

MPEG 및 H.264 압축

2006년 11월

동국대학교

전자공학과

원치선

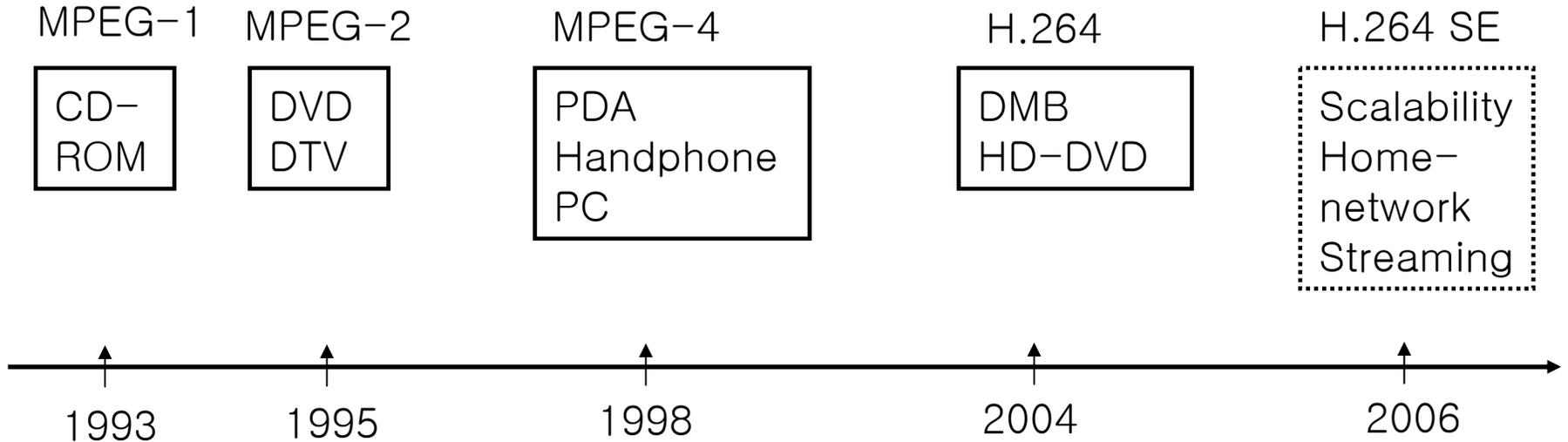
cswon@dongguk.edu

기존의 압축 표준 비교

Communication ← → Storage/Entertainment

	<i>H.261</i>	<i>H.263</i>	<i>MPEG-4</i>	<i>MPEG-2</i>	<i>MPEG-1</i>
<i>Date of Standard</i>	1990 5/94 revised	5/96 Version 1 1/98 Version 2	2/99 Version 1 2/00 Version 2	11/94	11/92
<i>Primary Applications</i>	Wireline Video Conferencing	Desktop/Wireless Videophone & conferencing	Web Authoring, DVD, Wireless Videophone, HDTV	Broadcast, DVD, HDTV	Digital, Storage Media
<i>Typical Video Rate</i>	128-384 kbps	20-384 kbps	20 kbps-6Mbps	4-6Mbps	1.5Mbps
<i>Typical Video Frame Size</i>	176x144 (QCIF) 352x288 (CIF)	176x144 (QCIF) 352x288 (CIF)	176x144 (QCIF) 352x288 (CIF) 720x480 (ITU-R 601) 1920x1080	720x480 (ITU-R 601) 1920x1080	352x240 (SIF)
<i>Typical Associated Audio Quality</i>	Speech	Speech	Speech, Music Stereo CD, Surround Sound	Surround Sound	Stereo CD Quality

MPEG Applications



Scalability?

- 기존의 비디오 코딩: 1번의 인코딩으로 1가지의 수신조건만 만족
- SVC: 1번의 인코딩으로 여러 해상도, 프레임율, 비트율 복원 가능
- Multi-type network channel, multiple device 환경에 최적

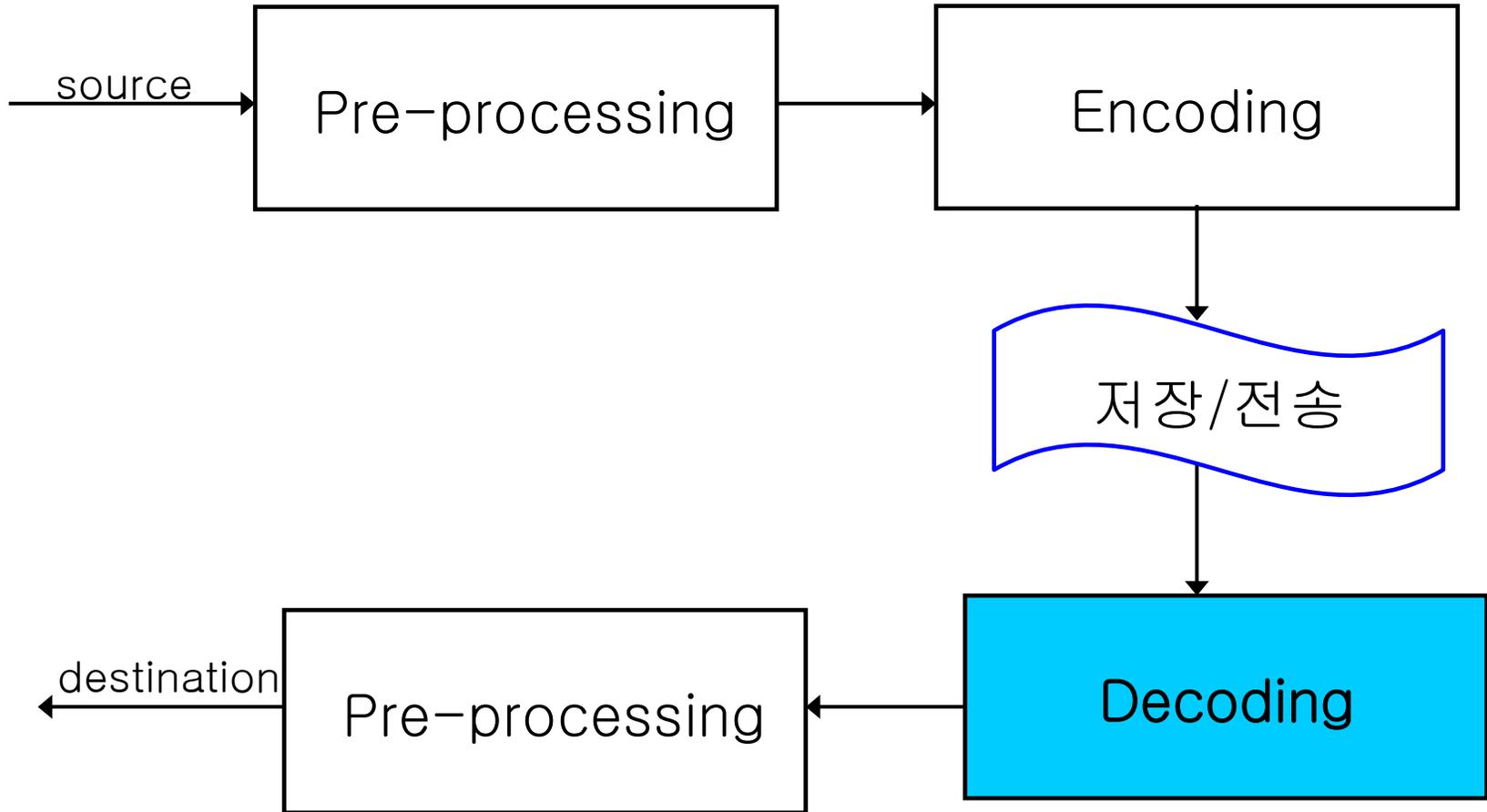
비디오 데이터 압축의 핵심 기술

기본원리: 비디오 데이터에 존재하는 공간적, 시간적, 그리고 통계적 중복성을 인간 시각적인 특성을 고려하여 제거

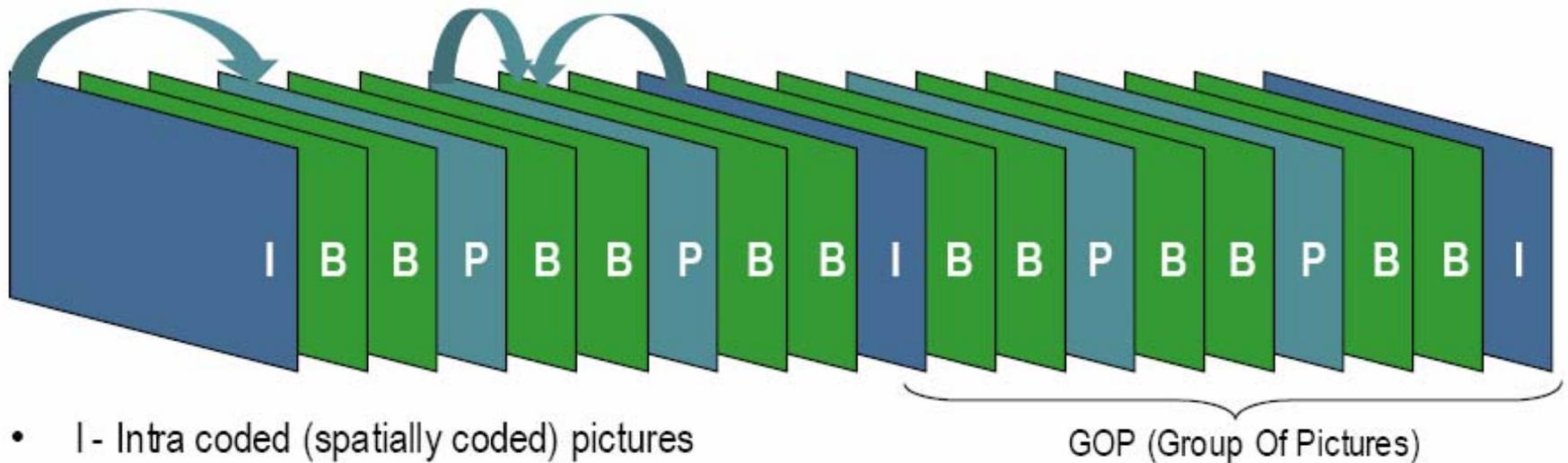
관련기술

- Prediction (예측)
 - Transformation (변환)
 - Quantization (양자화)
 - Entropy Coding (엔트로피 부호화)
- Hybrid Coding (MC + T + VLC)

영상 압축 표준의 범위

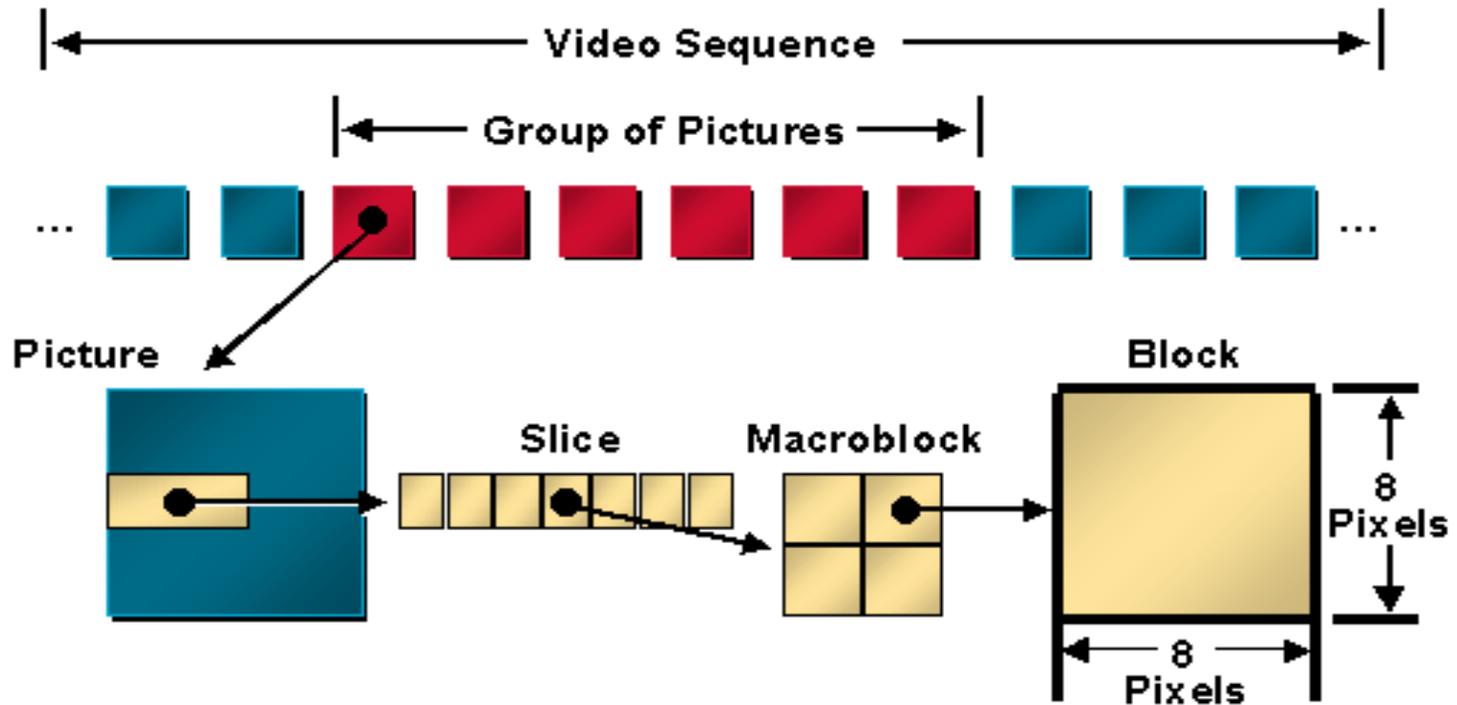


MPEG-1,2 Video Structure (1)



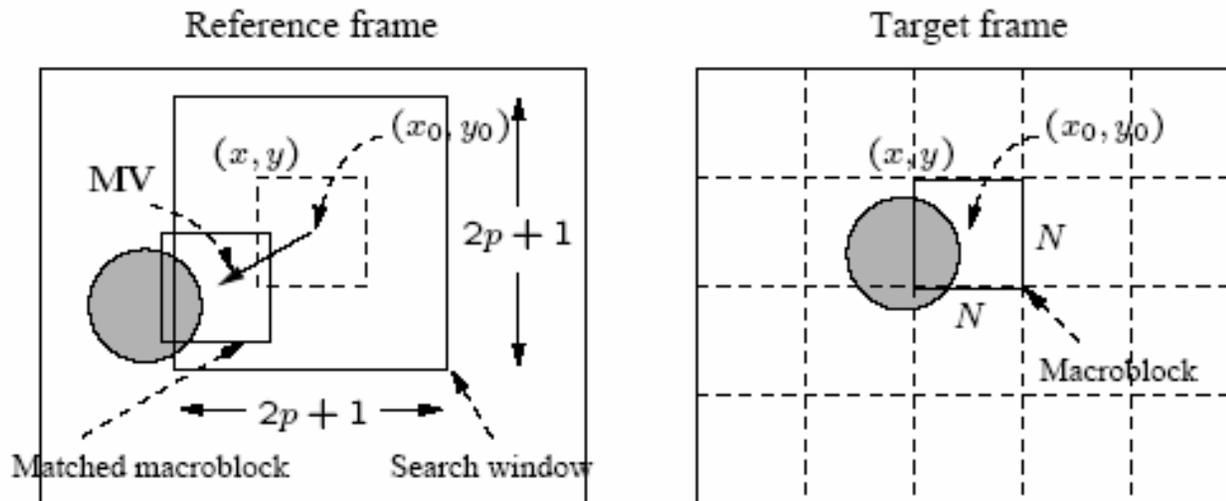
- I - Intra coded (spatially coded) pictures
 - Forms the anchor for a GOP
- P - Forward Predicted pictures
 - Predicted from previous I or P pictures
 - P picture made up of vectors showing where to get pixel data from in previous pictures and/or values that must be added to previous picture to get current pixel value
- B - Bidirectional Predicted pictures
 - Predicted from previous or later I or P pictures (never from other B pictures)
 - Made up of vectors showing where to get pixel data from in previous pictures

MPEG-1,2 Video Structure (2)

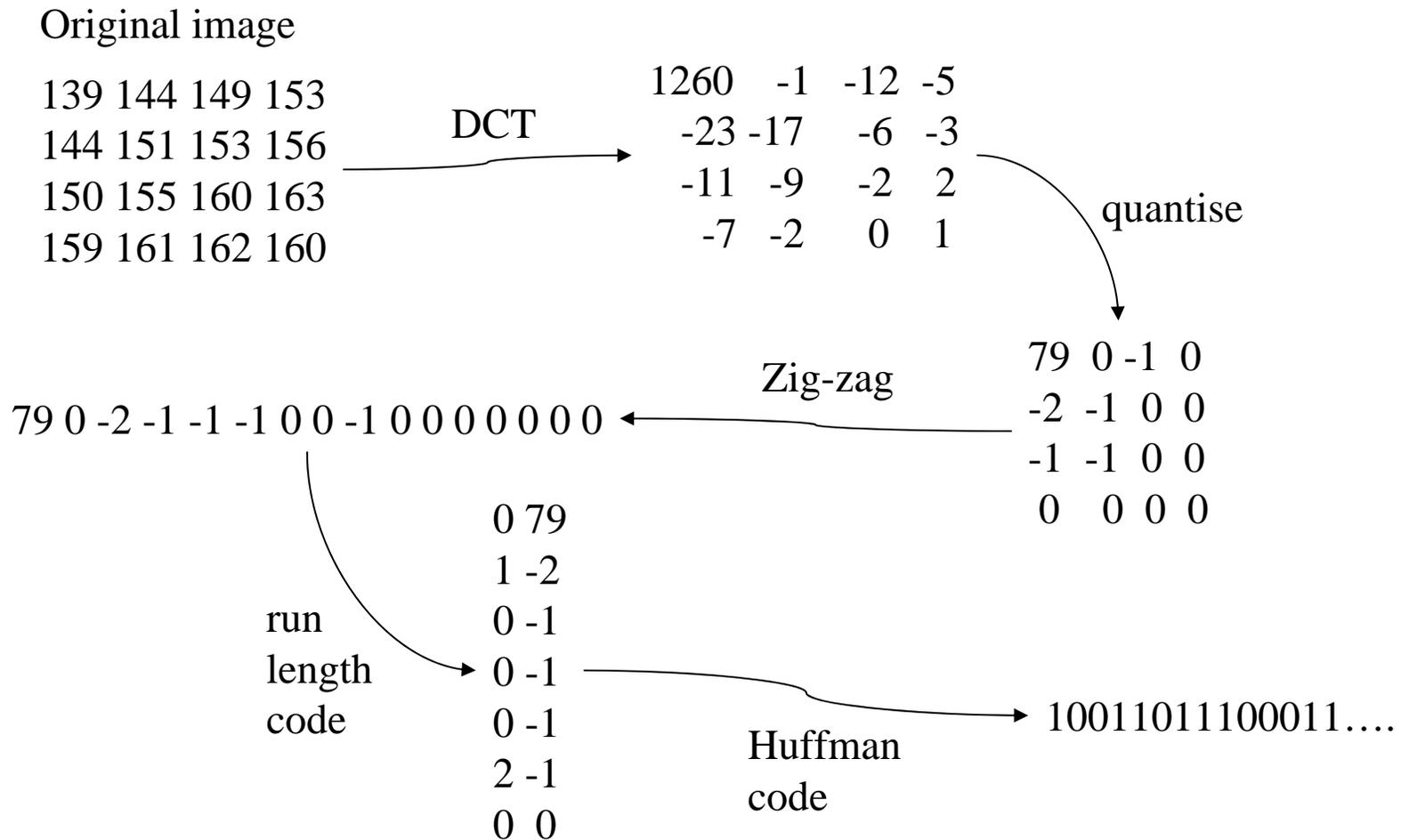


움직임 보상 (Motion Compensation)

- 각 프레임의 영상 공간을 $N \times N$ (luminance 영상은 $N=16$, 4:2:0의 부표본화가 적용되었을 때 chrominance 영상은 $N=8$)의 매크로블록 (macroblock) 단위로 분할
- 현재의 프레임을 타깃 프레임(target frame)으로 설정
- 타깃 프레임의 매크로블록과 이전 또는 이후 프레임 (참조 프레임 (reference frames))들 사이의 정합을 계산한다. 타깃 매크로 블록과 참조 프레임의 가장 잘 정합된 매크로블록의 변위가 움직임 벡터 (motion vector)이다.
- 전방향 예측 (forward prediction): 참조 프레임이 이전 영상인 경우
후방위 예측 (backward prediction): 참조 영상이 미래 영상인 경우
- 움직임 벡터 탐색 영역: 타깃 매크로 블록을 중심으로 $(2p+1) \times (2p+1)$



Block Transform Encoding



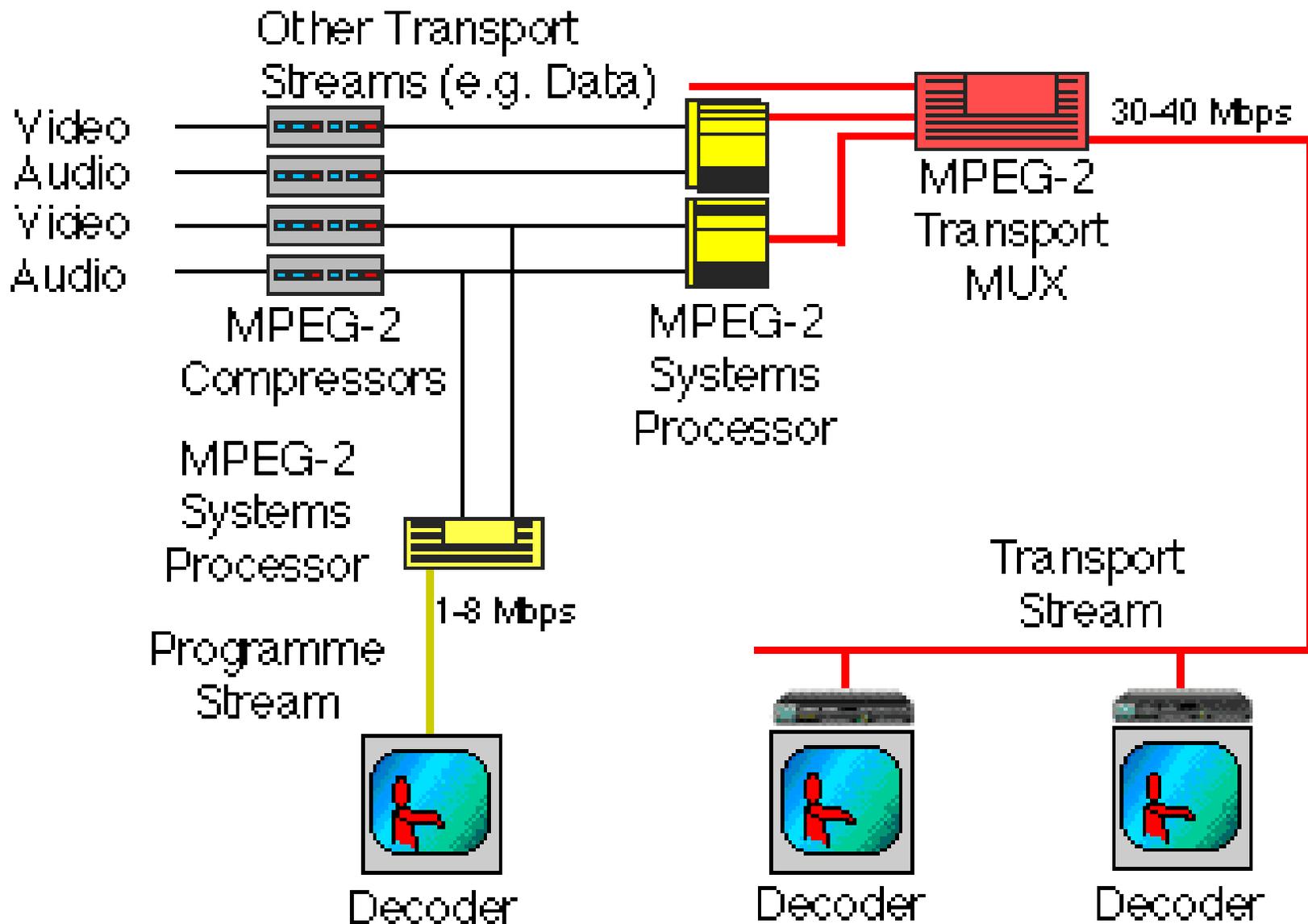
비디오 압축의 특성 파라미터

- Throughput of the channel: Data rate + ECC
- Distortion of the decoded video
- Delay (latency)
- Complexity

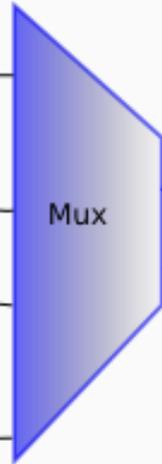
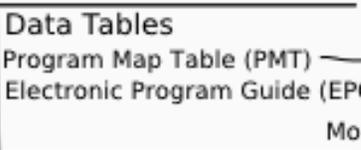
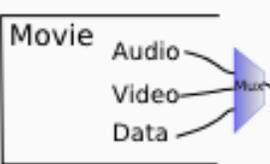
Problem Formulation

최대 허용 delay와 complexity하에서 bit-rate와 distortion의 Trade-off 최적화

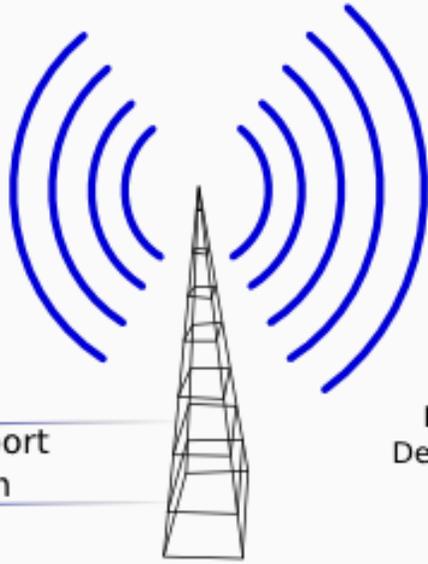
MPEG-2 SYSTEM



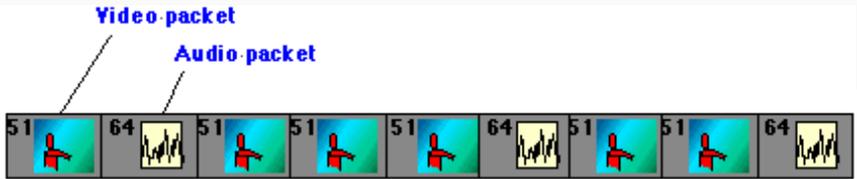
MPEG-2 SYSTEM (TS) (1)



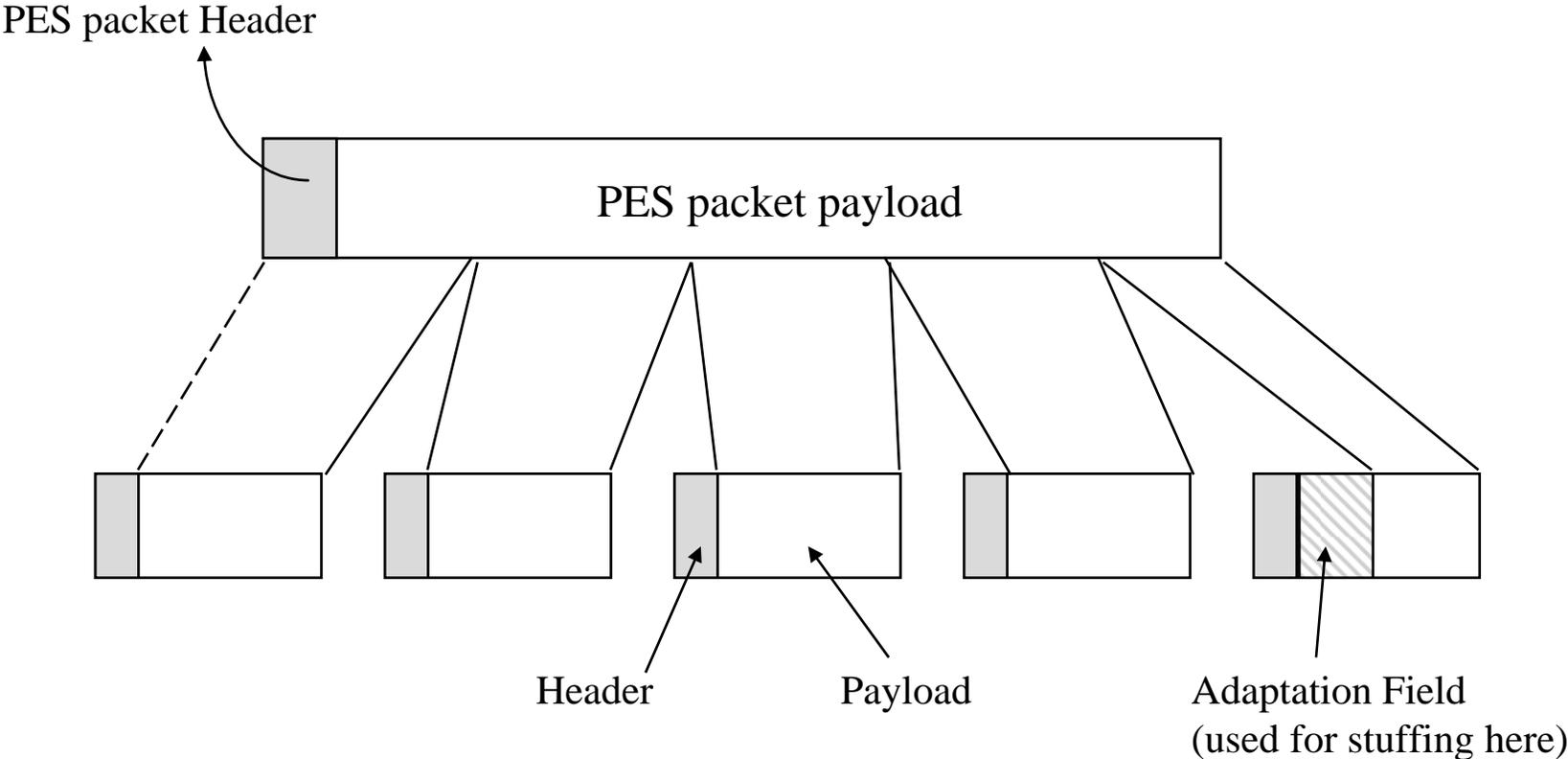
Transport Stream



ATSC Receiver
Receives the Transport Stream
Decodes it and displays on your TV



MPEG-2 SYSTEM (TS) (2)



Video packet
Audio packet



MPEG-2 SYSTEM (TS) (3)

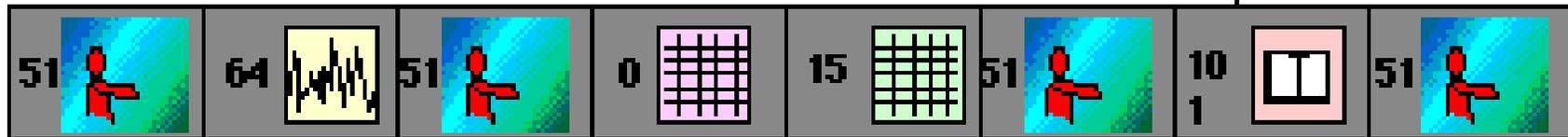
Video packet

Audio packet

Programme Association Table (PAT)

Programme Map Table (PMT)

Other packets



Packet header includes a unique Packet ID (PID) for each stream

PAT lists PIDs for programme map tables

Network Info = 10
Prog 1 = 15
Prog 2 = 301
Prog 3 = 511
etc

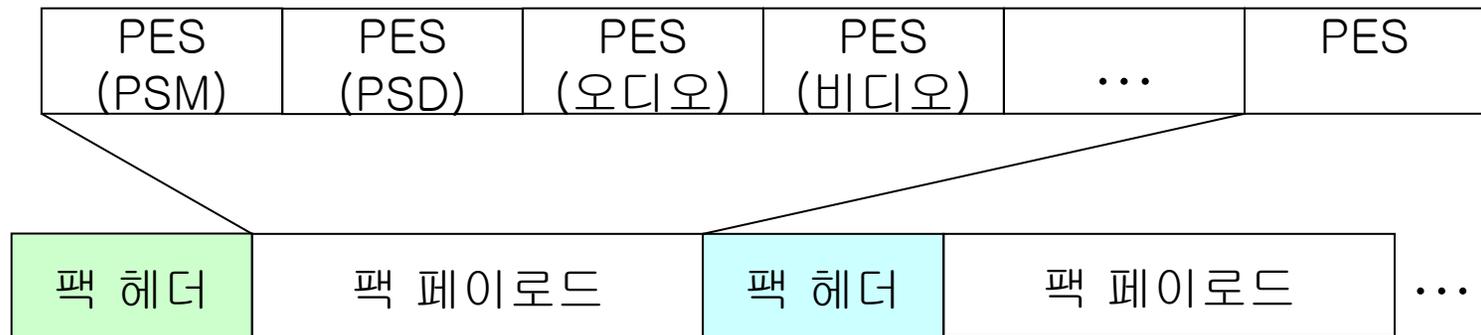
PMT lists PIDs associated with a particular programme

Video = 51
Audio (English) = 64
Audio (French) = 66
Subtitle = 101 etc.

Programme guides, Subtitles, Multimedia data, Internet packets, etc

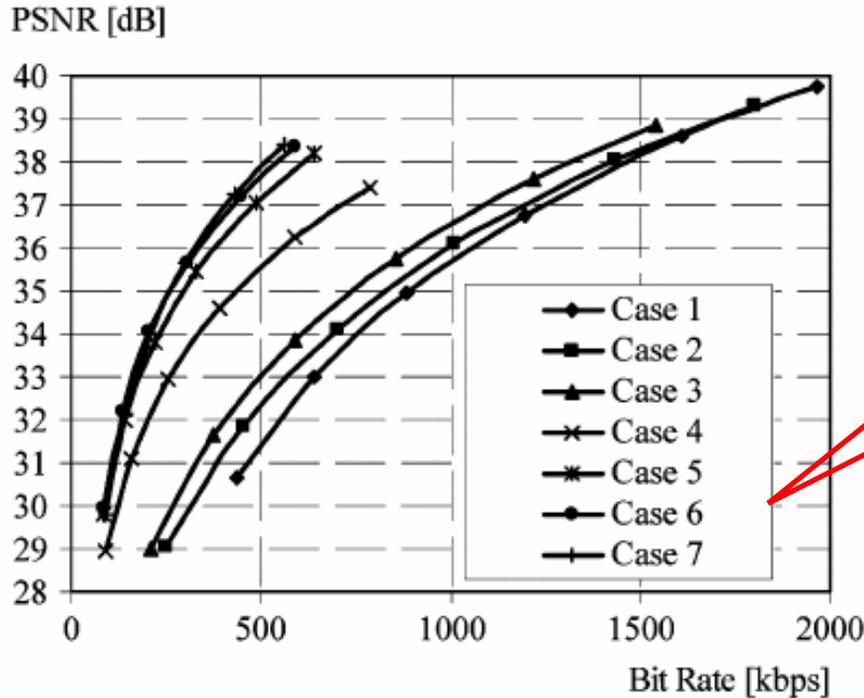
MPEG-2 SYSTEM (PS)

- Program Stream
 - Primarily intended for storage and retrieval from storage media
 - Grouping of video, audio, and data elementary streams that have a common time base
 - Each program stream consists of only one program
 - Useful in error free environments
 - Large packet size
 - Packets size may be variable (hard for decoder to predict start and end of packets)
 - DVD standard uses the MPEG-2 Program Stream



PSM (Program Stream Map) : 프로그램을 구성하는 비디오와 오디오의 속성 정보
PSD (Program Stream Directory): 임의 액세스가 가능한 액세스 단위의 주소정보 및 표시시각을 저장한 표

핵심 기술의 압축 성능 비교



새로운 기능을 추가할 때 복잡도의 증가나 추가 정보를 나타내야 할 비트 수 증가에 대해 주의 요망

IEEE Proceedings,
pp.18-31, Jan 2005

- Case 1) Spatial-transform + Intra coding only (e.g., JPEG)
- Case 2) Adding Skip mode to form a CR coder
- Case 3) Adding residual difference coding, but with only zero-valued MVs
- Case 4) Adding integer-precision MC with blocks of size 16x16 luma samples
- Case 5) Adding half-sample-precision MC
- Case 6) Allowing some 16x16 regions to be split into four blocks of 8x8 luma samples each for MC
- Case 7) Increasing MV precision to quarter-sample

MCP (Motion-Compensated Prediction)

항목	내용	비고
Fractional-sample-accuracy	<ul style="list-style-type: none"> 정수 픽셀 단위 이하의 MV 정밀도 픽셀간 가상의 픽셀은 보간 공간 정밀도가 높을 수록 MV의 정확도 향상 	<ul style="list-style-type: none"> 1/2 샘플: MPEG-1,2, H.263 1/4 샘플: MPEG-4 ASP (luma), H.264 1/8 샘플: H.264 Chroma
영상 경계에서의 MV	영상 경계 확장 (extrapolation)	H.263부터 채택
Bipredictive MCP	<ul style="list-style-type: none"> 2개의 MV의 평균치 Uncovered 영역이나 평탄하고 지속적인 움직임의 경우 유리 	MPEG-1 이후 모든 표준
Variable block size MCP	<ul style="list-style-type: none"> MV 당 픽셀의 면적 움직임 필드의 정확도와 MV를 표현하기 위한 비트 수 사이의 trade-off 	H.264에서 16x16 이하의 움직임 벡터 단위 (macroblock) 허용
Multipicture MCP	<ul style="list-style-type: none"> MV를 검출하기 위해 참조하는 앞뒤 영상의 수 많을수록 장거리 통계적 상관성 반영 	H.264에서 5개까지의 복호화된 참조 프레임 허용

H.264/AVC 개요

- main goals
 - enhanced compression performance
 - network-friendly video applications
- Features
 - 4 x 4 integer DCT (avoid mismatch problem)
 - variable block-size motion compensation
 - quarter-sample-accuracy motion compensation
 - multiple reference frame motion compensation
 - directional spatial prediction for intra coding
 - more effective entropy coding
 - context-adaptive variable-length coding (CAVLC)
 - context-adaptive binary arithmetic coding (CABAC)
 - more flexible encoding features
 - flexible slice size
 - flexible macroblock ordering (FMO)
 - arbitrary slice ordering (ASO)

MPEG-4와 H.264/AVC

Functionalities	MPEG-4 Visual	H.264
Supported data type	Rectangular video frames and fields, arbitrary-shaped video objects, still texture and sprites, synthetic or synthetic-natural hybrid video objects, 2D and 3D mesh objects	Rectangular video frames and fields
Number of profiles	19	3 (+ High Profiles for FExt)
Compression Efficiency	Medium	High
Support for video streaming	Scalable coding	Switching slices
Motion compensation minimum block size	8 x 8	4 x 4
MV accuracy	$\frac{1}{2}$ or $\frac{1}{4}$	$\frac{1}{4}$
Transform	8 x 8 DCT	4 x 4 DCT approximation
Built-in deblocking filter	No	Yes

H.264 구조

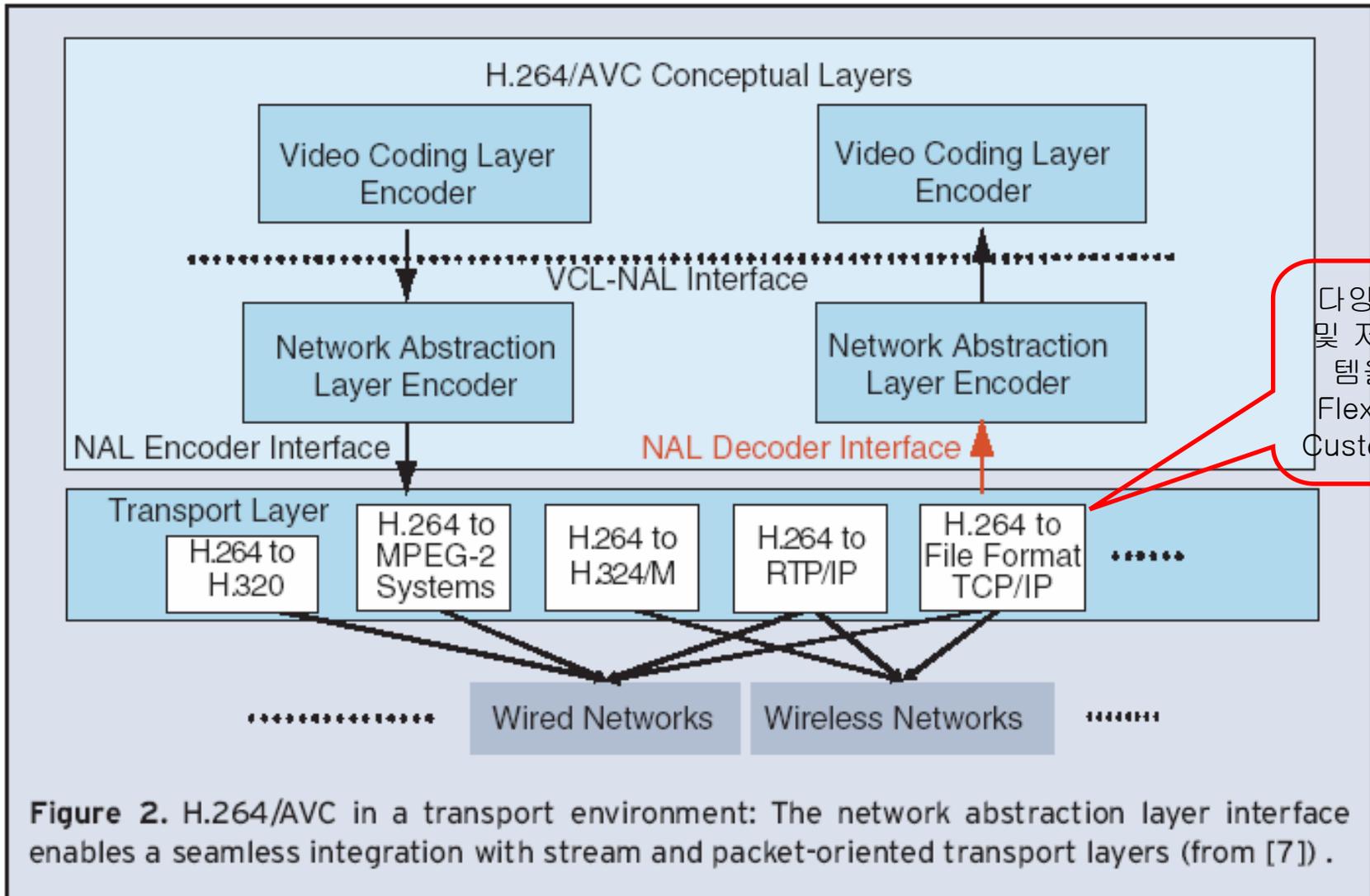


Figure 2. H.264/AVC in a transport environment: The network abstraction layer interface enables a seamless integration with stream and packet-oriented transport layers (from [7]).

NAL Structure

항목	기능
NAL Unit	<ul style="list-style-type: none"> • 부호화가 완료된 비디오 데이터는 정수 개의 바이트 단위인 NAL Unit로 패킷화 됨 • NAL의 첫번째 바이트는 데이터 형을 나타내는 헤더 • VCL NAL Unit → 비디오 데이터 포함 Non-VCL NAL Unit → 파라메타나 enhancement 정보를 위한 보조 데이터 • 손실에 대한 강인성을 위해 중요한 데이터 (MV 값 및 type)와 덜 중요한 데이터 (residual DCT 계수)를 서로 다른 NAL Unit에 배치하여 서로 다른 중요도로 전송 → 데이터 partitioning
Parameter Sets	<ul style="list-style-type: none"> • Sequential Parameter Sets (SPS): 부호화된 일련의 비디오 픽처에 적용 • Picture Parameter Sets (PPS): 하나 혹은 그 이상의 개별 픽처의 복호에 적용 • SPS와 PPS의 분리로 robustness 증대 • Robustness를 위해 SPS와 PPS를 반복적으로 전송 가능
Access Units	<ul style="list-style-type: none"> • 단일 복호된 픽처와 관련된 VCL 및 non-VCL의 셋을 말함 • 픽처의 모든 MB + 에러 극복을 위한 추가 데이터 (redundant slices)

H.264 구조 (비디오 포맷)

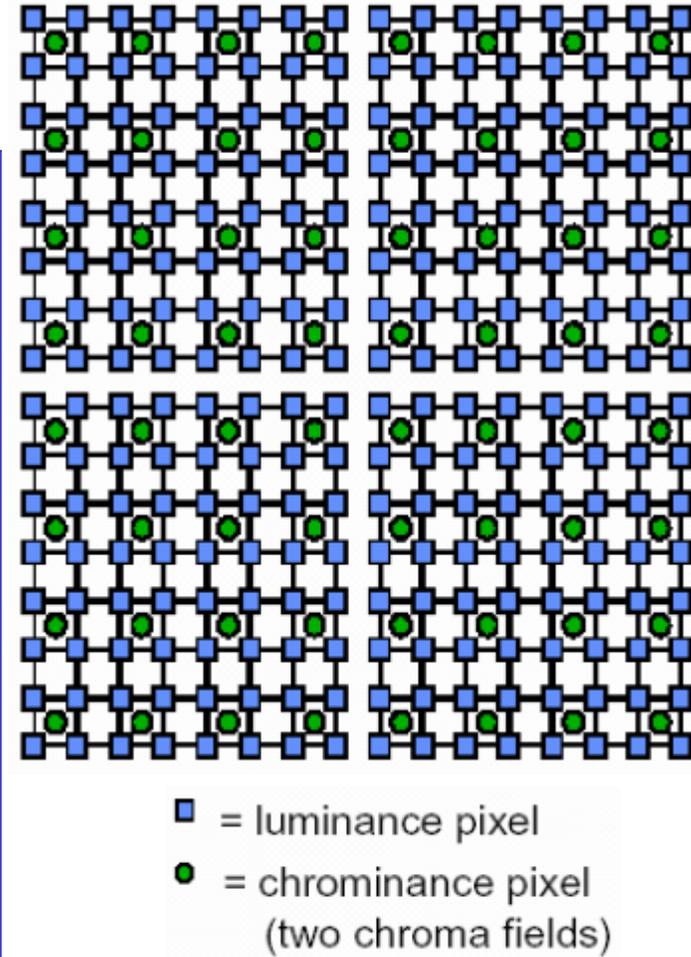
비디오 포맷

- 4:2:0 순차 및 비월주사 비디오 지원

비월주사 관련 모드

- **Frame mode**: 두 개의 필드를 하나의 프레임으로 통합하여 하나의 픽처로 부호화
- **Field mode**: 각 필드를 독립된 픽처로 간주
- **Macroblock-adaptive frame/field mode (MBAFF)**
전체 프레임을 하나의 픽처로 부호화하지만 Prediction과 Residual 부호화를 위해 수직방향으로 인접한 두 개의 MB를 선택적으로 두 개의 필드로 분리 가능

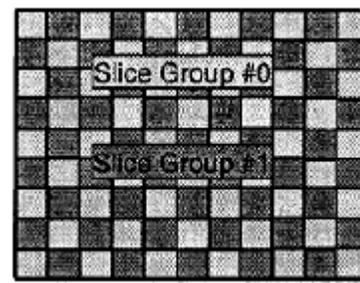
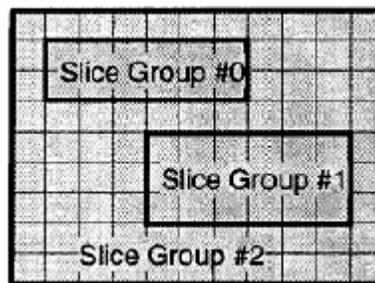
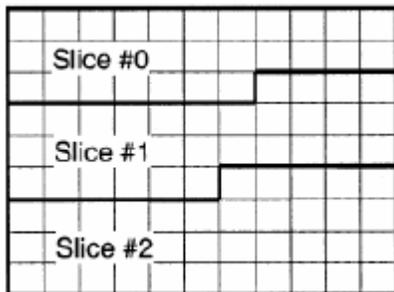
- 위의 세 가지 모드의 선택은 각 프레임 별로 적응적 수행
- 첫 번째 두개의 모드 만을 선택하는 경우 Picture-adaptive frame/field (**PAFF**)라 부름
- 필드모드인 경우: MCP는 참조 필드 사용, 계수 스캔이 지그재그와 다름, 수평에지에 강한 de-blocking 사용 안함



H.264 구조 (Slices)

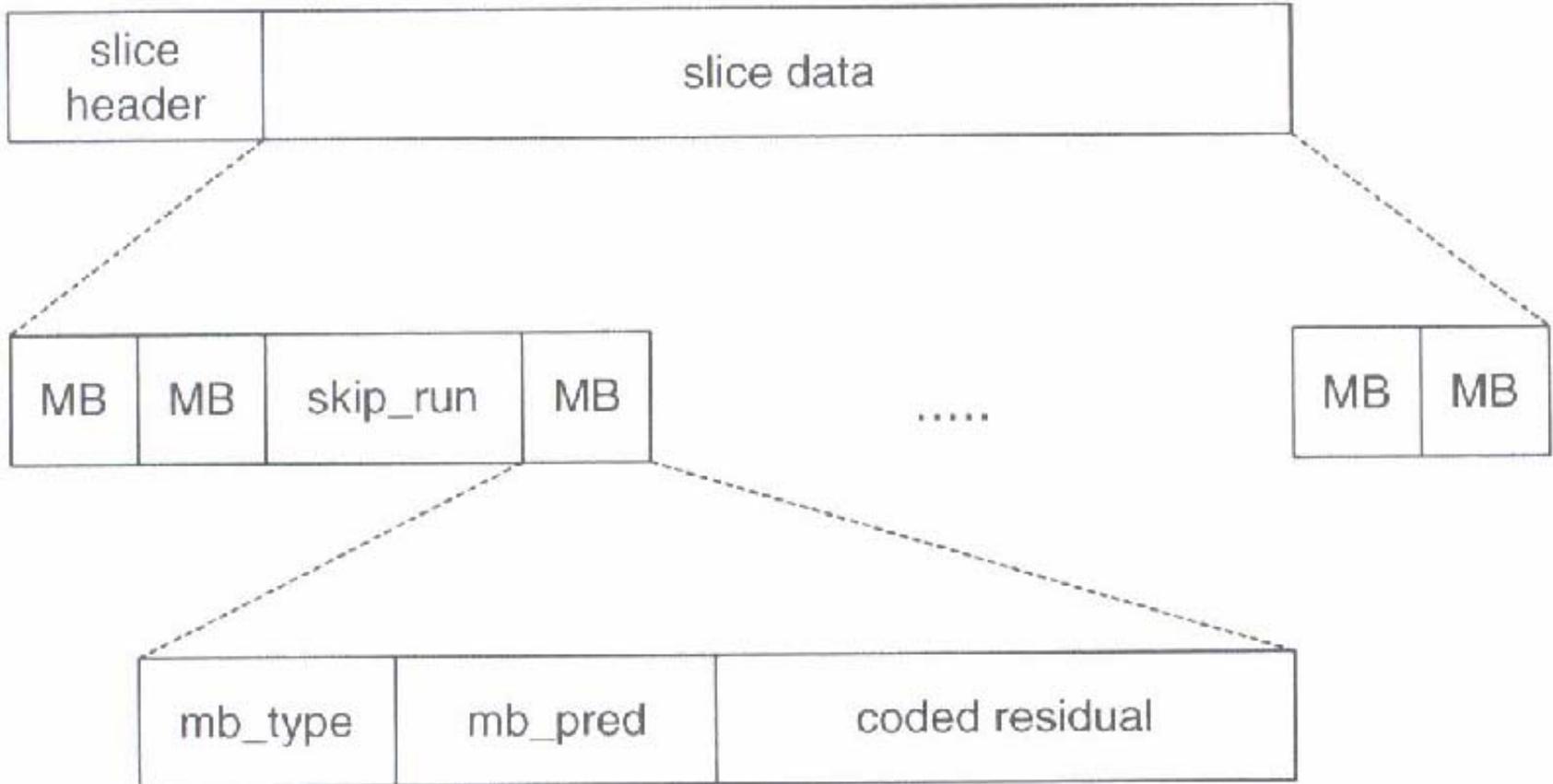
Slices

- 픽처 내의 raster scan 방향의 MB의 집합임.
- 영상 복호화의 기본 독립 단위 → error resilience (resynchronization)
 - Error resilience 기능은 FMO (Flexible Macroblock Ordering) 기술을 사용하여 더 강화시킬 수 있음
 - FMO는 픽처 내의 MB를 복수의 집단으로 나누어 각각을 다른 슬라이스에 배정함 (예, “foreground” slice group and “leftover”, checkerboard 등)
 - ASO (Arbitrary Slice Ordering): FMO와 관련된 것으로 ASO가 사용되지 않으면 손실 강인성과 지연 감소를 위해 슬라이스 순서대로 첫번째 MB은 증가하는 순서로 배열되어야 함
- 슬라이스들 사이의 독립성을 이용하여 병렬처리 가능.



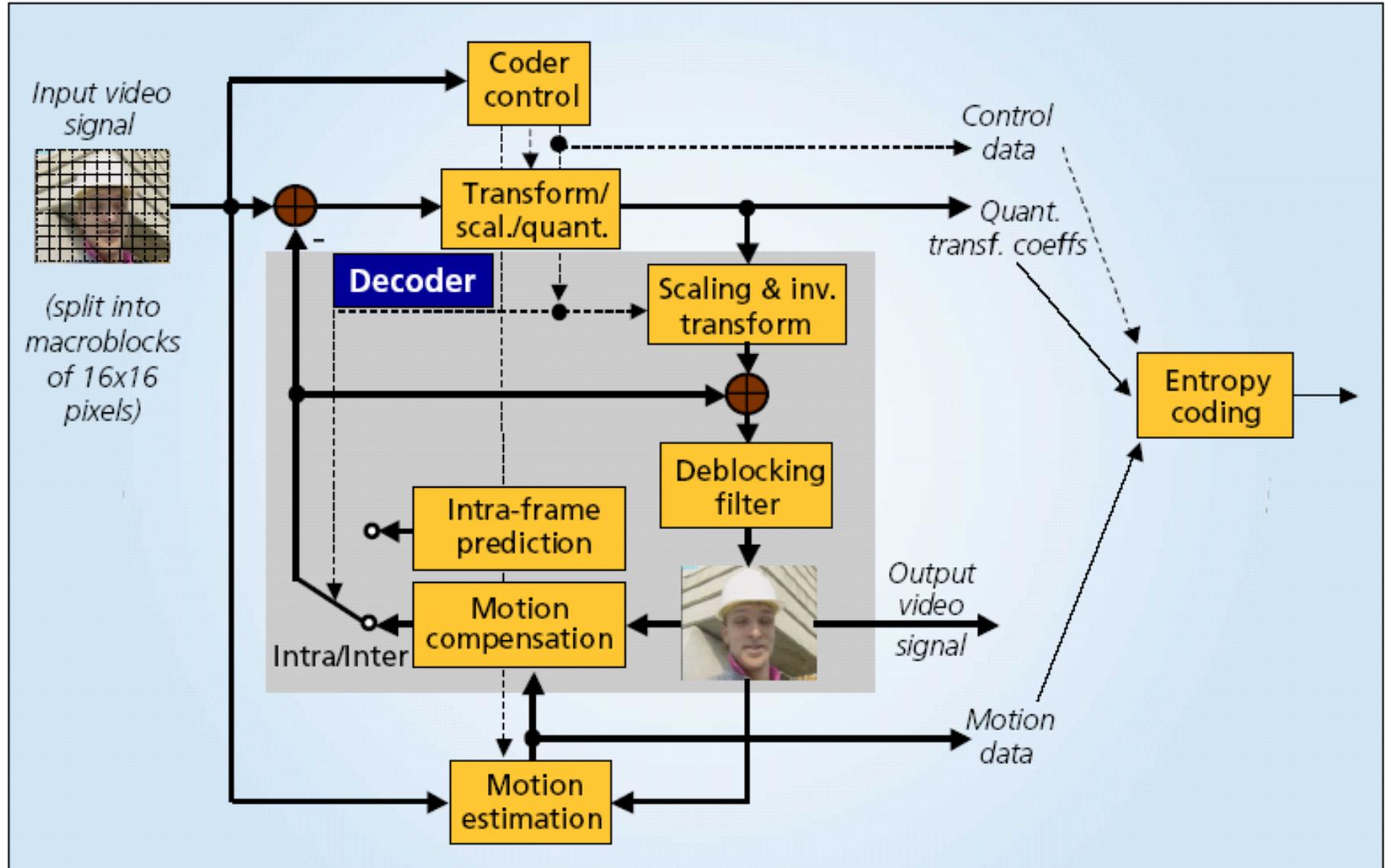
H.264 구조 (Slices)

Slice Syntax

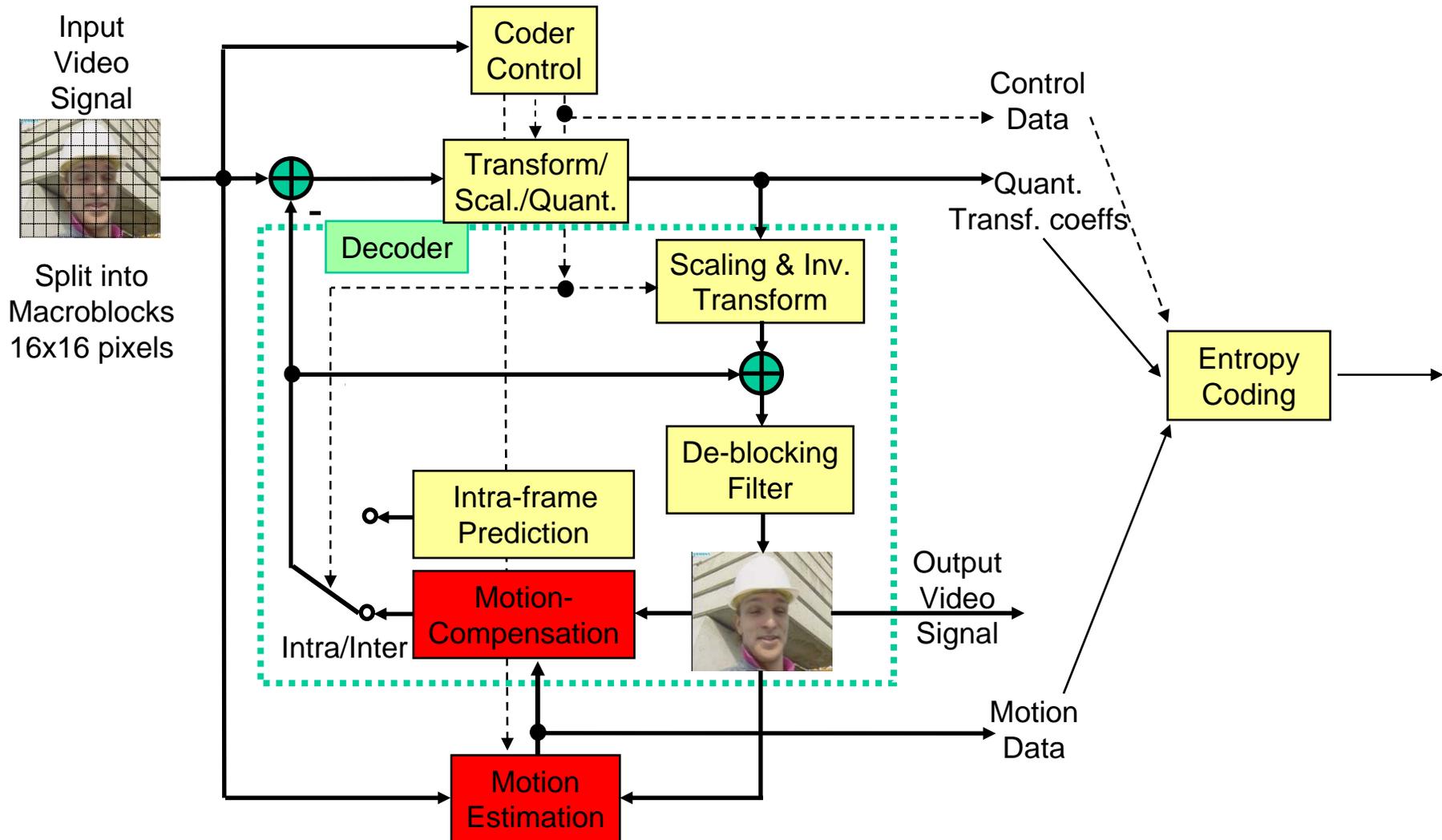


H.264/AVC CODEC Structure

Basic structure of H.264/AVC for a macroblock

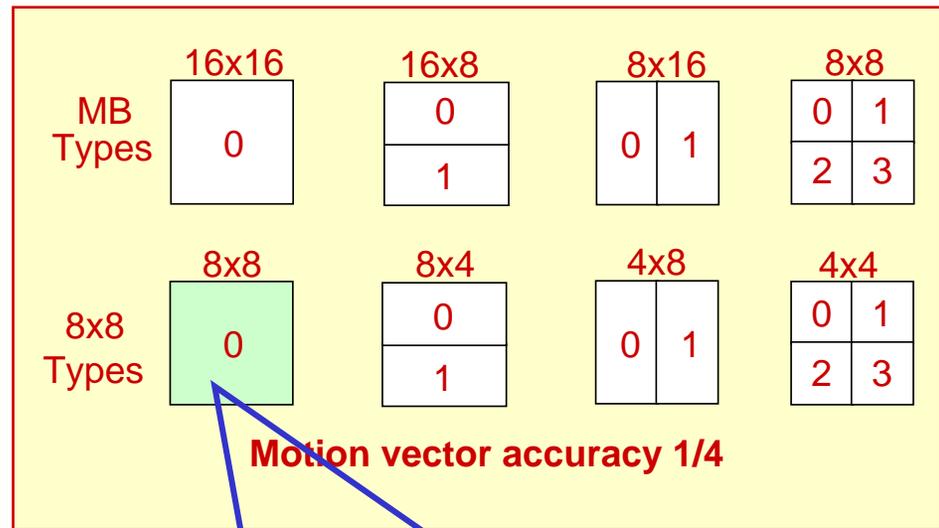


MC/ME

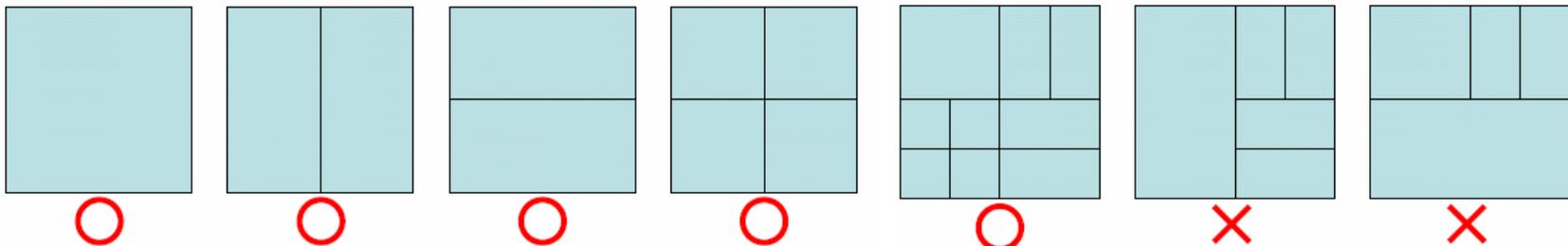


Inter Frame Prediction (P Slices)

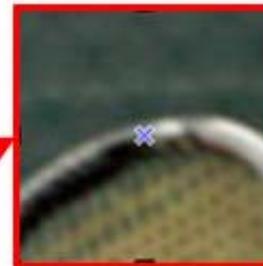
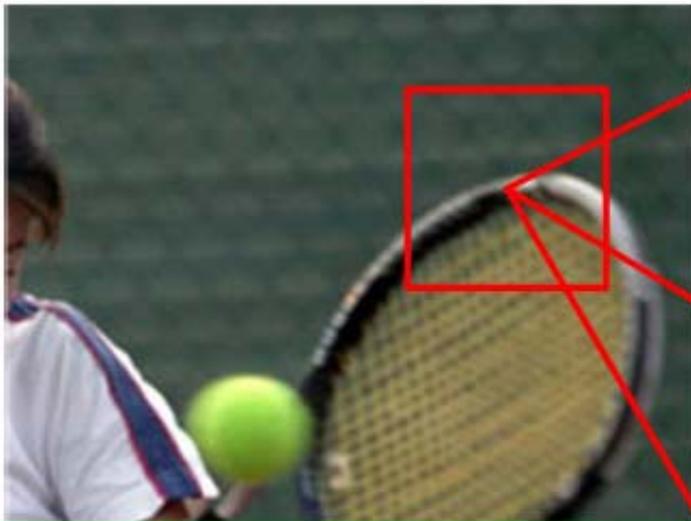
- 다양한 블록 사이즈
- 1/4 sample accuracy
 - 6 tap FIR filtering to 1/2 sample accuracy (수평 및 수직방향)
 - simplified filtering (정수 위치와 1/2 위치의 값을 평균) to 1/4 sample accuracy
- Picture boundary에서도 MV 계산
- MV는 이웃 블록으로부터의 예측이나 중간값으로 차분 부호화
- Multiple reference pictures
- P_Skip 모드: 양자화된 예측에러도 MV도 보내지 않음, index 0 (list 0)에 위치한 영상을 참조로 P_16x16 처럼 복원 → 변화가 없거나 일정한 움직임 (slow panning) 영역에서 유리



8x8이 선택되면 각 8x8 블록마다 그것이 더 작은 블록 8x4, 4x8, 혹은 4x4로 나누어 지는지 나타내는 추가적인 선택스 필요

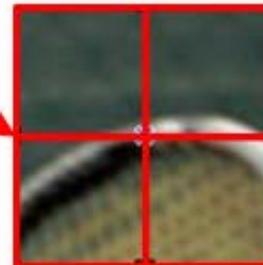


Motion Compensation Accuracy



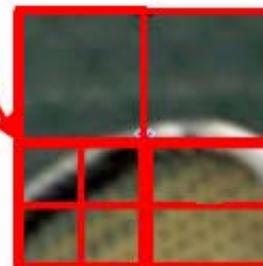
MPEG-2

- 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

Multiframe Motion Compensation

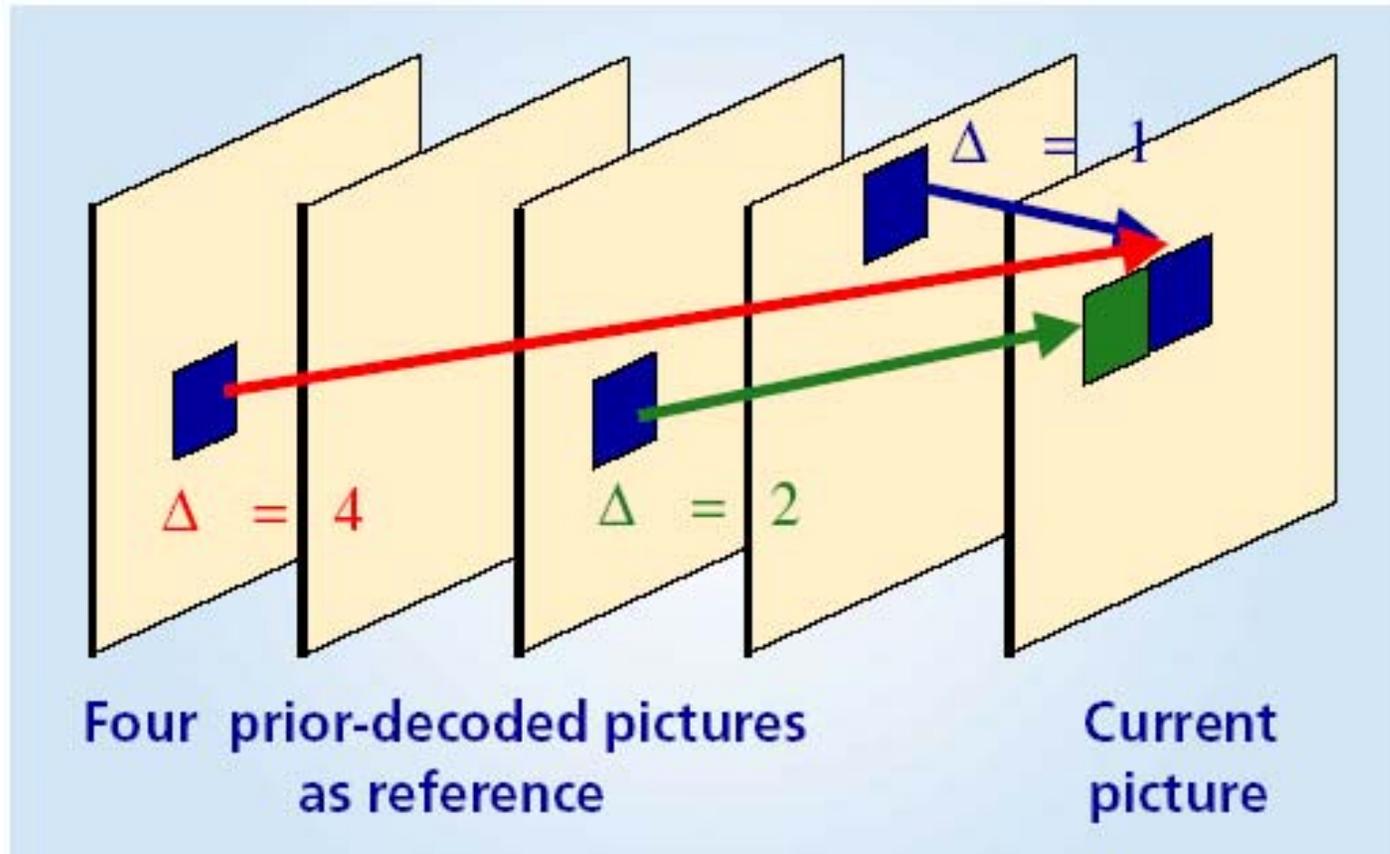


Figure 3

Multi-frame motion compensation. In addition to the motion vector, also picture reference parameters (Δ) are transmitted. The concept is also extended to B pictures as described below.

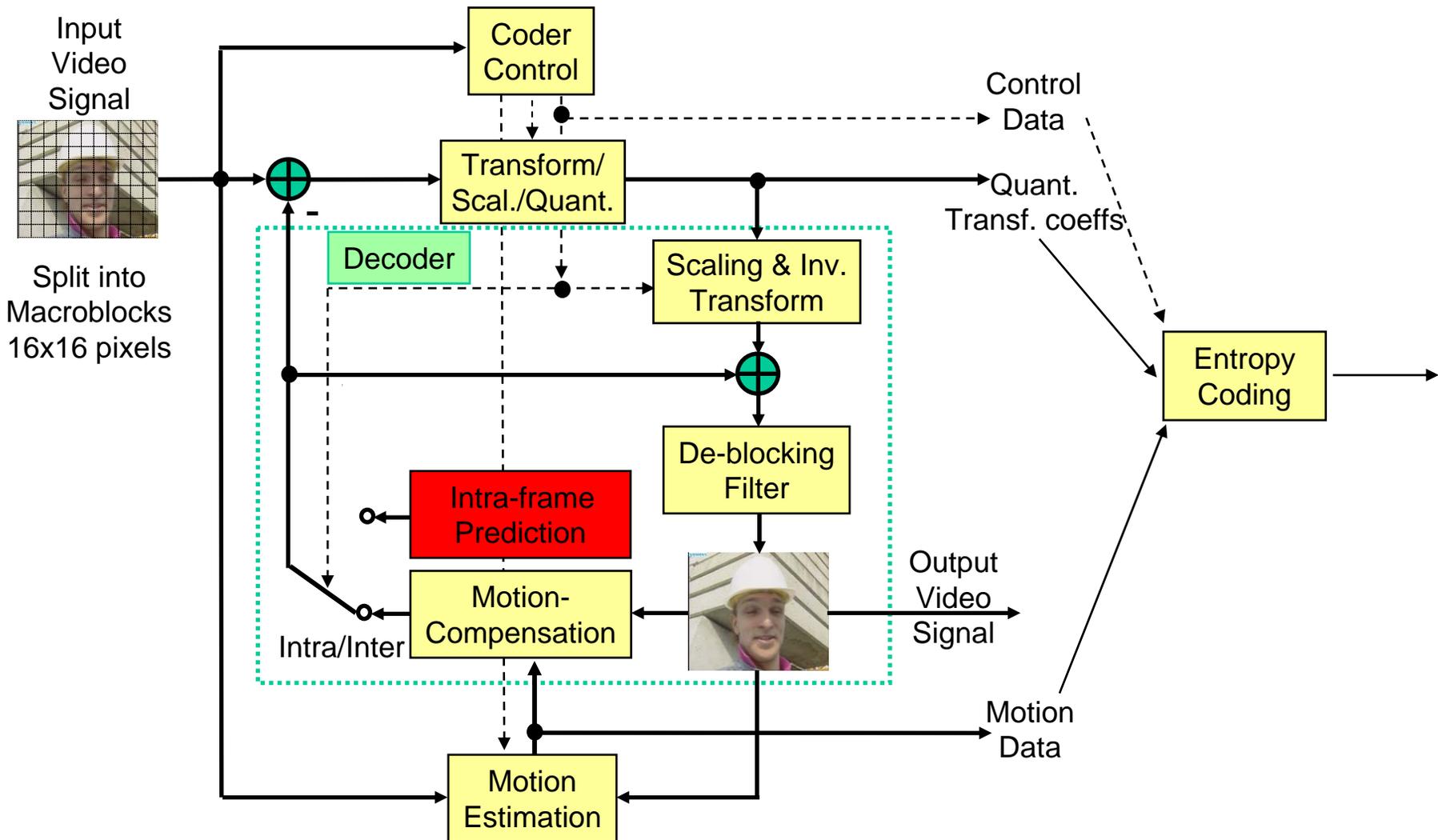
Multiframe Motion Compensation (계속)

- 전방향 참조 프레임 최대 5장 후방향 1장
- Reference Picture ListX
 - Encoder와 Decoder에 함께 존재
 - 전방향 List0, 후방향 List1
 - DPB(Decoded Picture Buffer)에 Reference Picture 저장
 - Short term or long term reference picture

Multiframe의 장단점

- 장점
 - 반복적인 움직임에 효과적
 - 객체의 움직임으로 가려진 영역에 효과적
 - 부화소 단위의 움직임에 효과적
 - 주기적인 빛의 변화에 효과적
 - 바로 이전의 영상에 잡음이 있는 경우 효과적
- 단점
 - 인코더 복잡도 증가
 - 많은 메모리 요구

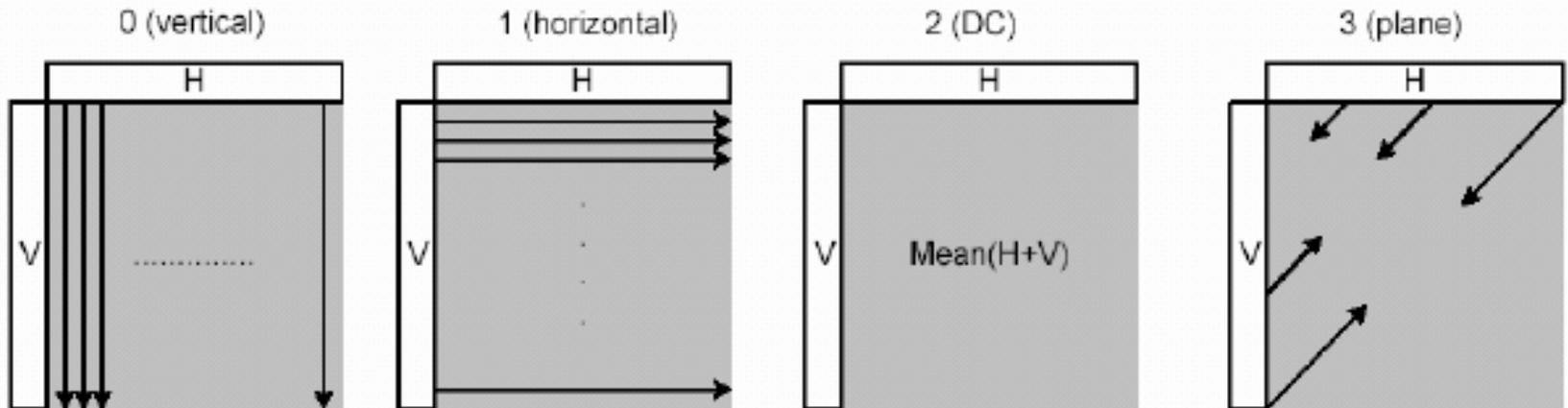
Intra Prediction



Intra Prediction (계속)

- Two modes for luma block
 - Intra 4x4 : 9 modes, texture area
 - Intra 16x16 : 4 modes, flat area
 - I_PCM
- One mode for chroma block
 - Similar to intra 16x16

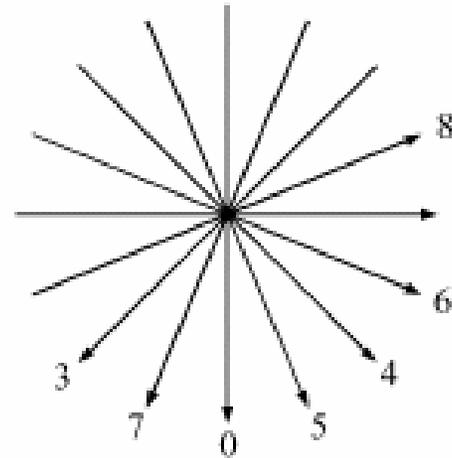
Intra 16x16



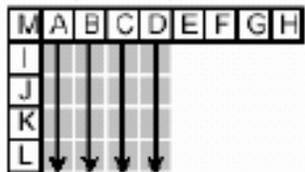
Intra Prediction (계속)

Intra 4x4

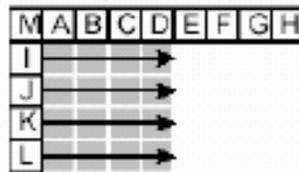
M	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				



0 (vertical)



1 (horizontal)



2 (DC)



3 (diagonal down-right)



4 (diagonal down-left)



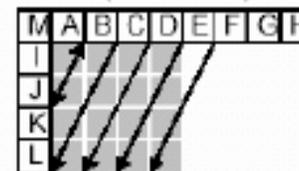
5 (vertical-right)



6 (horizontal-down)



7 (vertical-left)



8 (horizontal-up)



Intra Prediction (계속)

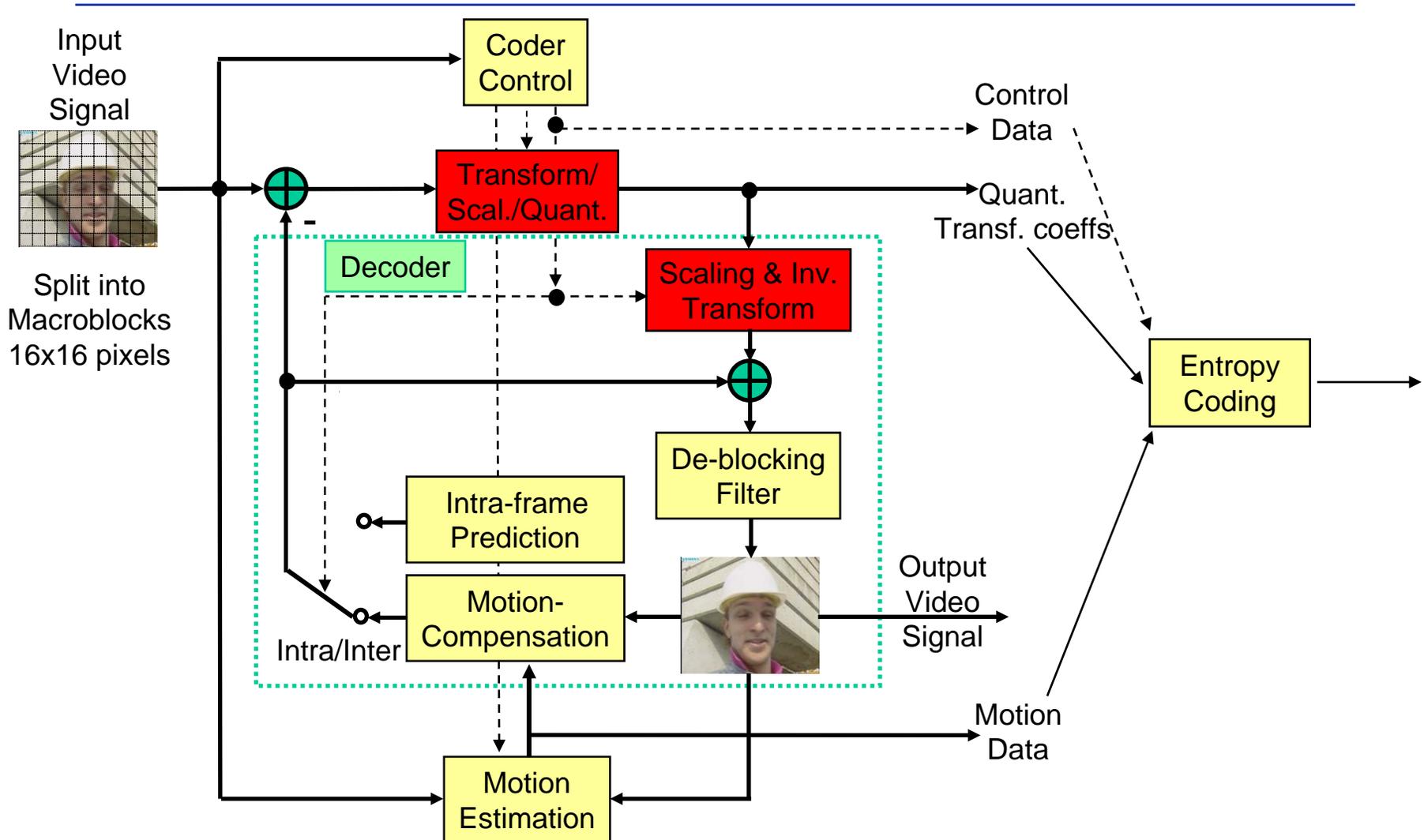
- 모든 슬라이스 type에 대해 2가지의 기본적인 Intra Prediction: Intra_4x4, Intra_16x16 (특별한 경우 I_PCM이라 불리는 3번째의 인트라 코딩 존재)
- Intra_4x4는 각 4x4 luma block에 사용되며 상세한 영상 내용 부위에 적합
- Intra_16x16은 상당히 평탄한 영상 부위에 적합
- 공간영역에서의 Intra coding이 이웃의 블록에 Inter prediction과 함께 쓰일 때 에러가 시공간적으로 전파될 수 있으므로 constrained Intra coding mode를 사용할 수 있음.
- Intra_16x16 mode: 16x16 luma 성분이 일시에 예측됨
- Intra_16x16 mode의 4가지 mode: vertical, horizontal, DC, plane
- plane mode: position-specific linear combination, 천천히 변하는 영역에 유리
- Intra MB의 chroma sample은 luma의 Intra_16x16과 유사
- I_PCM Intra MB type은 prediction이 수행되지 않고 보상없이 원 데이터가 그대로 전송됨 → 1) 부호화를 위해 필요한 비트 수가 원데이터의 비트 수보다 많을 수 없다, 2) 특정 부위의 Lossless 부호화 효과

- Rate distortion optimization (RDO)을 위한 모든 가능한 모드의 수는 $4 \times (9 \times 16 + 4) = 592$ different RDO calculations

Rate-Distortion Optimization

- 최적 Intra & Inter Prediction Mode와 블록 사이즈의 결정을 위해 Rate-Distortion Optimization (RDO) 알고리즘 사용
- Rate and distortion are function of QS.
- Rate and distortion are also functions of the percentage of null quantized coefficients (“zeros”) in the block.

Transform Coding



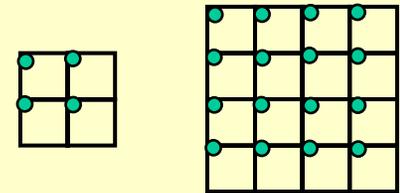
Transform Coding (계속)

- DCT의 역변환이 정수 계산이므로 역변환 mismatch 문제 해소 및 복호화 복잡도를 낮춤
- DCT 블록의 양자화 계수는 zig-zag 스캔
- Chroma의 2x2 DC 계수는 raster-scan

4x4 Block Integer Transform

$$H_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

- Repeated transform of DC coeffs for 8x8 chroma and 16x16 Intra luma blocks (Intra_16x16)



Hadamard Transform for luma

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$H_3 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Hadamard Transform for chroma

Quantization

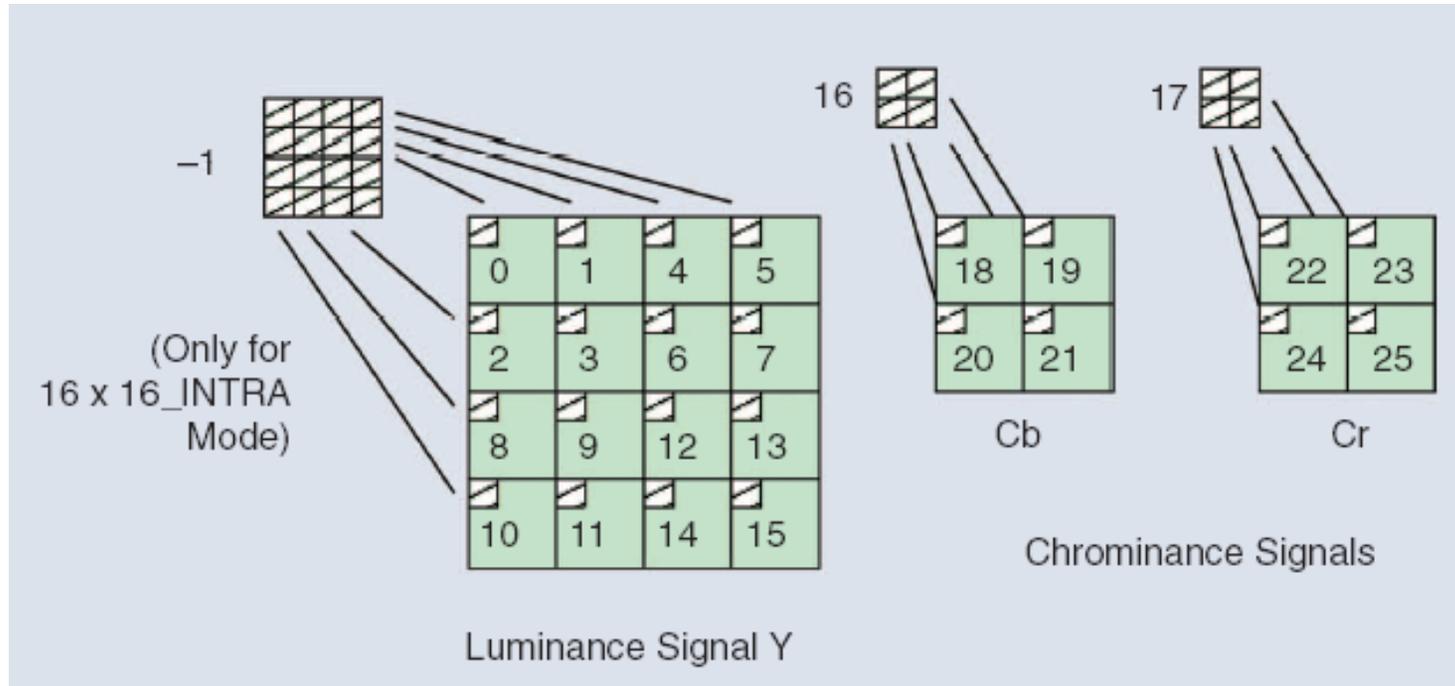
- 모든 DCT 계수들은 스칼라 양자화됨
 - Logarithmic step size control
 - Extended range of step sizes
 - Smaller step size for chroma (cf. H.263 Annex T)
 - Table-driven: 12.5% increase in qStep per 1-QP increase

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

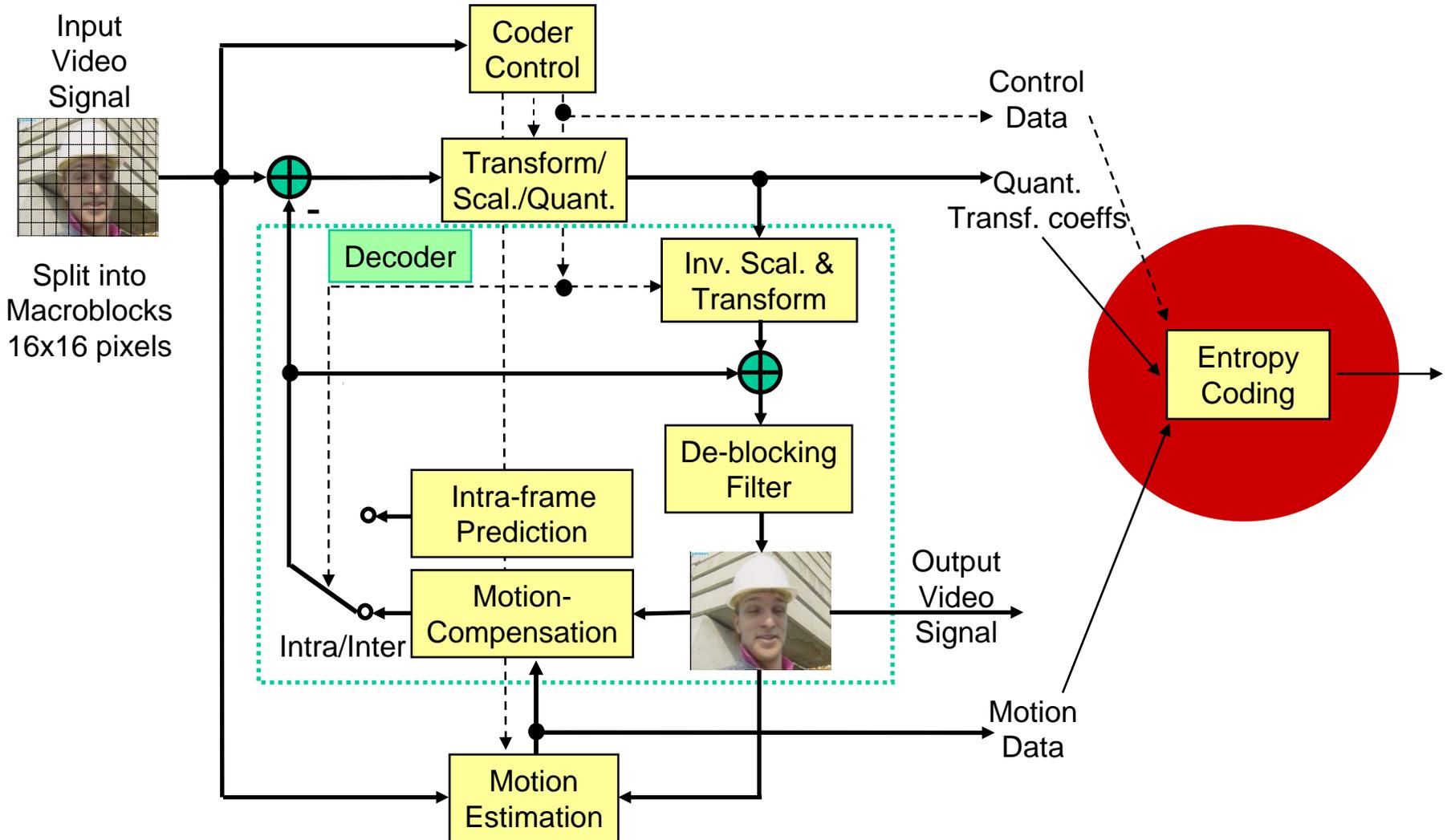
- Reconstruction is 16-bit multiply, add, shift

Transmission Order

- DC 계수 (-1, 16, 17) 우선 전송



Entropy Coding

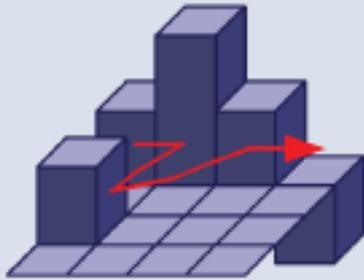


CAVLC

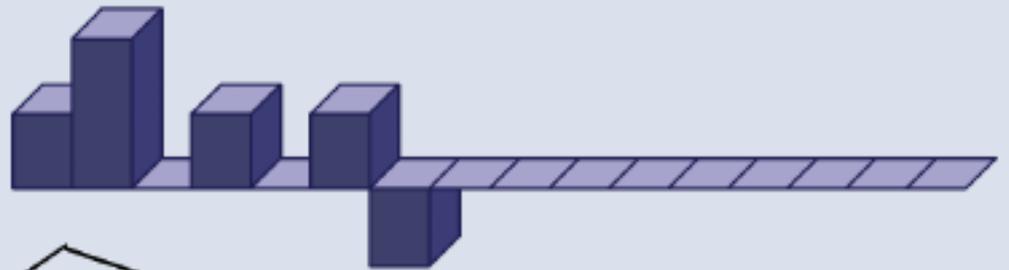
- Preamble: # of nonzero coefficients, # of T1 (Trailing 1's) (4개의 VLC 테이블 중 한 개 사용)
- Significant coefficient들의 부호와 레벨 값이 스캔의 역 순으로 부호화 (각 레벨 값을 부호화할 때 6개의 테이블 중 하나를 선택하여 이전에 레벨에 따라 부호화 됨)
- 마지막 non-zero값 이전의 zero의 수
- 각 significant level 당 이전에 발생한 연속적인 zero의 수
- 각 부호화 단계 마다 최대 가능 zero의 값에 따라 run 값의 부호화를 위한 테이블 선택
- 전체 32개의 VLC 테이블 사용
- 테이블의 일부는 저장 공간이 필요없이 on-line 계산에 의해 결정
- 기존의 run-length 부호화에 비해 2-7%의 비트율 절감

CAVLC (계속)

4 x 4 Block of Quantized Transform Coefficients

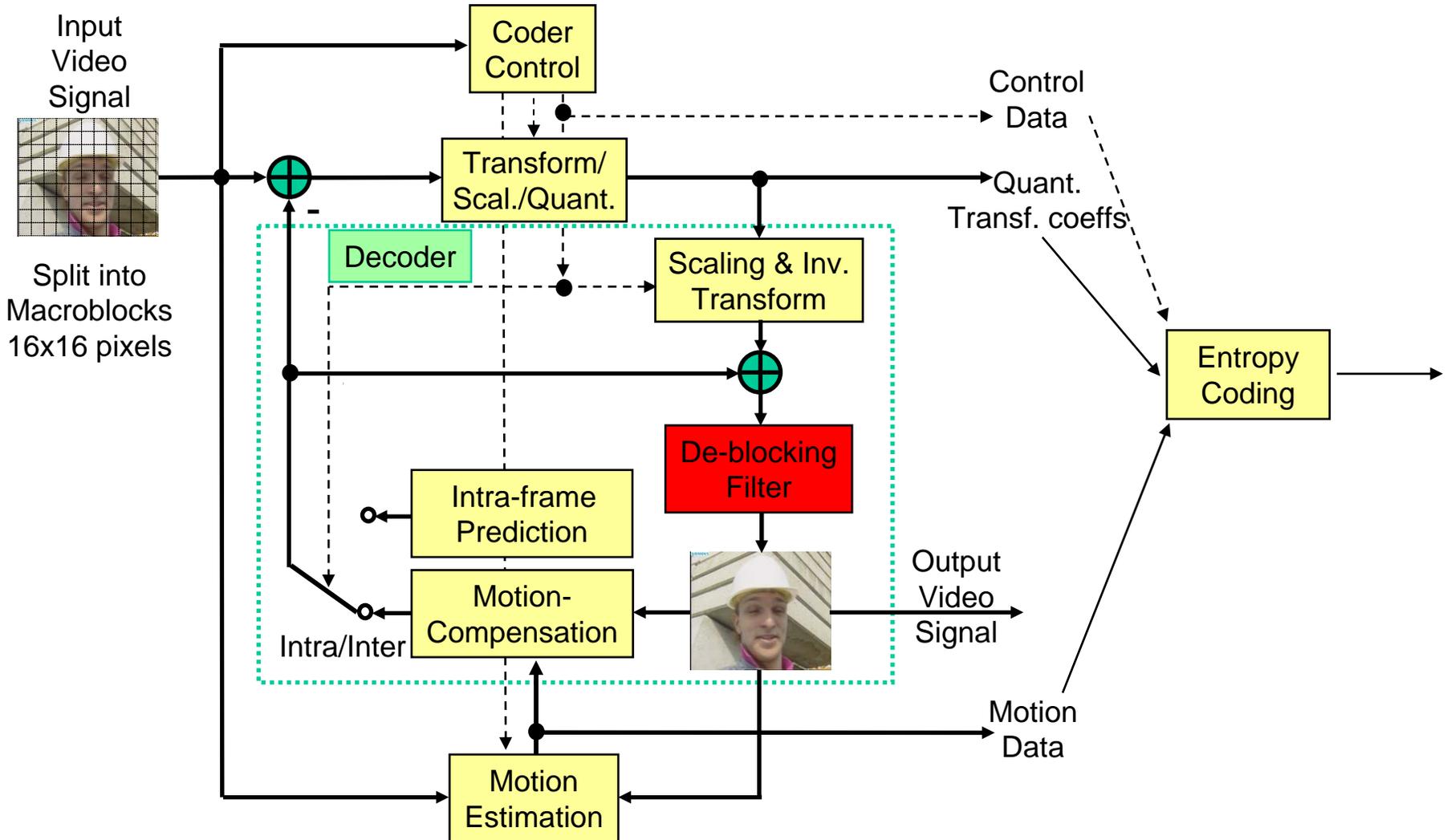


Array of Scanned Quantized Transform Coefficients



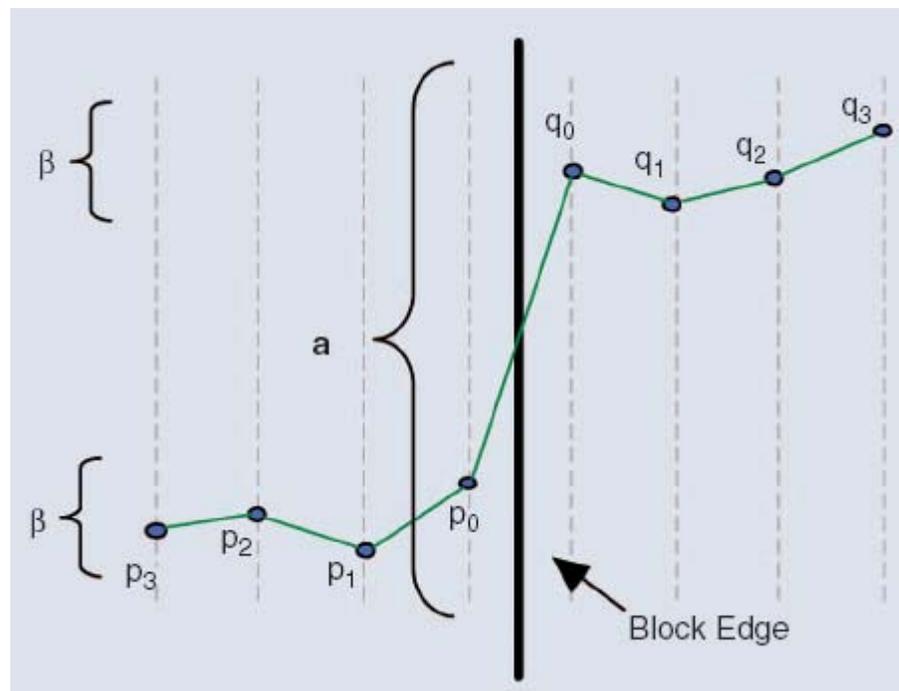
	CAVLC
Preamble	Number signif. coeff: 5 Trailing 1's (T1): 3
Reverse Pre-coded Syntax Elements	Sign T1: -1, 1, 1 Levels: 2, 1 Total Zeros: 2 Run Before: 0, 1, 1, (0)

De-blocking Filter



De-blocking Filters (계속)

- De-blocking 필터를 Coding Loop내에 놓은 이유: (1) MC의 참조 영상을 모두 필터링 된 깨끗한 영상으로 사용 (2) 복호기에 De-blocking 필터를 의무적으로 설치토록하여 최소 화질 보장
- 3개의 레벨에 따른 적응 처리
 - (1) On Slice Level: global strength가 비디오 시퀀스의 개별 특성에 의해 조절
 - (2) On block edge level: inter/intra 예측 결정, 움직임 차이, 두 개의 참여 블록의 부호화된 residual의 존재 여부에 따라 필터 강도 조절 (0~4)
 - (3) On sample level: 모든 에지를 가로지르는 샘플값들을 분석하여 실제 에지인지 블록 경계인지 구분
- De-blocking 필터는 필터처리를 하지 않은 것에 비해 객관적인 화질을 유지 하면서 5-10%의 비트율을 줄임.
(주관적 화질 향상은 더 큼)



De-blocking Filters (계속)

Deblocking filter: Highly compressed decoded inter picture



1) Without Filter



2) with H264/AVC Deblocking

H.264 프로파일

High Compression quality, e.g. broadcast, but high complexity

Extended Profile

Streaming

High Profile

CABAC

Main Profile

SP/SI Slices

Data Partitioning

MB-level frame/
Field switching

B Slices

Weighted
Prediction

Interlace

I/P Slices

De-blocking
filter

CAVLC

Multiple Ref.
Frames

1/4 sample MC

Intra Predict.

FMO
and ASO

Redundant
Slices

Adaptive Tr
4x4 or 8x8

HVS weighting
matrices

Predictive
Lossless

Monochrome
format

Low complexity and low delay: Lower capability plus error resilience (video conferencing wireless)

Baseline Profile

FRExt Amendment

- Fidelity Range Extensions (FRExt) Amendment 1: 2004년 7월 완성
- Content-contribution, content-distribution, Studio editing and post-processing의 응용을 위해
 - Use more than 8 bits per sample of source video accuracy
 - Use 4:2:2 or 4:4:4
 - Alpha-blending
 - Use very high bit-rates
 - Use very high resolution
 - Achieve very high fidelity – even representing some parts of the video lossless
 - Avoid color-space transformation rounding error
 - Use RGB color representation
- Produce the *High* profiles: support all features of the prior Main profile and additionally support an adaptive transform coding and perceptual quantization scaling matrices
- Applications: HD-DVD, BD-ROM, DVB standards for European broadcasting television

H.264 Levels (1)

레벨

- 주어진 프로파일 내에서도 모든 가능한 신택스를 사용할 수 있도록 구현하는 것은 비경제적이므로 각 프로파일 당 레벨 (Level)을 정의
- 레벨은 비트 스트림 내의 신택스 요소의 값에 대한 제한 (한정)
- H.264/AVC에 각 프로파일에 대해 15개의 레벨이 정의 됨
- 각 레벨은 비트 스트림의 상한치 혹은 디코더 처리 율의 하한치를 규정
 - QCIF로 부터 4kx2k의 영상의 크기
 - 디코더 처리율: 초당 1485 에서 983040 블록 처리
 - 움직임 벡터 범위: $[-64, +63.75]$ 에서 $[-512, +511.75]$

H.264/AVC Levels (2)

Level Number	Typical Picture Size	Typical frame rate	Maximum compressed bit rate (for VCL) in Non-FRExt profiles	Maximum number of reference frames for typical picture size
1	QCIF	15	64 kbps	4
1b	QCIF	15	128 kbps	4
1.1	CIF or QCIF	7.5 (CIF) / 30 (QCIF)	192 kbps	2 (CIF) / 9 (QCIF)
1.2	CIF	15	384 kbps	6
1.3	CIF	30	768 kbps	6
2	CIF	30	2 Mbps	6
2.1	HHR (480i or 576i)	30 / 25	4 Mbps	6
2.2	SD	15	4 Mbps	5
3	SD	30 / 25	10 Mbps	5
3.1	1280x720p	30	14 Mbps	5
3.2	1280x720p	60	20 Mbps	4
4	HD Formats (720p or 1080i)	60p / 30i	20 Mbps	4
4.1	HD Formats (720p or 1080i)	60p / 30i	50 Mbps	4
4.2	1920x1080p	60p	50 Mbps	4
5	2kx1k	72	135 Mbps	5
5.1	2kx1k or 4kx2k	120 / 30	240 Mbps	5

Complexity of Codec Design

- Codec design includes relaxation of traditional bounds on complexity (memory & computation) – rough guess 2–3x decoding power increase relative to MPEG-2, 3–4x encoding
- Problem areas:
 - Smaller block sizes for motion compensation (**cache access issues**)
 - Longer filters for motion compensation (**more memory access**)
 - Multi-frame motion compensation (**more memory for reference frame storage**)
 - More segmentations of macroblock to choose from (**more searching in the encoder**)
 - More methods of predicting intra data (**more searching**)
 - Arithmetic coding (**adaptivity, computation on output bits**)

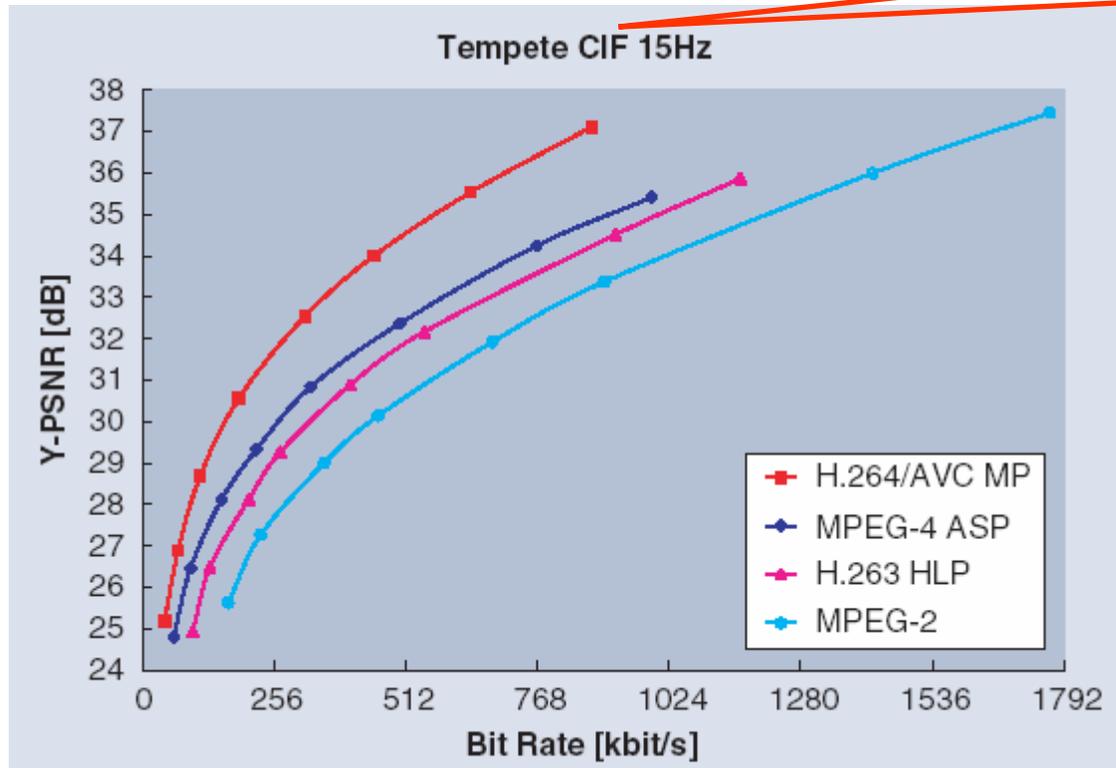
Error Detection for H.264/AVC

- Constraints imposed on the H.264/AVC video bitstream syntax
 - 1) Codeword for the VLC code, the transform coefficient, the motion vector code, CBP, DQUANT, the MBTYPE code, or the REFFRAME code.
 - 2) The total number of decoded MBs in a slice should be equal to the size of the slice.
 - 3) The number of the decoded transform coefficients within a 4x4 block should be smaller than 16.
 - 4) Invalid video data are detected. For example, the prediction error between the predictive block and the current block is an invalid value.

Comparisons (Coding Efficiency)(1)

Video Streaming Applications

352x288
progressive

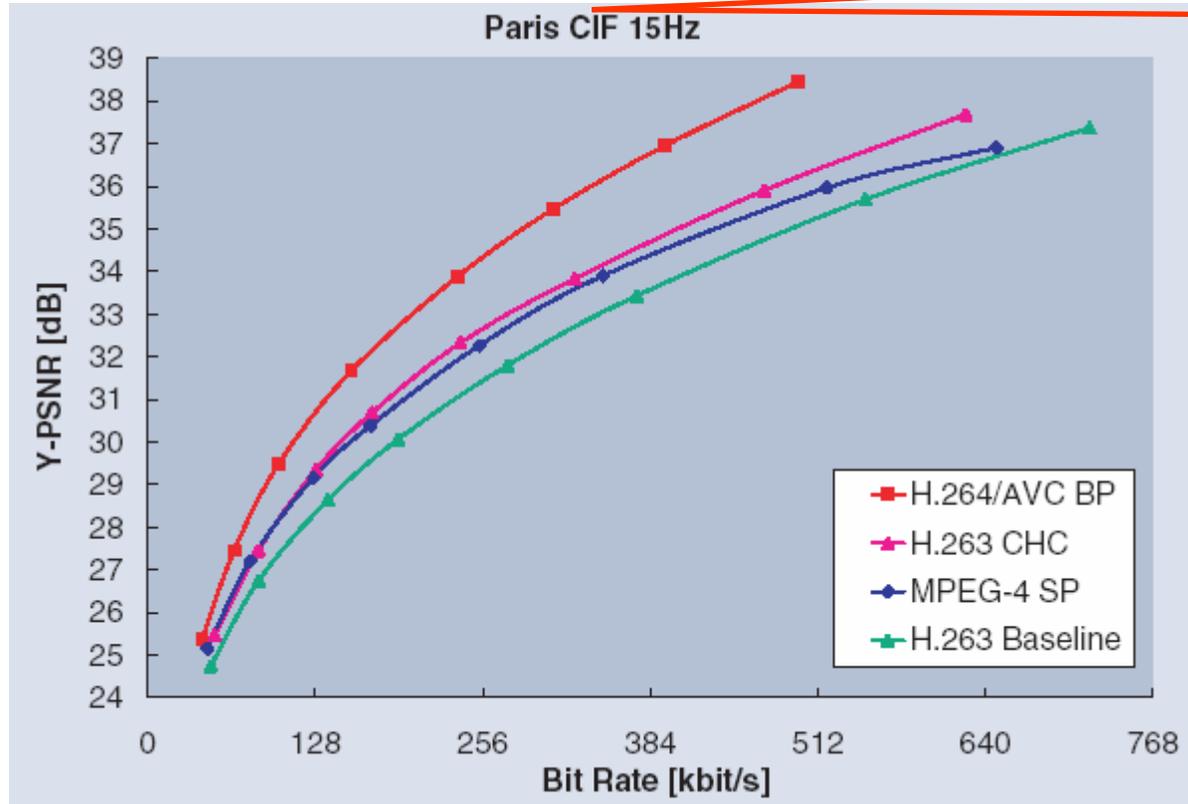


Coder	Average Bit Rate Savings Relative To:		
	MPEG-4 ASP	H.263 HLP	MPEG-2
H.264/AVC MP	37.44%	47.58%	63.57%
MPEG-4 ASP	-	16.65%	42.95%
H.263 HLP	-	-	30.61%

Comparisons (Coding Efficiency)(2)

Video Conferencing Applications

352x288
progressive

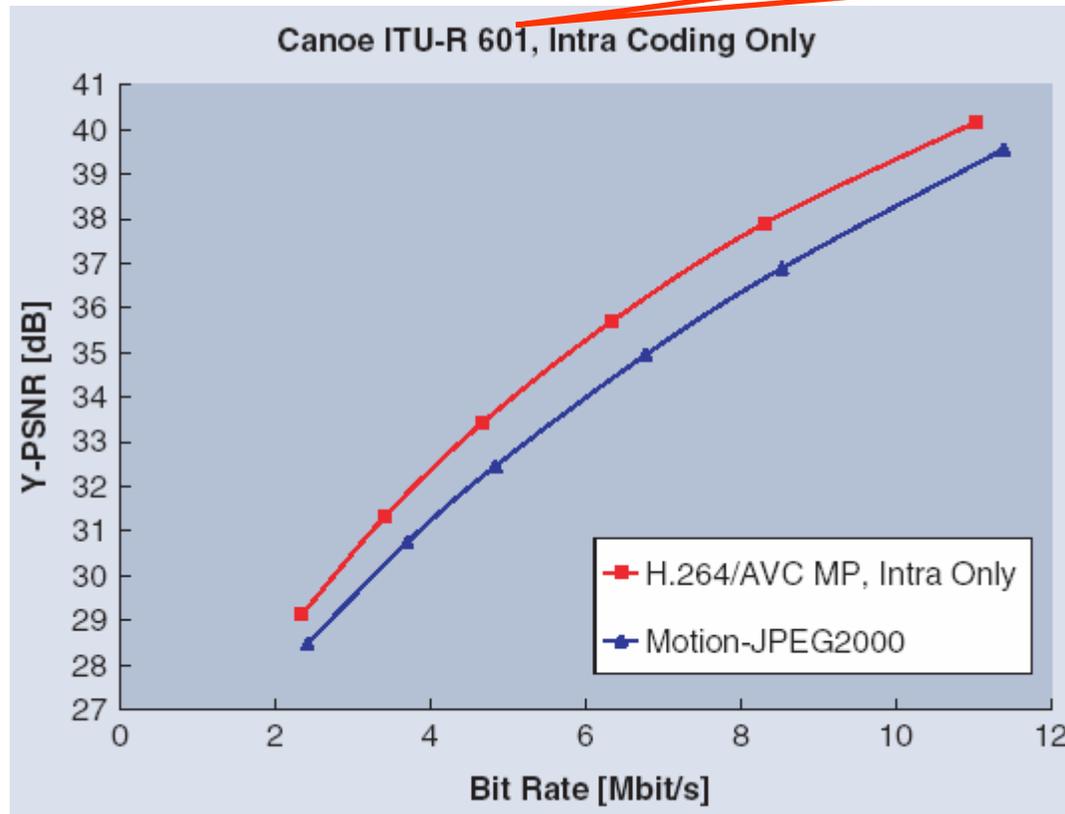


Coder	Average Bit Rate Savings Relative To:		
	H.263 CHC	MPEG-4 SP	H.263 Base
H.264/AVC BP	27.69%	29.37%	40.59%
H.263 CHC	-	2.04%	17.63%
MPEG-4 SP	-	-	15.69%

Comparisons (Coding Efficiency)(3)

Entertainment Applications

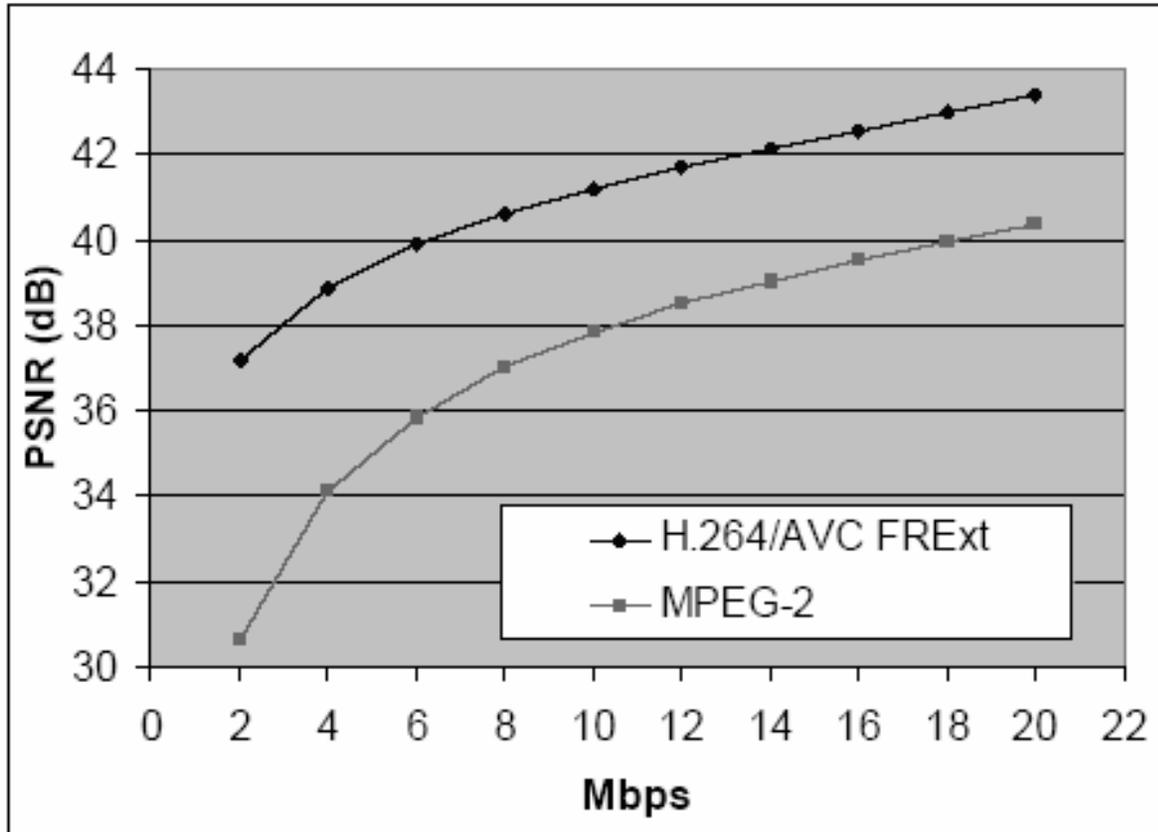
720x576
Interlaced



Comparisons (Coding Efficiency)(4)

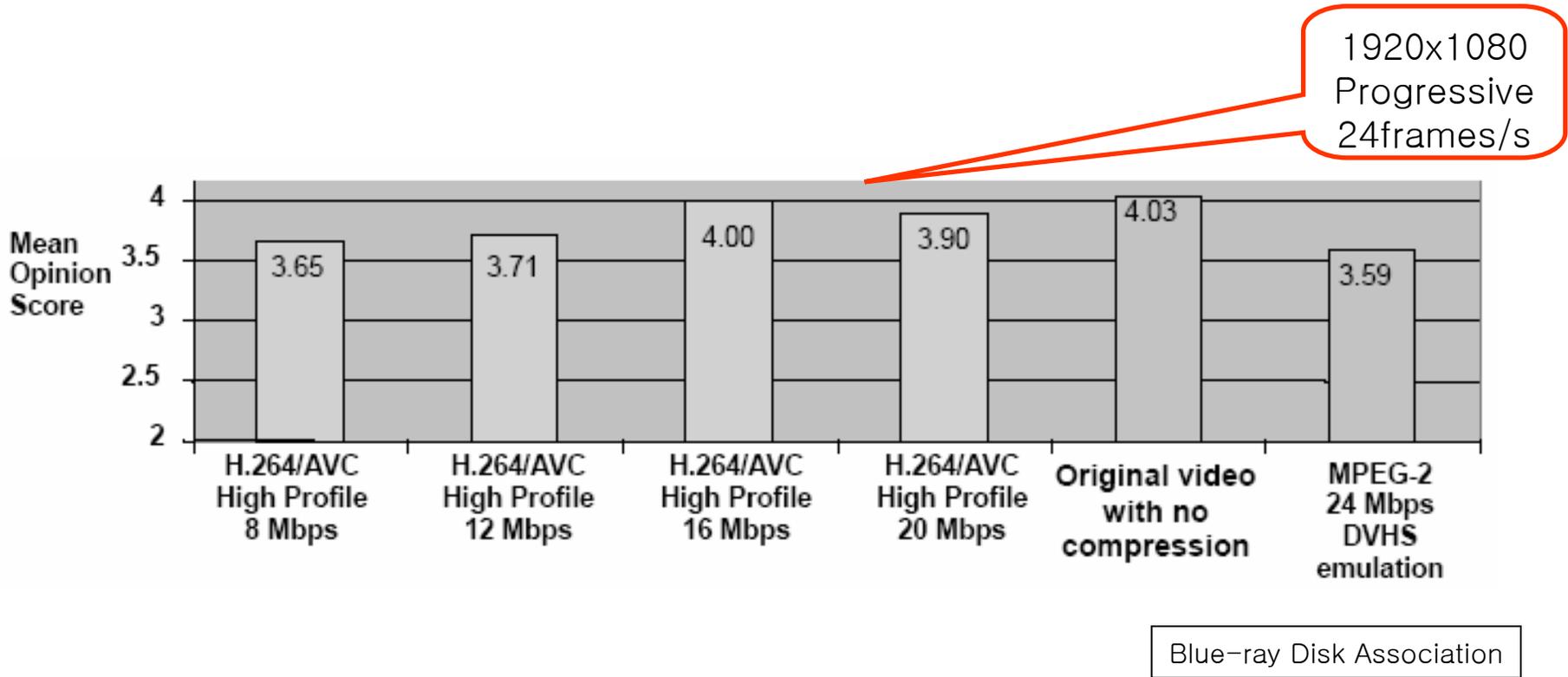
Digital TV Applications

1280x720
progressive



G.J.Sullivan, et al., SPIE August 2004

Performance of FExt High Profile



Comparisons (Hardware Complexity)(1)

Encoder Complexity

기능	복잡도
Variable Block Size	<ul style="list-style-type: none"> • 메모리 액세스 빈도를 선형적으로 증가시킴 • 이 기능을 사용하면 화질을 유지하면서 비트율을 약 4-20% 감축 • 추가적인 모드마다 복잡도가 2.5% 수준으로 선형증가하나 해당 압축 이득은 포화
Hadamard Transform	<ul style="list-style-type: none"> • 액세스 주기를 약 20% 증가 • 화질 대 비트율의 변화는 작음
RD-Lagrangian OPTimization	<ul style="list-style-type: none"> • 데이터 전송증가 120% • PSNR 향상 0.35dB 까지 • 비트율 9% 절략
B-frames	<ul style="list-style-type: none"> • 액세스 주기 -16에서 +12% • 비트율 10%까지 감축
CABAC	<ul style="list-style-type: none"> • 액세스 주기 25-30% 증가 (단일 가역 VLC 데이블 채택시에 비해) • 비트율 16%까지 감축
Displacement vector resolution	<ul style="list-style-type: none"> • ¼ pel position을 안쓰고 ½ pel 만 쓰는 경우 액세스 주기와 처리시간 약 10% 감축 • ¼ pel MB을 사용하는 경우 아주 낮은 비트율을 제외하고 약 30%까지 부호화 효율 증가
Search Range	<ul style="list-style-type: none"> • 참조프레임 수와 탐색 영역을 모두 증가시키면 액세스주기가 약 60배까지 증가 • 반면 PSNR과 비트율 성능에는 영향이 미약함
Multiple Reference Frames	<ul style="list-style-type: none"> • 프레임의 수가 증가할수록 약 25%의 복잡도 선형증가 • 중간 이하의 비트율에서는 무시할 정도 (약 2%이하)의 비트율 이득 • 높은 비트율에서는 좀 더 많은 비트율 이득 달성
Deblocking Filter	<ul style="list-style-type: none"> • 부호화 복잡도에 대한 영향은 거의 없음 • 주관적인 화질을 상당히 향상시킴

Comparisons (Hardware Complexity)(2)

Decoder Complexity

기능	복잡도
CABAC	<ul style="list-style-type: none">• 액세스 주기 12%까지 증가 (단일 가역 VLC에 비해)• 비트율이 높을수록 증가율이 높아짐
RD-Lagrangian Optimization	<ul style="list-style-type: none">• 중간 및 낮은 비트율에서 평균 약 5% 복잡도 증가• 높은 비트율에서는 영향력 미약
B-frames	<ul style="list-style-type: none">• 11-29% 복잡도 증가• 첫번째 B-frame의 경우 낮은 비트율에서 50%의 추가 복잡도 증가, 중간에서 고비트율에서 20-35%• 두번째 B-frame 이후의 추가 시간 증가는 몇 %에 불과
Hadamard Transform	<ul style="list-style-type: none">• 부호화 시간은 약 5% 증가• 메모리 액세스는 무시할 정도
Deblocking Filter	<ul style="list-style-type: none">• 필수적 채택 사양• 복호기 액세스 주기 6%까지 증가
Displacement vector resolution	<ul style="list-style-type: none">• Encoder가 ½ pel position의 MV만 보내면 부호화 시간은 약 15% 감소함

- 상대적으로 H.264/AVC Encoder의 복잡도는 MPEG-4 Part 2 (SP)의 약 10배 이상 Decoder는 약 2배 이상 증가
- H.264/AVC Profile의 Tool과 파라미터 값의 선택에 따라 압축성능을 유지하면서 하드웨어 복잡도를 Encoder 6.5배 Decoder 1.5배 까지 줄일 수 있음

인코더의 입력파라미터와 성능지수 관계

파라미터	사용용도	주의
Hadamard on/off	왜곡 측정시 Hadamard 변환의 사용 여부	약 20%의 복잡도 증가
Rd-optimization on/off	VLC 부호화를 수행하고 실제 발생될 비트 량에 따른 rd-cost 계산	아주 큰 복잡도를 갖고 있기 때문에 실제 하드웨어에 적용한 예는 없음. 적용시 최대 9%의 비트율 향상
Reference Frame	Reference frame의 개수는 압축의 효율을 높여주나 회로의 복잡도가 높아짐	한 프레임 증가에 대해 25%의 복잡도 증가 발생 비트율은 높은 비트율에서 14%의 향상을 가져옴
FMO/ASO	채널 상에 에러가 발생할 경우 에러 감춤을 위한 방법	비트율 향상 없음 디코딩 효율을 위한 파라미터
Variable Block Size	16x16, 16x8, 8x8, 8x4, 4x8, 4x4 단위 예측 보상	각 모드 마다 2.5%의 복잡도 증가 비트율 향상 4-20% 향상
Motion vector resolution	¼ pel 움직임 벡터 사용 여부	½ pel 움직임 벡터 경우보다 10% 복잡도 증가 저 전송율시 30% 정도의 비트율 상승
움직임 벡터 탐색 영역	움직임 벡터의 크기 결정	

H.264 Competitors

Name	Group	Standard	Strong Points	Notes
H.264	MPEG+ITU	HD-DVD DMB (main)	<ul style="list-style-type: none">• High performance (on low-bit rate)• MPEG standard	<ul style="list-style-type: none">• High Complexity• MPEG-LA licence
WMV9 (VC1)	Microsoft	HD-DVD (main)	<ul style="list-style-type: none">• High performance (on low-bit rate)• Moderate complexity• Compatible to PC	<ul style="list-style-type: none">• MS solution• Patent issue• CE vs IT issue
AVS	China	EVD2	<ul style="list-style-type: none">• High performance (similar to H.264)• moderate complexity	<ul style="list-style-type: none">• Very large market• China's Standard• Patent issue

WMV9 & AVS: Low complexity + Comparable performance to H.264

맺음말

- H.264/AVC의 부호화 툴은 기존의 MPEG-4와 MPEG-2와 비교해서 약 50%의 비트 절약이 가능
- 그러나 H.264/AVC의 Decoder 복잡도는 MPEG-4 Visual SP에 비해 2배 Encoder는 약 10배 더 복잡
- H.264/AVC main profile decoder는 MPEG-2에 비해 약 4배 더 복잡, Encoder의 복잡도는 rate-constrained encoder control과 움직임 검출 알고리즘에 따라 상당히 다를 수 있음
- VLSI와 CPU의 성능향상과 메모리의 가격하락이 H.264/AVC의 구현을 가능하게 함

Reference S/W

<http://bs.hhi.de/~suehring/tml/download>