

エッジ方向に基づく適応イントラ予測に関する検討

A Study on Adaptive Intra Prediction based on the Edge Direction

筑波 健史 徳毛 靖昭 青野 友子
Takeshi Tsukuba Yasuaki Tokumo Tomoko Aono

1. はじめに

国際標準化機関 ITU-T と ISO/IEC により、2003 年に動画像符号化方式の国際標準として H.264 / MPEG-4 Part 10: AVC(以下、H.264 / AVC)[1]が規格化された。H.264 / AVC は、現在、放送・通信・蓄積の各方面において利用され始め、今後の更なる普及が期待されている。一方、近年はますます動画像の高精細化が進み、動画像のデータ量が増加しているため、蓄積・伝送のために動画像符号化方式の符号化効率向上が必要である。

本稿では、符号化効率の向上を目指し、予め予測方向を特定方向に偏らせた予測セット（予測方法の集合）を複数定義しておき、M×M画素ブロック単位で予測セットを切り替えるイントラ予測の検討を行った。評価実験の結果、H.264 / AVC と比較して本方式による符号量削減の効果を確認したので報告する。

2. 従来手法とその問題点

H.264 / AVC の符号化モードには、符号化済み隣接ブロックの局所復号画像から原画像を予測する N×N 画素ブロック(N=4, 8, 16)のイントラ予測がある。図 1 に 4×4 画素ブロック、8×8 画素ブロックのイントラ予測において規定されている予測方法を示す。上記イントラ予測では、DC 予測（モード 2）、ほぼ等角度に割当てられた 8 種の方向に対応する予測（モード 0, 1, 3~8）、計 9 通りの予測方法が定義されている。上記モード 0, 1, 3~8 は、特定方向のエッジを有効に予測可能な予測方法である。

定性的には、エッジの方向に沿って予測することで、予測誤差を低減させ、符号量を削減し、画質が向上すると考えられる。

しかし、上記イントラ予測は、8 種類の予測方向に限定されており、原画像のエッジ方向がいずれの予測方向とも異なる場合、原画像と予測画像の残差には新たなエッジ輪郭が発生し、高周波成分が増加するため、予測効率が低下する。そのため、予測効率を向上させるためには、より多方向のエッジを予測可能なイントラ予測が必要である。

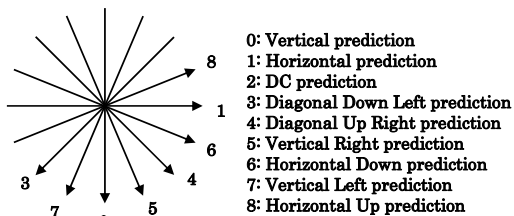


図 1 : H.264/AVC におけるイントラ予測(4×4, 8×8)

シャープ (株) 技術本部 先端映像技術研究所,
Advanced Image Research Laboratories, Corporate R&D Group, Sharp Corporation

3. エッジ方向に基づく適応イントラ予測

3.1 概要

より細かい角度からの予測を可能とするため、本方式では、画像の空間領域におけるエッジが特定方向へ偏る性質を利用して、予め特定方向の予測を重視した予測セットを複数定義しておき、M×M画素ブロックサイズ毎に予測セットを切り替えてイントラ予測を行う。本方式は、4×4, 8×8 画素ブロックにおける予測方向を単純に従来の 9 種類から増やす場合と比較して、予測方法を表すのに要する符号量の増加、および予測方法の決定に要する演算回数の増加を抑えることが可能である。

例えば、図 2 (a)のような横縞模様の画像には、水平方向(0[deg])のエッジが多いため、図 3 (a)に示すエッジ方向分布となる。エッジ方向が水平方向へ集中する場合は、図 4(a)に示す水平方向の予測を細かく定義した予測セットを用いてイントラ予測を行う。

また、図 2 (b)のような縦縞模様の画像には、垂直方向(90[deg])のエッジが多いため、図 3 (b)に示すエッジ方向分布となる。エッジ方向が垂直方向へ集中する場合は、図 4(b)に示す垂直方向の予測を細かく定義した予測セットを用いてイントラ予測を行う。ここで、エッジ方向分布とは、小矩形領域(4×4 画素ブロック)毎の支配的なエッジ方向の出現頻度分布を指す。また、エッジ方向の算出には[4]の手法を利用した。

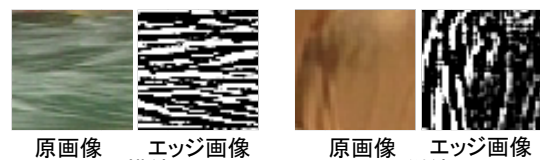
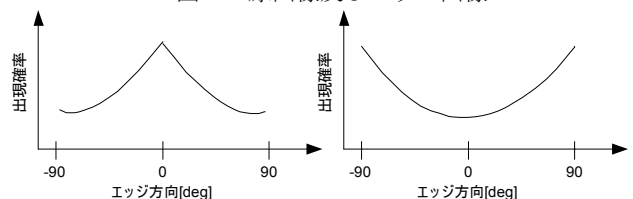


図 2 : 原画像及びエッジ画像



(a) 水平方向にエッジが集中 (b) 垂直方向にエッジが集中

図 3 : エッジ方向分布の例

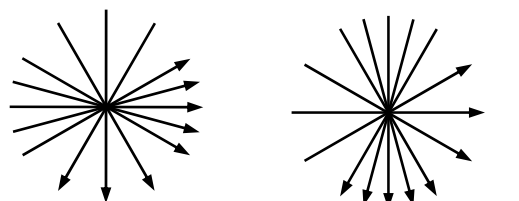


図 4 : 予測セットの例

3. 2 符号化処理の流れ

本方式における符号化処理の流れについて、予測セットの切り替え単位を 16×16 画素ブロックとし、 4×4 画素ブロックのイントラ予測を行う場合を例に説明する。

本方式では、予め定めた複数の予測セットの中より、 16×16 画素ブロック単位に式(1)に示す符号化コスト $J(set)$ を各予測セットについて算出し、その符号化コストを最小化する予測セット $best_set$ を用いて符号化を行う。ここで、式(2)はある予測セット set 適用時、 16×16 画素ブロックに含まれる 4×4 画素ブロック bk を最良予測モード $best_mode(bk)$ で符号化する場合の符号化コスト $J(bk|set)$ を表す。また、 $D(bk|set)$ 、 $R(bk|set)$ は 4×4 画素ブロックの予測誤差、発生符号量を表し、 R_{set} は予測セット set を表す情報の符号量、 λ_{mode} は予測モード、 λ_{set} は予測セットのラグランジュ乗数をそれぞれ表す。

$$J(set) = \lambda_{set} \times R_{set} + \sum_{bk \in block} J(bk|set) \quad (1)$$

$$J(bk|set) = D(bk|set) + \lambda_{mode} \times R(bk|set) \quad (2)$$

以下に、本方式における詳細な符号化処理の流れを示す。
Step1: 最小符号化コスト関数 J_{min} を十分大きな値で初期化する。

Step2: 予測セット set を用いて、 16×16 画素ブロックに含まれる全 4×4 画素ブロックを符号化する場合の符号化コスト $J(set)$ を求める。

Step3: もし、 $J(set) < J_{min}$ を満たすならば、 $\{J_{min}, best_set, best_mode(bk) (bk=0,1,2,\dots,15)\}$ を、それぞれ $\{J(set), set, mode(bk) (bk=0,1,2,\dots,15)\}$ に更新する。それ以外ならば、更新をしない。

Step4: Step2~Step3 の処理を全ての予測セットに関して繰り返す。

Step5: Step2~Step4 の処理を経て取得された 16×16 画素ブロックの符号化コストを最小化するパラメータ $\{best_set, best_mode(bk), (bk=0,1,2,\dots,15)\}$ を用いて符号化を行う。

上記 Step1~Step5 までの処理により、 16×16 画素ブロック毎にエッジ方向の分布に応じた予測セットを用いて符号化することが可能となる。

4. 評価実験

本方式を H.264 / AVC 参照ソフトウェア [2] 上の輝度成分のイントラ予測 (4×4 画素ブロック) に実装し、評価実験を行った。本実験で用いた予測セットを表 1 に、実験条件を表 2 に、および実験結果を表 3 に示す。なお、H.264/AVC に対する符号量削減量 $\Delta RATE[\%]$ と画質 (PSNR) の利得 $\Delta PSNR[\text{dB}]$ は BD-SNR[3] をもとに計算した。

表 3 の実験結果より、QCIF 及び CIF の計 7 シーケンスにおいて、平均 3.38[%] の符号量が削減された。符号量削減幅が最も大きい 6.84[%] であった Foreman はエッジの多い画像である。一方、符号量削減幅が最も小さい 1.34[%] であった Silent はエッジの少ない画像である。

以上から、H.264 / AVC と比較して、本方式はエッジの多い画像において有効な予測方式であることが確認できた。

表 1 : 予測セットの構成

		予測セット								
		0	1	2	3	4	5	6	7	8
シーケンス	0	90	90	90	90	90	90	90	90	90
	1	0	0	0	0	0	0	0	0	0
	2	DC	DC	DC	DC	DC	DC	DC	DC	DC
	3	-45	15	75	16.8	73.2	18.6	71.4	20.4	69.6
	4	45	-15	-75	-16.8	-73.2	-18.6	-71.4	-20.4	-69.6
	5	67.5	30	60	33.6	56.4	37.2	52.8	40.7	49.3
	6	22.5	-30	-60	-33.6	-56.4	-37.2	-52.8	-40.7	-49.3
	7	-67.5	60	-30	67.2	22.8	74.4	15.6	81.5	8.5
	8	-22.5	-60	30	-67.2	-22.8	-74.4	-15.6	-81.5	-8.5

※DC : DC 予測, それ以外の数値 : 予測方向[deg]

表 2 : 実験条件

プロファイル	Baseline(I only)
量子化パラメータ	22,27,32,37(固定)
RD最適化	ON
フレーム数	150
補間方法	線形補間

表 3 : 実験結果

シーケンス	画像サイズ	$\Delta PSNR[\text{dB}]$	$\Delta RATE[\%]$
Container	QCIF	0.203	2.56
Foreman	QCIF	0.540	6.84
Paris	QCIF	0.286	2.77
Silent	QCIF	0.094	1.34
Foreman	CIF	0.321	5.20
Mobile	CIF	0.225	2.00
Tempete	CIF	0.268	2.93

5. まとめ

本稿では、エッジ方向に基づく適応イントラ予測に関する基礎検討を行った。エッジ方向の偏りに応じた予測セットを考慮する必要があると考え、予め予測方向を特定方向に偏らせた予測セットを複数定義しておき、 $M \times M$ 画素ブロック単位で予測セットを切り替えるイントラ予測を検討した。評価実験より、本方式は H.264 の Baseline と比較して、平均 3.38[%] の符号量削減の効果を得られ、特にエッジの多い画像において予測効率の向上が期待できる。

参考文献

- [1] ITU-T Rec. H.264 | ISO/IEC 14496-10, "Advanced video coding for generic audio-visual services," May 2003
- [2] H.264/AVC Reference Software (JM ver12.0), <http://iphome.hhi.de/suehfig/tml/>
- [3] G. Bjontegaard, "Calculation of average PSNR differences between RD-Curves," ITU-T Q.6/SG16 VCEG, VCEG-M33, April 2001
- [4] Asker M. Bazen, Sabish H. Gerez, "Systematic Methods for the Computation of the Directional Fields and Singular Points of Fingerprints," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.24, No.7, July 2002.