

Selective Enhancementと基本レイヤの縮小を利用したMPEG-4 ス ケーラブル符号化に関する一検討

秦泉寺久美[†] 中嶋 淳一[†] 上倉 一人[†] 八島 由幸[†]

[†] 日本電信電話株式会社, NTT サイバースペース研究所
〒239-0847 神奈川県横須賀市光の丘1-1

E-mail: †{jinzenji.kumi,nakajima.junichi,kamikura.kazuto,yashima.yoshiyuki}@lab.ntt.co.jp

あらまし 本論文は、一つのビットストリームから低いレートでビットストリームが切り出された場合でも、主観的な品質の劣化を抑制するスケーラブルビットストリームの構築を目的とする。スケーラブル符号化として、基本レイヤにMPEG-4 Advanced Simple Profile、拡張レイヤにMPEG-4 Fine Granularity Scalable (FGS) Profileを用い、拡張レイヤの符号化の優先順位を変更する手法であるSelective Enhancement(SE)を利用して、品質の向上を図る。本論文では、差分DCT係数と画像の特徴の関係を明らかにし、画像の主観品質に寄与するブロックノイズおよびモスキートノイズが出現するブロックを差分DCT係数およびローカルデコード画像から推定する画像特徴抽出法を提案した。また、本画像特徴抽出法と符号化される差分DCT係数の特徴から、SEにおける2つの自由度である(1)どの領域に適用するか、(2)どのくらい優先するかを決定するビットプレーンシフト法を提案した。また、本提案手法を自然画像およびセルアニメーション画像に適用し、スケーラブル符号化実験を行い、PSNRとノイズの現れ方から画像の品質の評価を行った。その結果、従来のFGS符号化を用いた場合に比べて、主観品質を向上させることを実証した。

キーワード MPEG-4, DCT係数, 特徴抽出, スケーラブル符号化

A Study of MPEG-4 scalable video coding using Selective Enhancement and resolution reduction of base layer

Kumi JINZENJI[†], Junichi NAKAJIMA[†], Kazuto KAMIKURA[†], and Yoshiyuki YASHIMA[†]

[†] Nihon Telegraph and Telephone Corporation, NTT Cyber Space Labs.
Hikarinooka 1-1, Yokosuka-Shi, Kanagawa, 239-0847 Japan

E-mail: †{jinzenji.kumi,nakajima.junichi,kamikura.kazuto,yashima.yoshiyuki}@lab.ntt.co.jp

Abstract This paper aimed to improve subjective quality of scalable video under low bit-rate conditions in One Source Multi Use. We use MPEG-4 Streaming Profile, which has Advanced Simple Profile as a base layer coding and Fine granularity Scalable (FGS) profile as an enhancement layer, with Selective Enhancement (SE). SE has two major parameters, such as "domain" and "priority". We suggest that FGS bitstream is closely related to DCT coefficients difference before and after quantization in base layer. We classify the DCT coefficients' feature in relating to the subjective quality. Then we propose feature extraction algorithm for Improving Subjective quality from DCT coefficients and local decode images in the base layer. Then, we propose a bit plane shift algorithm that specifies the two parameters of SE using the result of feature extraction. Finally, we applied the algorithm to scalable video coding. Experimental results show that it effectively improves the subjective quality of video at the low bitrate in comparison with usual scalable coding.

Key words MPEG-4, DCT Coefficients, Feature Extraction, Scalable Video Coding

1. はじめに

本論文では、ビットストリームの一部がカットされても高い主観品質を維持するスケーラブル符号化方式の構築を目的とし、MPEG-4 ASP/FGS 符号化を基本としたビットストリーム生成の検討を行う。

FGS 符号化を用いたスケーラブル符号化における画質を向上させる方法のひとつとして、Selective Enhancement(SE) という符号化優先順位を変更する手段がサポートされている。SE は (1) どの領域を、(2) どの程度優先させるかという 2 つの自由度を持ち、これらを適切に設定する必要がある。

本論文は、FGS 符号化において、画像の主観品質を左右する 2 つの歪であるブロックノイズおよびモスキートノイズが出現するブロックを差分 DCT 係数およびローカルデコード画像から推定する画像特徴抽出法を提案する。また、提案手法によって領域分割された結果を用いて、SE を行うアルゴリズムを提案する。また、自然画像およびセルアニメーション画像にて、主観品質に寄与する領域を優先的に FGS 符号化することで、主観品質を向上させる条件について考察する。

以降、第 2 章にて FGS を用いた MPEG-4 スケーラブル符号化の概略、第 3 章にて差分 DCT 係数の特徴について、第 4 章にて主観画質に寄与する画像特徴の定義とその画像特徴を抽出するアルゴリズムの提案について、第 5 章では画像特徴抽出法によって領域分割された結果を用いたビットプレーンシフト方法の提案と MPEG-4 スケーラブル符号化に提案手法を適用した符号化実験について、第 6 章では、結論について述べる。

2. MPEG-4 スケーラブル符号化

2.1 MPEG-4 スケーラブル符号化器構成

図 1-(a) に通常の FGS 符号化器を示す。また、図 1-(b) および (c) に本論文で使用した符号化器の構成を示す。

図 1-(b) は、量子化前の DCT 係数と量子化後の DCT 係数の差分 (差分 DCT 係数) をそのままビットプレーン展開し、VLC 符号化することで、図 1-(a) とほぼ同じビットストリームを生成できる。

図 1-(c) は、図 1-(a) において、基本レイヤを縮小画像にて符号化し、そのローカルデコード画像を拡大したものと、原画像の差分を拡張レイヤにて符号化したものである。再生側では、基本レイヤのビットストリームを復号し、拡大処理をし、拡張レイヤ

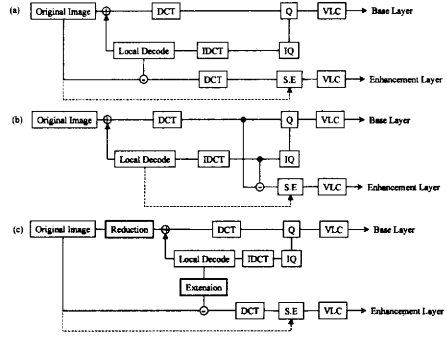


図 1 FGS を用いた MPEG-4 スケーラブル符号化器構成: (a) 一般的な構成 (その 1), (b) 一般的な構成 (その 2), (c) 基本レイヤに拡大・縮小処理を加えた構成。

Fig. 1 MPEG-4 Scalable Encoder using FGS: (a) General encoder 1, (b) General encoder 2, (c) Encoder with base layer scaling .

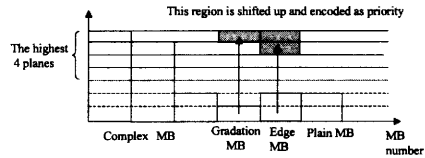


図 2 MPEG-4 FGS 符号化における Selective Enhancement の概念。

Fig. 2 Concept of "Selective Enhancement" in MPEG-4 FGS Encoder.

のデコード画像と重畳する。

2.2 FGS における Selective Enhancement

図 2 に SE の概念を示す。SE によって、MB 毎にビットプレーンのシフトを行うことが可能である。差分 DCT 係数は図 2 のようにビットプレーン展開され、上位プレーンから順番に符号化される。ここで、任意に優先したい MB のビットを何ビットかシフトする。これによって、差分 DCT 係数の振幅の小さな MB でも、SE によって上位のプレーンにて符号化、伝送される。どの MB をどの程度シフトするかは標準では規定されていない。

3. 主観画質に寄与する画像特徴抽出アルゴリズム

3.1 画像特徴抽出アルゴリズム

図 3 に主観画質を考慮した画像特徴の抽出アルゴリズムを示す。本アルゴリズムは、差分 DCT 係数を用いて、画像を前節で定義した特徴を反映する領域に分類する。以下に、主観品質に影響があるとと思われる順に、領域を列挙する。

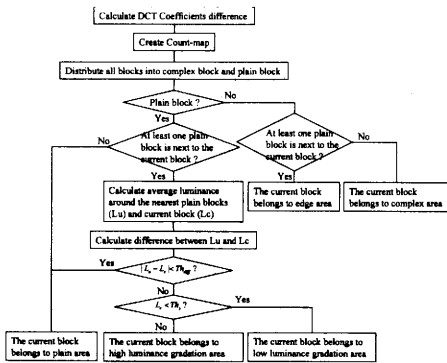


図3 差分DCT係数を用いた画像特徴抽出アルゴリズム。

Fig.3 Feature extraction algorithm using DCT coefficients.

- (1) エッジ領域
- (2) 高輝度グラデーション領域
- (3) 複雑領域
- (4) 低輝度グラデーション領域
- (5) 平坦領域

エッジ領域、高輝度グラデーション領域はそのノイズの出現が主観品質に大きな影響を与える領域である。複雑領域においては、フレームの一部に集中して現れる場合は、被写体等である可能性があり、この領域の品質の劣化は、主観品質に影響を与えると考えられる。低輝度グラデーション、平坦領域ではノイズが主観品質に影響を与えにくいものとする。

以下に、アルゴリズムの詳細を記す。最初に、差分DCT係数を算出する。差分DCT係数を整数 N で除し、“0”になった係数をカウントする。このカウント値をマップ化したものをカウントマップと呼ぶ。原画像の任意のブロックにおけるカウント値がカウントマップの任意の座標における値となる。カウントマップの任意の座標 (x, y) において取りうるカウント値 $Val_{count}(x, y)$ の範囲は $0 \leq Val_{count}(x, y) \leq 63$ である。このカウント値 $Val_{count}(x, y)$ を閾値 Th_c で2値化し、“0”の場合を複雑ブロック、“1”の場合を平坦ブロックと分類する。

任意の着目ブロックが複雑ブロックであった場合、着目ブロックの近傍に一つでも平坦ブロックがある場合、このブロックをエッジ領域に属するとする。それ以外は複雑領域とする。ここで、2値化閾値 $Th_c = 50$ とした。

平坦ブロックと判定されたブロックにおいて、近傍のブロックに平坦ブロックがあり、周囲の平坦ブロックの輝度平均を各々算出し、その差分絶対値が

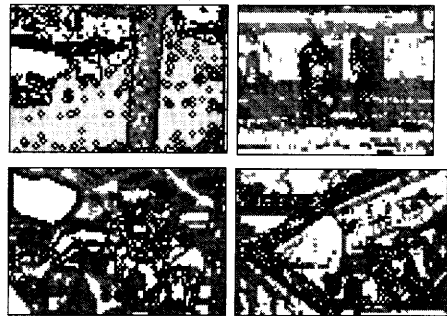


図4 画像特徴抽出実験結果。左上: Flower garden, 右上: Baseball, 左下: Robot1, 右下: Robot2

Fig.4 Results of feature extraction experiment. Upper left: Flower garden, Upper right: Baseball, Lower left: Robot1, Lower right: Robot2

閾値 Th_{diff} を超えた場合にグラデーション領域とする。またそれ以外を平坦領域とする。ここで、ブロックにおける輝度平均は基本レイヤにおけるローカルデコード画像の値を用いる。

グラデーション領域と判定されたブロックにおいて、輝度平均が Th_y を超える場合を高輝度グラデーション領域とする。また、それ以外を低輝度グラデーション領域とする。

このように、全てのブロックがエッジ領域、複雑領域、高輝度グラデーション領域、低輝度グラデーション領域、平坦領域に分類される。

3.2 画像特徴抽出実験

以下の4つの画像^(注1)に画像特徴抽出アルゴリズムを適用し、画像特徴抽出実験を行った。

- "Flower Garden" VGA(720 × 480 画素)
- "Baseball" VGA(640 × 480 画素)
- "Robot1" VGA(720 × 480 画素)
- "Robot2" VGA(720 × 480 画素)

実験で用いたパラメータは、 $N = 8, Th_c = 50, Th_{diff} = 50, Th_y = 150$ である。基本レイヤは1 Mbpsにてレート制御の下、符号化したものを使用した。

図4に実験結果を示す。ブロック毎に領域を分類する。明度が低い順に、エッジ領域、高輝度グラデーション領域、低輝度グラデーション領域、複雑領域、平坦領域を示す。

一般的なエッジ検出フィルタ^(注2)を用いて画像を処理するとエッジ部分が抽出されるが、「花畑」はもエッ

(注1) : 原画像の概観については紙面の都合上割愛する。

(注2) : Sobel Filter など。

ジとして抽出される。これに対して、本アルゴリズムでは、「花畑」は複雑領域と判断し、「花畑」と「空」の境界をエッジ領域と判断している。また、「Baseball」においては、バッター、審判、キャッチャーの輪郭部分がエッジ領域と判定され、背景部分はグラデーション領域として判定されている。「Robot1」、「Robot2」においても、線画の部分にエッジ領域が出現し、均等色領域には、高輝度および低輝度グラデーション領域と判定され、ほぼ、直感による判別と同程度の判定結果となった。

4. スケーラブル符号化実験

本節では、提案画像特徴抽出方法を符号化に適用してスケーラブルなビットストリームを作成し、その品質を PSNR および符号化画像のノイズの現れ方から評価する。

4.1 ビットプレーンシフトにおける基本方針

本節では、SE における (1) シフトすべき MB と (2) シフトするビット量について、ブロックノイズの目立つ高輝度グラデーション領域と、モスキートノイズの目立つエッジ領域のビットプレーンをそのフレームにおける最大ビットプレーンまでシフトする方針とする。最初に、拡張レイヤのビットレートから、1 フレーム当たりのビット数を算出する。算出されたビット数に達するまで、FGS 符号におけるビットプレーンを符号化する。差分 DCT 係数から、そのフレームにおける最大ビットプレーン値を算出する。本提案画像特徴抽出法にて領域分割された結果から、エッジ領域および高輝度グラデーション領域の各 MB における最大振幅値を最大ビットプレーン値にあわせてビットプレーンシフトする。

4.2 符号化実験

本提案画像特徴抽出アルゴリズムを図 1-(b) および (c) の符号化器に適用し、符号化実験を行い、画像品質を PSNR と主観画質から評価を行った。

便宜上、一般的な MPEG-4 スケーラブル符号化器構成である図 1-(b) を通常スケーラブル符号化方式 (Normal)、基本レイヤにおいて画像の縮小拡大を行う図 1-(c) を拡大縮小スケーラブル方式 (Scaling) と呼ぶ。2 つのスケーラブル符号化方法において、SE を用いない方式 (no-SE) と本画像特徴アルゴリズムとビットプレーンシフト方式を用いた方式 (提案 SE 方式, Proposed) の 2 つの方法を適用し、合計 4 つのパターンのスケーラブル符号化を比較検討する。符号化パラメータは以下のとおりである。

- サイズ：VGA

- ビットレート：

- 基本レイヤ：500kbps, 拡張レイヤ：500kbps
- 基本レイヤ：500kbps, 拡張レイヤ：1Mbps
- 基本レイヤ：500kbps, 拡張レイヤ：1.5Mbps

基本レイヤを 500kbps で符号化し、拡張レイヤは、500kbps, 1Mbps, 1.5Mbps の 3 通りのビットレートで符号化した。こうすることで、基本レイヤ 500kbps, 拡張レイヤ 1.5Mbps, 合計 2Mbps で符号化した一つのビットストリームから 1Mbps, 1.5Mbps, 2Mbps の 3 種類のビットストリームを取り出すこと (One source multi use) と等価な評価が可能である。符号化実験には前章と同じく「Flower garden」、「Baseball」、「Robot1」、「Robot2」の 4 つの画像を用いる。

最初に、通常スケーラブル方式について、本提案 SE 方式と SE を行わない方法について比較、考察を行う。「Flower garden」において、提案 SE 方式の顕著な品質の向上が見られた。拡張レイヤを 500kbps および 1.5Mbps で符号化した画像を図 5 に示す。

SE を行わない方法に比べて提案 SE 方式の「空」の領域のブロックノイズが軽減されていることが確認できる。また、SE を用いない場合では、「木」の一部が欠落したように見えるのに対し、提案 SE 方式ではつながって見える。複雑領域（「花畑」等）の差分 DCT 係数の絶対値が大きく、「空」の差分 DCT 係数が現れるのが相対的に低い順位のビットプレーンであるため、拡張レイヤを高いビットレートで符号化しても、十分に表現できるビットプレーンまで符号化することが困難であったためと考えられる。両方式の平均 PSNRh は提案 SE 方式が 0.3dB 以下の劣化にとどまった。「Robot1」においては、動きが激しいフレームにおいて、提案 SE 方式の顕著な品質の向上が見られた。拡張レイヤを 1.5Mbps で符号化した場合の様子を図 6 に示す。提案 SE 方式においてブロックノイズが低減されていることが確認できる。平均 PSNR は、提案 SE 方式が SE を行わない場合に比べて、0.25~0.4dB 程度の劣化が認められた。「Baseball」および「Robot2」においては、両方式とも主観品質において顕著な差は見られなかった^(注3)。「Baseball」における両方式の平均 PSNR の差はほとんどゼロであった。「Robot2」における、提案 SE 方式の平均 PSNR は SE を用いない場合に比べて 0.25~0.3dB 程度の劣化にとどまった。

次に、提案 SE 方式における通常スケーラブル方式と拡大縮小スケーラブル方式の品質の差異につい

(注3)：紙面の関係上、品質に差異のない画像の例は割愛する。

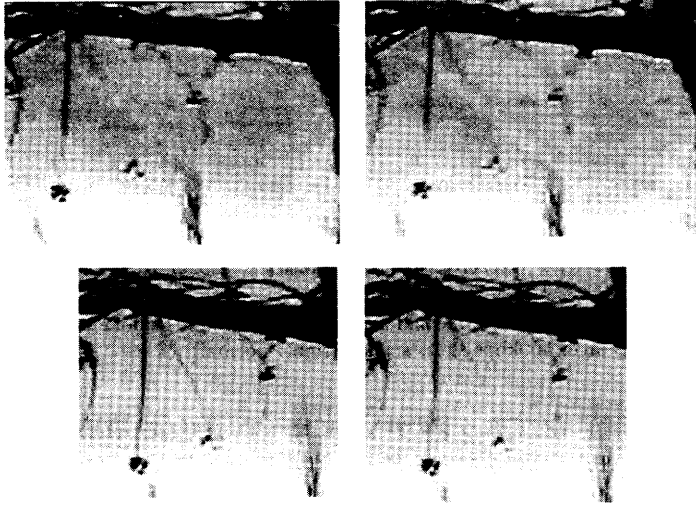


図5 "Flower garden"における符号化画像の比較 (Normal). 上段: 拡張レイヤ 500kbps における符号化画像 (左: 提案方式, 右: 従来方式, 以下同様), 下段: 拡張レイヤ 1.5Mbps における符号化画像.

Fig.5 Coding examples of normal method: "Flower garden". Upper: enhancement layer 500kbps (left: proposed, right:conventional), lower: enhancement layer 1.5Mbps.



図6 "Robot1"における符号化画像の比較 (Normal). 拡張レイヤ 1.5Mbps における符号化画像 (左: 提案方式, 右: 従来方式).

Fig.6 Coding examples of normal method: "Robot1".Enhancement layer 1.5Mbps (left: proposed, right:conventional).



図7 "Baseball"における符号化画像の比較. 拡張レイヤ 500kbps における符号化画像 (左: 拡大縮小方式, 右: 通常方式).

Fig.7 Coding examples of normal and scaling method: "Baseball". Enhancement layer 500kbps (left: scaling, right: normal).

て考察を行う。

"Baseball"および"Robot2"にて、両スケーラブル方式に顕著な差異が見られた。拡張レイヤを500kbps

で符号化した場合の様子を図7("Baseball")および図8("Robot2")に示す。"baseball"においては、「バックネット」の部分において、顕著な差が見られる。ま

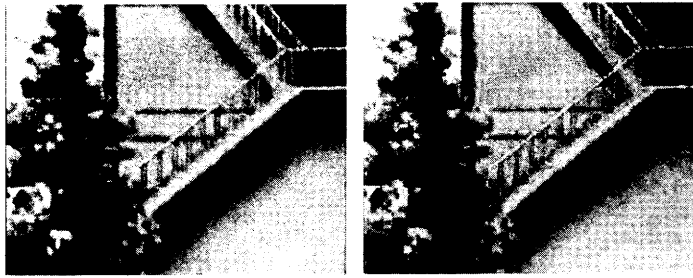


図8 "Robot2"における符号化画像の比較. 拡張レイヤ 500kbps における符号化画像 (左: 提案方式, 右: 従来方式).

Fig.8 Coding examples of normal and scaling method: "Robot2". Enhancement layer 500kbps (left: scaling, right: normal).

た,"Robot2"においては、「階段」部分のエッジの鮮明度と「壁」の部分のブロックノイズの現れ方において、拡大縮小方式に明らかな改善が見られた。ブロックノイズが低減された領域は、主に低輝度グラデーション領域である。この領域は、基本レイヤを縮小することで、少ない符号量で十分な品質を達成することができる。また、この領域は高周波成分を含まない領域であるので、縮小拡大処理に伴う高周波成分がカットされることによる主観品質の劣化を伴わない。エッジなどの高周波は拡張レイヤにて符号化されたため、品質が向上したと考えられる。一方で,"Flower garden"や"Robot1"では、高輝度グラデーション領域のブロックノイズの改善が見られたが、エッジ領域や複雑領域の品質劣化が見られた^(注4)。これは、両画像ともエッジ領域、複雑領域の占める割合が多い画像であり、加えて、基本レイヤが500kbpsと低レートであったことが品質の劣化を招いたと考えられる。すなわち、縮小拡大処理によって高周波が消去された画像から差分DCT係数をもとめるため、大きな振幅をもつ差分DCT係数の割合が高くなり、本実験に用いた条件下では十分な符号量が与えられなかったためと考えられる。全画像において、PSNRは通常スケラブル方式を用いた場合よりも、さらに2dB程度劣化する傾向があった。基本レイヤの縮小拡大処理が効果をもたらすのは、エッジ領域や複雑領域が十分少なく、グラデーション領域の面積が比較的多い画像であると言える。反対に、画像品質の劣化をもたらすのは、グラデーション領域に比べてエッジ領域や複雑領域の多い画像であると言える。

本提案方式は低いビットレートにてビットストリームがカットされた場合でも、主観品質に寄与するノイズの現れ方から、提案SE手法はSEを用いない方法

に比べて、良好な主観品質を達成できる見通しを得た。また、画像によっては、基本レイヤを一度縮小して符号化し、符号化画像を拡大して原画像と差分をとる場合が、主観品質に顕著な向上をもたらす場合があることを示した。実際的主観評価をまじえた考察は今後の課題としたい。

5. おわりに

本論文は、差分DCT係数の性質から、画像の主観品質を左右する2つの歪であるブロックノイズおよびモスキートノイズが出現するブロックを差分DCT係数およびローカルデコード画像から推定する画像特徴抽出法を提案した。また、本画像特徴抽出法によって領域分割された結果から、ビットプレーンシフトを行う領域の選択と、ビットシフト量を決定するビットプレーンシフトアルゴリズムを提案した。また、本提案手法をスケラブル符号化に適用し、主観品質に寄与するブロックノイズとモスキートノイズの現れ方から考察を行った。その結果、一つのビットストリームからSEを行わない方法に比べて高い主観品質のビットストリームを取り出せることを示した。

本研究は、通信・放送機構の委託研究「通信ネットワーク利用放送技術に関する研究開発」による。

文 献

- [1] Weiping Li, "Overview of fine granularity Scalability in MPEG-4 Video Standard," IEEE Trans. on CSVT, Vol.11, No.3, March 2001.
- [2] 秦泉寺久美, 中嶋淳一, 上倉一人, 八島由幸, "DCT係数を用いた主観品質に寄与する特徴の推定とMPEG-4スケラブル符号化への適用," 信学技報 IE2003-63, pp.45-50.

(注4) : 画像は紙面の都合上割愛する。