

独立階層による空間スケーラブル符号化方式に関する検討

筑波 健史[†] 永吉 功^{†,††} 花村 剛^{††,†} 富永 英義^{†††,††††}

[†] 早稲田大学大学院 国際情報通信研究科 〒169-0051 東京都新宿区西早稲田 1-3-10

^{††} 株式会社 メディアグルー 〒169-0072 東京都新宿区大久保 2-4-12 新宿ラムダックスビル 8 階

^{†††} 早稲田大学 国際情報通信研究センター 〒169-0051 東京都新宿区西早稲田 1-3-10

^{††††} 早稲田大学 理工学部 コンピュータ・ネットワーク工学科 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: †{tsukuba,isao,hana,tominaga}@tom.comm.waseda.ac.jp

あらまし 本稿では、階層間の依存関係を排除したスケーラブル符号化方式を検討する。まず、従来のスケーラブル符号化方式における階層間の依存関係による課題点を指摘し、符号化方法に求められる基本要件を示す。次に、上記課題を解決するため、空間領域において入力映像を等分割することにより階層間の依存関係を排除し、任意の階層の組み合わせによる映像再生を実現するスケーラブル符号化方式について述べる。最後に評価実験により提案方式において、画素分割の有効性、位相のずれの予測に動き補償を適用することの有効性を確認した。

キーワード 階層間独立性, スケーラブル符号化, トランスコーダ, MPEG

A Study on Spatial Scalable Coding Architecture with Independent Layers

Takeshi TSUKUBA[†], Isao NAGAYOSHI^{†,††}, Tsuyoshi HANAMURA^{††,†}, and Hideyoshi
TOMINAGA^{†††,††††}

[†] GITS, WASEDA University, 29-7 building 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo 169-0051 Japan

^{††} Media Glue Corp. Ramdax Bldg. 8th floor, 2-4-12 Ohkubo, Shinjuku-ku, Tokyo, 169-0072 Japan

^{†††} GITI Waseda University 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

^{††††} Dept.of Computer Science. Eng., WASEDA Univ.,3-4-1 Ohkubo, Shinjuku-ku, Tokyo, 169-8555 Japan

E-mail: †{tsukuba,isao,hana,tominaga}@tom.comm.waseda.ac.jp

Abstract In this paper, we describe a study on a spatial scalable video coding architecture with independent layers. First, we discuss about scalable video transmission by traditional scalable video coding architecture. Second, we explain the proposed method. It splits an input image into some equal parts in a space domain, and converts splitted images to independent layer bitstreams. At last, from the simulation experiments, in the proposed scheme, we show the validity of pixel division and that of the application of a motion compensation to prediction of the difference in a phase.

Key words Layer Independency, Scalable Video Coding, Transcoder, MPEG

1. はじめに

ネットワークを介して MPEG などの国際標準符号化方式を利用した動画映像配信サービスを行う場合、再生環境や伝送条件が異なるため様々な品質の映像が要求される。これらの要求を満たすために、1つのソースから異なる利用環境に応じてさまざまな品質を選択的に提供するスケーラブル映像配信システムが検討されている [1]。これらのシステムを実現する技術として、スケーラブル符号化方式 [2] [3] が検討されている。これらの方式は、基本映像品質を与える基本階層と基本品質を向上させる拡張階層を生成し伝送することでスケーラビリティを実現する。一方、階層構造を生成する場合、符号化器・復号器の演算量の増大、ネットワークルータなどにおける伝送優先制御の複雑性が課題であることが知られている。

そこで本稿では、従来のスケーラブル符号化方式における階層間の依存関係による課題を指摘し、スケーラブル符号化方式における基本要件を整理する。また、拡張階層間での依存関係を排除することにより、上記課題を解決するスケーラブル符号化方式について検討し、評価実験により有効性を示す。

2. 階層間の依存関係による課題

階層間の依存関係による課題を以下に示す。

- 階層間の依存関係により下位階層から順に優先伝送する必要性
- 階層数の増加に伴うサーバ、ネットワーク負荷の増大
- 中間階層損失時における損失階層以上の階層符号の復号不可能性

上記課題は、伝送するストリームの情報構造に起因すると考えられる [4]。スケーラブル映像配信を実現する技術として、MPEG-2 では SNR スケーラブル符号化方式が標準化され、MPEG-4 では画像品質の選択に柔軟性を実現した FGS (Fine Granularity Scalability) [5] が標準化された。また、多階層符号化トランスコーディング技術 [6] などが現在検討されている。しかし、いずれの方式においても階層間に依存関係が存在し、図 1 のように中間階層を損失した場合に、受信符号量に相当する映像品質は提供されない。このような中間階層損失時における耐性の低減を回避するためには複雑な伝送優先制御が必要となる。

そこで、上記課題を解決するために階層間の依存関係を排除した情報構造を生成することによりスケーラビリティを実現する符号化方式を検討する必要がある。本方式を独立階層による空間スケーラブル符号化方式と定義し、基本要件を以下にまとめる。

- 任意の階層の組み合わせによる映像再生が可能
- 符号化効率の向上

3. 独立階層による空間スケーラブル符号化方式

本節では、上記基本要件を満たす具体的な符号化処理を述べる。本方式は、原画像を空間領域において、等分割することにより階層間の依存関係を排除した階層構造を有したスケーラブル符号化方式である。

3.1 基本原理

単独階層での映像再生を実現し、任意の階層の組み合わせから映像再生を可能とするため、各階層は、映像再生に最低限必要となる基本情報と基本情報の品質を向上させる拡張階層から構成される。基本情報はすべての階層が共通にもち、拡張階層

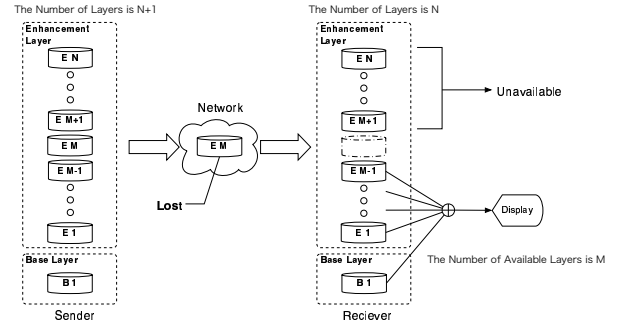


図 1 従来のスケーラブル符号化方式

はそれぞれ互いに独立した情報である。

3.2 処理の流れ

本方式では、符号化側において、原画像から、分離器 (Down-sampler) により N 個の分割画像を生成する (STEP1)。分割画像の 1 つを選択し、低品質に符号化し基本階層 *Base* とする (STEP2)。また N 個の各分割画像と基本階層の復号映像との差分信号を符号化し、拡張階層 *Enh(n)* とする (STEP3)。多重化器 (MUX) において、*Base* を複製し、それぞれを *Enh(n)* と多重化して N 本のストリーム $St(n)$ を生成する。

復号側では、分離器において $St(n)$ をそれぞれ、*Base* と *Enh(n)* に分割し、復号器において *Base* の復号信号と *Enh(n)* の差分信号を合成することにより映像再生を行う。

利用環境の帯域に合わせて m 本 ($m \leq N$) のストリームを配信し、ストリームを復号、合成することにより様々な品質の映像を実現する。

3.3 符号化処理方式

本節では、本方式の符号化器の構成を図 2 に示し、*Base* と *Enh(n)* の具体的な生成処理方法について述べる。符号化器の処理は以下に示す 3 つの Step から構成される。

Step1: 分割画像の生成

Step2: 基本階層 *Base* の生成

Step3: 拡張階層 *Enh(n)* の生成

図 2 において、Step1 では、原画像をベースバンド上において、標本点を変更して 2:1 ダウンサンプリングを行うことにより、階層信号の元となる 4 種類の分割画像を生成する。Step2 では、分割画像の 1 つを選択し、DCT 変換の後、粗く量子化後符号化し、基本階層 *Base* を生成する。Step3 では、画素値まで復号した基本階層と 4 つの分割画像との差分をとり、各差分信号を DCT 変換の後、細かく量子化し符号化することで拡張階層 *Enh(n)* を生成する。

3.3.1 分割画像の生成

本節では、分割画像を生成する手法とその課題点を述べる。原画像 $f(x, y)$ から分割画像 $g_i(x, y)$ ($i = 1, 2, 3, 4$) を生成する方法は、図 3 に示すように、水平方向 (x 軸方向)、垂直方向 (y 軸方向) に関して、原画像 $f(x, y)$ の画素の座標 (x, y) の奇数・偶数により場合分けして間引く 4 種類のサンプリング方法により行う。つまり、分割画像 $g_i(x, y)$ ($i = 1, 2, 3, 4$) は、次式で表される。

$$g_1(x, y) = f(2x, 2y) \quad (1)$$

$$g_2(x, y) = f(2x + 1, 2y) \quad (2)$$

$$g_3(x, y) = f(2x, 2y + 1) \quad (3)$$

$$g_4(x, y) = f(2x + 1, 2y + 1) \quad (4)$$

ここで、 $x = 0, 1, 2, \dots, N - 1, y = 0, 1, 2, \dots, M - 1$ である。

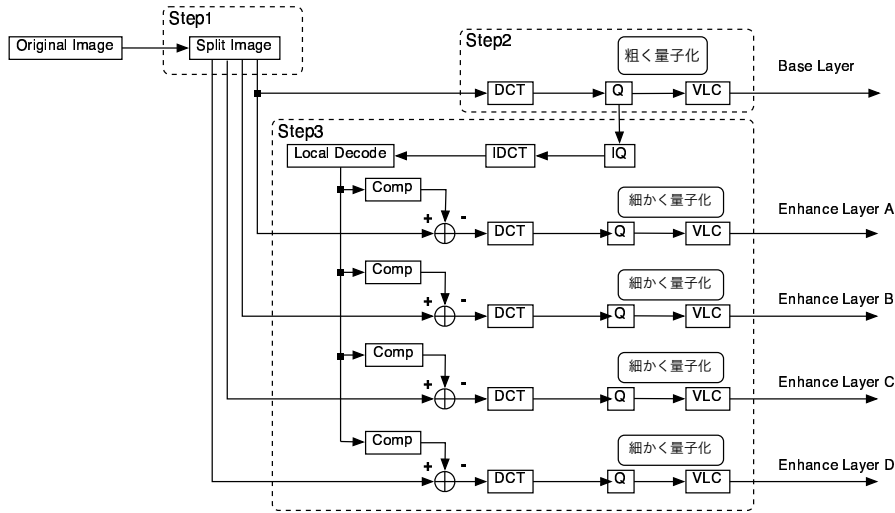


図2 符号化器の構成

画像のダウンサンプリングの課題点として、単純に周波数を半分にしたのでは、エイリアシングが発生し、視覚的に画質が劣化してしまうことが挙げられる。そこで、エイリアシングを防ぐために、ローパスフィルタにより高周波成分をカットしてサンプリングを行う方法についても検討する。特にインターレース画像において、垂直方向 (y 軸方向) に関して 2:1 ダウンサンプリングすることは、トップフィールドとボトムフィールドに分けることと等価であり、フィールド間では 1/60 秒差があるので、フィルタ処理は行わない。水平方向 (x 軸方向) に関してのみ、フィルタ処理を行う。以下ではフィルタ処理による画素分割の手法の述べる。

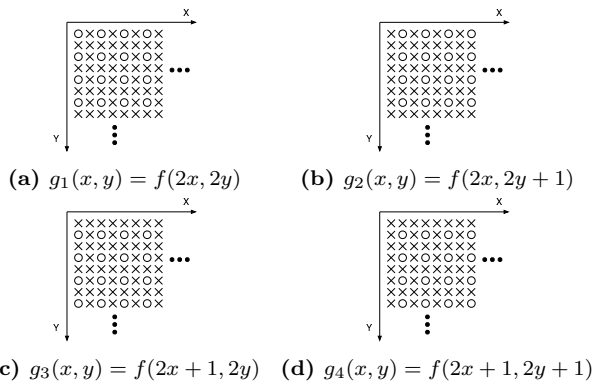


図3 サンプリング方法

フィルタ処理による分割方法

分割画像の x 座標がとる値の範囲を $0 \leq N - 1$ とすると、 $x = 0, N - 1$ においては、 $g_1(x, y), g_3(x, y)$ に対しては、 $\{1, 1\}/2$ 、また $g_2(x, y), g_4(x, y)$ に対しては、 $\{1, 2, 1\}/4$ の重み付け平均フィルタによりダウンサンプリングを行う。

$x = 1, N - 2$ においては、 $g_i(x, y), i = 1, 2, 3, 4$ に対して、 $\{1, 2, 1\}/4$ の重み付け平均フィルタによりダウンサンプリングを行う。

$2 \leq x \leq N - 3$ においては、 $g_i(x, y), i = 1, 2, 3, 4$ に対して、 $\{-29, 88, 138, 88, -29\}/256$ の重み付け平均フィルタによりダウンサンプリングを行う。ここで、 $//$ は小数点の丸込め操作を含む割り算である。

3.3.2 拡張階層の生成

本節では、拡張階層となる差分信号を生成する手法とその課題点を述べる。分割画像間の相関が強いことを利用して、基本階層となる低品質な画像と、各分割画像とのベースバンド上での差分信号を拡張階層とする。

基本階層の復号画像から分割画像を見ると、拡張階層の信号は、隣接画像間の差分となるため、分割画像毎で位相が異なるなどの信号特性において差異がある。この信号特性の差異を補償することにより、符号化時における拡張階層の差分信号の発生符号量を低減し、かつ画質を向上させることができると考えられる。そこで本方式では、信号特性の差異を補償するために、マクロブロック毎に予測誤差信号の絶対値平均誤差 (MAE) が最小となるような参照部位をシフトさせて差分信号を求めて符号化する。

4. 実験

本節では第3節において提案した独立階層による空間スケラブル符号化方式を実験により評価する。特に以下の点に着目する。

- (1) 非分割、分割における信号特性
- (2) 拡張階層における非補償、補償操作による差分信号の特性
- (3) 非分割、分割・階層構造における符号化効率について実験をおこなった。また、共通の実験条件を表1に示した。

表1 共通の実験条件

テストシーケンス	Bus, Table Tennis
画像フォーマット	ITU-R BT.601 4:2:0 Format 輝度信号 704[pe] × 480[line] 色差信号 352[pe] × 240[line]
符号化フレーム数	150[Frame]

4.1 画素分割の信号特性

4.1.1 実験内容と実験条件

画素分割による信号特性の変化を確認するため、分割しない場合 (方式1)、フィルタなしの分割 (方法2) とフィルタによる分割 (方法3) について、 8×8 ブロック内の輝度成分の DCT 係数の標準偏差を求め、結果を図5に示す。また図4のよう

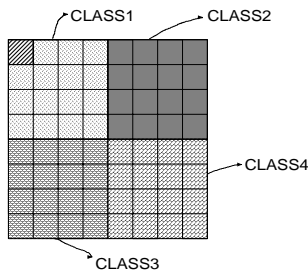


図4 DCT係数のクラス分け

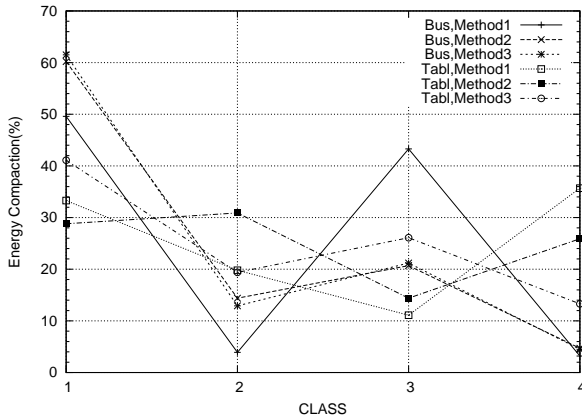


図6 エネルギー比率

にDCT係数のクラス分けを行い、各クラスに含まれるエネルギー比率を求め、結果を図6。それぞれクラスは、低周波数成分の領域 (CLASS1)、水平方向の周波数成分を多く含んだ領域 (CLASS2)、垂直方向の周波数成分 (CLASS3)、および高周波数成分 (CLASS4) である。

4.1.2 結果と考察

図5より、DCT係数の標準偏差から、“Bus”と“Table Tennis”の両方も方式2、方式3は、方式1と比較して、水平方向の周波数成分を含む領域に係数の分布が増加し、垂直方向の周波数成分を含む領域の係数の分布は減少する変化を確認できる。

図6より、エネルギー比率の特性からは、“Bus”では、方式2、方式3は、方式1と比較して、CLASS1, CLASS2のエネルギー比率が増加し、CLASS3のエネルギー比率が減少し、CLASS4のエネルギー比率がわずかに増加することが確認できる。一方、“Table Tennis”では、方式2、方式3は、方式1と比較して、CLASS1が増加し、CLASS4が減少しており、方式3の方が変化は顕著である。しかし、方式2はCLASS2が増加するのに対して、方式3はCLASS3が増加する差異が確認できる。

以上から、垂直方向に2:1にダウンサンプリングすることは、インターレース画像である“bus”のような水平方向の相関が強く垂直方向の相関が弱い横動きが多い映像において、フィールド毎に分けることと等価となる。一方、同じインターレース画像でも“Table Tennis”のようにズームやシーンチェンジが激しい映像では、垂直方向、水平方向の両方に関して相関の変化が激しく、高周波成分を多く含むため符号化効率が低下すると予測される。また水平方向に2:1にダウンサンプリングすることは、水平方向の相関が弱まり、高周波成分の増加することになる。つまり、インターレース画像に関しては、画素分割を

することで、垂直方向の周波数のエネルギーを減少させ、より低周波部分にエネルギーを集中させることが確認できる。同時に、CLASS3における水平方向の周波数成分が増加することが確認できる。しかし、方式3のように、水平方向に対して、ローパスフィルタを使用することにより、水平方向の周波数成分のエネルギーの増加を抑えることができ、より低周波成分へエネルギーを集中させることができ、有効性を確認できる。

4.2 拡張階層の信号特性

4.2.1 実験内容と実験条件

基本階層の復号画像と分割画像との差分信号の特性を確認するために、

Method1 動き補償なし・フィルタなし

Method2 動き補償なし・フィルタあり

Method3 動き補償あり・フィルタなし

Method4 動き補償あり・フィルタあり

の場合について、 8×8 のブロック内の差分信号を輝度成分のDCT係数の標準偏差を求め、結果を図7に示す。標準偏差の値は、各差分信号の標準偏差の平均をとっている。また、実験で用いた動き補償の方法を表2に示す。ただし、基本階層は量子化係数 $MQ = 10$ として固定量子化した。

表2 動き補償の方法

探索範囲	$\pm 15[\text{pixel}]$
探索方法	全探索
演算方法	MAE (絶対値平均誤差)

4.2.2 結果と考察

図7より、“Bus”について、動き補償を行わない場合、DCT係数の分散特性はDC成分とその近傍のAC成分に集中しているが、動き補償を行った場合、DCT係数の分布が平坦であることが確認できる。また、“Table Tennis”について、動き補償を行わない場合、DCT係数の分布は主に水平方向の周波数を多く含む領域、CLASS2とCLASS4の中間領域に集中し、動き補償を行った場合、“Bus”と同様にDCT係数の分布が平坦になることが確認できる。

以上から、本方式では拡張階層の差分信号を求める場合、動き補償を行うことで、画像によらずブロック内のDCT係数の分散特性をほぼ一様に平坦になることが確認できる。これはブロック内のDCT係数毎の出現頻度に差がないことを意味している。また符号化する際、量子化マトリックスをこの拡張階層の信号特性に合わせて最適化する必要がある。

4.3 符号化特性

4.3.1 実験内容と実験条件

本方式の符号化特性を検証するために、基本階層、拡張階層のそれぞれのレート歪み特性を調べた。基本階層については、分割の際のフィルタなし (方式1) とフィルタあり (方式2) の場合について比較し結果を図8に示す。拡張階層については、

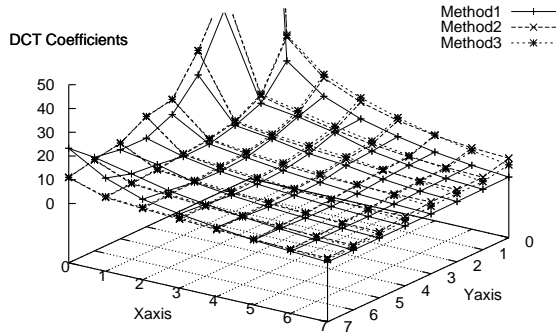
Method1A フィルタなし・非補償

Method1B フィルタなし・動き補償

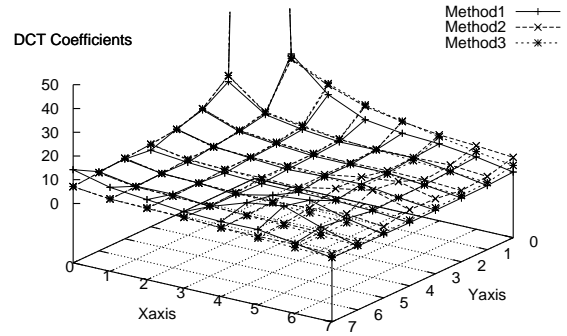
Method2A フィルタあり・非補償

Method2B フィルタあり・動き補償

の場合について比較し結果を図9に示す。ただし基本階層は $MQ = 10$ として固定量子化した。また、基本階層と拡張階層をあわせた場合のレート歪み特性を図10に示す。本実験では、MPEG-2のイントラ符号化に限定し、DCT係数のみの情報量を調べた。以下に符号量の演算方法を示す。情報源 S におけるエントロピー $H(S)$ は、

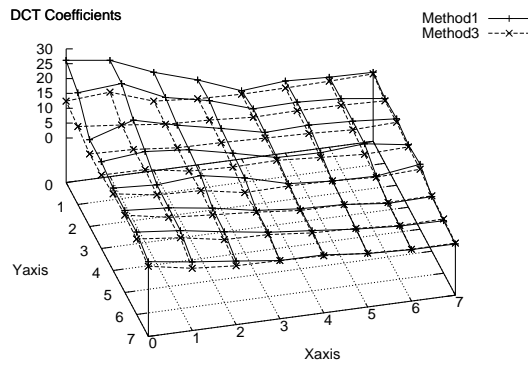


(a) Bus

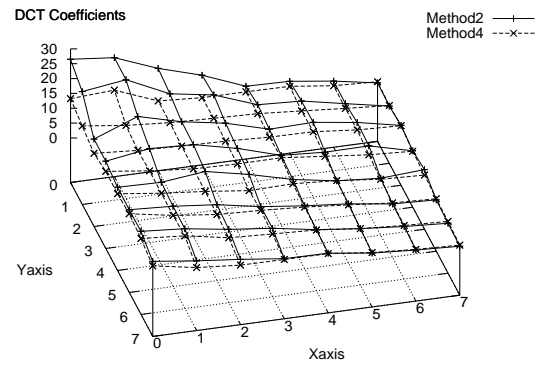


(b) Table Tennis

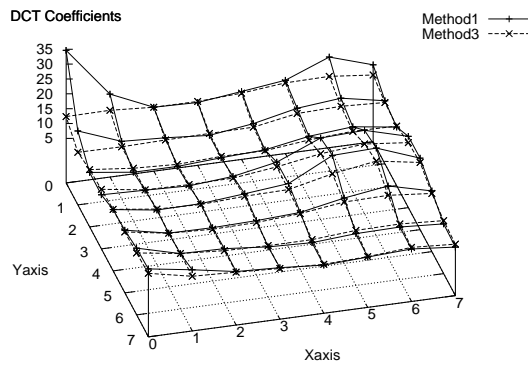
図5 非分割・分割における輝度成分の DCT 係数の標準偏差



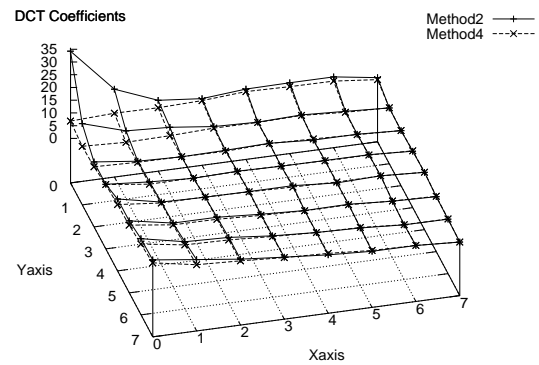
(a) フィルタなし (Bus, Base:mq=10)



(b) フィルタあり (Bus, Base:mq=10)



(c) フィルタなし (Table Tennis, Base:mq=10)



(d) フィルタあり (Table Tennis, Base:mq=10)

図7 差分信号の輝度成分の DCT 係数の標準偏差 (Buss, Table Tennis)

$$H(\mathbf{S}) = - \sum_{n=1}^M p_n \log_2 p_n \quad (5)$$

となる。この式より、算出する符号量 $I(\mathbf{S})$ を、

$$I(\mathbf{S}) = N \times H(\mathbf{S}) \quad (6)$$

として求めた。ここで、 N は情報源 \mathbf{S} に含まれる全要素の出現回数の合計であり、 P_n はシンボルの出現確率である。

4.3.2 結果と考察

図8の基本階層のレート対歪み特性から、フィルタを適用し

た場合、符号化効率が低下していることが確認できる。“Bus”に関しては、その差はあまり見られないが、“Table Tennis”に関しては、その差が顕著である。

図9の拡張階層のレート対歪み特性から、位相のずれを予測するために動き補償を行うことにより、予測誤差を低減し符号化効率を増加する効果があることを確認できる。フィルタを適用した場合に、“Table Tennis”では、符号化効率が大きく低下していることが確認できる。また、基本階層のレート歪み特性と比較すると、曲線の傾きが緩やかであることがわかる。

図10より、“Bus”に関しては、原画像をそのまま符号化した

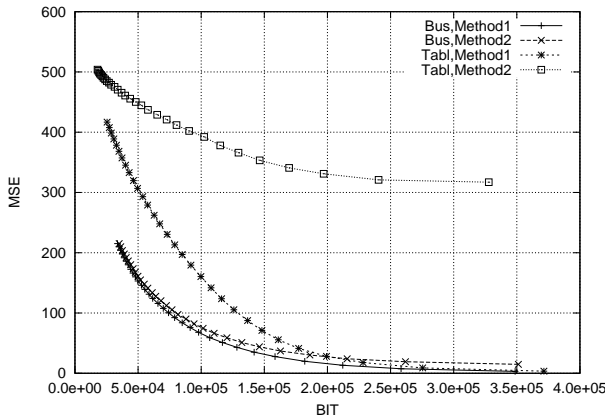


図8 基本階層のレート対歪み特性

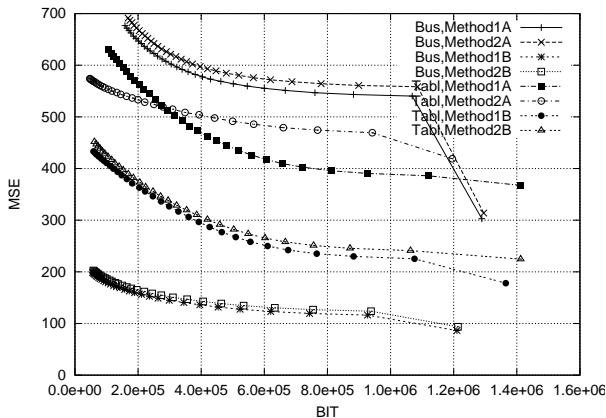


図9 拡張階層のレート対歪み特性

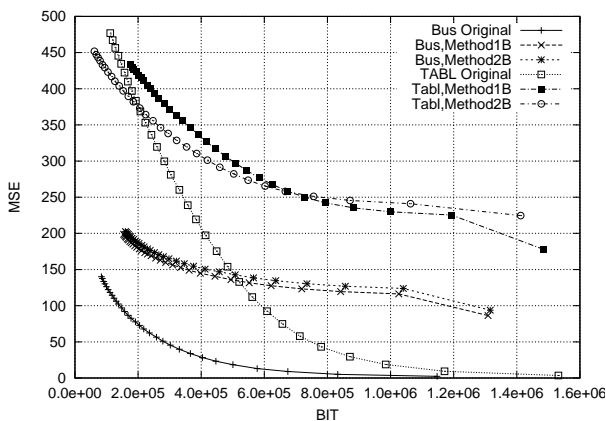


図10 提案方式のレート対歪み特性

場合に比べて、提案方式はいずれも符号化効率が低いことが確認できる。またフィルタを適用する場合は、フィルタを適用しない場合に比べ、発生符号量に対してMSEがわずかに上回り、符号化効率が低下していることが確認できる。"Table Tennis"に関しては、フィルタを適用を適用しない場合に比べ、フィルタを適用する場合は、符号量に対するMSEが顕著に大きくなっていることがわかる。ローパスフィルタを適用したことが画像の劣化につながったと考える。しかし、フィルタ使用した場合、低レート時における符号量に対するMSEは原画像をそのまま符号化した場合に非常に近づいており、有効性を確認できる。

以上より提案方式の符号化効率の原因として、拡張階層とな

る差分信号が周波数領域において値の小さいDCT係数が低周波から高周波まで一様に分布していることが考えられる。これにより、符号化時において高周波成分を荒く量子化すると、量子化誤差が大きくなる。しかし、MSEを改善するために量子化パラメータを小さい値にしても、符号量が増加するため根本的な改善とならない。符号化特性を改善するためには、基本画像と分割画像間の位相のずれをいかに誤差を少なく予測する必要がある。

5. 今後の課題

以下に今後の検討課題を示す。

- (1) 基本階層から拡張階層を効率よく予測する手法
- (2) 拡張階層の符号化方法に関する課題
- (3) 基本階層と各拡張階層に割り当てる符号量制御方法
- (4) インター符号化における符号化特性

本方式では基本階層から拡張階層を予測するために、動き補償を用いているが、画素分割により生じる基本階層と各階層間での位相のずれを補償するものではないため、位相のずれを補償するフィルタ設計が必要である。また拡張階層を符号化した際に、発生符号量に対してMSEが大きく符号化効率が低いという問題がある。原因は、拡張階層となる差分信号の特性、DCT係数の分散特性が一様に平坦であるため、高周波数成分にもDCT係数が残っているため、ランレグス符号化の利点を活用できないためである。よって信号特性に適した符号化方法を検討する必要がある。本実験から、基本階層と拡張階層のレート対歪み特性が確認できたので、この結果より、基本階層と拡張階層に割り当てる符号量を制御する方法を検討する。最後に、本実験では符号化する際、フレーム内でのイントラ符号化のみであった。しかし、動画像を効率よく符号化するためには、フレーム間のインター符号化を検討し、符号化特性を調べる必要がある。

6. まとめ

本稿では、映像配信サービスを行う上で、多様化する再生環境、伝送条件に柔軟に対応できるスケーラブル符号化による映像配信システムに着目した。従来のスケーラブル符号化の問題点を指摘した。拡張階層間の依存関係を排除し、基本階層と各拡張階層を多重化することにより、単独階層のみで再生可能なスケーラブル符号化方式を検討した。また提案方式の特性を評価する実験を行い、その結果、拡張階層の符号化に関して、現行のイントラ符号化ままでは満足する符号化効率を得る事ができないことが確認できた。最後に、今後の検討すべき課題を述べた。

文 献

- [1] 小林, "映像配信技術の動向," 2000年映像情報メディア学会冬期大会, Dec. 2000
- [2] ISO/IEC 13182-2:2000, "Generic Coding of Moving Picture and Associated Audio, Part-2 Video", 2000
- [3] 永吉, 花村, 笠井, 富永, "MPEG-2ビットストリーム分離・合成機能によるスケーラブル映像符号化方式," 電子情報通信学会論文誌 D-2 Vol.J84-D-2 No.12 pp.2525-2540 2001年12月
- [4] 涌井, 永吉, 花村, 富永, "ビットストリーム変換によるスケーラブル映像符号化方式に関する検討," PCSJ2003, P-2.01, pp.13-14, Nov.2003
- [5] ISO/IEC 14496-2:2001, "Generic Coding of Audio-Visual objects, Part-2 Visual,"
- [6] 角田, 永吉, 花村, 富永, "ビットストリーム変換処理による多階層映像符号化方式に関する検討," 信学春季総大, 2003年3月