
The H.26L Video Coding Project

- New ITU-T Q.6/SG16 (**VCEG - Video Coding Experts Group**) standardization activity for video compression
- **August 1999:** 1st test model (TML-1)
- **December 2001:** 10th test model (TML-10)
- **December 2001:** Formation of the **Joint Video Team (JVT)** between VCEG and MPEG to finalize H.26L as a joint project (similar to MPEG-2)
- Schedule:
 - **February 2002:** Last major feature adoptions
 - **November 2002:** Final approval



Goals of the H.26L Project

- **Simple syntax specification**
 - Targeting simple and clean solutions
 - Avoiding any excessive quantity of optional features or profile configurations
- **Improved Coding Efficiency**
 - Average bit rate reduction of 50% given fixed fidelity compared to any other standard
- **Improved Network Friendliness**
 - Issues examined in H.263 and MPEG-4 are further improved
 - Major targets are mobile networks and Internet

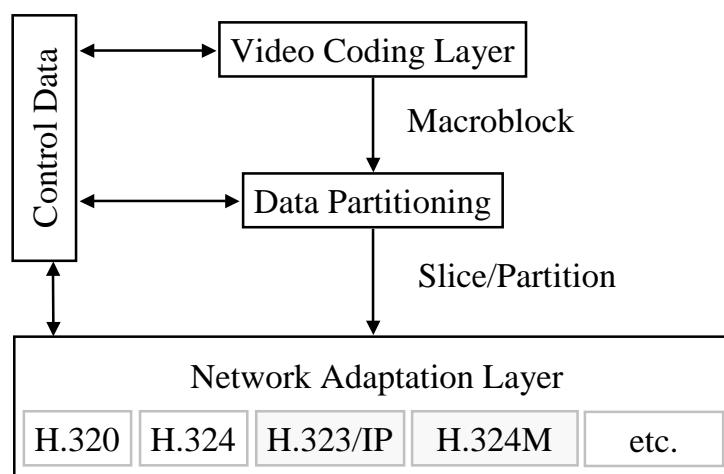


Applications

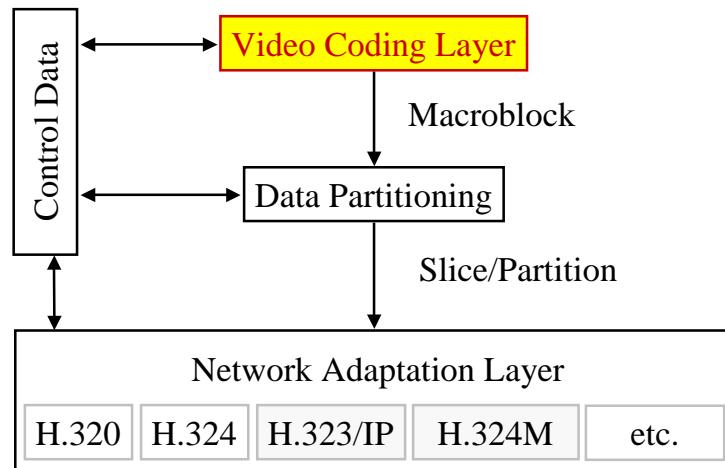
- **Conversational H.32X Services**
 - H.320 Conversational
 - *3GPP Conversational H.324/M*
 - *3GPP Conversational IP/RTP/SIP*
 - H.323 Conversational Internet/unmanaged/best effort IP/RTP
- **Streaming Services**
 - *3GPP Streaming IP/RTP/RTSP*
 - Streaming IP/RTP/RTSP (without TCP fallback)
- **Other Services**
 - Entertainment Satellite/Cable/DVD, 0.5 – 8 Mbit/s
 - Digital Cinema Application
 - *3GPP Multimedia Messaging Services*



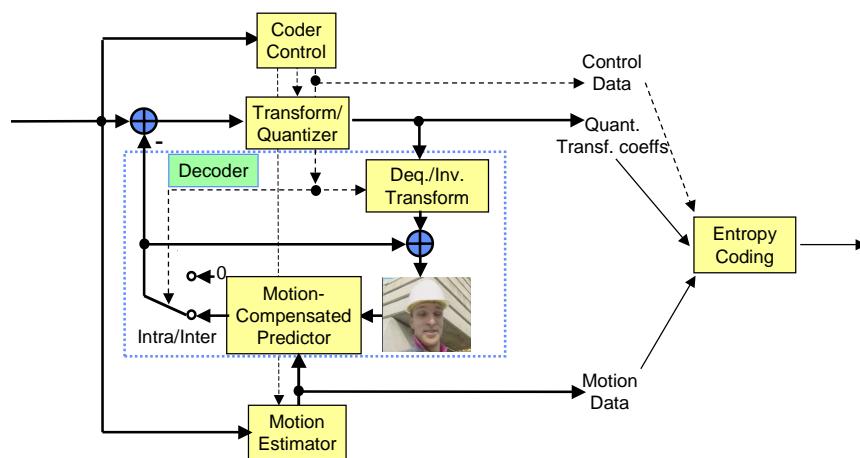
H.26L Layer Structure



H.26L Layer Structure



H.26L Video Coding Layer

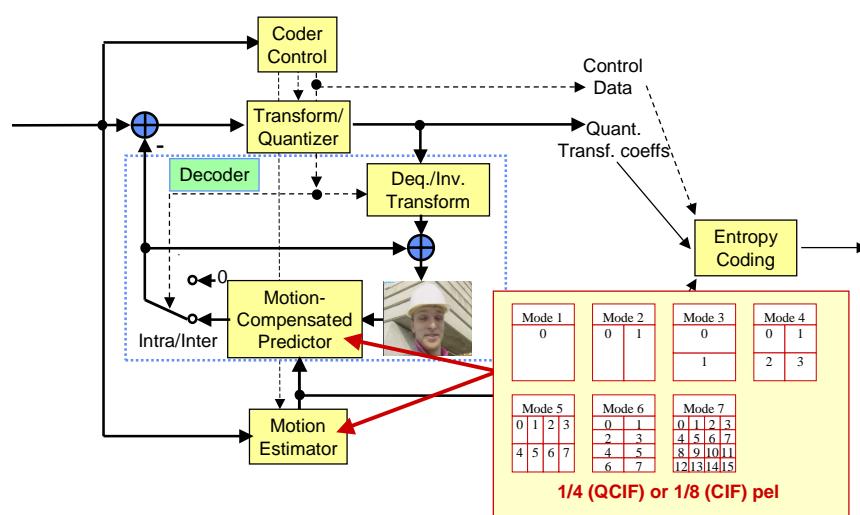


Common Elements with other Standards

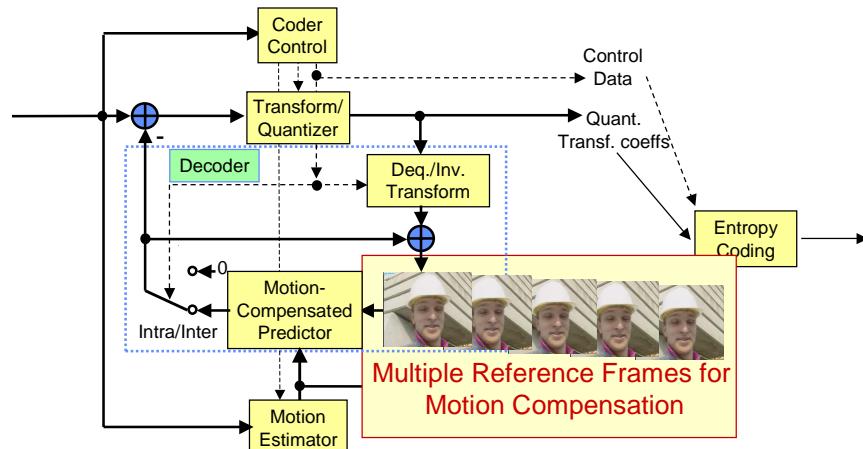
- 16x16 macroblocks
- Conventional sampling of chrominance and association of luminance and chrominance data
- Block motion displacement
- Motion vectors over picture boundaries
- Variable block-size motion
- Block transforms (not wavelets or fractals)
- Run-length coding of transform coefficients
- Scalar quantization
- I-, P-, and B-Picture types



Motion Compensation Accuracy



Multiple Reference Frames

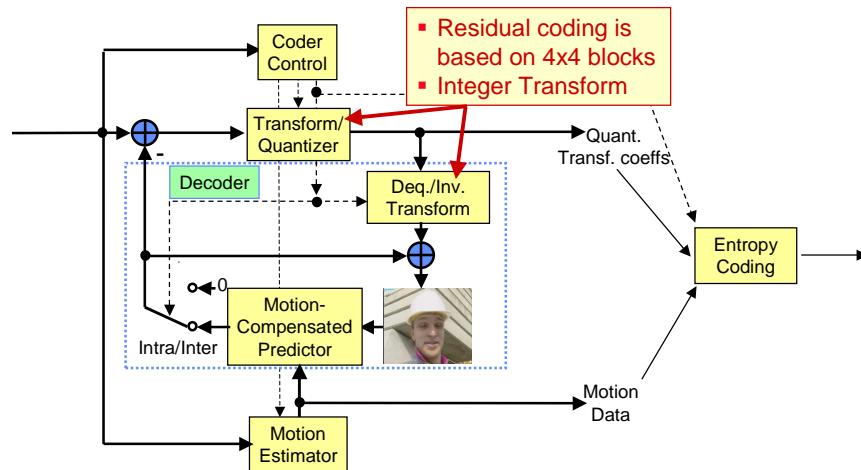


Motion Compensation

- Various block sizes and shapes for motion compensation (7 segmentations of the macroblock: 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4)
- 1/4 sample (sort of per MPEG-4) and 1/8 sample accuracy motion
 - 6x6 tap filtering to 1/2 sample accuracy, bilinear filtering to 1/4 sample accuracy, special position with heavier filtering
 - 8x8 tap filtering applied repeatedly for 1/8 pel motion
- Multiple reference pictures (per H.263++ Annex U)
- Temporally-reversed motion and generalized B-frames
- B-frame prediction weighting



Residual Coding



Residual and Intra Coding

- Transform
 - Integer transform approximating a DCT
 - Matrix is obtained by $T=\text{round}(26 \times H)$
 - Based primarily on 4x4 transform size (all prior standards used 8x8)
 - Expanded to 8x8 for chroma by 2x2 transform of the DC values
- Intra Coding Structure
 - Directional spatial prediction (6 types luma, 1 chroma)
 - Expanded to 16x16 for luma intra by 4x4 transform of the DC values

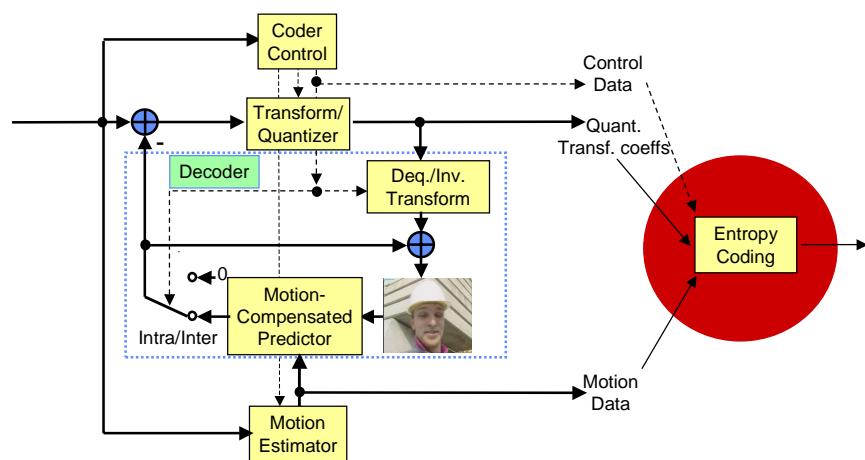


Quantization and Deblocking

- Quantization
 - Two inverse scan patterns
 - Logarithmic step size control
 - Smaller step size for chroma (per H.263 Annex T)
- Deblocking Filter (in loop)



Entropy Coding



Universal Variable Length Code (UVLC)

- One table that is used universally for all symbols
- Simple, but has the following disadvantages
 - Probability distribution may not be a good fit
 - Probability distribution is **static**
 - Correlations between symbols are ignored, i.e. **no conditional probabilities** are used
 - Code words must have **integer number of bits**
(Low coding efficiency for highly peaked pdfs)



Context-based Adaptive Binary Arithmetic Codes (CABAC)

- Usage of **adaptive** probability models
- Exploiting symbol correlations by using **contexts**
- **Non-integer** number of bits per symbol by using **arithmetic codes**
- Restriction to **binary arithmetic coding**
 - **Simple and fast adaptation** mechanism
 - Fast binary arithmetic coders are available
 - **Binarization** is done using the UVLC



Test Model Coder Control (1)

- Coder control is a non-normative part of H.26L but is used in VCEG to show H.26L encoder performance and to make design decisions
- Rate-Constrained Mode Decision: minimize

$$J(MODE | QP, \lambda_{MODE}) = SSD(MODE | QP) + \lambda_{MODE} \cdot R(MODE | QP)$$

SSD - Sum of squared differences (luminance & chrominance)

R - Number of bits (MB-header, motion, all transform coefficients)

$MODE$ - Element of set of possible macroblock modes

- Set of possible macroblock modes

- Dependent on frame type

- For instance, P-frame in H.26L:

{SKIP, INTER_16x16, INTER_16x8, INTER_8x16, INTER_8x8,

INTER_8x4, INTER_4x8, INTER_4x4, INTRA_4x4, INTRA_16x16}



Test Model Coder Control (2)

- Rate-Constrained Motion Estimation:
Integer-pixel motion search as well as sub-pixel refinement is performed by minimizing

$$J(REF, \mathbf{m} | \lambda_{MOTION}) = SAD(REF, \mathbf{m}) + \lambda_{MOTION} \cdot \{R(\mathbf{m} - \mathbf{p}) + R(REF)\}$$

SAD - Sum of absolute differences (luminance)

R - Number of bits associated with motion information

REF - Reference frame

\mathbf{m} - Motion vector

\mathbf{p} - Prediction of motion vector

- Relationship between $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$

- Choice of $\lambda_{MODE} = 0.85 \cdot 2^{QP/3}$

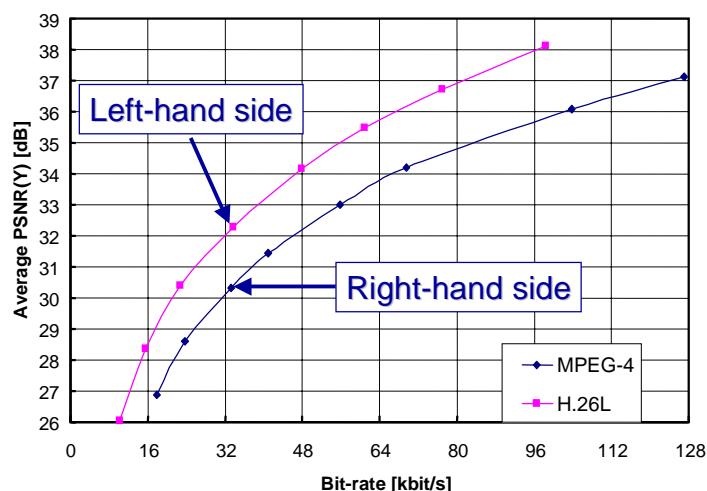


Comparison of H.26L and MPEG-4

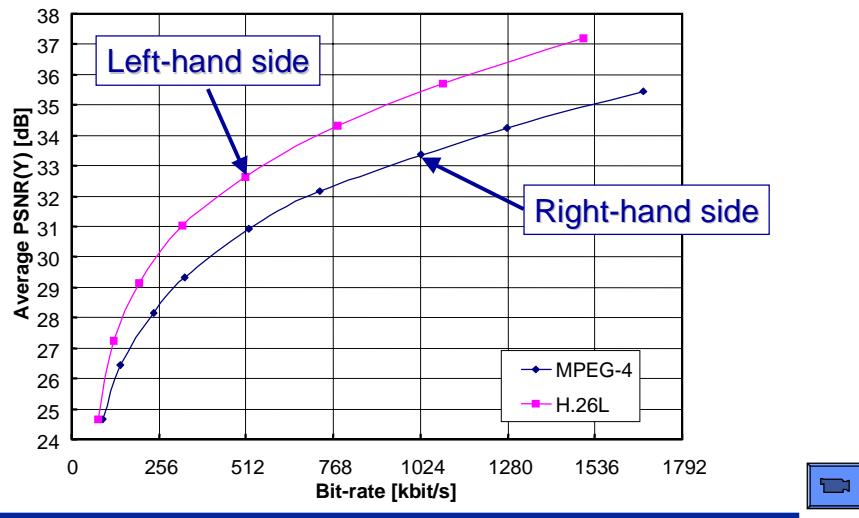
- Both:
 - Sequence structure $IBBPBBP\dots$
 - Search range: 32x32 around 16x16 predictor
 - Encoders use similar D+IR optimization techniques
- MPEG-4: Advanced Simple Profile (ASP)
 - Motion Compensation: 1/4 pel
 - Global Motion Compensation
 - $QP_B = 1.2 \times QP_P$
- H.26L:
 - Motion Compensation: 1/4 pel (QCIF), 1/8 pel (CIF)
 - Using CABAC entropy coding
 - 5 reference frames
 - $QP_B = QP_P + 2$



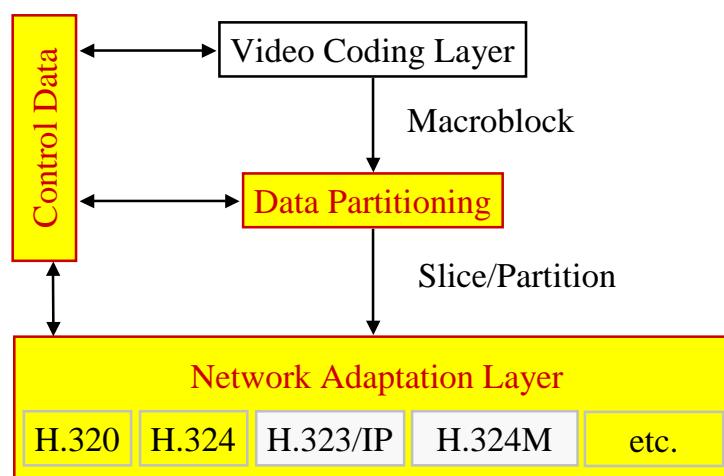
RD Curves: Foreman (QCIF, 10Hz)



RD Curves: Tempete (CIF, 30Hz)



H.26L Layer Structure



Network Adaptation Layer

- **Tasks**

- Mapping of slice structure on transport layer
- Setup, framing, encapsulation, interleaving, logical channels, closing, timing issues, synchronization, etc.
- Transport of control and header information
- Further network specific issues (feedback, prioritization,...)

- **The specification for each NAL includes**

- Verbal description
- Encapsulation process (processing of slice structure)
- Header and parameter set specification



Network Adaptation Features

- **Slice Structure Coding**

- Slices for a specified number of macroblocks
- Slices for a specified number of Bytes

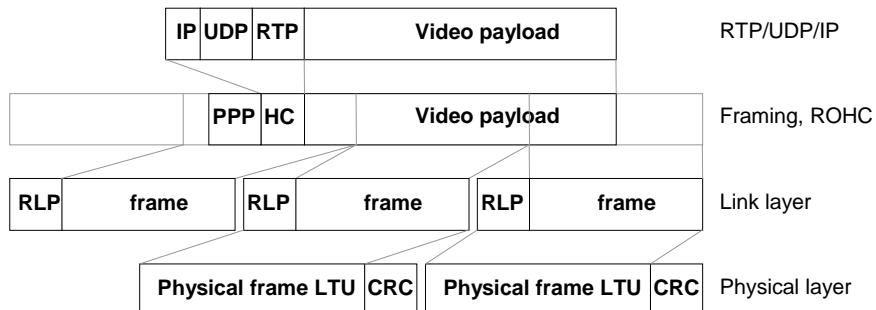
- **Data Partitioning:** header, motion vectors, Intra, and Inter transform coefficients

- Mitigating Error Propagation (with and without feedback):
 - **Intra** picture and Intra macroblock refresh
 - Use of **multiple** reference pictures
 - Use of **I-, P, and B-pictures**
- Switching between pre-coded sequences: **SP-pictures**



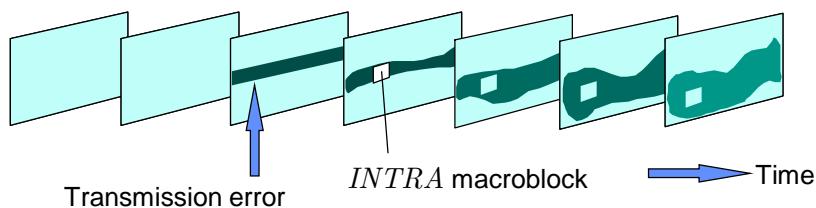
Common Test Conditions

- Mainly concentrating on RTP/IP over 3GPP/3GPP2 networks
- Packetization through the user plane protocol stack (CDMA-2000)
- IP/UDP/RTP header compression used: 3 bytes
- Loss of LTU leads to loss of IP packet
- Retransmission at RLP layer is possible



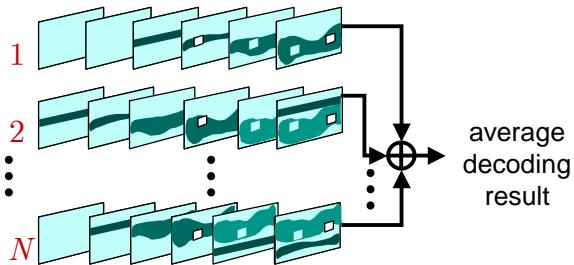
Impact of Transmission Errors

- If just one frame is missing → concealment at decoder side → reference pictures at coder and decoder differ → error propagation
- Error decays slowly → mitigate error propagation
- Mitigating error propagation:
 - Use of **multiple** reference pictures
 - **INTRA** picture and **INTRA macroblock** refresh



Assignment of Intra Macroblock Coding

- Intra coding provides lower coding efficiency than Inter coding
→ Trade-off between error resilience and coding efficiency
- Random transmission errors cause that decoding result becomes random variable
- Model decoding random variable using N sample functions
- Optimize encoding operation for average decoding result for a given packet-loss rate p



Modified Coder Control and Comparison

- Given the N decoded versions with random packet losses of probability p at the encoder
- Lagrangian Mode Decision

$$D_2(M | Q) + \lambda_M \cdot R(M | Q)$$

- With modified distortion measure

$$D_2(INTER | Q) = \frac{1}{N} \sum_{n=1}^N \sum_i (s_i - \hat{s}_{i,n}(p))^2$$

- Approach adaptively increases costs for Inter coding and therefore increases Intra coding rate
- Comparison against periodic slice-based Intra coding
- Both: 1 slice = 1 packet, previous frame error concealment
- Packet loss rate: 10 %



Summary

- Coding standards have been driving compression and transport of video signals in industry and universities
- First video coding standard: H.120
- Basis for all modern standards: H.261
- A major step forward: MPEG-1
- The most successful standard: MPEG-2
- The next generation: H.263
- Object-based coding with H.263 fall-back: MPEG-4
- A new exciting Standard: H.26L

