

## 低遅延高効率 HDTV 伝送のための JPEG2000 符号化制御の最適化

内藤 整<sup>†</sup>      佐野 雄磨<sup>††</sup>      渡辺 裕<sup>††</sup>      小池 淳<sup>†</sup>  
松本 修一<sup>†</sup>

Optimal Design of JPEG2000 Encoder Mechanism for Low Delay and Efficient Distribution of HDTV Programs

Sei NAITO<sup>†</sup>, Yuma SANO<sup>††</sup>, Hiroshi WATANABE<sup>††</sup>, Atsushi KOIKE<sup>†</sup>,  
and Shuichi MATSUMOTO<sup>†</sup>

あらまし 素材伝送クラスの圧縮伝送サービスに対する品質要件として、高画質はもちろんのこと、伝送遅延の低減が強く求められている。MPEG-2 に代表される従来のコーデック技術と比較し、大幅に低遅延でのコーデック処理を実現するためには、動き補償を適用しないイントラ符号化の採用が有効とされる。イントラ符号化方式の中でも、特に JPEG2000 は符号化効率の高さから動画像符号化への応用が有望視されている。しかしながら、JPEG2000 は静止画用圧縮符号化方式であり、同方式を単純に動画像符号化へ適用した場合に得られる符号化性能は十分ではなく、改善の余地が大いに見込まれる。本論文では代表的な映像フォーマットとして HDTV を対象とし、JPEG2000 をベースに低遅延かつ高効率な符号化を実現するための符号化制御技術を提案し、性能評価を行ったので報告する。

キーワード JPEG2000, 最適符号化構造, HDTV 伝送, 低遅延符号化

### 1. ま え が き

映像素材の伝送サービスに対するユーザ要求品質に関して、主に系列放送局間の番組配信や現場からのテレビ中継の用途では、画質はもちろんのこと、伝送遅延の低減が重要視されている。素材伝送クラスの圧縮伝送において、従来より MPEG-2 に準拠したコーデック技術の利用例が多く、その伝送遅延の大半をコーデック処理時間が占めていた。これは MPEG-2 が動き補償と DCT の併用を基本とし、一般的に適用される符号化パラメータでは、符号化効率を優先し、動き予測におけるフレームの並べ換え遅延や、CBR (Constant Bit-Rate) 符号化に適合させるためのバッファリング遅延が大きいことに起因する。

一方で従来技術と比較し大幅に低遅延での動画像

符号化を実現するためには、動き補償を適用せずに、フレームまたはフィールドといった画面内で処理が完結するイントラ符号化技術の採用が有効であると考えられる。イントラ符号化方式として、特に符号化効率の高さから、JPEG2000 を動画像符号化へ応用するためのファイルフォーマットとして、Motion JPEG2000 の国際標準化が既に完了しており [1]、デジタルシネマの規格に採用される見込みである [2]。ここで JPEG2000 は静止画用圧縮符号化方式の国際標準であり、動画像符号化に応用する上での問題として、以下が報告されている。

- 符号量制御として、静止画像用の従来技術である R-D 最適化手法をそのまま適用すると主観画質の低下を招く可能性がある [3]。

- 1080i などインタレース信号に対し、単純にフレーム単位の符号化処理を適用すると、符号化効率の低下を招く可能性がある [4]。

これらの問題解決により符号化性能を改善し、かつ低遅延での符号化を実現する目的から、本論文では代表的な映像フォーマットとして HDTV を対象とし、JPEG2000 をベースに低遅延かつ高効率な符号化を

<sup>†</sup>(株) KDDI 研究所, 上福岡市

KDDI R&D Laboratories Inc., Visual Communications Laboratory, 2-1-15 Ohara, Kamifukuoka-shi, 356-8502 Japan

<sup>††</sup> 早稲田大学大学院国際情報通信研究科, 東京都

Graduate School of Global Information and Telecommunication Studies, Waseda University, 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

現するための符号化制御技術を提案し、性能評価を行う。本論文では、2. で提案方式を検討する上での前提条件につき導入を行う。3. では符号化構造の最適化を実現する要素技術として、符号量制御方式、並びに符号化タイプの適応選択方式について、最適化技術の提案を行う。4. では基本性能評価として、提案要素技術の導入により得られる直接的な効果の検証を行う。5. では一連の要素技術の導入により達成されるコーデック遅延量につき考察を加える。最後に 6. で結論を述べる。

## 2. 検討の前提条件

### 2.1 想定する符号化器

本研究で想定する符号化器のブロック図を図 1 に示す。映像入力としては 1080i の HDTV を想定し、フィールドまたはフレーム単位に JPEG2000 符号化が順次適用される構成である。本論文では JPEG2000 符号化の適用対象となる映像信号の単位を単にピクチャと呼ぶ。動画像符号化方式において、エンコーダ出力が CBR の回線に適合することを裏づけるためには、MPEG-2 Video の VBV (Video Buffer Verifier) に代表されるバッファモデルの規定が不可欠である。バッファモデルの導入による問題として、検証用バッファ (以下、検証バッファと呼ぶ) のバッファサイズに比例してコーデック遅延の増加を招くため、同バッファサイズの低減は符号化遅延を短縮する上で必須の課題であるといえる。

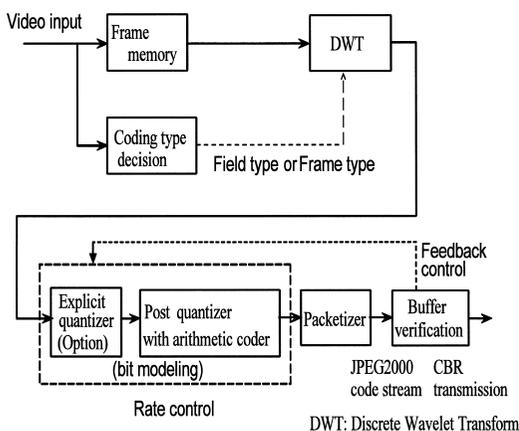


図 1 想定する符号化器のブロック図

Fig. 1 Block diagram of assumed encoder.

## 2.2 従来技術の問題点

### 2.2.1 符号量制御

JPEG2000 による静止画像符号化の符号量制御手法としては、Explicit 量子化を適用せずに、ビット切捨てによるポスト量子化を MSE 最小化規範で行うのが一般的であり、同手法の適用をベースとした動画像符号化の実装例が報告されている [5]。ここで Explicit 量子化とは、図 1 で示すとおりポスト量子化器の前段で適用されるスカラ量子化を指し、規格上はオプション扱いである。符号量制御をポスト量子化の単独適用により行う場合、R-D (Rate-Distortion) 最適化手法であるため MSE は最小となるが、サブバンド内でコードブロックごとに符号化された符号化パス数に開きが生じ、これが 2 のべき乗のオーダで量子化誤差に反映される結果、ひずみ量は空間的に偏った分布となり得る。このため動画像符号化への応用時には、同分布が時間的に変化することで、主観画質を低下させる可能性がある。以下に同弊害を定量的に評価する。まず動画像符号化においてポスト量子化の単独適用により符号量制御を実装した際の、コードブロックごとの切捨てパス数を観測した結果を表 1 に示す。同表はコードブロックごとの切捨てパス数について、時間的な変動度を評価する目的から、同パス数の時系列サンプルより分散値を求め、サブバンドごとに最大値を与えるコードブロックを最悪時として抽出した上で、最悪時の分散値と平均値を示している。符号化条件は後述の表 2 に従い、符号化対象として ITE HDTV 標準動画像の Opening Ceremony より 60 フレーム分を使用した。符号化タイプはフレームタイプとした。簡単のためピクチャごとの割当符号量は同一とした。結果より、ポスト量子化の単独適用に基づく符号量制御を施した

表 1 切捨てパス数の時間的な変動 (最悪時)

Table 1 Temporal fluctuation in the number of truncated coding passes observed under the worst case.

サブバンド	切捨てパス数の統計量	
	分散	平均
3LL	1.00	12.13
3HL	0.59	11.42
3LH	0.07	11.03
3HH	0.74	11.80
2HL	2.10	12.22
2LH	20.50	26.65
2HH	20.33	13.27
1HL	72.06	18.73
1LH	65.99	15.20
1HH	72.10	18.72

場合には、特定のコードブロックにて切捨てパス数の時間的変動が大きくなり、この傾向は高域サブバンドにおいて顕著であることが分かる。

### 2.2.2 インタレース信号の扱い

HDTV に代表されるインタレース信号を JPEG2000 で扱う場合、符号化タイプとしてフィールドタイプとフレームタイプのいずれかを選択可能である。一般に符号化効率の高さからフレームタイプを固定的に割り当てる例が多いが、フレームタイプは再生時刻が 1 フィールド間隔離れた二つのフィールドをマージし、一括して符号化処理を行うため、高速な動きを伴う素材に対しては効率が低下する。

## 3. 提案符号化制御

### 3.1 符号量制御

#### 3.1.1 基本原理

2.1 で挙げた低遅延符号化に向けた課題、並びに 2.2.1 で挙げた問題点を踏まえ、以下の前提に基づき符号量制御手法の検討を行った。

(a) Q-table による Explicit 量子化を基本とする。Q-table の算出式を式 (1) に示す。band, delta[band] はそれぞれ、サブバンド識別子、サブバンドごとの量子化ステップサイズである。cq は各サブバンドに一律で適用される重み係数であり、ピクチャの符号化が完了するつど更新される。dqt は量子化テーブル delta の初期値として参照される値を指す。

$$\text{delta}[\text{band}] = \text{dqt}[\text{band}] \times \text{cq} \quad (1)$$

(b) ポスト量子化において廃棄する符号化パス数はすべてのコードブロックで同一とする。

ここで、(a) にて Explicit 量子化の併用を基本構成としたのは以下の理由による。仮に (b) のみでの制御を考えた場合、ピクチャに対する符号量配分のパターン数は符号化パス数により制限され、符号量制御の観点からはとり得る選択肢が極めて限られており、粗い制御となる点が懸念される。よって、符号量制御において十分な精度を得る上で、Explicit 量子化の併用は必須であるといえる。

#### 3.1.2 処理手順

提案する符号量制御手法について、具体的な処理手順を以下に説明する。

##### (i) 目標符号量の算出

2.1 で述べたとおり、検証バッファサイズの低減により符号化遅延の短縮を図る目的から、検証バッファ

の占有量に関して理想値に対する誤差を吸収する仕組みを取り入れることとし、当該ピクチャ(時刻:  $t$ ) の目標符号量  $R(t)$  を式 (2) により算出する。目標符号量の導出は、MPEG-2 符号化における有効性が確認されているスライディングウィンドウ型の算出モデル [6] に基づき行った。

$$R(t) = \frac{v_0 + (L_W + t) \times B - \left( v_T + \sum_{i=0}^{t-1} S(i) \right)}{L_W} \quad (2)$$

$$B = \frac{\text{bit\_rate}}{\text{picture\_rate}} \quad (3)$$

$L_W$  : 符号量制御のウィンドウサイズ

$v_0$  : 検証バッファにて先頭ピクチャが引き抜かれる直前のバッファ残量

$v_T$  : ピクチャの引抜き直後における検証バッファ残量の目標値

$S(i)$  :  $i$  番目のピクチャにて発生した符号化ビット数

##### (ii) ポスト量子化

すべてのコードブロックで均等に廃棄する符号化パス数を  $np$  とし、当該ピクチャにおいて  $np$  に対する発生符号量  $r(t, np)$  はあらかじめ算出可能であるとする。このとき  $|r(t, np) - R(t)|$  を最小とする  $np$  を全探索により特定する。

##### (iii) Q-table の更新

ピクチャの符号化完了時には、後続ピクチャの符号化に備え、 $cq(t)$  の更新を式 (4)、式 (5) により行う。式 (5) において関数  $f(x)$  は式 (6) に示すとおり、当該ピクチャにおける  $np$  と  $r(t, np)$  の実測結果より、量子化ステップサイズ  $x$  と発生符号量  $y$  の関係を近似するために用いられる。式 (5) において関数  $f(x)$  は式 (6) により定義される関数であり、引数  $x$  はポスト量子化により切り捨てた符号化パス数を量子化ステップサイズに換算した値を指す。同関数は、当該ピクチャにおける  $np$  と  $r(t, np)$  の実測結果をもとに、 $x$  に対する発生符号量  $y$  の近似値を与える。次に式 (4)、式 (5) の導出根拠を示す。まず Explicit 量子化の直後で得られる発生ビット数が、既に目標符号量を下回るケースが発生する点を懸念し、Explicit 量子化の直後では目標値  $R(t)$  に比べて、1 ビットプレーン、つまり三つの符号化パスに相当する余剰ビットが得られることを理想とした。一方で当該ピクチャ  $t$  に対し  $cq(t-1)$  に

よる Explicit 量子化のもと目標符号量  $R(t)$  を達成するためには、 $x = x_T$  相当のポスト量子化が適用される。これに対し上記理想にかなうためには  $x = 2$  を満たすことが必要であるため、 $x_T$  の 2 に対する乖離分を Explicit 量子化の量子化ステップサイズで吸収することとした。

$$cq(t) = cq(t-1) \times \frac{x_T}{2} \quad (4)$$

$$f(x_T) = R(t) \quad (5)$$

$$\begin{cases} y = f(x) \\ x = 2^{\frac{np}{3}}, y = r(t-1, np) \end{cases} \quad (6)$$

### 3.1.3 タイル分割符号化への応用

JPEG2000 では符号化処理、及び復号処理を並列実装する上で有用な符号化ツールとして、タイル分割符号化がサポートされている。タイルとは符号化対象画像を固定サイズの方角エリアに分割したものを指す。符号化ツールとしてタイルを用いる場合、通常の符号化処理でピクチャに対して適用される符号化処理、及び復号処理が、それぞれタイル単位に個別に施される仕組みとなる。

提案した符号量制御手法をタイル分割符号化時にも同様に適用するため、(i) でピクチャに対して得た  $R(t)$  をタイルごとに分配した上で、(ii) のポスト量子化をタイルレベルで独立して行うこととする。 $R(t)$  の分配方法は式 (7)、式 (8) に従うこととする。 $r_j(t, 3)$  は  $r(t, 3)$  のうち  $j$  ( $j = 0, 1, \dots, \text{tile\_cnt} - 1$ ) 番目のタイルに相当する符号量であり、 $\text{tile\_cnt}$  はピクチャを構成するタイルの総数である。タイル化ひずみ抑制の観点から  $cq$  は全タイルで共通の値を使用するという前提のもと (iii) の Q-table 更新はピクチャレベルで行うこととする。

$$R_j(t) = R(t) \times \frac{r_j(t-1, 3)}{r(t-1, 3)} \quad (7)$$

$$r(t, 3) = \sum_{j=0}^{\text{tile\_cnt}-1} r_j(t, 3) \quad (8)$$

### 3.2 シーン適応型符号化タイプ選択

2.2.2 で挙げた問題に対処するアプローチとして、量子化制御や前処理フィルタによる対策に関する検討例がある [4]。これらの検討は、フレーム処理を前提としたアプローチであったが、本論文では符号化タイプ自体を最適値で更新することで更なる改善効果が得られると考え、以下に方式提案を行う。前提として図 1 で示したとおり、入力画像はまずフレーム単位で符号

化タイプ決定部に入力され、解析により、当該フレームに対する最適な符号化タイプが JPEG2000 符号化部に設定されることとする。提案方式において、符号化タイプの決定は式 (9) から式 (12) に従い行われる。 $I(x, y)$  は入力画像の輝度信号成分である。 $(x, y)$  は当該フレーム上の画素位置を指す ( $0 \leq x \leq \text{hsize} - 1$ ,  $0 \leq y \leq \text{vsize} - 1$ )。  $\text{hsize}$ ,  $\text{vsize}$  はそれぞれフレームを構成する水平画素数、垂直ライン数である。フレームは 1 ラインおきに第 1 フィールドと第 2 フィールドがマージされた構成をとる。基本原理として、提案方式は候補となる符号化タイプごとに、垂直方向の相関を推定した上で、より相関の高い一方を選択するアプローチである。ただし式 (11) に示すとおり、両符号化タイプにて発生するサイド情報量の違いを考慮し、重み係数  $a(R)$  ( $a(R) > 1.0$ ) を導入している。 $a(R)$  は符号化レートにより定まる定数である。

一般的な JPEG2000 エンコーダの実装において、符号量制御部では主観画質向上の目的から、Visual weighting と呼ばれる視覚感度を考慮したサブバンド単位の優先制御が導入される。一方でフレームタイプの符号化では、当該フレームを構成するフィールド間の動きが大きくなるにつれライン間相関が低下し、結果として 1LH サブバンドの係数レベルを増大させる [4]。典型例として、図 2 にカメラ操作を伴うテスト画像における DWT (Discrete Wavelet Transform) 係数の分布例を示す。Visual weighting の適用時には、ポスト量子化における R-D 最適化で評価するピクチャのひずみ量は、サブバンドごとのひずみ量に対して、重み付け総和をとったものである。ここで用いられるサブバンドごとの重み係数のうち、視覚的な優先度を反映させる成分は特に CSF (Contrast Sensitivity Function) 係数と呼ばれ、LL サブバンドを 1.0 とした相対値により表現される。通常 CSF 係数は 1.0 以下に設定され、高周波のサブバンドには低めの値が適用される結果、高周波成分を粗く量子化する作用が働く。以上を踏まえ、符号化タイプの選択は 1LH サブバンドに対する重み係数を加味して行うのが妥当であるといえる。提案方式では、係数  $a(R)$  を式 (12) に示す  $a(R, w_v)$  に置き換えることとする。 $w_v$  は 1LH サブバンドの CSF 係数を示し、 $b(R)$  は符号化レートに応じて定まる定数とする。

$$c_l = \sum_{y=0}^{\text{vsize}-3} \sum_{x=0}^{\text{hsize}-1} |I(x, y) - I(x, y+2)| \quad (9)$$

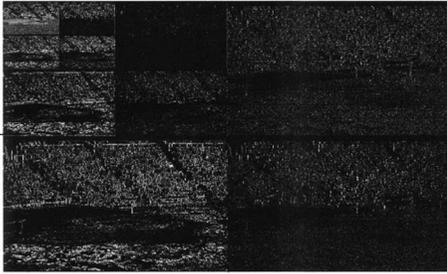


図 2 HDTV フレーム画像に対するサブバンド信号分布の例 (上段: 原画像, 下段: DWT 係数)  
 Fig. 2 Example of sub-band decomposition for interlaced HDTV frame (Top: Original signal, Bottom: DWT coefficients).

$$c_P = \sum_{y=0}^{vsize-3} \sum_{x=0}^{hsize-1} |I(x, y) - I(x, y + 1)| \quad (10)$$

$$(i) \quad c_P < a(R) \times c_I \quad (11)$$

フレーム処理を選択

$$(ii) \quad otherwise$$

フィールド処理を選択

$$a(R, w_v) = b(R) \times \log_2 w_v + a(R) \quad (12)$$

#### 4. 性能評価

3. で述べた提案手法の性能評価を行う目的から符号化シミュレーション実験を行った。本章で使用する共通の符号化条件を表 2 に示す。提案要素技術の符号化シミュレーションツールのソフトウェア実装は、Motion JPEG2000 の参照ソフトである VM version 4.1 [7] をベースに行った。入力には ITE HDTV 標準テスト画像より、以下の 3 種類を使用した。以後、特に言及しない限りタイル分割については非適用とする。

- European market

画面中央を汽車が移動する。カメラ操作はなし。汽車の色調が鮮明であるが精細部は少ない。

表 2 共通の符号化条件  
 Table 2 Condition of coding experiments.

入力画像	1920×1080/59.94 Hz (インタレース) 4 : 2 : 2 コンポーネント (8 ビット)
符号化レート	50 Mbit/s (0.8 bit/pixel)
DWT フィルタ	(9.7) 非可逆フィルタ
サブバンド分割レベル	3 レベル
コードブロックサイズ	64×64

- Whale show

シャチの動きを追う素材。水平方向に高速なパニングを伴う区間あり。観客席に精細部を含む。

- Opening ceremony

入場行進を撮影したもので、カメラ操作及び高速な動きをほとんど含まない。高精細部分が大半を占める。

#### 4.1 提案方式の性能評価

##### 4.1.1 符号量制御

提案符号量制御手法の安定度を評価するため、検証バッファの占有量の時間遷移を観測した結果を図 3 から図 5 に示す。符号化タイプはフレームタイプを用いた。ピクチャ単位の目標符号量の算出に際し、式 (2) の  $L_W$  は 15 に設定した。また理想的なケースに対する占有量の振幅を観測することで安定度の検証が行えると考えた。ここで提案方式の目的は、以下の両立であり、これを理想的に満足するケースとして、ポスト量子化は非適用とし、Explicit 量子化単独の制御により各ピクチャの符号化ビット数が互いに同一に保持される場合 (以下、最小ケースと呼ぶ) を想定した。

- ポスト量子化をコードブロックごとに一様に扱う。

- 所要の検証バッファサイズを低減する。

このため式 (2) における  $v_0, v_T$  は式 (13) に従い最小ケースで適用すべき設定を与えた。一連の図において、縦軸は占有量を 1 ピクチャ当りの平均符号化ビット数で正規化した値をパーセンテージで示す。

$$\begin{cases} v_0 = B \\ v_T = 0 \end{cases} \quad (13)$$

参考のため、1 フレームを 1920 画素 × 270 ラインの 4 タイル構成で符号化した際の結果も併せて示す。図中、タイル分割の非適用時、適用時の結果をそれぞれ、Proposed 1, Proposed 2 と表記する。結果より、検証バッファ占有量の変動分は ±6% に抑えられており、これは本論文で扱う以外の様々なテスト画像に対して行った実験結果においても、同様に成立していることを確認した。図 3 から図 5 で示される検証バッファ

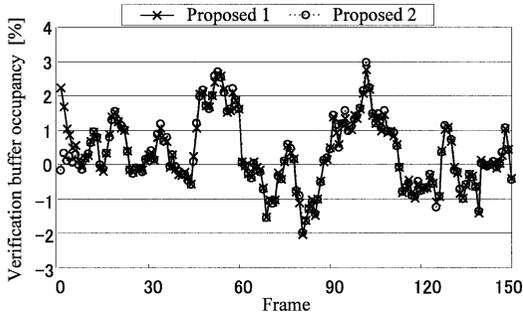


図 3 提案方式による検証バッファ占有量の結果 (European market)

Fig. 3 Verification buffer occupancy result obtained by proposed scheme (European market).

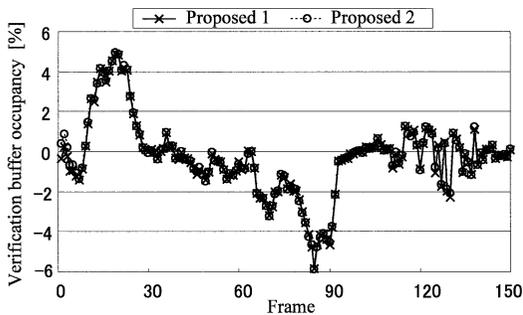


図 4 提案方式による検証バッファ占有量の結果 (Whale show)

Fig. 4 Verification buffer occupancy result obtained by proposed scheme (Whale show).

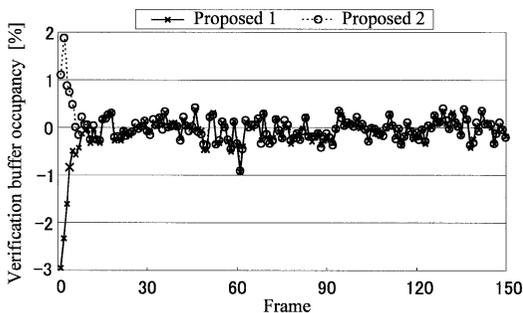


図 5 提案方式による検証バッファ占有量の結果 (Opening ceremony)

Fig. 5 Verification buffer occupancy result obtained by proposed scheme (Opening ceremony).

占有量の変動分のうち、マイナス分はアンダフロー防止のために確保すべき初期遅延量の増分に相当し、変動幅はアンダフロー及びオーバーフローを防止するための所要バッファサイズの増分に相当する。この点から提案方式の導入による所要バッファサイズの増分は、最

小ケースに比してたかだか 12%であると見込まれる。ただし最小ケースの符号化は、各ピクチャの Q-table として、全ピクチャに共通の目標符号量を達成するものを事前に探索することで実現したものであり、1パスの符号化システムにて実現することは困難である。ここでいう 1パスの符号化とは、各ピクチャの符号化にて、最適な符号化制御を決定するための予備的な符号化処理を伴わないことを意味する。また提案手法を用いない場合の符号量変動例として、Explicit 量子化において Q-table を固定的に与え、ポスト量子化を非適用とした場合の結果を測定した。結果、最小ケースに対する所要バッファサイズは、European market, Whale show, Opening ceremony に対してそれぞれ、6.3 倍、9.1 倍、1.2 倍であることを確認した。Q-table の固定値は先頭ピクチャの符号量が  $B$  に一致するように選定し、検証バッファに入力される符号化レートは総符号量をもとに算出した。以上より、提案した符号量制御手法はコーデック遅延を短縮する上で有効であるといえる。

以上の符号化実験において、符号化効率の評価のために取得した提案方式による PSNR のシーケンス平均を表 3 にまとめる。参考のため、上述の最小ケース、MSE 最小化による結果も併せて示す (同表にてそれぞれ、「最小ケース」、「MSE 最小」と表記)。結果より提案手法の適用による符号化効率の低下は少なく、最小ケースに対する PSNR の低下は無視できるレベルであることを確認した。

また提案方式により符号量制御を実装した際の切捨てパス数について、時間方向の分散値と平均値を観測した結果を表 4 に示す。提案方式の導入による効果を純粋に評価するため、符号化条件及び入力画像は表 1 の実験と全く同一に設定した。参考のため、1 フレームを 1920 画素  $\times$  270 ラインの 4 タイル構成で符号化した際の切捨てパス数について、各タイル別に集計した統計量を併せて示す。タイル番号は上端のタイルを 1 とし、下方向に順次連番を与えた。提案方式の原理上、サブバンドごとの結果はいずれも表中に記載の値と同一である。表 1 に対する比較結果より、提案手法の適用による効果として、3.1.1 で指摘のポスト量子化に起因する問題が解消されていることが分かる。更に主観品質の改善効果を確認する目的から主観評価試験を行った。同試験における評価条件は ITU-R 勧告 BT.500 に従い、画質評価には二重刺激連続品質尺度 (DSCQS) を用いた。モニタは 24 インチ CRT、被験

表 3 提案方式による PSNR の結果  
Table 3 PSNR result obtained by proposed scheme.

素材	Proposed 1	Proposed 2	最小ケース	MSE 最小
European market	38.26	38.21	38.26	38.48
Whale show	33.05	32.98	33.05	33.18
Opening ceremony	29.23	29.17	29.22	29.44

表 4 切捨てパス数の時間的な変動 (提案手法)  
Table 4 Temporal fluctuation in the number of truncated coding passes obtained by proposed scheme.

タイル分割	タイル番号	切捨てパス数の統計量	
		分散	平均
なし	—	0.15	3.05
4 分割	1	0.02	3.02
	2	0.27	3.07
	3	0.27	3.07
	4	0.15	3.05

表 5 主観画質評価試験結果

Table 5 Result of subjective picture quality evaluation.

素材	品質差平均 %		
	Proposed 1	Proposed 2	MSE 最小
European market	7.24	7.42	7.58
Whale show	10.71	10.88	12.68
Opening ceremony	16.02	16.38	22.63

者は非専門家 15 名とした。従来方式として、ポスト量子化の単独適用方式 (表 1 に相当) についても評価対象に加えた。表 5 に評価試験結果として、原画像に対する品質差平均を示す。結果より、いずれの素材においても提案方式により主観画質が改善しているといえる。Opening ceremony における効果が顕著である理由は、他素材に比べて静止部分が多く、特に芝生部分では、従来方式にて局所的な画質変動が目立っているのに対し、提案方式ではこれを抑制できていることによる。

#### 4.1.2 シーン適応型符号化タイプ選択

提案した符号化タイプ選択方式による PSNR の結果を図 6 から図 8 に示す。適応選択による符号化ゲインを評価する目的から、符号化タイプとしてフレームタイプ、フィールドタイプをそれぞれ固定的に割り当てた場合の結果を併せて示す。また一連の図は、Visual weighting を非適用とした場合、適用した場合の二つの結果を含んでいる。Visual weighting を適用する場合のサブバンドごとの CSF 係数は、JPEG2000 IS annex J [8] に従う値を用いた。事前実験の結果より、式 (11) 及び式 (12) における  $a(R)$ 、 $b(R)$  はそれぞれ、1.27、 $-0.12$  に設定した。

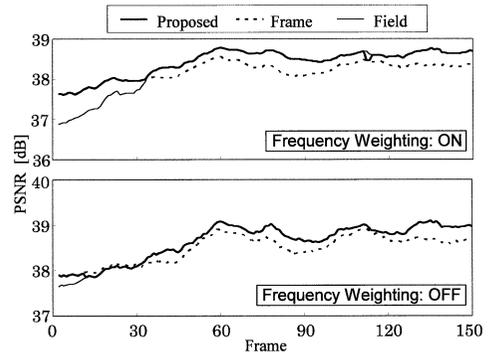


図 6 提案方式による PSNR の結果 (European market)  
Fig. 6 PSNR result obtained by proposed scheme (European market).

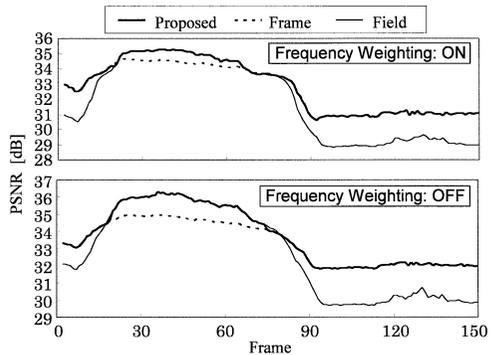


図 7 提案方式による PSNR の結果 (Whale show)  
Fig. 7 PSNR result obtained by proposed scheme (Whale show).

結果より提案方式の導入による画質劣化は特に発生していないことが分かる。Opening ceremony については、素材の性質上、性能でフレーム処理がフィールド処理を常に上回るため、提案方式とフレームタイプ固定方式の曲線が一致する結果となった。結論として提案方式の導入により、符号化タイプとしてより効率の高い一方が高い精度で選択される結果、従来型の固定割当方式に対する PSNR ゲインが確認できる。特にフレームタイプを固定的に適用したケースに対する PSNR ゲインは、フレームタイプの符号化において高速移動など、フィールド間変化が激しい場合に起こり

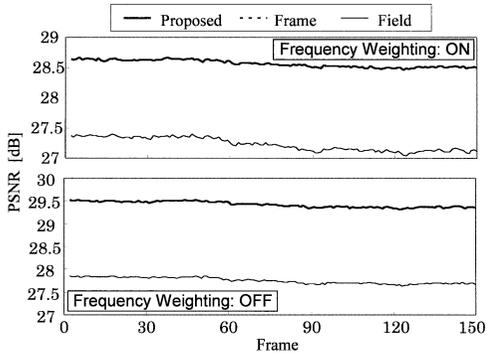


図 8 提案方式による PSNR の結果 (Opening ceremony)  
 Fig. 8 PSNR result obtained by proposed scheme (Opening ceremony).

得る 1LH サブバンド係数の増大に対応した適応制御が有効に働いていることを示すものである。

### 5. コーデック遅延量の試算

本論文で提案した要素技術の組合せに基づくエンコーダ及びデコーダのハードウェア実装を想定し、見込まれるコーデック遅延量を試算する。まずコーデック遅延は主に以下の処理時間の積み重ねにより発生する。

- 符号化単位 (画面またはタイル) の入力待ち時間 (エンコーダ)
- 符号化処理時間 (エンコーダ)
- ストリーム送待待ち時間 (エンコーダ)
- 復号開始待ち時間 (デコーダ)
- 復号処理時間 (デコーダ)

ここで提案した要素技術の組合せとコーデック遅延量の関係を表 6 に示す。遅延時間の単位は HDTV のフィールド周期 (約 16.68 ms) とした。先に列挙した遅延項目のうち、符号化単位ごとに施される符号化処理、及び復号処理は一般に専用 LSI により行われることを想定すると、その処理時間は無視できるため同表から除外した。またストリーム送待待ち時間と復号開始待ち時間の合計は、検証バッファサイズを符号化レートで除した値により表せる点から (導出根拠を付録に示す)、これら 2 項目の合算値を送受信バッファの欄に記載した。符号量制御に関して、提案手法を導入する際に確保すべき送受信バッファサイズの算出は、4.1.1 で得られた知見より最小ケースと比して 12%の増分を見込むという前提をもとに行った。

表 6 提案手法によるコーデック遅延時間  
 Table 6 Codec delay accompanied by proposed scheme.

符号化単位	符号量制御	遅延時間フィールド		
		符号化単位の 入力待ち時間	送受信 バッファ	合計
フィールド固定	提案	1.00	1.12	2.12
	最小ケース	1.00	1.00	2.00
フレーム固定	提案	2.00	2.24	4.24
	最小ケース	2.00	2.00	4.00
提案	提案	2.00	2.24	4.24
	最小ケース	2.00	2.00	4.00

表 6 より符号化単位について、提案手法を適用した際の遅延量は、従来手法として位置づけられるフレーム処理の固定適用時と同レベルであることが分かる。また符号量制御について、いずれの符号化単位を適用した条件においても、提案手法を適用するために見込むべき遅延量は、達成される再生品質及び遅延量に関し理想的なケースといえる最小ケースに比してたかだか 6%増である。参考までに従来手法として位置づけられるポスト量子化の単独適用に関し、毎ピクチャの符号量が同一となるよう制限を加えた場合の遅延量は最小ケースに一致するため、同方式と比較した提案手法の遅延量増加分も全く同レベルであるといえる。

以上の試算結果より明らかなどおり、低遅延性と符号化性能はトレードオフの関係にあり、符号化単位並びに符号量制御の組合せは、ユーザ要求に見合うものを選択すべきである。更なる低遅延化の要素として、符号化単位の入力待ち時間については、JPEG2000 でサポートされているタイルやプレシントといった、画面分割に基づく符号化ファンクションを効果的に活用することで、更なる改善が図れるものと予想される [9]。

### 6. む す び

HDTV の素材伝送を対象とし、JPEG2000 の枠組みの中で低遅延かつ高効率な圧縮符号化を実現するための要素技術として、符号量制御の最適化、符号化タイプの適応選択に着目し、提案方式の導入を行った。性能評価として、符号化シミュレーション結果より、提案技術の導入による効果を定量的に検証した上で、想定されるコーデック遅延量について考察を加えた。

HDTV を対象とした低遅延及び高圧縮を両立する圧縮符号化技術の確立は、FPU やワイヤレスカメラといった主に放送事業者をユーザとする素材伝送用途に加え、コンシューマ向けブロードバンド回線を介し

た HDTV テレカンファレンスなど、幅広い用途での需要が見込まれる。本研究に関連する要素技術の更なる性能改善の積み重ねを通じ、インタラクティブ性に優れた映像伝送サービス普及の一助としたい。

謝辞 日ごろ御指導頂く株式会社 KDDI 研究所代表取締役所長浅見徹氏に感謝致します。

文 献

- [1] ISO/IEC JTC1/SC29/WG1 N2250, "Motion JPEG-2000 final draft international standard 1.0," Sept. 2001.
- [2] Digital Cinema Initiatives, "Digital Cinema System Specification version 4.3," Dec. 2004.
- [3] S. Naito, A. Koike, and S. Matsumoto, "Optimal JPEG2000 encoder mechanism for low delay and efficient distribution of HDTV programs," Proc. ICASSP2004, vol.V, pp.81-84, May 2004.
- [4] T. Kuge, "Wavelet picture coding and its several problems of the application to the interlace HDTV and the ultra-high definition images," ICIP2002, vol.III, pp.217-220, Sept. 2002.
- [5] W. Chan and A. Becker, "Efficient rate control for motion JPEG2000," Proc. Data Compression Conference (DCC2004), p.529, March 2004.
- [6] 内藤 整, 小池 淳, 和田正裕, 松本修一, 羽鳥好律, "ひずみ最小レート制御の高度利用に基づく MPEG-2 ビット配分の最適化," 信学論 (D-II), vol.J86-D-II, no.11, pp.1565-1574, Nov. 2003.
- [7] ISO/IEC JTC1/SC29 WG1 N2144, "Motion JPEG-2000 verification model ver.4.1," June 2001.
- [8] ISO/IEC JTC1/SC29/WG1 N1890, "JPEG2000 part I Final draft international standard," Sept. 2000.
- [9] 佐野雄磨, 内藤 整, 渡辺 裕, "JPEG2000 Precinct 構造を用いた低遅延 HDTV 符号化方式," 情処学研報, 2004-AVM-46, Oct. 2004.

付 録

バッファリング遅延量の導出

以下に、バッファモデル導入時における送信側及び受信側で発生するバッファリング遅延量の総和を推定する。受信側において、検証バッファのサイズ、検証バッファより再生時刻が  $t (t = 0, 1, \dots)$  であるピクチャが引き抜かれた直後の占有量をそれぞれ、 $L_v$ 、 $v(t)$  とするとき式 (A.1)、式 (A.2) が成立する。 $v_0$ 、 $B$ 、 $S(i)$  の定義は本文中 3.1.2 に従うこととする。

$$v(t) = v_0 + t \times B - \sum_{i=0}^t S(i) \quad (A.1)$$

$$0 \leq v(t) \leq L_v \quad (A.2)$$

検証バッファとは一般に、受信側に実装される仮想

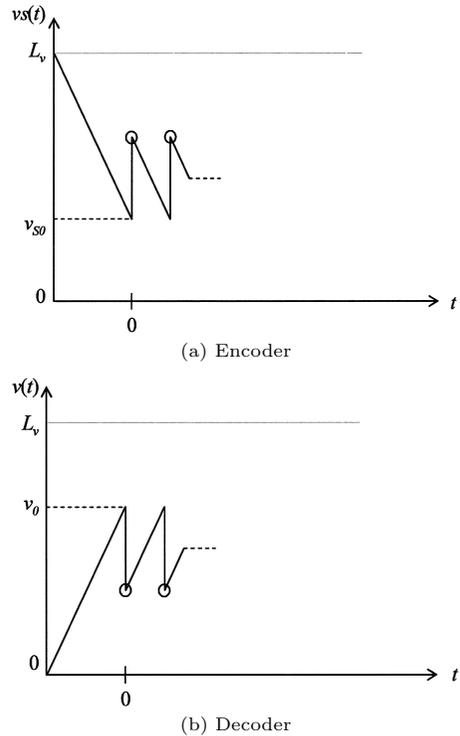


図 A.1  $v(t)$  と  $vs(t)$  の遷移  
Fig. A.1 Transition of  $v(t)$  and  $vs(t)$ .

的なバッファを意味するが、送信側及び受信側で発生するバッファリング遅延量の総和を導出する目的から、以下に送信側の検証バッファの振舞いを検証する。まず受信側の検証バッファと比較して、入力と出力の関係が逆になる点から、バッファサイズは受信側と同一であるとすると、検証バッファにピクチャ  $t (t = 0, 1, \dots)$  が書き込まれた直後の占有量  $vs(t)$  は式 (A.3) により算出される。 $v_{s0}$  は先頭ピクチャの書込みが行われる直前の占有量である。実際には  $v_{s0}$  に相当するデータは符号化ストリームに関係しないため、実装上は  $v_{s0}$  のデータ送信にかかる時間だけ送信待ちを入れるなどの処理が想定される。参考のため、 $v(t)$  と  $vs(t)$  の遷移イメージを図 A.1 に示す。図中、実線はバッファ占有量の遷移を示し、丸印は  $v(t)$  と  $vs(t)$  のサンプルを示す。検証バッファの破綻防止のため、式 (A.4) に示す条件が常に満足される必要があるが、これは  $v_{s0}$  として式 (A.5) に従う値を選択することで実現できる。よって、送信側でのストリーム送待時間、及び受信側での復号開始待ち時間の合計を  $T_w$  とおくと、同値は式 (A.6) に示すとおり、検証バッファサイズを符

号化レートで除した値により表せる．

$$vs(t) = vs_0 + \sum_{i=0}^{t-1} S(i) - t \times B \quad (\text{A}\cdot3)$$

$$0 \leq vs(t) \leq L_v \quad (\text{A}\cdot4)$$

$$vs_0 = L_v - v_0 \quad (\text{A}\cdot5)$$

$$\begin{aligned} T_w &= \frac{v_0 + vs_0}{B} \\ &= \frac{L_v}{B} \end{aligned} \quad (\text{A}\cdot6)$$

(平成 17 年 2 月 7 日受付, 6 月 27 日再受付)



内藤 整 (正員)

平 6 早大・理工・通信卒．平 8 同大学院修士課程了．同年国際電信電話(株)入社．以来，同社研究所にて主に動画像符号化の研究に従事．現在，(株)KDDI 研究所映像通信グループ研究主査．平 11 年度学術奨励賞受賞．



佐野 雄磨 (学生員)

平 16 早大・理工・通信卒．同年同大学院国際情報通信研究科修士課程入学，現在に至る．



渡辺 裕 (正員)

昭 55 北大・工・電子卒．昭 60 同大学院博士課程修了．同年 NTT 入社．NTT 研究所において画像の高効率符号化に関する研究及び動画像通信システムの開発に従事．平 12 早稲田大学国際情報通信研究センター教授．平 14 早稲田大学大学院国際情報通信研究科教授．マルチメディア符号化，処理及び配信に関する研究に興味をもつ．平成 11 年より ISO/IEC JTC 1/SC 29 議長．工博．IEEE，情報処理学会，映像情報メディア学会，日本音響学会各会員．



小池 淳 (正員)

昭 58 新潟大・工・情報卒．昭 60 東北大学院修士課程了．同年国際電信電話(株)入社．以来，テレビ会議・電話システム，動画像の高効率符号化方式の研究・開発に従事．現在，(株)KDDI 研究所映像通信グループリーダー．工博．



松本 修一 (正員)

昭 52 北大・工・電子卒．同 54 同大学院修士課程了．同年国際電信電話(株)入社．以来，同社研究所にて HDTV，SDTV の高効率符号化方式の研究・開発に従事．現在，(株)KDDI 研究所取締役．工博．昭 58 年度本会論文賞，昭 61 年度篠原記念学術奨励賞受賞，平 11 年度本会業績賞受賞．