

# Video Compression Standards (II)

A/Prof. Jian Zhang

NICTA & CSE UNSW  
COMP9519 Multimedia Systems  
S2 2009  
[izhang@cse.unsw.edu.au](mailto:izhang@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 NxN discrete cosine transform matrix C=c(k,n) is defined as:

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

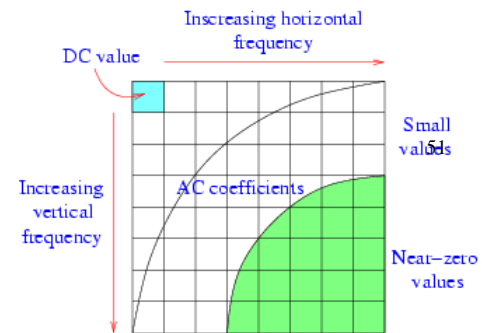


## Basic Transform coding *Tutorial 2*



- The distribution of 2-D DCT Coefficients

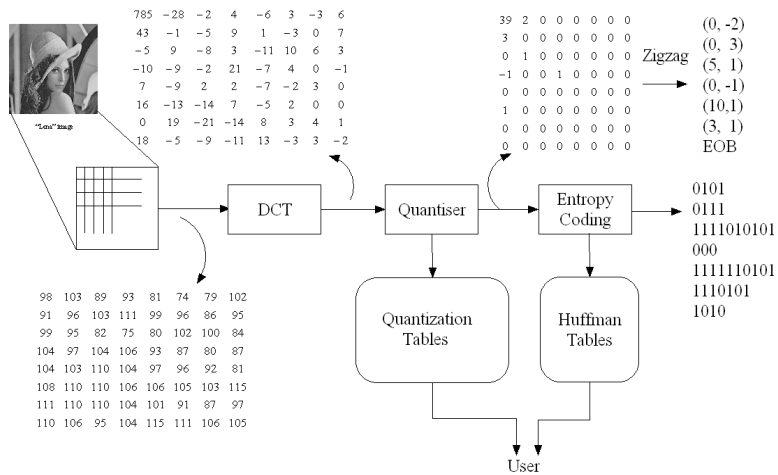
Ref: H. Wu



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

[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

Size: 2

-3	00
-2	01
2	10
3	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*

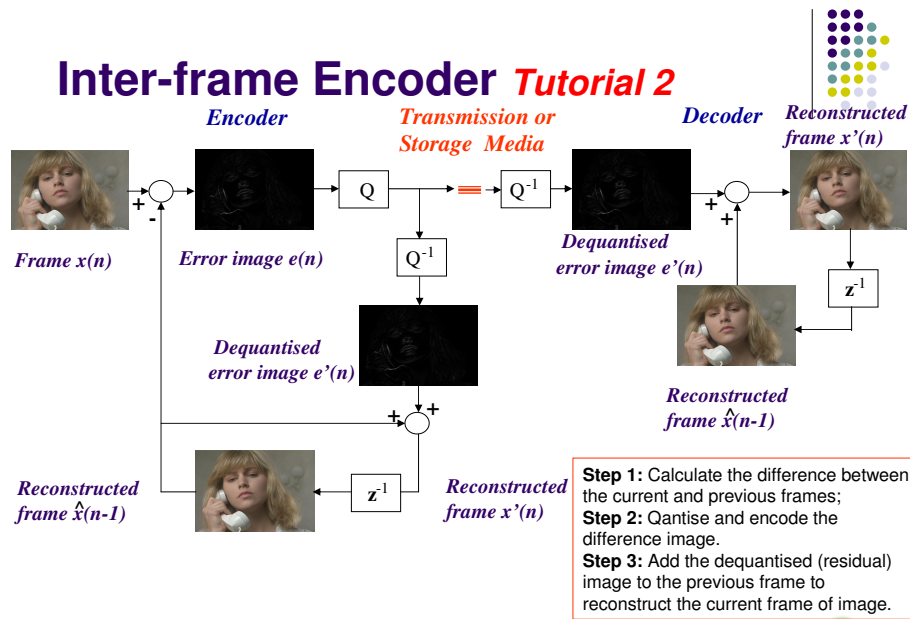
Zig-Zag scan

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	111001
1	3	7	1111001
1	4	9	111110110
2	1	5	11011
2	2	8	11111000
3	1	6	111010
3	2	9	111110111
4	1	6	111011
5	1	7	1111010
6	1	7	111011
7	1	8	11111001
8	1	8	11111010
9	1	9	111111000
10	1	9	111111001
11	1	9	111111010
EOB	4	1010	

Composite values

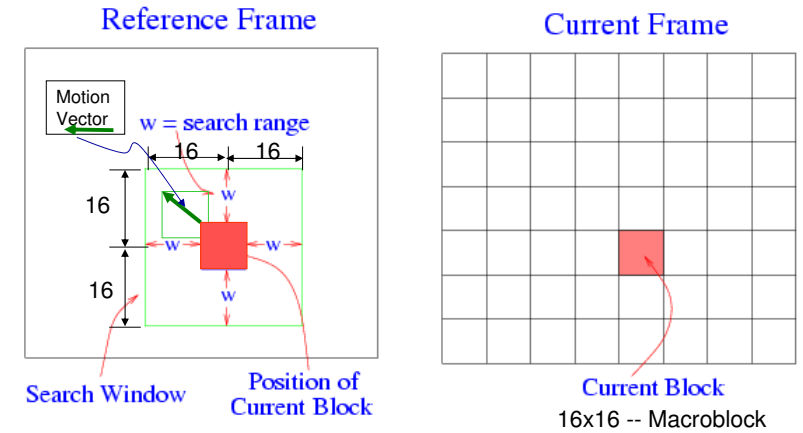
RRRR	SSSS	0	1	2	9	10	11
0	EOB						
.	N/A						
.	N/A						
15	ZRL						

# Inter-frame Encoder *Tutorial 2*



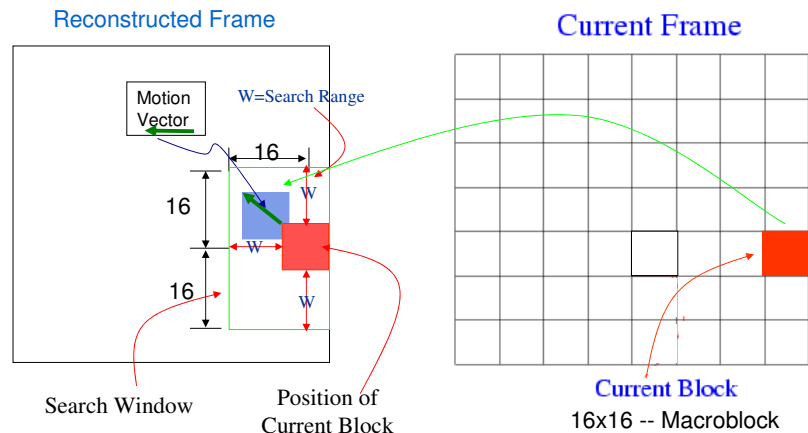
# Block Based Motion Estimation *Tutorial 2*

- Block base search



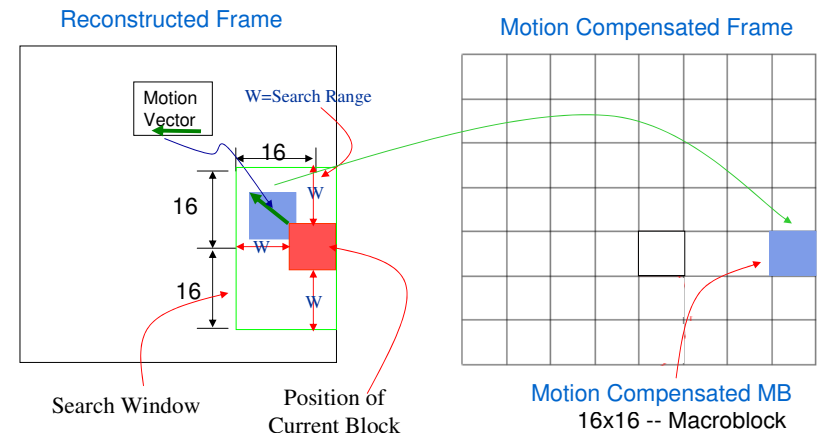
# Block Based Motion Estimation *Tutorial 2*

- Block base search



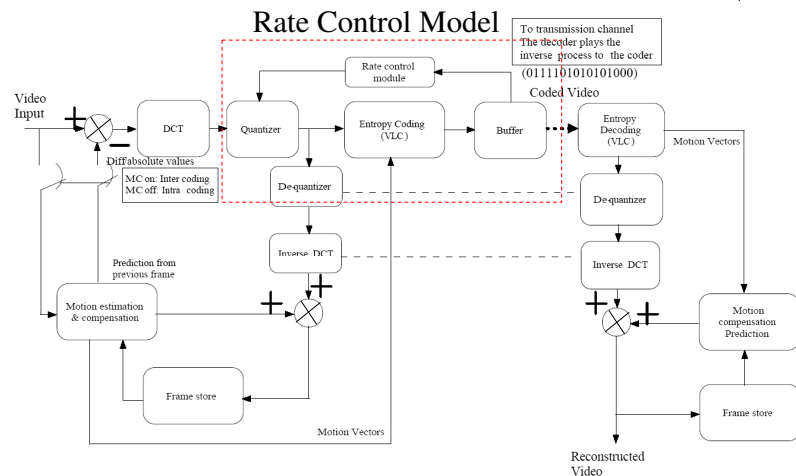
# Block Based Motion Estimation *Tutorial 2*

- Block base search



# Digital Video Coding (DVC) Structure

## – Hybrid MC/DPCM/DCT *Tutorial 2*

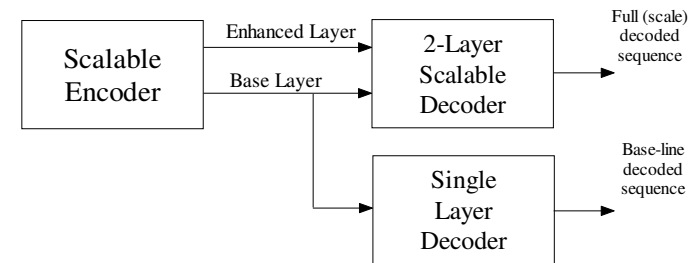


Codec = encoder/decoder

# 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

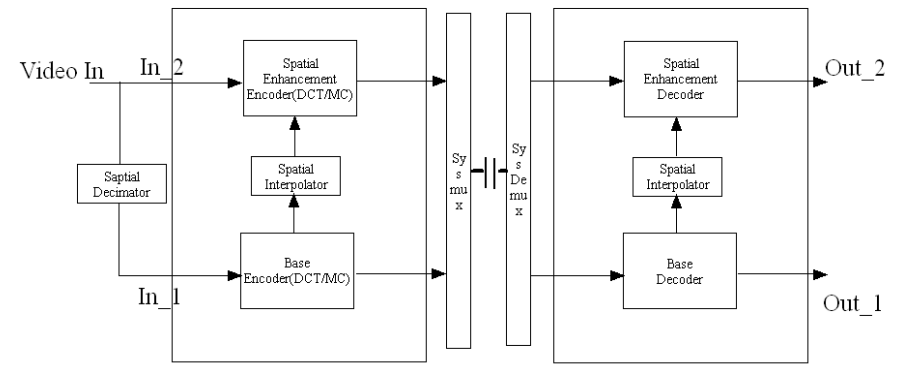


- Spatial Scalability
  - A spatially scalable coder operates by filtering and decimating a video sequence to a smaller size prior to coding.
  - An up-sampled version of this coded base layer representation is then available as a predictor for the enhanced layer
  - As prediction is performed in the 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



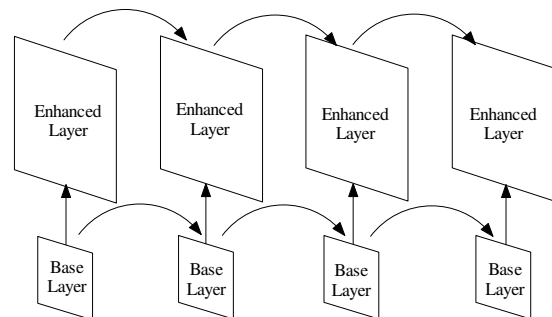
- Spatial Scalability – Spatial Scalability Codec



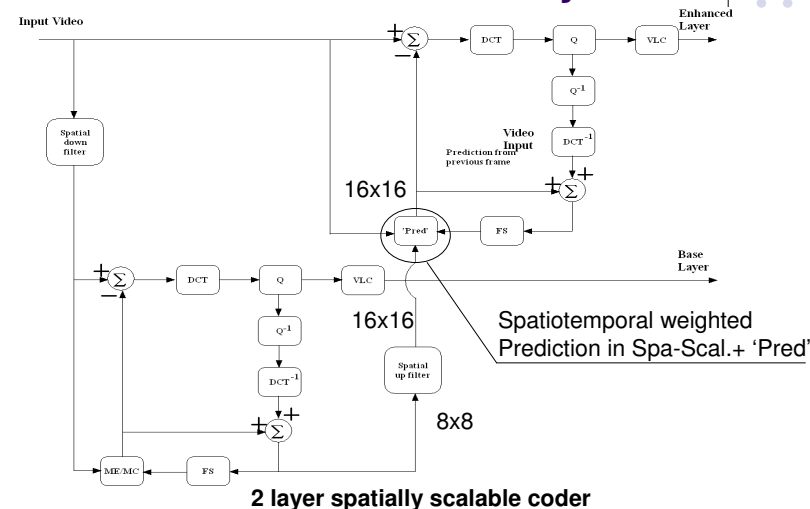
## 4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability

### • Spatial Scalability Types

- Progress to progress
- Progress to interlaced
- Interlaced to progress
- Interlaced to interlaced

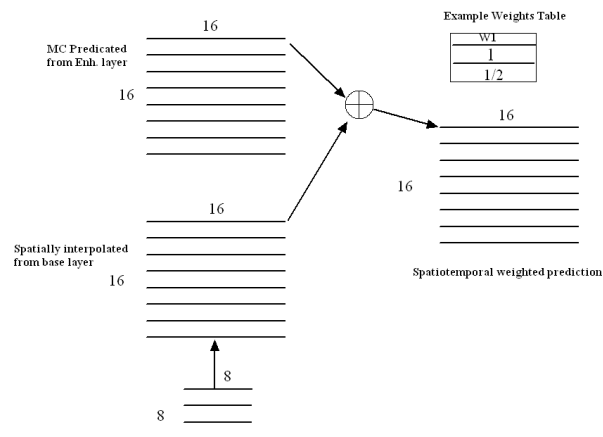


## 4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability



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

### • Spatiotemporal weighted Prediction



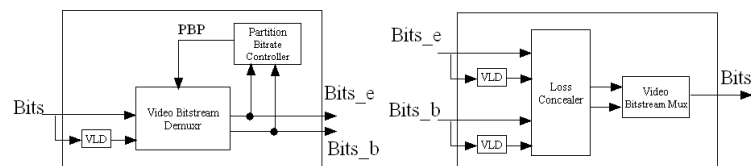
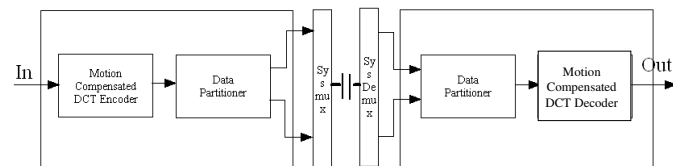
## 4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability

### • 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.

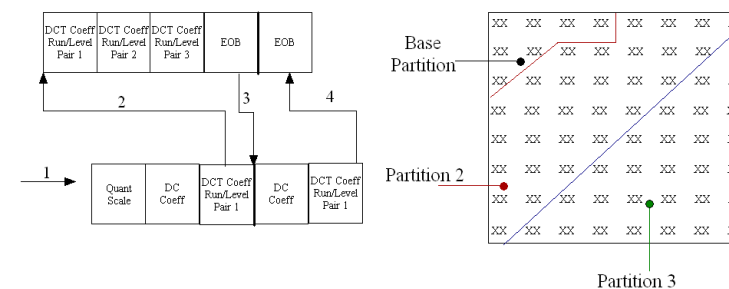
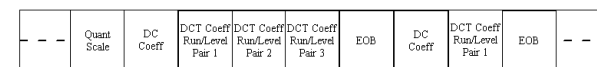
## 4.1 Digital Video Coding (DVC) Standards– MPEG-2 Scalability

- Data partitioning



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

- Data partitioning – bitstream example (PBP = 64)



## 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 <sup>st</sup> 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 <sup>nd</sup> coeff in the scan order
j	PBP=0 plus to first non-zero coeff after the j <sup>th</sup> coeff in the scan order

## 4.2 MPEG-4 visual standard

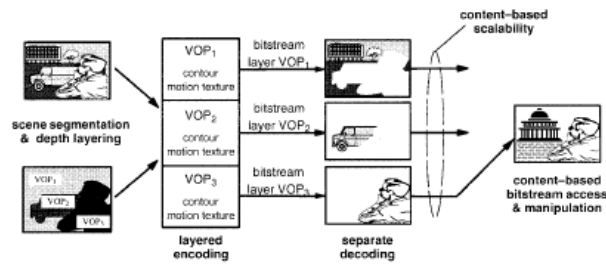
- Video Coding and Communication

- 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

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



COMP9519 Multimedia Systems – Lecture 4 – Slide 25 – J Zhang



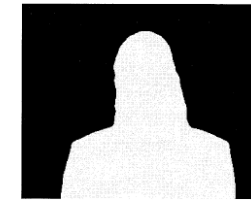
## 4.2 MPEG-4 Visual Standard

- 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



COMP9519 Multimedia Systems – Lecture 4 – Slide 26 – J Zhang



## 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



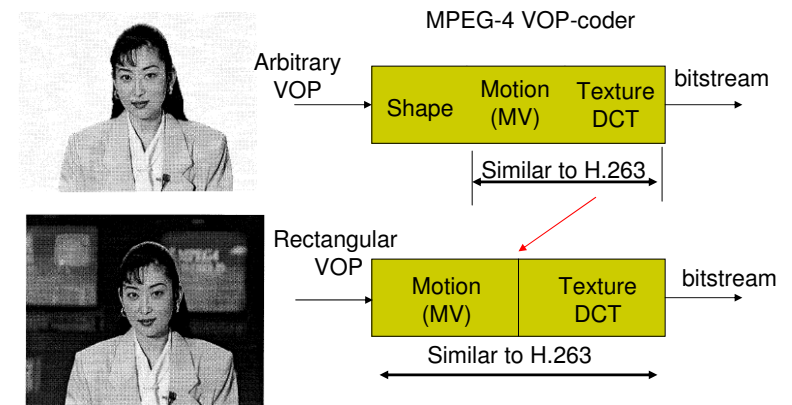
COMP9519 Multimedia Systems – Lecture 4 – Slide 27 – J Zhang



## 4.2 MPEG-4 Visual Standard

- "Video Object Plane" layered coding

Ref: Thomas Sikora



COMP9519 Multimedia Systems – Lecture 4 – Slide 28 – J Zhang

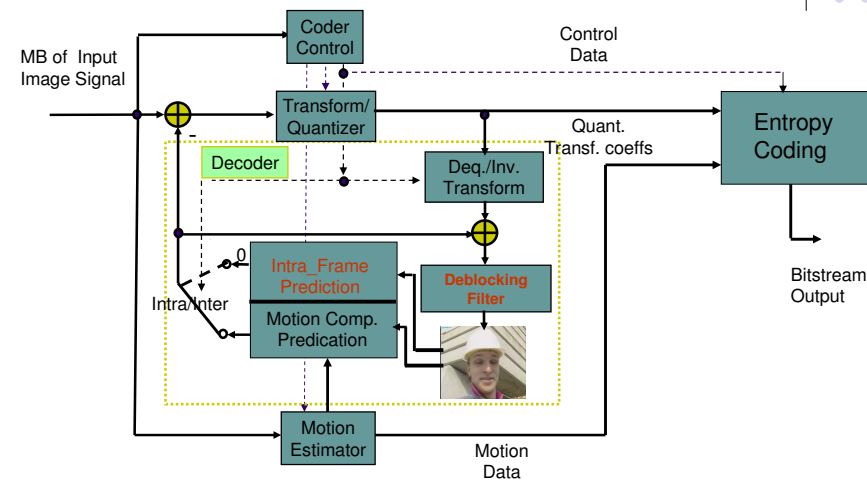




## 4.3 Introduction to H.264 Video Coding Standard

- 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
  - Prediction 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

