Video Compression

Zafar Javed SHAHID, Marc CHAUMONT and William PUECH
Laboratoire LIRMM
VOODDO project

Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier
LIRMM UMR 5506 Université Montpellier II
34392 Montpellier Cedex 5 France
www.lirmm.fr

Outline

- ✓ Introduction to Video compression
- ✓ Brief history of video coding standards
- ✓ Technical overview of H.264/AVC
- ✓ H.264: feature highlights
- ✓ Watermarking Scheme for H264 Video
- ✓ Conclusion

INTRODUCTION TO VIDEO COMPRESSION

The need for compression
- Image -

- CIF image: 352x288 pixel
- RGB representation: 24 bpp (8bits for red,green,blue)
- Total amount of bytes: > 300K
- JPEG: common image compression standard, < 20K, similar quality

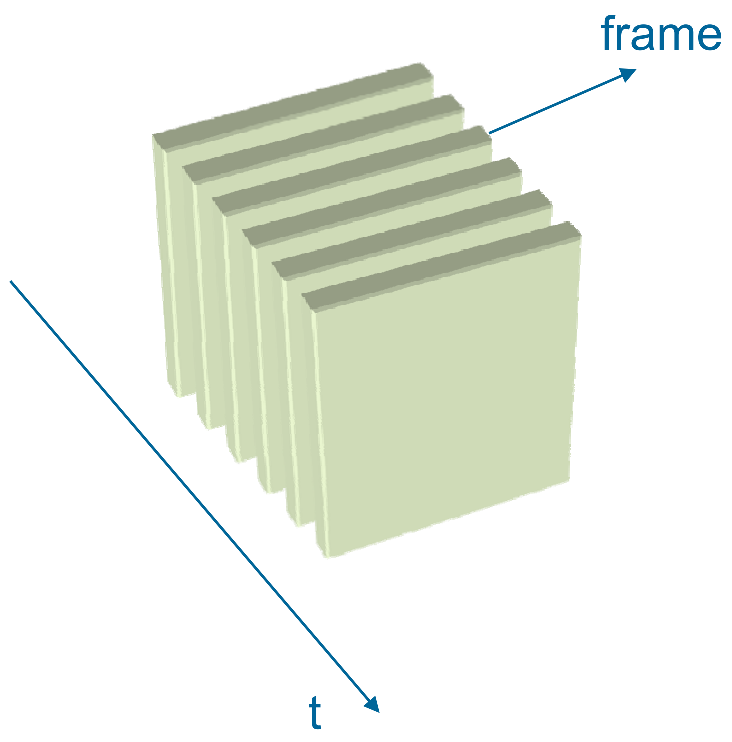
Original Image
Size: 300k

Compressed Image
Size: 20k
PSNR: > 30dB

The need for compression
- Video -

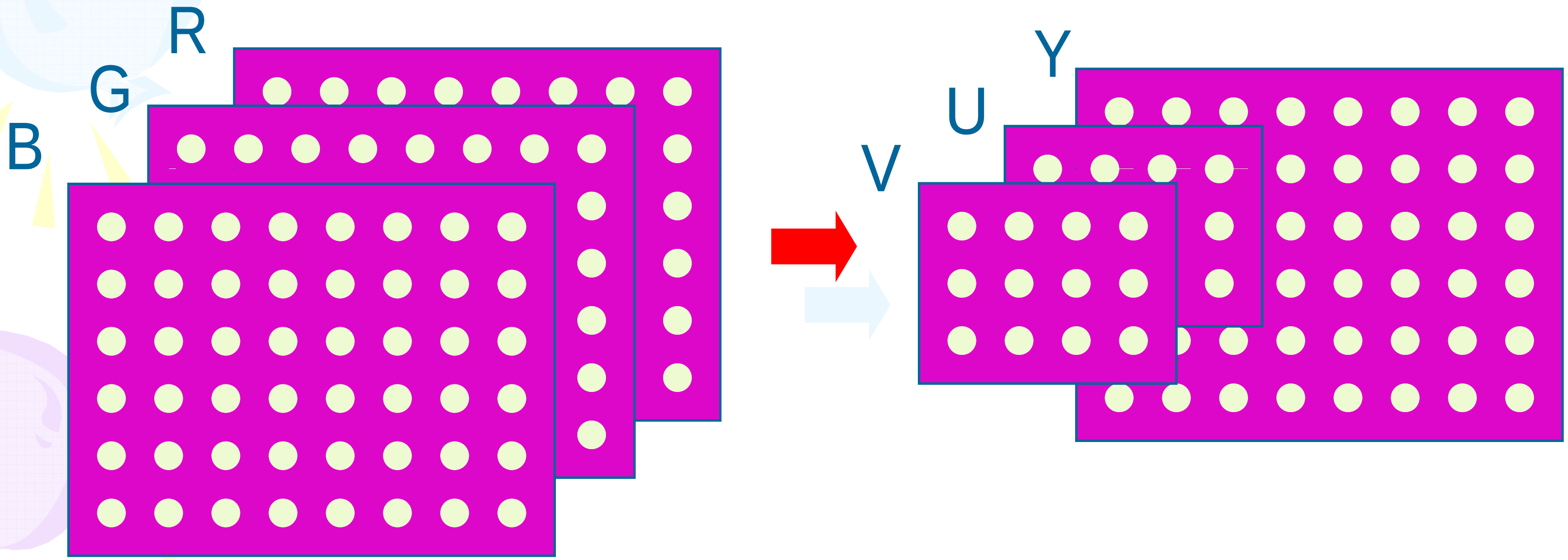
- Video signal:sequence of frames (images) related among temporal dimension
- TV video quality: 704x576 pixels per frame, 12 bpp, 25 frames per second

  > 121 Mbps
- Too much data for video transmission or storage
- Increasing importance of multimedia communication:

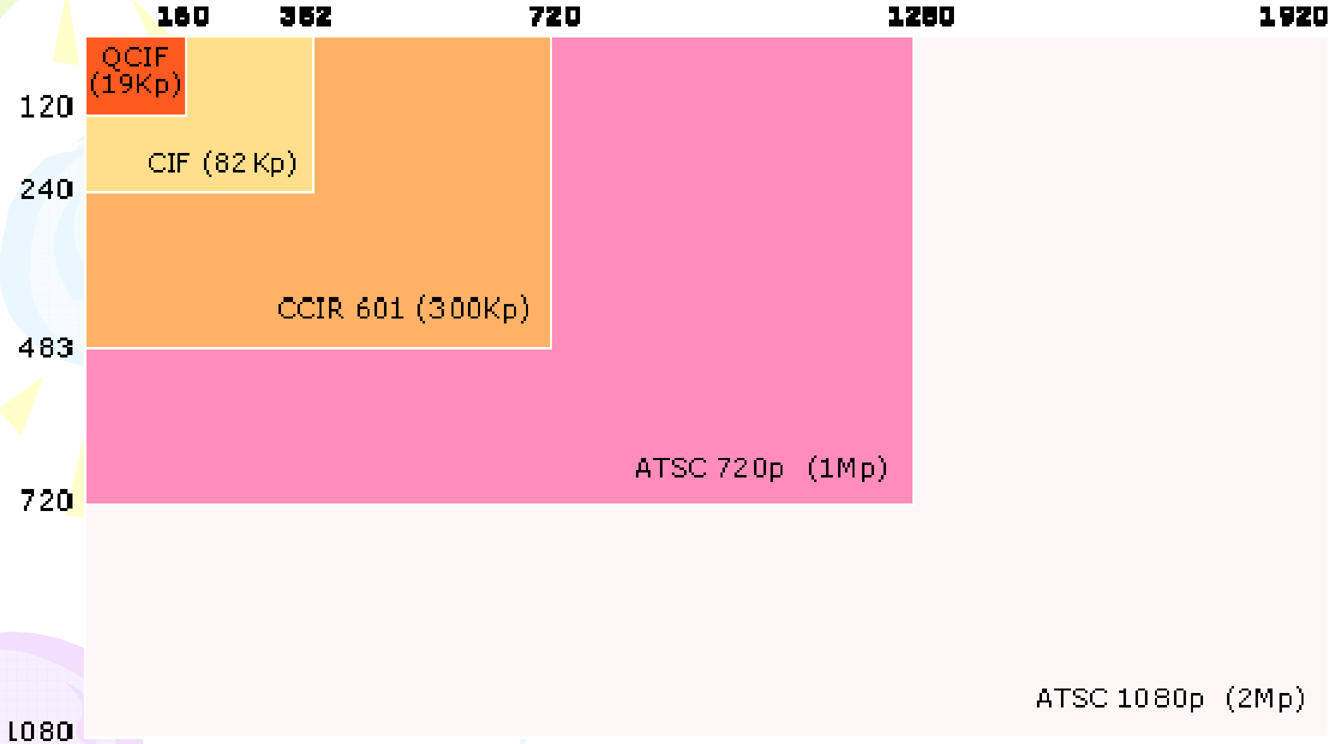


Video color space

- Because human visual system is more sensitive to *luma* than *chroma*, in H264/AVC C_b, C_r components are subsampled and have ¼ of the number of samples than the Y component in 4:2:0 sampling



Video formats



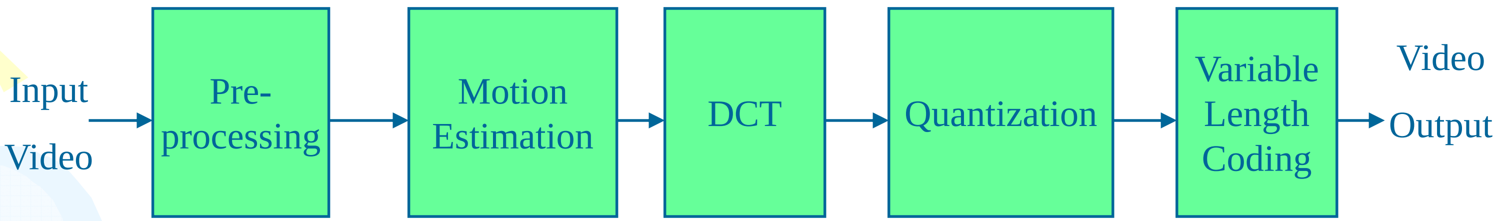
Qcif: Mobile video communication

Cif: Videoconference

CCIR: Standard Definition TV

ATSC: High Definition TV

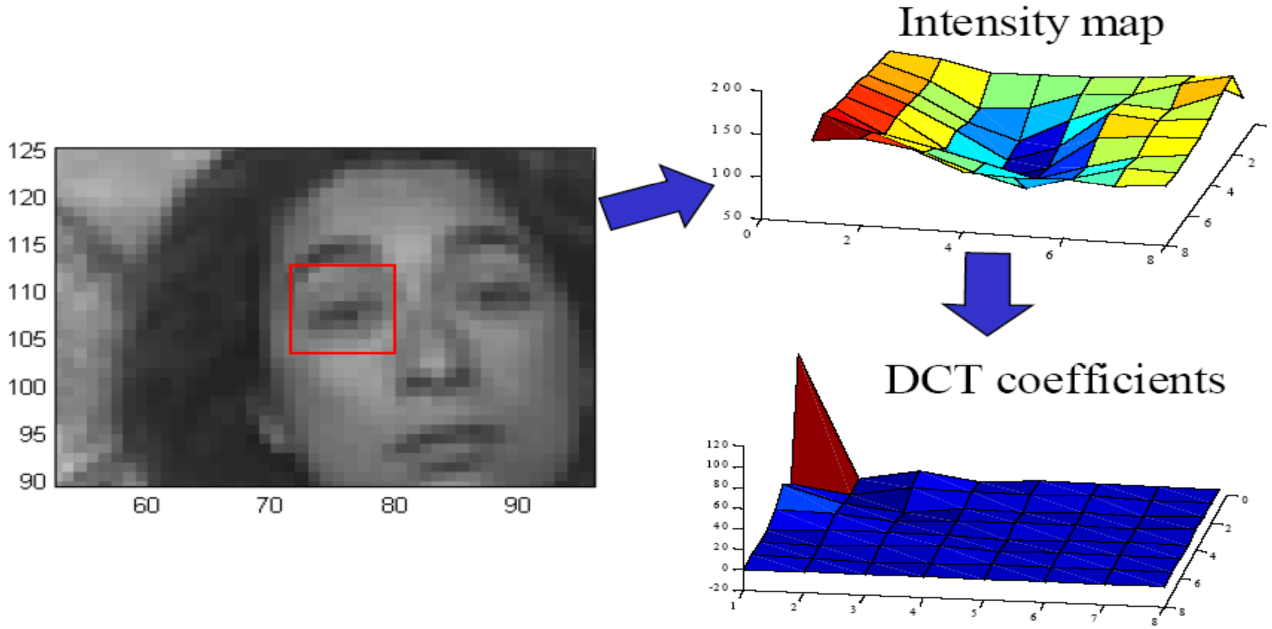
Video Encoder Overview



Example of standard Video encoder

Encoding intra frames DCT

- Transform a block of *n*x*n* pixels into a block of *n*x*n* spatial frequency coefficients
- Energy tends to be concentrated into a few significant coefficients
- Other coefficients are close to zero and not significative

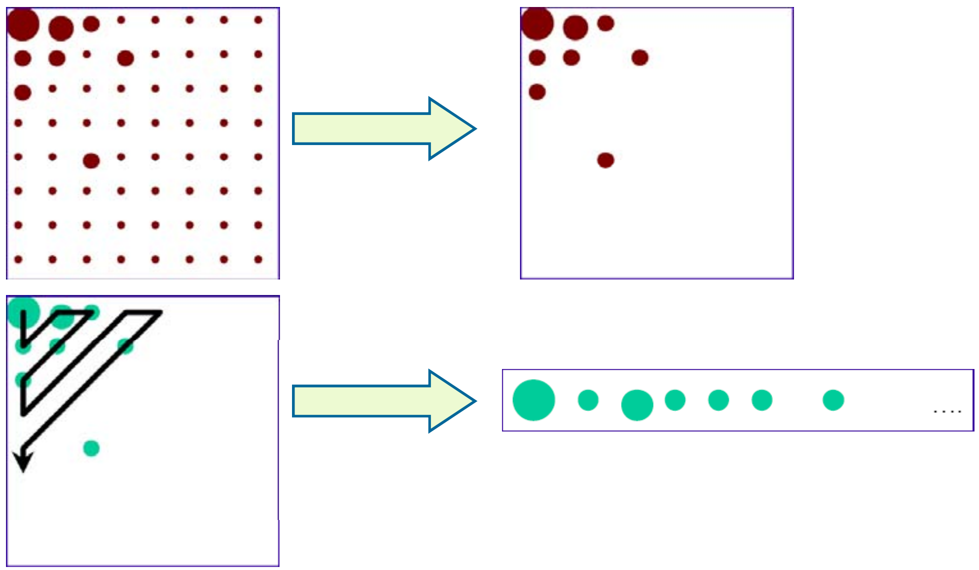


Encoding intra frames - II

- Weighted scalar quantization: loss of precision, few non-zero coefficients are left

- Zig-zag scan: non-zero coefficients tend to be grouped together

- Run-Level encoding: encode each coefficient value as a (run,level) pair
  - run: number of zeros preceding values
  - lenght: non-zero value



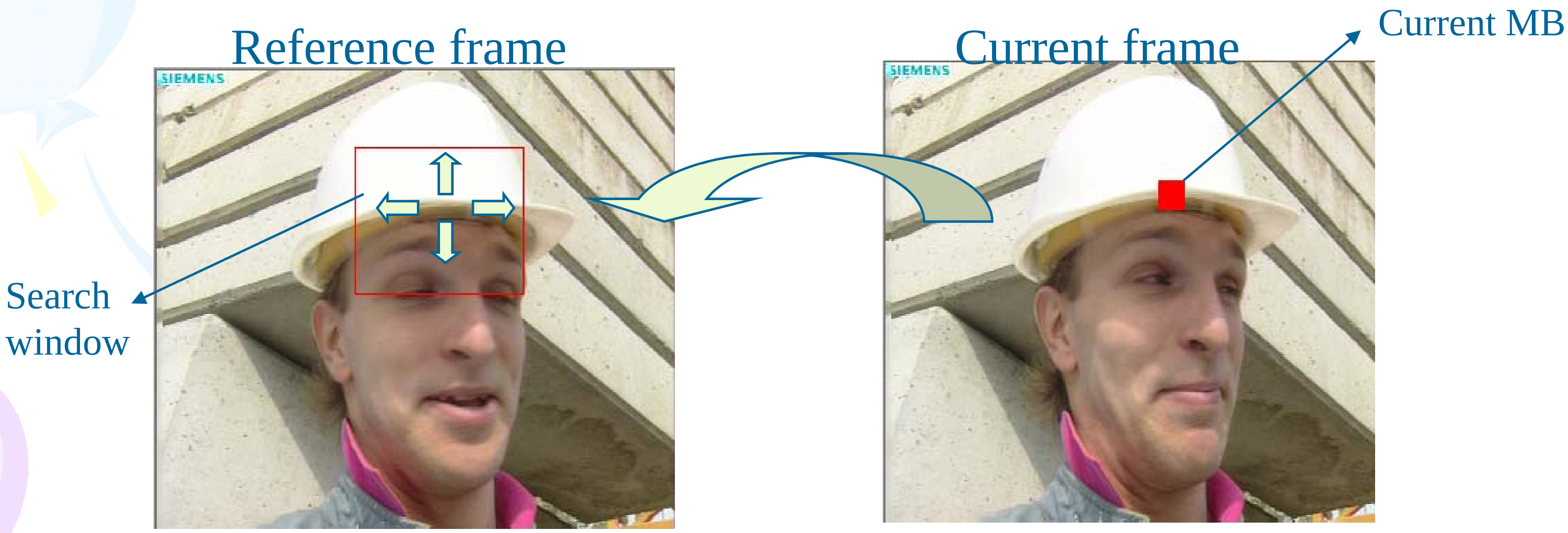
Example:
Original data 14,3,4,0,0,-3,0,0,0,0,0,14,…
(Run,level) (0,14)(0,3)(0,4)(2,-3)(5,14)...

Encoding inter frames

- Main idea: predict current frame using previously coded one (*reference frame*)
- For each MB motion information is extracted from current and reference frame
- Temporal predicted frame is obtained from reference frame using motion information estimated
- The residual (original-MC) are coded using transform coding like for intra frame
- The prediction can be improved by using bi-directional prediction

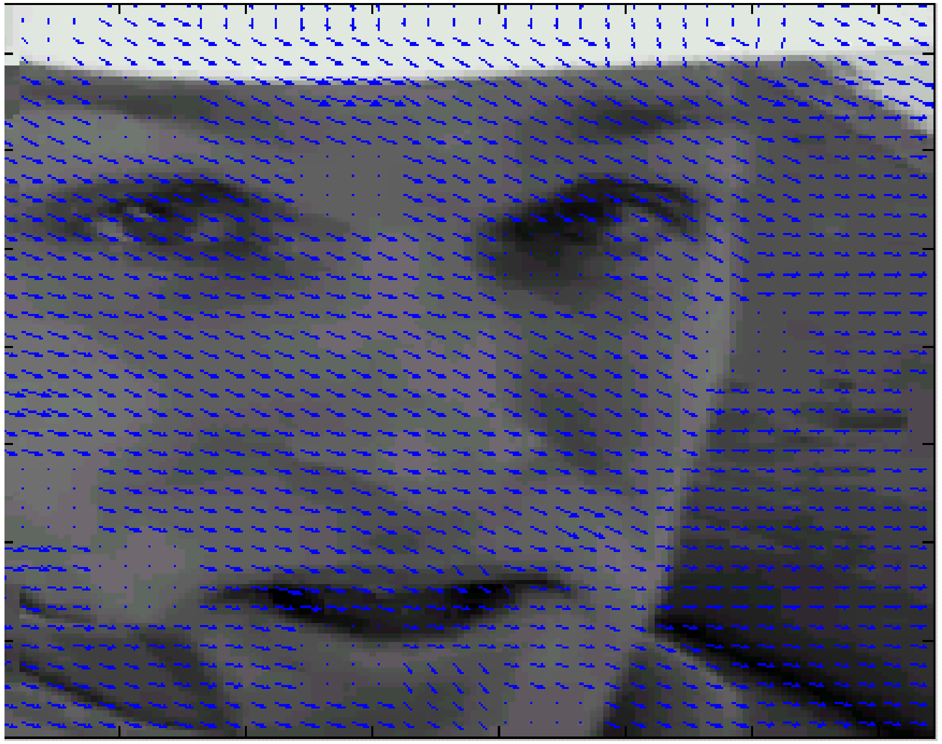
Encoding inter frames
Motion Estimation

- Find the MB in the reference frame which is the most similar to the current MB
- Error metric: MSE or SAD (*Sum of Absloute Differences*)



Encoding inter frame
Motion Estimation

- Assuming a simplified translational model, motion information is described with 2 parameters for each block (*motion vector*)
- *Motion vector:* relative horizontal and vertical offsets (mv_1, mv_2) of a given macroblock from one frame to another
- *Motion vector field:* collection of motion vectors for all the macroblocks in a frame

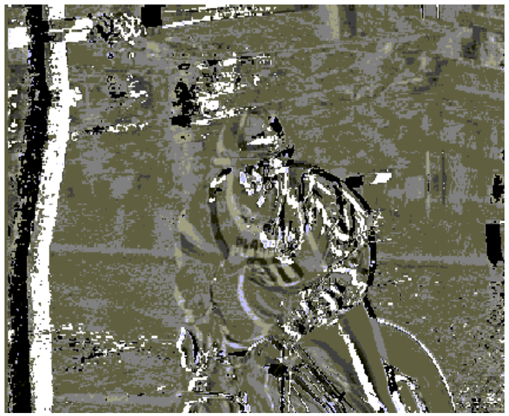


Encoding inter frames
Motion Compensation

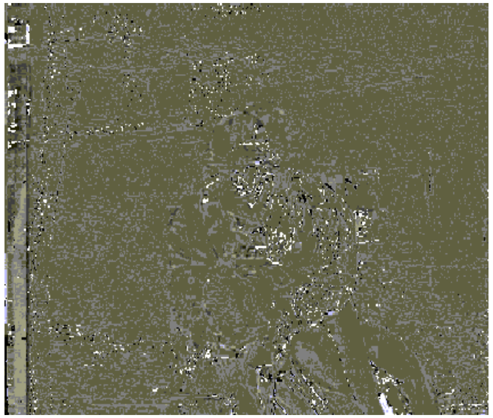
- Adding estimated motion information to the reference frame a *motion compensated* predicted frame is obtained
- The *Displaced Frame Differences* (*DFD = [original – MC]*) is calculated and encoded like for intra frames

Original Frame(N)

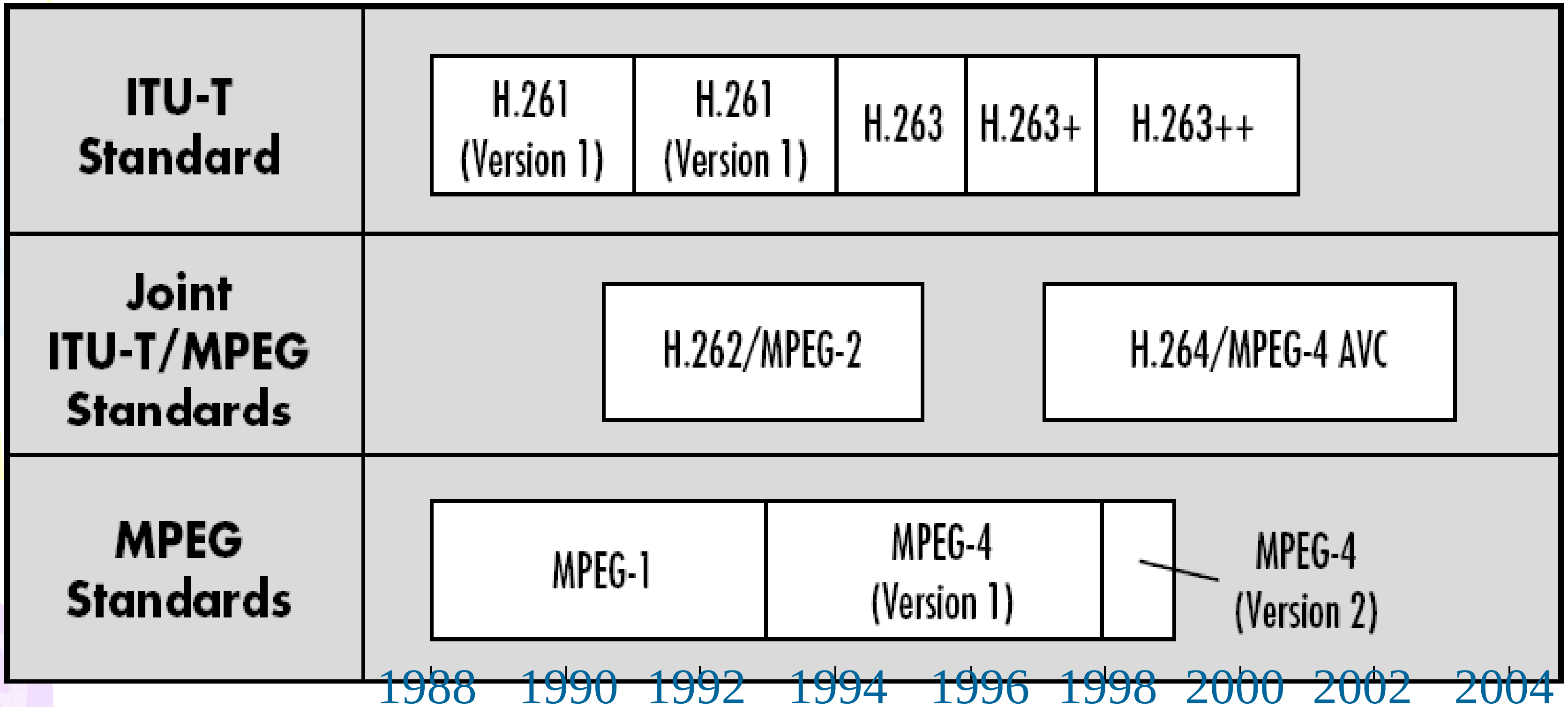
Frame(N)-Frame(N-1)



DFD



Video compression standards– II



H.264

Comparison of video coders (QCIF, 30 fps, 100 kbit/s)

Original

H.263 baseline (33 dB)

H.263+ (33.5 dB)

MPEG-4 core (33.5 dB)

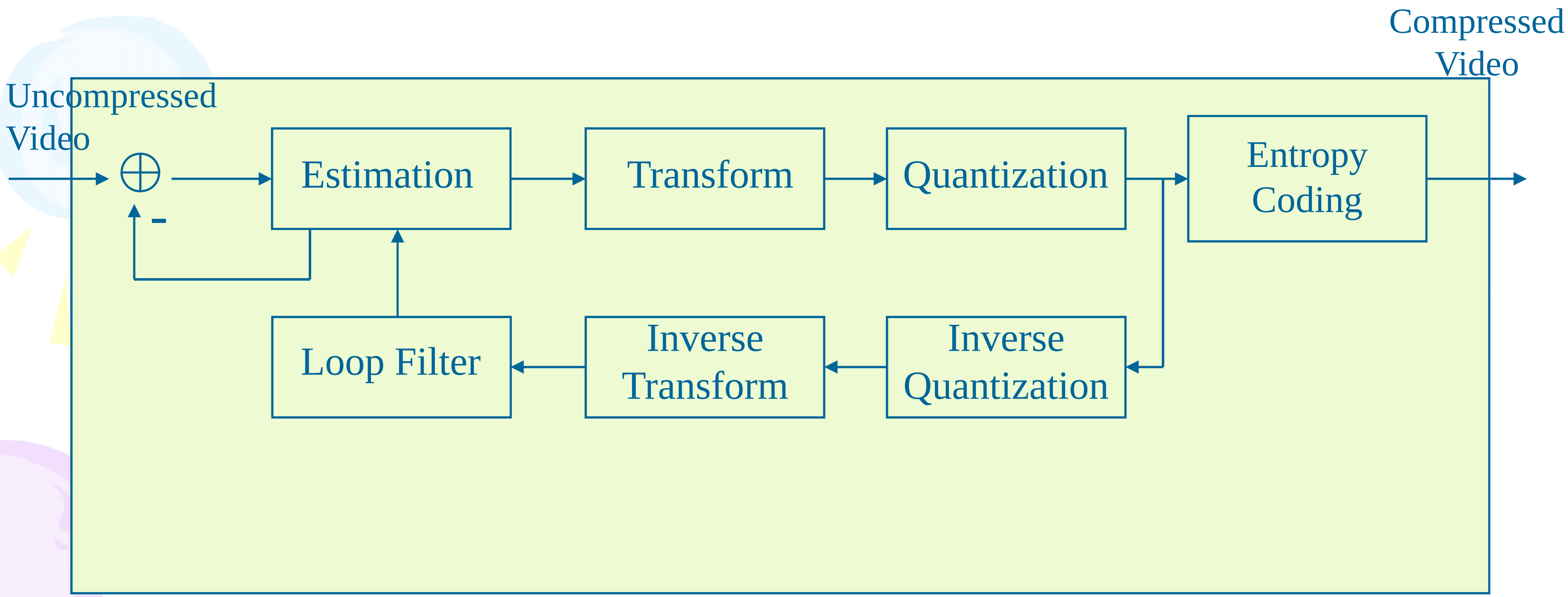
H.264 (42 dB)

H.264 / MPEG4 AVC

Network Abstraction Layer

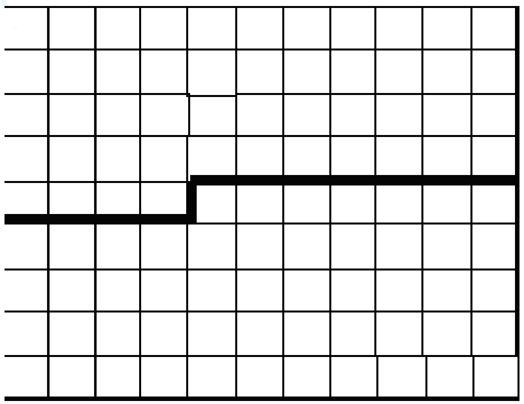
- For braod variety of applications
- *NAL units*: packets that contains an integer number of bytes
- NAL units are classified as VCL and non-VCL:
  - VCL NALs contain data representing the values of the samples in the video pictures
  - Non VCL NALs contain additional information (header or enhancement information)
- *Acces unit*: set of NAL units corrisponding to one decoded picture
- *IDR*: instantaneous decoding refresh access unit; contains an intra picture. Won't be a picture in the stream that will require a reference to any previous picture in the NAL unit stream.

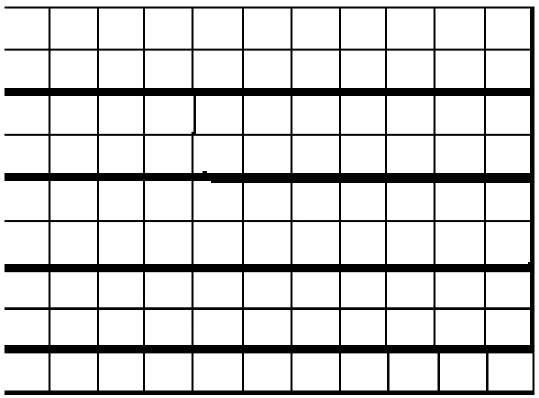
H.264 technical overview - I



Macroblocks and slices

- A picture to code is partitioned into fixed-size area of 16x16 pixels (macroblocks, MB)
- MB is the basic coding element
- MBs are grouped into slice
- A slice is defined as self-contained: can be decoded without using data from other slices (i.e. no intra prediction between slice boundaries)
- Slice size is very flexible in H.264





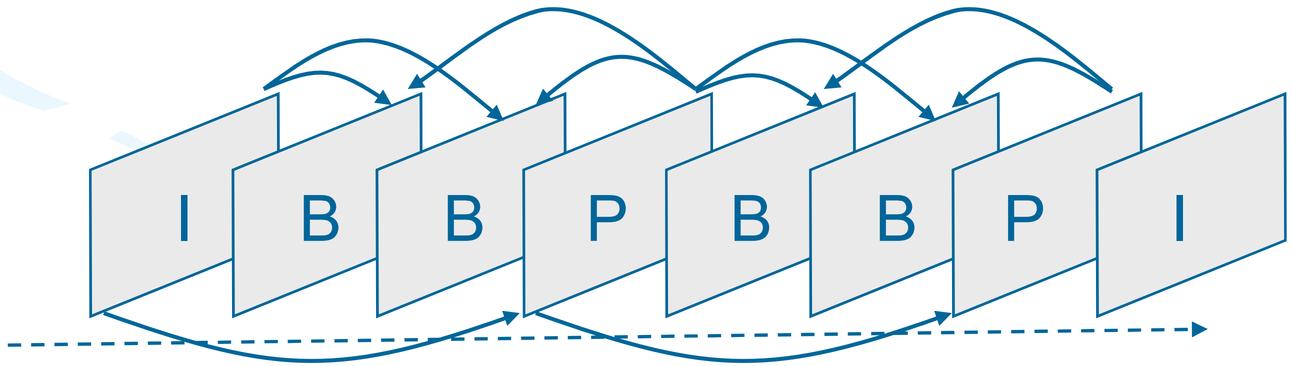
Qcif frame:

-176x144

-11x9 MB

Three Type of Frames

- I-Type:
  - Only Intra-Frame Prediction is used
- P-Type (Predictive):
  - Uni-directional Inter-Frame Prediction is used
- B-type (Bidirection Predictive):
  - Both Past & Future Frames are used for Inter-Frame Prediction

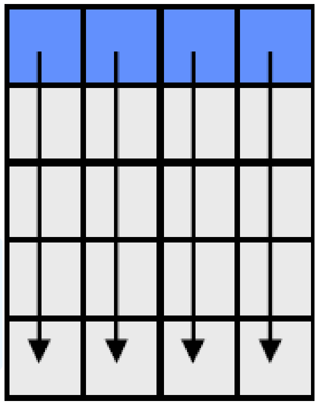


H.264 transform coding

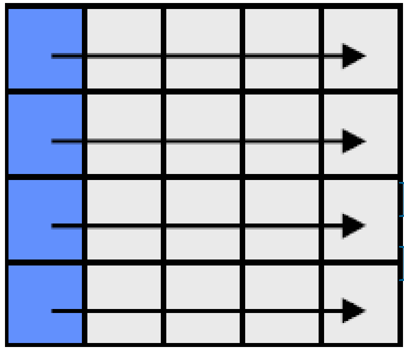
- As in previous standards transform coding of the prediction residual is used
- 4x4 Integer Transform
- blocking and ringing artifacts are reduced

Smaller transform => less computation

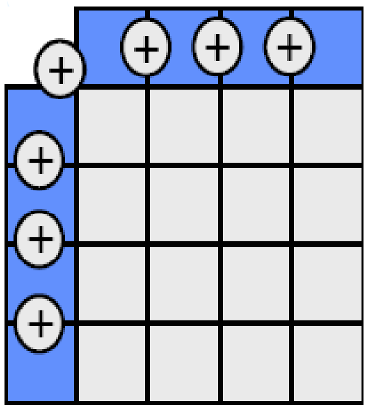
Intra frame prediction
Prediction modes (Intra_4x4)



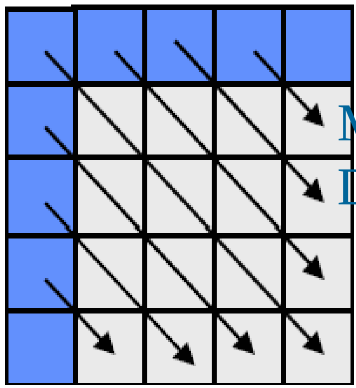
Mode 0
Vertical



Mode 1
Horizontal



Mode 2
DC



Mode 4
Diagonal Down/Right

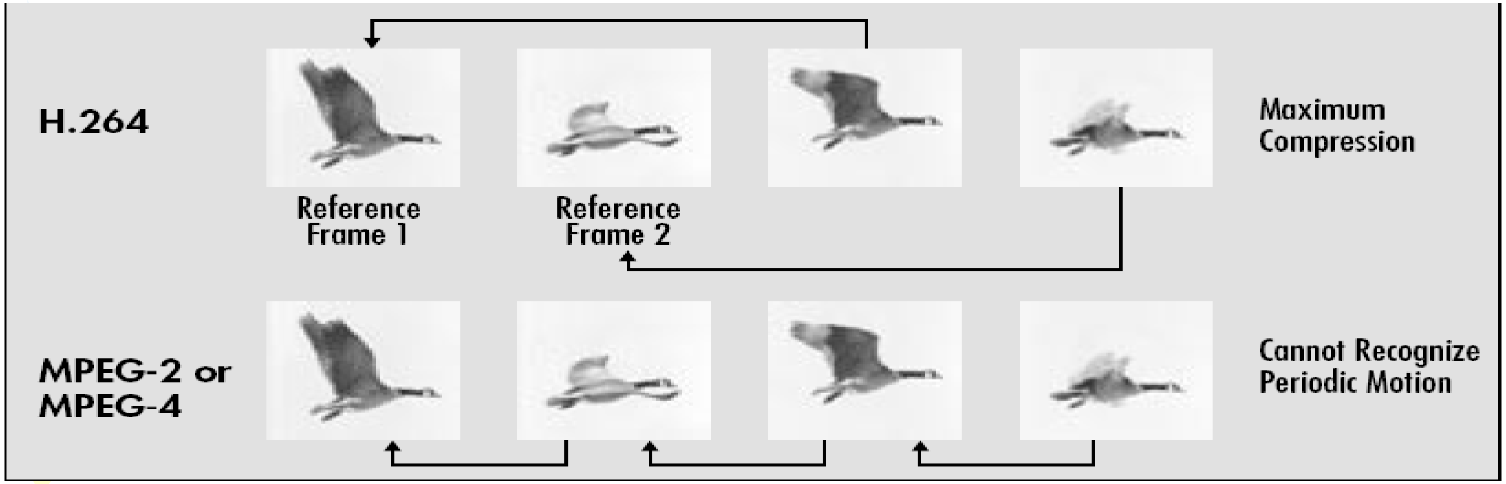
Inter frame prediction Half/Quarter pixel ME

- Variable Block size (16x16,16x8,..,8x8,4x8,...,minimun size 4x4)
- half/quarter pel sample motion estimation is enabled
- **Using *full / half / one-quarter* represents a great improvement compared to earlier standards:**

  More accurate motion rappresentation

Inter frame prediction
Multiple Reference Frames

- Multiple Ref Frames
- Useful when dealing with motion that is periodic or in presence of camera switching between 2 scenes

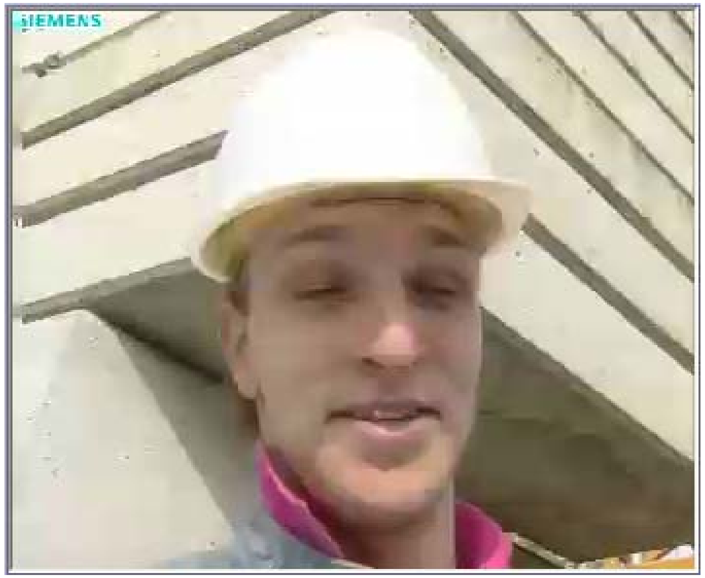


Quantization

- Reduces precision of integer coefficients and tends to eliminate high frequency coefficients
- A quantization parameter (*qp*) is used to determine perceptual quality (52 values)

qp 10

4.2 Mbps



qp 36

160 Kbps

Deblocking filter

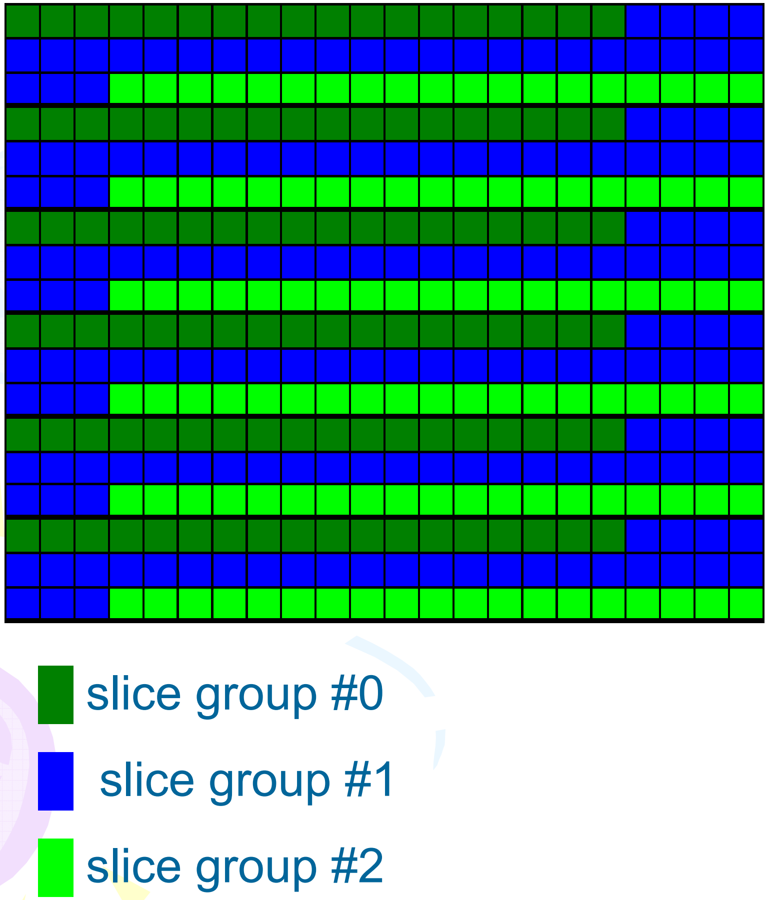
- In loop deblocking filter
- The strength of the filter is controlled by the value of several syntax element (*qp*)

No deblock

Deblock

Flexible macroblock ordering

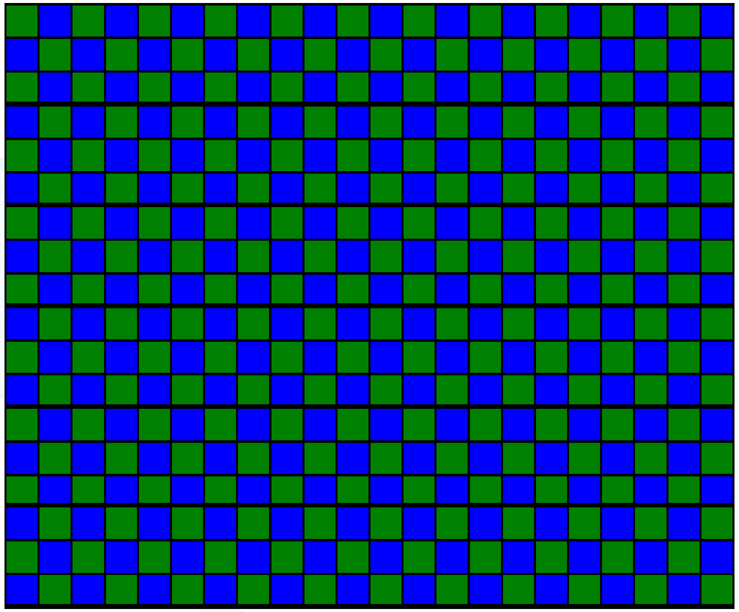
Interleaved slices map group



- CIF frame (22x18 MBs)
- 3 Slice groups
- Run length slice group 0 -> 18
- Run length slice group 1 -> 29
- Run lengtht slice group 2-> 19

Flexible macroblock ordering

Scattered slices map group

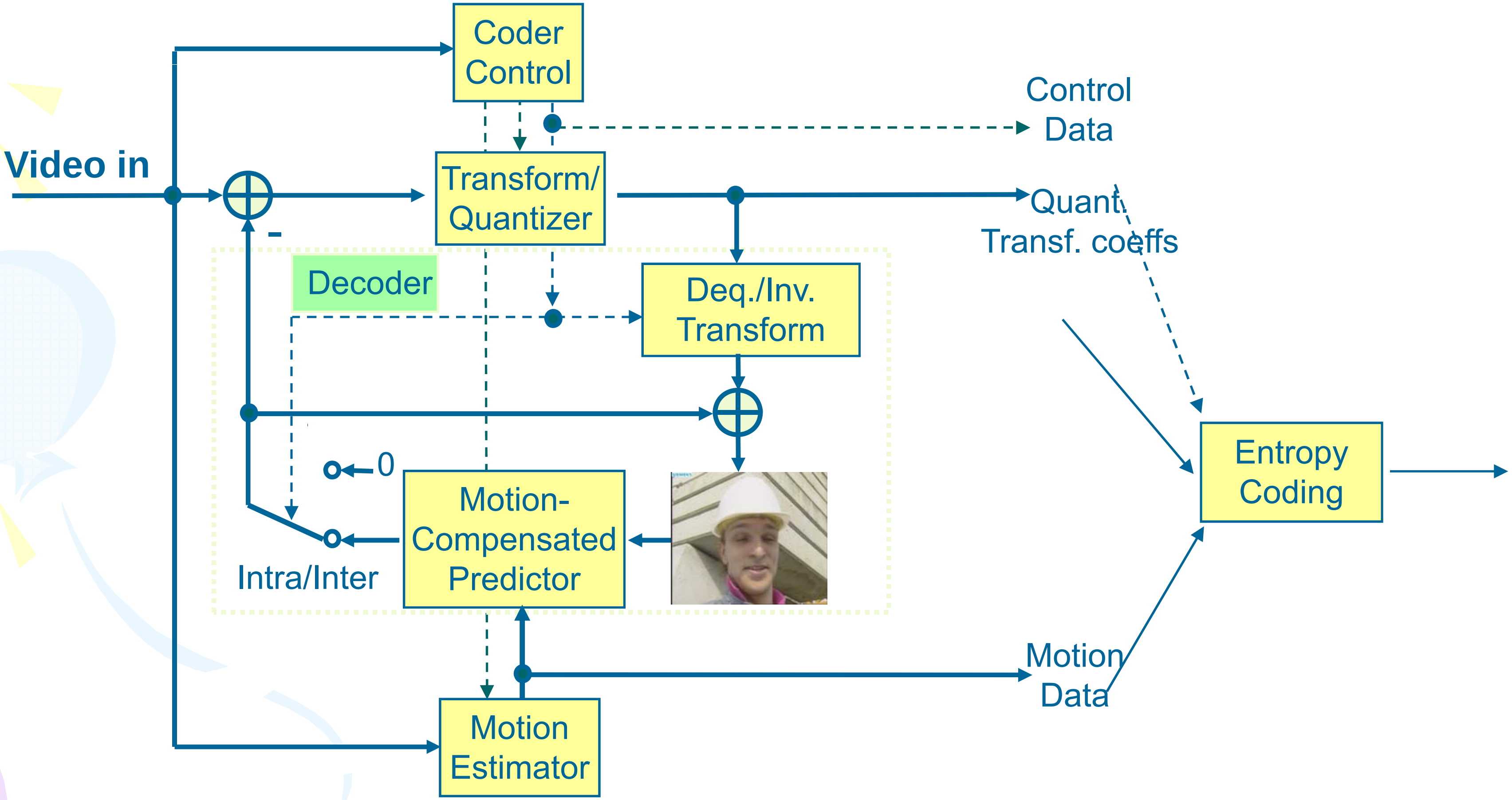


=> slice group #0

=> slice group #1

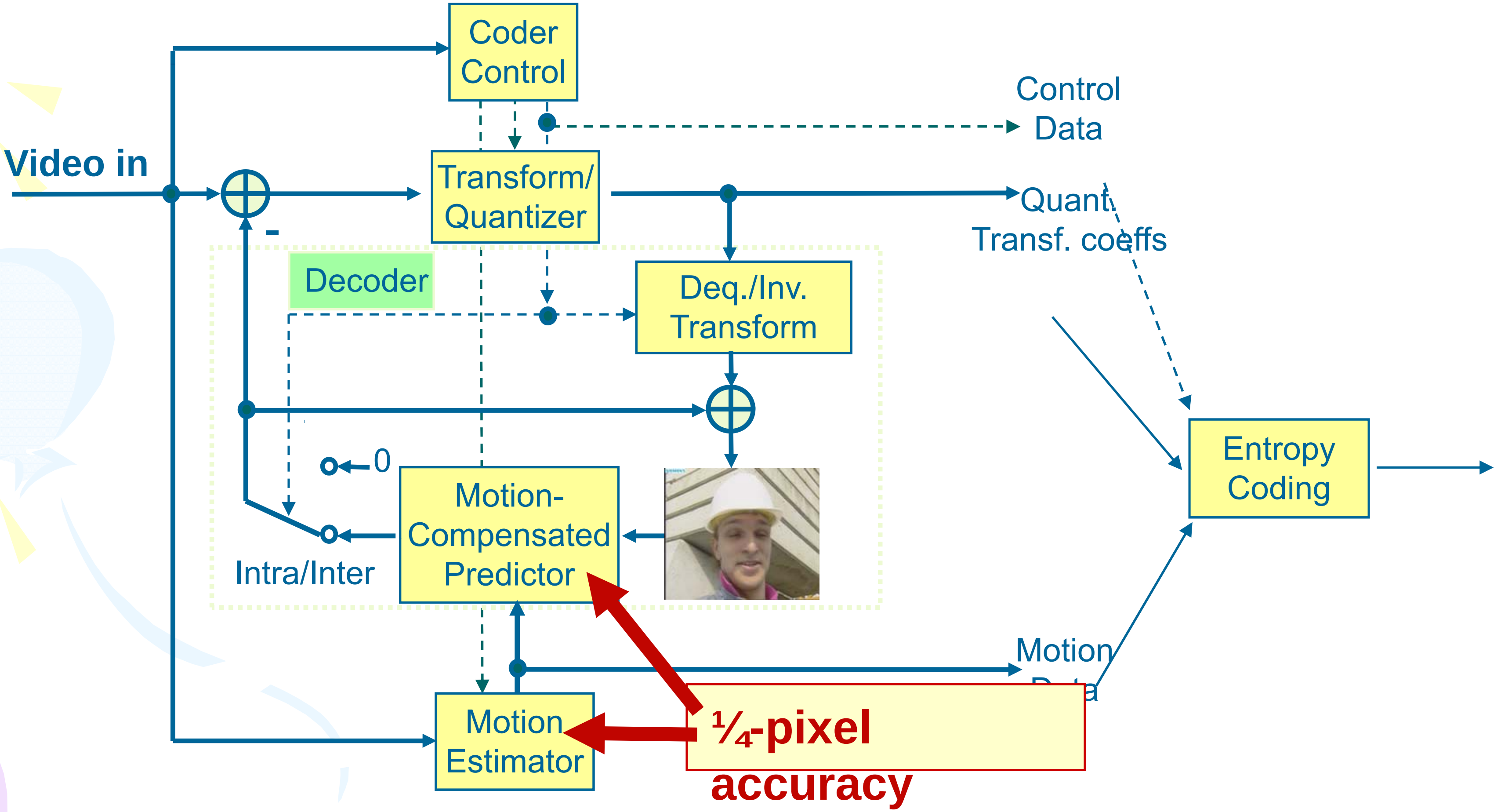
- CIF (22x18 MBs)
- # of slice groups -> 2

H.264 features

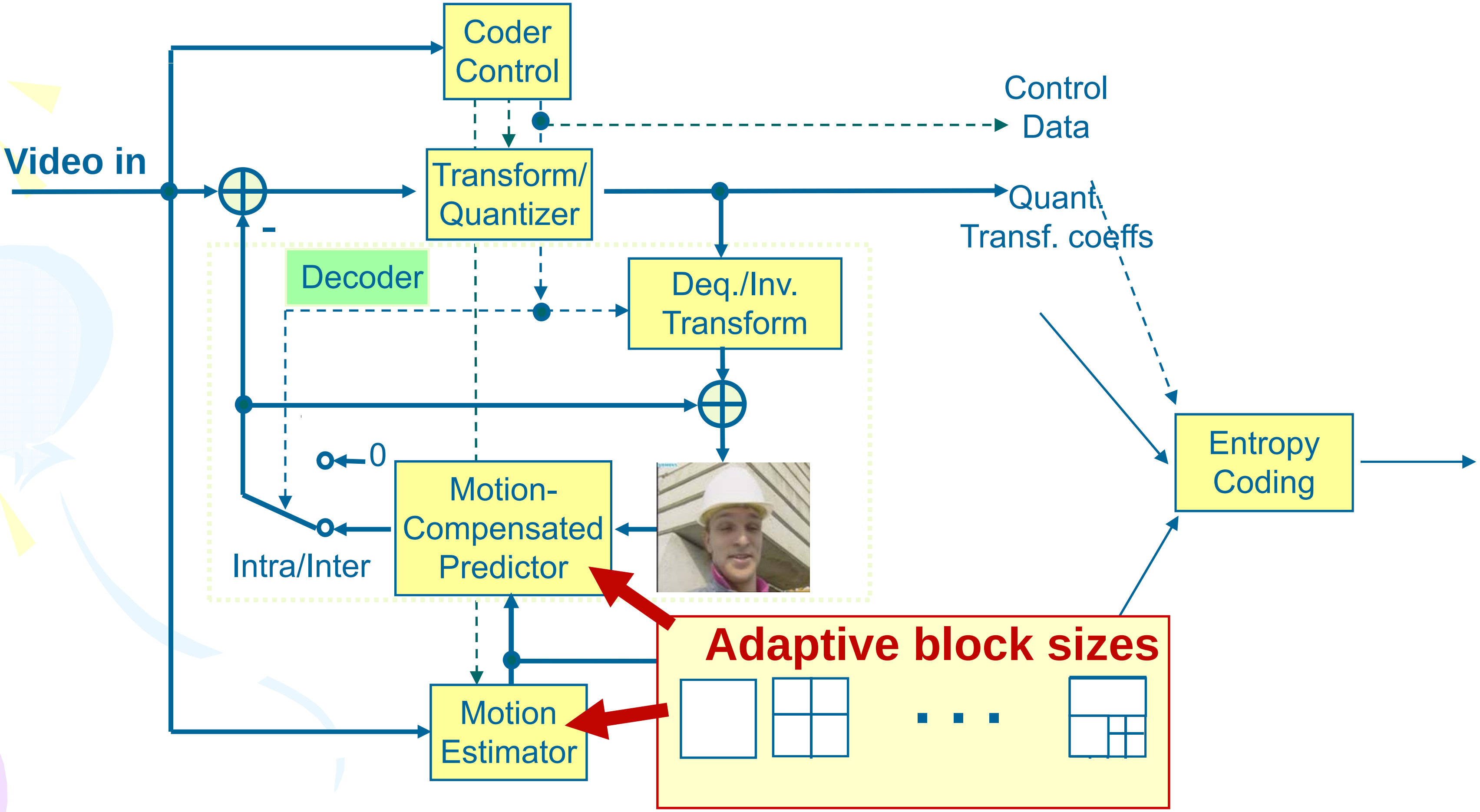


<u>Standards:</u> H.261, MPEG-1, MPEG-2, H.263, MPEG-4, H.264/AVC

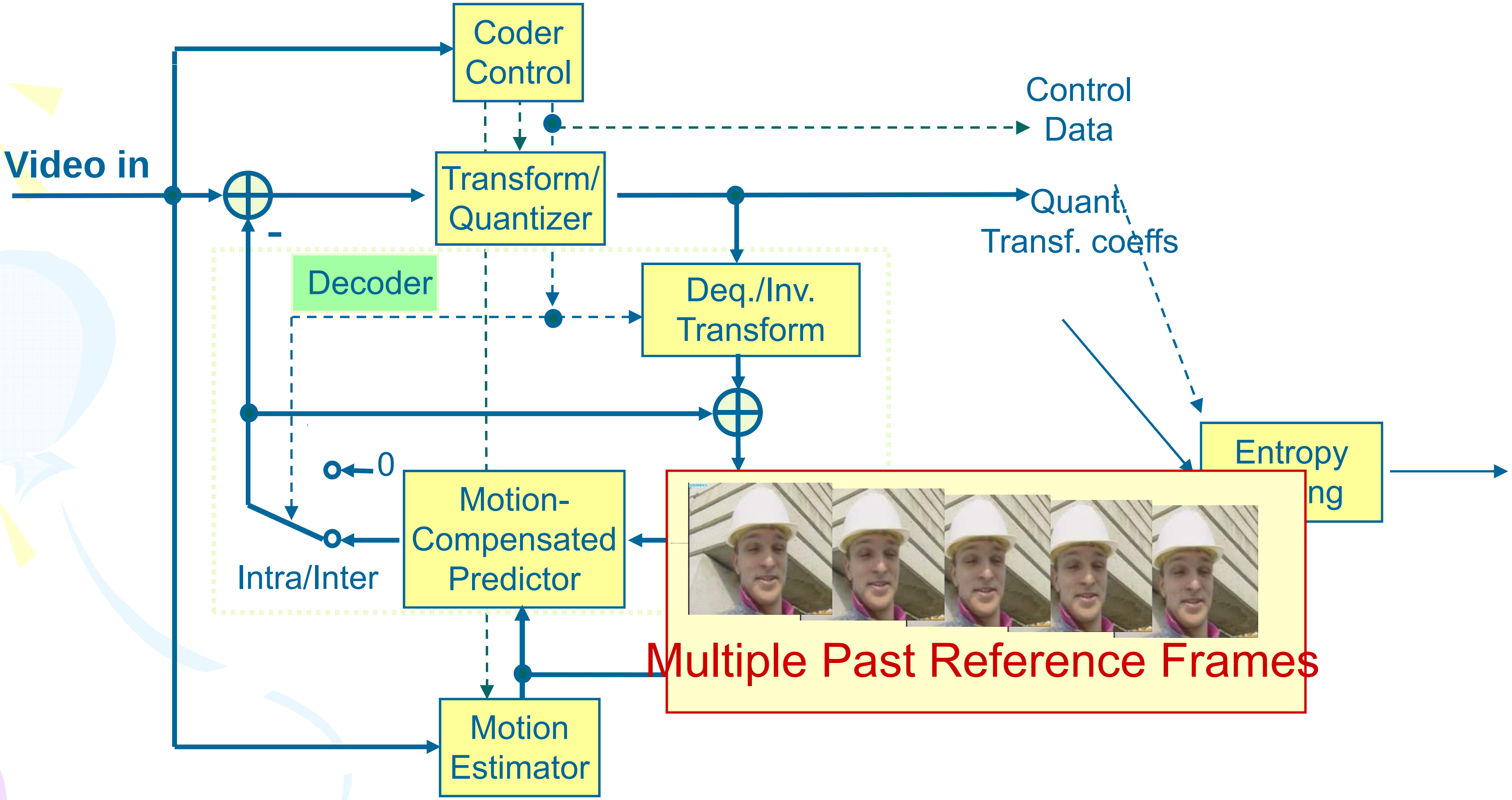
H.264 features



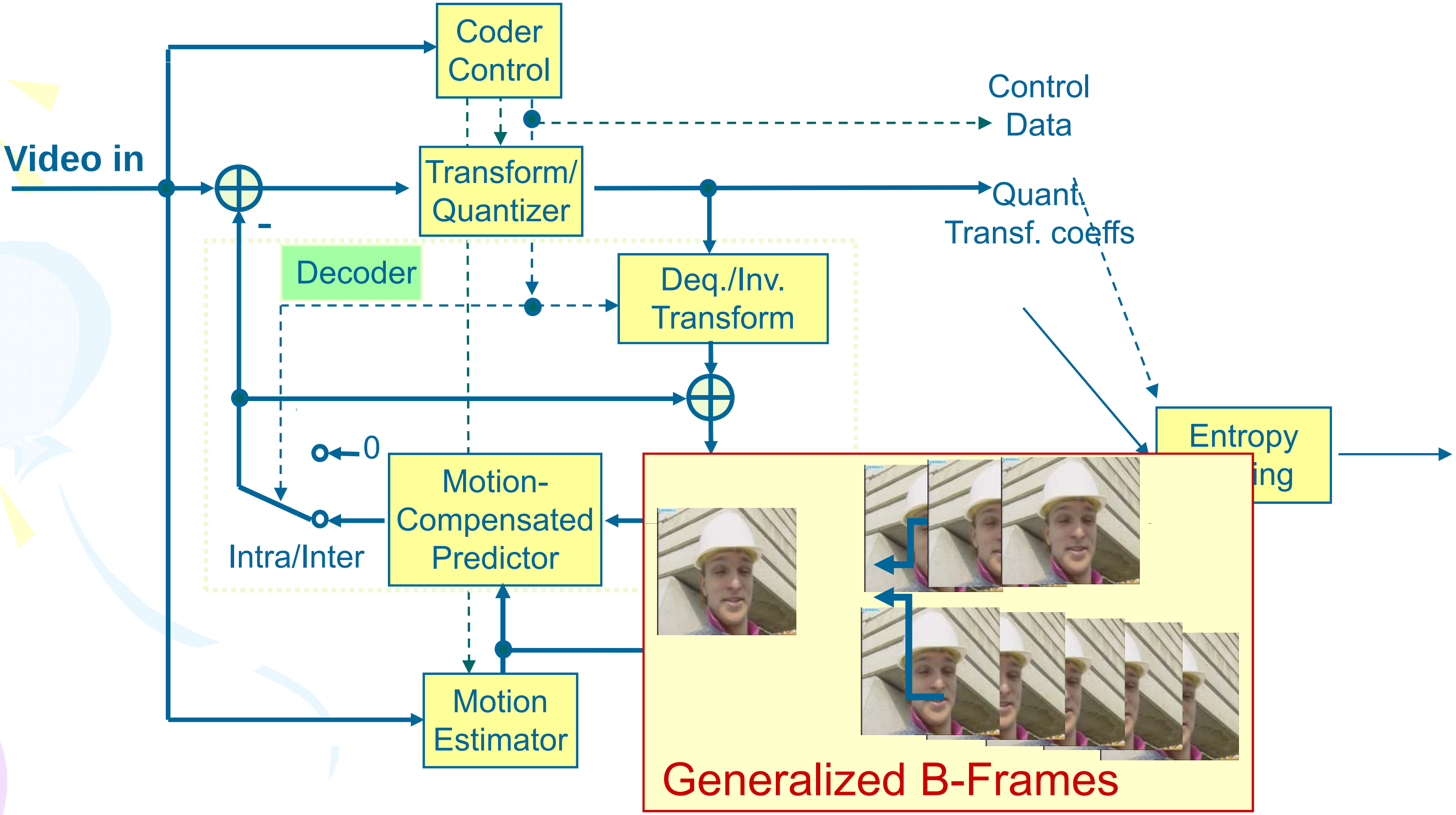
H.264 features



H.264 features



H.264 features



Conclusion

- **H.264/AVC offers:**
  - significant bit rate and quality
  - Better network adaptation

- **Important differences:**
  - Enhanced motion prediction capability
  - Use of a small block-size, exact-match transform
  - Adaptive in-loop deblocking filter
  - Enhanced entropy coding methods

Example : *foreman video sequence (100 frames)*

QP 18, 813 kB, 44 dB

QP 24, 479 kB, 39 dB

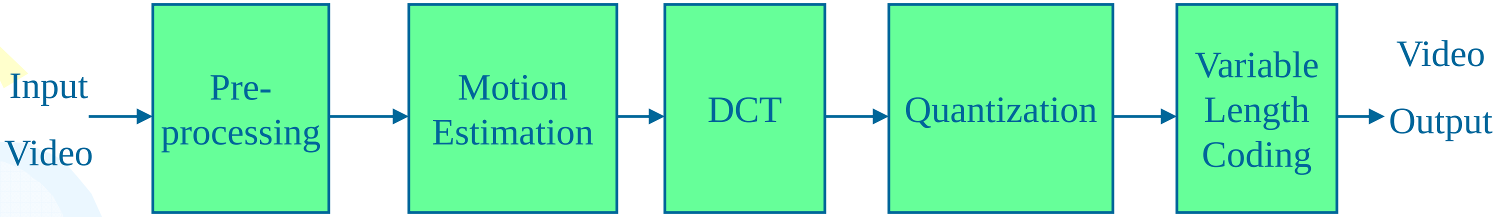
QP 30, 268 kB, 35 dB

QP 36, 145 kB, 31 dB

LIRMM

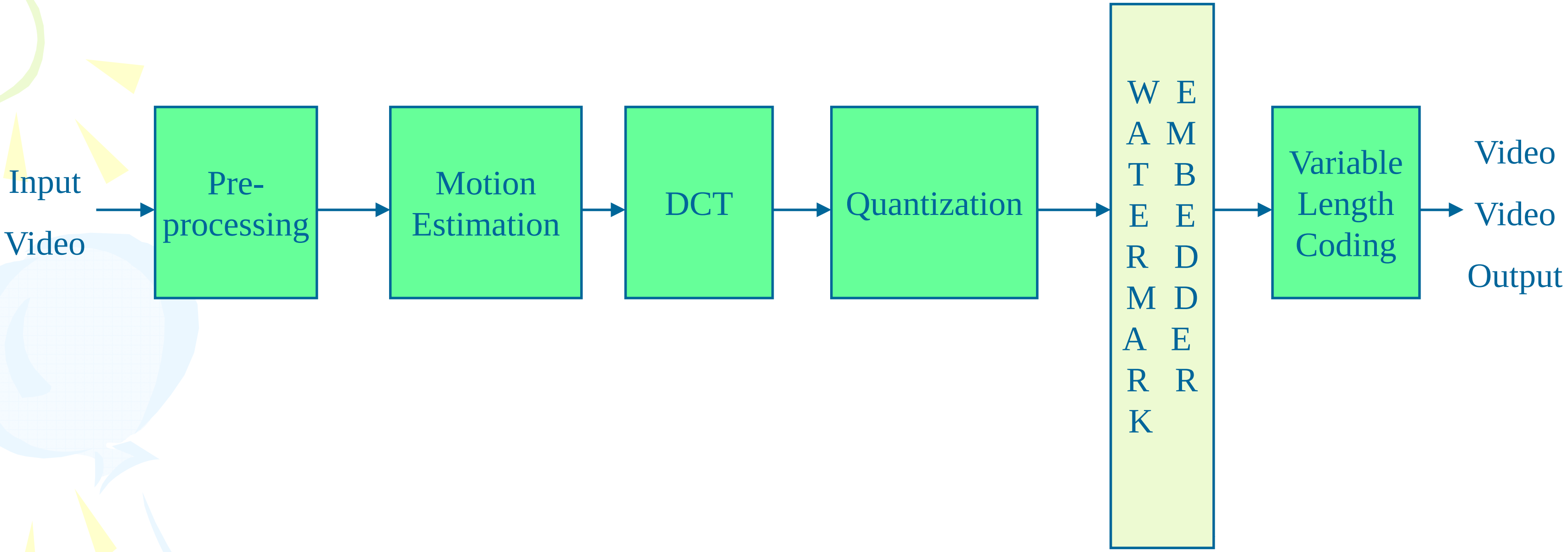
Watermarking for H.264/AVC

Video Encoder Overview



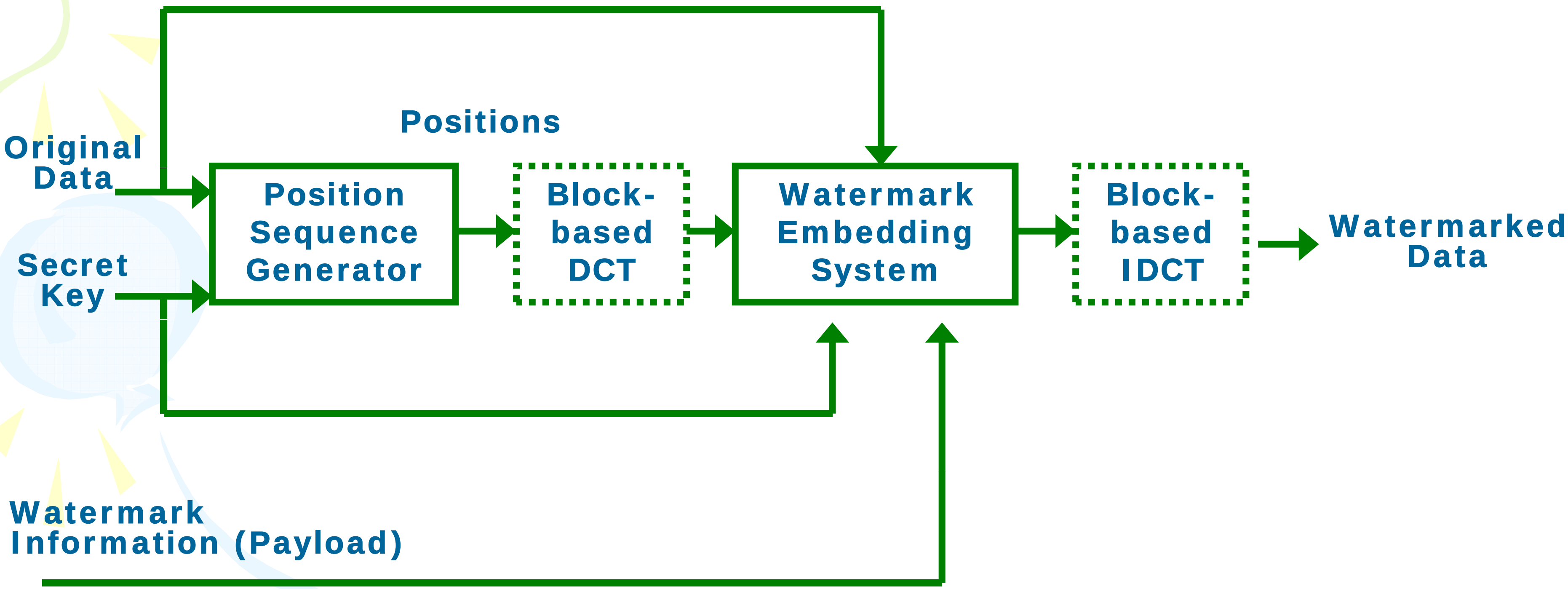
Example of standard Video encoder

Video Encoder with Watermark Insertion



Video encoder with Watermark Inserter

Watermark Embedding



Watermark insertion using DCT coefficients

LIRMM

DCT Coefficients (post Quantization)

212	65	0	0	312	785	0	0
0	0	0	244	72	**16**	**17**	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

→ WATERMARK EMBEDDER →

DCT Coefficients (post Watermarking)

212	65	0	0	312	785	0	0
0	0	0	244	72	18	17	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Watermark detection using DCT coefficients

LIRMM

DCT Coefficients (post inverse quantization)

212	65	0	0	312	785	0	0
0	0	0	244	72	18	17	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

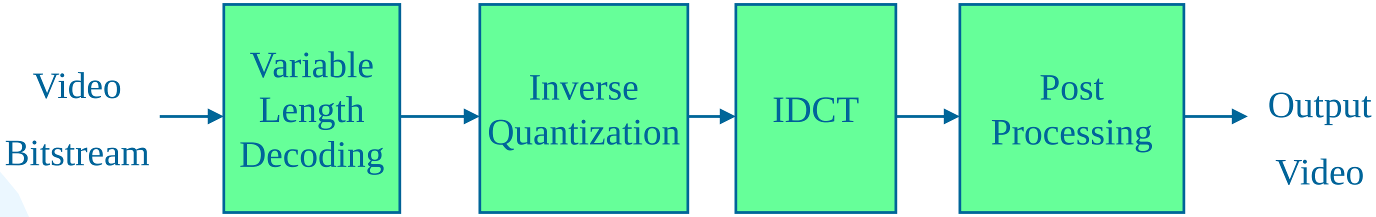
→ DETECTOR →

Extracted watermark payload bits

is based on the ratio of the two highlighted

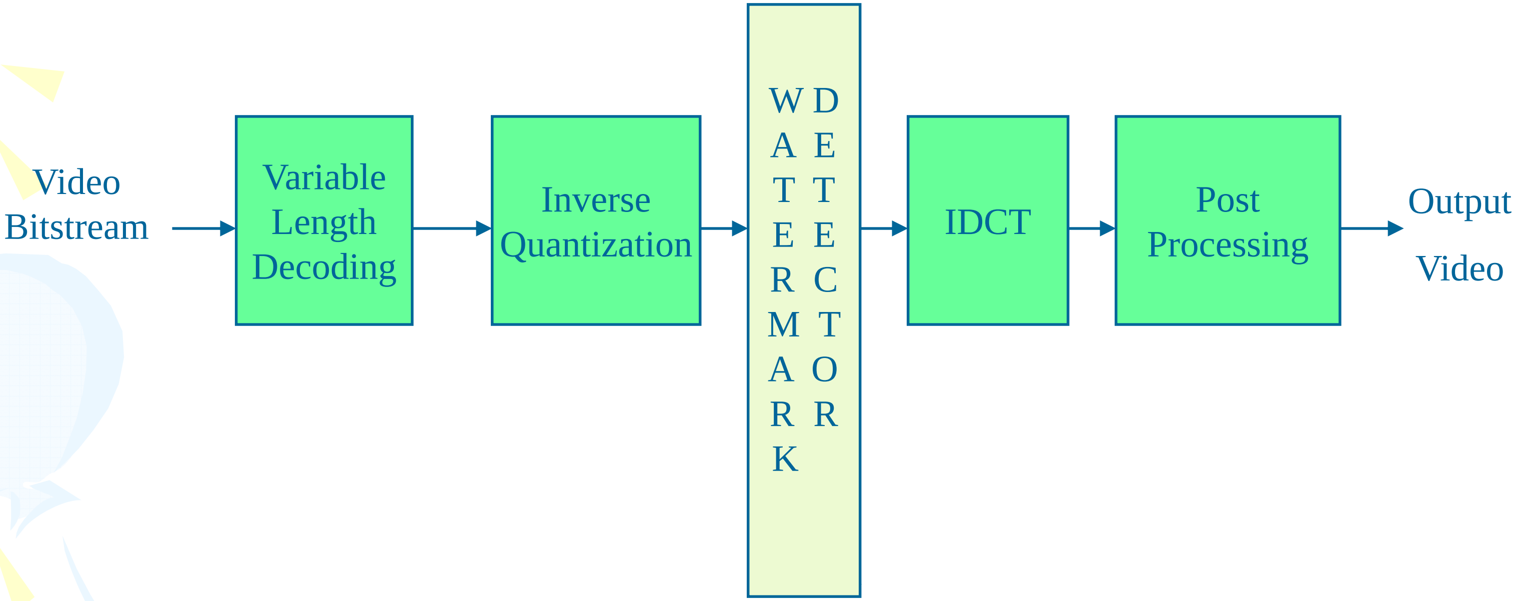
DCT coefficients

Video Decoder Overview



Standard Video decoder

Video Decoder with Watermark Detector



Video decoder with Watermark Detector

Thank you!