Multimedia Coding and Transmission

Video Coding

Ifi, UiO

Norsk Regnesentral

Vårsemester 2005

Wolfgang Leister

This part of the course ...

- ... is held at Ifi, UiO ...
  (Wolfgang Leister)
- ... and at University College Karlsruhe
  (Peter Oel, Clemens Knoerzer)

The story so far ...

- Data compression
  - information theory
  - run length encoding
  - Huffman coding
  - Zif-Lempel(-Welch) algorithm
  - Arithmetic coding
- Colour coding and raster images

Overview

- Video Formats
  - MJPEG
  - H.261
  - MPEG I, II
  - MPEG TS
  - H.263
  - AVI,QuickTime

MJPEG

- Motion-JPEG
- Sequence of JPEG-Frames
- not a standard
- many proprietary formats
  - e.g., AVI and QuickTime
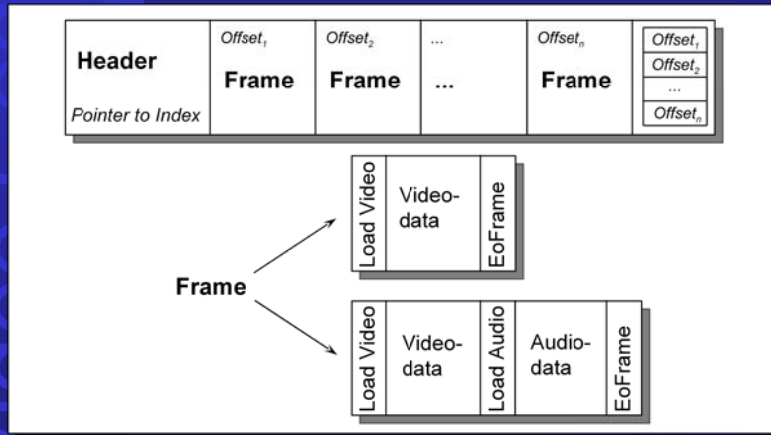- only Baseline-JPEG
- audio track(s)

MJPEG

- Compression as in JPEG
- Temporal dependencies are not used.
- suitable for video cutting software:
  - cut is possible at each frame
  - no quality decrease when cutting
- Hardware support possible!

MJPEG by Parallax

Header	Offset₁	Offset₂	...	Offsetₙ	Offset₁
	Frame	Frame	...	Frame	Offset₂
Pointer to Index					...
					Offsetₙ

Frame →

Load Video	Video-data	EoFrame

Load Video	Video-data	Load Audio	Audio-data	EoFrame

MJPEG by Parallax

- Header
  - Id / Version
  - Frames per second / number of frames
  - Width / Height
  - Bandwidth
  - Quantising factor
  - Number of Audio Tracks / Sampling Rate
  - Offset of frame index

H.261

- CCITT (Commité Consultatif International de Télécommunications et Télégraphique)
- ITU-T (International Telecommunication Union)
- Video Codec for Audiovisual Services at p x 64 kbit/s
- P64
- Developed for image telephony and video conferences

H.261

- Adapted for ISDN (p x 64 kbit/s)
- Const. data rate by feedback
- 4:2:0 Sampling (Chrominance-channels with half the resolution of luminance channel (number of rows and columns))
- Two Resolutions:
  - CIF: 352 x 288 (Common Intermediate Format)
  - QCIF: 176 x 144 (Quarter CIF)

Subsampling

- A:B:C Notion of CCIR-601
- **4:2:2** horizontal 2:1 downsampling of colour channels
- **4:1:1** horizontal 4:1 downsampling of colour channels
- **4:2:0** horizontal and vertical 2:1 downsampling of colour channels

Example

	Y	C_b, C_r each
4:2:2	352 x 288	176 x 288
4:1:1	352 x 288	88 x 288
4:2:0	352 x 288	176 x 144

H.261

Two frame types:

- IntraFrames
  - Very similar to JPEG-Image (DCT, Quantising, Coding)
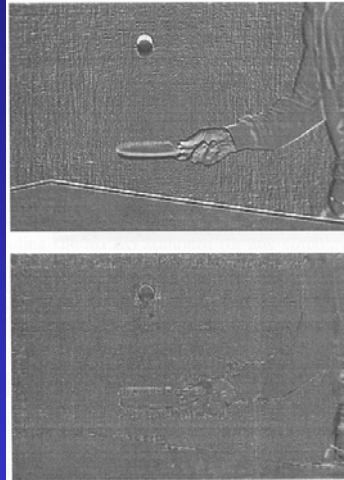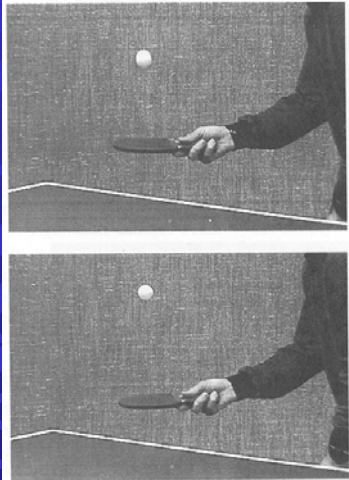- InterFrames
  - Code differences to previous frame.
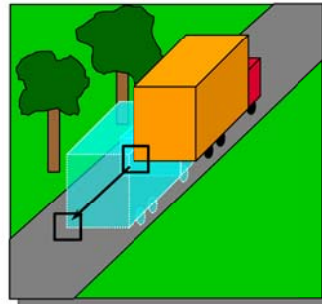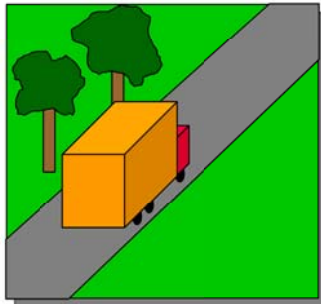 - Movements are compensated by motion-estimation.

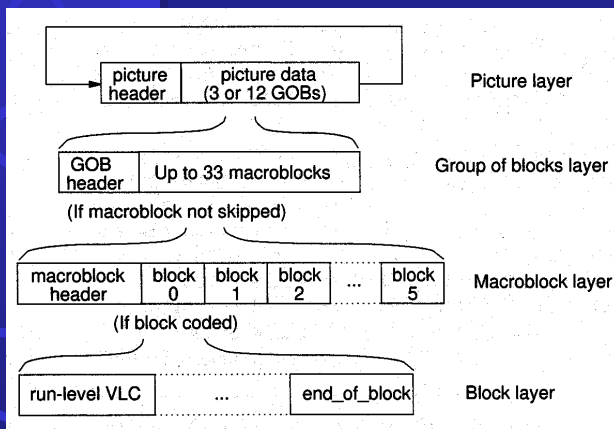## Intra- / InterFrames
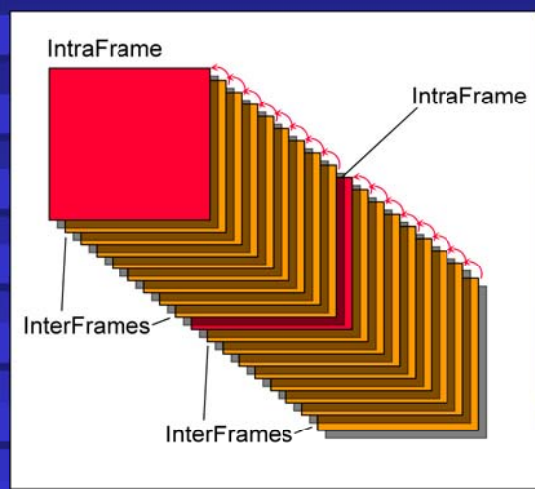
Motion Estimation

Motion-Estimation

H.261

- Motion-Vector max. ±15 Pixels
- all DCT-Coefficients are quantised with same value
- Quantising controled by output stream (feedback).
- Quantising has dead zone

H.261 data layout

MPEG

- Motion Picture Expert Group
- ISO/IEC 11172 (MPEG-1) / 13818 (MPEG-2)
- Video / Audio compression and coding

MPEG

- Differences to H.261
  - Motion-Vectors not limited to ±15 Pixels
  - Motion-Vektors not necessarily integer numbers
  - 3 (4) Frame-Types (I-, P-, B-Frames)
  - Data stream not limited to p x 64 kbit/s.
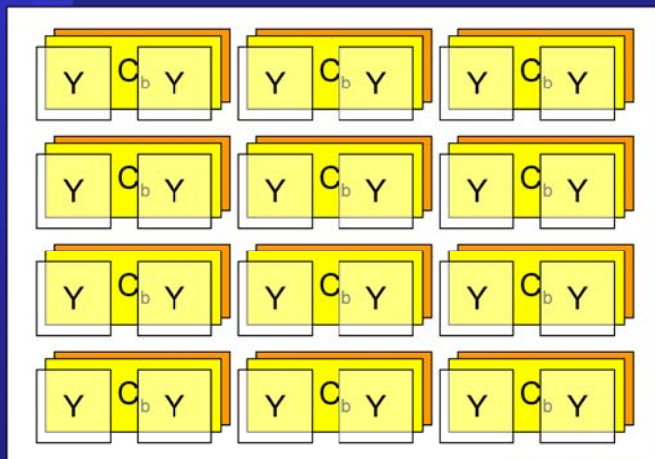 - Quantising of coefficients with matrix
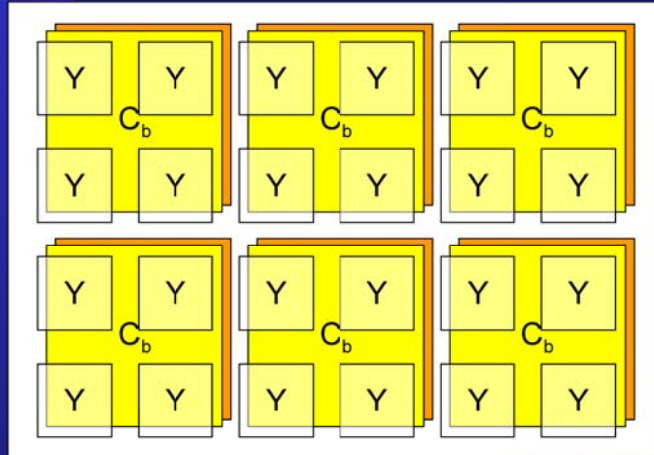
Macro blocks

- Adjacent 8x8-Blocks of channels are joined to macro blocks.
- Depending on sub sampling several 8x8 blocks are in one macro block
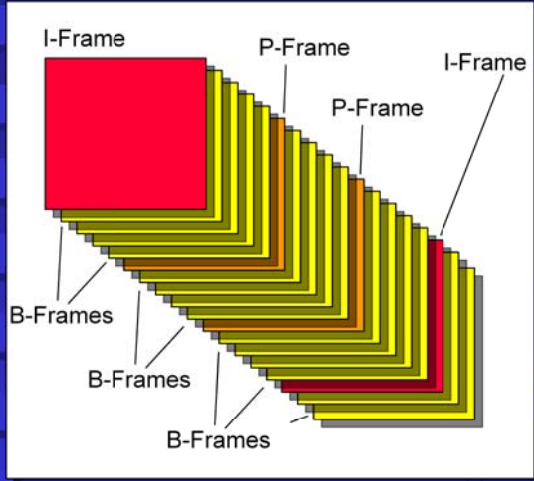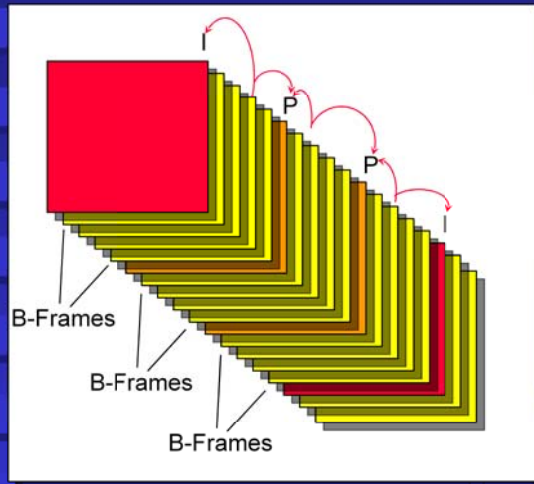

Example 4:2:2

Example 4:2:0

MPEG

- Frame-Types:
  - **I-Frame**: Like IntraFrames of H.261 (ca. every 15. Frame)
  - **P-Frame**: Like InterFrames of H.261 (Predicted Frame). Related to previous I or P-Frame
  - **B-Frame**: (Bidirectional predicted Frame) No equivalent of H.261. Related to previous _and_ successor I- or P-Frame

I-, P- and B-Frames

I-, P- and B-Frames

Motion

- Motion compensation
  - normative part of MPEG
  - Decoder's point of view
- Motion estimation
  - NOT normative part of MPEG
  - Encoder's point of view

Motion Compensation

- 1 motion vector for each region
- region = macro block
- $\forall \Rightarrow$ 1 motion vector per macro block
- Precision: 1 Pel, ½ Pel
- motion vector coded differentially
  (prediction derived from preceding macro block)
- Rules for resetting motion displacements

Motion Compensation

- P-Pictures:
  - forward motion vectors
- B-Pictures:
  - forward motion vectors
  - backward motion vectors
  - if both used: average of pel values from forward and backward motion-compensated reference picture
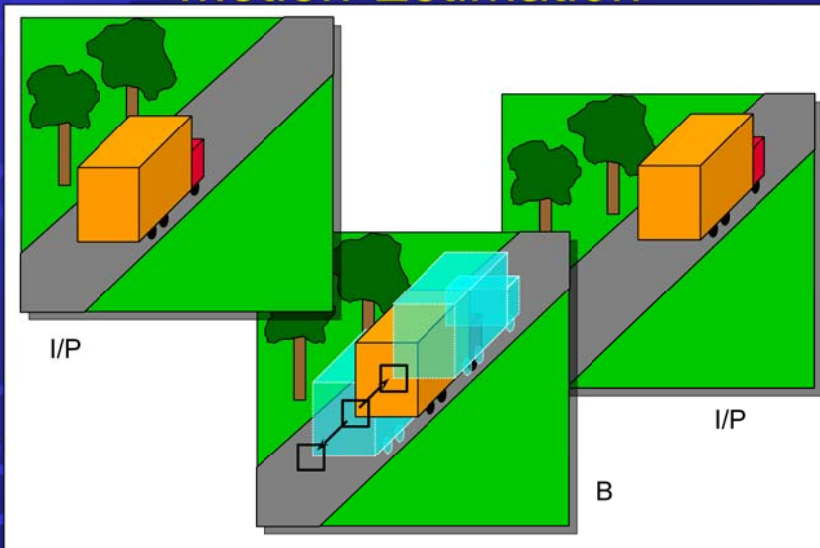
Motion Compensation

- MPEG-2:
  - as in MPEG-1 (previous slide)
  - uses different names
  - Dual Prime Motion
    - for interlaced video
    - averaging predictions from two adjacent fields of opposite parity

---

Motion-Estimation

I/P

I/P

B

Motion-Estimation

- Criterion for block matching
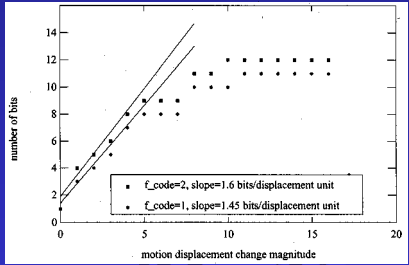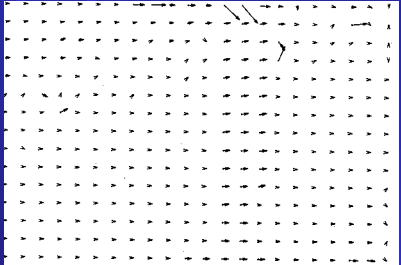 - Mean Square Error
  - Mean Absolute Distortion

$$MAD(x,y) = (1/256)\sum_{i=0}^{15}\sum_{j=0}^{15} | V_n(x+i, y+i) - V_m(x+dx+i, y+dy+i) |$$

  - Sum of Absolute Distortions (SAD)
  - Minimization of the bitstream

Motion Estimation

- Correlation between motion vectors

Motion Estimation

- Motion displacement search algorithms
  - pel-recursive
    - iterative process
    - use intensity gradient and frame difference
  - block matching
    - compute measure of distortion
    - select vector that minimizes distortion

Motion Estimation

- Fast search algo - sparse sampling
- Variable resolution search techniques
- Statistically sparse searches
- Spatial continuity
- Telescopic search
- 3D spatial/temporal estimation
- Phase correlation
- others ...

Motion Estimation

Overview in:

Mitchell, Pennebaker, Fogg, and LeGall:
MPEG Video Compression Standard,
Chapman&Hall, 1996
p 301 ff.

Frame Order

Display order

$I_0B_1B_2P_3B_4B_5P_6B_7B_8I_9...$

$I_0P_3B_1B_2P_6B_4B_5I_9B_7B_8$ …

Coding order

MPEG data layout
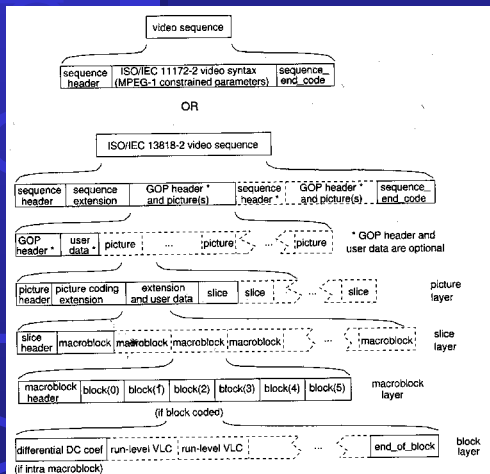
MPEG Encoder
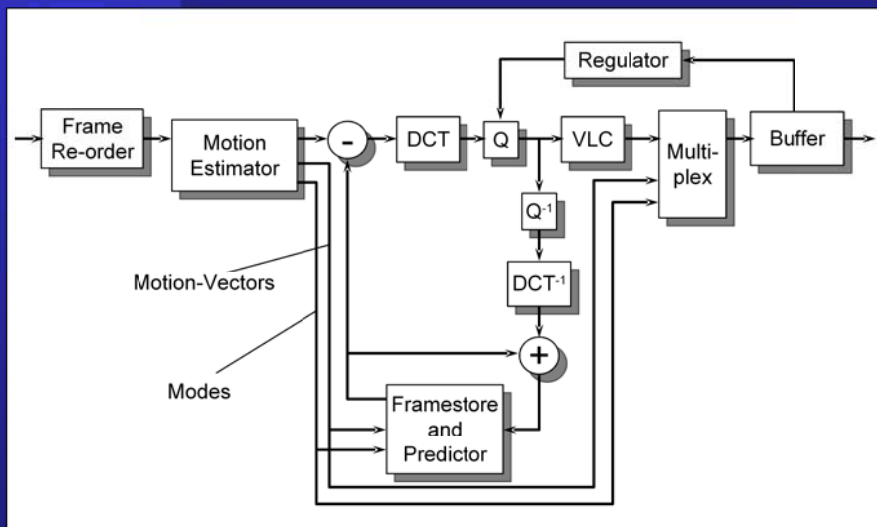
MPEG Decoder
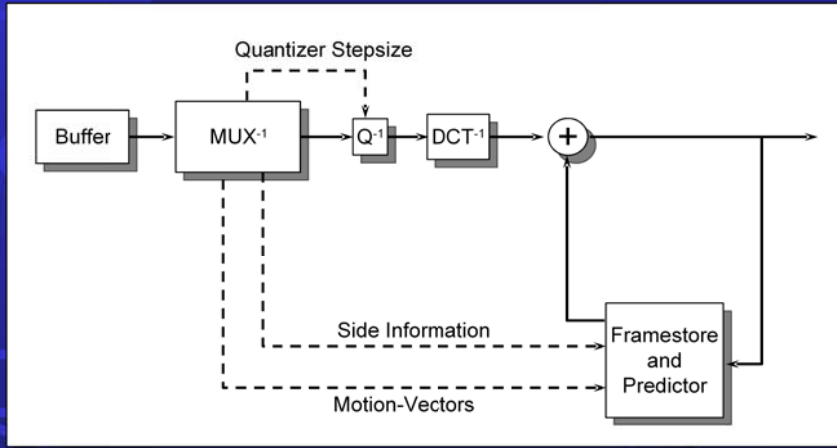
MPEG TS

- Packetised Elementary Stream (PES)
- MPEG-2 Programme Stream ——— MPEG-1 compatibility
- MPEG-2 Transport Stream (TS)
    - fixed size 188 bytes packets
- DSM-CC: Digital Storage Medium Command and Control
- Service Information Tables
    - Information on streams, programmes, networks, conditional access, textual description, rating, ...

MPEG-2 TS
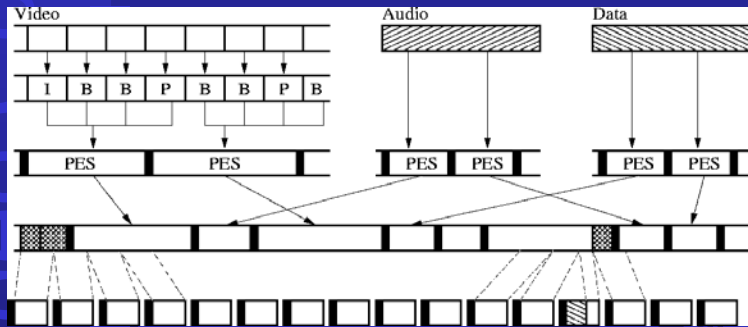
MPEG-2 Multiplexer
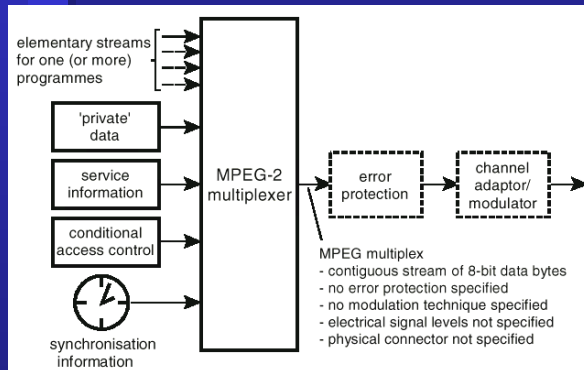
Presentation Unit
Access Unit
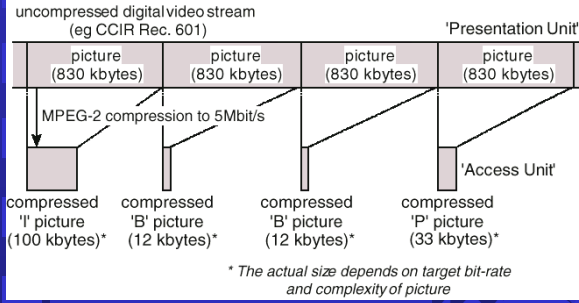
PES
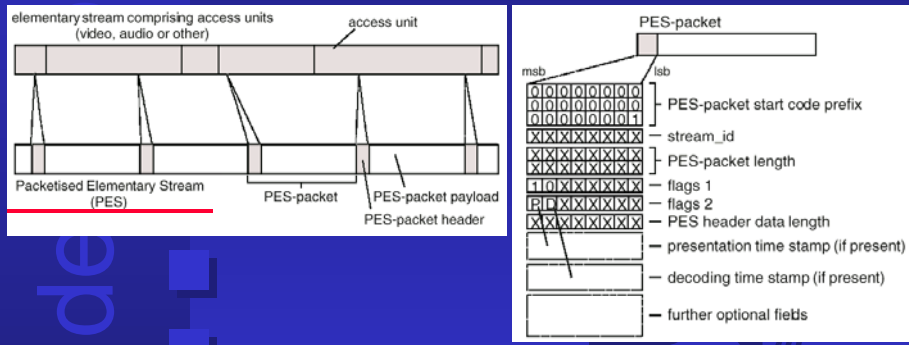
Programme Stream Multiplex

- Pack Header
  - contains system clock reference
  - must occur every 0.7 sec.
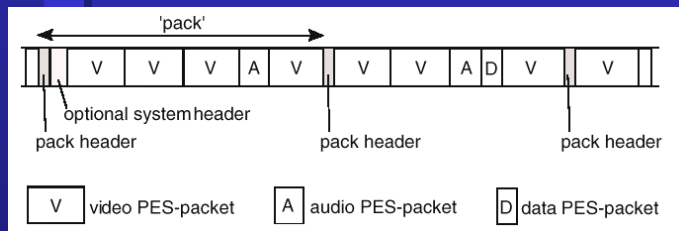
MPEG-2 TS
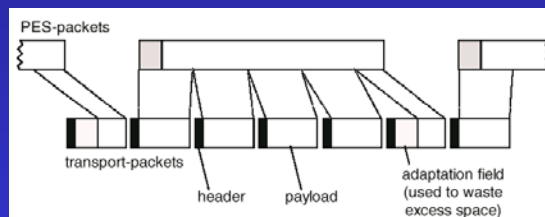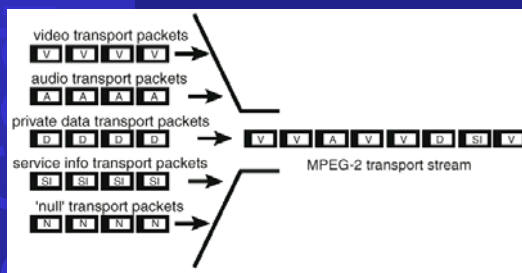
MPEG-2 Transport Packet
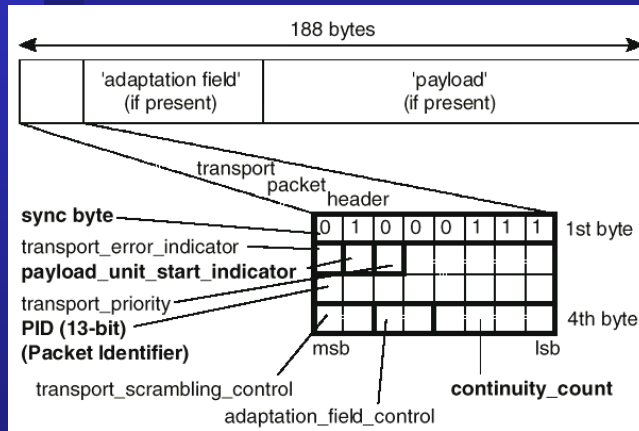
MPEG-2 TS

PSI
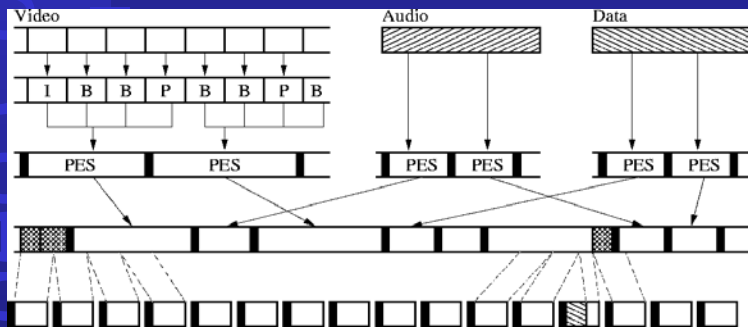
- PSI = Programme Specific Information
- Programme Map Table (PMT)
  - what belongs to a programme?
  -
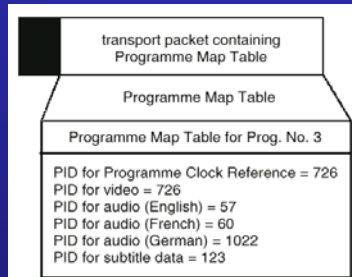
transport packet containing
Programme Map Table

Programme Map Table

Programme Map Table for Prog. No. 3

PID for Programme Clock Reference = 726
PID for video = 726
PID for audio (English) = 57
PID for audio (French) = 60
PID for audio (German) = 1022
PID for subtitle data = 123

PSI

- PSI = Programme Specific Information
- Programme Map Table (PMT)
- Programme Association Table (PAT)
  - which PMT are availab
  - PID=0

PID=0

transport packet containing
Programme Association Table

Programme Association Table

Network Information Table

programme number	Programme Map Table PID value
0	16
1	306
3	1127
18	17
7	244
8	244
10	244
4	87

Programme Map Table for Prog. 1

Programme Map Table for Prog. 3

PSI

transport packet containing Programme Map Table

Programme Map Table

Programme Map Table for Prog. No. 3

PID for Programme Clock Reference = 726
PID for video = 726
PID for audio (English) = 57
PID for audio (French) = 60
PID for audio (German) = 1022
PID for subtitle data = 123

PID=0

transport packet containing Programme Association Table

Programme Association Table

programme number	Programme Map Table PID value
0	16
1	306
3	1127
18	17
7	244
8	244
10	244
4	87

Network Information Table

Programme Map Table for Prog. 1

Programme Map Table for Prog. 3

PMT: PID=1127   PAT: PID=0 (always)

PSI

- PSI = Programme Specific Information
- Programme Map Table (PMT)
- Programme Association Table (PAT)
- Network Information Table (NIT)
  - PMT=0
  - optional, content not defined by MPEG
  - channel frequencies
  - satellite transponder
  - modulation characteristics
  - service originator, service name,

PSI

- PSI = Programme Specific Information
- Programme Map Table (PMT)
- Programme Association Table (PAT)
- Network Information Table (NIT)
- Conditional Access Table (CAT)
  - scrambling system(s) in use
  - PID for conditional access management
  - PIC for entitlement information
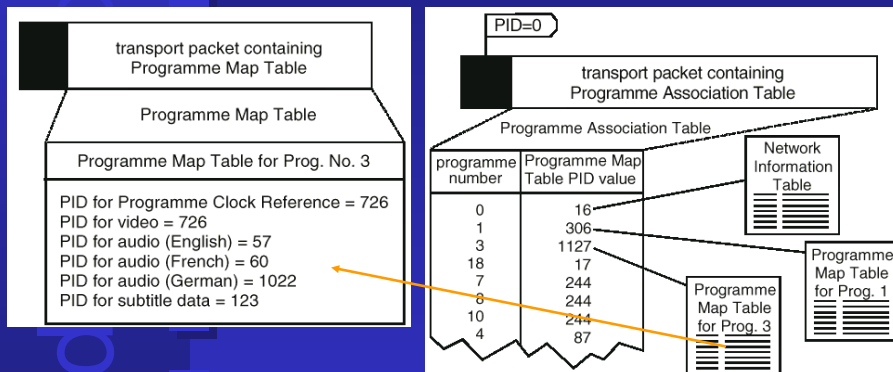 - not defined by MPEG

PSI

- PSI = Programme Specific Information
- Programme Map Table (PMT)
- Programme Association Table (PAT)
- Network Information Table (NIT)
- Conditional Access Table (CAT)
- Programme Stream Map
  - for programme streams

PSI

- Programme Map Table (PMT)
- Programme Association Table (PAT)
- Network Information Table (NIT)
- Conditional Access Table (CAT)

MPEG and IP

- MPEG is both storage and transport medium
- MPEG streams can be transported on IP
- UDP packets can be transported on MPEG
- Unidirectional application protocols
- Broadcast
- Multicast
- Unicast
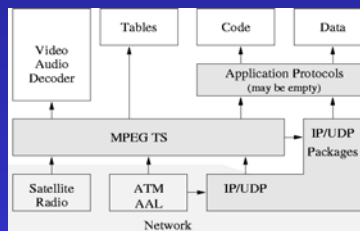
Time

- Multiplexer / Decoder Clock measure time in units of 27 MHz, 42 bit
- Time Stamps expressed in units of 90 kHz, 33 bit

Programme Stream:

- System Clock Reference (SCR)
  - at least every 0.7 sec

Transport Stream:

- Programme Clock Reference (PCR)
  - at least every 0.1 sec

Time

- Presentation Time Stamp (PTS)
- Decoding Time Stamp (DTS)
  - temporarily stored at DTS, not shown
  - I and P pictures (accessed for B pictures)
  - PTS > DTS

MPEG - I

- Resolution: 360 x 288 x 25 (CIF Eur.)
                352 x 240 x 30 (CIF US)
- Bitrate <= 1,862 Mbit/s
- Macroblocks/Frame <= 396
- Macroblocks/Second <= 9900
  - (396 x 25, 330 x 30)
- No Interlace
- Audio: 2 Channels (Stereo)

MPEG - II

- Resolutions
  - low: 360 x 288 (352 x 240)
  - main: 720 x 596 (704 x 480)
  - high1440:   1440 x 1152 (HDTV 4:3)
  - high: 1920 x 1080 (HDTV 16:9)
- Bitrate: 2 - 80 Mbit/s
- Interlace Support
- 5 Audio-Channels

Data rates

	Frametyp			
	I	P	B	Ø
MPEG-1 SIF 1.15 Mbit/s	150.000	50.000	20.000	38.000
MPEG-2 601 4.00 Mbit/s	400.000	200.000	80.000	130.000

I-Distance : 15
P-Distance: 3

~~MPEG-3~~

MPEG-4

MPEG-7

MPEG-21

These are subject to another session …
– be patient …

H.263

- ITU-T approved standard
  - "Video coding for low bit rate communication"
  - v1 1996
  - v2 1998 = H.263+; focus today
  - v3 ? H.263++/H.26L
- Unspecified, variable bit rate (<28.8 kbps)
- Developed for low delay environments

H.263 block diagram

|	
T	Transform
Q	Quantizer
P	Picture Memory with motion compensated variable delay
CC	Coding control
p	Flag for INTRA/INTER
t	Flag for transmitted or not
qz	Quantizer indication
q	Quantizing index for transform coefficients
v	Motion vector

H.263 sampling blocks

- 4:2:0 sampling
  - luminance Y to chrominance C_B, C_R
- Block:
  - 8 x 8 pixels
- Macroblock (MB):
  - 4 Y + C_B + C_R blocks
- Group of blocks (GOB):
  - One or more rows of MBs
  - In GOB header: resynchronization

Y	Y
Y	Y
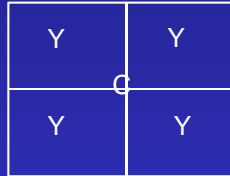
H.263 resolutions

- Five standardized resolutions
  - CIF: 352 x 288 (*as in H.261*)
  - QCIF: 176 x 144 (*as in H.261*)
  - Sub-QCIF: 128 x 96
  - 4CIF: 704 x 576 (little used?)
  - 16CIF: 1408 x 1152 (little used?)
- Custom resolutions negotiable
  - multiple of 4 in both directions

H.263 frames

- Two (six) frame types:
  - I-frames: intra
  - P-frames: predictive (inter)
  - B-frames (*optional*): bidirectional predicted
  - PB-frames (*optional*): decoded B and P frame as one unit
  - EI-frames (*optional*): enhanced I-frame
  - EP-frames (*optional*): enhanced P-frame

H.263 coding

- Spatial redundancy (intra coding):
  - DCT
  - Variable length coding (Huffman)
  - Quantisation
- Temporal redundancy (inter coding):
  - Motion compensation
    - Block-based comparison (MB or block)
    - Nonintegral motion vector values (half-pixel)
  - Motion estimation

H.263 coding options (1)

- 16 modes, *negotiable* at session start:
  - Efficiency / improved picture quality (10)
    - Unrestricted Motion Vector
    - Syntax-Based Arithmetic Coding
    - Advanced Prediction!
      - Four Motion Vectors per Macroblock
      - Overlapped Block Motion Compensation
    - PB Frame
    - Advanced Intra Coding
    - Alternate Inter VLC
    - Modified Quantization
    - Deblocking Filter
    - Improved PB Frame

H.263 coding options (2)

- Error robustness (3) (lossy channel)
  - Slice-Structured
  - Reference Picture Selection
  - Independent Segment Decoding
- Scalability!
  - Temporal, signal-to-noise ratio (SNR), spatial
- Reference Picture Resampling
- Reduced Resolution Update

H.263 levels

- Preferred combination of supported options (profiles)
  - Level 1: advanced intra coding, deblocking filter, full-frame freeze, modified quantization
  - Level 2: unrestricted motion vector, slice-structured, reference picture resampling + level 1
  - Level 3: advanced prediction, improved PB frames, independent segment decoding, alternate VLC + level 2

AVI

- Audio Video Interleaved, by Microsoft Corp.
- specialisation of RIFF (Resource Interchange File Format)
- Container-Format
- File consists of blocks (chunks)
- Each chunk characterised by 4 letters
- format: ID (4Byte) length (4 Byte) data
- Each chunk can contain sub-chunks.

AVI

- Supported compressors:
  (Status 1994/1995)
  - Cinepack Codec by SuperMatch
  - Intel Indeo Video R2.1 , R3.1 and Raw
  - Microsoft Video 1
  - Microsoft RLE
  - Motion JPEG
  - MPEG

QuickTime

- Charles Wiltgen
- Version 2.1 (of Juli 1996)
- Originally for Macintosh
- available for MS-windows and others
- Container-Format
- File built of blocks (tracks)
- Each block can contain sub-blocks
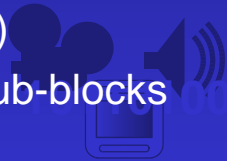
QuickTime

	QuickTime for Macintosh 2.5	QuickTime for MS-Win. 2.1.1	QuickTime for MS-Win. 2.5
Animation	r/w	r	r/w
Cinepak	r/w	r	r/w
Component Video	r/w		r/w
Graphics	r/w	r	r/w
M-JPEG A	r/w		r/w
M-JPEG B	r/w		r/w
None	r/w	r	r/w
Photo-JPEG	r/w	r	r/w
Video	r/w	r	r/w
Photo-CD	r	r	r
Intel Indeo	r	r	r

Literature

- Joan L. Mitchell, William B. Pennebaker, Chad E. Fogg, and Didier J. LeGall: "MPEG Video Compression Standard", Chapman&Hill, ISBN 0-412-08771-5, 1996
- Ming-Ting Sun, and Amy R. Reibman:"Compressed Video over Networks", Marcel Dekker, ISBN 0-8247-9423-0, 2001, Kap. 2
- Wolfgang Leister, Svetlana Boudko, Ole Aamodt, Peter Holmes: "Digital TV - a survey", NR Report 988, 2002. http://publications.nr.no/digitv.pdf

Literature

- MPEG System:
  - P.A. Sarginson: "MPEG-2: Overview of the System Layer", 1996, http://www.bbc.co.uk/rd/pubs/reports/1996-02.pdf

The End of this Lecture