Multimedia Coding and Transmission

Video Coding

Ifi, UiO
Norsk Regnesentral
Vårsemester 2003
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

The story so far ...

• Lossy Image Coding
  – JPEG (JIFF)
  – Wavelet-Coding
  – Fractal Coding
Overview

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

![Diagram of MJPEG by Parallax]
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 (Comité 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

<table>
<thead>
<tr>
<th></th>
<th>Y</th>
<th>C_b, C_r each</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:2:2</td>
<td>352 x 288</td>
<td>176 x 288</td>
</tr>
<tr>
<td>4:1:1</td>
<td>352 x 288</td>
<td>88 x 288</td>
</tr>
<tr>
<td>4:2:0</td>
<td>352 x 288</td>
<td>176 x 144</td>
</tr>
</tbody>
</table>

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
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
- ⇒ 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

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

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

• Correlation between motion vectors

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 TS

- Packetised Elementary Stream (PES)
- MPEG-2 Programme Stream
- 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

[Diagram showing the structure of MPEG-2 Transport Stream, including Video, Audio, Data, PES packets, and MPEG-1 compatibility indication.]
MPEG-2 Multiplexer

- Elementary stream for one (or more) programmes
- Private data
- Service information
- Conditional access control

Presentation Unit
Access Unit

Uncompressed digital video stream (e.g., CCIR Rec. 601)

- MPEG-2 compression to 5Mbits
- Compresed 1 picture (100 kbytes)∗
- Compresed 1F picture (50 kbytes)∗
- Compresed IF picture (70 kbytes)∗
- Compresed P picture (33 kbytes)∗

∗ The actual size depends on target bit-rate and complexity of picture
Programme Stream Multiplex

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

MPEG-2 Transport Packet

188 bytes

adaptation field
(payload if present)

sync byte
transport_error_indicator
payload_unit_start_indicator
transport_priority
PID (13-bit)
(Packet identifier)
transport_scrambling_control
adaptation_field_control
continuity_count

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

4th byte

1st byte
MPEG-2 TS

PSI

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

• PSI = Programme Specific Information
• Programme Map Table (PMT)
• Programme Association Table (PAT)
  – which PMT are available
  – PID=0

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
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
Time

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

<table>
<thead>
<tr>
<th>Frametyp</th>
<th>I</th>
<th>P</th>
<th>B</th>
<th>∅</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG-1 SIF 1.15 Mbit/s</td>
<td>150.000</td>
<td>50.000</td>
<td>20.000</td>
<td>38.000</td>
</tr>
<tr>
<td>MPEG-2 601 4.00 Mbit/s</td>
<td>400.000</td>
<td>200.000</td>
<td>80.000</td>
<td>130.000</td>
</tr>
</tbody>
</table>

I-Distance: 15
P-Distance: 3

MPEG - III

- Developed for HDTV
- However, HDTV is covered by MPEG-II.
- Development stopped!
MPEG - IV

Subject to another session …
– be patient ...

AVI

• Audio Video Interleaved
• Microsoft Corp.
• specialisation of RIFF (Resource Interchange File Format)
• Container-Format
AVI

• File put together from blocks (chunks)
• Each chunk characterised by 4 letters
• format: ID (4Byte) length (4 Byte) Daten
• 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
- Originally for Macintosh
- available for MS-windows and others
- Container-Format

QuickTime

- File built of blocks (tracks)
- Each block can contain sub-blocks
## QuickTime

<table>
<thead>
<tr>
<th>Format</th>
<th>QuickTime for Macintosh 2.5</th>
<th>QuickTime for MS-Win. 2.1.1</th>
<th>QuickTime for MS-Win. 2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animation</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Cinepak</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Component Video</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Graphics</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>M-JPEG A</td>
<td>r/w</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-JPEG B</td>
<td>r/w</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Photo-JPEG</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Video</td>
<td>r/w</td>
<td>r</td>
<td>r/w</td>
</tr>
<tr>
<td>Photo-CD</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>Intel Indeo</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
</tbody>
</table>

## Literature

Literature

• MPEG System:
  – P.A. Sarginson: "MPEG-2: Overview of the System Layer", 1996,

The End of this Lecture