Multimedia Video Coding & Architectures (5LSE0), Module 09

Interframe Coding,

MPEG-1/2 Standards: Systems & Video

Peter H.N. de With

(p.h.n.de.with@tue.nl)

slides version 1.0

5LSE0 - Mod 09
Part 1
MPEG Systems Standard

MPEG System / Standardization

* **MPEG = Motion Picture Experts Group**
  - joint standardization of ISO and IEC
  - cooperation with CCITT
* **Objective (initial, MPEG-1)**
  - definition of a generic standard for coding of digital video and associated audio and data for digital storage media (DSM)
* **Objective (MPEG-2)**
  - a generic standard for various applications. such as DSM, television broadcasting, and communication
* **„Generic: wide range of bit rates, variable resolution and quality, flexible for different services**

MPEG System / Applications – (1)

* **Purpose of standardization**
  - improvement of interoperability
  - common technology in SW and HW
  - lower manufacturing costs
  - convergence in AV applications
* **Major applications of MPEG coding**
  - BSS Broadcasting Satellite Service (home use)
  - CATV Cable Television Distribution
  - CDAD Cable Digital Audio Distribution
  - DAB Digital Audio Broadcasting

MPEG System / Applications – (2)

- **Major applications of MPEG coding (continued)**
  - DTTB Digital Terrestrial Television Broadcast
  - EC Electronic Cinema
  - ENG Electronic News Gathering (+ satellite SNG)
  - FSS Fixed Satellite Services (to headends)
  - HTT Home Television Theatre
  - IPC Interpersonal Communications (video phone, conf.)
  - ISM Interactive Storage Media
  - MMM Multi-Media Mailing
  - NDB Networked Database Services (via ATM; etc.)
  - SSM Serial Storage Media (digital VTR, etc.)

MPEG System / Milestones – (1)

- **MPEG-1**
  - Generic coding of moving pictures and associated audio at a throughput rate of up to 1.5 Mbit/s
  - Input is SIF format
  - Related standard H.261 for audiovisual services at px64 kbit/s
  - Final standard described in ISO-11172
  - Applications areas: CD-i, Video CD, video on PCs
- **MPEG-2**
  - More wide application area than MPEG-1
  - Extensions for interlaced video signals (TV, VCR, ...)
  - Bit rates up to 100 Mbit/s

MPEG System / Milestones – (2)

- **MPEG-2 (continued)**
  - final standard described in ISO 11383
  - different descriptions for Audio, Systems
  - draft of standard in November 1993, ratification in progress
- **MPEG-4**
  - advanced extensions of MPEG-2 with respect to block coding
  - New: **model-based** or **object-oriented** coding
  - very low bit rates (e.g. 10 kbit/s ... 100 kbit/s), draft in 1998-99
- **MPEG-7**
  - for archiving of video sequences
  - database management, standard 2001

MPEG System / Application bit rates

- **MPEG-1 bit rates**
  - Video decoder rates up to 1.856 Mbit/s (telecomm. channel 31 x 64 kbit/s = 1.984 Mbit/s - 0.128 Mbit/s audio
  - CD-i, Video CD have bit rates of 1.2 Mbit/s (audio about 200 kbit/s)
- **MPEG-2 bit rates**
  - 4-5 Mbit/s PAL TV quality
  - 6-9 Mbit/s CCIR-601 component video quality (approach studio quality)
  - 19 Mbit/s ATV standard for HDTV in the USA
  - 20-40 Mbit/s for HDTV contribution (studio to studio)

9

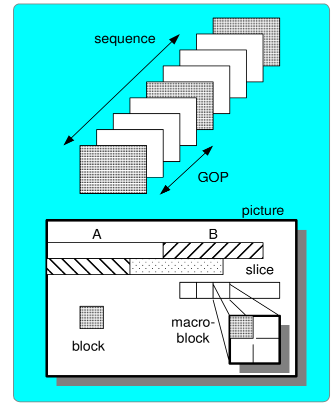
MPEG System / Structured data –(1)

* **Structure of MPEG formatted data, applied codecs**
* **System**
  – Multiplexing, packetizing of multiple compressed data streams
  – Synchronization and timing of individual data contributions
* **Video coding using hybrid compression**
  – Motion compensation in the temporal domain
  – DCT coding in the spatial domain (in the image)
* **Audio coding**
  – Subband coding at 64, 128, or 192 kbit/s
  – Audio subband masking of inaudible components

10

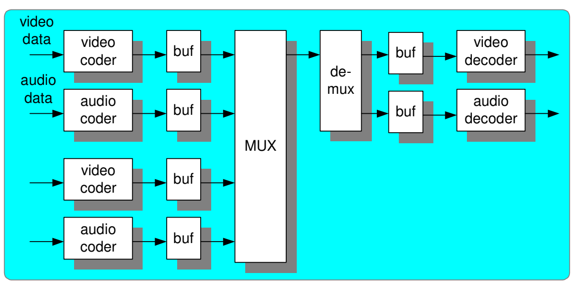
MPEG System / Structured data –(2)

* **MPEG bit stream Structure**
* Division in layers
  – Sequence layer
  – Group Of Pictures (GOP)
  – Picture layer
  – Slice layer
  – Macroblock layer
  – Block layer



11

MPEG System / Block diagram



* **MPEG System block diagram**
  – timing information of synchronous video and audio
  – timing of multiple MPEG-formatted data streams

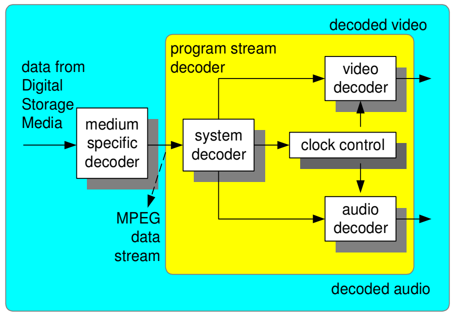
12

MPEG System / Program stream

* **MPEG Program stream**
  – Combines one or more **Packetized Elementary Streams** (PES) into one single multiplexed stream.
  – All streams have a ***common time base***.
* **Requirements Program Stream**
  – Define ***multiplex*** for audio and video streams
  – ***Control buffering*** of data (over- and underflow)
  – Enable ***start of decoding*** after random access
  – Supply ***timing information***
  – Low overhead rate

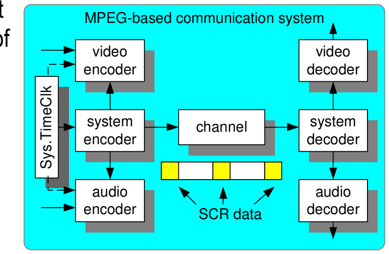
MPEG System / Decoder program stream

- Typical Decoder Architecture for MPEG progr. stream
- Operate modes as
  - **MULTIPLEX-WIDE**: the program itself in the pack layer
  - **STREAM-SPECIFIC**: one of the elementary streams, in the PES packet layer



MPEG System/ Timing system level –(1)

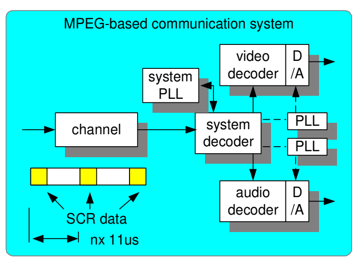
- Timing model
  - End-to-End delay from input of encoder to signal output of decoder is constant
  - All samples are presented only once
- Transport Mechanism
  - System Clock Reference (SCR) is sent to decoder
  - SCR specifies intended time at which SCR is entered in the system decoder



- All time units expressed in terms of common System Time Clock (STC)

MPEG System / Timing system level –(2)

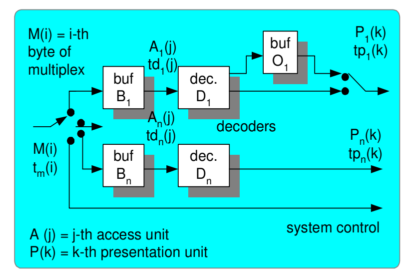
- Accuracy
  - Units of 90 kHz or 11 microsec.
  - Repetition intervals no longer than 700 ms (stable control)
- System remarks
  - For **transmission** systems: PLLs are required to recover master clock for D/A conversions



-- For **recording sytems**: stand-alone video and audio clock for D/A

MPEG System / System Target decoder

- STD is a model!
  - Describes timing and buffering of decoder exactly
  - Parameterized in MPEG (-1 or -2) fields
  - Encoder responsible that STD can decode all
- Physical decoder
  - Must compensate for its differences with the STD

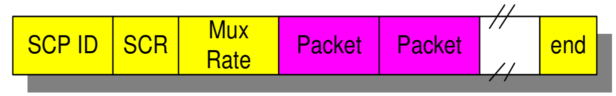


td(j) = DTS decoding time stamp
tp(j) = PTS presentation time stamp

MPEG System / Synchronization

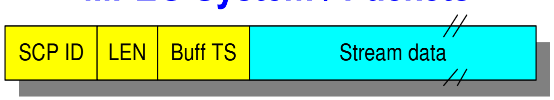
* **Presentation Time Stamps enable synchronization in program stream**
  - **PTS** apply to the **presentation time of compression** layer constructs (video packet can start anywhere in the bit stream
  - **E-to-E synchronization** occurs when
    1. encoder saves time stamps at capture time
    2. PTS propagate with corresponding data to the decoders
    3. decoder uses the PTS to start presentation
* **PTS is sent to the decoder**
* Accuracy
  - units of 90 kHz
  - SCR+PCR extensions with resolution of 27 MHz (MPEG-2)

MPEG System / Packs (Program Str.)



* **Data units of Pack layer (MPEG-1)**
  - SCP Start Code Prefix (unique code in AV stream @ system level)
  - ID Pack Identifier
  - SCR System Clock Reference, 33-bit counter, incr. at 90 kHz
  - MUX RATE multiplied bit rate in units 50 Bytes/s, 22-bit field which can vary each pack (VBR support). Often not req. by receiver

MPEG System / Packets



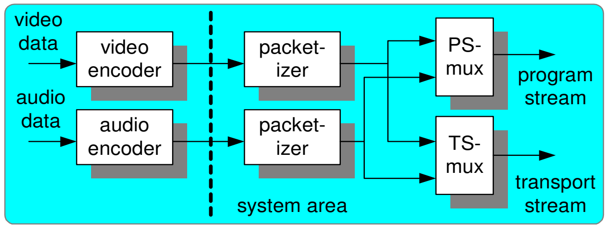
* **Data units of Packet layer (MPEG-1)**
  - Packet has only data of one input stream, chronological order
  - Total number of packets per pack is not defined
  - ID Stream ID, 110X XXXX Audio (32 available), 1110 XXXX Video (16 available), 1111 XXXX Data stream (16 available)
  - LEN, Distance to start code next packet in Bytes (16 bits)
  - BUFF TS, System Target Decoder buffer size information
  - STREAM DATA, Data for Video, Audio, data decoder, or specific data, (may include PTS and/or decoding time stamps)

MPEG System / MPEG-2 System – (1)

New system specification**: based on MPEG-1, but extended**

* **New requirements**
  - Fixed-length data packets option for error-sensitive media
  - Super MUX: program or channel MUX enabling more programs
  - Lower multiplexing complexity
* Solution: **TWO STREAMS** for different applications
  1. **Program Stream**
     - Special AV programs,
     - Common time base for error-free environments (variable packets)
  2. **Transport stream**
     - Multi-program broadcast, program storage, allows editing
     - One or more distinct time bases
     - Fixed-length packets of 188 Bytes for error-sensitive systems

MPEG Sys./ MPEG-2 system diagram



- Distinction between program stream and transport stream
- Packetizing formats optimal for different environments

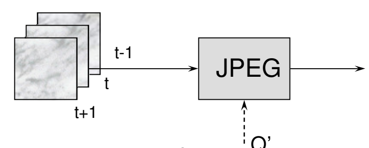
MPEG System / MPEG-2 Transport Str.

- **Transport Stream (TS) definition**
  - One or more programs, each containing one or more elementary streams multiplexed together
  - Fixed-length 188 Byte packets, with usually incl. 4-Byte header (Sync, PID, control-bit parameters, etc.)
- **Transport stream extraction possibilities (examples)**
  - Select data one program, decode, and presentation of results
  - Extract TS packets of one or more programs from multiple TS and construct new TS
  - Extract 1 program and constitute new program stream (DVD record)
  - Convert program stream to TS, transport it over error-sensitive media, and recover valid program stream (DVD+ DVB broadcast)
- **Conclusion**: **important for broadcasting applications**

5LSE0 - Mod 09
Part 1
Interframe: Hybrid Video Coding

Intraframe Coding

- **Encode frame-by-frame, disregarding all temporal information**
  - Example: Motion-JPEG (AVI compressed)

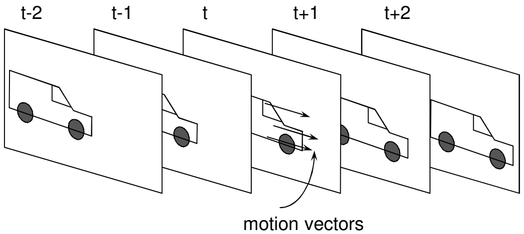


  - Easy bit allocation per frame
  - Random access is possible
  - Robust to transmission/decompression problems
  - **But** … Moderate compression capabilities

Interframe Coding of temporal differences

25

- **Encode differences between frames (temporal DPCM); consider *motion* of parts of the frames**



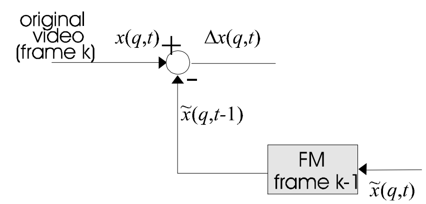
Complexity is between intra-frame and 3-D coding

- Can always fall back onto intra-frame coding
- Moderate delay

Principles of Hybrid Coding – (1)

26

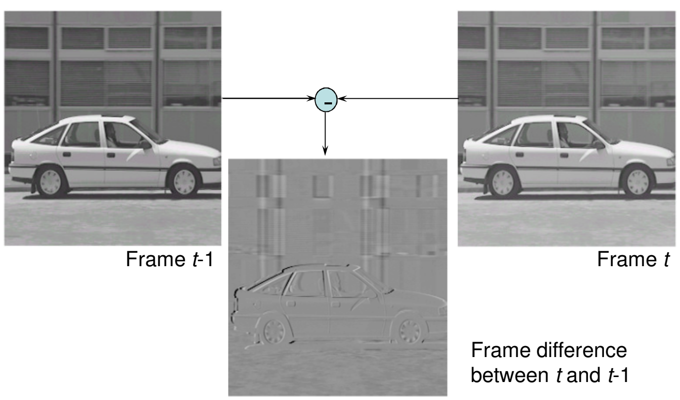
- **Basic idea**
  - Predict current frame on basis of (coded) previous one
  - Transmit only quantized prediction differences
  - Usually done on 8x8 blocks



- Prediction difference: $\Delta x(q,t) = x(q,t) - \tilde{x}(q,t-1)$

Example Frame Differences

27



Frame *t*-1

Frame *t*

Frame difference between *t* and *t*-1

Temporal Prediction Gain

28

- **Like normal DPCM, assess the effect of interframe prediction by the prediction gain**

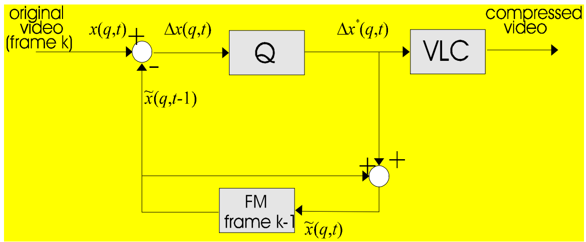
$$G_P = \frac{\text{variance of the original frame}}{\text{variance of the prediction difference}}$$

Frame number	σ_x^2	$\sigma_{\Delta x}^2$	G_P
1	1888.2	282.0	6.7
2	1885.9	225.7	8.4
3	1873.6	265.7	7.1
4	1884.6	329.1	5.7
5	1889.4	342.6	5.5
6	1901.1	368.9	5.2

around 1.2 - 1.5 bit

Principles of Hybrid Coding – (2)

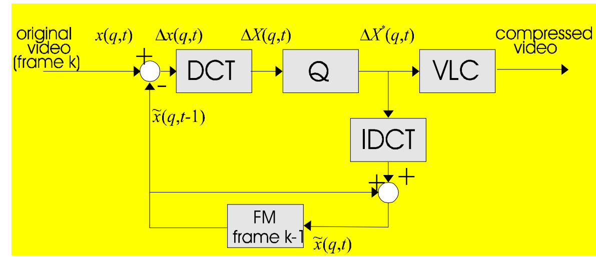
* Complete temp. DPCM system (1-st ord. predictor, h_1=1)



* Quantized prediction difference: $\Delta x^*(q,t) = Q[\Delta x(q,t)]$
* Reconstruction: $\tilde{x}(q,t) = \Delta x^*(q,t) + \tilde{x}(q,t-1)$

Principles of Hybrid Coding – (3)

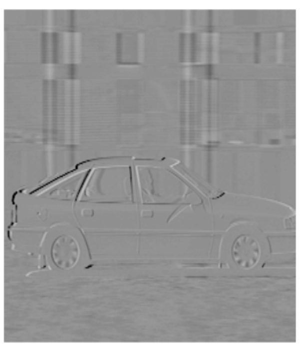
* **Prediction difference *locally*** still contains a lot of *spatial* correlation and lots of zeroes: **Decorrelate via 8x8 DCT**



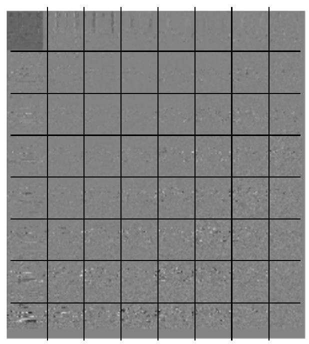
* Quantization in DCT domain $\Delta X^*(q,t) = Q[\Delta X(q,t)]$
* Reconstruction $\tilde{x}(q,t) = \mathrm{DCT}^{-1}(\Delta X^*(q,t)) + \tilde{x}(q,t-1)$

Example: DCT of differences

Frame difference



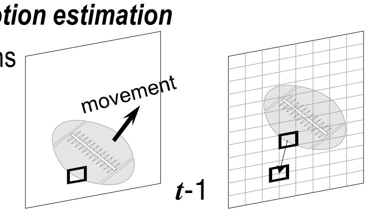
DCT of frame difference in bands



- Higher DCT coefficients contain more variance than lower ones
- Taking frame differences removes spatial low-frequency components

Motion Compensation

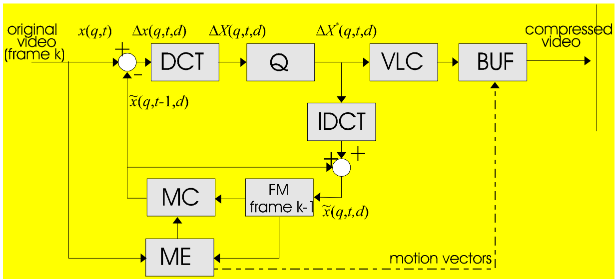
* **Part of the spatial correlation is due to unsuccessful temporal prediction**
  - Unpredictable information (occluded regions)
  - Moving spatial information (object movement)
* **Find for each block $x(q,t)$ a corresponding block in encoded frame t-1: *Motion estimation***
  - Difference in positions is called ***motion*** or ***displacement*** vector



Principles of Hybrid Coding – (4)

33

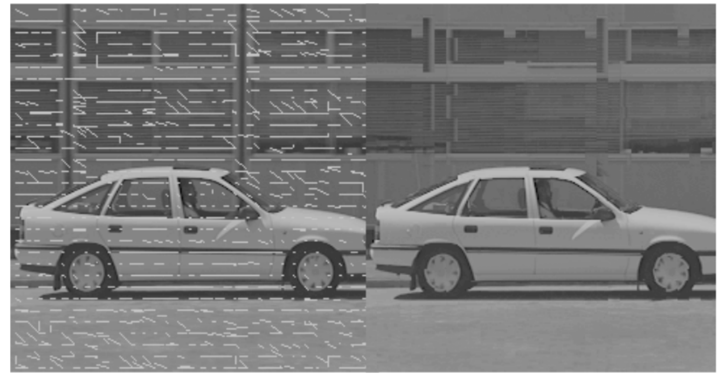
- Form difference between $x(q,t)$ and the corresponding block found in encoded frame t-1: *Motion-Comp. Predict.*



- Motion-compensated prediction difference
- Overhead: 1 motion v. / block $\Delta x(q,t,d) = x(q,t) - \tilde{x}(q,t-1,d)$

Example – Motion-Compens. Prediction

34



Frame with motion vectors — Motion-compensated prediction

Example Motion-Compens. Predict.

35

Frame difference — Motion-compensated frame difference

Motion Comp. / Prediction Gain

36

		No motion compensation		Motion compensated prediction	
Frame number	σ_x^2	$\sigma_{\Delta x}^2$	G_P	$\sigma_{\Delta x}^2$	G_P
1	1888.2	282.0	6.7	73.5	25.7
2	1885.9	225.7	8.4	86.8	21.7
3	1873.6	265.7	7.1	89.5	20.9
4	1884.6	329.1	5.7	91.9	20.5
5	1889.4	342.6	5.5	96.0	19.7
6	1901.1	368.9	5.2	99.5	19.1

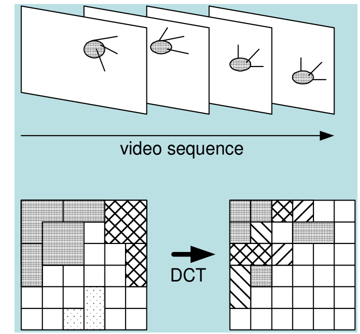
around 0.2 - 0.4 bit

5LSE0 - Mod 09
Part 3

MPEG Video Coding: Architecture and Properties

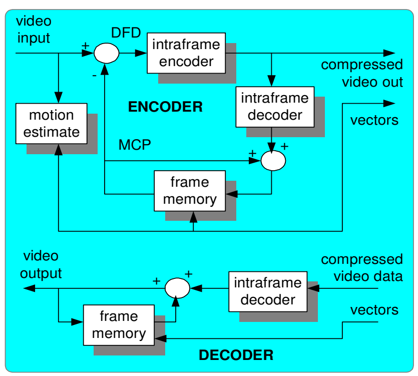
MPEG Video / Principles

- **MPEG video exploits both** spatial **and** temporal **redundancy**
- Temporal **redundancy**
  - Motion estimation
  - Motion compensation
- Spatial **redundancy**
  - Block transformation
  - Variable-length coding



MPEG Video / Hybrid coder architect.

- **MPEG hybrid coding uses**
  - Motion estimation in temporal encoder
  - Intraframe transform coding in spatial encoder/decoder



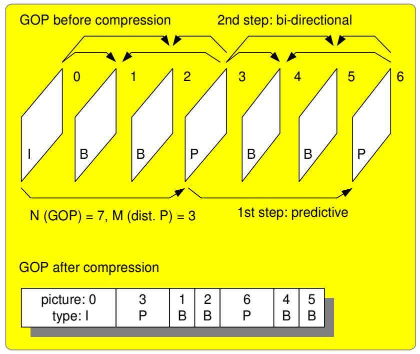
MPEG Video / Picture types: I, P, B

- **Intraframe (I) coded pictures**
  - serve as starting point for a group of pictures
  - modest compression
  - reference picture for other picture types
- **Predictive (P) coded pictures**
  - coded with reference of past pictures (I or P)
  - reference of remaining B-pictures
- **Bi-directional (B) coded pictures**
  - coded with reference of past and future pictures (I,P and P)
  - never used as a reference
  - highest compression
- DC (D) intra-coded pictures (not used, spec. modes)

41

MPEG Video / Group Of Pictures (GOP)

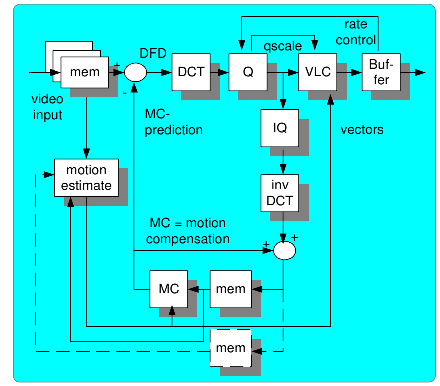
- **GOP in normal order at input**
  - Step 1: code next P picture
  - Step 2: code intermediate B picts.
- **GOP after compression**
  - Non-chronological order
  - Unequal bit size



42

MPEG Video / Predict, Bidirectional

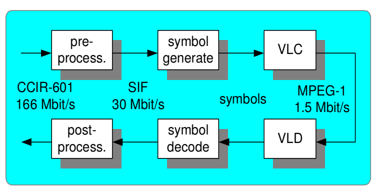
- **MPEG encoder block diagram**
- **ME process requires**
  - Reorder memories
  - ME unit
  - Two-sided extensions
- **VLC at output**
  - Rate-control buffer
  - *qscale* parameter coding



43

MPEG Video / MPEG-1 sampling – (1)

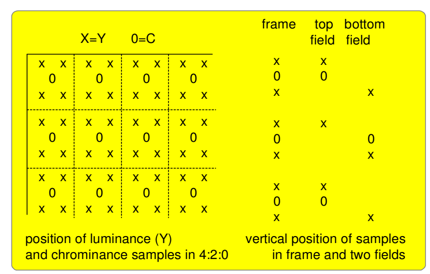
- **MPEG-1 resampling is required**
  - target 1.5 Mbit/s too low
  - factor of 5 compression by SIF at 30 Mbit/s
- **CCIR-601 (4:2:2 sampling)**
  - Y : 720x576 frame, 2:1 interlace, 50 Hz
  - U,V : 360x576 frame, id.
- **SIF („2:1:0“ sampling)**
  - Y : 352x288 lines
  - U,V: 176x144 lines
  - 25 frame/s
  - 1:1 progressive



44

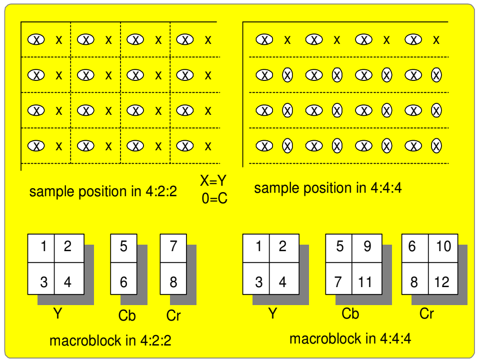
MPEG Video / MPEG-1 sampling – (2)

- **MPEG sampling, mainly on 4:2:0**
  - MPEG-1: line sequential on frame basis
  - MPEG-2: line sequential also with interlaced fields



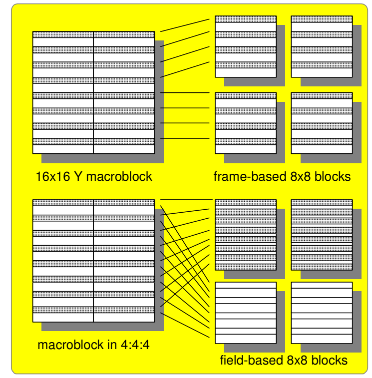
MPEG Video / MPEG-2 sampling

* **Alternative sampling in MPEG-2**
  – 4:2:2, The CCIR-601 studio standard
  – 4:4:4, for High-quality RGB applications



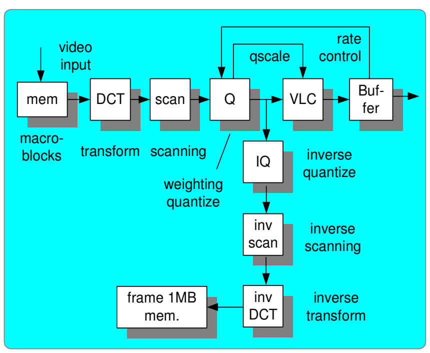
MPEG Vid./ MPEG-2 motion adaptivity

* **MPEG-2 accepts interlacing, thus requires motion-adaptive DCT**
  – Keep fixed 8x8 blocks in both modes
  – Static 8x8 **frame-based** and moving 8x8 **field-based** blocks



MPEG Video / Intraframe coding part

* **MPEG intraframe coder / decoder block diagram**
  – Local encoding
  – Reconstruction for motion compensation
  – Based 8x8 DCT, adapt. Quantization and 2-D VLC
  – Feedback coding with rate buffer control



MPEG Video / Quantization – (1)

MPEG Quantization for inter- / intraframe data

* **DC coefficient**
  – Human eye very sensitive for DC errors, thus **fixed** quantizer
  – MPEG-1*: DC**Q**=DC / 8.0* and inverse *DC = 8.0xDC**Q***
  – MPEG-2: higher DC precision 8-11 bits (*n x DC**Q***)
* **AC coefficients**
  – Weighting $W(u,v)$ according to perception

intra block weighting

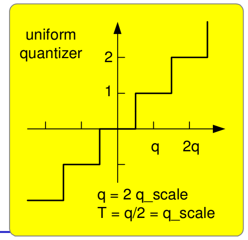
08	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

inter block weighting

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

MPEG Video / Quantization – (2)

- **AC coefficients (cont.)**
  - **MPEG-1 encoder** formula
    $FQ(u,v) = 16\ F(u,v) \,/\, (2\ q\_scale\ W(u,v))$
  - **MPEG-1 decoder** formula
    $F(u,v) = 2\ (FQ(u,v) + k)\ q\_scale\ W(u,v)\ \ /\ 16$
  - k = 0 for intrablocks, and
    $k = sign(FQ(u,v))$ for non-intra blocks
  - **Mismatch control** (value closest to zero):
    if $F(u,v)$ even, then
    $F(u,v) = F(u,v) - sign(\ F(u,v)\ )$



MPEG Video / Quantization-(3) MPEG-2

- **MPEG-2 has more precise quantization**
- **DC coefficients:** up to 11 bits precision
- **AC coefficients**
  - **MPEG-2 decoder** formula
    $F(u,v) = 2\ (FQ(u,v) + k)\ q\_scale\ W(w,u,v)\ \ /\ 32$
  - q_scale is mapped onto larger range than 0...31
  - w is defined by intra / non-intra and colour sampling
  - k = 0 for intrablocks, and $k = sign(FQ(u,v))$ for non-intra blocks
  - special additional mismatch control: $F(7,7) = F(7,7)$ if SUM $ac(F(u,v))$ is odd, and $F(7,7) = F(7,7)$ +/- 1 if $F(7,7)$ is even/odd and SUM is even.

MPEG Video / Quantization-(4) MPEG-2

MPEG-2 AC coefficients (cont.)

- **Extra adaptivity possibilities**
  - Quantizer matrix $W(w,u,v)$ can be reloaded in frame header, giving **adaptive weighting** on sequence or application
  - q_scale Parameter can be modified on macroblock basis, enables smooth regulation of bit rate locally in the image
  - In any case, MPEG-2 different weighting for Y and C

MPEG Video / Quantization – (5)

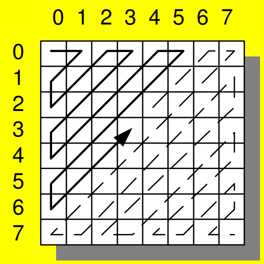
MPEG-2 AC coefficients

- Larger range of q_scale by mapping of transmit code
  - Two characteristics: uniform and non-uniform
  - Non-uniform curve enables different control for low bit-rates

q_scale	quantize type	
	type=0	type=1
0	forbidden	
1	2	1
2	4	2
3	6	3
4	8	4
5	10	5
6	12	6
7	14	7
8	16	8
9	18	10
10	20	12
11	22	14
12	24	16
13	26	18
14	28	20
15	30	22
16	32	24
17	34	28
18	36	32
19	38	36
20	40	40
21	42	44
22	44	48
23	46	52
24	48	56
25	50	64
..	...	...
31	62	112

MPEG Video / VLC Scanning – (1)

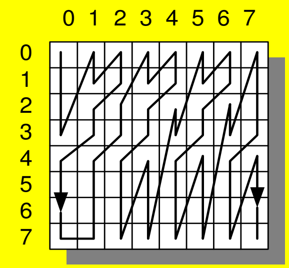
* **Scanning of transform coefficients (MPEG-1/2)**
  – Preprocessing step for variable-length coding
  – Scanning functions reorders coefficients to cluster zeros for runlength coding
  – Start with „low-frequency“ coefficients
  – Fundamental scanning pattern is diagonal



zigzag scanning

MPEG Video / VLC Scanning – (2)

* **MPEG-2 extension of scanning function**
  – Picture header extension: indicate the use of alternative scanning pattern (on picture basis)
  – In case of quantizer matrix download: use always diagonal (zigzag) scanning



MPEG-2 alternative scanning

MPEG Video / Var. Length Coding – (1)

* **Variable-length coding of AC coefficients: algorithm of (runlength, amplitude) coding**
  – STEP 1: (load coefficient), test of coefficient is zero
  – STEP 2: (update runlength), if zero coefficient, increment zero counter, go to STEP 4
  – STEP 3: (jointly code), if non-zero coefficient, then
  – 3a. jointly code [runlength, amplitude] in one codeword
  – 3b. reset runlength counter
  – STEP 4: (do next coefficient), go to STEP 1. If last coefficient, then go to STEP 5.
  – STEP 5: (EOB) Terminate block with EOB-word, ignore runlength value. Codetable is modified Huffman code.

MPEG Video / Var. Length Coding – (2)

* **2-D VLC table of codewords**
  1. Unlikely symbols are coded by [esc. code]+[fixed suffix]
  2. Also VLC coding of macroblock address, motion vectors,...

zero run ↓ / amplitude →	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	11	2	3	4	4	5	5	5	6	6	6	7	7	7	8	8	8	8	9	9	14
1	12	4	6	7	7	8	9	9	10	10	10	11	11	12	12	12	12	12	12		
2	12	5	7	9	10	10	11	12	12	13											
3	12	6	8	10	12	12	13														
4	12	6	9	11	12																
5	12	7	10	11	12																
6	13	7	10	12																	
7	13	8	12	12																	
8	13	8	12																		
9	13	9	12																		
10	13	9	13																		

EOB = 4

Example of wordlength table

• MPEG-2 has alternative encoding table for intrablocks. For non-intra blocks, always the same table is used.

MPEG Video / Var. Length Coding – (3)

- **2-D. VLC Table of code words**
  - Special code for 1st coeff.
  - Escape code to avoid long code words
  - Appended sign bit

code	runlength	amplitude
10	EOB	
1s (note2)	0	1
11s (note3)	0	1
011s	1	1
0100s	0	2
0101s	2	1
0010 1s	0	3
0011 1s	3	1
0011 0s	4	1
0001 10s	1	2
0001 11s	5	1
0001 01s	6	1
0001 00s	7	1
0000 110s	0	4
0000 100s	2	2
0000 111s	8	1
0000 101s	9	1
0000 01	escape	-

code	runlength	amplitude
0010 0110s	0	5
0010 0001s	0	6
0010 0101s	1	3
0010 0100s	3	2
0010 0111s	10	1
0010 0011s	11	1
0010 0010s	12	1
0010 0000s	13	1
0000 001010s	0	7
0000 001100s	1	4
0000 001011s	2	3
0000 001111s	4	2
0000 001001s	5	2
0000 001110s	14	1
0000 001101s	15	1
0000 001000s	16	1

Note1: s=sign bit, 0=pos/1=neg.
Note 2: code for dct_coeff_first
Note 3: code for dct_coeff_next

MPEG Video / Var. Length Coding – (4)

- **VLC table for motion vectors**
  - Sym-metrical
  - Special code
  - Appended sign bit except 0

VL code	motion code
0000 0011 001	-16
0000 0011 011	-15
0000 0011 101	-14
0000 0011 111	-13
0000 0100 001	-12
0000 0100 011	-11
0000 0100 11	-10
0000 0101 01	-9
0000 0101 11	-8
0000 0111	-7
0000 1001	-6
0000 1011	-5
0000 111	-4
0001 1	-3
0011	-2
011	-1
1	0

VL code	motion code
0000 0011 000	+16
0000 0011 010	+15
0000 0011 100	+14
0000 0011 110	+13
0000 0100 000	+12
0000 0100 010	+11
0000 0100 10	+10
0000 0101 00	+9
0000 0101 10	+8
0000 0110	+7
0000 1000	+6
0000 1010	+5
0000 110	+4
0001 0	+3
0010	+2
010	+1
1	0