

LI-002

画素サンプリング画面内予測を用いた画像符号化の効率改善 Improving Efficiency of Image Coding Using Sampling Intra Prediction

和田 直史†
Naofumi Wada

中條 健†
Takeshi Chujoh

1. まえがき

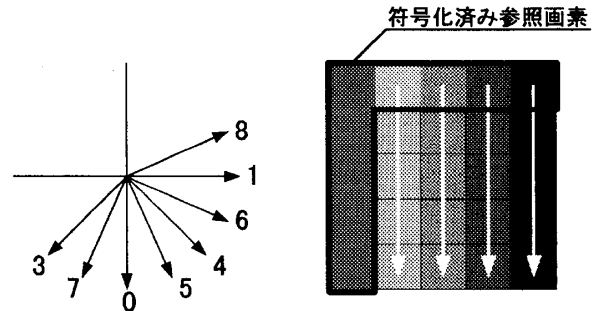
2003年に規格化された動画画像符号化方式である H.264/MPEG-4 AVC (以下 H.264) [1]は、従来の地上波テレビ放送などで使われている MPEG-2 の約 2 倍の符号化効率を実現できるとして、現在ワンセグ放送や次世代光ディスクの動画画像符号化方式として採用されている。H.264 では、大きく分けて I フレームのようにフレーム内で閉じた処理を行う画面内符号化と、P フレーム、B フレームのようにフレーム間の相関を利用する画面間符号化が存在する。H.264 の画面内符号化においては、静止画像の符号化標準規格である JPEG の性能を上回り、さらには JPEG2000 と比較してもほぼ同等の性能が得られるという報告がある[2]。また、近年 H.264 Intra-only Profile と呼ばれる画面内符号化のみの規格が検討されるなど、映像の加工・編集の容易性をメリットとして持つ画面内符号化の重要性とその性能向上が注目されている。

本稿では、特に画面内符号化における予測処理である画面内予測 (以下、イントラ予測) に着目し、H.264 の画面内符号化に対して符号化効率をさらに向上させるための予測方式の検討を行った。ここでは、新たなイントラ予測方式として画素サンプリング画面内予測 (Sampling Intra Prediction: SIP) を提案し、実験によりその有効性を示す。

2. H.264 におけるイントラ予測の概要と問題点

H.264 や MPEG-2 など従来の動画画像符号化における画面内符号化では、画像を $N \times N$ 画素の矩形ブロックに分割し、一般的にはラスタスキャン順にブロック単位で符号化処理を行う。H.264 High Profile では、輝度信号に対してブロックサイズの異なる 3 種類のイントラ予測 (4×4 予測, 8×8 予測, 16×16 予測) が規定されており、それらを 16×16 画素から成るマクロブロック (以下、MB) 単位で切り替える可変ブロックサイズ方式が採用されている。 4×4 , 8×8 予測では、利用可能な参照画素の平均値で予測する DC 予測 (モード 2) と、図 1(a) のようにそれぞれ 22.5 度ずつの方向から予測する方向予測の計 9 つの予測モードが規定されている。また、 16×16 予測では、垂直予測、水平予測、DC 予測、平面予測の 4 つの予測モードが規定されている。方向性を持つ予測モードでは、ブロック周辺の符号化済み参照画素を用いて予測方向に外挿補間を行なって予測値を生成する。

ここで、一般的に画像は空間的距離に近い画素ほど近い値となる傾向があり、従来イントラ予測のようなブロック周辺画素からの外挿予測では参照画素からの距離が遠い画素 (例えば、図 1(b) に示した予測モード 0 の場合、ブロック下に位置する画素) において、予測誤差が増大するという問題がある。



(a) 各予測モードの予測方向 (b) 予測モード 0
図 1 H.264 におけるイントラ予測

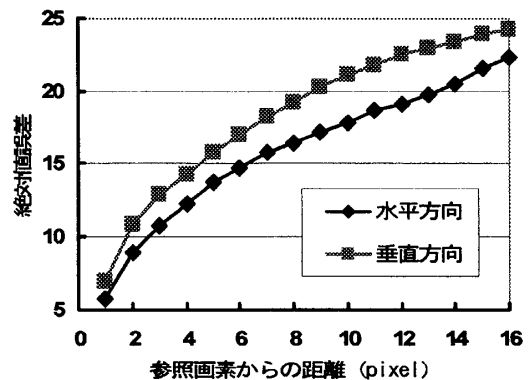


図 2 参照画素からの距離と絶対値誤差の関係

上記の問題点を確認するため簡単な実験を行なった。図 2 は、符号化前の原画像に対して、MB 内の画素と MB 周辺の参照画素との絶対値誤差を参照画素からの距離毎にプロットしたグラフである。画像は QCIF (176×144 pixel) からフル HD (1920×1080 pixel) までのサイズの画像を 17 種類用い、縦軸を全ての画像における絶対値誤差の平均値としてプロットした。図 2 より、参照画素からの距離に対して水平方向、垂直方向共に予測誤差は単調増加していることが確認できる。

3. Sampling Intra Prediction (SIP)

本稿では、2 章で述べた問題点を改善するため、新たな予測方式として画素サンプリング画面内予測 (Sampling Intra Prediction: SIP) を提案する。

画素サンプリングを用いた画像符号化方式については、過去にも多くの研究がなされている。例えば、HDTV 用のアナログ帯域圧縮技術として検討された MUSE (Multiple Sub-Nyquist Sampling Encoding)[3] や、TAT (Time-Axis Transform)[4] などは、空間方向に画素をサブサンプリングすることにより伝送する情報量を削減し、受信側で間引かれた画素を補間することによって元のサイズの再生画像を

† 株式会社 東芝 研究開発センター

得る方式である。また、インタレース方式におけるフィールド・フレーム予測や文献[5]のように、画面内を一様にサブサンプリングし、サブサンプリングした画素をフレーム内又はフレーム間で予測する方式がある。

本稿の提案方式である SIP は、画素サンプリングを H.264 における画面内符号化の枠組みで用い、イントラ予測の予測効率を向上させることを目的としたものである。以下、3.1 ではまず SIP の基本となる処理について述べる。次に、3.2 では MB 毎にサンプリングパターンを適応的に選択する方法について述べ、3.3 では SIP の性質に適した量子化パラメータの制御方法について述べる。

3.1 SIP の基本処理

SIP は、所定のパターンに従ってサンプリング(ここでは、単なる間引き選択を意味する)した MB 内の画素を、ブロック周辺画素からの外挿予測によって先に予測符号化し、そのローカルデコードを用いて他の画素を内挿予測する。これにより、従来 H.264 のイントラ予測に対して、より空間的に相関の高い隣接画素から予測を行うことで予測誤差の低減を図る。以下では SIP の基本となる具体的な処理の流れを図 3 の概念図を用いて述べる。図 3 は縦横 1 画素おきに MB 内の画素をサンプリングした場合の例であり、図 3 中の色付きの画素 d を基準画素と呼び、その他の色無しの画素 a~c を補間画素と呼ぶ。

[STEP1]

MB 内の画素を所定のサンプリングパターン(図 3 では縦横 1 画素ずつ間引くパターン)に従って間引き選択し、基準画素と補間画素に分類する。

[STEP2]

基準画素 d を既に符号化済みのブロック周辺画素から外挿予測する。

[STEP3]

基準画素 d の予測誤差を再ブロック化し、基準画素ブロック(図 3 中の D)として離散コサイン変換(DCT)、量子化、エントロピー符号化すると共に、ローカルデコード画像を作成する。

[STEP4]

STEP3 で生成した基準画素 d のローカルデコードを含む既に符号化済みの画素を参照画素として用い、補間画素 a~c を線形補間フィルタにより内挿予測する。

[STEP5]

補間画素の予測誤差を基準画素と同様に再ブロック化し、補間画素ブロック(図 3 中の A~C)として DCT、量子化、エントロピー符号化を行なう。

ここで、STEP2 の基準画素の予測では、従来 H.264 の 4×4 予測、 8×8 予測、 16×16 予測をそれぞれ適用することが可能であり、これらの予測ブロックサイズは H.264 と同様のモード判定により MB 毎に適応的に切り替える。STEP4 では、補間画素を予測するフィルタとして、動き補償における小数画素生成フィルタとしても用いられている $\{1/32, -5/32, 20/32, 20/32, -5/32, 1/32\}$ の 6 タップの線形補間フィルタを用いる。未符号化領域と隣接するブロック境界では、境界に位置する参照画素を延長することによってフィルタ処理を行う。

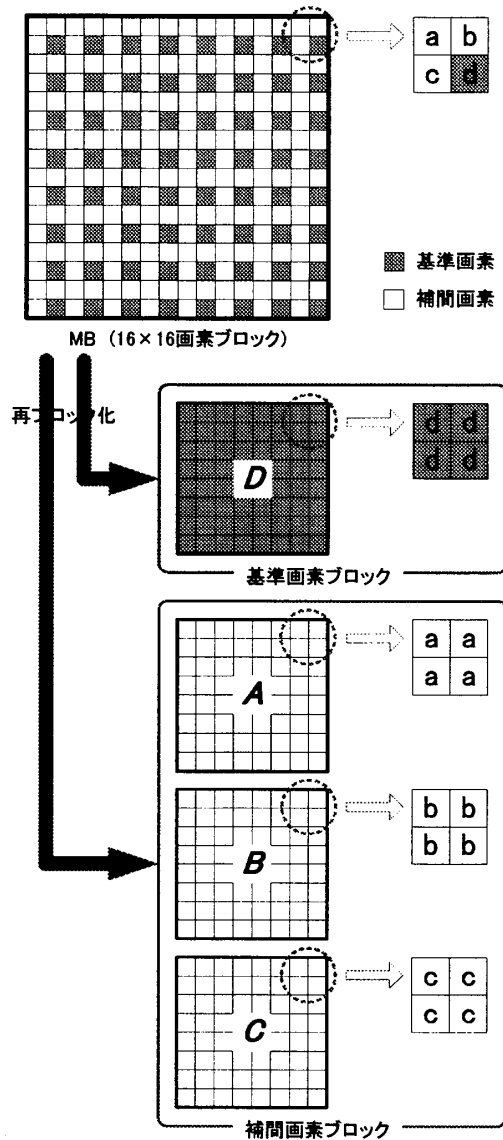


図 3 SIP の基本処理を表す概念図

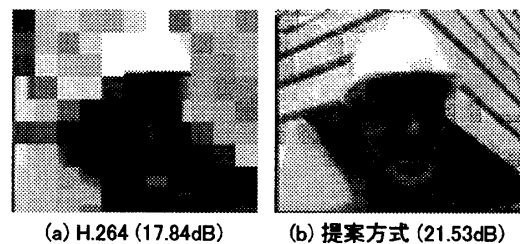


図 4 16×16 予測を用いた場合の予測画像の比較

また、図 3 のサンプリングパターンにおいては、補間画素 a~c の符号化順を $b \rightarrow c \rightarrow a$ とし、順次ローカルデコードを作成し、後続する補間画素の予測に用いる。最後に符号化する補間画素 a の予測は、既に符号化済みの b, c を参照画素として用いる。

上記処理によって予測誤差が減少することを確認するため、QCIF(176×144 pixel)サイズの標準画像である“Foreman”を用いて H.264 のイントラ予測と提案方式の予測画像を比較した例を図 4 に示す。図 4(a)は、H.264 に

において 16×16 予測のみを用いて符号化した場合の予測画像である。また、図 4(b)は提案方式を用いた場合の予測画像であり、図 3 のサンプリングパターンに基づいて基準画素及び補間画素を設定し、基準画素は(a)と同様に 16×16 予測のみで外挿予測したものである。図 4 より、提案方式では、より原画に近い予測画像が生成されていることが分かる。また、図 4 の画像の原画に対する輝度の PSNR を測定した結果、(a)では 17.84dB、(b)では 21.53dB となり、予測誤差が低減できていることを確認した。

3.2 MB 適応サンプリングモード選択

前節で述べた処理では、間引き方向にある程度画素の相関が高いか、画素値が線形に変化している場合に補間画素の予測誤差を低減する効果が高い。逆に、複雑テクスチャ領域など隣接画素の相関が低い領域では補間画素の予測精度が低下し、期待する効果が得られ難い。したがって、SIP では、隣接画素間の相関が高い方向にのみ間引きを行ない、画素間の相関が低い方向に対しては画素を間引かないことにより、性能低下を防ぐ。

ここで、図 2 のグラフより、画像は大局的に横方向の相関が強い傾向があると言えるが、局所的に見ると画像の方向性は一様ではない。そのため、本方式では、図 5 のように水平方向にのみ 1 画素ずつサンプリングするモード(モード 1)と、垂直方向にのみ 1 画素ずつサンプリングするモード(モード 2)、水平垂直両方向に 1 画素ずつサンプリングするモード(モード 3)、それに従来の H.264 のモード(モード 0)を加えた 4 つのモードを MB 毎にレート歪最適化などのモード判定により適応的に切り替える。これにより、特に垂直または水平方向に画素の相関が低い領域ではモード 1 やモード 2 が選択され、水平・垂直方向共に画素間の相関が低い場合は、従来の H.264 と同様の予測であるモード 0 が選択されることによって、符号化効率の低下を防ぐことができる。

一方、上記のようにモードを切り替えるためには、モードを示す情報を MB 毎にデコーダ側に伝送する必要がある。そのため、従来 H.264 の外挿予測で十分予測可能である場合(サンプリングしない場合もした場合も予測誤差がほとんど発生しない場合)には、追加したモード情報の分だけオーバーヘッドが増加し、符号化効率の低下を招く。また、このような領域は、MB 内の輝度変化が少ない平坦な領域であり、そのような領域では比較的 16×16 予測が選択されやすい傾向がある。したがって、 16×16 予測が選択された場合には、サンプリングモード情報を送らずに、エンコーダ・デコーダ共に固定のモード(ここでは、モード 0)を用いることによって、オーバーヘッドを削減する。

3.3 量子化パラメータ制御

SIP では、補間画素の予測に用いる参照画素に含まれる符号化歪が大きい場合、隣接する補間画素との相関が低くなり、予測誤差が十分に小さくならない可能性がある。この問題を改善する方法として、基準画素の量子化パラメータ(以下、QP)に小さい値を設定することにより、基準画素の符号化歪を小さくし、補間画素の予測精度を向上させる方法が考えられる。

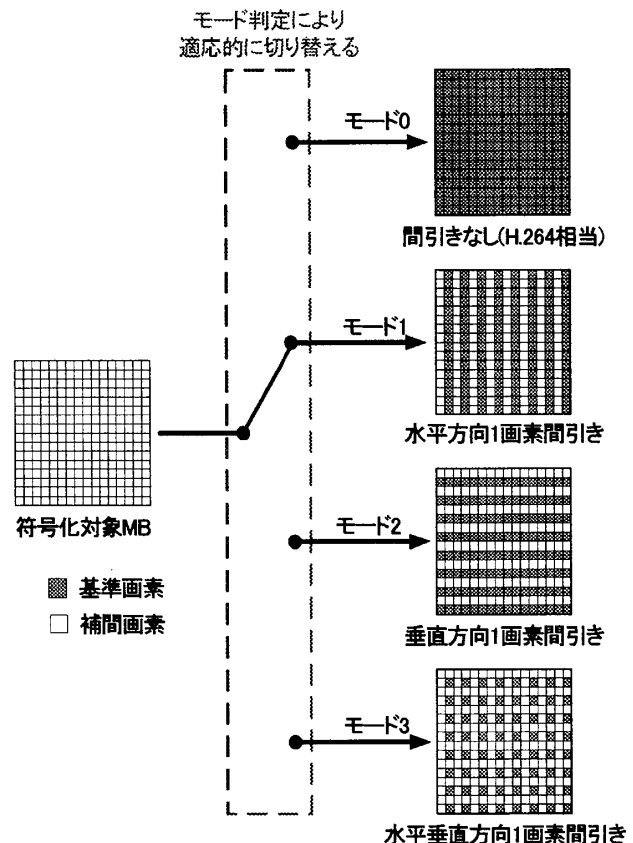


図 5 MB 適応サンプリングモード選択の概念図

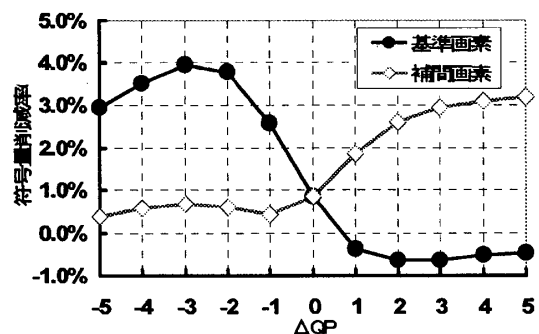


図 6 ΔQP の変化に伴う符号量削減率の変動

また、内挿予測によって予測する補間画素では、外挿予測によって予測する基準画素に比べて予測誤差が小さくなる傾向がある。予測誤差値の分布がゼロを中心としたラプラス分布であると仮定すると、補間画素ではよりゼロ付近に集中した分布となるため、補間画素では QP に大きい値を設定することによって、ある程度の画質を維持したまま発生符号量の削減が可能である。

よって、ここでは基準となる量子化パラメータ QP に対してオフセット ΔQP を設け、SIP における基準画素または補間画素の QP を $QP + \Delta QP$ として制御する(図 5 のモード 0 は除く)。基準画素および補間画素における ΔQP を変動させた場合の同一 PSNR での符号量削減率の変化を示したグラフを図 6 に示す。

図 6 のグラフより、基準画素の ΔQP を小さくし、補間画素の ΔQP を大きくすることによって、基準画素の PSNR

向上による補間画素の予測精度向上効果, 及び補間画素の符号量削減効果が得られ, 符号量削減率が向上することが確認できる. 一方, ΔQP がある一定の大きさを超えると逆に符号化効率が低下していることから, 基準画素での発生符号量の増加, 及び補間画素の PSNR 低下がデメリットとして存在すると考えられる. したがって, 上記トレードオフ関係を考慮し, 適切な基準画素と補間画素の QP を設定することによって, MB 全体としての符号化効率が向上するように符号量配分比の調整を行なう.

4. 性能評価

提案方式の有効性を確認するため, 符号化実験を行なった. 実験に用いた画像は, 720p(1280×720 pixel)の画像 4 種類と, フル HD(1920×1080 pixel)の画像 3 種類の計 7 種類である. 実験条件を表 1 に示す. 表 2 は H.264 を比較方式とした提案方式の客観評価の結果である. 評価指標は BD-SNR[6]を用いた. 表 2 における Δ Bitrate は同一 PSNR での符号量削減率(%), Δ PSNR は同一 Bitrate での輝度信号に関する画質改善度合(dB)を示している. これにより, 提案方式は H.264 High Profile と比較して同一 PSNR の下, 平均 5.8%, 最大 8.8%の符号量削減効果が得られることを確認した.

また, 図 7 は, H.264 High Profile 及び提案方式である SIP に加えて, JPEG2000 を用いて画像 “Cyclists”, “CardToss” を符号化した場合のレート歪曲線を示したものである. ここでは, JPEG2000 の Codec として Kakadu Software[7]を用いた. 図 7 より, 通常の H.264 の画面内符号化は, 720p の画像である “Cyclists” で JPEG2000 とほぼ同等の性能であるのに対し, フル HD の画像である “CardToss” においては JPEG2000 に比べて性能が下回っている. 一方, 提案方式では, “Cyclists” で H.264 及び JPEG2000 よりも高い性能を示し, さらに, “CardToss” においても JPEG2000 とほぼ同等の性能を得ることができた.

5. まとめ

本稿では, 画面内符号化におけるイントラ予測に着目し, 画素サンプリング画面内予測 (Sampling Intra Prediction : SIP) を新たに提案した. 本方式では, 特定の画素サンプリングパターンに基づいて選択した画素(基準画素)を外挿予測によって予測符号化し, 符号化済み基準画素を用いて残りの画素(補間画素)を内挿予測することにより, 予測誤差を低減する. また, サンプリングパターンをマクロブロック毎に適応的に切り替えることで局所領域毎の方向性を考慮したサンプリングを行ない, さらに, 提案方式の性質に基づいて量子化パラメータに適切なオフセット値を設定することにより, 基準画素と補間画素における符号量配分比の調整を行なった. 実験の結果, H.264 High Profile と比較し, 同一 PSNR の下, 平均 5.8%, 最大 8.8%の符号量削減効果があることを確認した.

表 1 実験条件

モード判定	レート歪最適化
量子化パラメータ	22, 27, 32, 37
フレーム予測構造	III... (I-Slice only)
エントロピー符号化	CABAC
符号化フレーム数	10 フレーム
フレームレート	30 fps
ΔQP	基準画素 : -3, 補間画素 : +3

表 2 実験結果

画像	画像サイズ	Δ Bitrate (%)	Δ PSNR (dB)
Cyclists	720p	6.76	0.260
Optis	(1280×720)	4.66	0.216
Raven		4.11	0.161
Sailormen		5.69	0.246
CardToss	HD	8.76	0.332
DinnerTable	(1920×1080)	7.09	0.228
Waves		3.61	0.128
Average		5.81	0.225

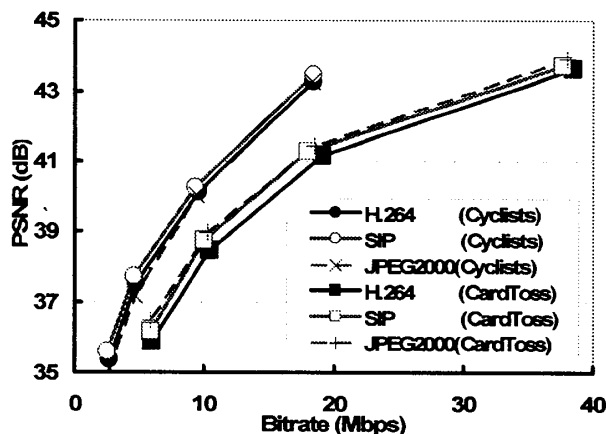


図 7 レート歪曲線

参考文献

- [1] “Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding,” ISO/IEC 14496-10, 2005.
- [2] M. Ouaret, etc., “On Comparing JPEG2000 and Intraframe AVC,” SPIE Applications of Digital Image Processing XXIX, Aug. 2006.
- [3] 二宮, “講座 HDTV 第 3 回 MUSE 方式,” テレビ誌, 42, 8, pp.832-840, Aug. 1988.
- [4] M.Tanimoto, etc., “TAT(time-axis transform) bandwidth compression system of picture signals,” IEEE Trans. Commun, vol.36, no.3, pp.347-354, Mar. 1988.
- [5] Y.Yashima, K.Sawada, “An Extrapolative-Interpolative Coding Method for HDTV Signals,” IEEE Transactions on Communications, Vol.38, No.10, pp.1779-1785, 1990.
- [6] G.Bjontegaard, “Calculation of average PSNR differences between RD-Curves,” ITU-T VCEG, VCEG-M33, April, 2001.
- [7] <http://www.kakadusoftware.com/>