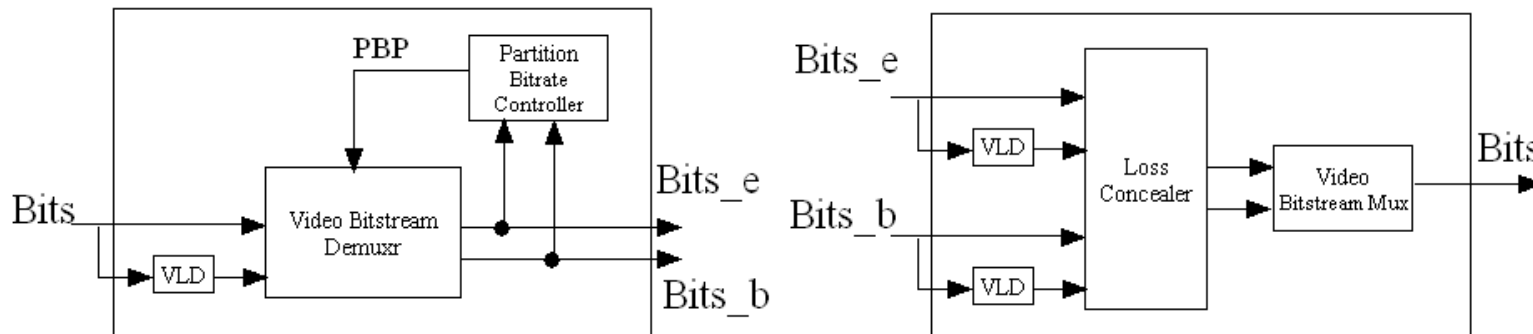
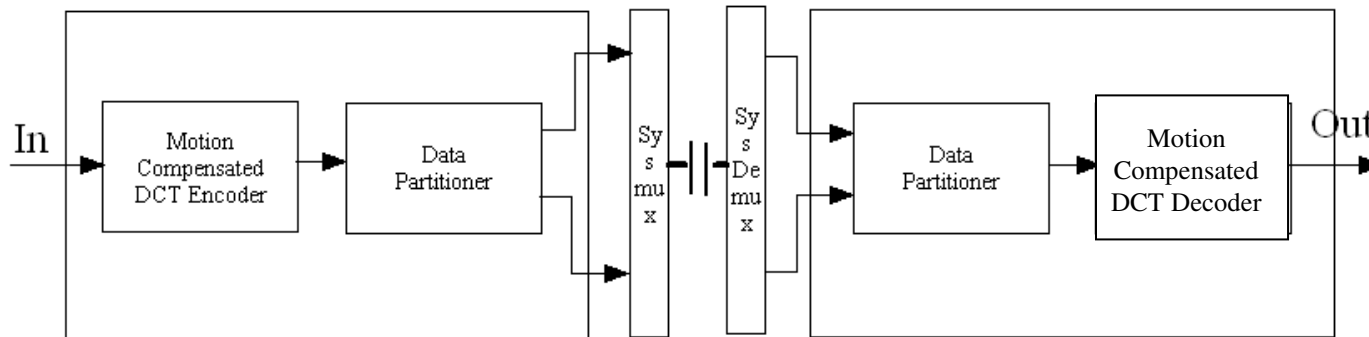


- Data partitioning
 - Data partitioning permits a video bitstream to be divided into two separate bitstreams
 - The BL contains the more info. including address and control info. as well as lower order DCT coef.
 - The HL contains the rest info. of the bitstream
 - The syntax elements in BL are indicated by propriety breakpoint (PBP)
 - Some syntax elements in BL are redundant in HL to facilitate error recovery
 - It has the advantage to introduce almost no additional overhead
 - The disadvantage of this scheme: considerable drift occurs if only the BL is available to a decoder.

4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



- Data partitioning



4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



- Data partitioning

Priority Break Point	Definition
65	All data at sequence, GOP, Pic and slice layers
66	PBP=65 plus MB data to MB type
67	PBP=66 plus data to MB motion Vectors
0	PBP=67 plus MB data from CBP to DC (or 1 st non-zero) Coeff.
1	PBP=0 plus to first coeff. Following DC to first non-zero coeff after the first coeff. in the scan order
2	PBP=0 plus up to first non-zero coeff after the 2 nd coeff in the scan order
j	PBP=0 plus to first non-zero coeff after the jth coeff in the scan order



4.2 MPEG-4 visual standard

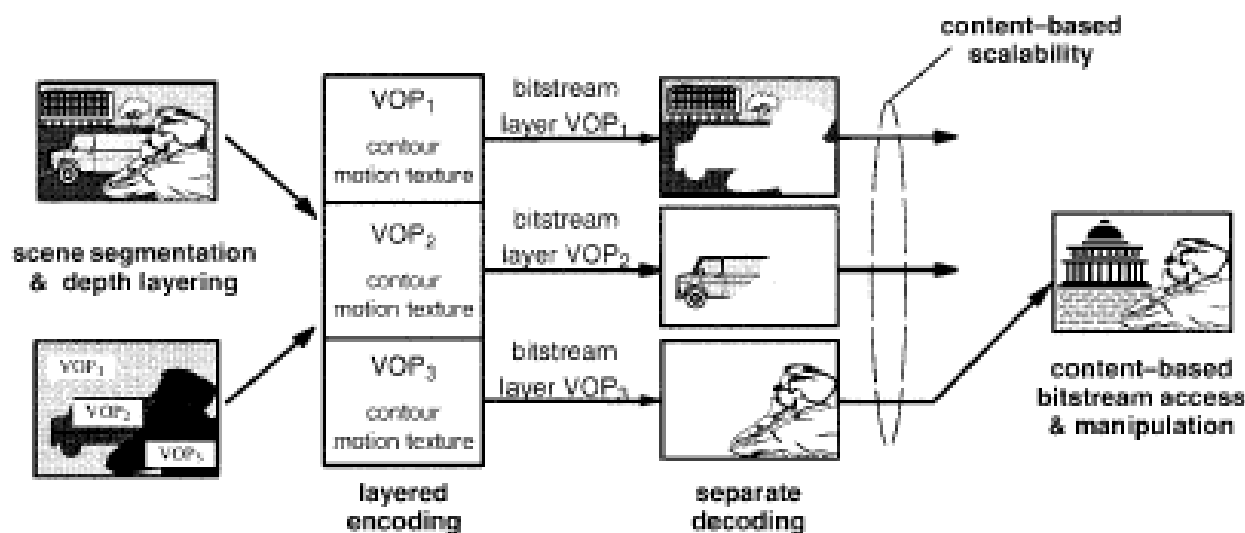
- Video Coding and Communication
 - MPEG-4 standard: video part -- content based video coding scheme
 - To enable all these content-based functionalities, MPEG-4 relies on a revolutionary, content based representation of audiovisual objects.
 - As opposed to classical rectangular video (eg: MPEG1/2), MPEG-4 treats a scene as a composition of several objects that are separately encoded and decoded
 - The scalability at the object or content level enables to distribute the available bit-rate among the objects in the scene
 - Visually, more important objects are allocated more bits.
 - Encoded once and automatically played out at different rates with acceptable quality for the communication environment and bandwidth at hand.



4.2 MPEG-4 Visual Standard

- Access and manipulation of arbitrarily shaped images

Ref: Thomas Sikora



Object Based MPEG-4 Video Verification Model

1. In MPEG-4, scenes are composed of different objects to enable content-based functionalities.
2. Flexible coding of video objects
3. Coding of a “Video Object Plane” (VOP) Layer



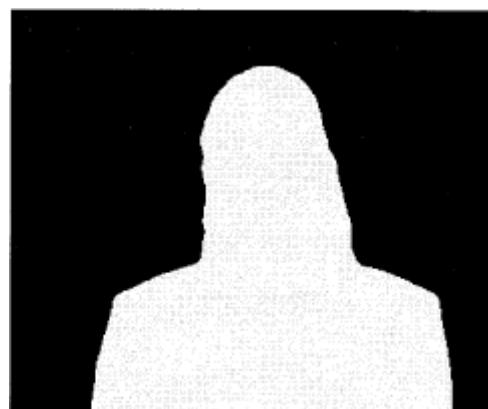
4.2 MPEG-4 Visual Standard

- Video Object Planes (VOP's)

Ref: Thomas Sikora



Original



Binary Segmentation Mask

The binary segmentation Mask is to extract the back/fore-ground layers

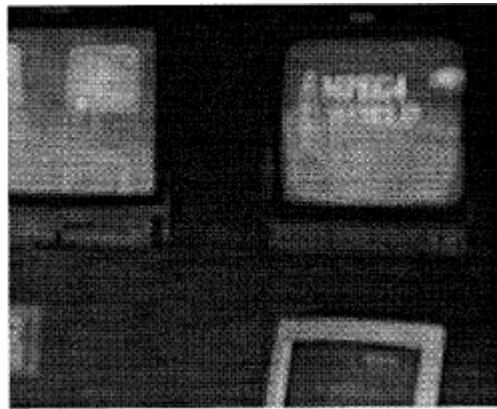
Ref: MPEG-4 AKIYO testing video sequence

4.2 MPEG-4 Visual Standard



- Decomposition into VOP's

Ref: Thomas Sikora



Background Layer VOP



Foreground Layer VOP

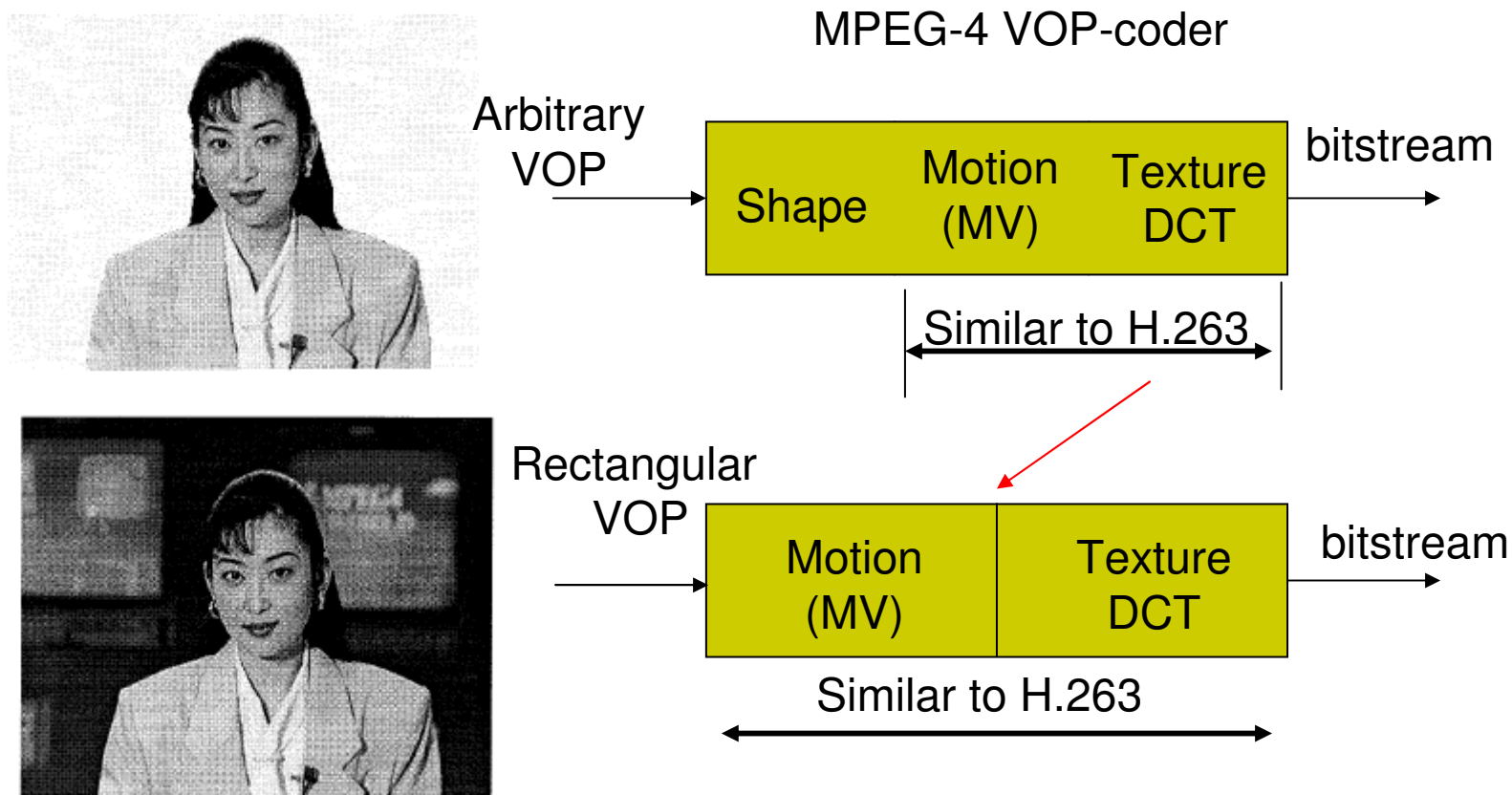
The overlapping VOP's bring the opportunity to do the manipulation of Scene content



4.2 MPEG-4 Visual Standard

- Video Object Plane” layered coding

Ref: Thomas Sikora

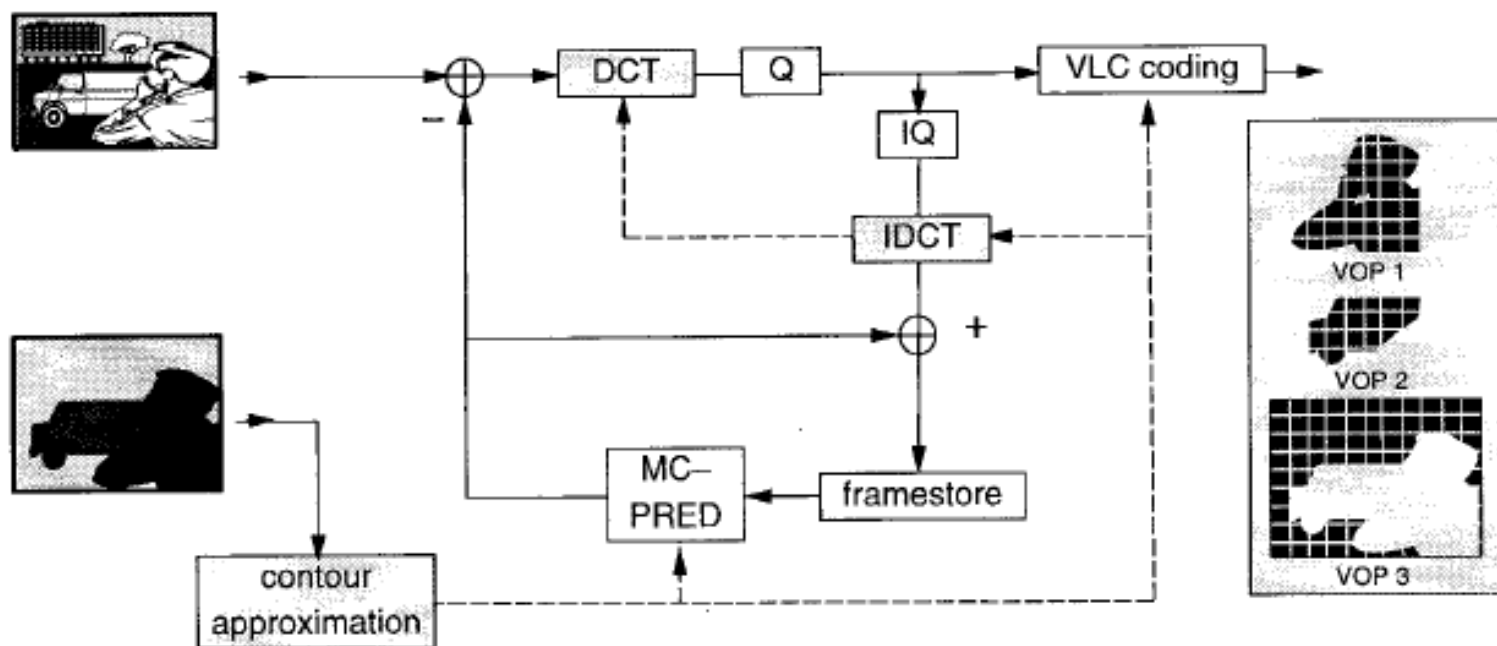




4.2 MPEG-4 Visual Standard

- DCT-Based Approach for Coding VOP's

Ref: Thomas Sikora

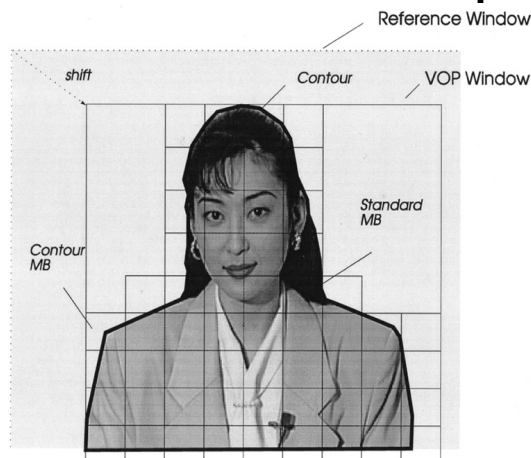


Block diagram of the basic MPEG-4 hybrid DPCM/transform codec structure



4.2 MPEG-4 Visual Standard

- Coding of a “Video Object Plane”



Ref: Thomas Sikora



4.2 MPEG-4 Visual Standard

- Background Padding for Motion Compensation

Ref: Thomas Sikora

Padded background



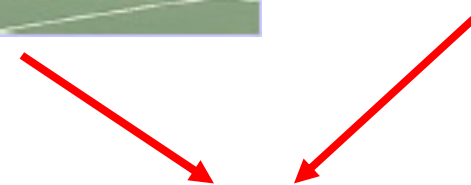
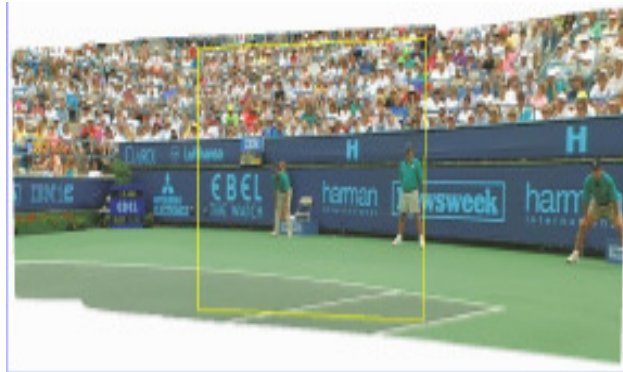
Previous Frame



Current Frame

4.2 MPEG-4 Visual Standard

One Typical Example -- Sprite Coding



1. A non-changing background only has to be transmitted once
2. Only foreground objects transmitted and re-Inserted at the decoder
3. Object are much smaller than full video

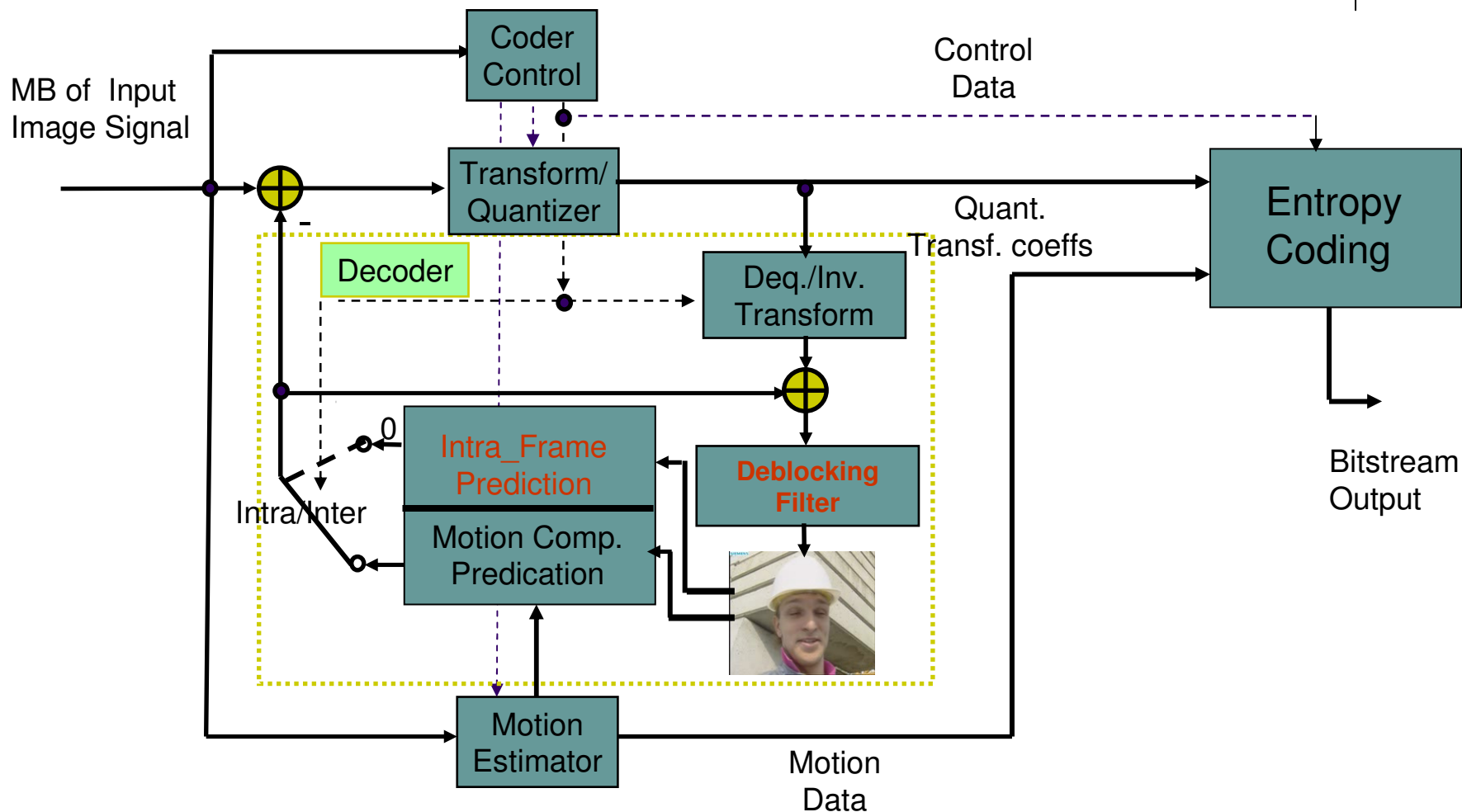
4.3 Introduction to H.264 Video Coding Standard



- It started from the ITU-T H.26L Project (Long term)
- It aims to improve the coding efficiency up to 50% compared to MPEG-4 video coding standard
- In Dec. 2001, MPEG and ITU-T experts set up joint video team (JVT) to focus on this new standard.
- The final version of the standard has been approved by ITU-T 2003. H.264 video coding standard or MPEG-4 Part 10.
- The new technical approaches:
 - An Adaptive deblocking loop filter to remove the artifacts
 - Multiple frame for ME/MC
 - Predication in Intra mode
 - Integer transform
 - Optimized rate control strategy (my opinion)



4.3 Video Codec Structure of H.264

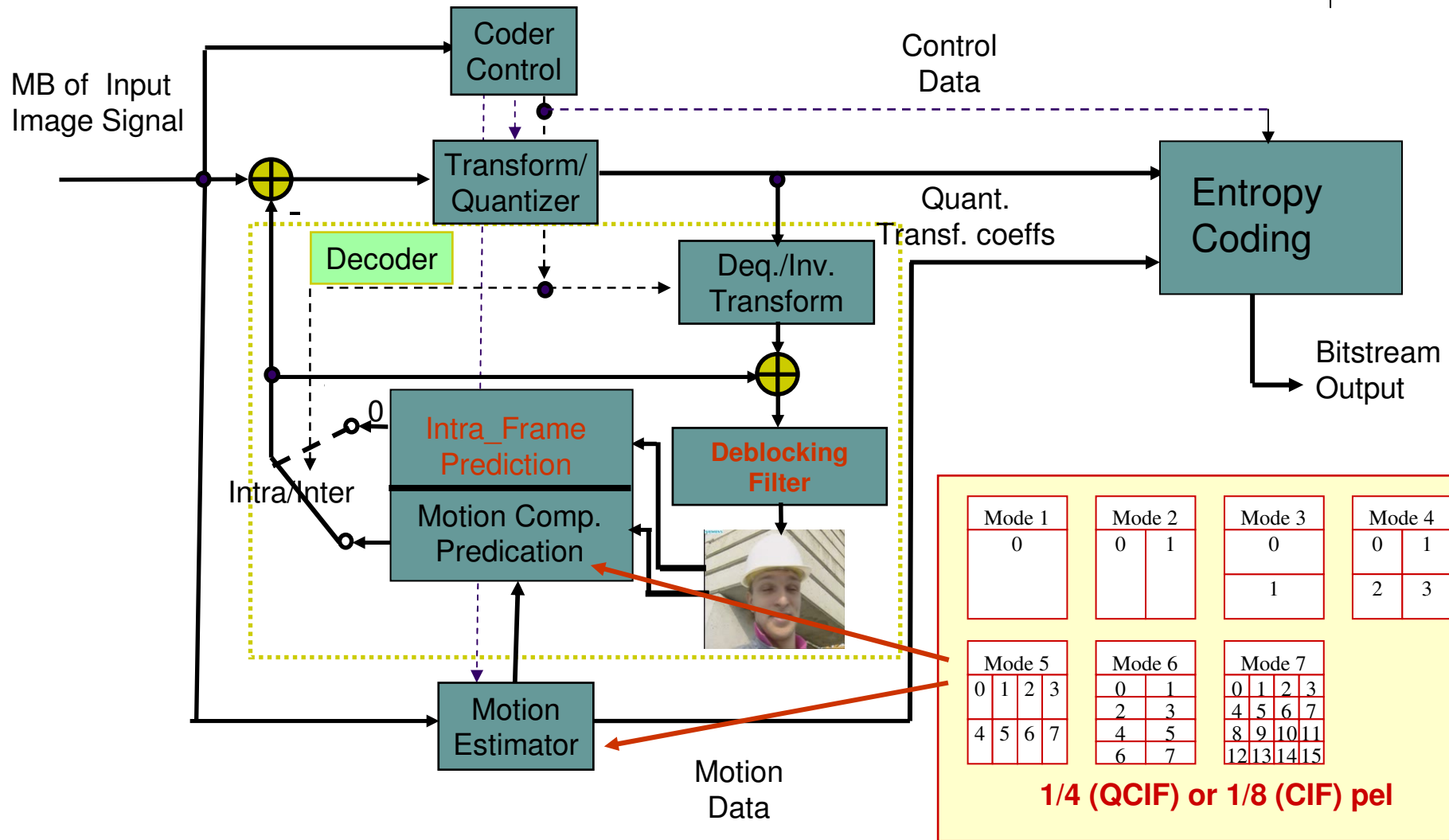


4.3 Video Codec Structure of H.264 (H.26L TML-8 Design Part 1 of 4)



- Hybrid of DPCM/MC/Trans coding as in Prior standards. Common elements include:
 - 16x16 macroblocks
 - Conventional sampling of chrominance and association of luminance and chrominance data
 - Block motion displacement
 - Motion vectors over picture boundaries
 - Variable block-size motion
 - Block transforms (not DCT, wavelets or fractals)
 - Scalar quantization (weighted)

4.3 H.264: Motion Compensation Accuracy





4.3 H.264: Multiple Reference Frames

- Motion Compensation:
 - Multiple reference pictures (per H.263++ Annex U)
 - B picture prediction weighting
 - New “SP” transition pictures for sequence switching
 - Various block sizes and shapes for motion compensation (7 segmentations of the macroblock: 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4)
 - 1/4 sample (sort of per MPEG-4) and 1/8 sample accuracy motion