

周辺近傍画素を利用したイントラ予測改善手法の検討

松尾 翔平[†], 高村 誠之[†], 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 : {matsuo.shohei, takamura.seishi, kamikura.kazuto,
yashima.yoshiyuki}@lab.ntt.co.jp

あらまし 最新映像符号化標準 H.264/MPEG-4 Part 10: AVC では、従来の映像符号化標準にある周波数領域での画面内 (イントラ) 予測ではなく、空間画素領域におけるイントラ予測が利用可能となった。イントラ予測に用いる参照画素は、符号化対象ブロックに隣接するブロック内に存在する画素の中で、符号化対象ブロックに対して最近接の画素に限定されていた。参照先の画素がノイズを含んだり、またはその位置にオクルージョンが発生する場合や、画像の特徴として縞模様のような空間周波数が類似する信号が周期的に現れる場合などには、最近傍の画素のみの参照では予測の精度が落ちる場合があると予想される。そこで本稿では、最近傍を含む周辺近傍画素も参照の対象に含めることで、上記の場合に柔軟に対応可能なイントラ予測改善手法を提案する。評価実験の結果、本提案手法により H.264/AVC より最大で約 4 % の符号化効率改善が得られた。

キーワード H.264/AVC、イントラ予測、周辺近傍画素、複数参照ライン

A Study on Improved Intra Prediction Using Multiple Reference Lines

Shohei MATSUO[†], Seishi TAKAMURA[†], Kazuto KAMIKURA[†] and
Yoshiyuki YASHIMA[†]

[†] NTT Cyber Space Laboratories, NTT Corporation

1-1 Hikarino-oka, Yokosuka, Kanagawa 239-0847, JAPAN

E-mail : {matsuo.shohei, takamura.seishi, kamikura.kazuto,
yashima.yoshiyuki}@lab.ntt.co.jp

Abstract: Intra prediction in the spatial domain was newly introduced into a state-of-the-art video coding standard “H.264/MPEG-4 Part 10: AVC”, while that in the frequency domain was available in previous video coding standards such as MPEG-4 Part 2: Visual. Since the reference pixels for intra prediction are restricted to the nearest ones to the block to be encoded, it is expected that the coding efficiency could deteriorate if they include noise or occlusion and if the periodic signals such as a stripe exist in the picture. In this paper, we show the improved intra prediction method which allows utilization of distant adjacent pixels as reference and the rate reduction of our method turns out to be up to 4 % according to the simulation results.

Key words: H.264/AVC, Intra prediction, Distant adjacent pixels, Multiple reference lines

1 はじめに

2003年3月に国際標準化機関ITU-TとISO/IEC JTC1の2つの機関によって、最新映像符号化標準であるH.264/MPEG-4 Part 10: AVC (以下、H.264/AVC)[1]は規格化された。本方式は、放送・通信・蓄積という幅広い範囲を包括し、非常に多種多様なアプリケーションに対応可能である。2006年4月1日からサービスが開始された携帯・移動端末向けワンセグ放送や、次世代蓄積メディアBlu-ray DiscおよびHD DVDにも採用されており、今後さらに普及が期待される映像符号化標準である。

MPEG-2やMPEG-4 Part 2: Visual (以下、MPEG-4)[2]などの従来の映像符号化標準に比べ、H.264/AVCの符号化効率は大きく向上している。例えば、複数参照フレーム、可変ブロックサイズ動き補償、デブロックフィルタ、CAVLC・CABAC等のコンテキスト適応型エントロピー符号化など、有効な符号化ツールが数多く導入されているためである[3]。同一画面内から予測を行うイントラ予測はMPEG-4にも導入されているが、MPEG-4では予測を行う対象が直交変換された変換係数の直流成分と交流成分になり、周波数領域での予測であった。これに対し、H.264/AVCでは空間画素が予測の対象となっており、直交変換を経ないため、より柔軟な予測が可能となっている。参照先のブロックも符号化対象ブロックの左、左上、上の3種類から右上を加えた4種類に増えており、ブロックサイズも8×8のみから、4×4, 8×8, 16×16の3種類から選択可能となった。

上述した機能追加の結果として、H.264/AVCのイントラ予測の性能はMPEG-4より上回り、JPEGやJPEG-2000と同等な性能を有していると報告されている[4]。しかしながら、時間方向の予測を行う画面間(インター)予測に比べると予測の効果は小さく、イントラ予測にて得られるIピクチャの符号量は、インター予測にて得られるPおよびBピクチャに比べると大きくなるのが一般的である。したがって、イントラ予測の更なる改善が望まれている。

イントラ予測の改善手法として、文献[5]では、モード決定のコスト関数を見直すことで予測性能を向上し、文献[6]では、イントラ予測の符号化順序を変更することで改善を施していた。しかし、これらの手法では、予測に利用する画素が符号化対象ブロックに対して最近傍の位置に限定されており、参照される領域がノイズやオクルージョンを含む場合は予測

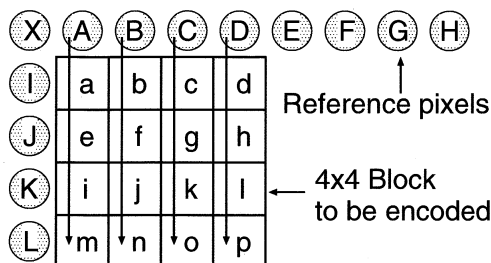


図1: 従来のイントラ予測の参照画素(符号化対象ブロックが4×4輝度信号、垂直予測の場合)

精度の低減を免れない。そこで本稿では、それらの問題を解決するイントラ予測手法を提案する。

2 従来予測手法とその問題点

2.1 H.264/AVCでのイントラ予測

空間画素領域で同一の画面内から予測を行うイントラ予測は、H.264/AVCから導入された。予測はブロック単位で行われ、輝度信号に対して3種類のブロックサイズ(4×4, 8×8, 16×16)が利用可能となっている。また、各ブロックサイズではそれぞれ複数の予測モードが選択可能である。4×4と8×8のブロックサイズの場合は9種類のモード、16×16のブロックサイズの場合は4種類のモードが用意されている。色差信号に対しては8×8のブロックサイズのみが利用可能であり、予測方向は輝度信号での16×16ブロックの場合と同じである。

これらのブロックサイズとモードにおいて、どの場合でも例外なく、イントラ予測に用いる参照画素は、隣接するブロック上の符号化対象ブロックに最近傍の画素のみに限定されている。具体的な例として、符号化対象ブロックが輝度信号の4×4ブロック、垂直予測(予測モード0)の場合を図1に示す。図1に示されるように、符号化対象ブロックの左上にあるブロックからX、上にあるブロックからA, B, C, D、右上にあるブロックからE, F, G, H、そして左にあるブロックからI, J, K, Lに示す画素の値を予測に用いる。垂直予測では、Aの値をa, e, i, mに、Bの値をb, f, j, nに、Cの値をc, g, k, oに、Dの値をd, h, l, pに適用する。以上のように、4×4ブロックの場合は(4×3+1=)13の画素から、8×8ブロックの場合は(8×3+1=)25の画素から予測を行う。16×16は予測

モードの違いにより、右上のブロックを予測に使用しないため、 $(16 \times 2 + 1) = 33$ の画素を用いる。

2.2 従来予測手法の問題点

従来のイントラ予測では、予測に利用する参照画素は全て符号化対象ブロックに隣接する位置に限定されていた。予測を行う際に、なるべく高い相関性を有する対象を利用すれば効果が上がるのは明らかなことであり、例えば、インター予測では一般に符号化対象フレームの直前のフレームを利用すれば効率が高い。しかしながら例外として、直前よりも前のフレームを参照した方が効率が高くなる場合がある。例えば、

- オクルージョンが発生し、直前フレームではうまく参照先が見つからない場合
- 物体が周期運動していて、直前以外のフレームを参照したほうが予測誤差が小さくなる場合

などが挙げられる [7, 8]。

これと同様の概念がイントラ予測にも適用できる。参照先の画素が符号化対象ブロックに最近傍の位置にあれば、空間相関性の観点から、相関性が最も高い(符号化対象ブロックに近ければ近いほど相関が高い)と考えられるため、予測誤差が小さくなると思われる。しかしながら例外として、下記のような場合には必ずしも符号化対象ブロックに近いほど予測誤差が小さくなると言い切れず、逆に大きくなる可能性がある。

- オクルージョンや雑音が符号化対象ブロックの最近傍の位置に存在する場合
- 縞模様などの周期的信号が現れる場合
- 参照する画素は復号画像であるため、符号化対象ブロックに接近するほど、重畳雑音が大きいため、その予測誤差が非常に大きくなる場合
- イントラ予測の方向が9種類と粗いため、最近傍画素よりも良い予測信号が周辺近傍画素から生成できる場合

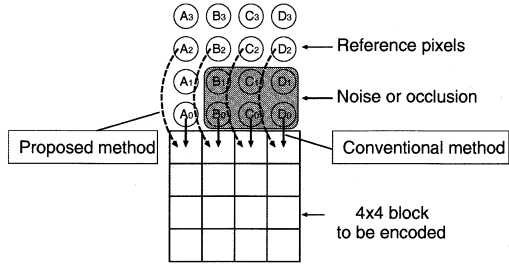


図 2: ノイズなどが含まれる場合

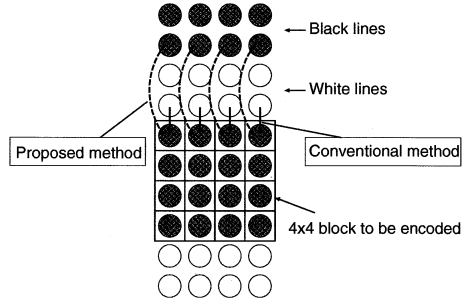


図 3: 周期的に同様の信号が発生する場合

3 周辺近傍画素を利用したイントラ予測手法

3.1 概要

2.2 に示した状況で、予測誤差の増大による符号化効率低減の解決を目的とし、効率的なイントラ予測手法を紹介する。本目的を達成するための方法として、最近傍の位置の画素以外にも、既に符号化済みの画素であれば全て使用可能とすることにより、より予測誤差の少ないイントラ予測を実現する。

図 2 に示すように、符号化対象ブロックの参照画素に雑音が含まれている場合を想定する。この場合、図 2 内における A_2, B_2, C_2, D_2 のように、距離が離れている位置から予測誤差を低減し得る画素を利用することで、イントラ予測の効率向上が期待できる。また、図 3 に示すように、画像信号の特徴として、黒と白の縞が交互に現れる場合を想定する。この場合も同様に距離が離れている位置から予測誤差を低減し得る画素を利用し、イントラ予測の効率を上げられる。他の場合も同様に対応できる。

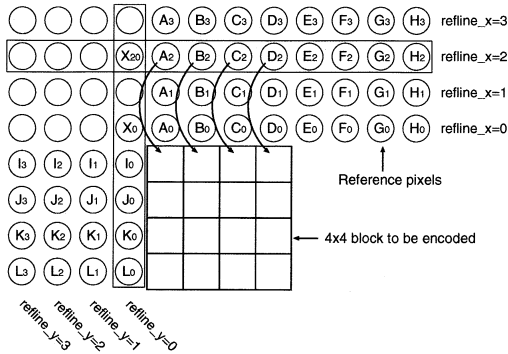


図4: 提案イントラ予測手法 (m=n=4 の場合)

3.2 複数参照ラインによるイントラ予測

最近傍画素を含めた周辺近傍画素も参照に利用する本提案手法の処理について説明する。例として、図4に符号化対象ブロックのサイズが4×4の場合を示す。従来利用していた最近傍画素のうち、 X_0 から H_0 を含む水平方向の画素集合、および I_0 から L_0 を含む垂直方向の画素集合を参照ライン (Reference Line) と定義し、それぞれ本稿では $refine_x$ と $refine_y$ と呼び、符号化対象ブロックに近いラインから順番に $0, 1, 2, \dots$ とする。また、それぞれの取りうる範囲を $0 \leq refine_x \leq m, 0 \leq refine_y \leq n$ とする。図4は $m=n=4$ の場合を示している。なお、従来のH.264/AVCの参照方法は $refine_x=refine_y=0$ を用いた場合に相当する。また、下記処理ステップにおいて、式(1)に示すコスト関数 J を用いる。

$$J = SSD + \lambda_{MODE} \times (R_{BLOCK} + R_{REF}) \quad (1)$$

ここで、SSD(Sum of Squared Difference) は二乗誤差和、 λ_{MODE} はモードに関するラグランジュ乗数、 R_{BLOCK} は符号化対象ブロックを符号化した際の発生符号量、 R_{REF} は参照ラインの推定符号量をそれぞれ示す。 R_{REF} は、事前に複数の画像から得た平均的な $refine_x$ と $refine_y$ の生起頻度を基に算出した。

提案手法の処理は以下のステップから構成される。

Step 1: 最小コスト関数 J_{min} を十分大きな値で初期化する。

Step 2: (mode, $refine_x$, $refine_y$) のそれぞれに初期値を代入して、ループを開始する。

Step 3: $refine_x$ および $refine_y$ から一意に定まる13の参照画素 (A から L および X) を取得する。例えば、図4に示されるように $refine_x=2$ と $refine_y=0$ の場合は、 $A=A_2, B=B_2, \dots, L=L_0$ 、そして $X=X_{20}$ と

なる。取得した参照画素から予測信号を作成し、符号化を行い、コスト関数 J を得る。

Step 4: もしも $J < J_{min}$ を満たす場合は、(best_mode, best_refine_x, best_refine_y, J_{min}) をそれぞれ (mode, $refine_x$, $refine_y$, J) に更新する。満たさない場合は、更新処理を行わない。

Step 5: Step 2 における3つのループを、取り得る全ての (mode, $refine_x$, $refine_y$) に関して繰り返す。本処理により、最小コストを実現する (best_mode, best_refine_x, best_refine_y) の組み合わせを取得できる。

Step 6: 3つの最良パラメータを符号化し、そのパラメータを利用して符号化対象ブロックの符号化を実行する。

Step 1 から 6 の処理により、符号化対象ブロックに最近接の画素を含めた周辺近傍画素もイントラ予測に利用可能となる。参照可能な画素数は従来の13画素に限定された状態から、本提案手法により $8m+4n+mn$ 画素に増える。結果として、2.2に示された状況に対応可能となり、予測誤差をより小さくできる。ここでは、輝度信号の4×4ブロックについての処理を述べたが、言うまでもなく、本提案手法は8×8, 16×16ブロックサイズや色差信号についても適用可能である。

4 評価実験

4.1 実験内容

本提案手法をJM11.0[9]をベースとした次世代符号化向け参照ソフトウェアKTA(Key Technical Area)[10]へ実装を行い、そのレート歪み性能を評価した。CIF(352×288)サイズの3種類の画像 (Bus, Foreman, Mobile) を30fps、符号化フレーム枚数1枚、イントラ予測ブロックサイズ3種類 (4×4, 8×8, 16×16)、量子化パラメータ20種 (8, 10, 12, ..., 46)、CABACとRD最適化を使用、という条件下で本実験を実施した。比較する方法を下記に示す。

- 従来手法 (JM11.0 と同等)
- 提案手法1 (オーバヘッド加味前)
- 提案手法2 (オーバヘッド加味後)

ここで、オーバヘッドは参照元情報 ($refine_x$ と $refine_y$) の推定符号量を示しており、フレーム単位での0次エントロピーから算出している。

		refine_x=0	refine_x=1	refine_x=2	refine_x=3
refine_y=0	p_0	p_1	p_3	p_5	
refine_y=1	p_2	N/A	N/A	N/A	
refine_y=2	p_4	N/A	N/A	N/A	
refine_y=3	p_6	N/A	N/A	N/A	

図 5: p_i の定義

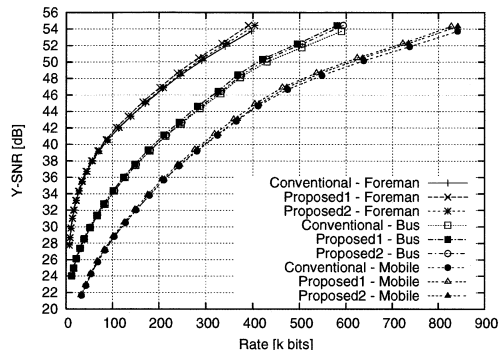
本実験では $m=n=4$ とし、 $1 \leq \text{refine}_x$ かつ $1 \leq \text{refine}_y$ は常に非選択とした。これは予備実験から上記範囲の選択される確率が 1 割未満となっており、本提案手法の性能に対して影響をほとんど与えないと判断したためである。

選ばれる参照ラインに関する事象を図 5 に示す。 p_i は選択された参照ラインの組み合わせを示しており、例えば p_1 ならば、 $\text{refine}_x=1$ と $\text{refine}_y=0$ が選択されたことを示し、 p_0 は $\text{refine}_x=\text{refine}_y=0$ となり、従来手法に相当する。オーバーヘッドを考慮しない場合 (提案手法 1)、 $p_i(i \neq 0)$ が選択されれば、常に従来手法の性能を上回ると予想される。レート歪み性能と同時に、この参照ラインの選択率も調査を行った。

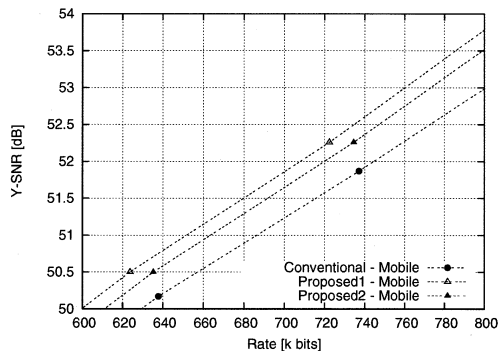
4.2 実験結果と考察

レート歪み性能の結果を図 6 に示す。オーバーヘッドを考慮しない提案手法 1 は全てのビットレートにおいて従来手法を上回る結果となっていた。オーバーヘッドを考慮する提案手法 2 については、高ビットレート (QP=8~24 前後、図 6(b) 参照) において従来手法よりも上回り、約 4% の符号量削減が確認できたが、低ビットレート (QP=30 以上、図 6(c) 参照) になると従来手法とほぼ同等になった。また、高ビットレートのほうが、提案手法 1 も含めて利得向上率が大きく、低ビットレートになると小さくなった。

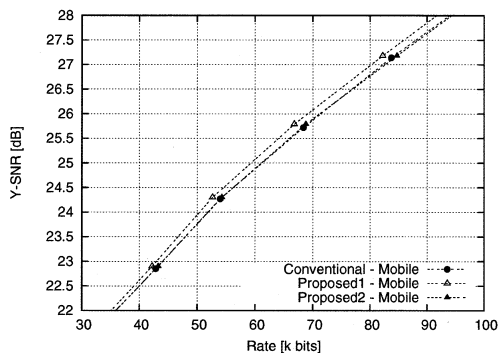
高ビットレートにて提案手法の性能が向上する理由としては、オーバーヘッドが QP に依らずほぼ一定であることが挙げられる。ゆえに、高ビットレートではオーバーヘッドの全符号量に占める割合が相対的に小さくなる。



(a)



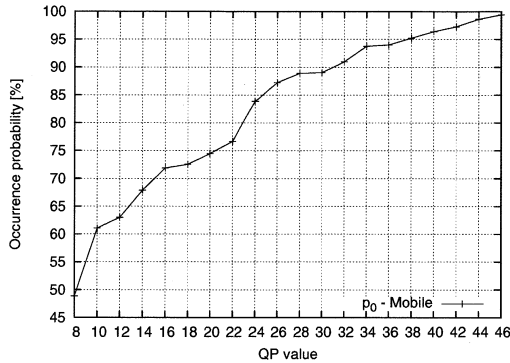
(b)



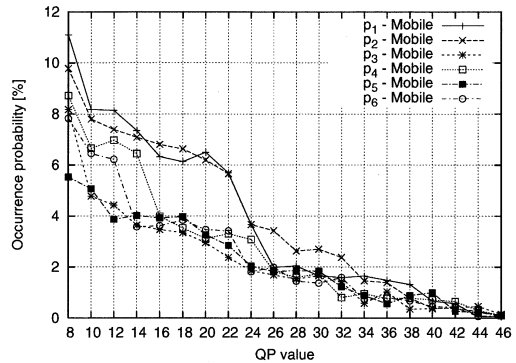
(c)

図 6: レート歪み性能 (Bus, Foreman, Mobile) (a) 概観、(b) 高ビットレート拡大図: Mobile、(c) 低ビットレート拡大図: Mobile

p_i の確率分布の推移の結果を図 7 に示す。上述した内容を裏付ける結果が、図 7(a) に示されている。QP の増加に伴って p_0 (従来手法に相当) の選択率が上昇している。この原因の 1 つとしては、画像がぼや



(a)



(b)

図 7: p_i の推移 (Mobile、4×4 ブロック) (a) $i=0$ の分布、(b) $i \neq 0$ の分布

けてくるために、参照ラインを変えても同じような画素値が出現し、距離の離れた画素を参照する必要がなくなるためと考えられる。

p_i ($i \neq 0$) の選択率に関しては、図 7(b) に示されるように、QP が上がると p_i の選択率は下がる傾向にあった。 $p_i > p_j$ ($i < j$) という傾向も観察される。 p_j ($j > 2$) については、距離が離れても選択率に大きな差が生じないのは興味深い結果である。

5 おわりに

本稿では、イントラ予測に関して、符号化対象ブロックに対する最近接画素を含んだ周辺近傍画素も予測参照の対象にすることで、予測精度の向上したイントラ予測手法を提案した。評価実験の結果、従来手法に比べて最大で約 4% の符号量削減を確認した。今後は、オーバヘッドの削減手法や効率的な参照ラインの選択手法の検討を行う予定である。

参考文献

- [1] ISO/IEC 14496-10:2003, Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding, Dec. 2003.
- [2] ISO/IEC 14496-2:2004, Information technology – Coding of audio-visual objects – Part 2: Visual, May 2004.
- [3] A. Tamhankar et al., “An overview of H.264/MPEG-4 Part 10,” in Proceedings of EURASIP, July 2003, vol. 1, pp. 1–51.
- [4] A. Aravind et al., “Quality and complexity comparison of H.264 intra mode with JPEG2000 and JPEG,” in Proceedings of ICIP, Oct. 2004, vol. 1, pp. 525–528.
- [5] C. H. Tseng et al., “Improved and fast algorithms for intra 4×4 mode decision in H.264/AVC,” in Proceedings of ISCAS, May 2005, vol. 1, pp. 2128–2131.
- [6] Z. Gang et al., “The intra prediction based on sub block,” in Proceedings of ICSP, Aug. 2004, vol. 1, pp. 467–469.
- [7] Y. Su and M. T. Sun, “Fast multiple reference frame motion estimation for H.264,” in Proceedings of ICME, June 2004, vol. 1, pp. 695–698.
- [8] A. Chang et al., “A novel approach to fast multi-frame selection for H.264 video coding,” in Proceedings of ICASSP, Apr. 2003, vol. 3, pp. 413–416.
- [9] “JM reference software ver11.0,” Jan. 2007, <http://iphome.hhi.de/suehring/tml/download/>.
- [10] “KTA reference software ver1.2,” Nov. 2006, <http://iphome.hhi.de/suehring/tml/download/KTA>.