

H.264 Overview

Video Team

R&D Center

AddPac Technology Co. Ltd.

sales@addpac.com



Contents

- **History of Video Coding Standards**
- **H.264 Features**
- **H.264 Technical Tools**



Video Codec Applications

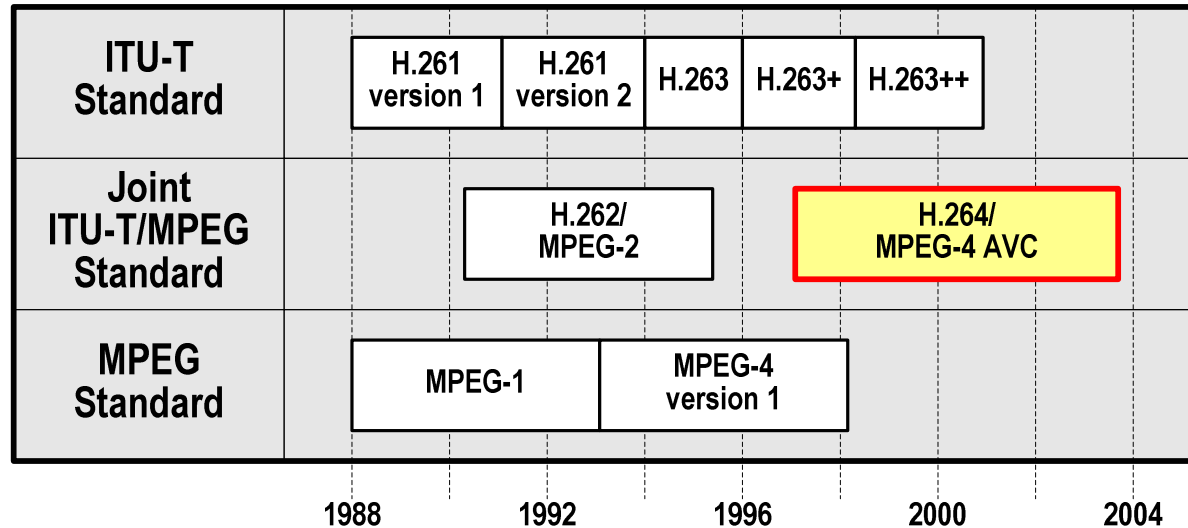
■ Codec (Encoder & Decoder)



■ Digital Video Applications

셋탑 박스	디지털 TV, 위성 네트워크 브로드밴드 네트워크(Cable, xDSL) ATM 네트워크
비디오 스트리밍	VoD
비디오 폰 및 비디오 컨퍼런싱	H.320, H.323, SIP 기반의 제품
(모바일) 멀티미디어 단말기	셀룰라 핸드폰, PDA, DMB 단말기
비디오 모니터링 시스템	비디오 감시 및 녹화 시스템 웹 카메라 네트워크 카메라 개인용 비디오 레코더(PVR: Personal Video Recorder)

History of Video Standards



■ ITU VCEG

- H.26L → [H.264](#)

■ ISO MPEG

- [MPEG-4 AVC](#) (advanced video coding) or MPEG-4 Part 10

■ VCEG-MPEG

- [JVT](#) (joint video team)
- JMx.x (joint model: reference software)

H.264 Features

■ High compression performance

- Many advanced compression techniques are adopted.
 - ✓ Advanced Intra-Prediction
 - ✓ Strong Motion Isolation (4x4, 1/4-pel resolution)
 - ✓ Multiple Reference Frames
 - ✓ Weighted Bi-Prediction
 - ✓ Context-adaptive VLC/BAC
- **Average bit rate reduction of 50%** given fixed fidelity compared to any other video standard

} Very High Complexity!

■ Exact match decoding

- Integer Transform

■ Improved Perceptual Quality

- In-Loop Deblocking Filter

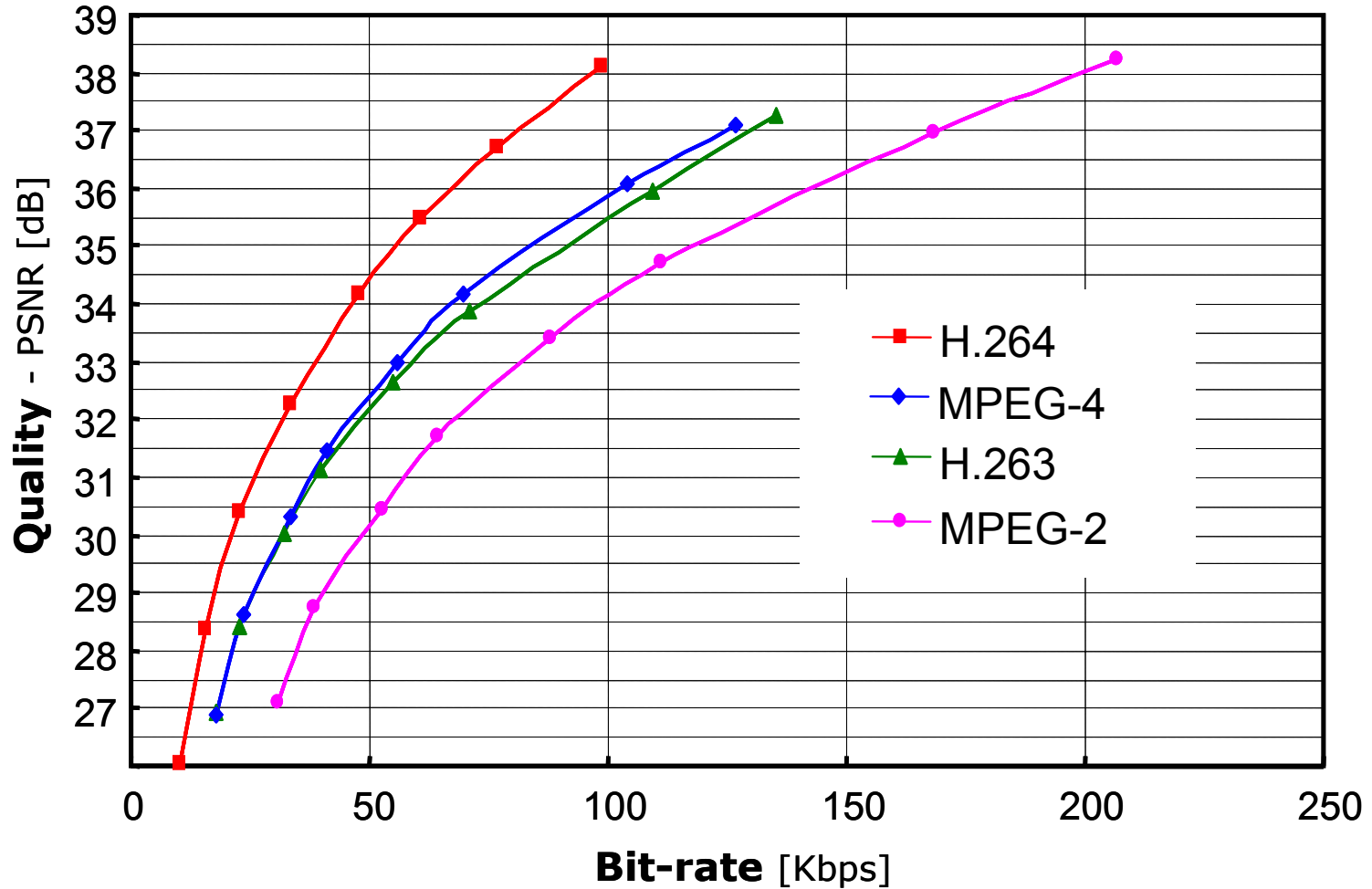
■ Network friendliness

- NAL (Network abstraction layer)
- Enhanced Error Resilience



Comparison to Other Standards

Foreman QCIF (at 10 frames/s)



Comparison to MPEG-4

■ Environment

- Matrix 20sec
- 320x240
- 30fps
- 128Kbps (Low Bit-rate)



H.264 - Technical Tools

- Sequence→GOP→Picture→Slice→MB→Block
- Chroma format – 4:2:0
- Picture type – I, P, B, SI, SP
- Frame structure – Interlaced, Progressive
- Adaptive frame/field – per picture, per MB
- Deblocking filter – in-loop
- MV resolution – $\frac{1}{4}$ pixel
- Tree-like motion segmentation – 16x16 to 4x4
- Entropy coding – EG and CAVLC/CABAC
- Data partitioning – NAL_unit, priority
- ASO (arbitrary slice order) – independently decodable
- FMO (flexible macroblock order) – map
- Redundant slice
- ABP (adaptive bi-prediction) – adaptive weighting

H.264 Profiles

■ **Baseline (Videoconferencing & Wireless)**

- I and P picture types (not B)
- Interlace
- Per-picture adaptive frame/field
- In-loop deblocking filter
- 1/4-sample motion compensation
- Tree-structured motion segmentation down to 4x4 block size
- VLC-based entropy coding (CAVLC)
- Some enhanced error resilience features: FMO, ASO, Redundant slices

■ **Main Profile (esp. Broadcast)**

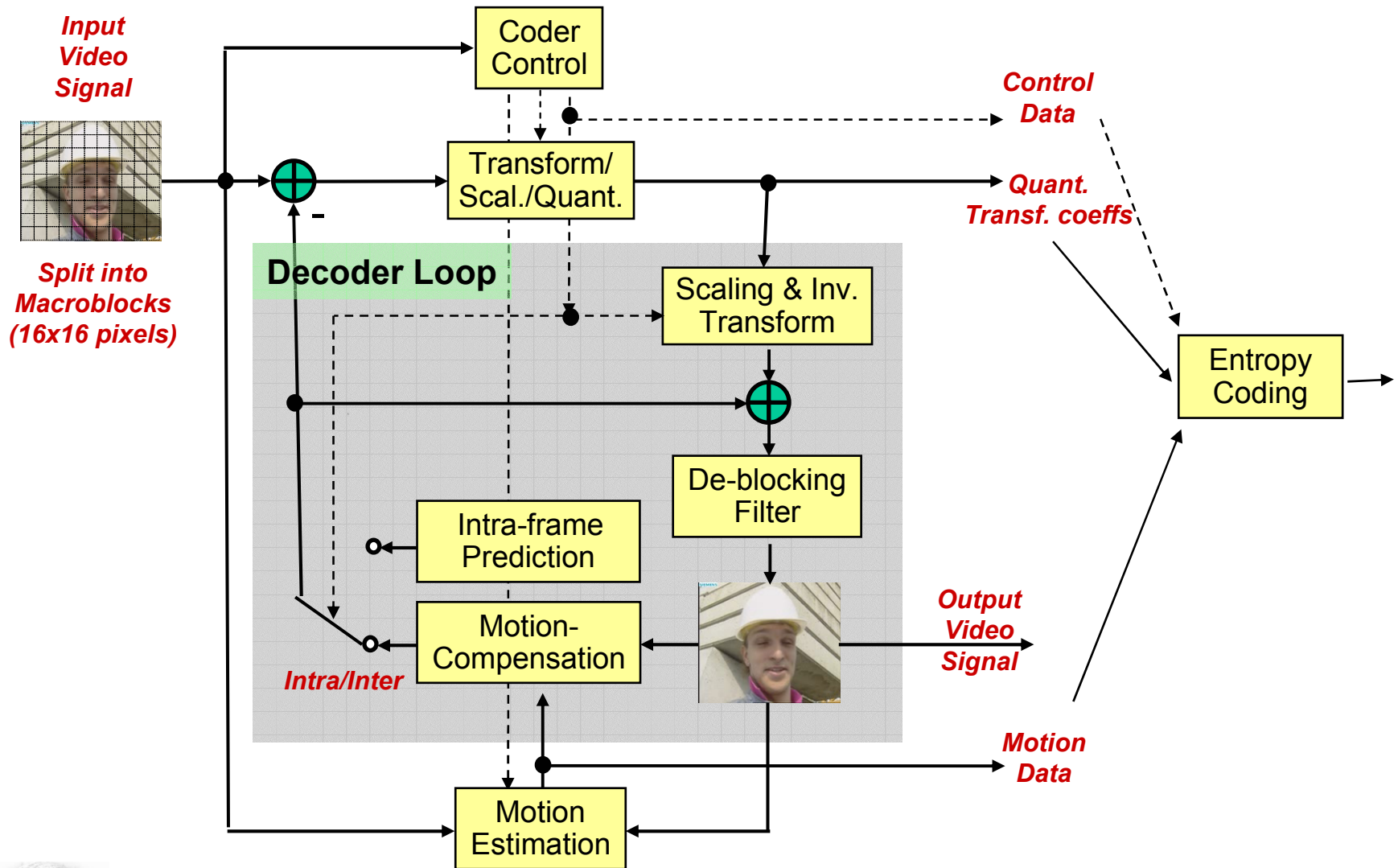
- All Baseline features except enhanced error resilience features
- B pictures
- CABAC (context-adaptive binary arithmetic coding)
- MB-level frame/field switching
- Adaptive weighting for B and P picture prediction
- Interlace

■ **Profile X (esp. Streaming)**

- All Baseline features
- B pictures
- More error resilience: Data partitioning
- SP/SI switching pictures

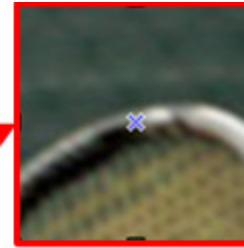
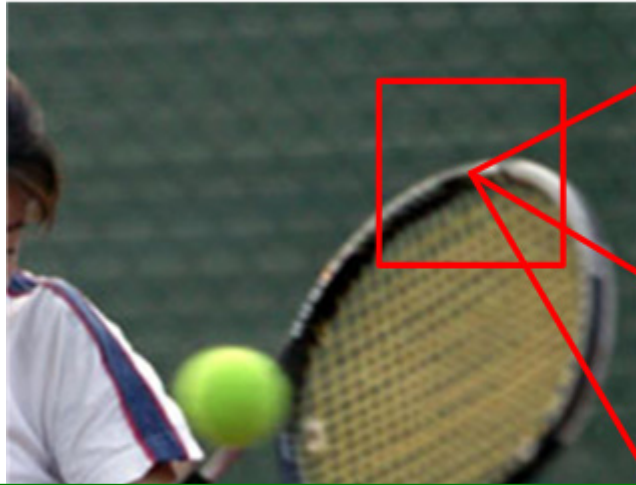


H.264: Encoder Structure



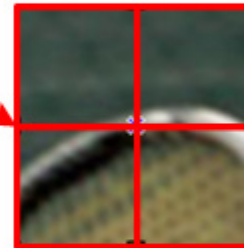
H.264: Motion Compensation Accuracy

Macroblock Partitioning



MPEG-2

- 16x16 block size
- Square shape
- 1/2 pel motion vector
- Weak Motion Isolation !



MPEG-4

- 8x8 block size
- Square shapes
- 1/2 pel motion vector
- Moderate Motion Isolation !!



H.264

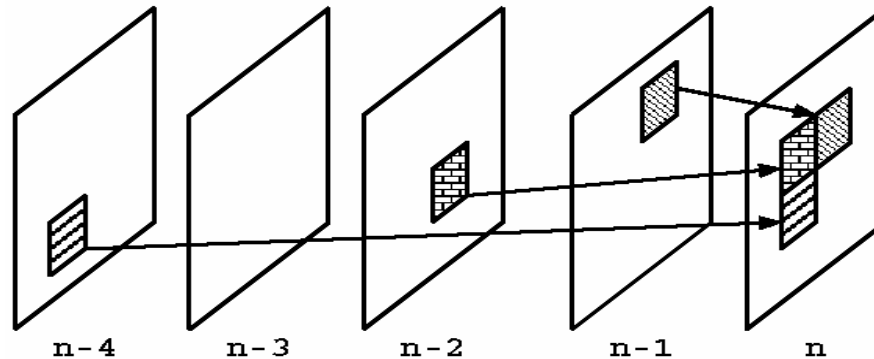
- 4x4 block size
- Arbitrary shapes
- 1/4 pel motion vector
- Strong Motion Isolation !!!

	16x16	16x8	8x16	8x8
MB Types	0	0 1	0 1	0 1 2 3
	8x8	8x4	4x8	4x4
8x8 Types	0	0 1	0 1	0 1 2 3

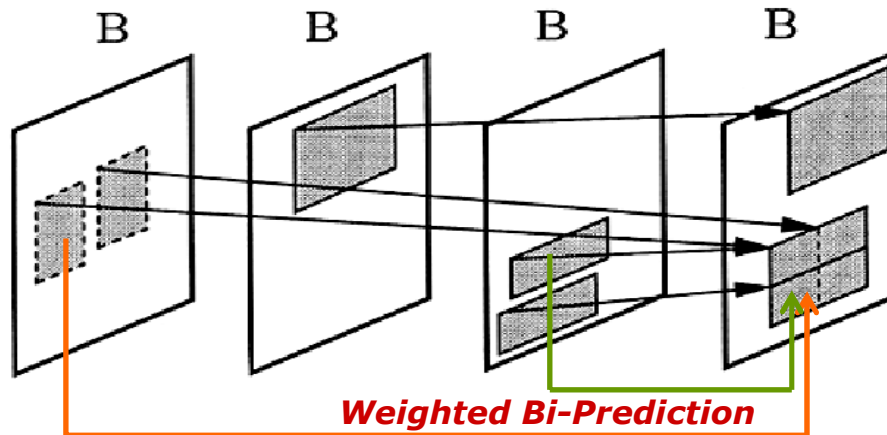
Motion vector accuracy 1/4 pixel

H.264: Multiple Reference Frames

■ Per-MB reference control

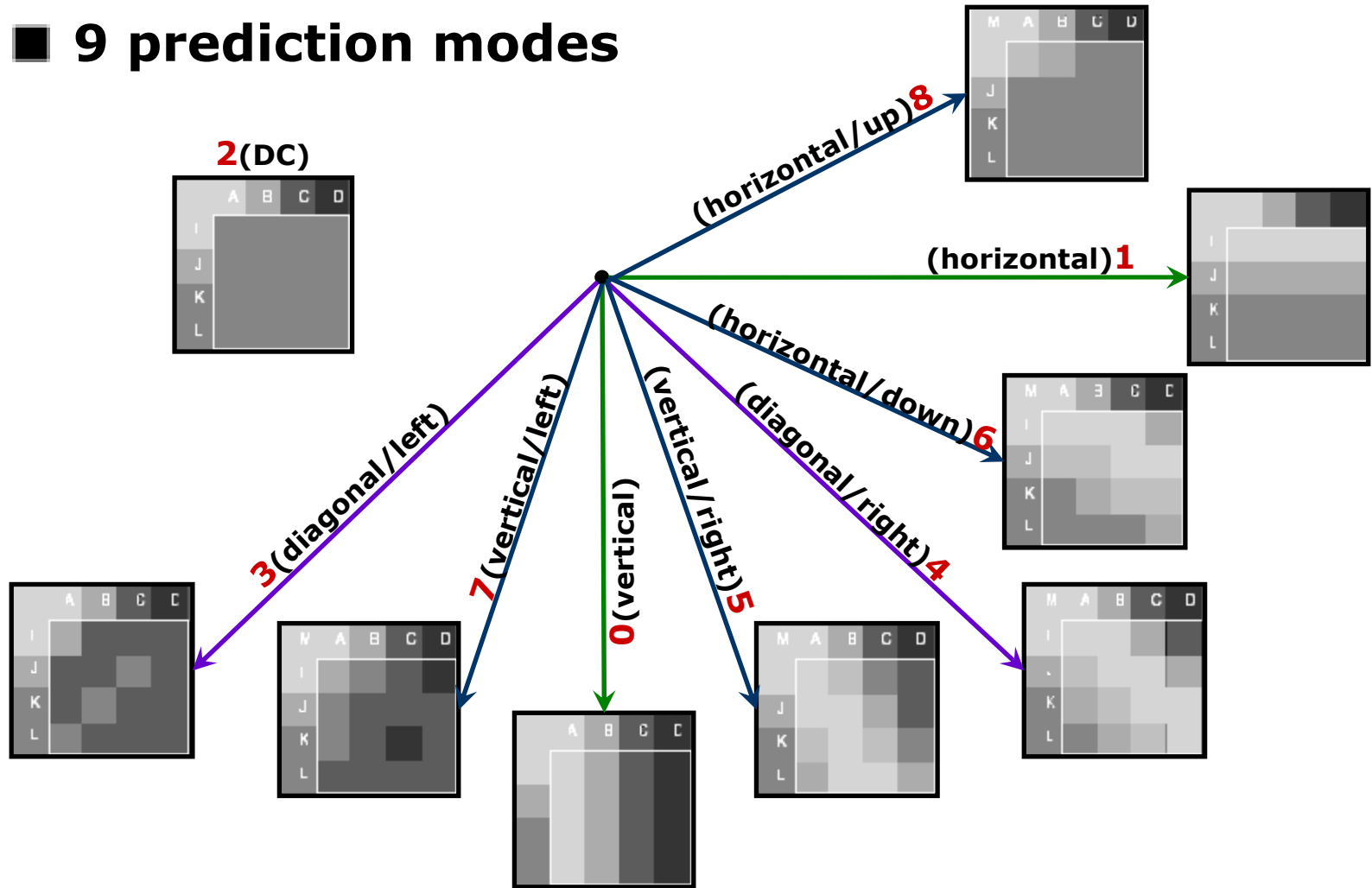


■ Generalized B(bi-predictive) picture



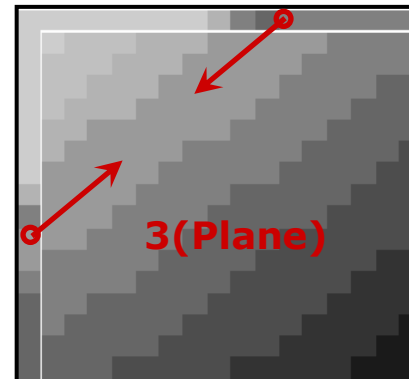
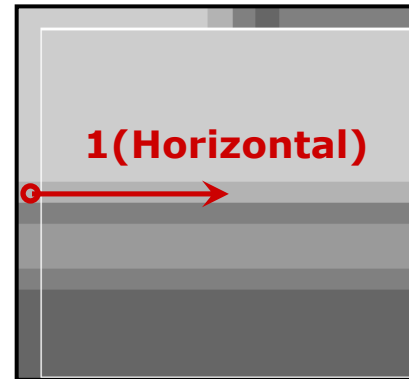
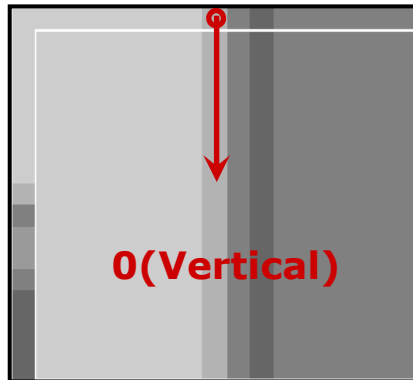
H.264: 4x4 Intra-Prediction

9 prediction modes



H.264: 16x16 Intra-Prediction

■ 4 prediction modes



H.264: Transform

■ EXACT MATCH Simplified Transform

- Based primarily on 4x4 transform

4x4 DCT

$$H = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

where $a = 1/2$, $b = \sqrt{1/2} \times \cos(\pi / 8)$,

and $c = \sqrt{1/2} \times \cos(3\pi / 8)$.

H.264

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

- Requires only 16 bit arithmetic (including intermediate values)
- Expanded to 8x8 for chroma by 2x2 transform of the DC values

H.264: Quantization

- Quantization of transform coefficients
 - **Logarithmic step size control**
 - **Extended range of step sizes**
 - **Smaller step size for chroma (cf. H.263 Annex T)**
 - **Table-driven (12.5%-increase in QStep per 1-QP increase)**
→ 2-times in QStep for every 6th increment in QP

QP	0	1	2	3	4	5	6	7	8	9	10	11	12	...
qStep	0.625	0.6875	0.8125	0.875	1	1.125	1.25	1.675	1.625	1.75	2	2.25	2.5	...
QP	...	18	...	24	...	30	...	36	...	42	...	48	...	51
qStep	...	5	...	10	...	20	...	40	...	80	...	160	...	224

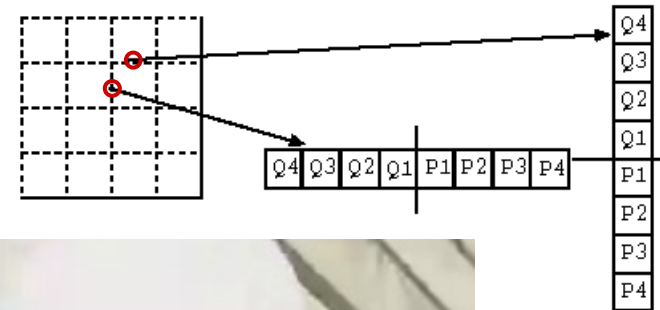
- **Reconstruction is 16-bit multiply, add, shift**
- Actual definition is very complicated due to
 - **The non-orthonormality of the integer transform**
→ position dependent scaling
 - **Reconstruction simplicity**



H.264: In-loop Deblocking Filter

- Improves subjective visual quality
- Much better than post filtering
- Highly context adaptive

16x16 Macroblock



1) Without Filter



2) With H264/AVC Deblocking

H.264: Entropy Coding

■ **Exp-Golomb Code**

- Universally for all symbols except for transform coefficients

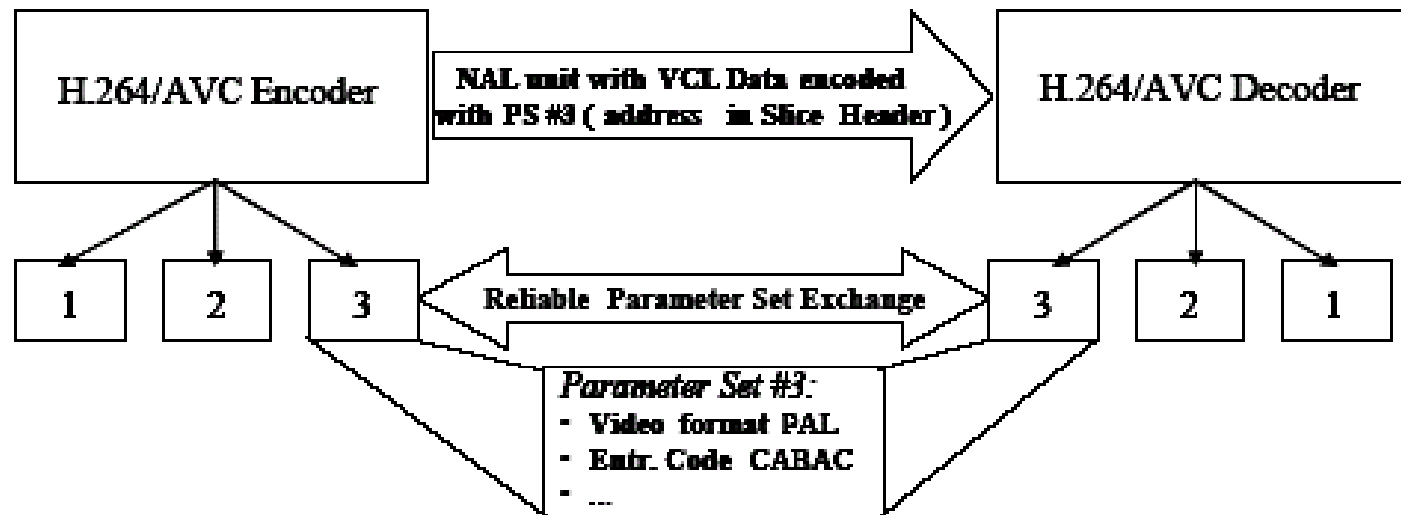
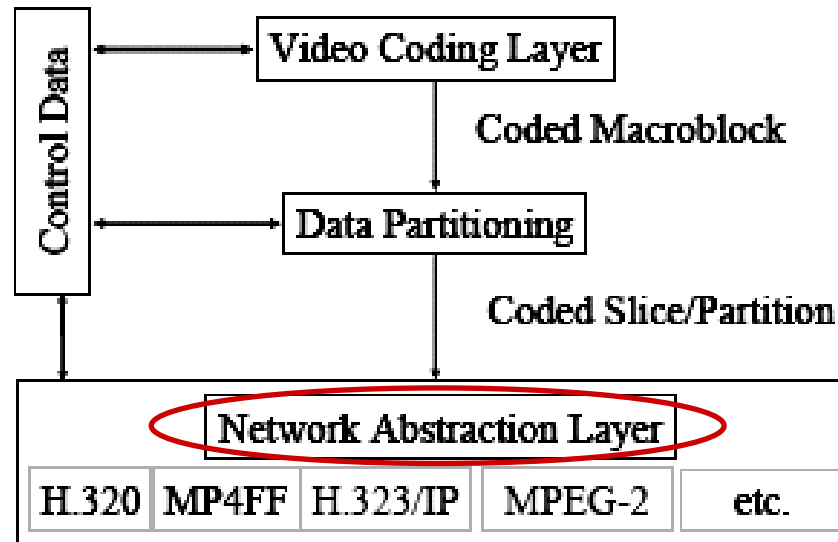
■ **CAVLC** (Context adaptive VLC)

- For transform coefficients
- No end-of-block, but number of coefficients is decoded
- Coefficients are scanned backwards
- Contexts are built dependent on transform coefficients

■ **CABAC** (Context-based binary arithmetic coding)

- For transform coefficients
- Usage of adaptive probability models for most symbols
- Exploiting symbol correlations by using contexts
- **Average bit-rate saving over CAVLC 10-15%**

H.264: Network Abstraction Layer (1)



H.264: Network Abstraction Layer (2)

- **NAL unit**: transport entity, e.g., contain one slice
- H.264/AVC standard defines a byte-stream format to transmit a sequence of NAL units as an ordered stream of bytes
- NAL unit boundaries need to be identified to obtain NAL units with correct size to guarantee integrity
- A byte-oriented HDLC-like framing including start codes (1 or 2 bytes) and emulation prevention is specified
- For simplified gateway operation, the emulation prevention on byte basis is applied to all raw byte sequence payloads (**RBSPs**).
- **MPEG-2 systems support!**

Thank You!

www.addpac.com

Technical Sales and Marketing

Phone : +82 2 568 3848

Fax : +82 2 568 3847

sales@addpac.com