Recent Advances in Video Coding

Thomas Wiegand

Heiko Schwarz

TU berlin

Fraunhofer
Heinrich Hertz Institute

Overview

- Video Coding – State-of-the-Art
- High-Efficiency Video Coding
  - Motion representation
  - Motion compensation
  - Transform coding
  - Intra coding
  - Entropy Coding

[Second Edition.]

PATENT SPECIFICATION

Convention Date (United States): April 25, 1929. **341,811**

Application Date (in United Kingdom): April 25, 1930. No. 12,805/30.

Complete Specification Accepted: Jan. 22, 1931.

COMPLETE SPECIFICATION.

Improvements relating to Electric Picture Transmission Systems.

We, THE BRITISH THOMSON-HOUSTON COMPANY LIMITED, a British Company, having its registered office at Crown House, Aldwych, London, W.C. 2, (Assignees of RAY DAVIS KELL, of 111, Sanders Avenue, Scotia, County of Schenectady, State of New York, United States of America, a citizen of the United States of America), do hereby declare the nature of this invention and in what manner the same is to be performed, to be particularly described and ascertained in and by the following

fineness of detail is limited only by the speed of the action to be transmitted.

The invention will be better understood from the following description when considered in connection with the accompanying drawings in which Fig. 1 illustrates a picture transmitting apparatus wherein the invention has been embodied; and Figs. 2 to 5 illustrate various details of an apparatus which may be utilised to receive the difference between the successive images of a picture or moving object.

"It has been customary in the past to transmit successive complete images of the transmitted picture."

[...]

"In accordance with this invention, this difficulty is avoided by transmitting only the difference between successive images of the object."

The Digital Media Revolution

1989: Digital TV – Digital Broadcast, DVD

1999: TML-1 of H.264/AVC

DVB Digital Video Broadcasting — Every HDTV Receiver

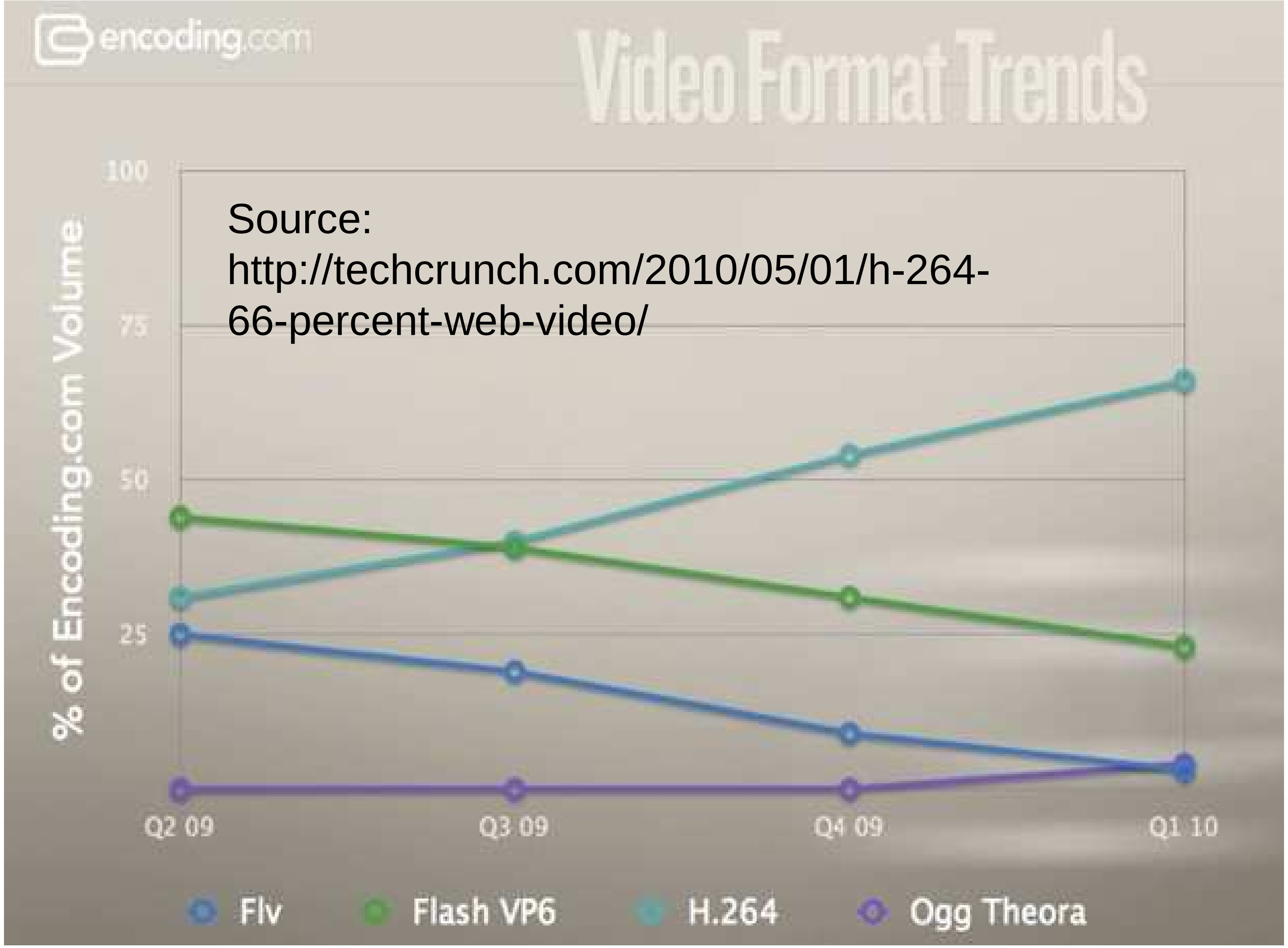
Blu-ray Disc — Every Blu-Ray Player

IETF — Most Internet Video

3GPP A GLOBAL INITIATIVE — Countless Mobile Video

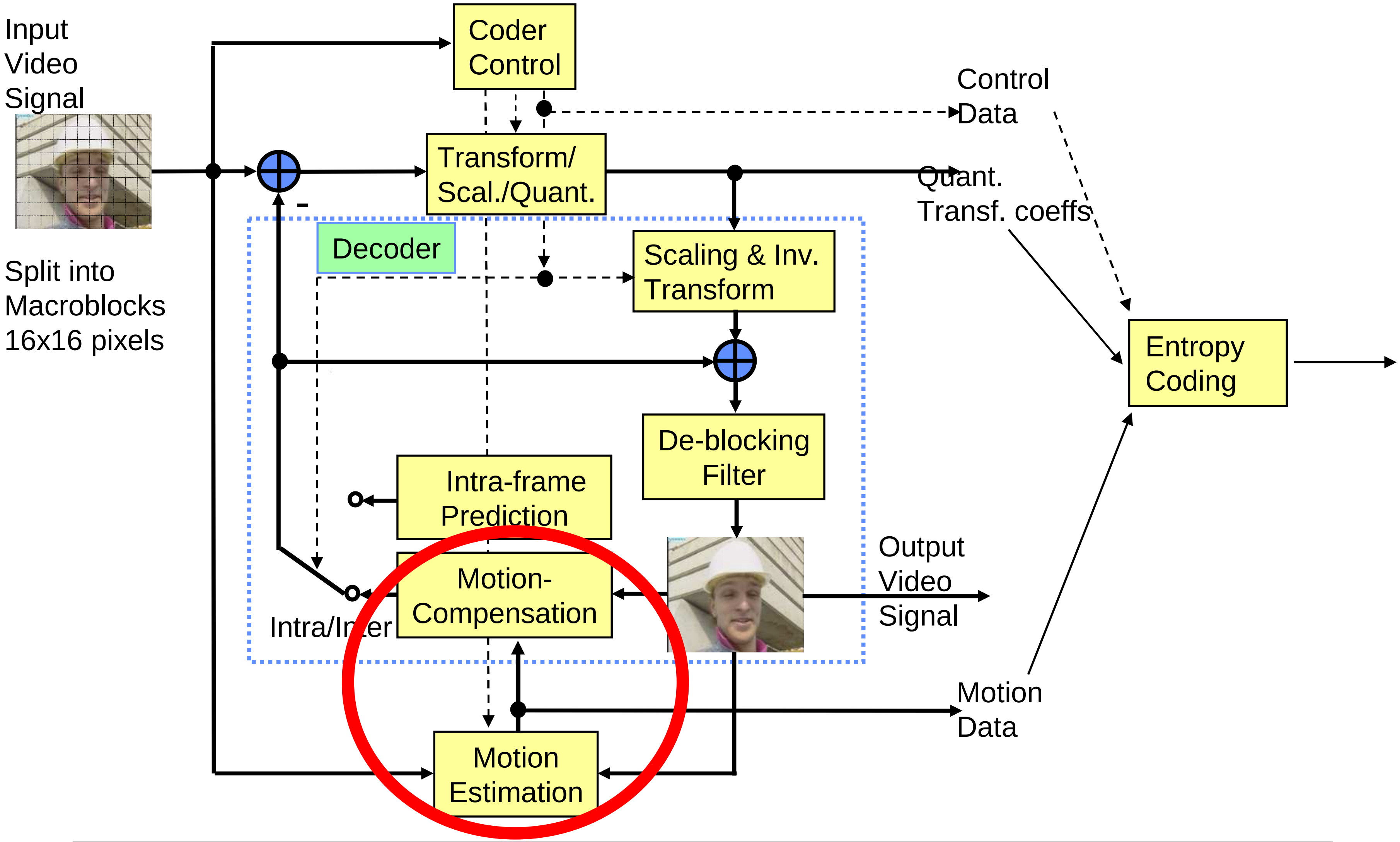
Today: more than 1 Billion devices with H.264/AVC

H.264/AVC: 66% of Web Video



Source:
http://techcrunch.com/2010/05/01/h-264-66-percent-web-video/

Motion-Compensated Hybrid Coding



Efficiency of Motion Compensation

- Intraframe coding: only spatial correlation exploited
  ➔ DCT [Ahmed, Natarajan, Rao 1974], JPEG [1992]

- Motion compensation (MC): integer-pel accuracy, 16x16
  ➔ H.261 [1991]

- Half-pel accurate MC
  ➔ MPEG-1 [1993], MPEG-2/H.262 [1994]

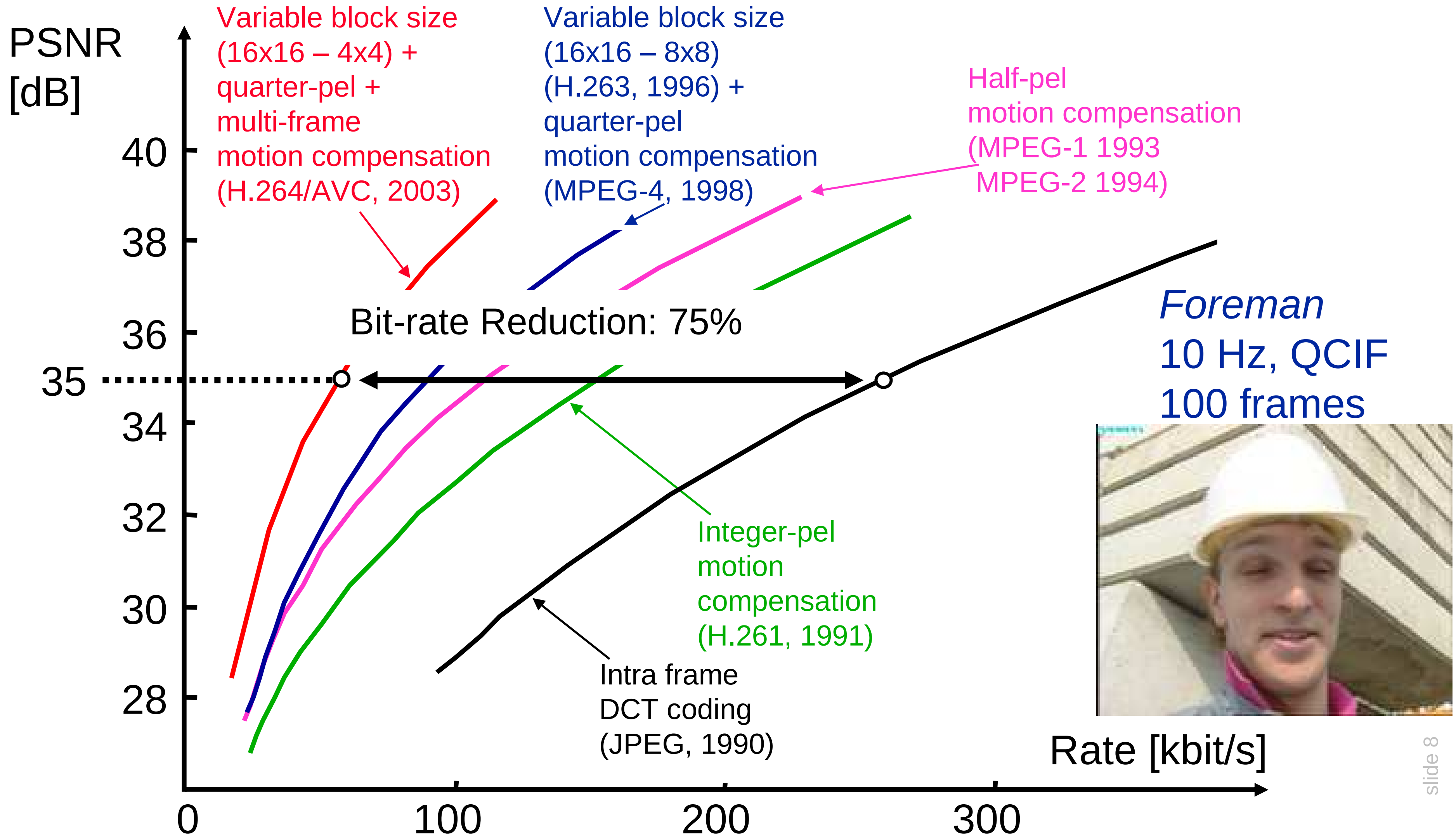
- 16x16 & 8x8 block size MC, quarter-pel
  ➔ H.263 [1996], MPEG-4 [1999]

- Quarter-pel & 16x16 - 4x4 block size MC and multiple reference pictures
  ➔ H.264/AVC [2003]

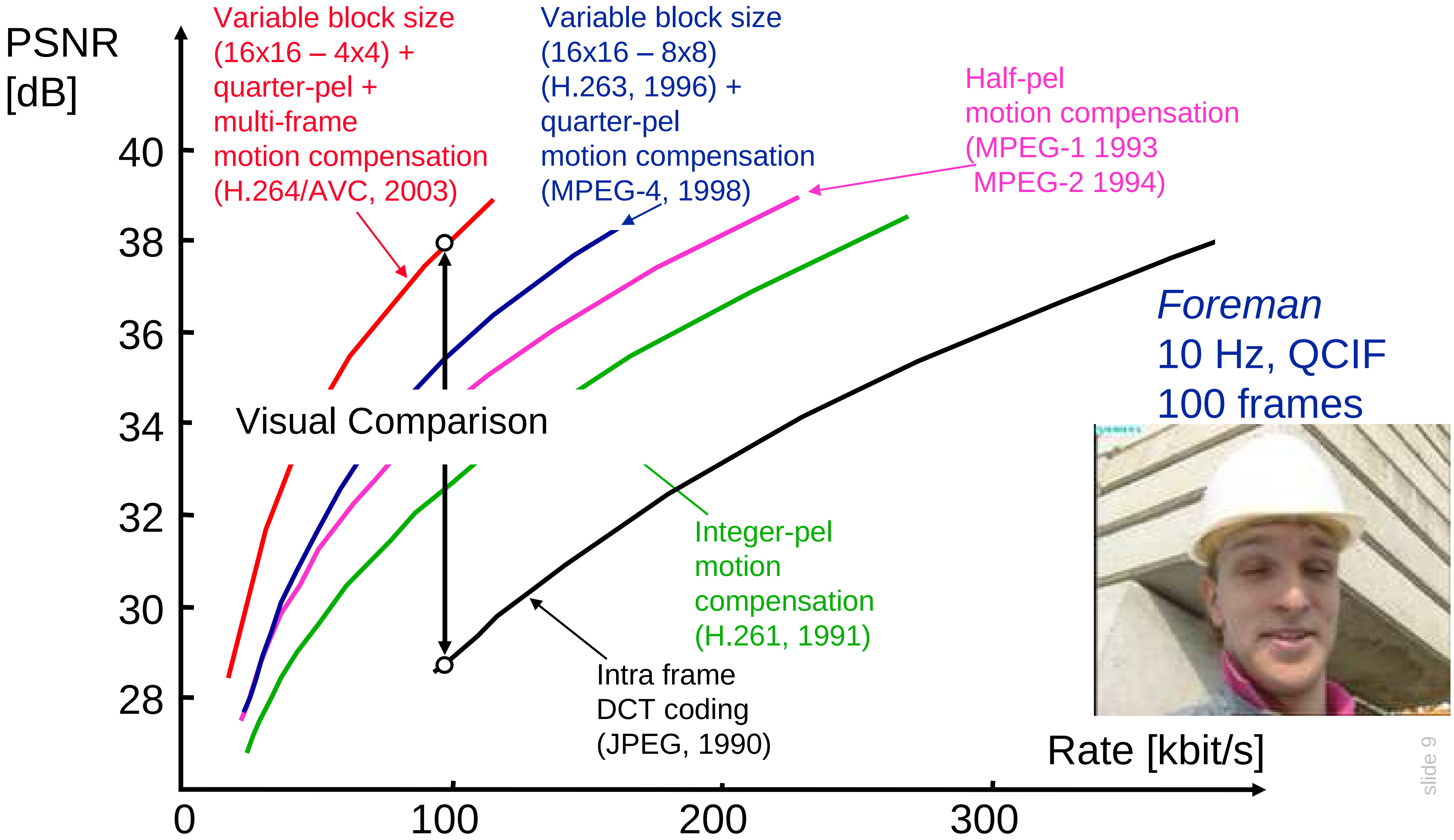
- 64x64 – 4x4 block sizes and new filters for MC
  ➔ HEVC [start 2010]

Complexity increase

Development of Video Coding



Development of Video Coding



Visual Comparison

Foreman, QCIF, 10 Hz, 100 kbit/s

JPEG

H.264/AVC

slid

Rate-Constrained Coder Control

- Video encoder needs to decide between many options denoted with parameter vector $\boldsymbol{p}$

- Constrained Problem:

$$\min_{\boldsymbol{p}} D(\boldsymbol{p}) \quad \text{s.t.} \quad R(\boldsymbol{p}) \le R_T$$

D - Distortion

R - Rate

R_T - Target Rate

$\boldsymbol{p}$ - Parameter Vector

- Unconstrained Lagrangian Formulation:

$$\min_{\boldsymbol{p}} D(\boldsymbol{p}) + \lambda \cdot R(\boldsymbol{p})$$

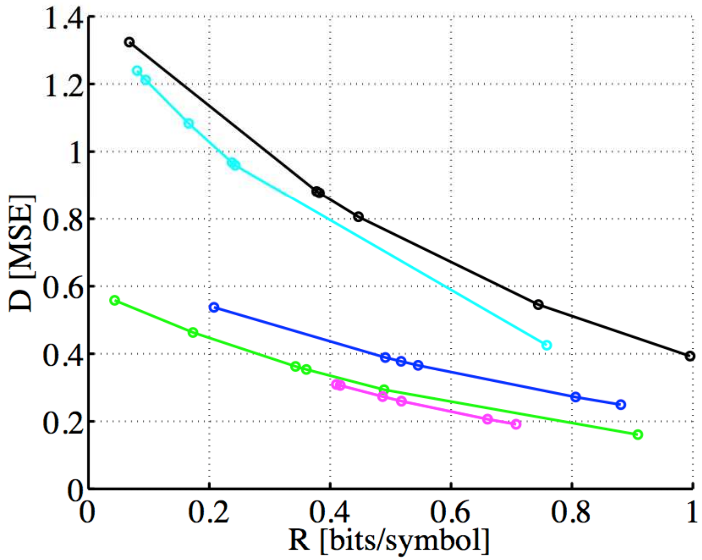
with λ controlling the rate-distortion trade-off

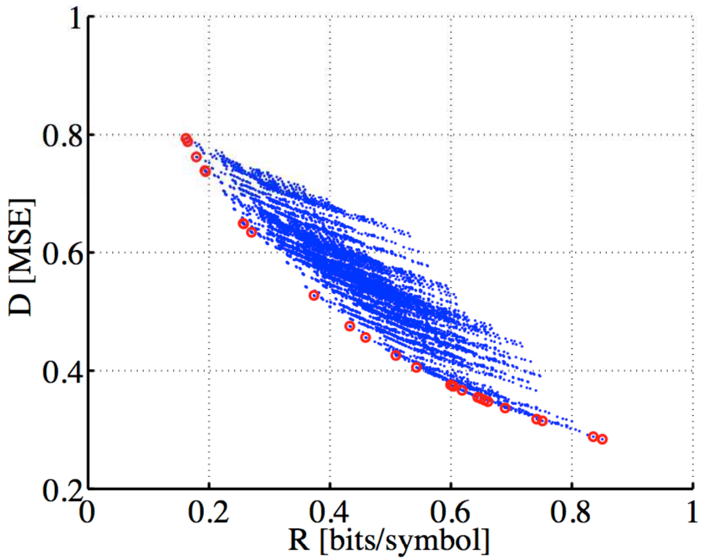
- Minimization tests the various modes in video coding

[Wiegand, et al., 1996]

Lagrangian Optimization

- 5 different quantizers with $D_i(R) = a_i^2 \sigma^2 2^{-2R}$ with a_i^2 and rate points R being randomly chosen

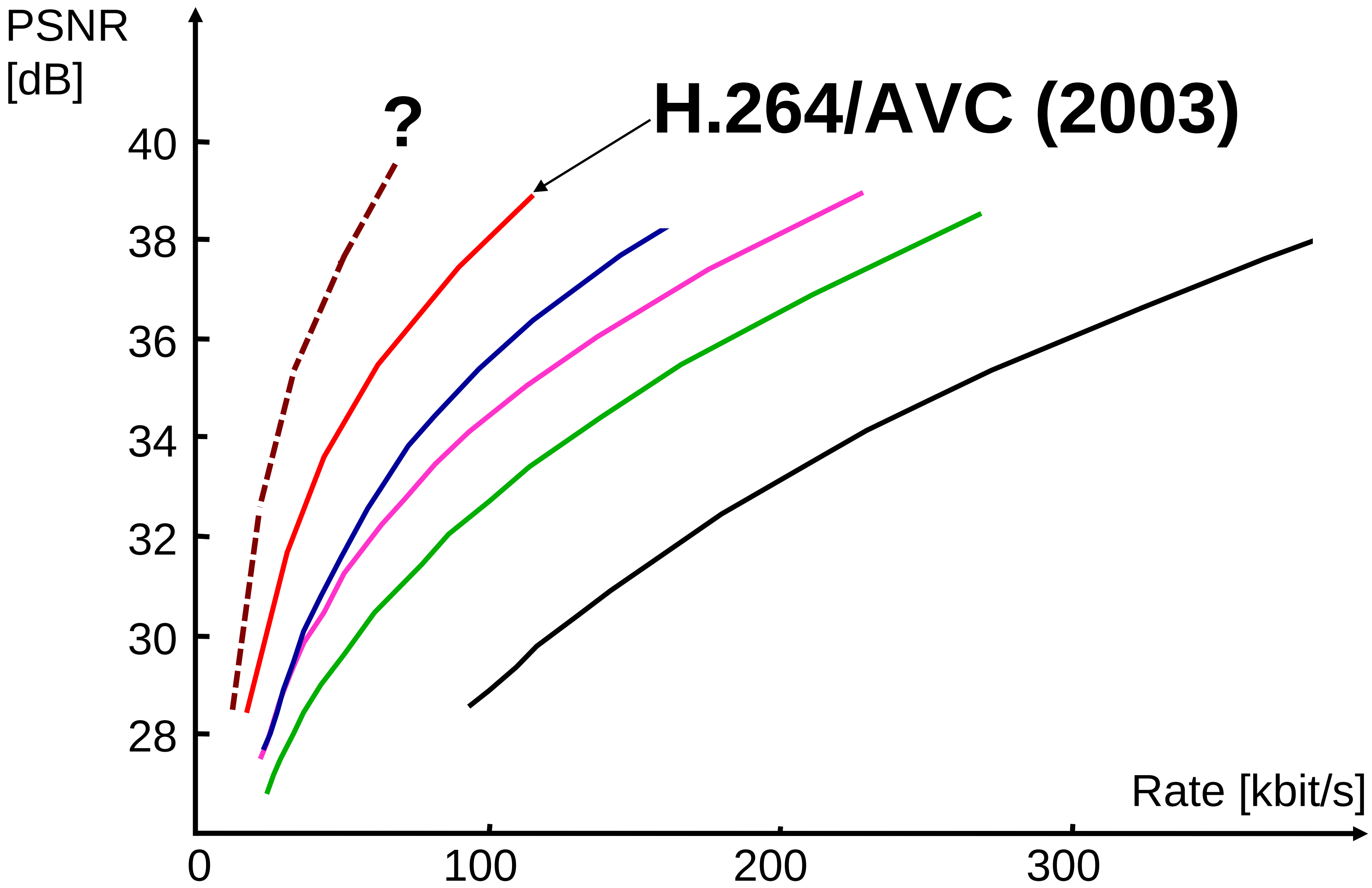




- LHS: Each circle is one of 6 rd points that is available for each quantizer
- RHS: Blue dots show mean distortion and rate for all possible combinations of the 5 different quantizers with their 6 rate distortion points
- RHS: Red circles show the solutions to the Lagrangian minimization problem

[Shoham & Gersho, 1989]

Next Steps in Video Coding



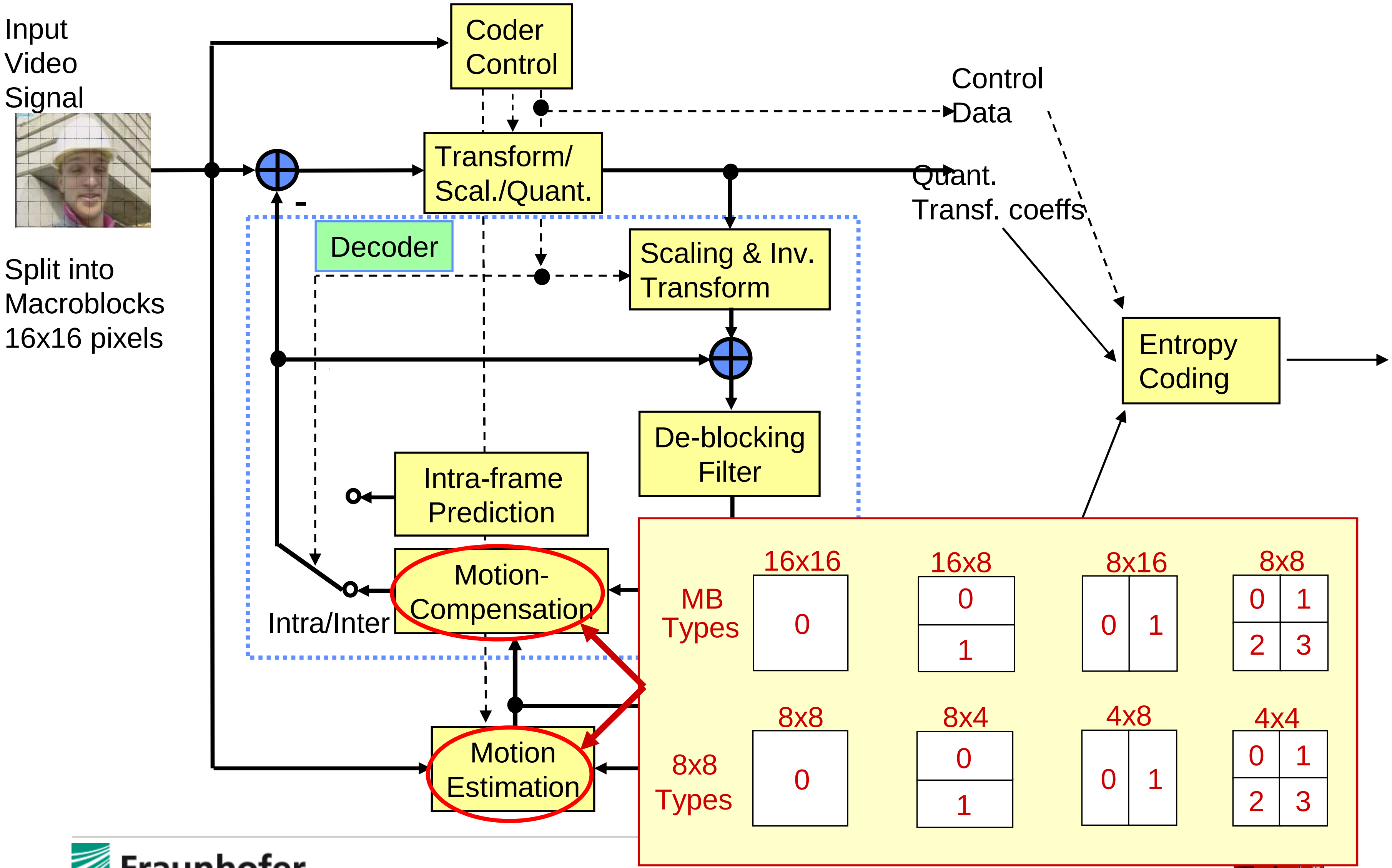
High Efficiency Video Coding

Fraunhofer
Heinrich Hertz Institute

High-Efficiency Video Coding

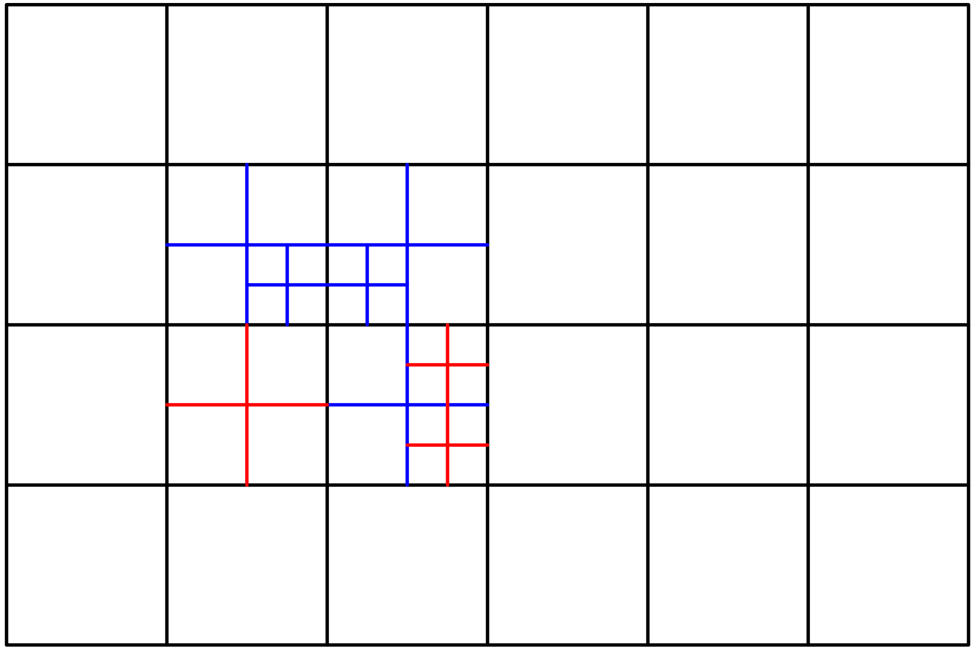
- New Joint Collaborative Team on Video Coding (JCT-VC) of ISO/IEC MPEG and ITU-T VCEG
- April 2010: Call for Proposals evaluated by JCT-VC
  - 27 proposals received
  - Significant video quality improvements relative to H.264/AVC High Profile for high and low delay
  - JCT-VC decided to create Test Model under Consideration (TMuC)
- October 2010: Creation of first Test Model
- Early 2013: intended finalization of standard

Motion Compensation Block Sizes

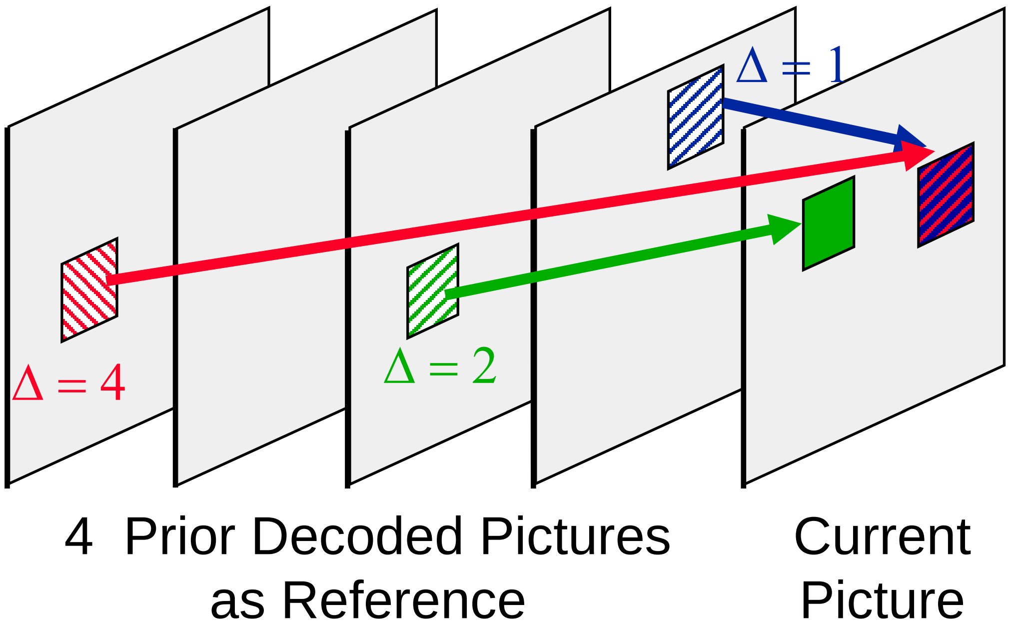


Picture Partioning

- Division of a picture into square blocks
- Blocks are assigned to quadtrees
- Maximum block size is signalled (e.g. 64x64)
- Quadtree-based subdivision of tree block into prediction and transform blocks
- Rectangular shapes
- Geometric partitions



Multiple Reference Pictures and Generalized B Pictures



1. Extend motion vector by reference picture index Δ
2. Provide reference pictures at decoder side
3. In case of bi-predictive coding (B pictures): decode 2 sets of motion parameters

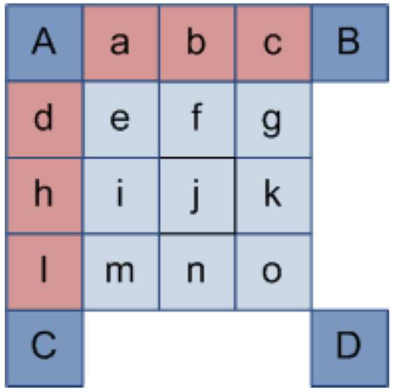
Flexible buffering of reference pictures generalizes B pictures:

- B pictures can be reference pictures (decoupling of concepts)
- Reference picture for a B picture can be any prior decoded picture

[Wiegand, et al., 1999]

Motion-Compensated Interpolation

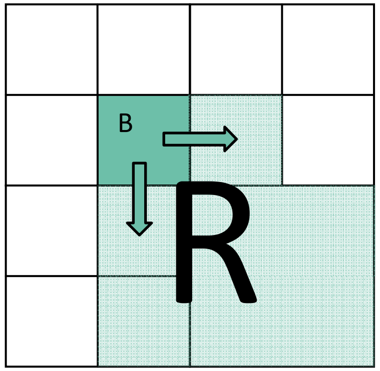
- **Longer Interpolation Filters**
  - In H.264/AVC: 6-tap filters + bi-linear
  - In HM: 12-tap interpolation filters
- **Generalized Interpolation**
  - MOMS basis functions
  - IIR pre-filtering of reference image (2-tap)
  - Short FIR filters (4-taps)

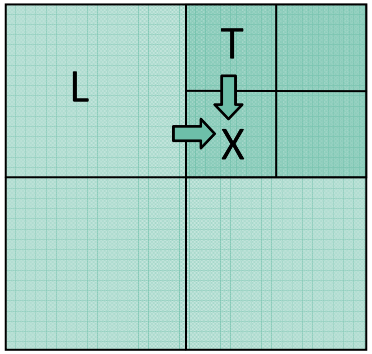


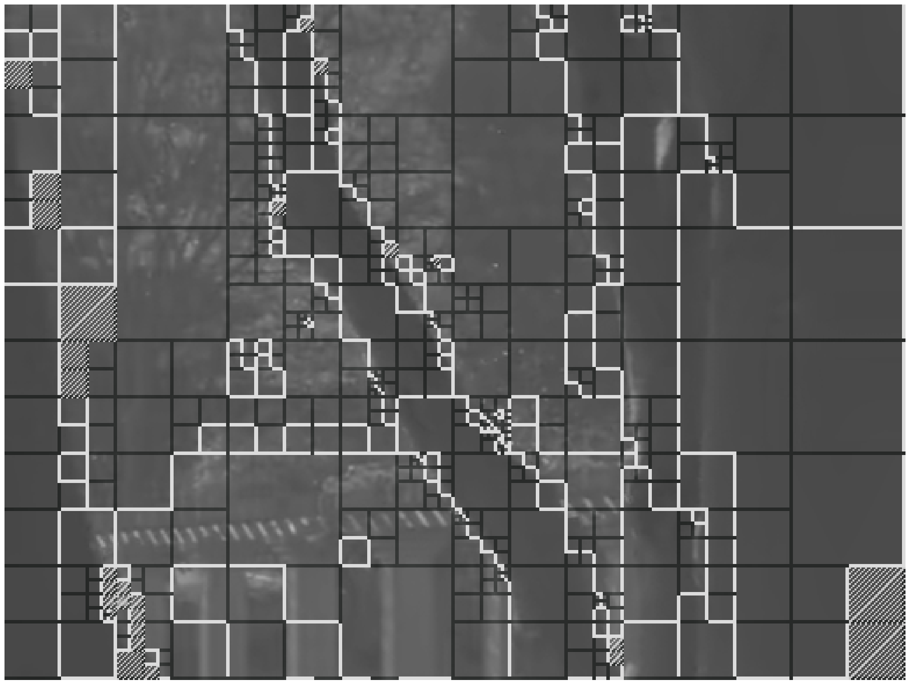
- **Adaptive Interpolation filters**
  - Transmission of filter coefficients in slice header
  - Transmission of an index into a list of predefined filters
- **Directional interpolation filters**
  - Low-complexity interpolation

Motion Representation

- **Quadtree-based splitting of blocks**
- **Motion Partition Merging**



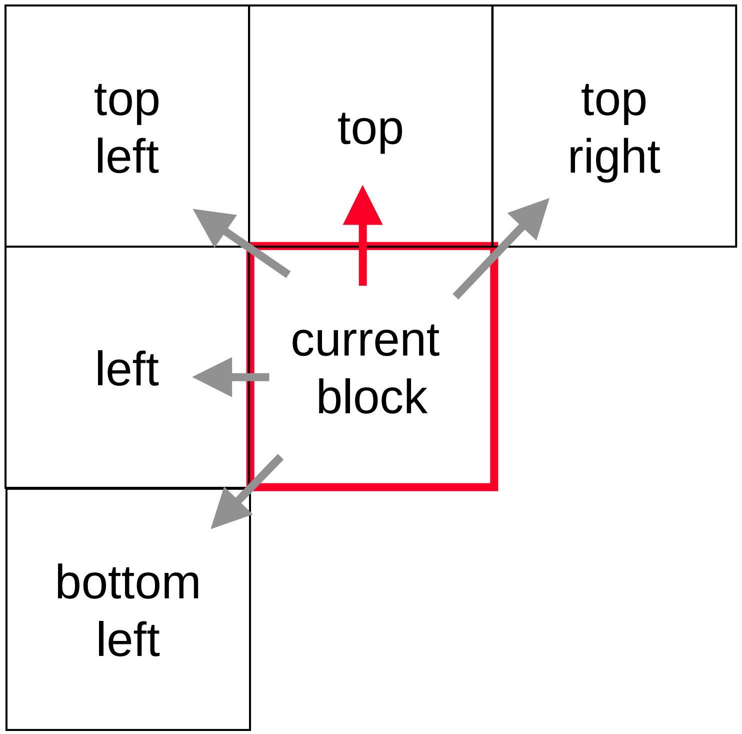




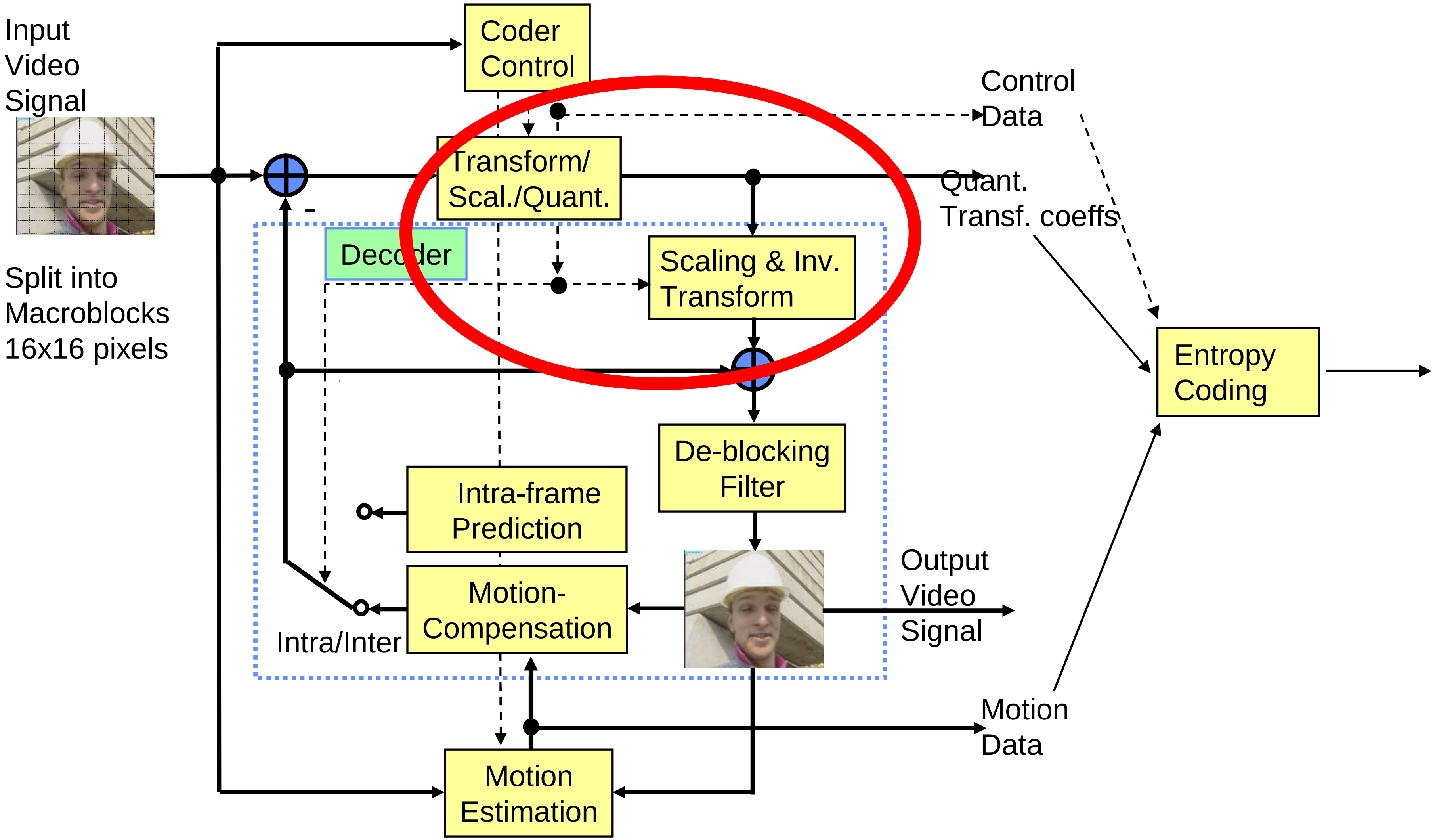
slide 20

Motion Vector Coding

- **Adaptive Motion Vector Coding**
  - Signaling of motion vector predictor

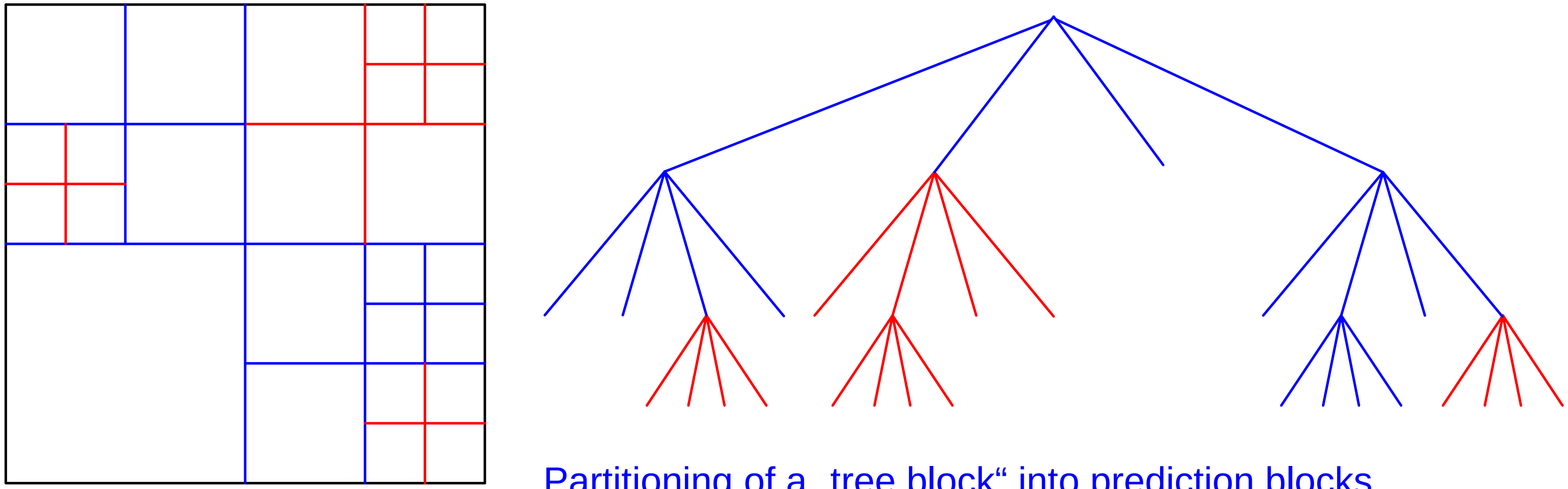


Transform Coding



Partitioning for Prediction and Transform

1. Partitioning into prediction blocks (intra or inter prediction)
2. Partitioning of prediction blocks into transform blocks (specifying the transform sizes)

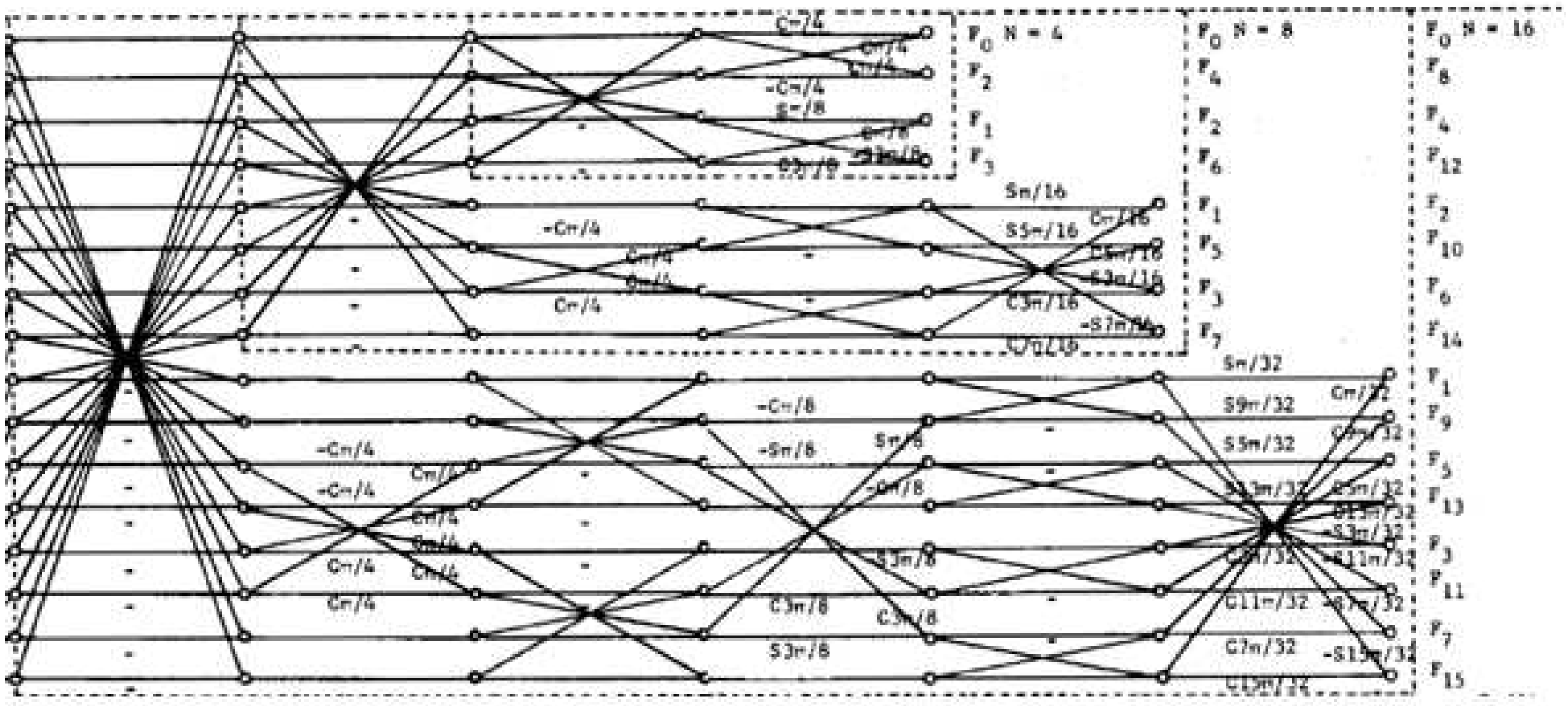


Partitioning of a „tree block“ into prediction blocks

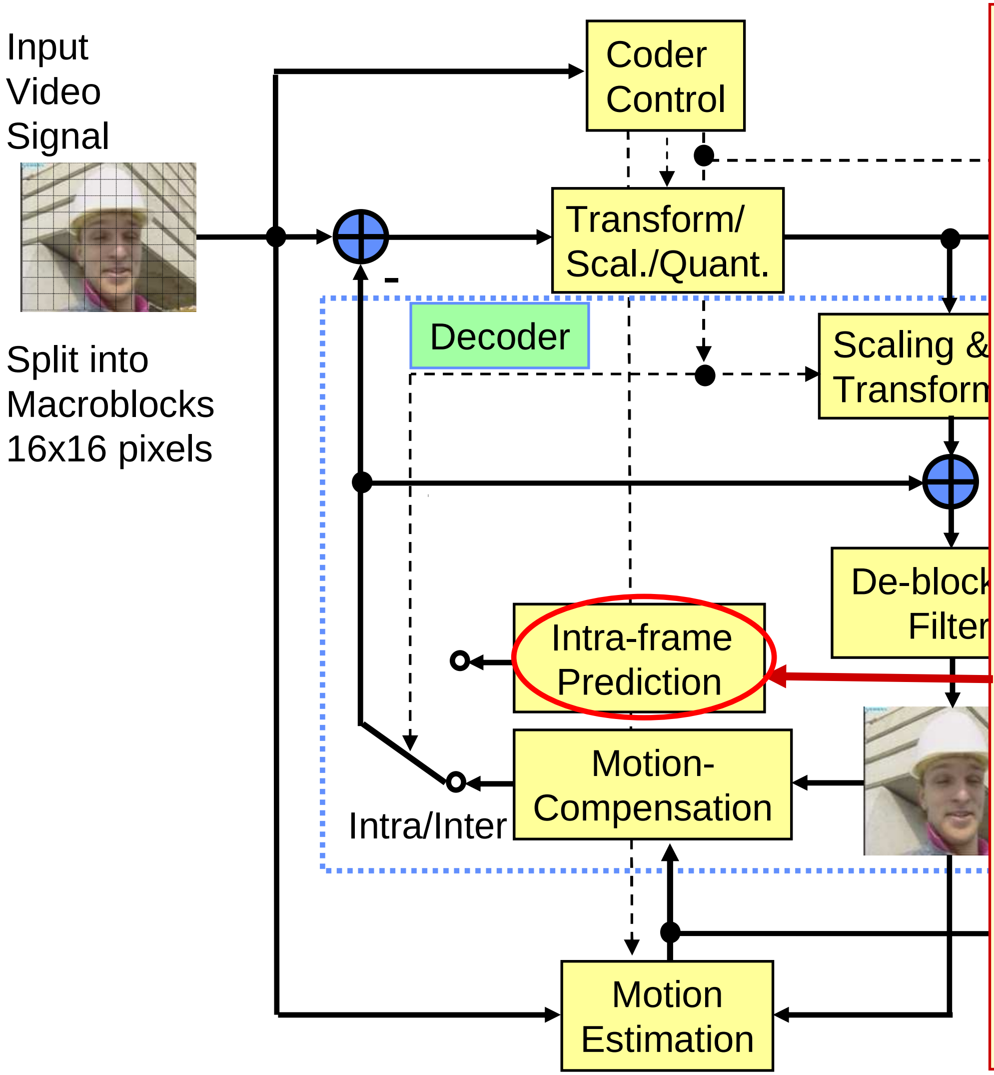
Partitioning of a prediction block into transform blocks

Variable Transform Block Sizes

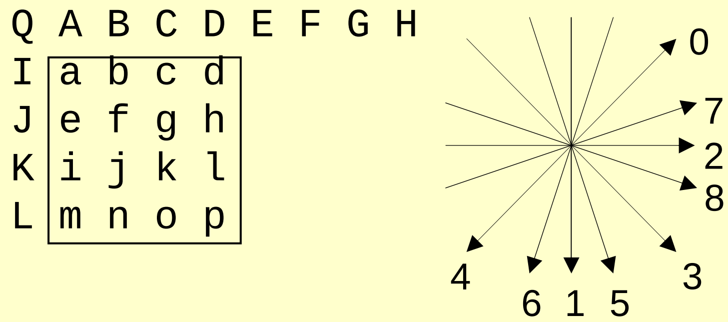
- Transform sizes range from 4x4 to 32x32
- Requirement on fast algorithms
- Additional new directional transforms proposed



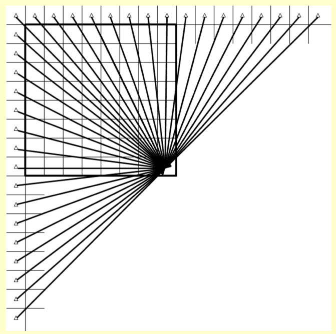
Intra Prediction



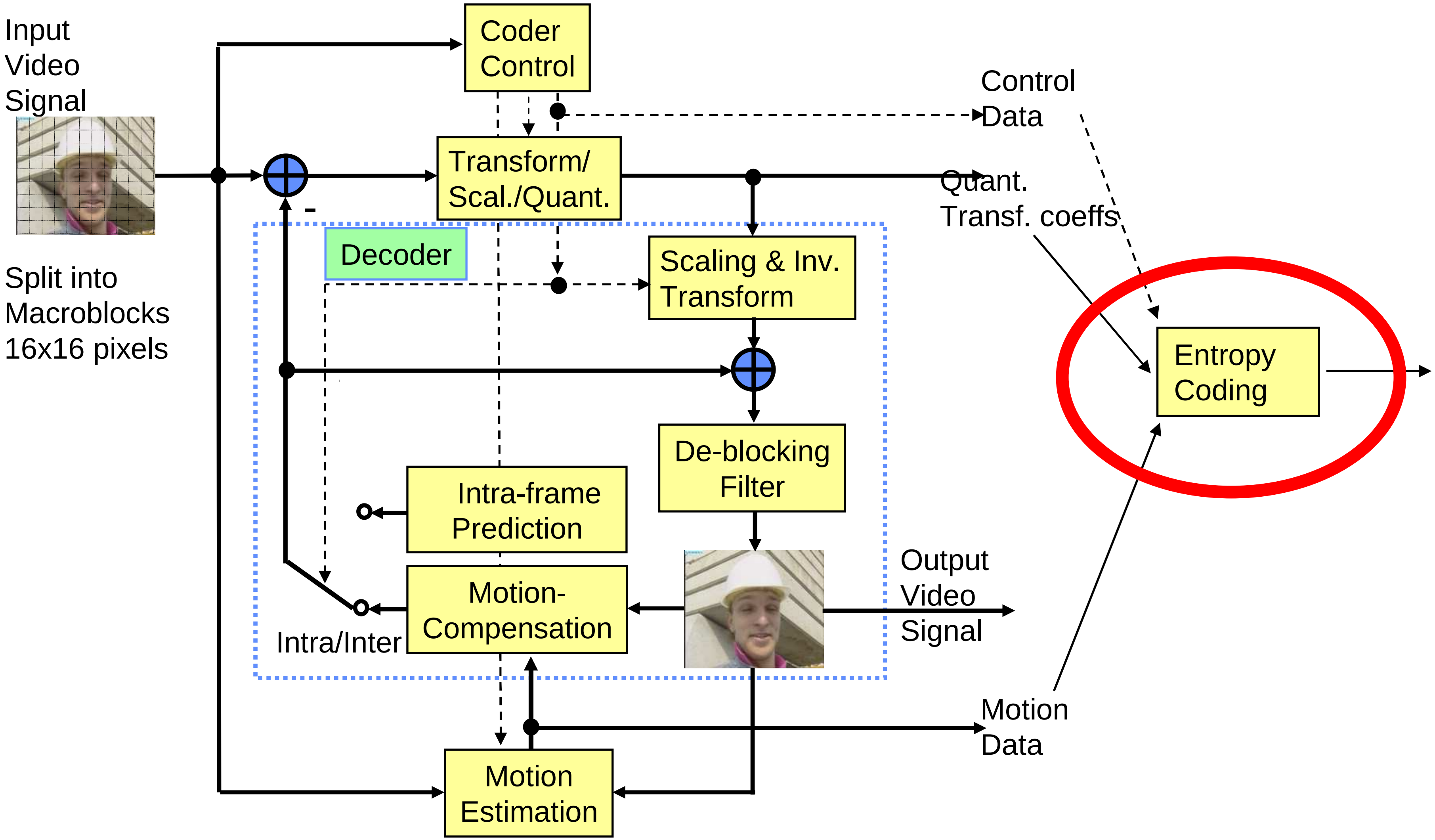
- H.264/AVC directional prediction



- New directional predictions and transforms



Entropy Coding



Entropy Coding

- Entropy coding: reversible mapping from one data representation to another, more compact representation
- Entropy is lower bound on transmission rate [Shannon, 1948]

$$H(S) = -\sum_{m=0}^{M-1} p_S(a_m)\ \log_2 p_S(a_m)$$

- Problem complicated in video coding
- Video encoder generates sequence of symbols
- Sequence of symbols contains a mix of data types: control data, motion vectors, transform coefficients

Entropy Coding in Video Compression

- Symbols in video codec correspond to a mix of data type: control data, motion vectors, transform coefficients
- Huffman coding: create a code for each data type
- Arithmetic coding: binarize data and arithmetically encode each binary symbol
- Example

$$\{s_1, s_2, \ldots, s_N\} \rightarrow \{b_1, b_2, b_3, b_4, \ldots, b_B\}$$

Symbol	Binary Symbol
0	0
1	10
2	11

Huffman Coding

- Simple construction method for entropy codes [Huffman, 1951]
- Minimizes average codeword length

$$\bar{\ell} = \sum_{k=0}^{K-1} p(a_k) \cdot \ell(a_k) \quad \text{with constraint} \quad \sum_{k=0}^{K-1} 2^{-\ell(a_k)} \leq 1$$

- Bounds on Rate $H(S) \leq R \leq H(S) + 1$
- Example

Symbol a_k	Probability $p(a_k)$	Huffman Code
a_0	0.75	0
a_1	0.125	10
a_2	0.125	11
	$H(S) = 1.061$ bit/s	$R = 1.25$ bit/s

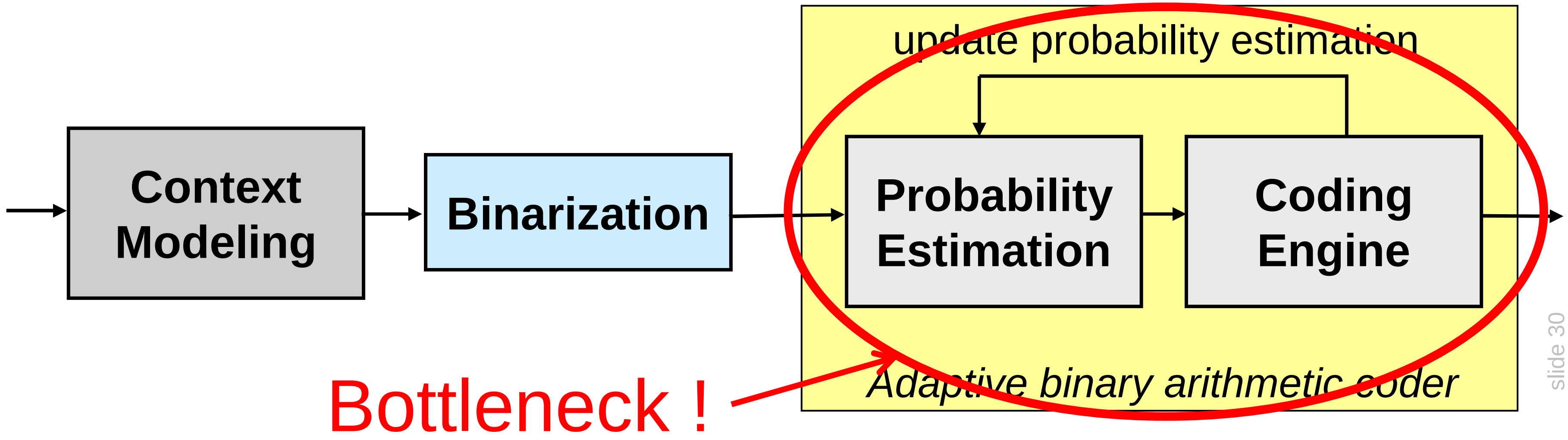
Redundancy 17.8 %

Arithmetic Coding

- Alternative construction method for entropy codes
  [Rissanen, 1976], [Pasco, 1976]
- Achieves entropy bound asymptotically for $S^n = \{S_0, S_1, \ldots, S_{n-1}\}$

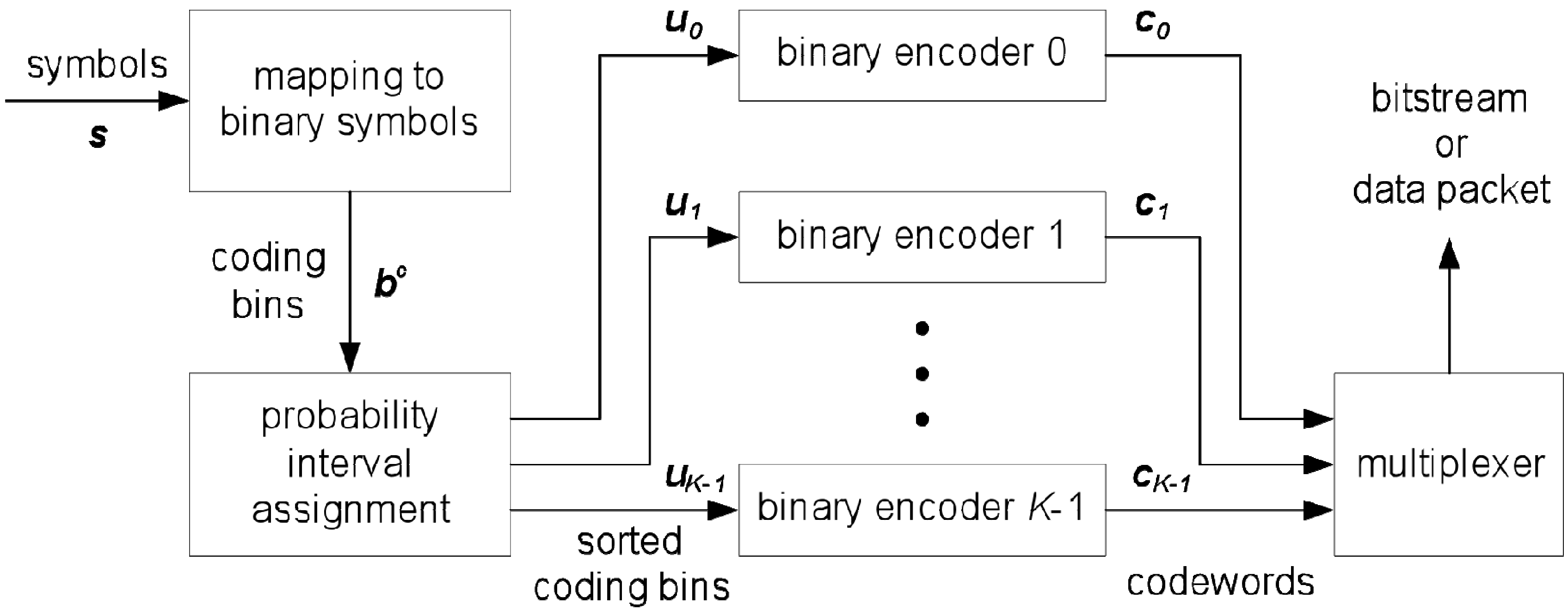
$$\frac{1}{n} H(S^n) \leq R \leq \frac{1}{n} H(S^n) + \frac{1}{n}$$

- In video coding: context-adaptive binary arithmetic coding(CABAC), [Marpe, Schwarz, Wiegand, 2003]



Probability Interval Partitioning Entropy Codes

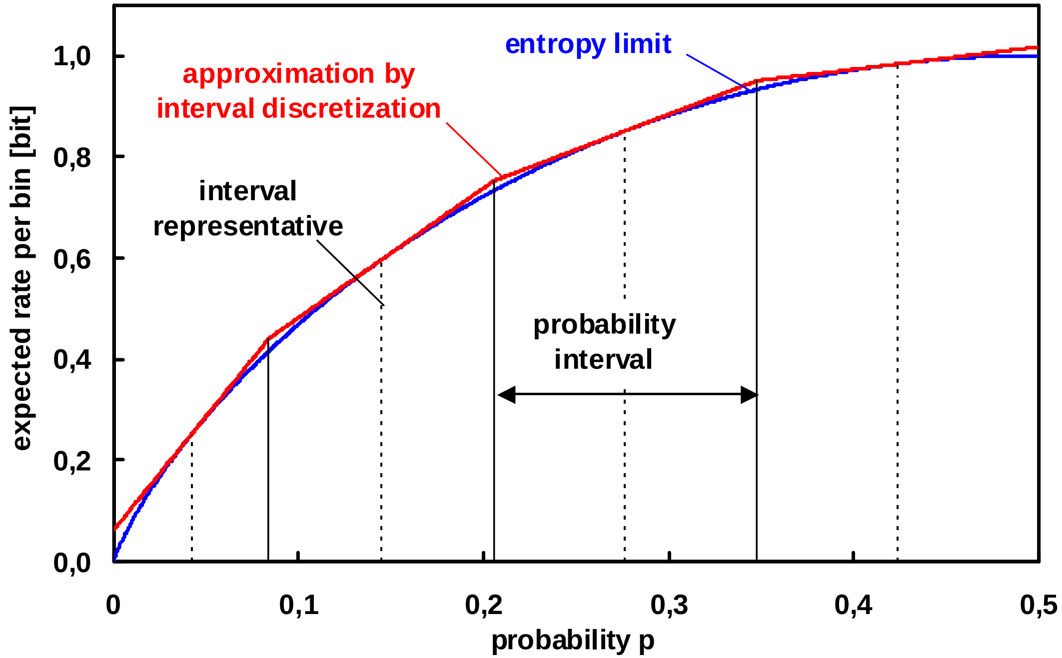
- Arithmetic coding represents probability of binary symbols
- New concept: PIPE coding [Marpe, Schwarz, Wiegand 2010]



Probability Interval Partitioning

- Uniform probability distribution of binary probabilities
- With slope of binary entropy function

$$H'(p) = \log_2\left(\frac{1-p}{p}\right)$$



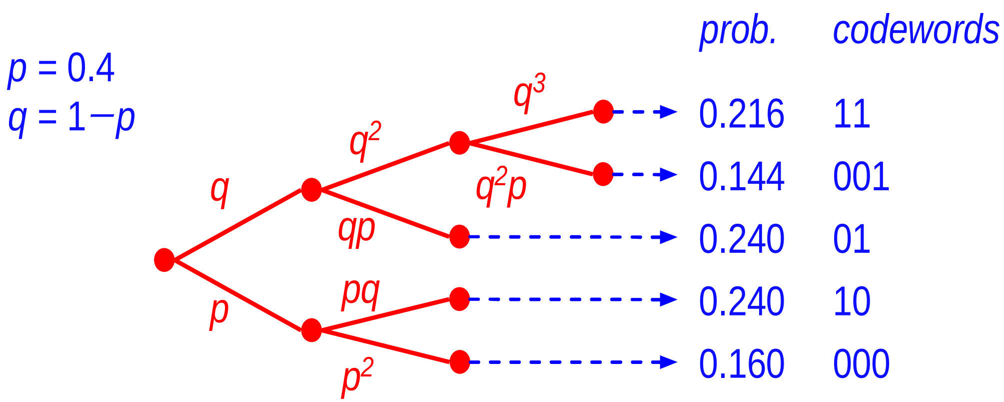
- Rate increase

$$\begin{aligned} R_b = R_b(p, p_{\mathcal{I}_k}) &= -p\ \log_2(p_{\mathcal{I}_k}) - (1-p)\log_2(1-p_{\mathcal{I}_k}) \\ &= H(p_{\mathcal{I}_k}) + (p - p_{\mathcal{I}_k})\ H'(p_{\mathcal{I}_k}), \end{aligned}$$

Average Rate Increase (relative to entropy limit)						
f(p)	K=1	K=2	K=4	K=8	K=12	K=16
uniform	12.47%	3.67%	1.01%	0.27%	0.12%	0.07%

Encoding of Binary Symbols in Each Interval

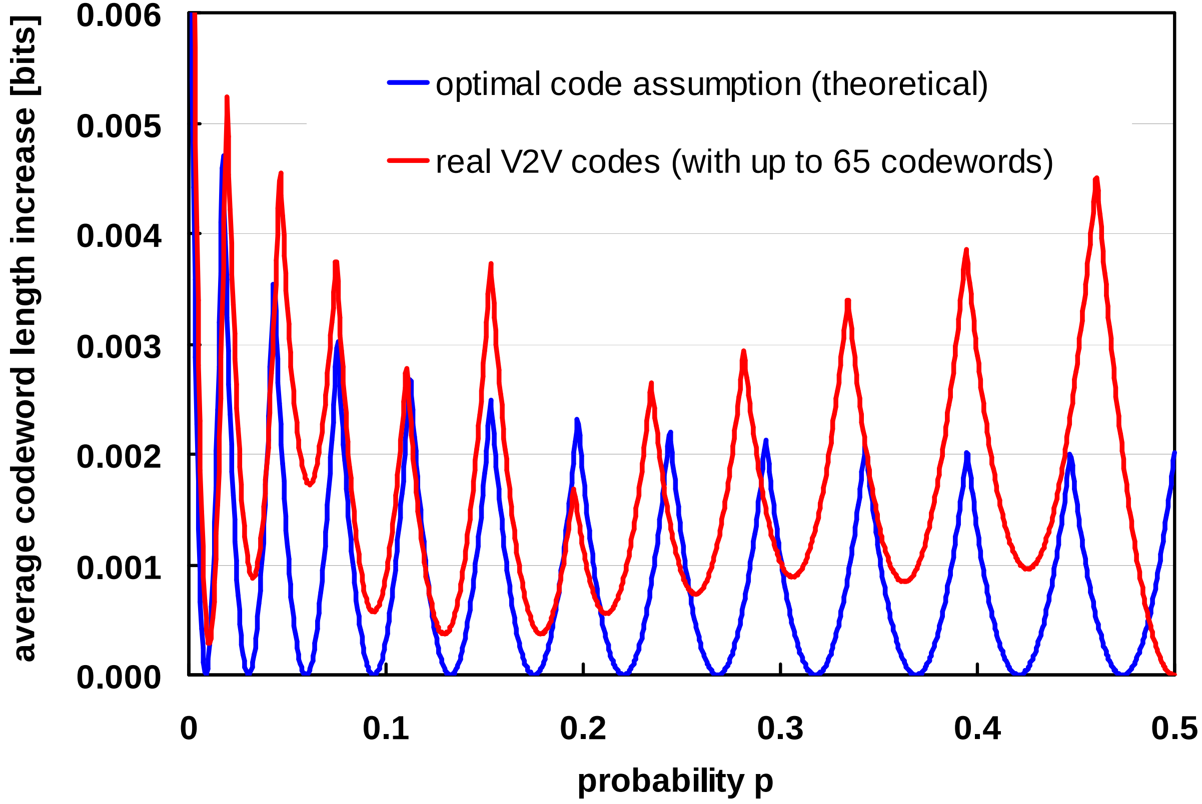
- Assume fixed probability for each interval
- Create combined events of binary symbols
- Probability of symbol *p*
- Leaf node probability is $p_{\mathcal{L}}(j) = p^{x_j}\,(1-p)^{y_j}$



bin sequence	codeword
000	11
001	001
01	01
10	10
11	000

Efficiency of PIPE coding

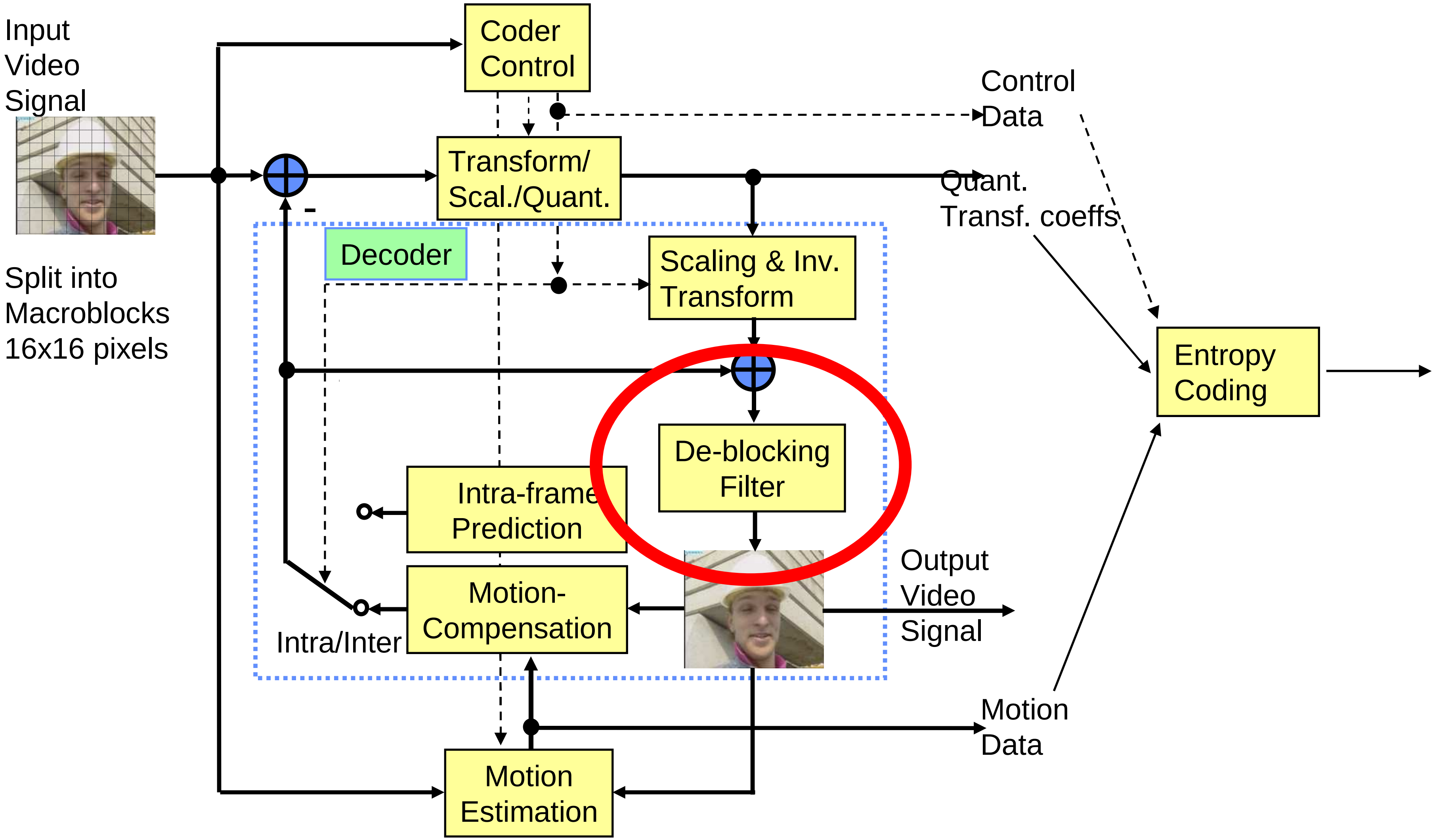
- Coding efficiency of arithmetic coding
- Complexity of Huffman coding



0.12 %

0.24 %

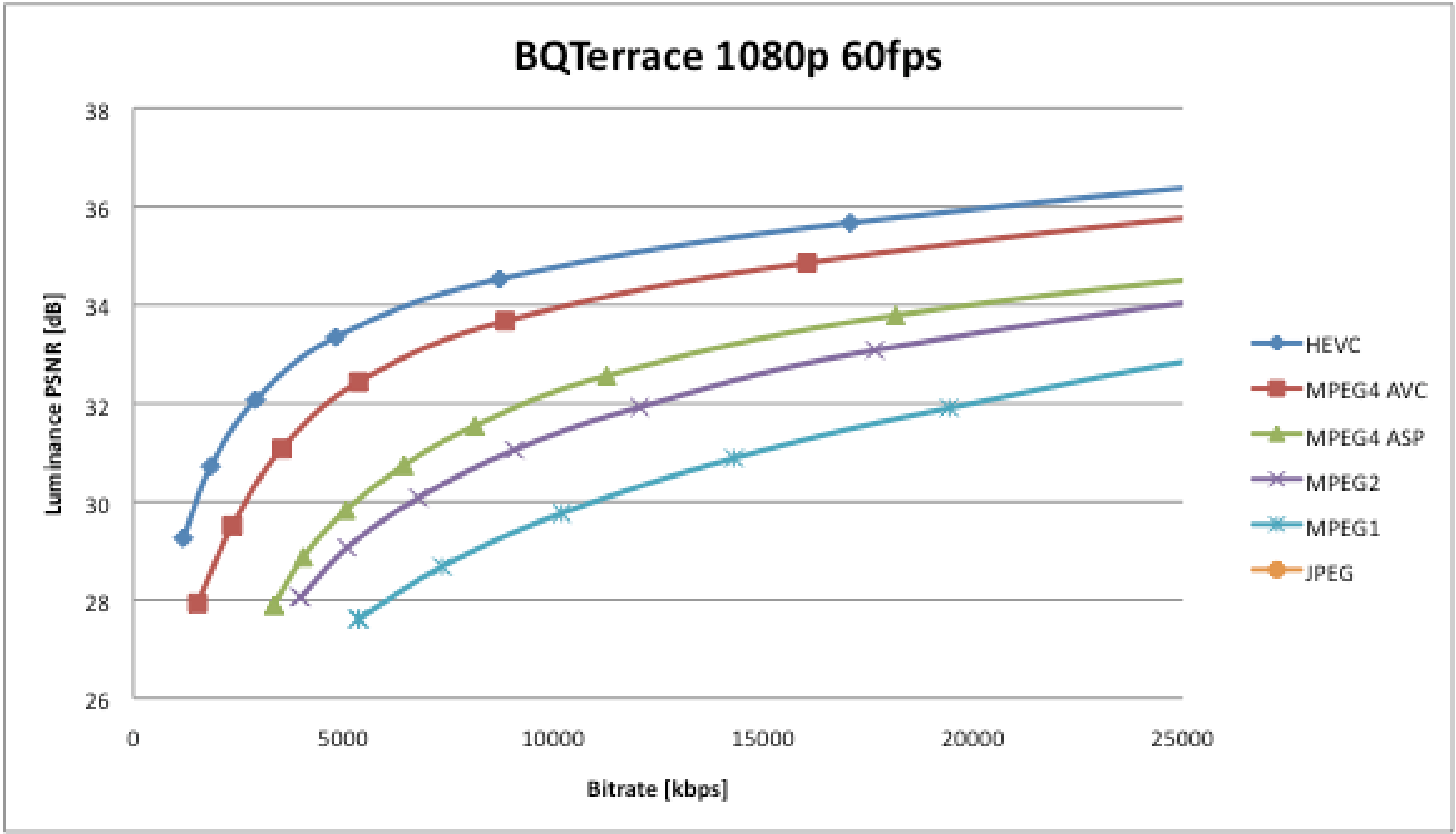
In-Loop Filtering



In-Loop Filtering

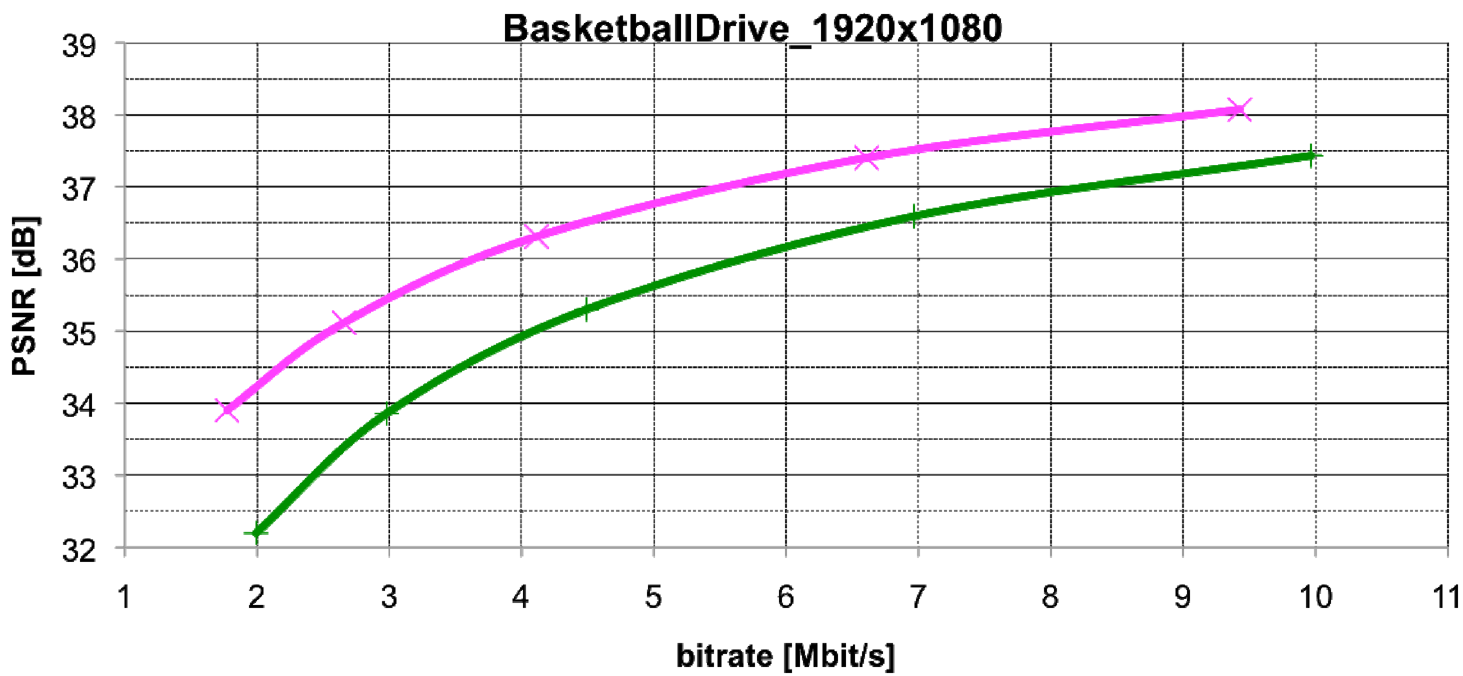
- **Deblocking filter**
- **Adaptive Loop Filter**
  - Wiener filter
  - Transmission of filter coefficients and indication for regions

Coding Efficiency



Coding Efficiency

- On average, current gains seem to be around 30% bit rate reduction
- Goal: 50% bit rate reduction against H.264/AVC HP at same quality
- Finalization: January 2013



Summary

- **H.264/AVC**: everywhere today
- **High-Efficiency Video Coding**
  - New project of MPEG and VCEG
  - Started in April 2010
  - First Test Model in Oct. 2010
  - Currently roughly 30% bit rate reduction
  - Goal: 50% bit rate reduction vs. H.264/AVC HP
  - Finalization intended for early 2013

Acknowledgement & Support

HHI & TUB members and research associates

- D. Marpe
- H. Schwarz
- K. Müller
- P. Kauff
- R. Schäfer
- other members of the HHI/TUB team

Supported by

Federal Ministry of Economics and Technology

SEVENTH FRAMEWORK PROGRAMME