

RI-005

メディアン予測の最適制御に基づく MPEG-4 AVC/H.264 動きベクトル予測の性能改善

Performance improvement of motion vector prediction based on MPEG-4 AVC/H.264 by optimal control for median prediction

吉野 知伸 内藤 整 酒澤 茂之 小池 淳
Tomonobu Yoshino Sei Naito Shigeyuki Sakazawa Atsushi Koike

1. まえがき

IPTV など近年の HDTV 放送サービスにおいては、優れた圧縮性能を有する H.264 が一般に適用されつつある。しかしながら、圧縮符号化性能のさらなる追求により、ADSL やモバイル回線での HDTV 配信など、HDTV サービスのさらなる普及促進を牽引できるため、H.264 を凌ぐ次世代符号化方式の確立が求められている。H.264 では、インター符号化において、有効な符号化ツールが数多く定義されており、イントラ符号化に比べ高い圧縮性能を有している。このため、H.264 の優れた符号化性能はインター符号化の寄与によるところが大きい。インター符号化を構成する符号化ツールの中でとりわけ、動きベクトル予測の効果は顕著である。同ツールにより動きベクトルの効率的な符号化が可能であるが、総符号量に占める動きベクトルの符号量の割合は依然として高く、HDTV を対象とした低ビットレート符号化において特に顕著である。H.264 の動きベクトル予測手法(以下、標準手法)では、予測ベクトルに対する参照ベクトルの適否を考慮していない。このため、予測ベクトルに不適切な参照ベクトルが含まれるとき、同手法の予測性能が低下する懸念がある。以上を踏まえ、H.264 の圧縮性能改善に向けて有望な技術要素として動きベクトル予測に着目することとする。対象とするアプリケーションは、HDTV の放送サービスに限定し、比較的低ビットレートでの符号化を想定する。

以下に本論文の構成を示す。2.では動きベクトル予測手法について、標準手法の導入を行った上で、同手法が抱える問題について説明する。3.では同問題を H.264 の独自拡張により解消するためのアプローチを提案する。4.では符号化シミュレーション実験を通じ、提案手法の性能評価を行う。5.において結果の詳細な分析を行い、提案手法を構成する技術要素の効果配分やこれらが有効に機能する条件を明らかにする。

2. 従来の動きベクトル予測手法

2.1 メディアン予測法

H.264 では、メディアン予測法に基づく動きベクトル予測生成手法が採用されている。図1に H.264 の予測ベクトル生成手法を示す。

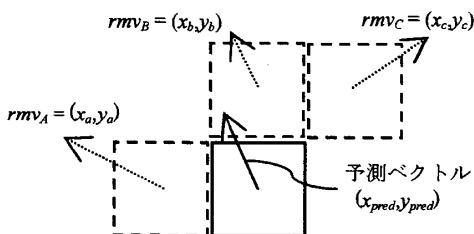


図1 標準手法の動きベクトル予測

メディアン予測法では、処理ブロックの左、上、右上の符号化済みブロックを参照ブロックとし、参照ベクトルのメディアンにより、予測ベクトルを得る。なお、以降、処理ブロックに対して左、上、右上の参照ブロックをそれぞれ、参照ブロック A, B, C, 各参照ブロックにおける参照ベクトルをそれぞれ、 rmv_A , rmv_B , rmv_C と呼ぶ。

2.2 標準手法における改善の余地

2.2.1 特定絵柄における予測性能低下

標準手法では、処理ブロックに対する参照ベクトルの適否を判断しておらず、不適切な参照ベクトルが含まれる場合、メディアン予測による予測ベクトルの精度は低下する。特に、処理ブロックもしくは参照ブロック B, C 付近に、異なる動きを持つオブジェクトのエッジ成分が水平方向に含まれるとき、 rmv_B , rmv_C は、同エッジを挟んで、処理ブロックの動きとは異なるものであるため、予測ベクトルとして不適切な候補となり得る。また、 rmv_B と rmv_C は類似したベクトルであるため、メディアン予測の結果は、 rmv_B , rmv_C の成分が支配的となり、予測ベクトル精度の低下が避けられない。以下、同問題を水平方向エッジ問題と呼ぶ。

ここで、水平方向エッジ問題が発生するとき、参照ブロック A と処理ブロックでは、絵柄が類似し、かつ類似した動きベクトルが得られる。したがって、処理ブロックでは、 rmv_A を予測ベクトルとすることで、高精度な予測ベクトルが得られ、水平方向エッジ問題は解決すると期待される。

文献[1]では、3つの参照ベクトルのうちの1つを予測ベクトルとして選択し、選択した参照ベクトルの識別情報を符号化する手法が提案されている。同手法により、水平方向エッジ問題を解決する予測ベクトルの生成が期待される。しかしながらすべてのブロックにおいて識別情報の符号化を伴うため、前記問題に該当しない箇所では、かえって符号化効率の低下が懸念される。

以上より、水平方向エッジ問題に対して、高精度な予測ベクトルを生成し、かつ、予測ベクトル生成における付加情報の増加を効率的に抑制することで、動きベクトル予測性能の改善が期待される。

2.2.2 候補数増加による予測性能改善

H.264 ではメディアン予測により予測ベクトルを得ている。しかしながら、同予測では、高々3つの参照ベクトルを入力としているに過ぎず、粗いサンプルに対する中間的な値を予測値としている。したがって、同予測値の精度は、必ずしも全てのケースにおいて厳密性が保証されるものではない。

上記課題を解決する手段の1つとして、予測ベクトル候補の数を増やし、メディアン演算の対象をブロック単位で適応的に選択する枠組みを採用することが挙げられる。新たな予測ベクトル候補には、生成において新たな付与情報が発生しないことが求められる。また、付加情報の発生を

抑える観点から、メディアン演算に与える候補ベクトルの判別をデコーダにて行えることが理想である。

3. 提案手法

動きベクトルの符号化に要する符号量の削減を目的とし、2.2 で示した課題を解決する手法を提案する。3.1 および 3.2 では、それぞれ 2.2.1 および 2.2.2 に基づく予測ベクトル生成手法を示す。3.3 では、標準手法を拡張して、両手法を適用する処理の流れを示す。

3.1 水平エッジ対処手法

以下、マクロブロックを MB、処理ブロックを含むマクロブロックを処理 MB と呼ぶ。水平方向エッジ問題が発生するブロックでは、 rmv_B , rmv_C に対して、 rmv_A が乖離する。提案手法では、処理 MB に隣接する符号化済み MB における動きベクトルを評価ベクトルとし、評価ベクトル間の距離を評価することで、水平方向エッジ問題が発生する危険性を推定する。評価ベクトルには、当該 MB 内で絶対値が最大の動きベクトルを代入する。処理 MB に対して、左、上、右上の MB が有する評価ベクトルをそれぞれ評価ベクトル A, B, C とする。提案手法の処理手順を示す。

る動きベクトルを評価ベクトルとし、評価ベクトル間の距離に基づき、上記ケースの該当有無を判定する。図 2 に評価ベクトルの決定例を示す。処理 MB の左に隣接する符号化済み MB では、最も右上に位置するブロックの動きベクトルを評価ベクトルとする。上および右上に隣接する符号化済み MB では、最も左下に位置するブロックの動きベクトルを評価ベクトルとする。提案手法の処理手順を以下に示す。

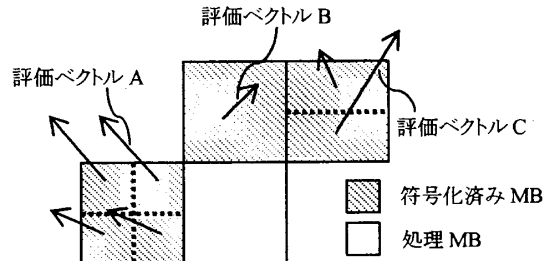


図2 評価ベクトル決定例

Step1. 処理 MB の左に隣接する符号化済み MB で、水平方向エッジ問題の危険性があると判断され、かつ、左側参照ベクトルを予測ベクトルとしている場合、Step3 へ。それ以外であれば、Step2 へ。

Step2. 評価ベクトル A から C について、互いの差分が成すベクトルの大きさを、それぞれ $|AB|$, $|BC|$, $|CA|$ とする。このとき、 $|AB|$ および $|CA|$ の小さい方の値が、 $|BC|$ に重み係数 w を乗じた値より大きいとき、Step3 へ。それ以外であれば、標準手法に基づき予測ベクトルを生成する。

Step3. 当該 MB では、水平方向エッジ問題が発生する危険性があるとみなし、予測ベクトルとして、 rmv_A もしくは標準手法の予測ベクトルを用いてそれぞれ符号化し、R-D 最適化法[2]に基づき予測ベクトルを決定する。ここで決定された予測ベクトル生成手法に従い符号化処理を進め、同時に、同生成手法を識別するため 1bit の付加情報を符号化結果に付与する。

Step1. 評価ベクトル A から C について、互いの差分が成すベクトルの大きさを、それぞれ $|AB|$, $|BC|$, $|CA|$ とする。 $|AB|$ から $|CA|$ の分散値が閾値 Th_{low} より大きく、かつ Th_{high} より小さいとき、Step2 へ。それ以外であれば、標準手法に基づき予測ベクトルを生成する。

Step2. 当該 MB では、新たな予測ベクトル候補が有効であるとみなし、予測ベクトルとして、標準方式、ならびに平均ベクトルのメディアン演算による方式を用いてそれぞれ符号化し、R-D 最適化法に基づき適切なものを選択する。ここで決定された予測ベクトル生成手法に従い符号化処理を進め、同時に、同生成手法を識別するため 1bit の付加情報を符号化結果に付与する。

3.2 平均ベクトル手法

提案手法では、2.2.2 の要件を満たしつつ、参照ベクトルの特徴を継承する新たな予測ベクトル候補として、3 つの参照ベクトルから対になるものを 3 組定義し、各組の平均値を新たな予測ベクトル候補 (以下、単に平均ベクトル) として扱い、計 3 つの平均ベクトルをメディアン演算の対象とできるようにする。簡単のため、メディアン演算の対象は、標準手法による参照ベクトル、平均ベクトルのいずれかであり、これらの混在によるケースは想定しないこととする。

ここで、平均ベクトルが、予測ベクトル候補として有効ではないケースについて考える。1 つの参照ベクトルが他の 2 つの参照ベクトルと乖離している場合、乖離する参照ベクトルを含む 2 つの平均ベクトルは、どの参照ベクトルからも乖離する。このとき、平均ベクトルに対するメディアン演算の結果は、いずれの参照ベクトルからも乖離する結果となる。また、全ての参照ベクトルが密集する場合、平均ベクトルの結果は、標準手法の結果とほぼ等価となる。

ここから上記 2 つのケースでは、平均ベクトルは有効な予測ベクトル候補と成り得ないため、候補から除外する。提案手法では、処理 MB に隣接する符号化済み MB にお

3.3 処理の流れ

標準手法を拡張して 3.1 および 3.2 の提案手法を実現する。提案手法の処理の流れを図 3 に示す。

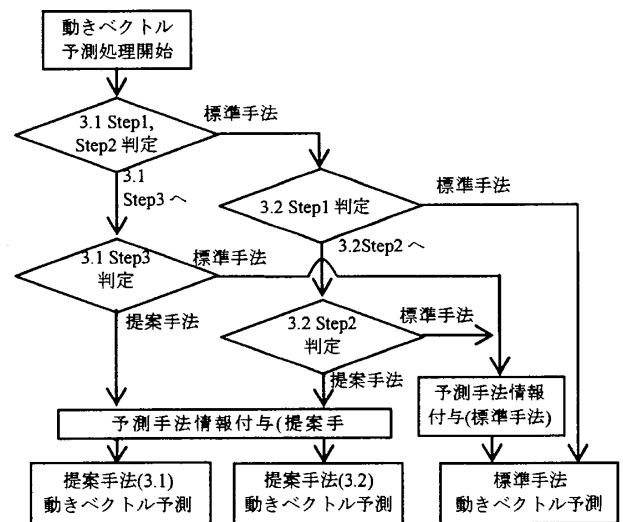


図3 提案手法の処理フロー

提案手法では、最初に 3.1 に基づき、水平方向エッジ問題に起因する予測性能低下の危険性を検知する。同危険性がある場合、予測ベクトル生成手法を表現する情報を付与する。同問題の危険性が無い場合、3.2 に基づき、予測ベクトル生成における平均ベクトルの適用可否を判断する。適用可能である場合、動きベクトル予測生成手法を表現する情報を付与する。適用不可である場合は、予測ベクトル生成手法を表現する情報を付与することなく、標準手法に基づく予測ベクトルを生成する。ここで、提案手法では、3.1 および 3.2 の判定は直列的かつ排他的に行われるため、共通の 1bit の情報で各手法の適用有無を表現することが可能である。また、デコーダでも、3.1 Step1, Step2 および 3.2 Step1 と同様に判定することで、1bit の情報で各手法の適用有無を判断可能である。

4. 符号化実験

提案手法の符号化性能を評価するために、符号化実験を行った。H.264 リファレンスエンコーダ(JM)[3]12.1 に、3.3 で示した提案手法を実装した。評価用素材として、ITE 標準動画 7 シーケンス(各 100 フレーム)を用いた。評価用素材の選択においては、提案手法の効果を明確にするため、カメラワークを伴うなど、動き特性の種類を多く網羅できるように配慮した。表 1 に各素材の特徴を示す。表 2 に符号化条件を示す。なお、提案手法のパラメータは、それぞれ、 $w=1.0$, $Th_{low}=5$, $Th_{high}=100$ とした。

表 1 評価用素材の特徴

素材	特徴
(a) Yachting	ヨットの動きに追従したカメラワーク(パン), 単調な動き
(b) Skyscrapers	街並みの望遠映像, 単調なカメラワーク(パン)
(c) Bronze with credits	緩やかなカメラワークとともに, 複数の字幕が水平方向に流れる
(d) Streetcar	固定カメラ映像, 左右に車や路面電車が横切る
(e) Driving	走行中の車内から前方の様子を写した映像, プレが含まれる
(f) Church	教会の建物を映した映像, 細かいテクスチャが含まれる, チルトカメラワーク
(g) Buddhist images	コントラストが乏しい映像, 斜め下方向への緩やかなカメラワーク

表 2 符号化条件

GOP 構造	IPPPP (以降, P のみ)
プロファイル	High Profile
エントロピー符号化	CABAC
R-D 最適化	有効
QP	30,33,36,39 (固定)
参照フレーム枚数	1
サーチレンジ	$\pm 64 \times \pm 64$
動き探索	UM Hexagon

表 3 に、符号化実験の結果を示す。同表では、符号量削減率について、BD 法[4]に基づく評価結果を示している。比較のため、標準手法として、JM12.1 を用いた結果を併せて示す。結果より、提案手法を用いることで、全素材の平均で約 3.53%、最大で 12.61% の符号量削減効果が得られた。最大の効果が得られた素材(b)の R-D 特性を図 4 に示す。

表 3 符号化結果

Test sequence	Bit-reduction ratio[%]
(a) Yachting	2.32
(b) Skyscrapers	12.61
(c) Bronze with credits	4.36
(d) Streetcar	1.72
(e) Driving	0.97
(f) Church	0.73
(g) Buddhist images	2.00
Average	3.53

PSNR[dB]

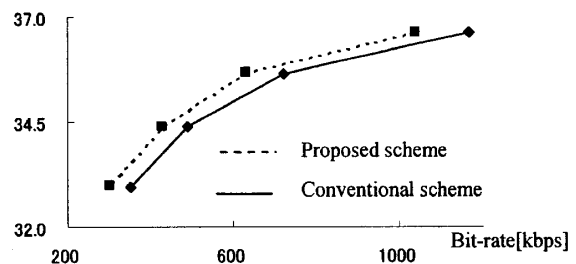


図 4 R-D 特性(素材(b))

表 4 に動きベクトルの符号化に要する符号量の計算結果を示す。同表では、フレーム単位での、総符号量に対する動きベクトル符号量の割合(標準手法 MV 率および提案手法 MV 率)、標準手法に対する提案手法による動きベクトル符号量の削減率(MV 削減率)、標準手法に対する提案手法による総符号量の削減率(Total 削減率)を示す。同結果は、QP=39 における先頭 9 フレーム (P ピクチャ) の結果である。結果より、提案手法により、動きベクトルの符号量が最大で約 24%削減されたことが分かる。

表 4 動きベクトル符号量削減[%]

素材	標準手法 MV 率	提案手法 MV 率	MV 削減率	Total 削減率
(a)	21.98	21.59	6.19	4.50
(b)	28.17	26.88	23.22	19.55
(c)	33.16	28.24	23.89	10.61
(d)	24.76	24.63	6.19	5.68
(e)	18.21	18.47	0.56	1.93
(f)	27.79	26.46	8.22	3.61
(g)	42.42	41.03	11.65	8.65

5. 考察

5.1 各改善要素の寄与

実験により確認された符号化性能改善に対し、3.1 および 3.2 の各手法の寄与度を確認する。表 1 の符号化条件の下、3.1 および 3.2 で示した手法をそれぞれ単独で適用した結果を表 5 に示す。比較のため標準手法として JM12.1 を用いており、結果は JM12.1 に対する平均符号量削減を表わしている。素材は特徴的なカメラワークを有する素材(b)(パン)、素材(e)(前進)を用いた。提案手法では、先に水平エッジ対処手法の適否を判断し、同手法が適さないと判断されたときのみ、平均ベクトル手法が適用される。したがって、全体の改善効果(表 3)のうち、水平エッジ対処手

法が寄与する改善効果は表5(ii)であり、平均ベクトル手法が寄与する改善効果は表5(i)-(ii)とほぼ等価となる。

表5 各改善要素の効果(平均符号料削減[%])

	(i) 4 節結果	(ii) 水平エッジ対処手法	(iii) 平均ベクトル手法	(i)-(ii)
(b)	12.61	12.53	1.83	0.08
(e)	0.97	0.75	0.31	0.22

結果より、まず素材(b)のようにカメラパンを含む素材では、水平エッジ対処手法による改善効果が極めて支配的である。パンシーンでは、画面全体が水平方向に動くが、画素値変化に乏しい平坦な領域では動き補償が不要であるため、動きベクトルが発生しない。したがって、パンシーンにおいて、平坦なテクスチャの領域と複雑なテクスチャの領域が接する境界では、水平エッジ対処手法が有効であり、表5からもこの特徴が読み取れる。一方、同素材に対し、平均ベクトル手法による全体の改善効果への寄与は極めて小さい。これは、同手法による改善点が、水平エッジ対処手法による改善点と重複しているためであると考えられる。 rmv_A の絶対値のみが大きいとき、平均ベクトル手法でも水平方向エッジ問題を緩和する予測ベクトルが生成される。このため、平均ベクトル手法単体では改善効果が得られたものの、全体では水平エッジ対処手法により解決されたために、寄与が小さくなったと考えられる。

次に素材(e)では、全体の改善効果に対して、平均ベクトル手法が23%程度寄与している。同素材では、カメラのブレなどにより、不特定方向の動きベクトルが発生する。このためメディアン予測から一意に高精度な予測ベクトルを生成することは困難であるが、提案手法により複数の予測ベクトル候補を適応的に選択することで、高精度な予測が可能となっている。

上記の改善効果への寄与傾向は、他の素材でも同様であり、パンもしくは斜め方向のカメラワークや水平方向の動きを伴う素材(a), (c), (d), (g)では、水平エッジ対処手法の効果が支配的である。一方、水平方向以外の動きを多く含む素材(f)では、平均ベクトル手法による寄与が増大する。なお、全素材を通して、水平エッジ対処手法に対する平均ベクトル手法の性能改善効果は低い。

5.2 判定処理性能

提案手法では、符号化済みMBにおける動きベクトルに基づき、処理MBに対する提案手法の適否を判定する。水平エッジ対処手法および平均ベクトル手法における適否判定の妥当性を検証するため、R-D最適化法に基づき標準手法もしくは提案手法の予測ベクトルを選択した結果と、提案手法の判定に基づき予測ベクトルの適否を判定した結果を比較した。ただしR-D最適化法による選択はあくまでも理想的な選択結果として引用したものであり、実際にこれを導入する場合、サイド情報を伴う。

表6に、素材(b), (e)に対するQP=39のときの先頭9フレーム(Pピクチャ)の結果を示す。各判定方法について、取り得る結果はそれぞれ2パターン(従来(C)および提案(P))あるため、両方式の結果の組合せは計4通りであり、これを順にCase(i)(i=0,1,2,3)と定義する。iと組み合わせの対応は表に示すとおりである。

表中の数値は、各Case(i)に該当するMB数のフレーム単位での平均値である。ここで、Case(0)およびCase(3)は、R-D最適化法と提案手法の適否判定が一致したケースである。また、Case(1)は提案手法における過剰検知であり、Case(2)は検出漏れに相当する。Case(2)が多いほど、提案手法による符号化性能改善が阻害されていることを示す。

表6 判定処理の妥当性評価

	R-D 結果, 提案判定結果	素材(b)		素材(e)	
		水平エッジ対処手法	平均ベクトル手法	水平エッジ対処手法	平均ベクトル手法
Case(0)	C, C	7130	7496	5592	5249
Case(1)	C, P	811	453	1758	2117
Case(2)	P, C	36	167	459	366
Case(3)	P, P	184	45	352	428

素材(b)では、Case(0), Case(3)が支配的であり、良好な適否判定が達成されている。また、素材(e)では、両手法ともCase(1)に該当するMB数が多く、符号化性能の改善が阻害されている懸念がある。以下、各素材での傾向について分析する。

素材(b)に関して、水平エッジ対処手法では、Case(1)が多く見られる。当該CaseのMBでは、動きベクトル予測手法を表現する冗長なサイド情報が発生する。しかし、Case(3)により動きベクトルの符号量が大きく削減されるため、符号化結果ではCase(1)の影響はCase(3)の改善効果に吸収される。一方、平均ベクトル手法では、そもそも理想ケースでの適用頻度が少ないため、Case(1)が低く抑えられているものの、符号化性能の改善に十分に寄与できていない。

素材(e)では、両手法ともCase(1)が多く見られる。同素材で理想ケースとのミスマッチが顕著となっているのは、同素材がカメラのブレを伴う素材であり、素材(b)と比べて各ブロックにおける動きベクトルが複雑である特徴を有しており、提案手法で期待しない箇所における過剰検知による影響が現れたものと推測される。

6. まとめ

HDTVに対する低ビットレート符号化を対象として、H.264の独自拡張による動きベクトル予測性能改善手法を提案した。提案手法は、特定の絵柄においてメディアン予測性能が低下する問題について、候補ベクトルを適応的に制御することで解決を図っている。また、予測ベクトルの精度追求を目指し、新たな予測ベクトル候補を導入し、標準方式との間で適応切替えを実現する手法を提案した。符号化実験結果より、平均で約3.53%、最大で12.61%の符号量削減を確認した。予測ベクトル候補の高精度化、Bピクチャへの拡張が今後の課題である。

参考文献

- [1] T. Yamamoto, "A new scheme for motion vector predictor encoding," ITU-T SG16/Q6, 32nd VCEG Meeting, San Jose, April 2007.
- [2] Garry Sullivan, Thomas Wiegand, "Rate-Distortion Optimization for Video Compression," IEEE Signal Processing Magazine, pp.74-90, Nov. 1998.
- [3] Karsten Suhring, H.264/AVC Software Coordination, <http://iphome.hhi.de/suehring/tml/>
- [4] Gisle Bjontegaard, "Calculation of Average PSNR Differences between RD curves," ITU-T SG16/Q6, 13th VCEG Meeting, Austin, Texas, USA, April 2001.