

信号理論

- No.13 画像/動画信号処理への応用 -

渡辺 裕

Signal Theory

- No.13 Application to Image/Video Signal Processing -

Hiroshi Watanabe

画像符号化標準

- グラフィクス(限定色)
 - GIF
 - PNG
- 2値画像
 - JBIG
- 自然画像
 - JPEG
 - JPEG2000

Image Coding Standard

- Graphics (Limited Color)
 - GIF
 - PNG
- Binary Image
 - JBIG
- Natural Image
 - JPEG
 - JPEG2000

GIF

- グラフィクス交換フォーマット
 - コンピューサーブで使われた
 - 256色以下の可逆符号化
 - LZW圧縮技術
 - GIF87a, GIF89a
- LZW
 - Lempel-Ziv-Welch
 - 辞書型のユニバーサル符号化
 - ユニシス特許問題 1984-2004

GIF

- Graphics Interchange Format
 - CompuServe
 - Lossless coding for 256 colors or less
 - LZW compression
 - GIF87a, GIF89a
- LZW
 - Lempel-Ziv-Welch
 - Dictionary type compression
 - Patented 1984-2004

PNG

- ポータブルネットワークグラフィクス
 - ISO/IEC JTC1/SC24で作成された標準
 - 8-24ビットカラーの可逆符号化
 - Deflate圧縮技術を使用
 - ISO/IEC 15948:2003
- Deflate
 - LZ77 + 動的ハフマン符号
 - 位置と長さ情報が分類されて符号化
 - IETF RFC1951

PNG

- Portable Network Graphics
 - ISO/IEC JTC1/SC24
 - Lossless for 8-24 bit colors
 - Deflate compression technology
 - ISO/IEC 15948:2003
- Deflate
 - LZ77 + Dynamic Huffman
 - Location and length are categorized
 - IETF RFC1951

JBIG

- Joint Bi-level Image Experts Group
 - ISO/IEC JTC1/SC29/WG1 で作成された標準
 - 二値画像が対象(白黒画像)
 - マルコフモデル + QM符号化器
 - ISO/IEC 11544
- QM coder
 - マルコフモデル
 - 算術符号化
 - LPS: 2^{-Q}

JBIG

- Joint Bi-level Image Experts Group
 - ISO/IEC JTC1/SC29/WG1
 - Bi-level image (black/white)
 - Markov model + QM coder
 - ISO/IEC 11544
- QM coder
 - Markov model
 - Arithmetic coding
 - LPS: 2^{-Q}

JPEG

- Joint Photographic Experts Group
 - ISO/IEC JTC1/SC29/WG1 で作成
 - 静止画像(RGB各8ビットの24ビットカラー画像)
 - 非可逆符号化/可逆符号化
 - 離散コサイン変換(DCT) + ハフマン符号化/QM符号化
 - ISO/IEC 10918: 1994 | ITU-T Rec. T-81
- 離散コサイン変換
 - 8x8 DCT, 線形量子化, 重み行列
 - ジグザグスキャン, 2次元ハフマン符号化
 - 算術符号(QM coder)はオプション

JPEG

- Joint Photographic Experts Group
 - ISO/IEC JTC1/SC29/WG1
 - Still Image
 - Lossy coding/lossless coding
 - DCT + Huffman/QM coder
 - ISO/IEC 10918: 1994 | ITU-T Rec. T-81

- DCT
 - 8x8 DCT, linear quantization, weight matrix
 - Zigzag scan, 2D-Huffman coding
 - Option: QM coder

JPEG 2000

- Joint Photographic Experts Group
 - ISO/IEC JTC1/SC29/WG1 で作成
 - 静止画像(RGB各8ビットの24ビットカラー画像)
 - 非可逆から可逆まで連続的に導出可能な符号化
 - ウェーブレット + EBCOT(算術符号化)
 - ISO/IEC 15444: 2000

- ウェーブレット
 - リフティング, 線形量子化
 - EBCOT(Embedded Block Coding with Optimized Truncation), 複数コンテキストの算術符号化(MQ-coder)

JPEG 2000

- Joint Photographic Experts Group
 - ISO/IEC JTC1/SC29/WG1
 - Still Image(RGB各8ビットの24ビットカラー画像)
 - From lossy to lossless
 - Wavelet + EBCOT(Arithmetic coding)
 - ISO/IEC 15444: 2000

- Wavelet
 - Lifting scheme, linear quantizer
 - EBCOT(Embedded Block Coding with Optimized Truncation), multi-context arithmetic coding(MQ-coder)

ビデオ符号化標準 (1)

- ISO/IEC
 - MPEG-1 (1992) Video CD, Web video
 - MPEG-2 (1994) DVD, HDTV
 - MPEG-4 (1999) (Mobile)
 - MPEG-4 AVC/H.264 (2005) Mobile TV, Blu-ray, ...
- ITU-T
 - H.261(1990), H.262(=MPEG2), H.263(1995), H.264(2005)

JPEG: Joint Photographic Experts Group

MPEG: Moving Picture Experts Group

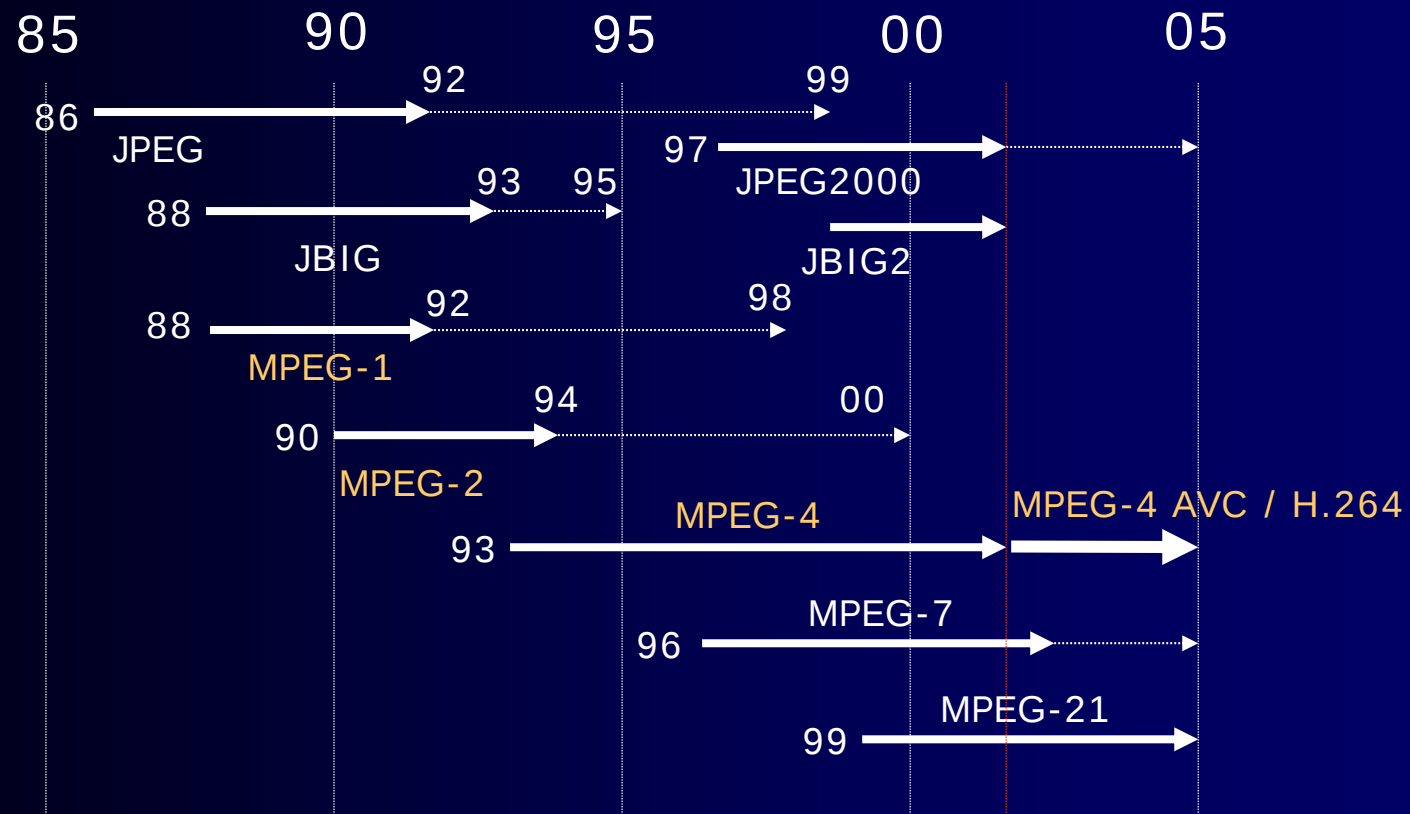
Video Coding Standard (1)

- ISO/IEC
 - MPEG-1 (1992) Video CD, Web video
 - MPEG-2 (1994) DVD, HDTV
 - MPEG-4 (1999) (Mobile)
 - MPEG-4 AVC/H.264 (2005) Mobile TV, Blu-ray, ...
- ITU-T
 - H.261(1990), H.262(=MPEG2), H.263(1995), H.264(2005)

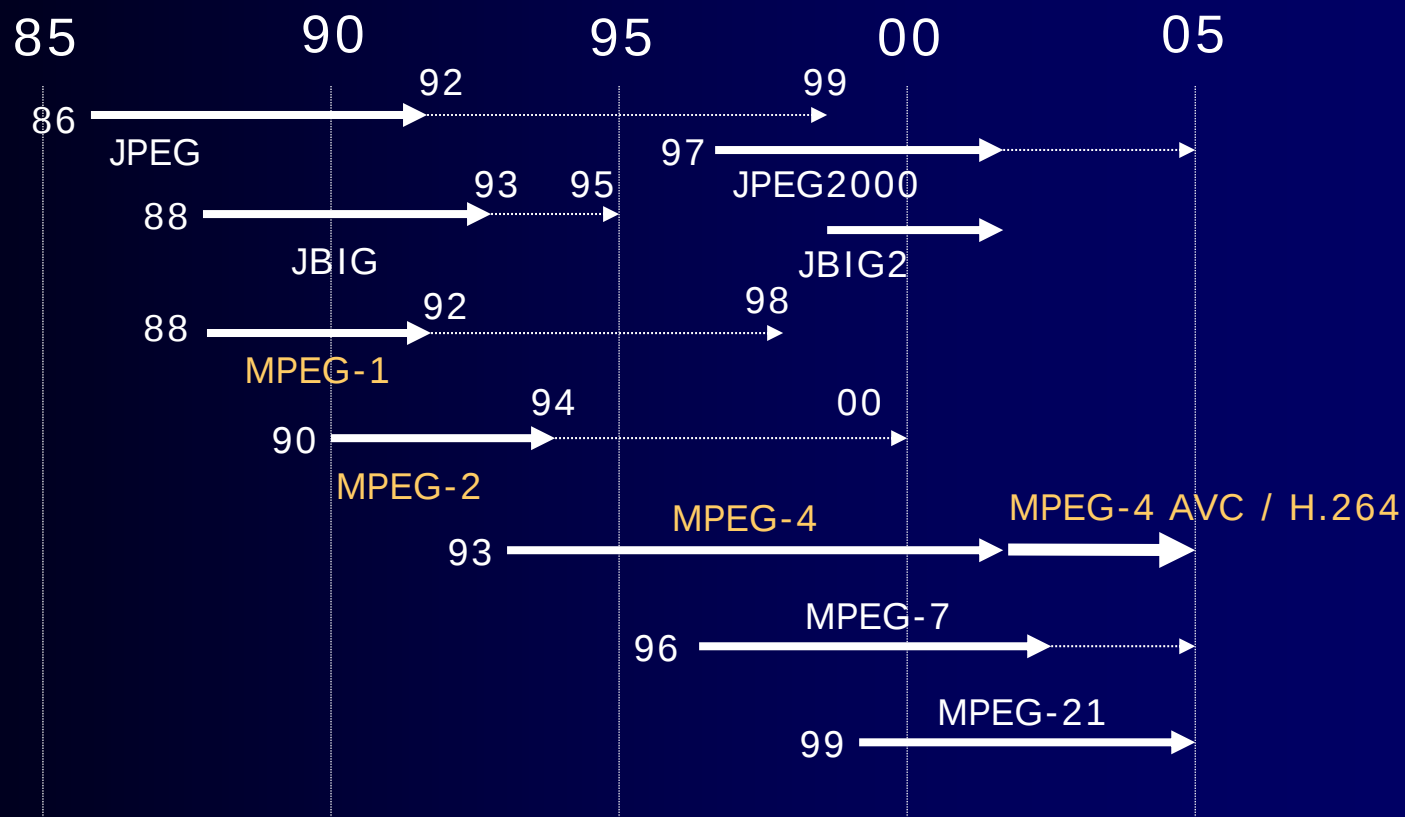
JPEG: Joint Photographic Experts Group

MPEG: Moving Picture Experts Group

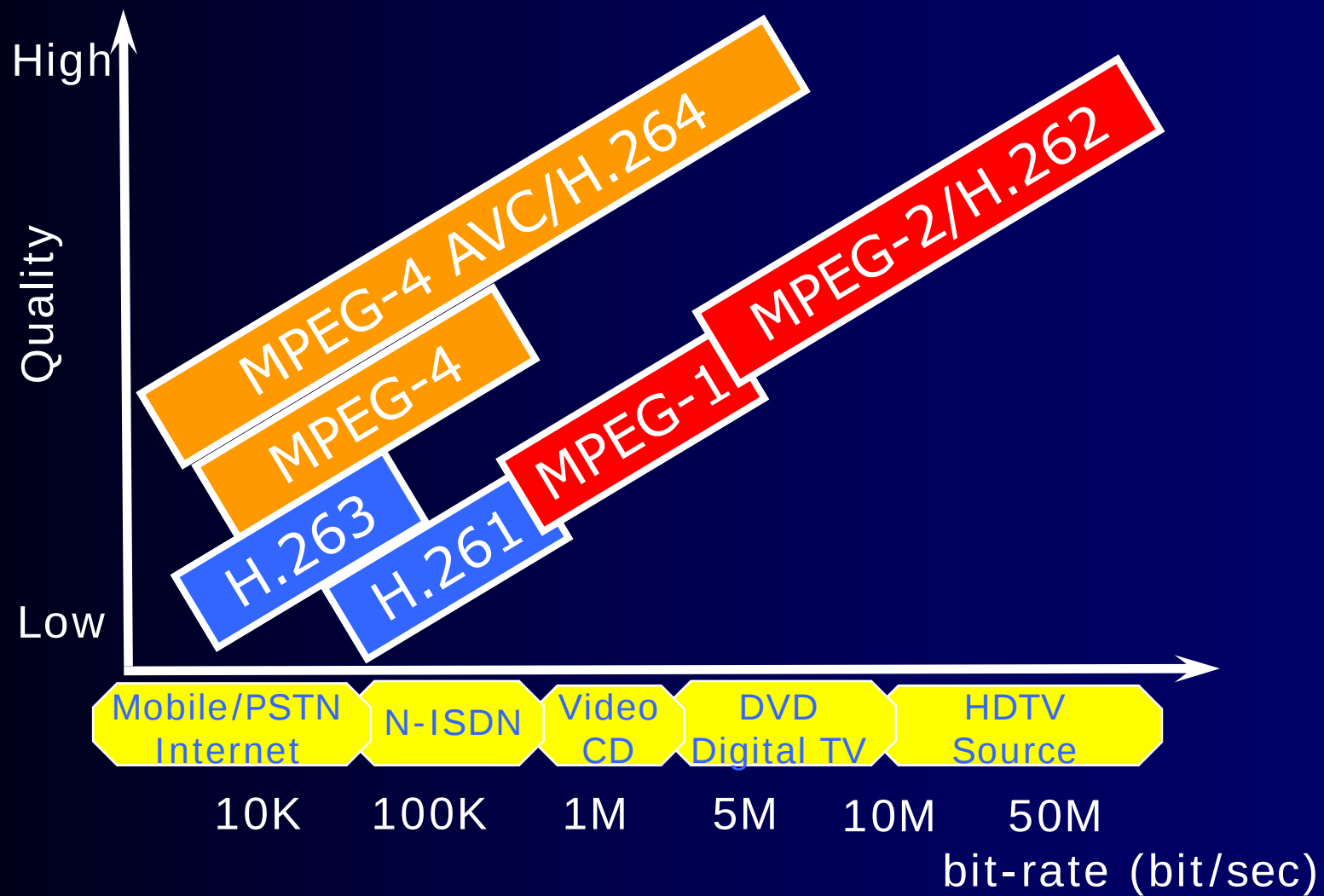
ビデオ符号化標準 (2)



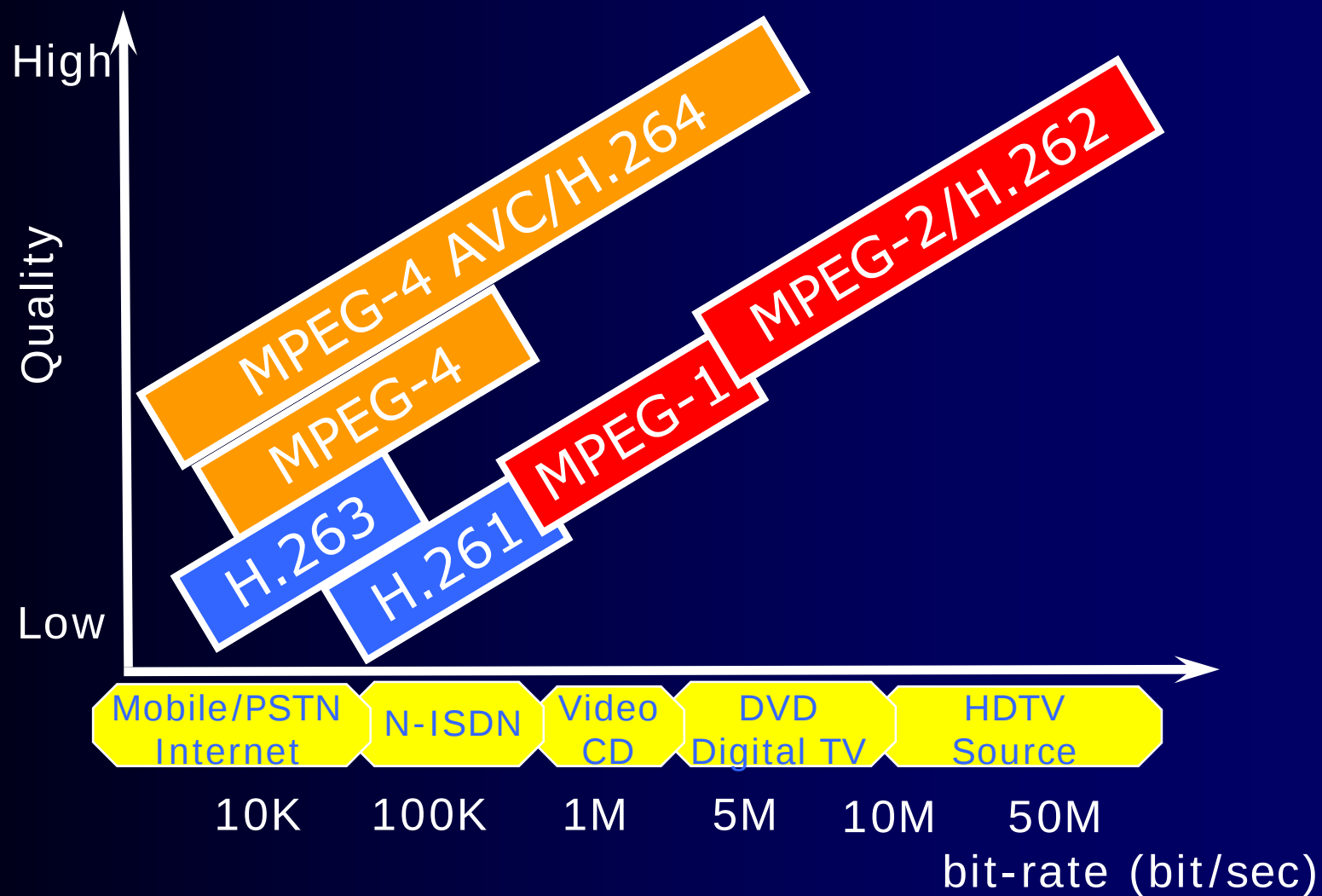
Video Coding Standard (2)



ビデオ符号化標準 (3)



Video Coding Standard (3)



H.261 (ITU-T, 1990)

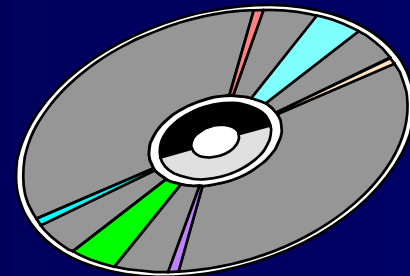
- ISDN用のテレビ電話・テレビ会議用ビデオ符号化
- ビットレート: $p \times 64\text{kbps}$ ($p = 1-30$)
- マクロブロック単位の動き補償と離散コサイン変換(DCT)
- 標準画像サイズはCIF, QCIF
- 順方向動き補償
- ループフィルタ

H.261 (ITU-T, 1990)

- Coding scheme for videophone, teleconferencing on digital network (ISDN)
- Bitrate: $p \times 64\text{kbps}$ ($p = 1-30$)
- DCT based (macroblock)
- Support of CIF, QCIF
- Forward MC
- Loop filter

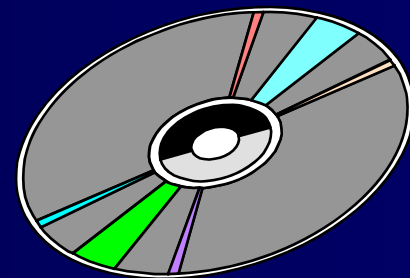
MPEG1 (ISO/IEC 11172, 1992)

- デジタル蓄積メディアのためのビデオ符号化
- 典型的なビットレート: 1.5Mbps (CD-ROM)
- 主な対象画像サイズSIF (NTSC, PAL)
- ランダムアクセス機能(GOP単位)
- 双方向動き補償, 半画素精度動き補償, 離散コサイン変換



MPEG1 (ISO/IEC 11172, 1992)

- Coding Scheme for Digital Storage Media
- Typical Bit rate: 1.5Mbps (CD-ROM)
- Support of SIF (NTSC, PAL)
- Random accessibility (GOP)
- Bi-directional MC, Half-pel MC, DCT



MPEG2 (ISO/IEC 13818, 1994)

- はん用符号化(蓄積メディア, 放送, 通信アプリケーション)
- ビットレート: 3-6Mbps (DVD, SDTV) 6-10Mbps(HDTV)
- 対象画像: SDTV, HDTVサイズ
- 蓄積メディア用機能, 放送用機能
- 双方向動き補償, 半画素精度動き補償, 離散コサイン変換
- インタレース対応予測方式



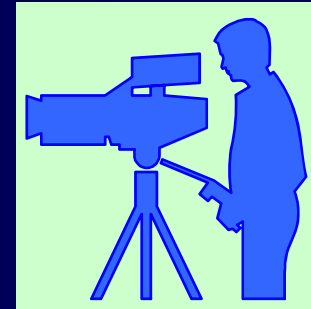
MPEG2 (ISO/IEC 13818, 1994)

- Generic Coding (Storage Media, Broadcasting, Communication)
- Bitrate: 3-6Mbps (DVD, SDTV) 6-10Mbps(HDTV)
- Video Size: SDTV, HDTV
- Functionality for Storage Media and Broadcasting
- Bi-directional MC, Half-pel MC, DCT
- Adaptive Prediction for Interlace Signal



MPEG-2ビデオへの要求条件

- 高品質
- マルチフォーマット (4:2:2, 4:4:4, インタレース)
- 動作ビットレートは固定(CBR)と可変(VBR)
- ランダムアクセス機能, 早送り/巻き戻し制御
- スケーラビリティ(空間, 時間, 画品質)
- MPEG1との互換性
- 編集機能
- 誤り耐性



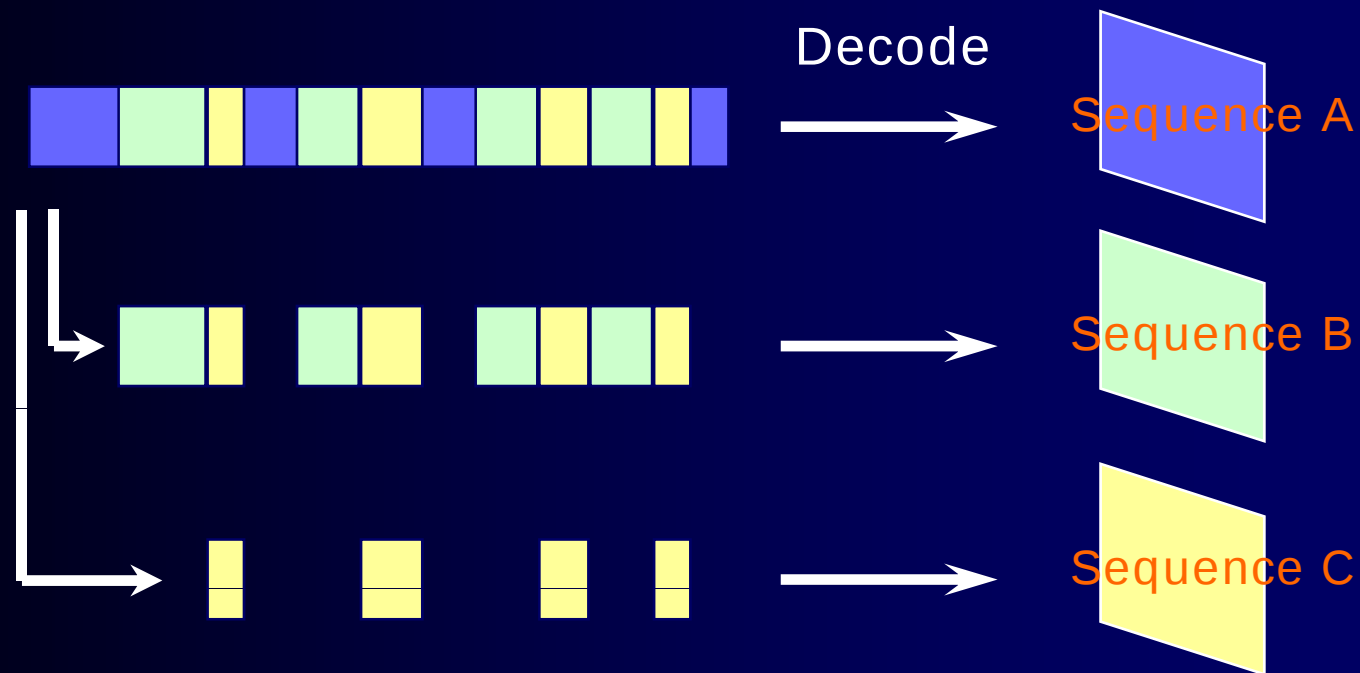
Requirements for MPEG2

- High quality
- Multi-format (4:2:2, 4:4:4, interlace)
- Constant bitrate (CBR) and Variable bitrate (VBR)
- Random access, FF/FR
- Scalability (Spatial, Temporal, SNR)
- Backward compatibility with MPEG1
- Editing functions
- Error tolerant



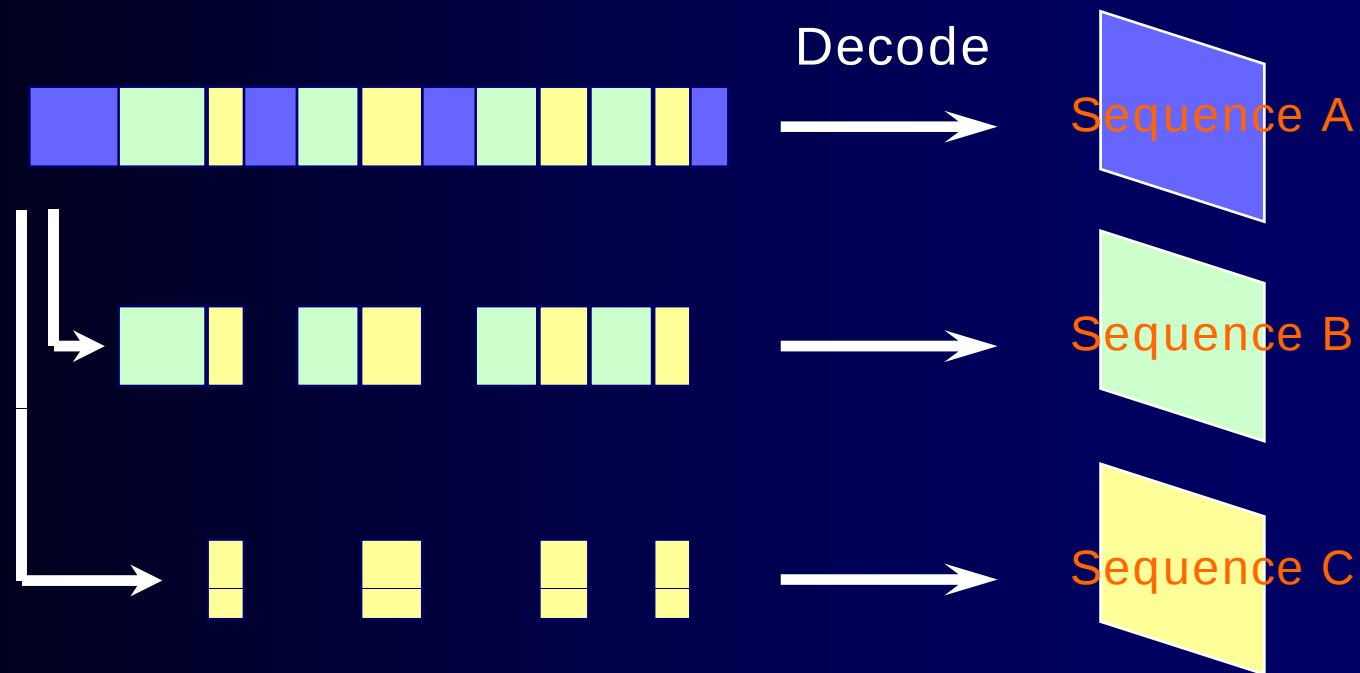
MPEG-2スケーラビリティ (1)

- MPEG-2ビットストリームの一部を復号して画像再生が可能



Scalability of MPEG-2 (1)

- Pictures can be decoded and reconstructed from a part of the MPEG2 bitstream



MPEG-2スケーラビリティ (2)

- 空間スケーラビリティ
 - “低解像度レイヤ” と “高解像度レイヤ”
- 画品質スケーラビリティ
 - “ぼけたレイヤ” と “高品質レイヤ”
- 時間スケーラビリティ
 - “ぎくしゃくしたレイヤ” と “滑らかなレイヤ”

Scalability of MPEG-2 (2)

- Spatial Scalability
 - “low-res layer” and “high-res layer”
- SNR Scalability
 - “blurred layer” and “high-definition layer”
- Temporal Scalability
 - “flickery layer” and “smooth layer”

MPEG-2のプロファイルとレベル

		Profile						
		Simple	Main	SNR	Spatial	High	422	Multi-view
Level	High 1920x1080		MP@HL			HP@HL	422P@HL	MVP@HL
	High1440 1440x1080		MP@H1440		SPP@H1440	HP@H1440		MVP@H1440
	Main 720x576	SP@ML	MP@ML	SNRP@ML		HP@ML	422P@ML	MVP@ML
	Low 352x288			SNRP@LL				MVP@LL

Profiles and Levels in MPEG-2

		Profile						
		Simple	Main	SNR	Spatial	High	422	Multi-view
Level	High 1920x1080		MP@HL			HP@HL	422P@HL	MVP@HL
	High1440 1440x1080		MP@H1440		SPP@H1440	HP@H1440		MVP@H1440
	Main 720x576	SP@ML	MP@ML	SNRP@ML		HP@ML	422P@ML	MVP@ML
	Low 352x288			SNRP@LL				MVP@LL

メインプロファイル@メインレベル仕様

- 処理画像 : 720 × 576,480 pels × 30 Hz
- ビットレート : 最大15 Mbits/s
- 形式 : 4:2:0
- ピクチャタイプ : I, P, B-ピクチャ
- 符号化単位 : フレーム構造/フィールド構造
- バッファサイズ : 最大 1.8 Mbits
- 可変ビットレート : 含まれる

MP@ML Specification

- Processing : 720 × 576,480 pels × 30 Hz
- Bitrate : Max. 15 Mbits/s
- Format : 4:2:0
- Picture type : I, P, B-pictures
- Coding Unit : Frame/Field Structure
- Buffer Size : Max. 1.8 Mbits
- VBR : Included

MPEG-4 (ISO/IEC 14496, 1998)

- オブジェクトベースのビデオ符号化
- ビットレート: 任意のレート
- 対象画像: 任意のプロGRESSIVE画像
- Webブラウザの概念に基づいたインタラクティブなオーディオビジュアルオブジェクトの圧縮と操作

MPEG-4 (ISO/IEC 14496, 1998)

- Object based video coding
- Bitrate: Any number
- Video: Progressive image
- Interactive audio-visual coding and manipulation based on the concept of Web browser operation

MPEG-4 ビデオ

- 自然画像符号化の特徴
 - ビデオオブジェクトプレーン(VOP)構造
 - 形状符号化, 半透明処理
 - エラー耐性
 - 可逆VLC, データ分割, スライス同期
 - 高圧縮
 - イントラAC/DC予測, ブロック歪除去フィルタ
 - オブジェクトスケーラビリティ
 - 空間, 時間, レイヤ構造のシンタクス

MPEG-4 Video

- Characteristics of natural video coding
 - Structure of video object plane (VOP)
 - Shape coding, alpha map processing
 - Error resilience
 - Reversible VLC, data partitioning, Slice header
 - High efficiency
 - Intra AC/DC prediction, De-blocking filter
 - Object-based scalability
 - Spatial-temporal layer structure syntax

ビデオオブジェクト符号化

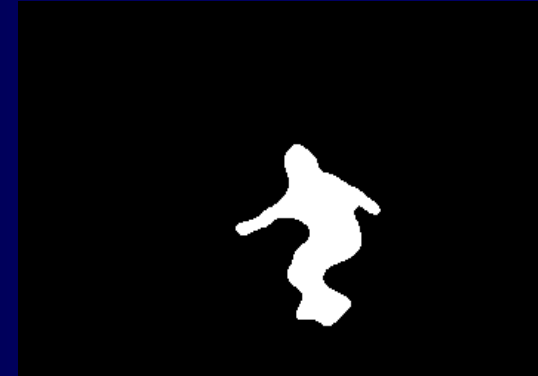
- Video Object Plane (VOP)
 - VOP0 と VOP1 を独立に符号化
 - 形状符号化



VOP0 (背景)



VOP1 (前景)



形状情報(α マップ)

Object Based Video Coding

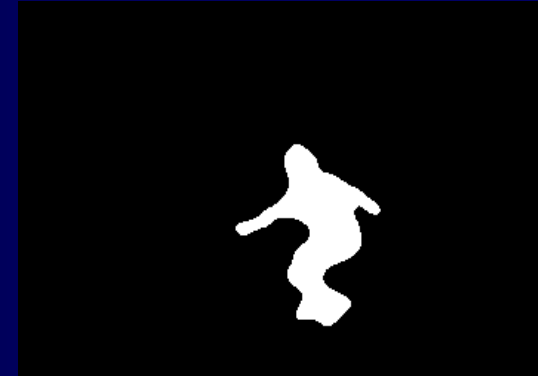
- Video Object Plane (VOP)
 - VOP0 and VOP1 are independently coded
 - Shape Coding



VOP0 (background)



VOP1 (object)



Shape data(α map)

VOP生成方法

- クロマキー分離
 - ブルーバック撮影, 後処理による物体分離
- アニメーション
 - 複数枚のセル画の重ね合わせ
- CG生成
- 自然画像からの切り出し
 - 手作業
 - 自動切り出し
(輪郭抽出, 境界追跡, テクスチャ解析, ...)

VOP Creation

- Separation by Chroma-key
 - Blue back video capture, separation by post processing
- Animation
 - Layered cell pictures
- CG generation
- Separation from natural images
 - By hand
 - Automatic separation
(Edge detection, contour tracking, texture analysis)

自動VOP検出方法の例

- エッジ検出と領域拡大/縮小の組み合わせ



Example of Automatic VOP Detection

- Combination with edge detection and region growing/shrinking method



VOP符号化の構成

■ Core Coder と Generic Coder

MPEG-4 VLBV Core Coder



Video
Object
Plane

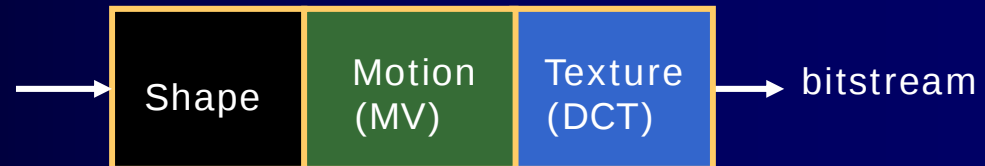


(Similar to H.263/MPEG-1)

Generic MPEG-4 Coder



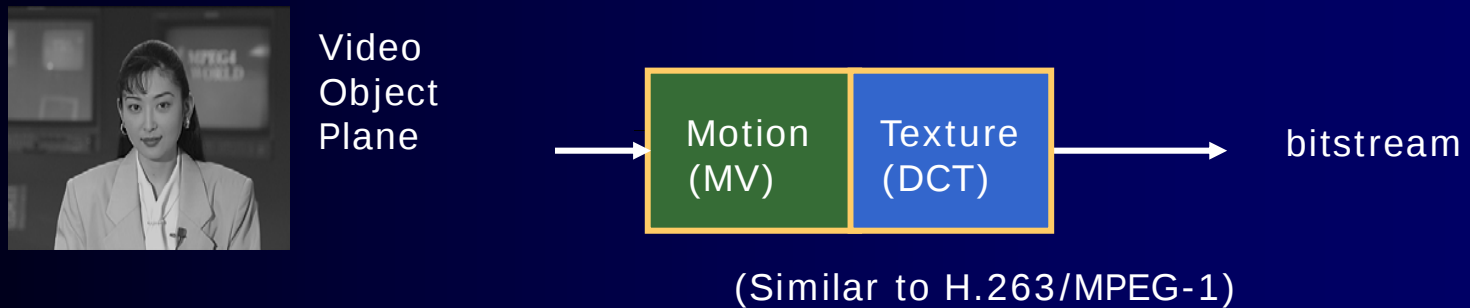
Video
Object
Plane



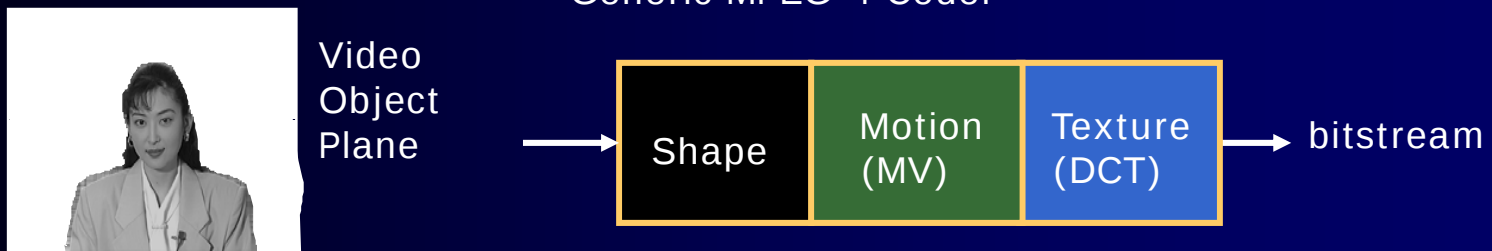
Structure of VOP Coding

■ Core Coder and Generic Coder

MPEG-4 VLBV Core Coder



Generic MPEG-4 Coder



VOPの信号

入力信号



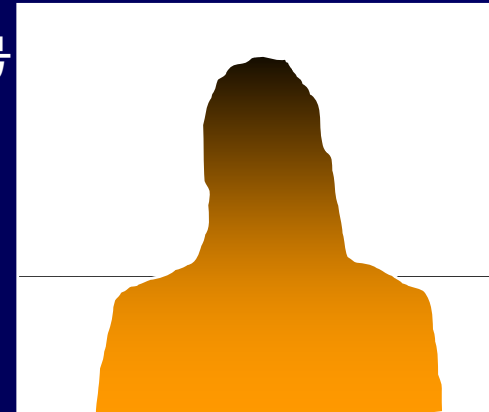
形状信号



画素信号



透過度信号

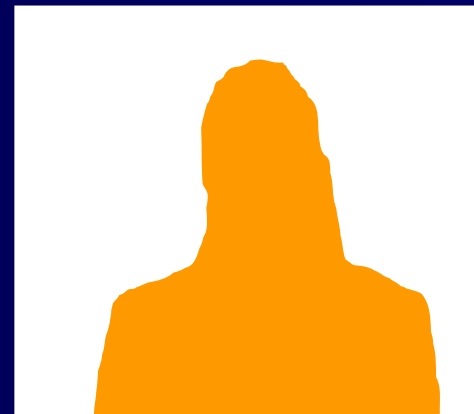


Signals in VOP Coding

Input



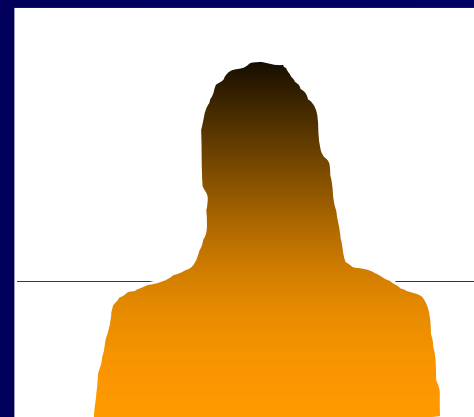
Shape



Pixel
value

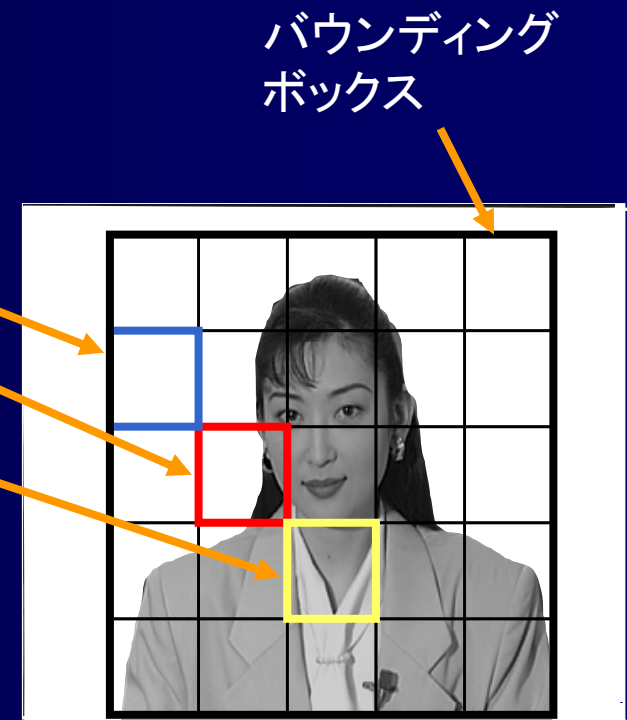


α map



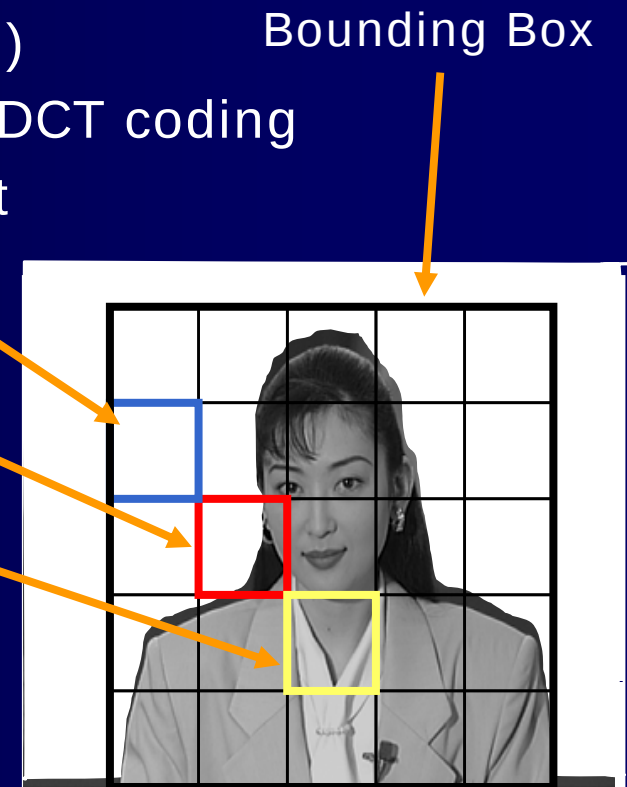
VOP符号化 -画素信号-

- 画素値情報
- 画素値そのもの(輝度・色差)
 - 動き補償DCT符号化
 - 物体外は符号化しない
 - 境界はパディング処理後に符号化
 - 物体内部のみ符号化



VOP Coding -Pixel Data-

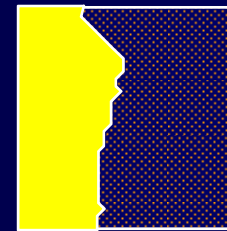
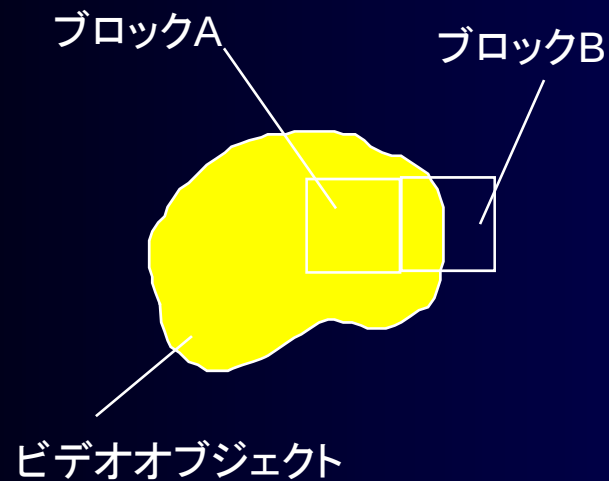
- Information of pixel location
- Pixel value (luminance, chrominance)
 - Motion compensated intrerframe DCT coding
 - No coding for outside of object
 - boundary block is coded after padding
 - Normal coding for inside of object



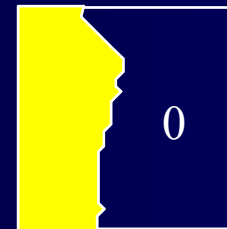
52

Padding DCT

- オブジェクト内 (ブロックA): 通常の8x8 DCT
- オブジェクト境界 (ブロックB): Padding後に 8x8 DCT
- オブジェクト外: スキップ



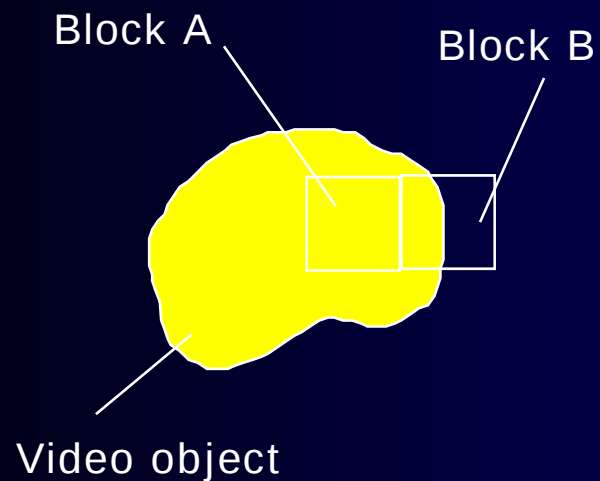
フレーム内ブロック:
ローパス外挿フィルタ



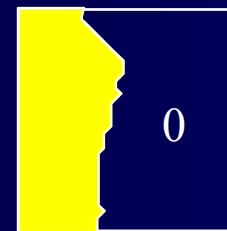
フレーム間ブロック:
零値外挿

Padding DCT

- Inside of object (Block A): Normal 8x8 DCT
- Boundary of object (Block B): 8x8 DCT after padding
- Outside of object: Skip



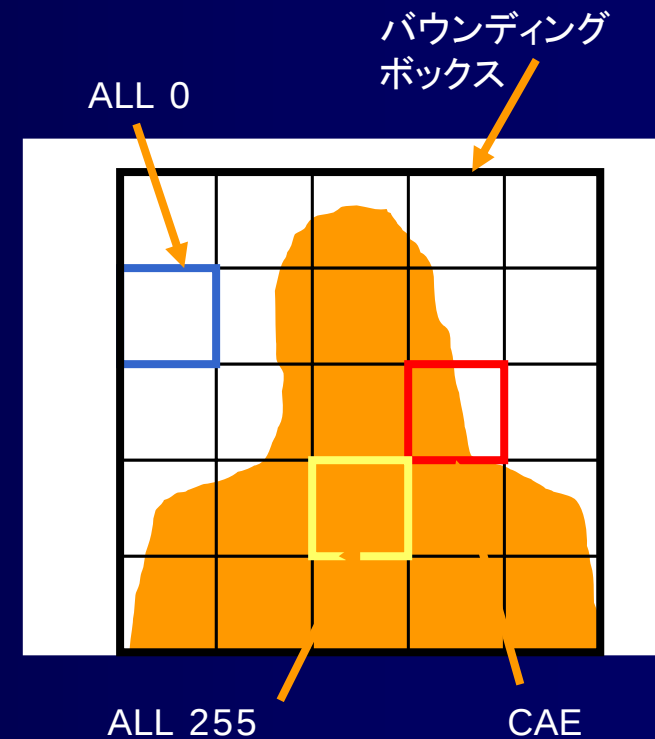
Intraframe block:
extrapolation by
lowpass filtering



Interframe block:
zero insertion

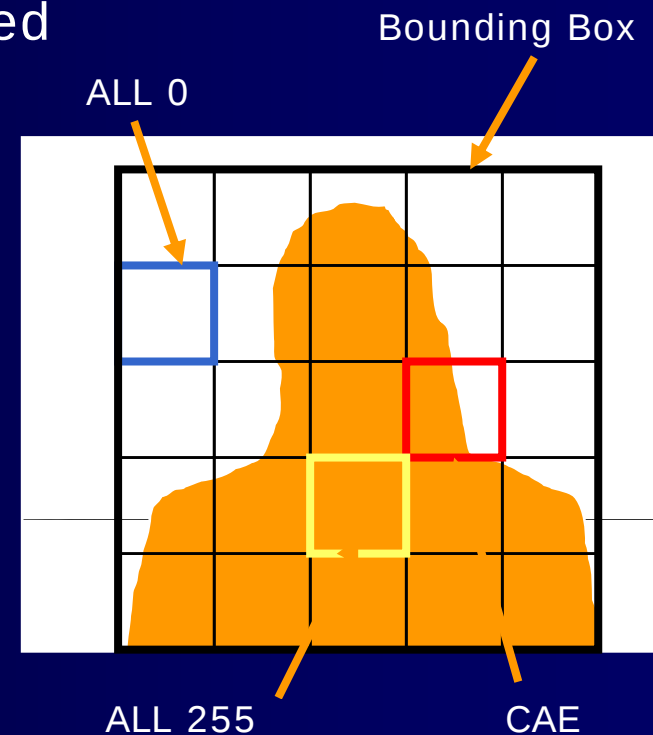
VOP符号化 –形状信号–

- 各画素が物体内部か外部かの情報 (0 or 255の2値)
- 2値形状符号化: CAE (Context-based Arithmetic Coding)
- 形状をマクロブロック単位に符号化
 - ALL “0”: VOP外
 - ALL “255”: VOP内
 - 境界ブロック:
 - フレーム内CAE
 - フレーム間CAE
 - (動き補償付き)



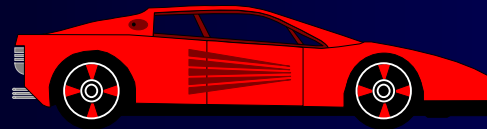
VOP coding –Shape Data-

- Flag to show inside/outside of object at pixel (0 or 255)
- Binary shape coding: CAE (Context-based Arithmetic Coding)
- Shape is coded macroblock based
 - ALL “0”: outside of VOP
 - ALL “255”: inside of VOP
 - Boundary block:
 - Intraframe CAE
 - Interframe CAE (with MC)



VOP符号化 -透過度信号-

- 各画素が透明か不透明か, どれくらい透明かの情報
- 0~255の8ビット
- 動き補償DCT符号化



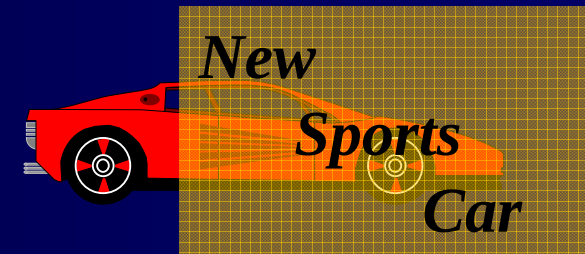
物体A



物体B



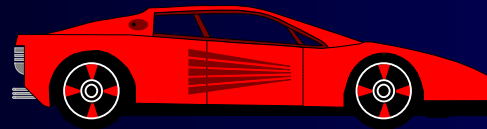
物体Bの透過度0%



物体Bの透過度50%

VOP Coding –Alpha Map-

- Information for level of transparency at each pixel
- 8 bit (0~255)
- MC-DCT coded



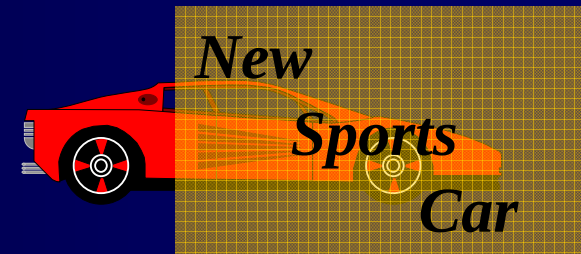
Object A



Object B



0% transparency of object B



50% transparency for object B

スプライト符号化

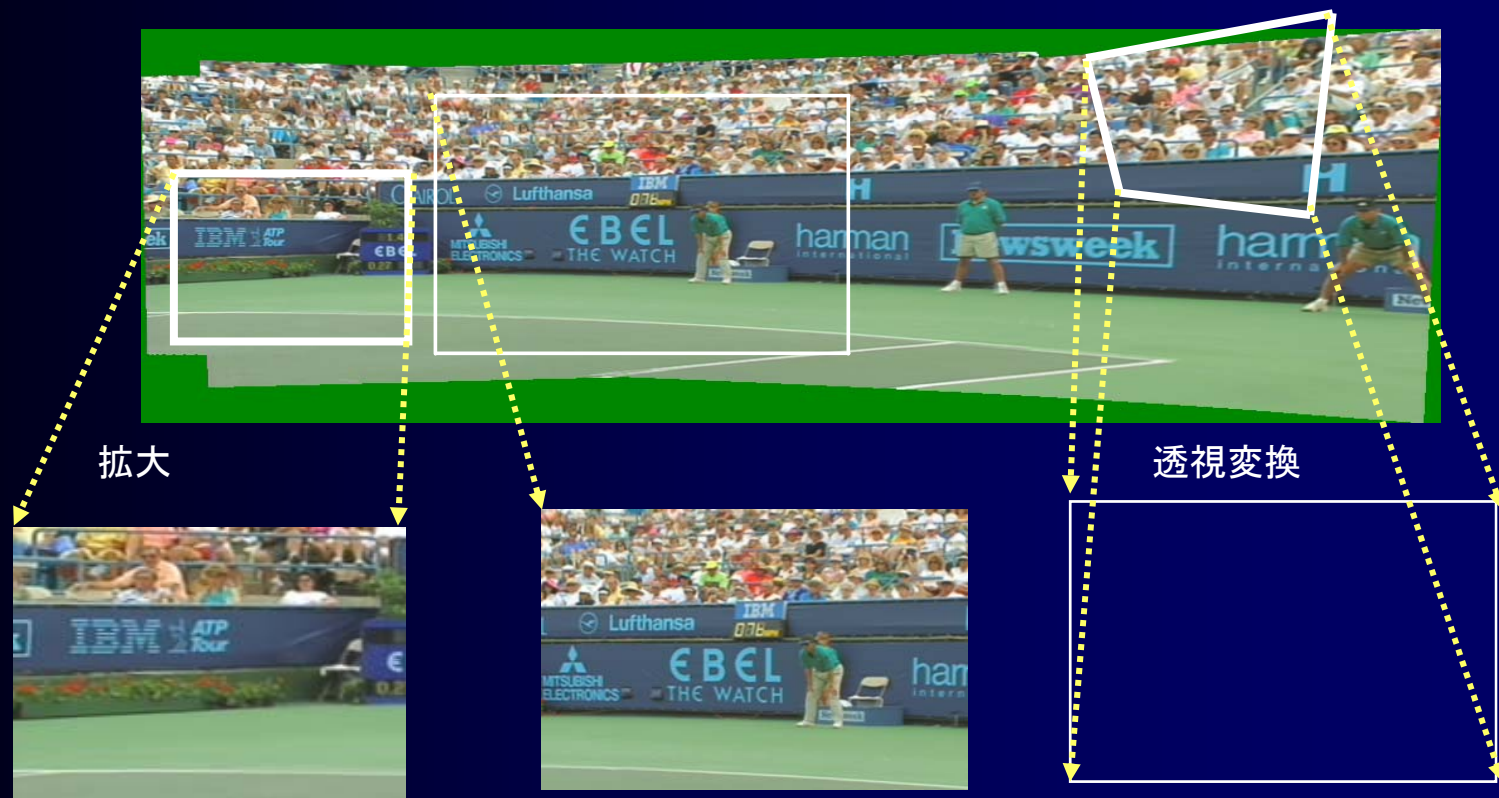
- ビデオクリップ全体の背景(スプライト)を予め作成
- スプライトをフレーム内符号化
- オブジェクトは, スプライトの幾何変換により予測
- 幾何変換は以下から選択可能
 - 静止
 - 平行移動
 - アフィン変換
 - 透視変換
 - ローカルアフィン変換
- 予測誤差は符号化しない

Sprite Coding

- Creation of Sprite (background) for entire video clip
- Intraframe coding for Sprite
- Object is predicted by Sprite's geometric transform
- Geometric transform can be selected
 - Still
 - Parallel motion
 - Affine transform
 - Projection transform
 - Local Affine transform
- Prediction error is not coded

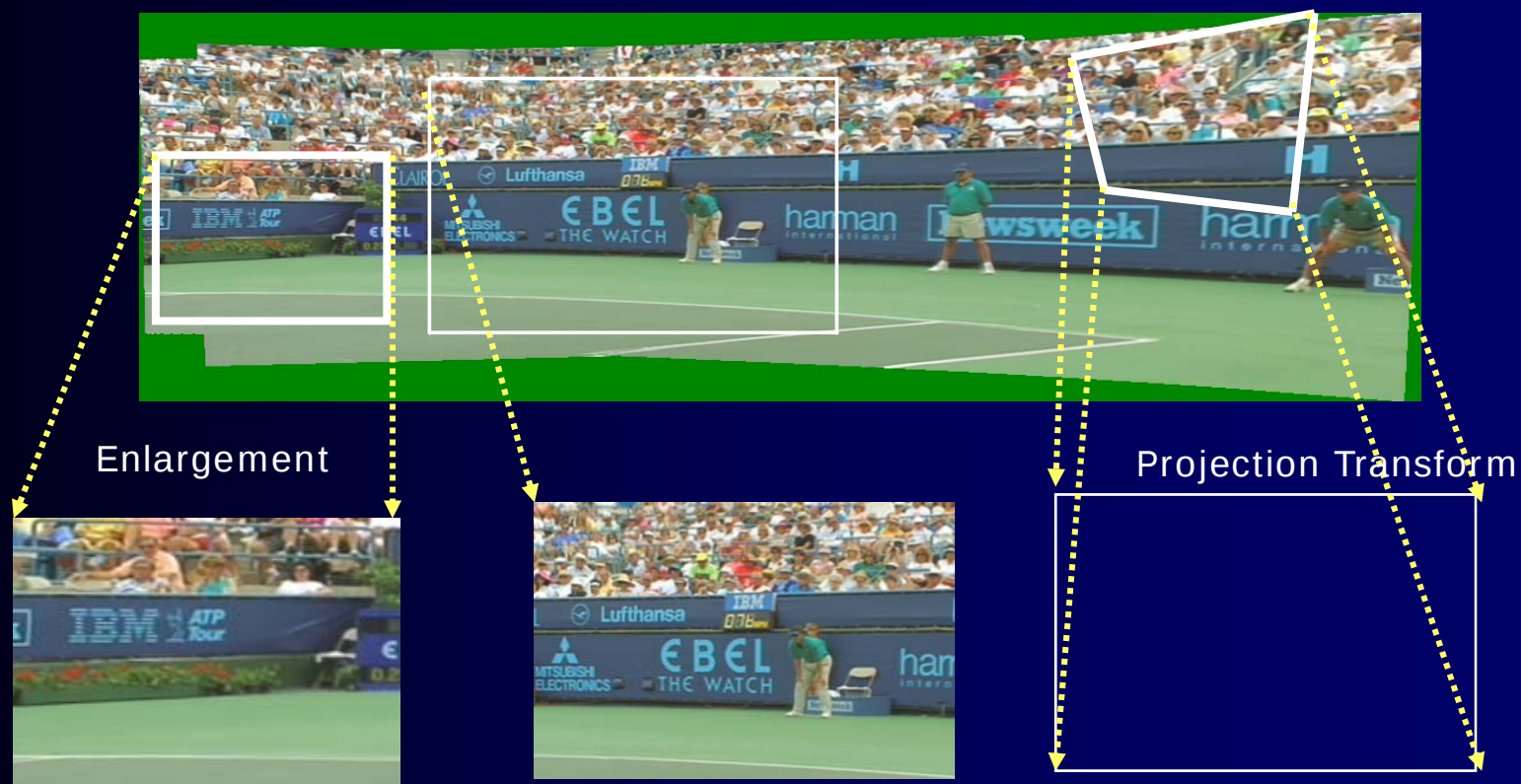
スプライト符号化 (2)

■ 幾何変換



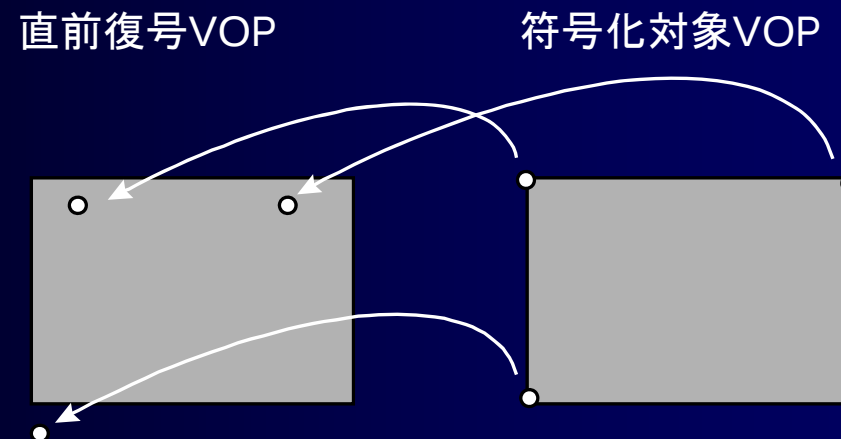
Sprite Coding (2)

- Geometric Transform



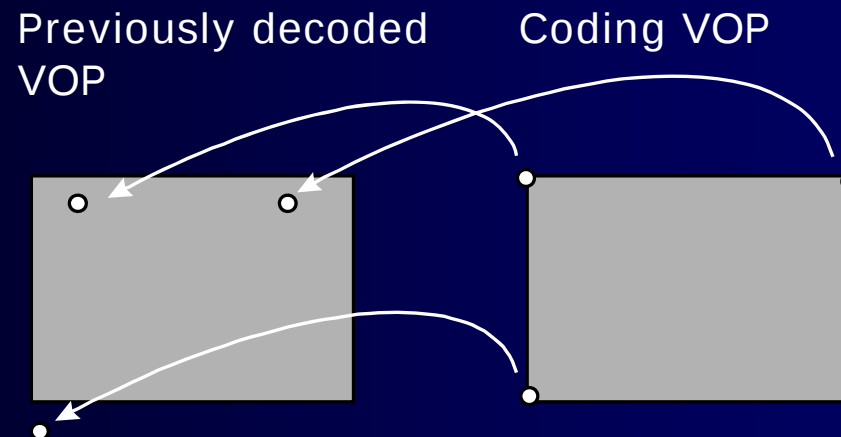
グローバル動き補償

- 画面全体の動きを幾何変換で予測
- スプライトの代わりに直前復号VOPを参照
- 予測誤差符号化



Global Motion Compensation

- Prediction of image by geometric transform
- Refer the previous VOP instead of Sprite
- Prediction error is coded

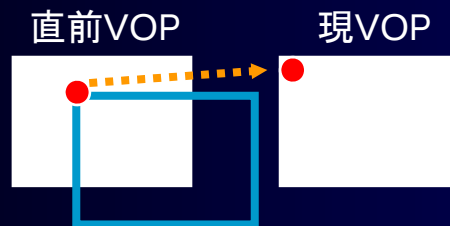


グローバル動き補償 (2)

■ 幾何学パラメータ

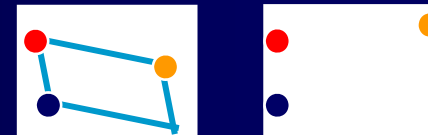
1点参照(平行移動)

$$x' = x + c, y' = y + f$$



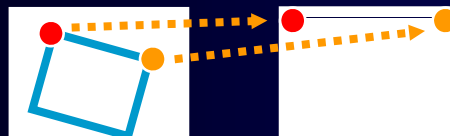
3点参照(アフィン変換)

$$x' = ax + by + c$$
$$y' = dx + ey + f$$



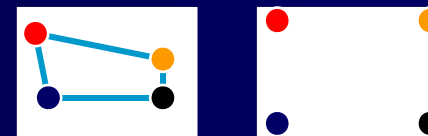
2点参照(拡大・縮小・回転)

$$x' = ax + by + c$$
$$y' = -bx + ay + f$$



4点参照(透視変換)

$$x' = (ax + by + c) / (gx + hy + 1)$$
$$y' = (dx + ey + f) / (gx + hy + 1)$$



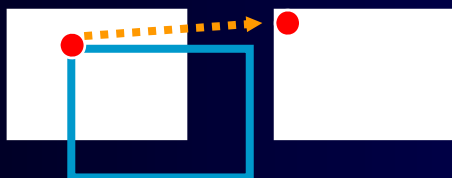
Global Motion Compensation (2)

■ Geometric parameters

1-point refer(parallel motion)

$$x' = x + c, \quad y' = y + f$$

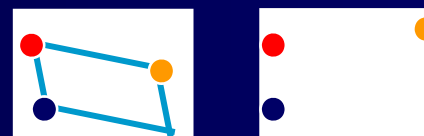
Previous VOP Current VOP



3-point refer(Affine transform)

$$x' = ax + by + c$$

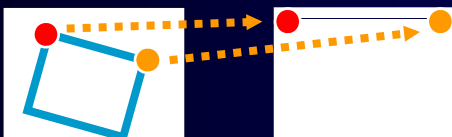
$$y' = dx + ey + f$$



2-point refer(enlarge/shrink, rotation)

$$x' = ax + by + c$$

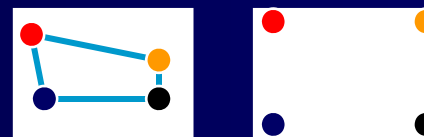
$$y' = -bx + ay + f$$



4-point refer(projection transform)

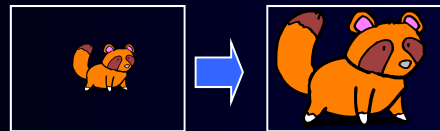
$$x' = (ax + by + c) / (gx + hy + 1)$$

$$y' = (dx + ey + f) / (gx + hy + 1)$$

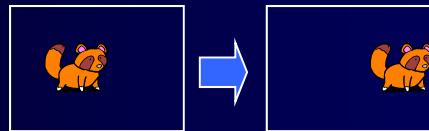
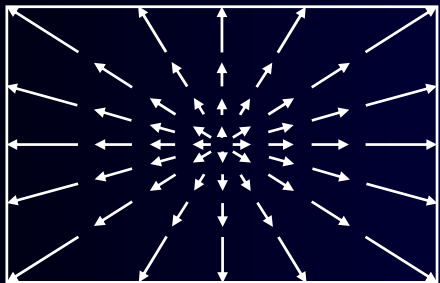


グローバル動き補償 (3)

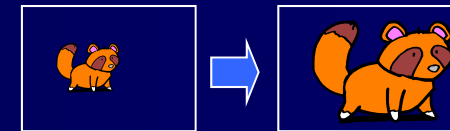
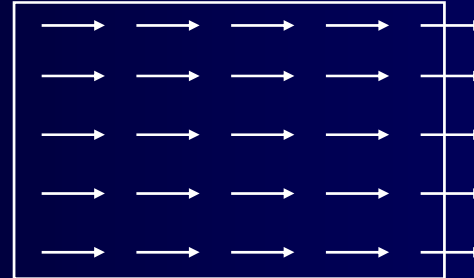
- カメラパラメータの推定
 - 従来法: 矢印の動きをすべて符号化
 - GMC: ズームパラメータおよびパニングパラメータの2個



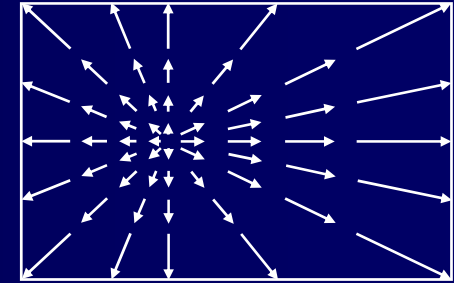
Zooming parameter Z



Panning parameter P

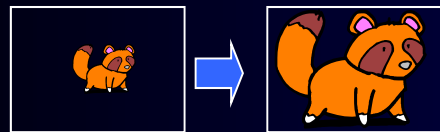


Zooming Z & Panning P

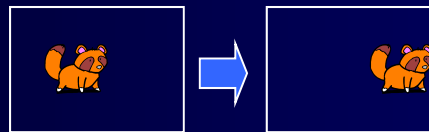
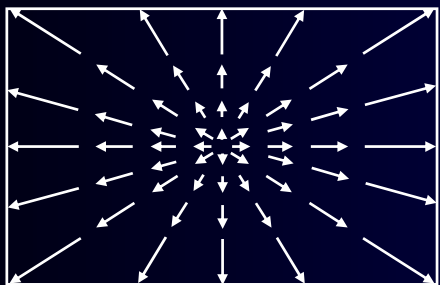


Global Motion Compensation (3)

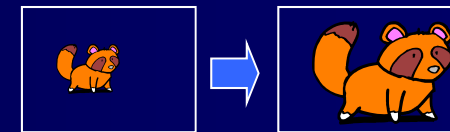
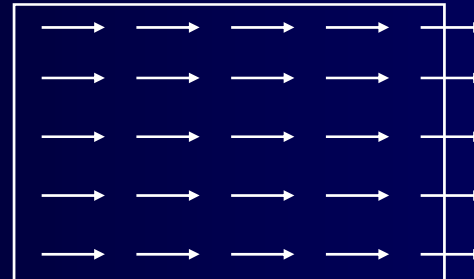
- Estimation of camera parameters
 - Conventional method: each motion is all coded
 - GMC: zooming and panning parameters only



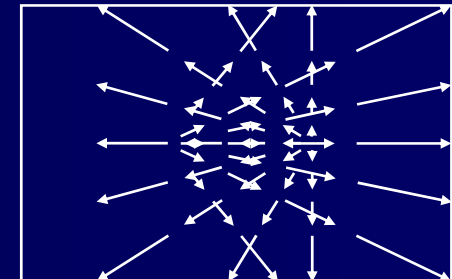
Zooming parameter Z



Panning parameter P



Zooming Z & Panning P



MPEG-4 AVC/H.264 (ISO/IEC 14496-10, 2003)

- Title: Coding of audio-visual objects,
 - Part 10: Advanced Video Coding
- プロジェクト
 - 2001年5月に開始, 最初のISは2003年に完成
- 目標
 - 超高圧縮ビデオ符号化
- アプローチ
 - $0.95^{14} < 0.5$

MPEG-4 AVC/H.264 (ISO/IEC 14496-10, 2003)

- Title: Coding of audio-visual objects,
 - Part 10: Advanced Video Coding
- Project
 - Started in May 2001, The first IS was finished in 2003
- Purpose
 - High Efficiency Video Coding
- Approach
 - $0.95^{14} < 0.5$

Main Technology in MPEG-4 AVC/H.264

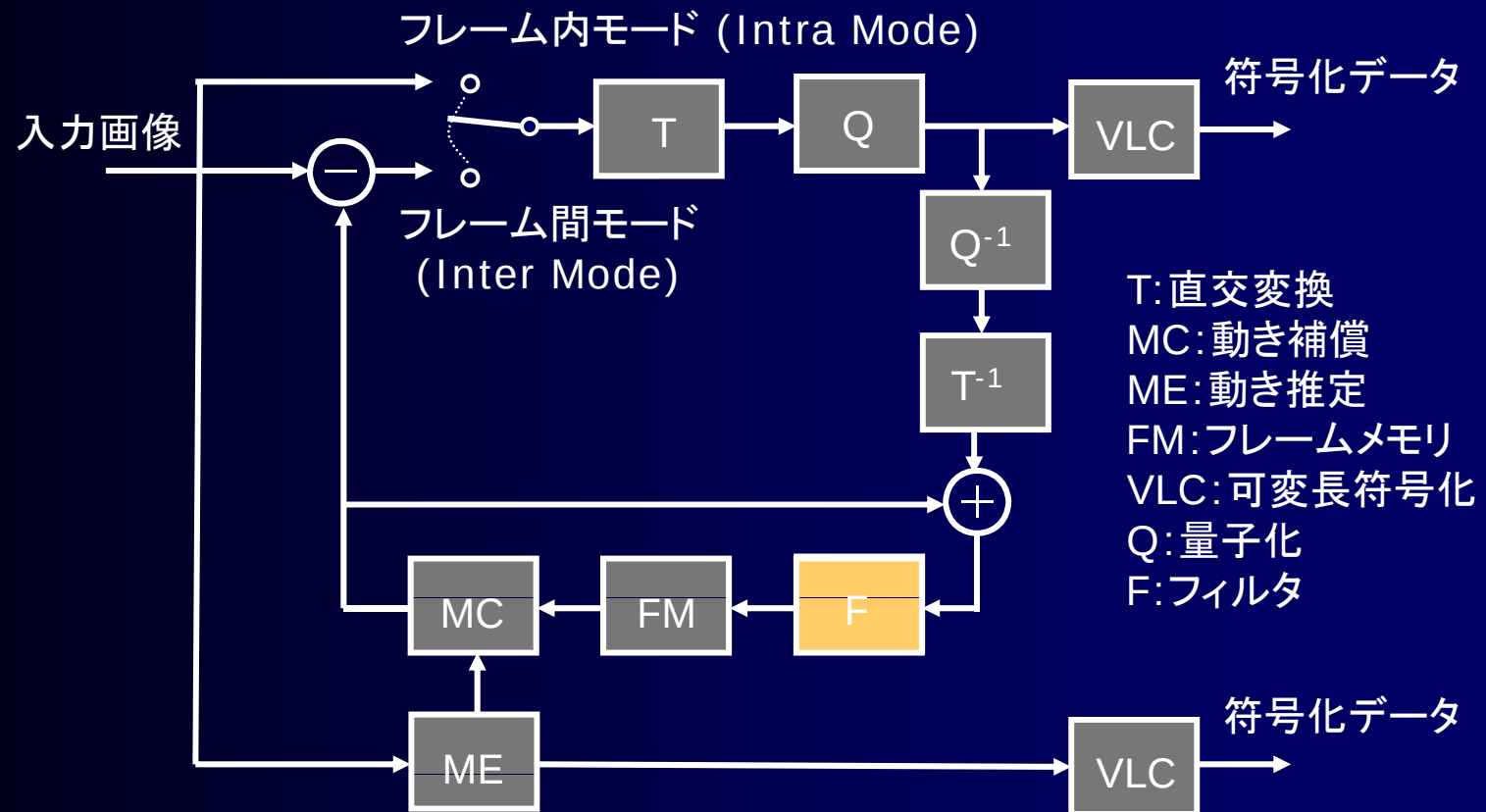
- 基本方式は動き補償フレーム間差分符号化
- ループ内にデブロッキングフィルタを使用
- ピクチャに代わりスライス単位の予測を採用
- 1/4画素精度の動きベクトル
- 直交変換に4x4整数精度DCTあるいはアダマール変換を採用
- 細かな量子化ステップサイズの採用
- エントロピー符号化
 - 誤り耐性の強い指数Golomb符号
 - 符号化効率の高いVLCとしてCAVLC (Context-based Adaptive Variable Length Coding)
 - 高圧縮を達成する算術符号化としてCABAC (Context-based Adaptive Binary Arithmetic Coding)

MPEG-4 AVC/H.264の主要技術

- Motion Compensated Interframe Prediction
- De-blocking filter as a loop filter
- Slice based prediction
- 1/4 pixel accuracy motion vector
- 4x4 integer DCT, Hadamard Transform
- Fine quantization step size
- Entropy coding
 - Error Resilient Golomb Code
 - Simple type: CAVLC (Context-based Adaptive Variable Length Coding)
 - Complex type: CABAC (Context-based Adaptive Binary Arithmetic Coding)

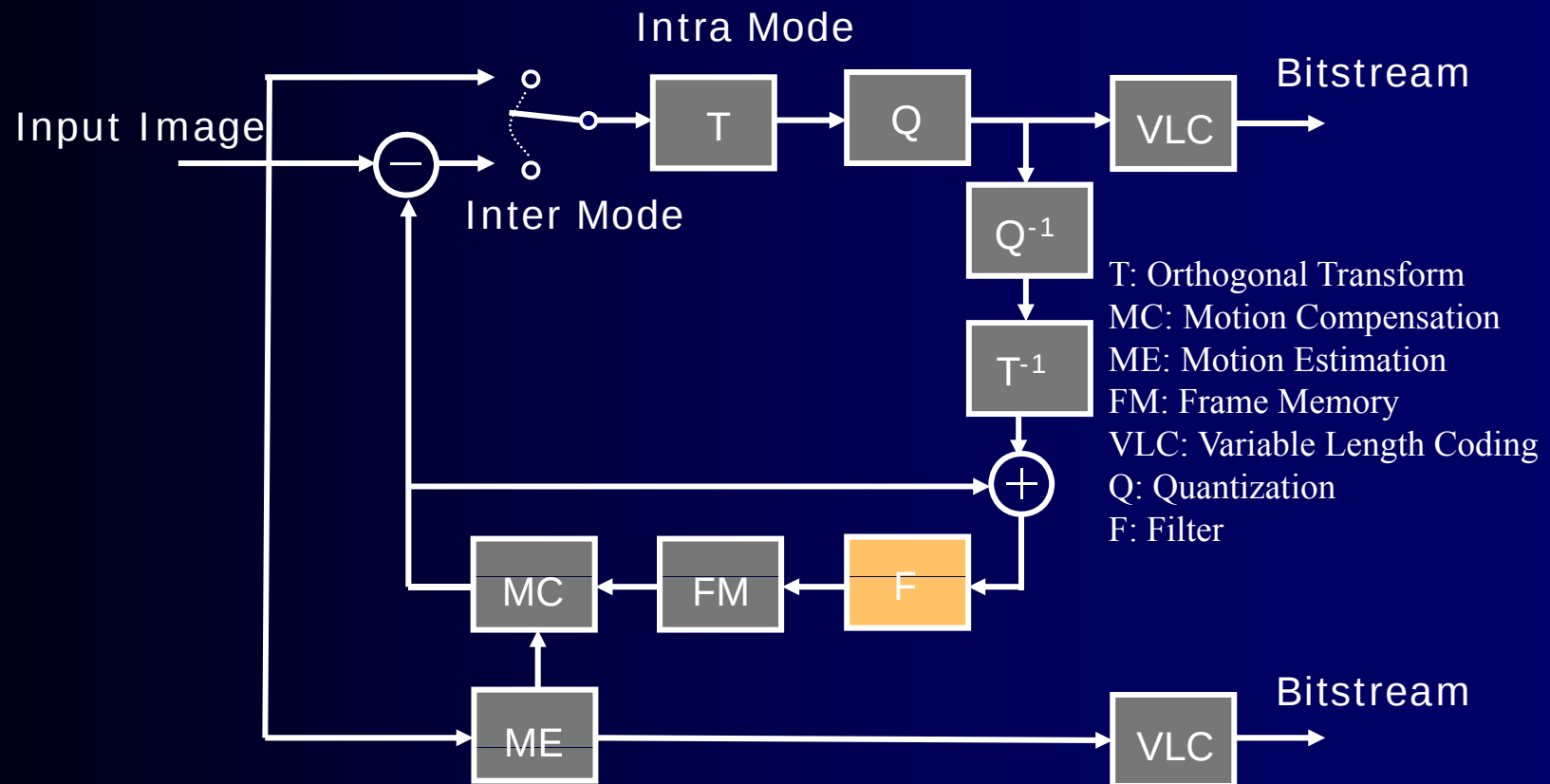
デブロッキングフィルタ (1)

■ 符号化器 (Encoder) の構成



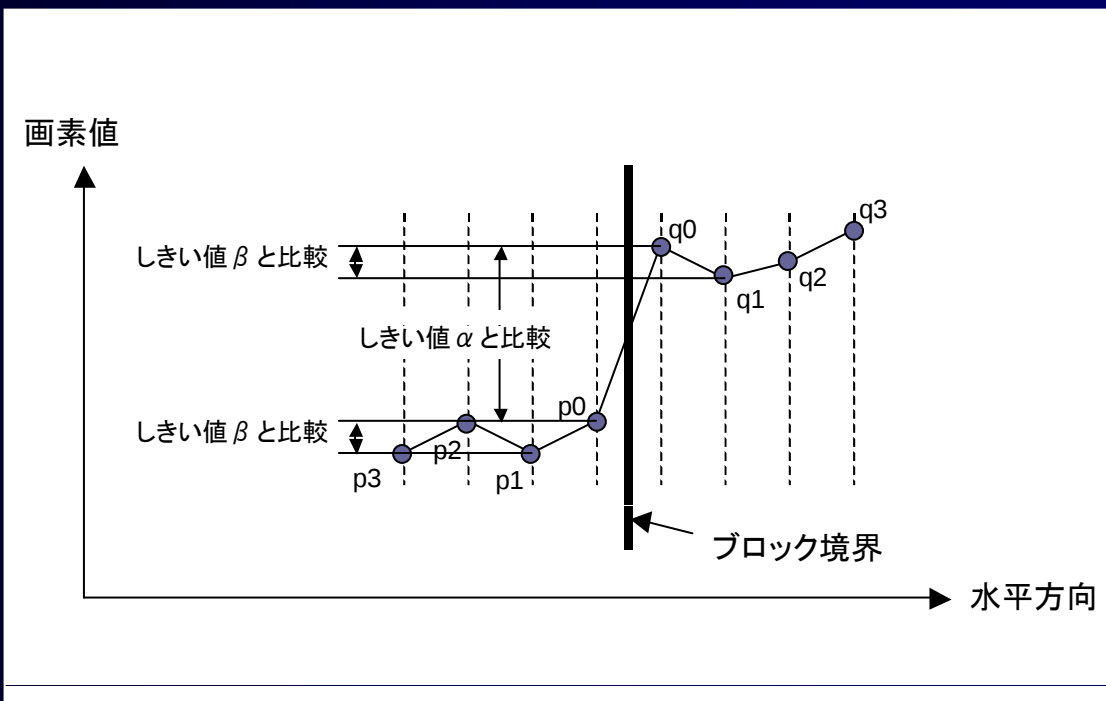
De-blocking Filter (1)

■ Encoder structure



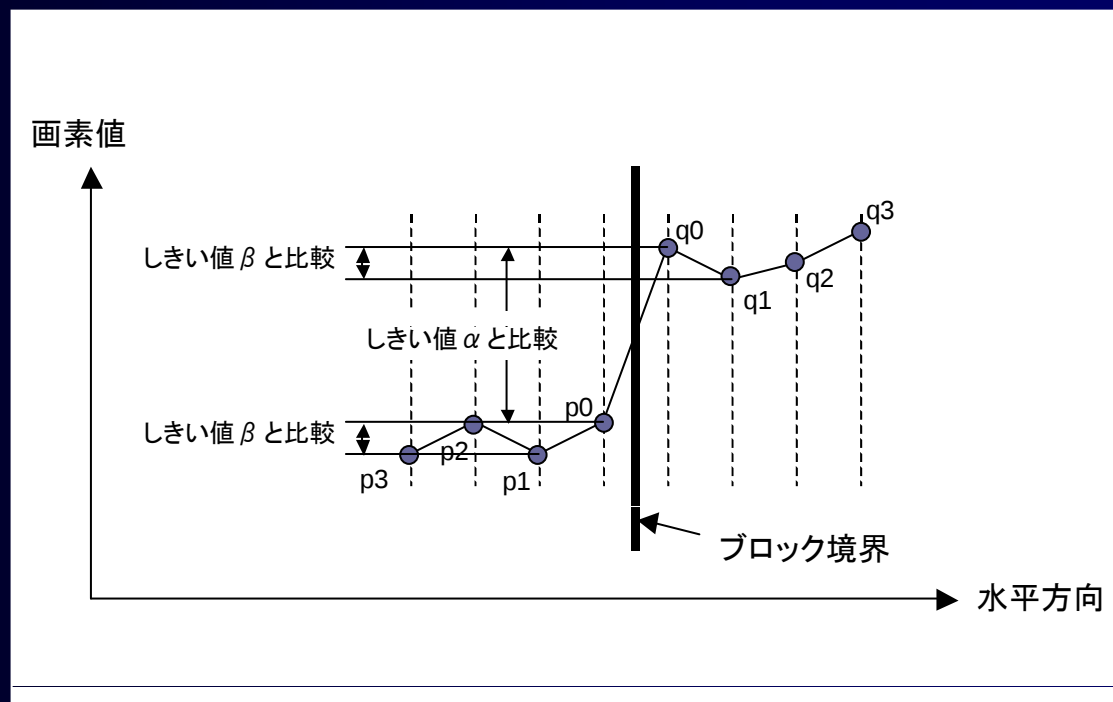
デブロッキングフィルタ (2)

- ブロック境界のみに適応的に処理
- ループ内フィルタ



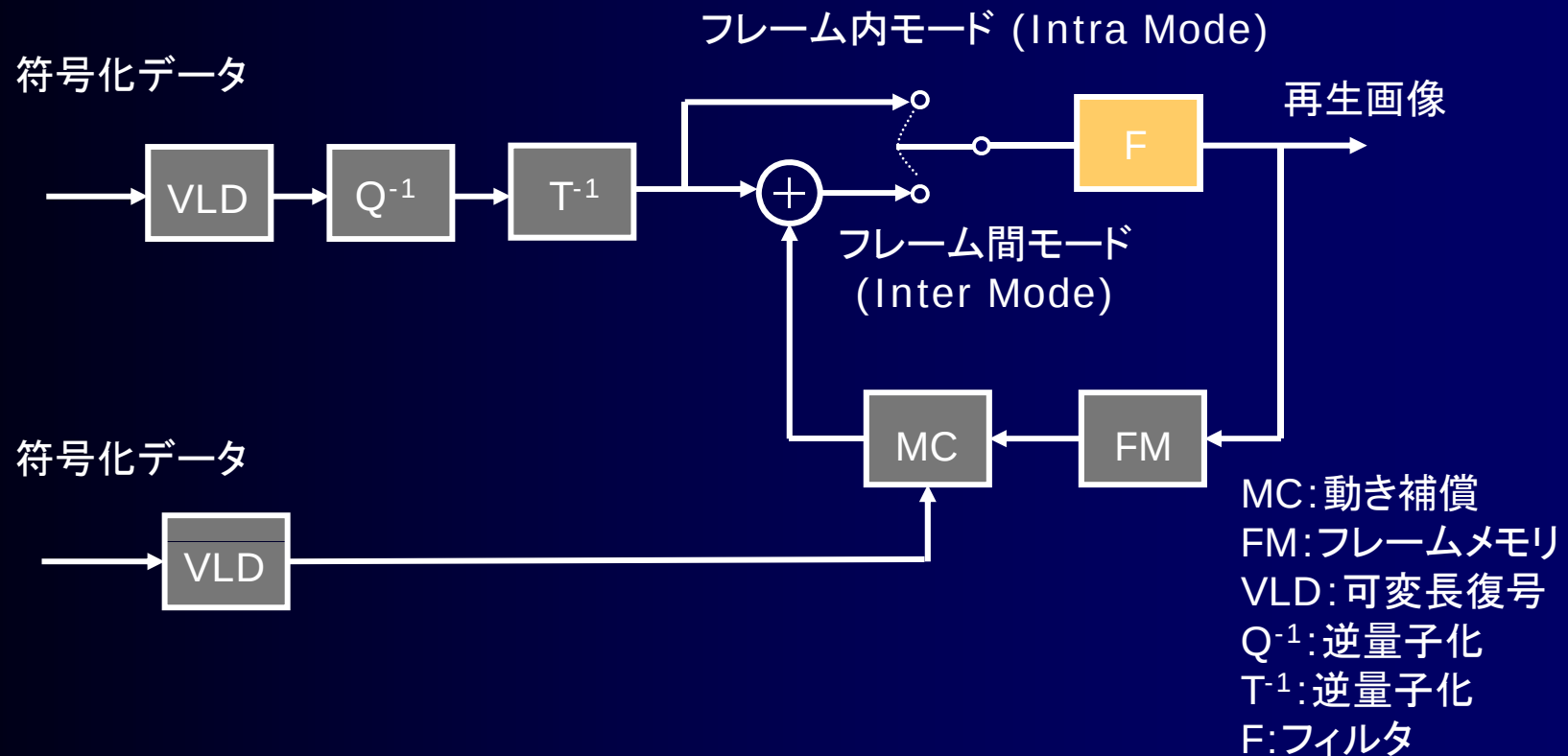
De-blocking Filter (2)

- Adaptive filtering to block boundary
- Inloop filter



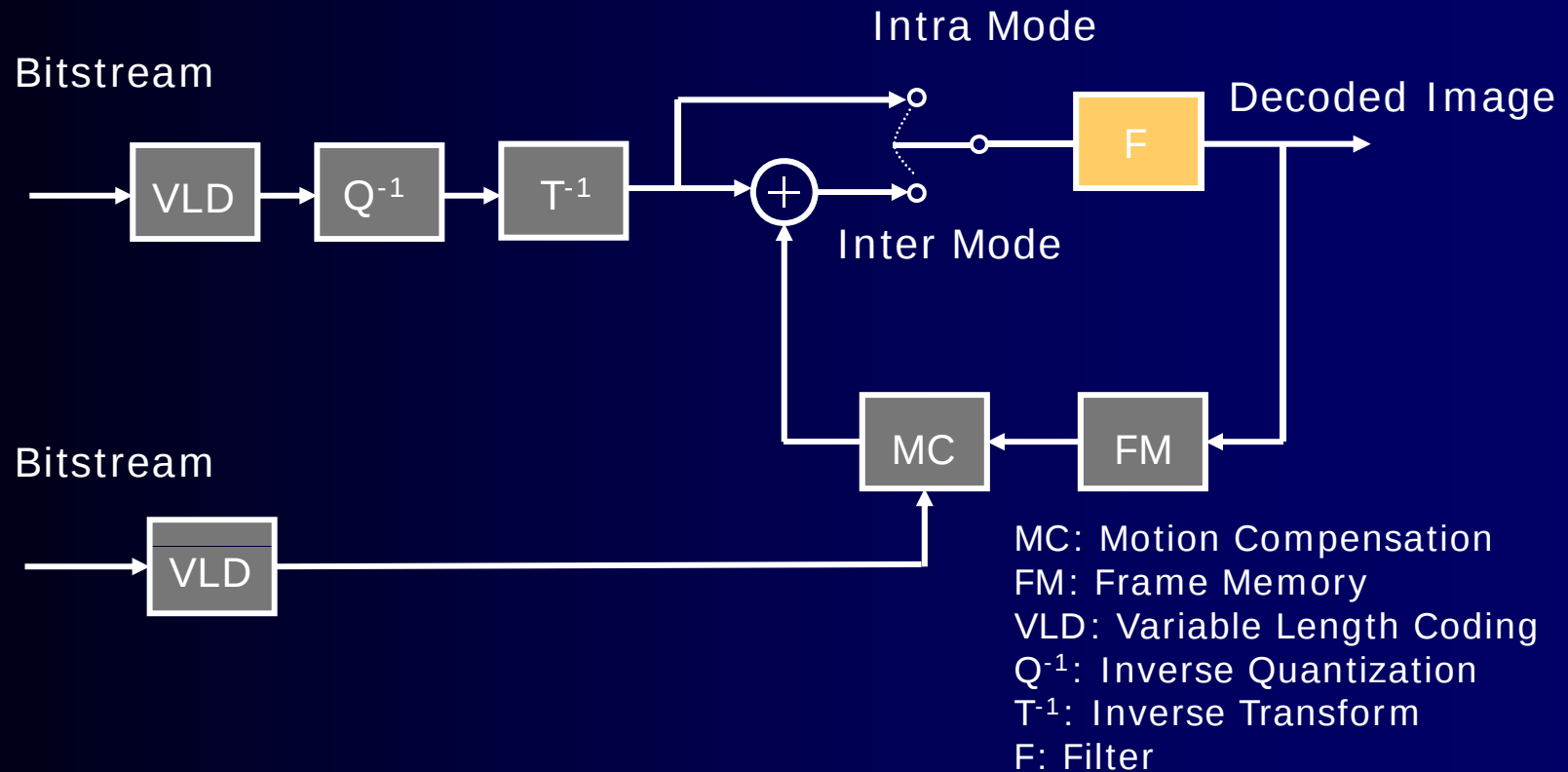
デブロッキングフィルタ (3)

■ 復号器 (Decoder) の構成



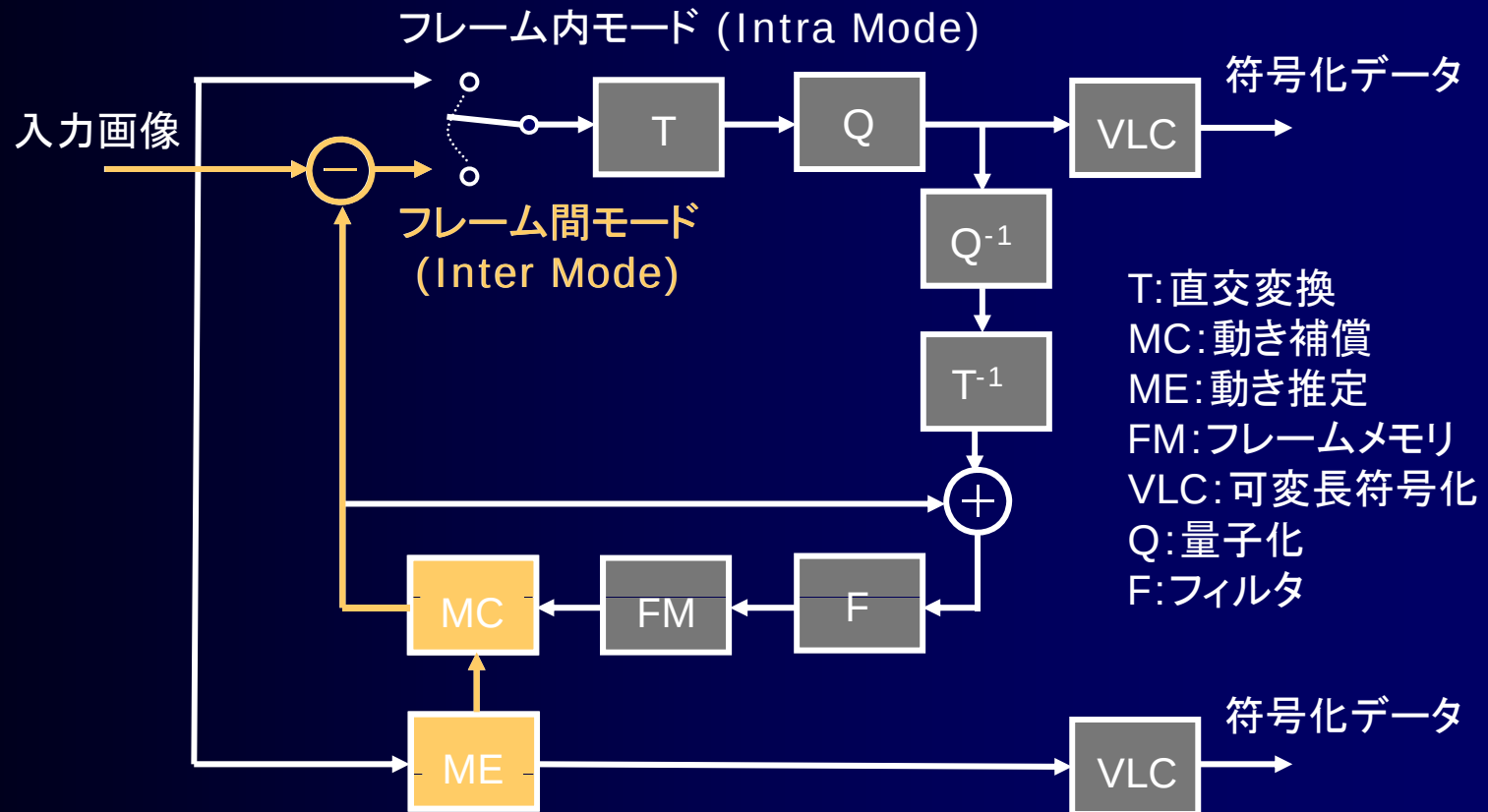
De-blocking Filter (3)

■ Decoder Structure



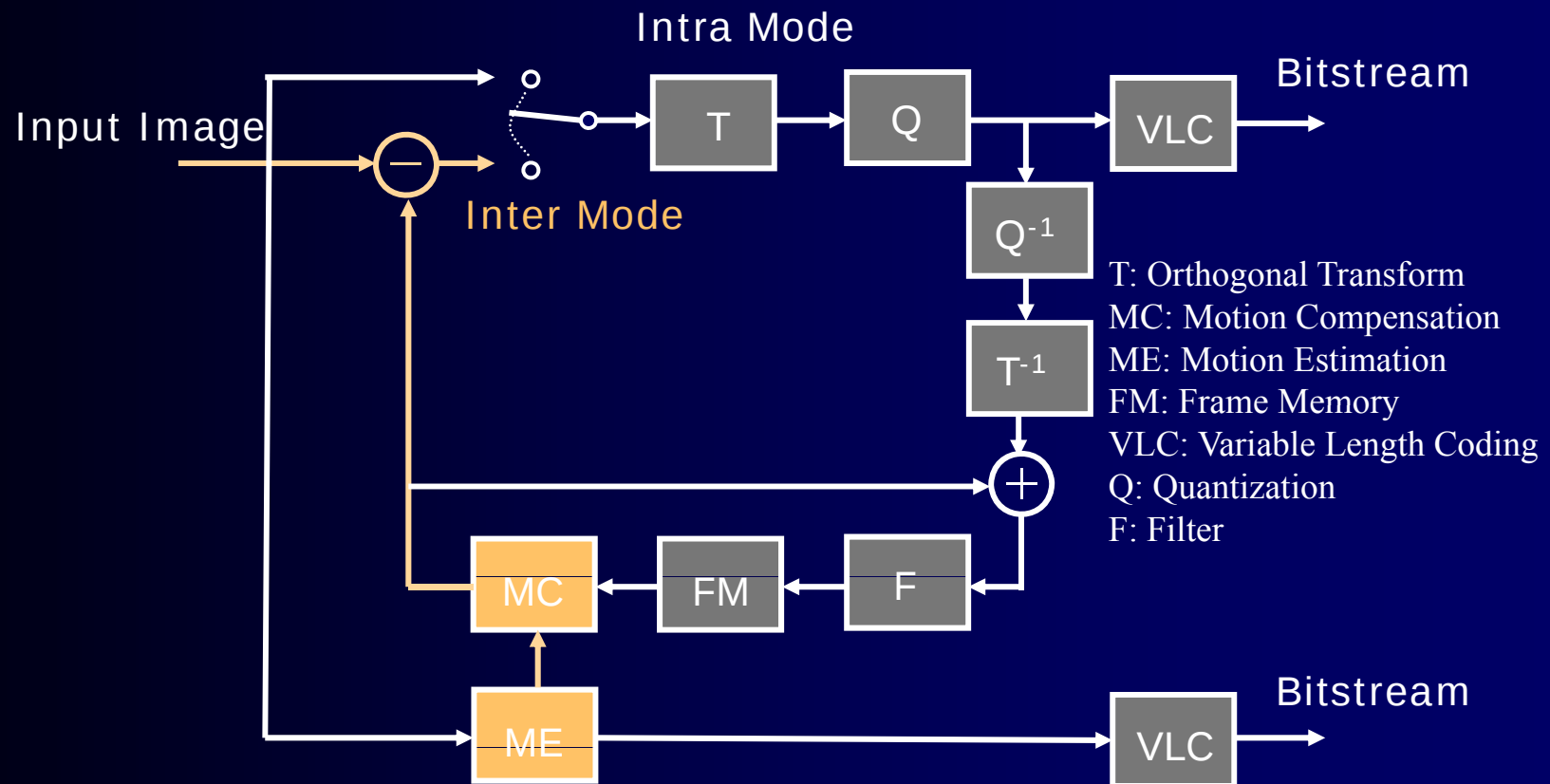
1/4画素精度動き補償 (1)

■ 符号化器 (Encoder) の構成



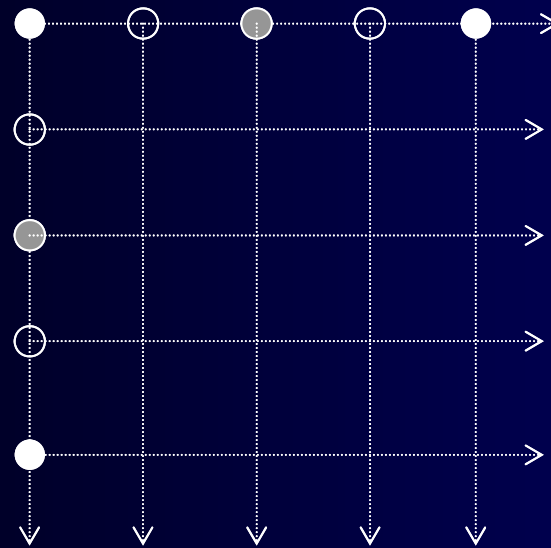
1/4 Pel Accuracy MC (1)

■ Encoder structure



1/4画素精度動き補償 (2)

- 補間画像の生成による高精度動き検出

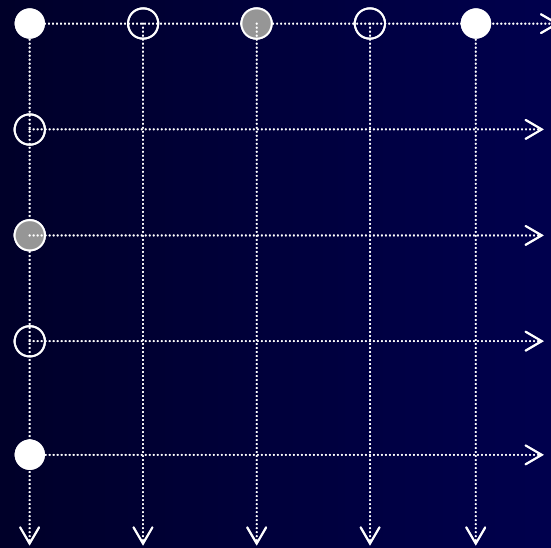


1/2画素補間フィルタ

1/4画素補間フィルタ

1/4 Pel Accuracy MC (2)

- Accurate motion estimation by fine interpolation

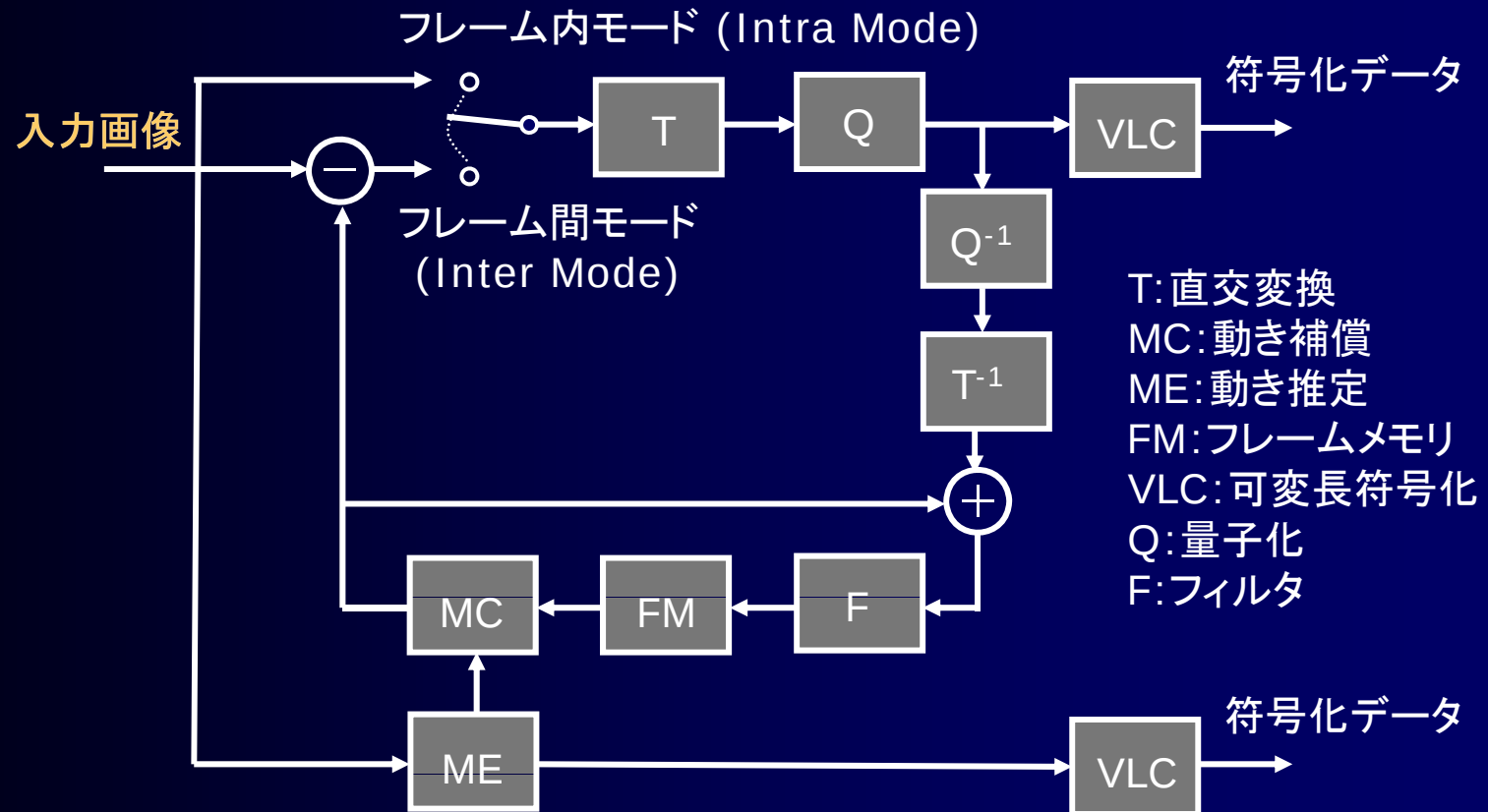


1/2 pel Interpolation filter: 6tap FIR

1/4 pel interpolation filter: 2tap FIR

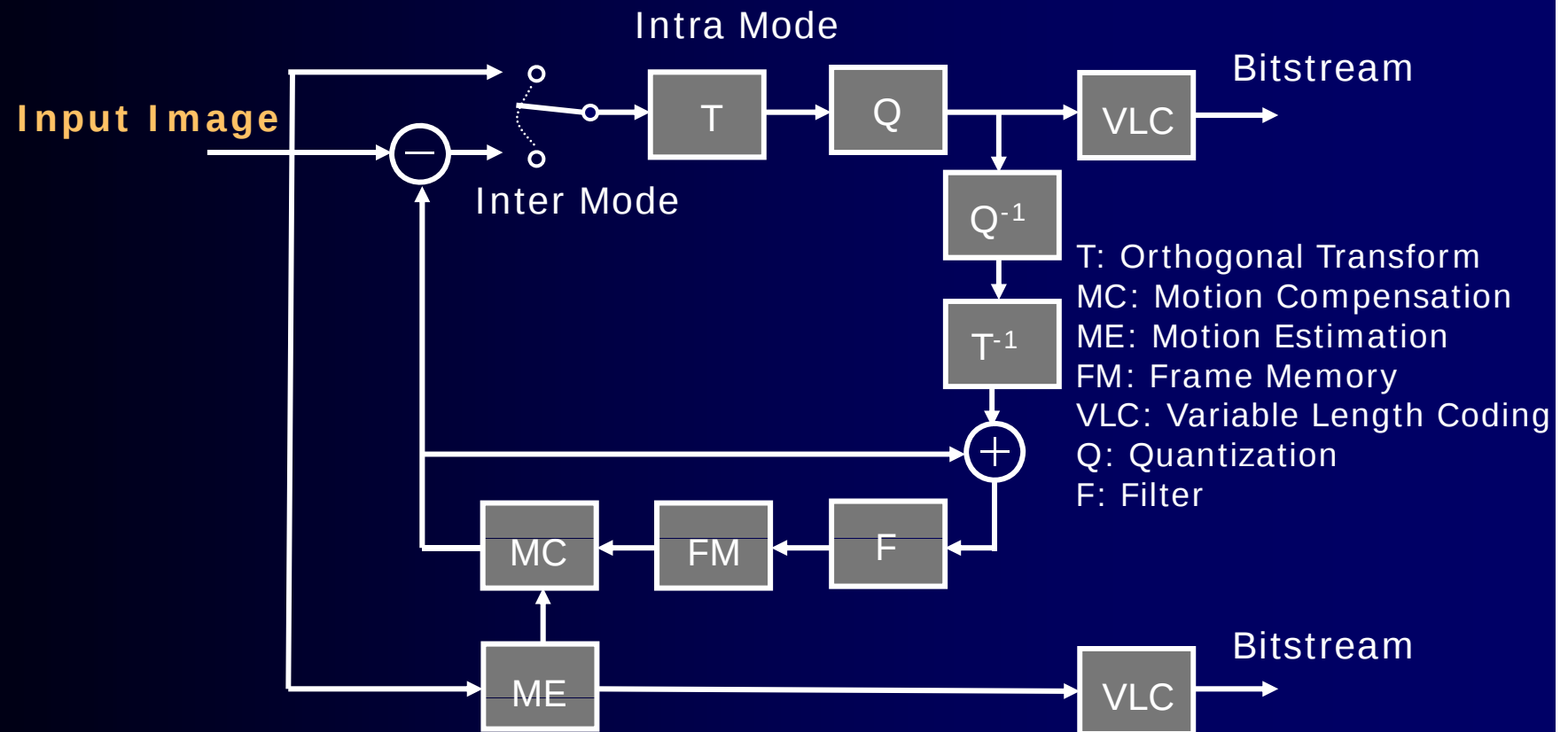
スライス単位の予測 (1)

■ 符号化器 (Encoder) の構成



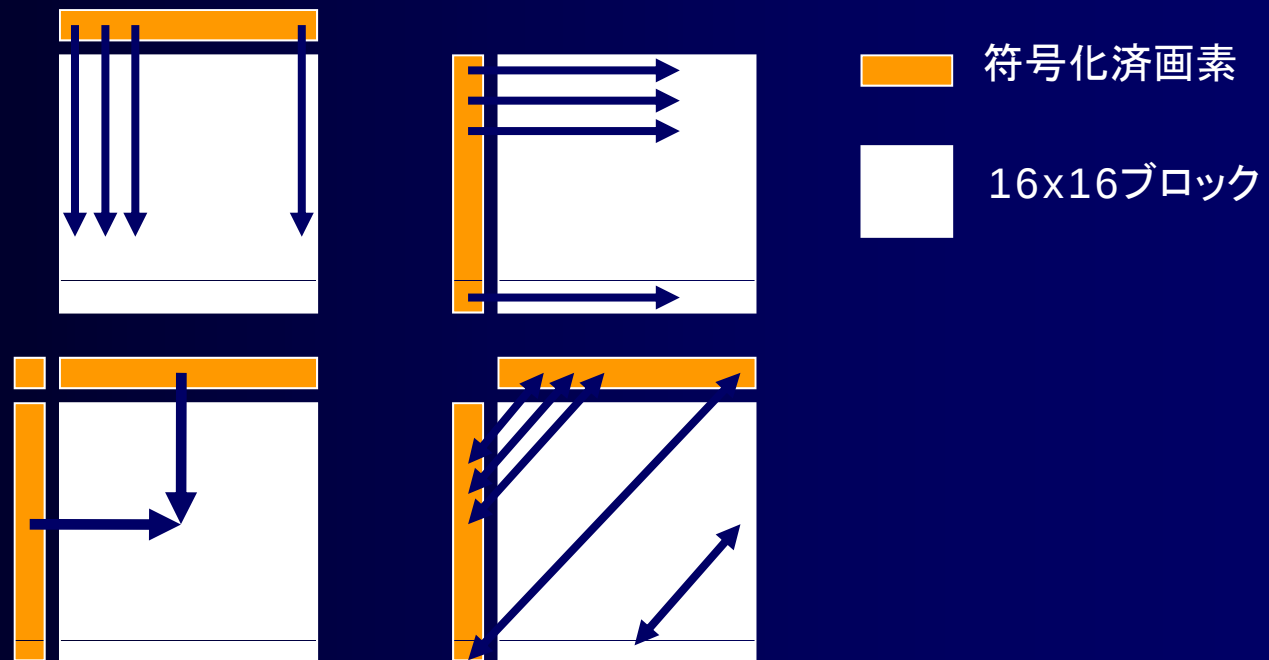
Slice Based Prediction (1)

■ Encoder structure



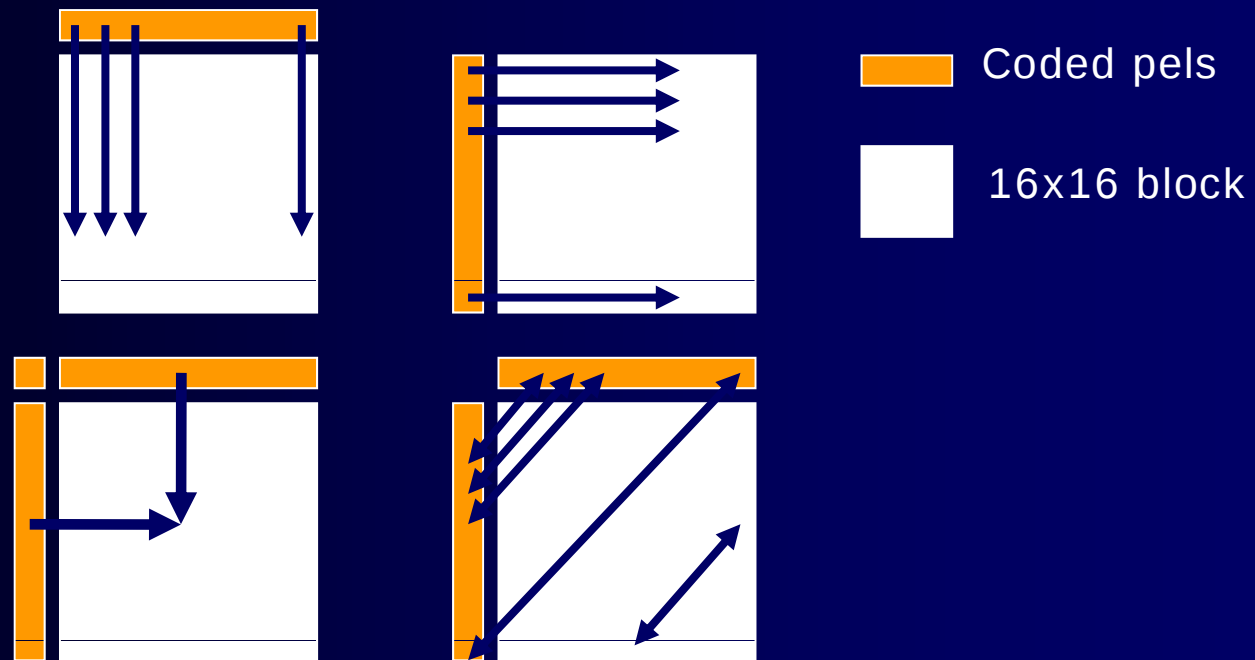
スライス単位の予測 (2)

- I-sliceでは16x16/8x8/4x4画素の3種類でフレーム内予測
- 16x16画素のブロックでは4つのモードでフレーム内予測 (垂直, 水平, 平均値, 平面予測)



Slice Based Prediction (2)

- I-slice employs intraframe prediction with 16x16/8x8/4x4 pel blocks
- In 16x16 pel block, 4 modes (vertical, horizontal, average, plane) can be selected

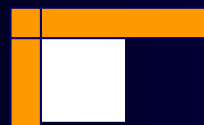


スライス単位の予測 (3)

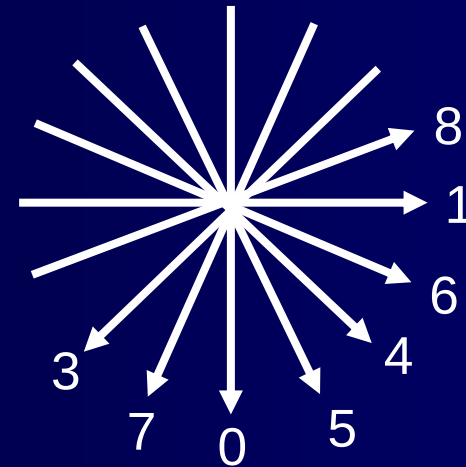
- 8x8/4x4画素のブロックでは9つのモードでフレーム内予測



8x8 ブロック



4x4 ブロック



2:DC

予測モード番号と予測方向

Slice Based Prediction (3)

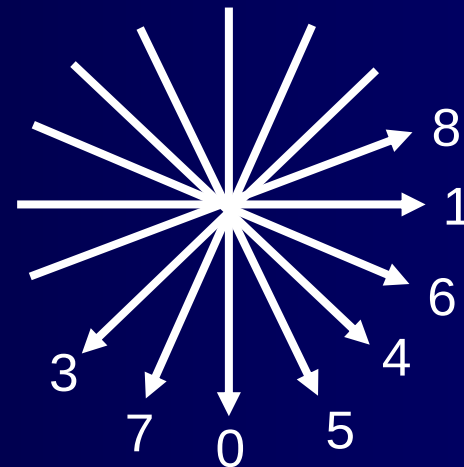
- In 8x8/4x4 pel block, 9 modes intraframe prediction are prepared



8x8 block



4x4 block

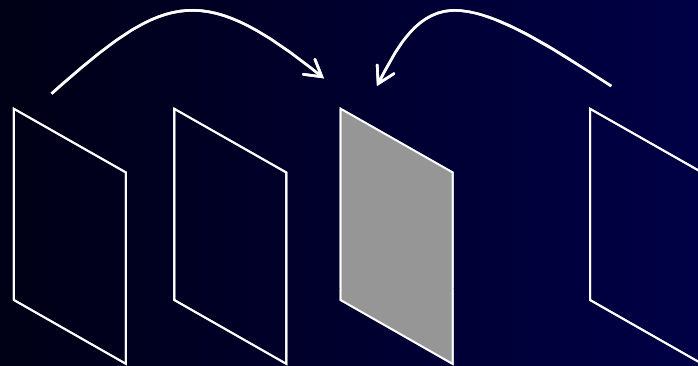


2:DC

Prediction mode number and
Its direction

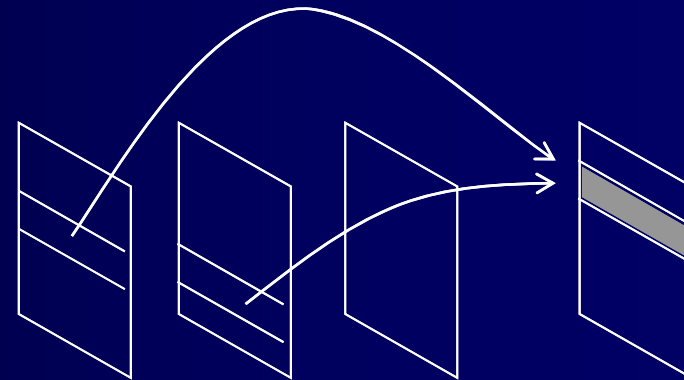
スライス単位の予測 (4)

フレーム単位の双方向予測



MPEG-2

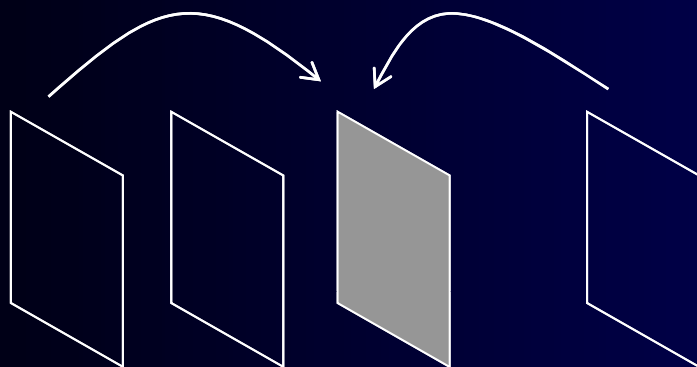
任意のスライスから予測



MPEG-4 AVC/H.264

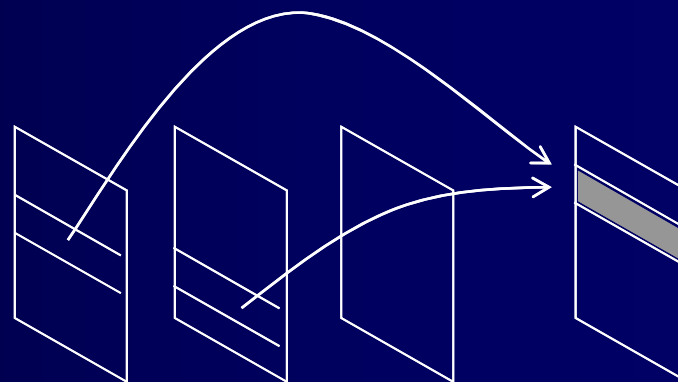
Slice Based Prediction (4)

Frame based bi-directional prediction



MPEG-2

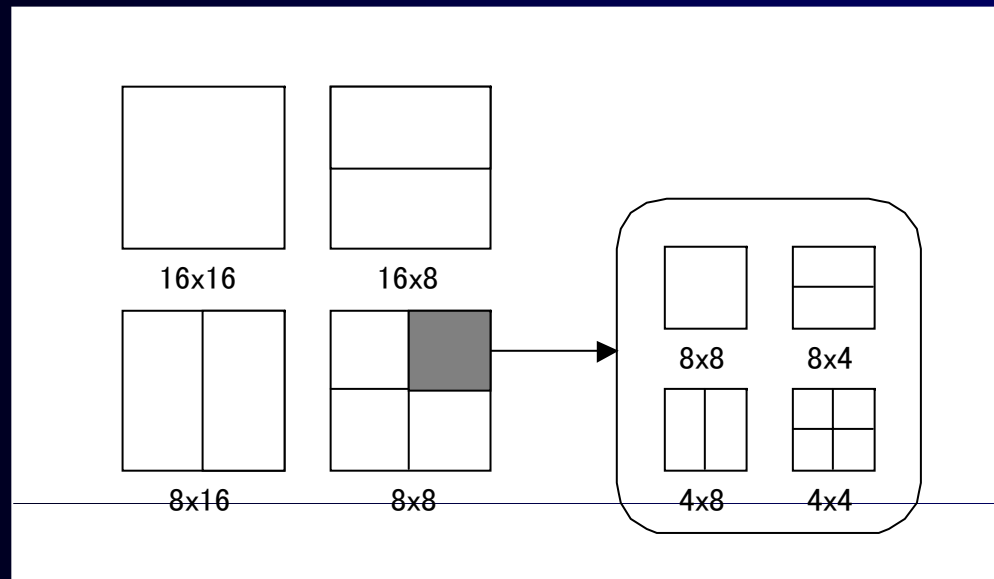
Prediction by arbitrary slices



MPEG-4 AVC/H.264

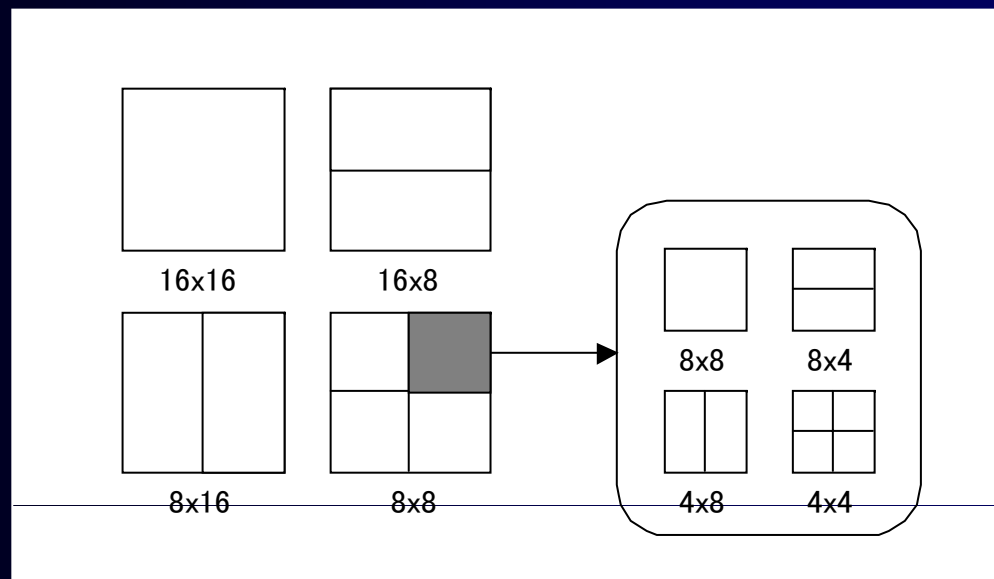
スライス単位の予測 (5)

- P-sliceでは16x16画素単位のマクロブロック
- 4種類のブロックサイズを切替えて動き補償
- 8x8画素ブロックはさらに3種類に分割可能



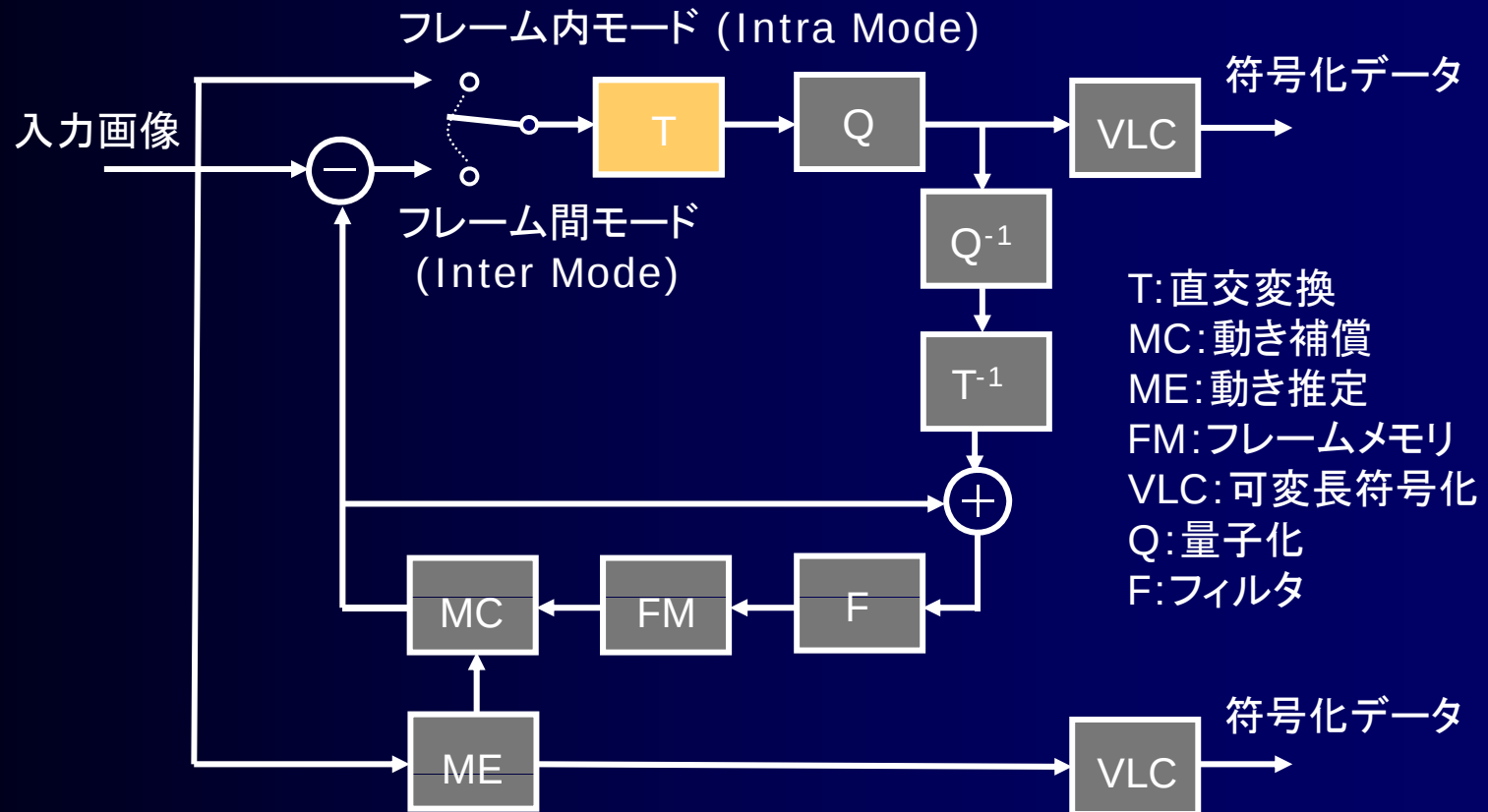
Slice Based Prediction (5)

- In P-slice, 16x16 pel macroblock
- 4 kinds of blocksize can be selected for MC
- 8x8 block can further be divided in 3 modes



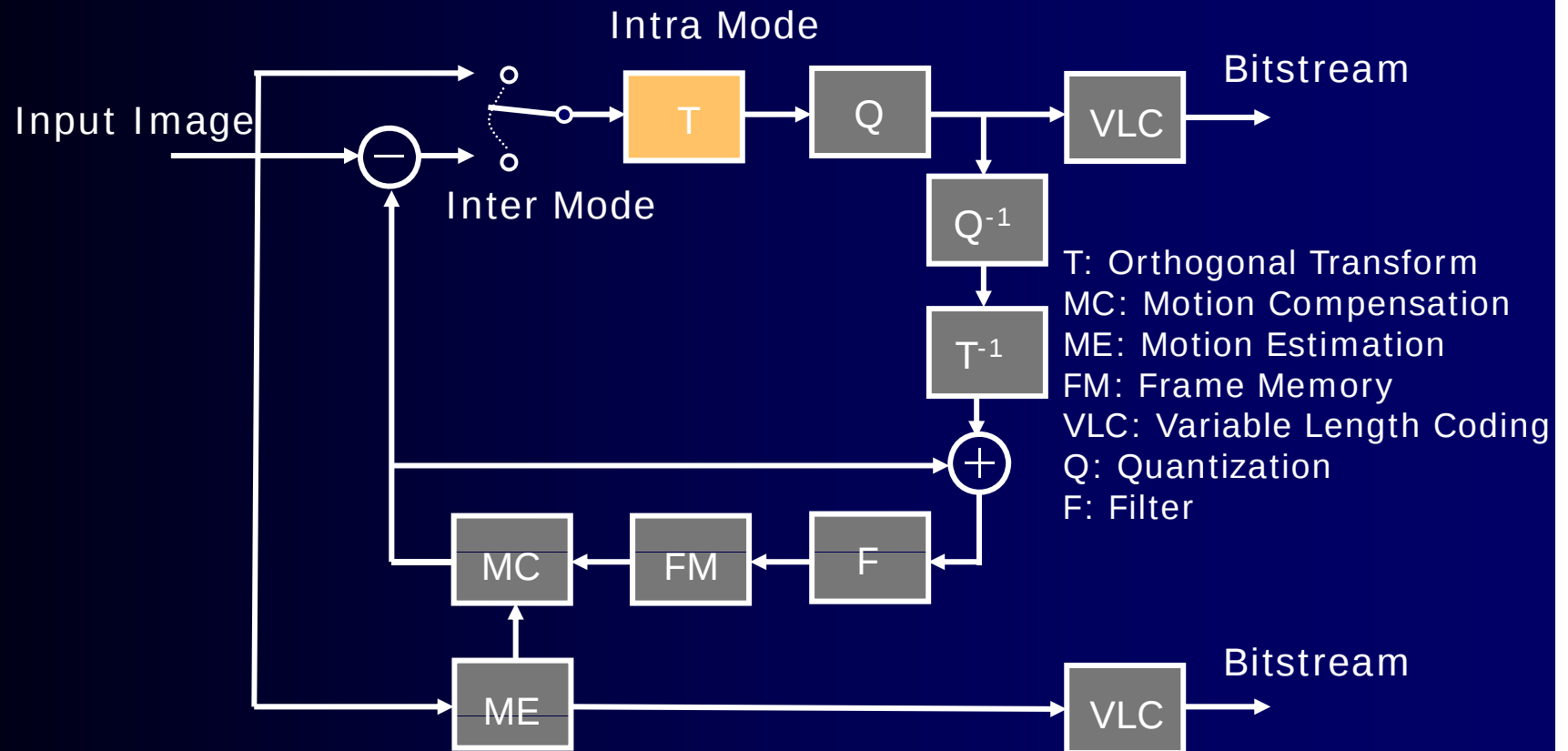
整数精度DCT (1)

■ 符号化器 (Encoder) の構成



Integer DCT (1)

■ Encoder structure



整数精度DCT (2)

- Inter-modeでは整数精度DCT (mismatch回避)

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{pmatrix}$$

- Intra-modeではアダマール変換

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Integer DCT (2)

- Integer DCT in Inter-mode to avoid mismatch

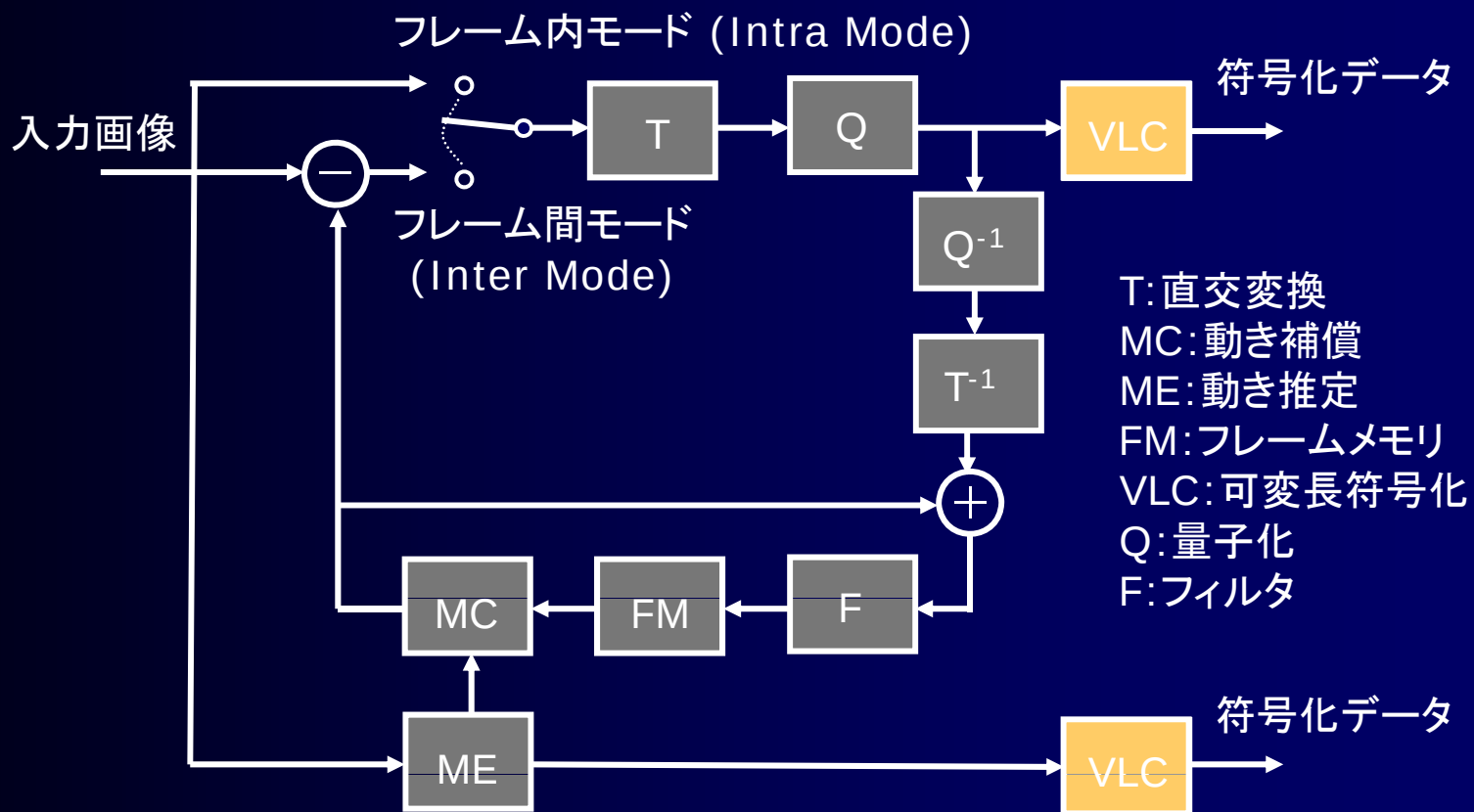
$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{pmatrix}$$

- Hadamard Transform in Intra-mode

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

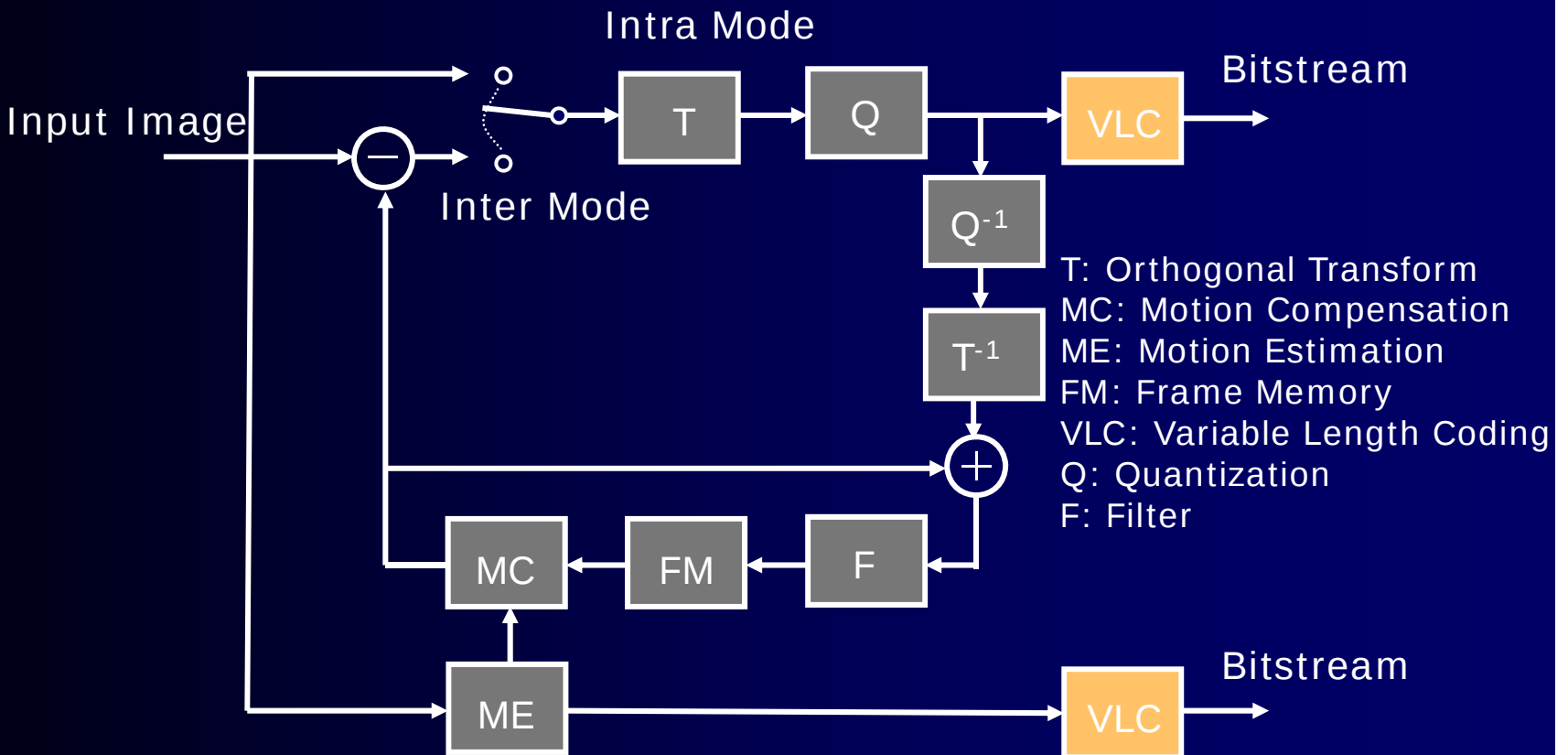
エントロピー符号化 (1)

■ 符号化器 (Encoder) の構成



Entropy Coding (1)

Encoder structure



エントロピー符号化 (2)

■ 指数Golomb符号

- ピクチャヘッダなど重要なデータに適用
- 誤り耐性を強化
- Prefix (“0”の連続) + Separator (“1”) + Suffix (“0”/”1”の組み合わせ)

指数Golomb符号		
	1	
0	1	x
0 0	1	x x
0 0 0	1	x x x
0 0 0 0	1	x x x x

Entropy Coding (2)

- Exponential Golomb Code
 - Apply to important data such as picture header
 - Enforce error resilience
 - Prefix (Consecutive “0”) + Separator (“1”) + Suffix (Combination of “0” and “1”)

Exponential Golomb Code		
	1	
0	1	x
0 0	1	x x
0 0 0	1	x x x
0 0 0 0	1	x x x x

エントロピー符号化 (3)

■ CAVLC

- コンテキスト適応型VLC
- 整数精度DCT係数の符号化に適用

種類	内容
1	DCT係数値の前の0の連続個数
2	DCT係数値
3	非0係数の個数
4	最後の非0係数以前の0係数の個数
5	最後に連続する絶対値1の係数の個数
6	最後に連続する絶対値1の係数の符号

Entropy Coding (3)

■ CAVLC

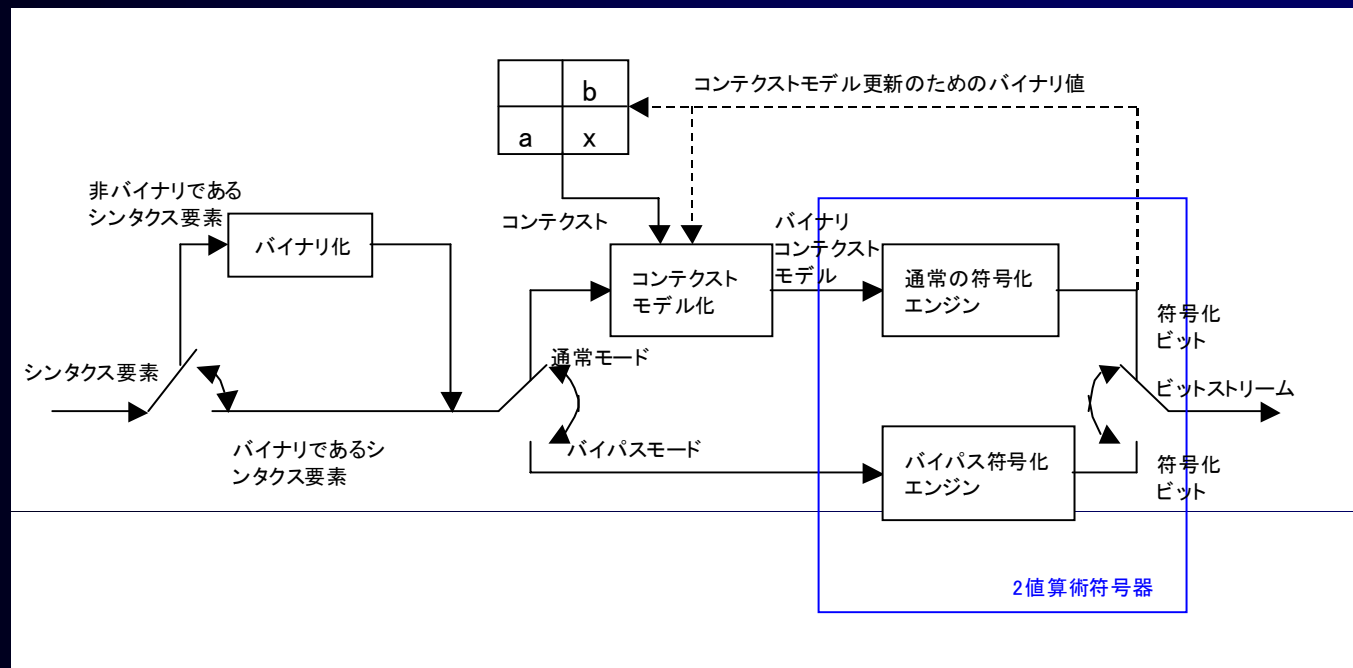
- Context Adaptive VLC
- Apply to integer accuracy DCT coefficient

No	Contxt
1	Consecutive “0” before DCT coefficient
2	DCT coefficient
3	Number of non-zero coefficient
4	Number of zero coeff before the last non-zero coeff.
5	Number of consecutive 1 coeff at the last
6	Flag of consecutive “1” coeff at the last

エントロピー符号化 (4)

■ CABAC

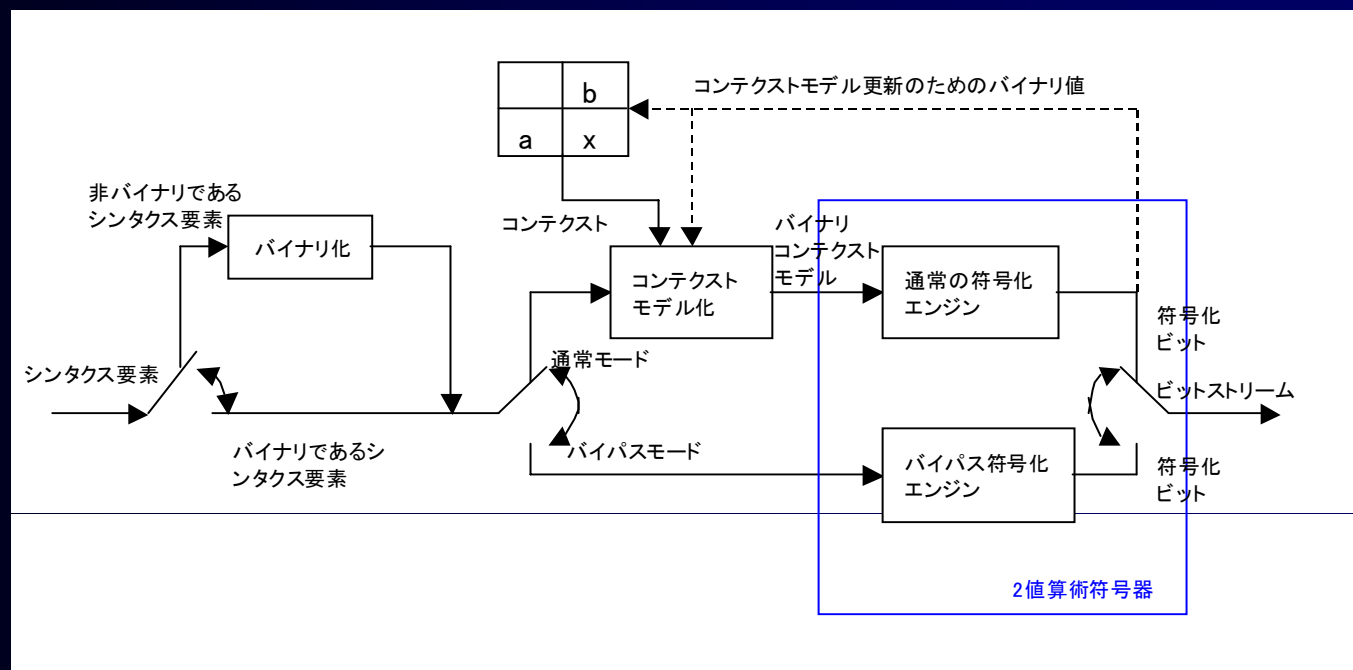
– コンテキスト適応型2値算術符号化



Entropy Coding (4)

■ CABAC

- Context Adaptive Binary Arithmetic Coding



プロファイルとレベル

- プロファイル
 - 3種類(スライス種類/エントロピー符号化/インタレース対応)
- レベル
 - 15種類(bitrate: 64-240kbps)

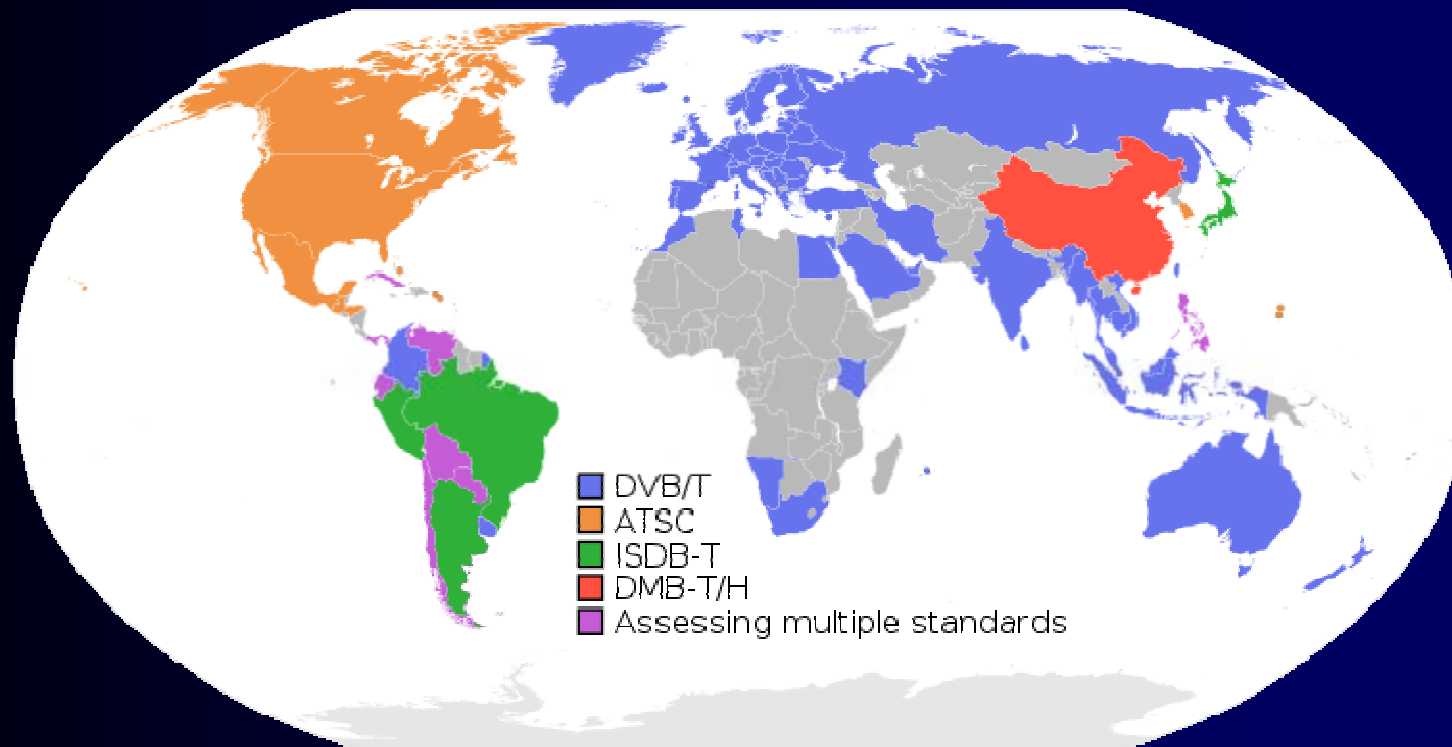
プロファイル	スライスの種類	エントロピー符号化	インタレース対応
ベースライン	I, P-slices	CAVLC	不可
拡張	I, P, B-slices	CAVLC	不可
メイン	I, P, B-slices	CAVLC, CABAC	可

Profile and Level

- Profile
 - 3 kinds (Slice type/Entropy coding type/Interlace capability)
- Level
 - 15 kinds (bitrate: 64-240kbps)

Profile	Slice Type	Entropy Coding	Interlace
Baseline	I, P-slices	CAVLC	No
Extention	I, P, B-slices	CAVLC	No
Main	I, P, B-slices	CAVLC, CABAC	OK

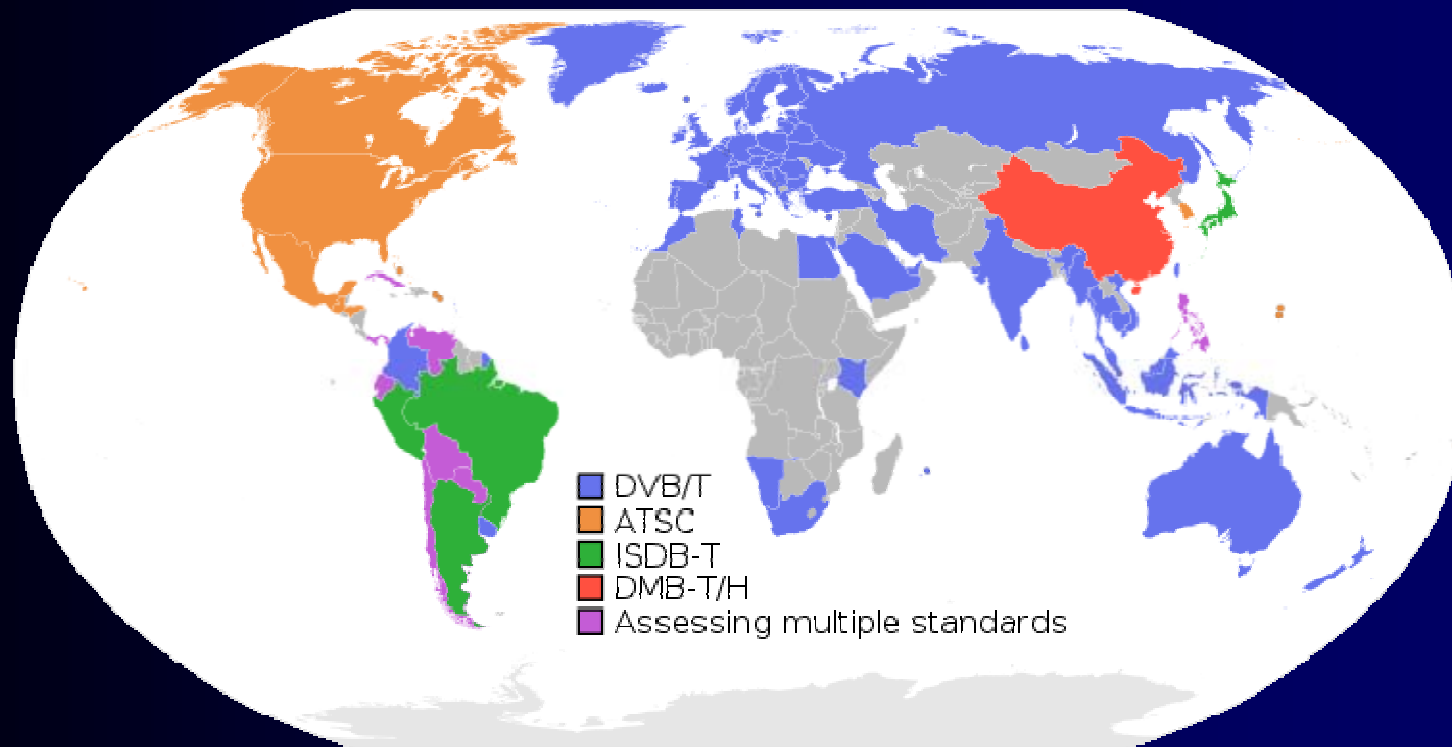
各国のデジタルテレビジョン (1)



From Wikipedia

http://en.wikipedia.org/wiki/File:Digital_broadcast_standards.svg

Digital Television in the World (1)



From Wikipedia

http://en.wikipedia.org/wiki/File:Digital_broadcast_standards.svg

各国のデジタルテレビジョン (2)

区別	名称	地域	ビデオ	オーディオ
地上波	DVB-H	欧州など	MPEG-4 AVC/H.264	MPEG-4 AAC MPEG-4 HE-AAC MPEG-2 HE-AACv2 AMR-WB AC3, Enhanced AC3
	ISDB-T	日本, ブラジル, アルゼンチン, ペルー	MPEG-2 (MP@ML, MP@HL)	MPEG-2 AAC
			MPEG-4 AVC/H.264	MPEG-2 AAC
	T-DMB	韓国	MPEG-4 AVC/H.264	MPEG-4 BSAC MPEG-4 HE-AAC
	DTMB	中国	-	-

Digital Television in the World (2)

Type	Name	Area	Video	Audio
Terrestrial	DVB-H	Europe etc	MPEG-4 AVC/H.264	MPEG-4 AAC MPEG-4 HE-AAC MPEG-2 HE- AACv2 AMR-WB AC3, Enhanced AC3
	ISDB-T	Japan, Brazil, Argentina, Peru	MPEG-2 (MP@ML, MP@HL)	MPEG-2 AAC
			MPEG-4 AVC/H.264	MPEG-2 AAC
	T-DMB	Korea	MPEG-4 AVC/H.264	MPEG-4 BSAC MPEG-4 HE-AAC
	DTMB	China	-	-

各国のデジタルテレビジョン (3)

区別	名称	地域	ビデオ	オーディオ
衛星	S-DMB	韓国	-	-
ハイブリッド	CMMB	中国	AVS	AVS
	DVB-SH	欧州など	MPEG-4 AVC/H.264	MPEG-4 AAC MPEG-4 HE-AAC など
携帯電話網	GPRS, EDGE WCDMA, CDMA EV	各地	-	-
	MBMS, BCMCS MediaFlo	欧州, 日本, 米国	-	-

Digital Television in the World (3)

Type	Name	Area	Video	Audio
Sattellite	S-DMB	Korea	-	-
Hybridツド	CMMB	China	AVS	AVS
	DVB-SH	Europe etc	MPEG-4 AVC/H.264	MPEG-4 AAC MPEG-4 HE-AAC など
Mobile Network	GPRS, EDGE WCDMA, CDMA EV	Many places	-	-
	MBMS, BCMCS MediaFlo	Europe, Japan, USA	-	-

ワンセグ仕様

ビデオ規格	MPEG-4 AVC/H.264 Baseline profile 1, 2
解像度	320x240 画素 (最大)
ビットレート	可変 ex. 180/200/240kbps
フレームレート	15 fps
オーディオ規格	MPEG-2 AAC モノラル, ステレオ, デュアルモノ
オーディオレート	可変 ex. 48/64kbps
データ放送記述	BML Cプロファイル
データ放送レート	約60kbps

One-Segment Specification

Video Coding	MPEG-4 AVC/H.264 Baseline profile 1, 2
Resolution	320x240 pel at the maximum
Bitrate	Variable ex. 180/200/240kbps
Framerate	15 fps
Audio Coding	MPEG-2 AAC mono, stereo, dual-mono
Audio Rate	Variable ex. 48/64kbps
Data Broadcasting	BML C Protocol
Data Broadcasting Rate	About 60kbps

モバイル用ビデオ符号化

- モバイルデジタルテレビジョンのためのビデオ符号化
 - 限られた帯域
 - 高い符号化効率
 - 端末側の消費パワー削減
- MPEG-4 AVC/H.264
 - 高い符号化効率(MPEG-2の2倍)
 - 演算量増加(MPEG-2の10倍以上、デブロッキングフィルタ)
 - 非常に多くの特許で構成されているためpatent poolが必須

Mobile Video Coding

- Video Coding for Mobile Digital Television
 - Limited Bandwidth
 - High Coding Efficiency
 - Low power consumption at mobile device side
- MPEG-4 AVC/H.264
 - Achieve high efficiency (2 times than MPEG-2)
 - High Computational Complexity (10 times than MPEG-2, especially de-blocking filter)
 - Contains a lot of patents, needs patent pool