

H.264/AVC における片方向予測を用いた 双方向動き予測の最適化技術

伊藤浩朗[†] 谷田部祐介^{†1} 溝添博樹^{†1}

概要: H.264/AVC をはじめとする動画像符号化では、複数のフレームを参照する双方向予測動き探索において、最適な予測方向とブロックサイズをいかに少ない処理量で求めるかが課題となる。本論文では、片方向予測の動き推定結果に基づいて双方向予測における最適なブロックサイズを推定する手法を検討する。提案手法により、わずかな符号化効率の低下で、双方向予測動き探索に要する処理増大を抑制する動き推定を実現した。

キーワード: 画像圧縮、H.264/AVC、動き予測、双方向予測

Effective Bi-predictive Motion Estimation considering Forward and Backward Motion Estimation for H.264/AVC

HIROAKI ITO^{†1} YUSUKE YATABE^{†1} HIROKI MIZOSOE^{†1}

Abstract: In moving picture encoding like H.264/AVC, it is one of the big issues to calculate effective bi-predictive motion estimation in low computational complexity. In this paper we focus on the encoding cost at forward motion estimation and backward motion estimation at every block size and estimate suitable block size in bi-predictive motion estimation. The proposed method realizes effective bi-predictive motion estimation with low computational complexity with a little loss in encoding efficiency.

Keywords: Image Compression, H.264/AVC, Motion Estimation, Bi-predictive prediction

1. はじめに

光ディスクへのHD(High Definition)映像記録といった高ビットレート用途からモバイル向け地上デジタル放送のような低ビットレートの用途まで幅広く利用可能な、高品質・高圧縮率を実現する画像符号化方式として、H.264/AVC(Advanced Video Coding)が普及しつつある。H.264/AVCは2003年12月にIS(International Standard)[1]が発行された規格であり、MPEG-4[2]やH.264[3]と比べて同じ画質で2倍以上の圧縮率の実現が可能な符号化方式である。

H.264/AVCでは、従来の符号化方式と比較して動き補償フレーム間予測符号化時に多数のブロックサイズが用意されており、画像の動きに応じてブロックサイズを切り替えることにより高圧縮を可能としている。またその反面、最適なブロックサイズを選択するための処理負荷が高くなっている。特に、Bピクチャでは、前方予測、後方予測、双方向予測の3種類の予測方向に対する動き予測を行う必要があり、リアルタイム符号化を実現する上で大きな課題となる。H.264/AVCに続いて2013年1月に規格化されたH.265/HEVCにおいても、更にブロックサイズの種類が多様化されており、上記課題はより重要度を増している。

そこで本論文では、上記の動き補償フレーム間予測符号化に関し、特に双方向予測において、圧縮率の低下を出来

るだけ抑制しながら、動き探索に要する処理負荷を下げることを目的として、前方予測と後方予測の各片方向予測の動き探索結果から最適なブロックサイズを求める方法を検討する。

以下、2章では関連する研究について述べ、3章では解決すべき課題を説明する。4章では提案アルゴリズムについて詳説し、5章で提案アルゴリズムの評価結果について述べる。最後に6章でまとめる。

2. 関連研究

2.1 動き探索の判定方法について

H.264/AVCをはじめとする動き補償フレーム間予測では、図1に示すように符号化/復号化済みの復号画像を参照フレーム(Ref. Picture)として符号化対象フレーム(Coding Picture)を符号化する。より具体的には、符号化対象ブロック(Coding Block)と関連の高い予測ブロック(Ref. Block)を参照フレーム中から探索し、符号化対象ブロックと予測ブロックの差分である予測差分を符号化する。その際、上記の予測差分に加えて、ブロックサイズ種別の情報と、予測ブロックと符号化対象ブロックの座標値の差分で表される動きベクトル(Motion Vector)をあわせて符号化する。H.264/AVCでは後述の図4に示すように複数の異なるブロックサイズが使用可能であり、一般に小さいブロックサイズを用いるほど予測差分は小さくなるものの、ブロックごとに動きベクトルを付加する必要があるため、動きベクトルの情報量が増加する。また予測差分情報は後述の図2に示すように周波数変換(DCT)された後、量子化パラメータ

^{†1}(株)日立製作所 横浜研究所
Yokohama Research Laboratory, Hitachi Ltd.

(QP)により量子化されるため画質劣化の要因となる。そのため、予測誤差と動きベクトルの双方の情報量を鑑みて最適な動きベクトルとブロックサイズを決定する必要がある。

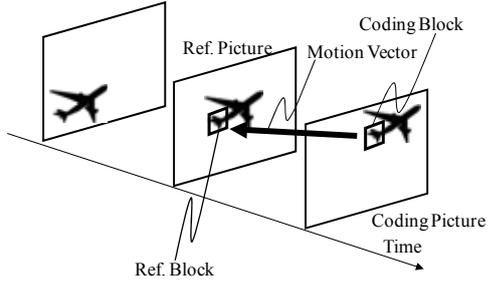


図1 動画像符号化における画面間予測の概要

Figure1 Inter-frame Coding Method in Moving Picture Coding

そのため、H.264/AVCではマクロブロック(MB)を単位として最適な符号化モードを選択するため式1に示す符号化コストをもうけている。式1右辺の第1項は符号化対象ブロックと予測ブロックとの予測誤差をアダマール変換し、変換後の各係数の絶対値和を算出したSATD(Sum of Absolute Transformed Difference)であり、第2項は符号化モード毎に定められるバイアス値(Bias)である。バイアス値はQPによって決定される。

$$Cost = SATD + Bias \quad \dots (式1)$$

本論文においても、上記の符号化コストを指針とし、本符号化コストが最小となるモードを選択する。

2.2 動き探索の演算量低減に向けた関連研究

動画像符号化における動き探索の演算量低減に向けては、これまで様々な手法が研究されている。

一つ目には、限定された領域について動き探索を行い、その結果から次の探索範囲を決定して徐々に細かい探索を行うステップサーチに関する研究が挙げられる。探索開始点を中心に上下左右の探索を行うダイヤモンドサーチ[4][5]や、探索開始点を中心として六角形となる点の探索を行うヘキサゴンサーチ[6]では、探索範囲内の全点に対して探索を行う全探索(Full Search)[7]と比べて探索箇所を限定したまま、PSNRの劣化を抑制する取り組みが研究されている。また、動き探索の途中であらかじめ定めた符号化コストを下回る動きベクトルが見つかった時点で動き探索の打ち切り(Early termination)を行う手法を併用する取り組みも検討されている。

次に、すべてのブロックサイズについて動き探索を行うのではなく、あるブロックでの探索結果を用いてブロックを統合/分割する手法についての研究が挙げられる。文献[8][9]では最初に4×4や8×8等小さいブロックサイズで動き探索を行い、その結果、動きベクトルの大きさと向きが類似したブロックを統合することで演算量を低減する手法が提案されている。また文献[10][11]では最初に8×8画素単位で動き探索を行い、その後ブロックを統合/分割する手法が提案されている。

3. 解決すべき課題

本章では、本論文で解決すべき課題を整理する。動画像符号化処理をハードウェアで実現する場合、画像入力、画面内予測、画面間予測、DCT変換、量子化、エントロピー符号化、デブロッキングフィルタ処理等、動画像符号化処理を構成する各処理ブロックを単位としたパイプライン処理構造での設計が広く用いられる。図2は符号化処理のブロック図を、図3はパイプラインの一例を示したものである。図3中の数字はMB番号を示し、例えば(A)のタイミングでは、MB番号=4の入力処理(Video Input)、MB番号=4に対応する参照画像取得(Ref Input)、MB番号=3に対する動き探索(Motion Estimation)等が並列して処理されることを示す。入力画像がフルHD画像(1,920×1,080)、フレームレートが30[frame/sec]の場合、1フレームを構成する8160個の各MBを33[msec]で処理する必要があることから、1MB期間は約4[μsec]となる。即ち、図3に示す各処理は[4μsec]以内に終了する必要がある。このような設計条件下においては、従来取り組みのステップサーチに示した打ち切り処理のように処理時間が変動するアーキテクチャではワーストケースを想定して設計を行う必要があり、処理時間の短縮を図ることが難しい。

加えて、H.264/AVCの動き補償フレーム間予測符号化では、図4に示すように動き補償を行う単位として16×16画素、16×8画素、8×16画素、8×8画素の4種類のブロックサイズを選択することが可能である。また、8×8画素のブロックサイズを選択した場合には、各々の8×8画素のブロックに対して、8×8画素、8×4画素、4×8画素、4×4画素のブロックサイズを選択することが可能であり、合計で7種類のブロックサイズが使用可能である。

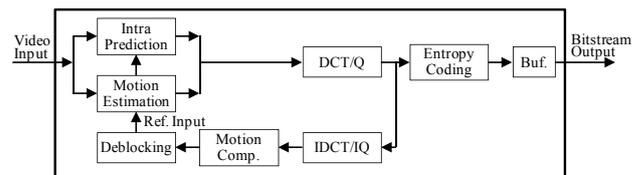


図2 符号化処理部ブロック図

Figure2 Encoder Block Diagram

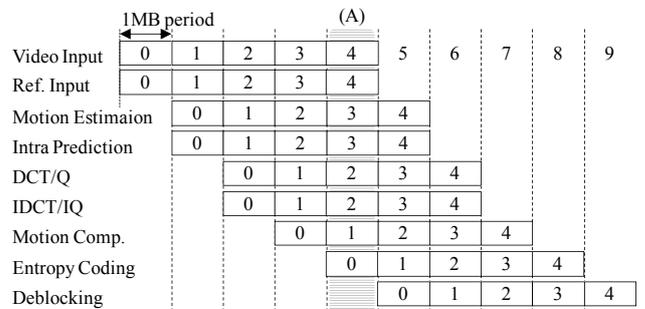


図3 符号化処理パイプライン処理例

Figure3 Encoder Pipeline Example

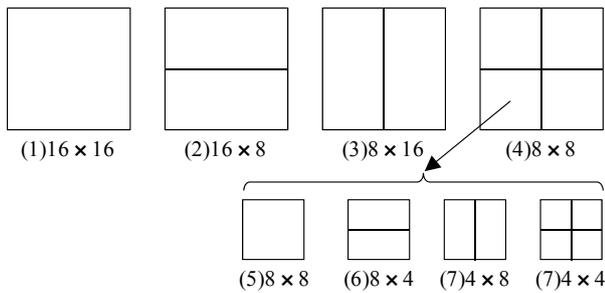


図4 H.264/AVCにおけるブロックサイズ種別

Figure4 Block Sizes in H.264/AVC

そのため、動き補償予測のうち、前方予測のみでなく、後方予測、双方向予測と参照画像の選択枝が増えるBピクチャにおいては、最適な動き予測の算出に多くの演算量が必要となる。一例として、図4に示したブロックサイズのうち(1)16×16～(4)8×8の4パタンのブロックサイズのみを用いた場合、各ブロックに対して前方予測、後方予測、双方向予測で最適な動きベクトルを算出する回数は表1のようになる。BピクチャではPピクチャ(表1中のForward-predのみ)と比較して演算回数が3倍になることが分かる。

表1 ブロックサイズ毎の動きベクトル算出回数

Table1 Motion Estimation Calc. Times for each block size

Block Size	Forward-pred	Backwad-pred	Bi-pred
16×16	1	1	1
16×8	2	2	2
8×16	2	2	2
8×8	4	4	4

従来の研究では、1つの予測方向に対してあるブロックサイズでの探索結果を用いてブロックを統合/分割する研究は多く行われていたが、双方向予測での最適なブロックサイズを推定する技術については報告されていない。そこで、本論文では、片方向予測での動き探索結果から双方向予測での最適なブロックサイズを推定することにより、双方向予測の符号化コスト算出の処理簡略化を目的とする。

4. 提案アルゴリズム

本章では、提案アルゴリズムの詳細について示す。

4.1 片方向予測と双方向予測の相関について

最初に、片方向予測において符号化コストが最小となるブロックサイズと、双方向予測において符号化コストが最小となるブロックサイズとの相関について予備検証を行った。具体的には、後述の図7(a)に示す標準動画「Walk through the Square」を後述のJM18.6にて符号化し、片方向予測(前方予測および後方予測)で符号化コストが最小となるブロックサイズについて、双方向予測ではどのブロックサイズで符号化コストが最小となるかを評価した。なお本検証においてはブロックサイズを16×16、16×8、8×16、8×8の4通りとする。結果を表2に示す。

表2 双方向予測の推定ブロックサイズと合致率

Table2 Estimated Block Size in Bi-precision and accuracy rate

Block Size	16×16	16×8	8×16	8×8	Total
EstimatedMBs	436	254	251	409	1350
Accuracy	80.0%	31.9%	37.1%	18.1%	44.2%

表2では、例えば片方向予測でブロックサイズ=16×16が符号化コスト最小となる場合は計436MBあり、これを双方向予測での最適ブロックサイズと推定した場合の合致率(実際に双方向動き予測を行った結果、ブロックサイズ=16×16の符号化コストが最小となる割合)は80.0[%]であることを示す。表2の結果より、片方向予測で符号化コストが最小となるブロックサイズをそのまま双方向予測での最適ブロックサイズと推定するアプローチでは、ブロックサイズ=16×16を除いて合致率が低いことが判明した。そこで、片方向予測で各ブロックサイズの符号化コストが最小となった場合に、双方向予測で符号化コストが最小となるブロックサイズの割合を調査した(表3)。

表3 双方向予測におけるブロックサイズの選択割合

Table3 Block Size ratio in Bi-prediction

(a)片方向予測での選択ブロックサイズ=16×16の時

(a)Block Size in Single-direction prediction=16×16

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	80.0%	8.5%	8.0%	3.5%

(b)片方向予測での選択ブロックサイズ=16×8の時

(a)Block Size in Single-direction prediction=16×8

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	48.0%	31.9%	13.0%	7.1%

(c)片方向予測での選択ブロックサイズ=8×16の時

(c)Block Size in Single-direction prediction=8×16

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	45.8%	10.8%	37.1%	6.3%

(d)片方向予測での選択ブロックサイズ=8×8の時

(d)Block Size in Single-direction prediction=8×8

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	35.2%	21.0%	25.7%	18.1%

調査の結果、双方向予測では片方向予測で選択されたブロックサイズよりも大きいブロックサイズで符号化コストが最小となる傾向であることが分かった。そのため、片方向予測の動き探索結果から双方向予測のブロックサイズを以下の式2にて求めることとした。

$$Bi_pred_blocksize = Blocksel(MVCost_{16 \times 16}, MVCost_{16 \times 8} * W_0, MVCost_{8 \times 16} * W_0, MVCost_{8 \times 8} * W_1)$$

・・・(式2)

関数 Blocksel()は片方向予測の符号化コスト(MVCost_{16×16}, MVCost_{16×8}, MVCost_{8×16}, MVCost_{8×8})の中から最小となるブロックサイズを選択する。なお、W₀、W₁は重み

付け係数である。

上記の式 2 に基づき再度、本節冒頭と同一の評価を行った。結果を表 4、表 5 に示す。

表 4 双方向予測の推定ブロックサイズと合致率(提案手法)

Table4 Estimated Block Size in Bi-precision and accuracy rate
 (Proposed)

Block Size	16×16	16×8	8×16	8×8	Total
Estimated MBs	857	193	213	87	1350
Accuracy	67.7%	40.9%	48.8%	37.9%	59.0%

表 5 双方向予測におけるブロックサイズの選択割合
 (提案手法)

Table5 Block Size ratio in Bi-prediction(Proposed)

(a)片方向予測での選択ブロックサイズ=16×16の時

(a)Block Size in Single-direction prediction=16×16

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	67.7%	14.1%	13.1%	5.1%

(b)片方向予測での選択ブロックサイズ=16×8の時

(a)Block Size in Single-direction prediction=16×8

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	36.3%	40.9%	13.5%	9.3%

(c)片方向予測での選択ブロックサイズ=8×16の時

(c)Block Size in Single-direction prediction=8×16

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	28.6%	9.4%	48.8%	13.1%

(d)片方向予測での選択ブロックサイズ=8×8の時

(d)Block Size in Single-direction prediction=8×8

Block Size in Bi-prediction	16×16	16×8	8×16	8×8
Selected Ratio	21.8%	12.6%	27.6%	37.9%

表 4 の結果より、提案手法では表 2 と比べて合致率が向上(44.2[%]→59.0[%])していることが分かる。また、表 5 に示す各ブロックサイズ別の合致率においても(a)16×16 を除いて合致率が向上していることが分かる。

4.2 動き探索の提案アルゴリズム

提案手法における動き探索処理を図 6 に示す。また比較のため、従来手法における動き探索処理を図 5 に示す。従来手法では、ブロックサイズごとに、前方予測(MV_fw_blksize)、後方予測(MV_bw_blksize)、双方向予測(MV_bid_blksize)の各符号化コストを算出し、その中で符号化コストが最小となる予測方向を選択(Sel)する。blksize は 16×16、16×8、8×16、8×8 のいずれかのブロックサイズを示す。またブロックサイズが 16×8、8×16、8×8 の際には各ブロックの符号化コストを合算(Sum)して 1MB の符号化コストを求める。最後に、16×16、16×8、8×16、8×8 の各ブロックサイズの中から符号化コストが最小となる符号化モードを選択(Sel)する。

これに対し提案手法では、前方予測、後方予測の符号化コストを算出する過程は従来手法と同じものの、16×16、16×8、8×16、8×8 の各ブロックサイズの片方向予測の中

から前述の式 2 に基づいて双方向予測のブロックサイズを選択(Bid Block Sel)し、そのブロックサイズでの符号化コストを算出(MV_bid)する。最後に、各ブロックでの符号化コスト 4 候補と双方向予測の符号化コスト 1 候補の中から符号化コストが最小となる符号化モードを選択(Sel)する。

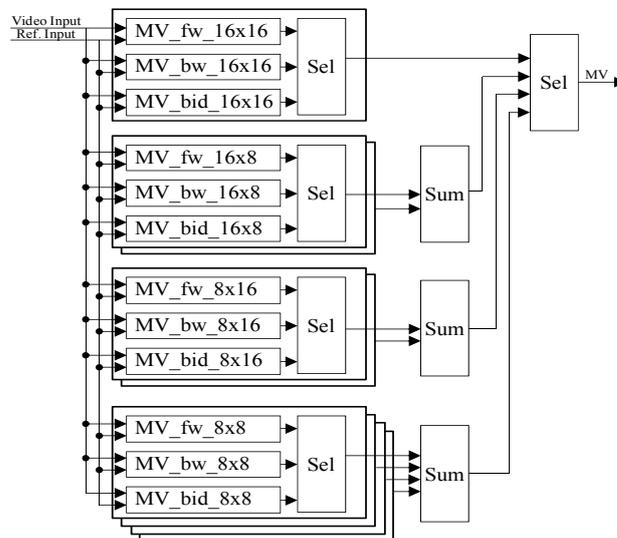


図 5 従来手法における動き探索処理

Figure5 Conventional Motion Estimation Block Diagram

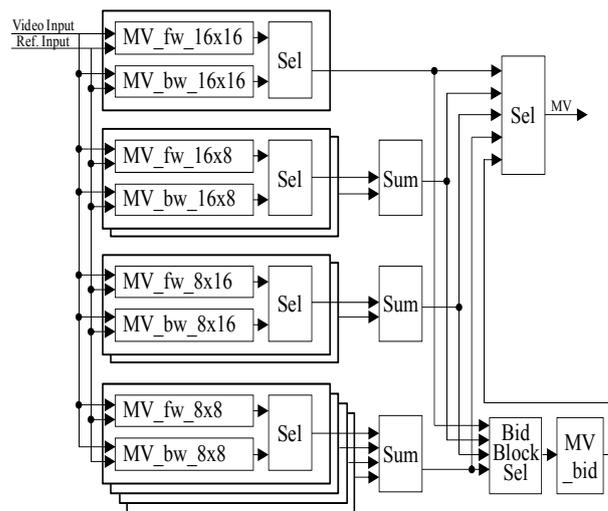


図 6 提案手法における動き探索処理

Figure6 Proposed Motion Estimation Block Diagram

5. 評価

提案手法の評価を行うにあたり、そのベースモデルとして JM18.6[12]を使用した。評価実験における基本的な符号化条件を表 6 に示す。また、評価実験に使用する入力画像として、図 7 に示す「Walk through the Square(以下、Square)」「Whale Show」、「Green Leaves」、「Crowded Crosswalk」[13]の 4 種類を使用した。「Square」は少女が花壇の周りを水平に動くとともに、カメラが水平にドリーするシーンであり、建物のレンガの壁や花壇の瓶および人物の顔にひずみが生じやすい[14]。「Whale Show」はシャチのジャン

プに合わせてカメラが上下左右にパンするシーンであり、背景の観客席には、ブロック歪などが生じやすい[14]。「Green Leaves」は葉の生い茂った並木道をカメラでゆっくりズームインするシーンであり、カメラのズームにより画面の中心部ではほぼ固定だが周辺では速く移動し、ブロック歪などの妨害が木の葉、道路に生じやすい[14]。「Crowded Crosswalk」は横断歩道をわたる人々をフルショットで撮影したシーンであり、ブロック歪などの妨害が背景の自動車などに生じやすい[14]。本評価実験では、上記のようにそれぞれ動きの特徴の異なる4種類のシーケンスを用いることで提案手法の汎用性を評価することとした。

表 6 符号化条件

JM Version	Ver.18.6
Profile	Main Profile
Input Picture Format	720×480
Frame Rate	60fps
Frame Number	140
Frame Structure	M=2,N=14
QP	25~32
Reference Frame Numer	2
Search Range	32
Optimization	RD-Optimization off

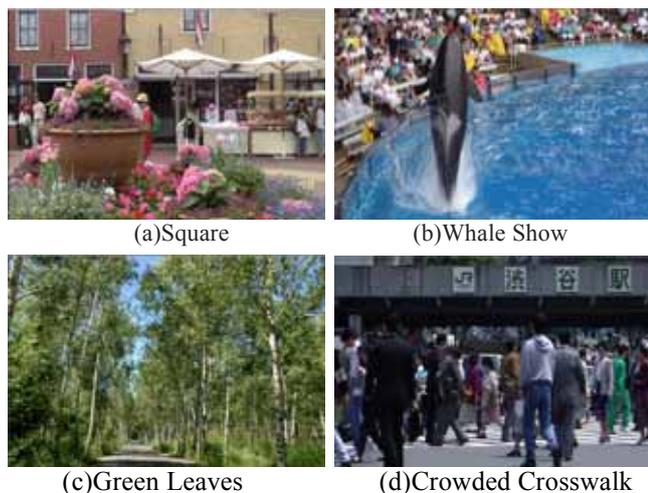


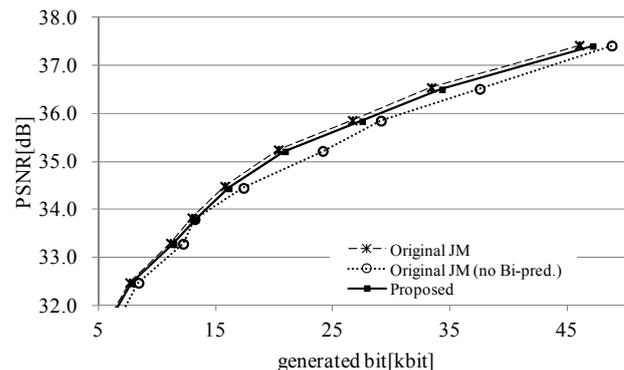
図 7 評価画像

Figure7 Test Sequences

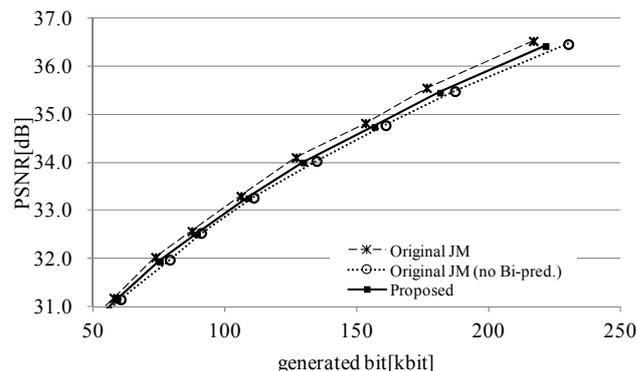
5.1 符号化効率に関する評価

提案手法に関して各 QP での発生符号量と PSNR を取得し、符号化効率の評価を行った。結果を図 8 および表 7 に示す。本評価では、JM18.6 を用いて図 5 の処理に基づいて動き探索を行う場合(図 8 中の"Original JM")、JM18.6 において強制的に双方向予測の処理をマスクし、前方予測および後方予測のみを有効として動き探索を行う場合(図 8 中の"Original JM(no Bi-pred.)")、および提案手法(図 8 中の"Proposed")の 3 手法について比較を行った。なお本提案手法は B ピクチャのみに影響を及ぼすため、図 8 では B ピクチャのみを評価対象として B ピクチャ 1 フレームあたりの

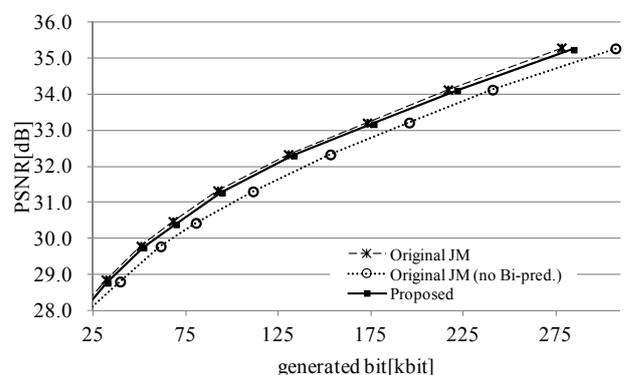
発生符号量と PSNR をそれぞれ横軸、縦軸に示す。



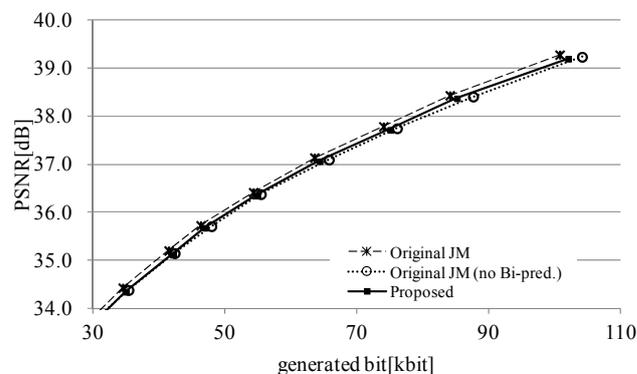
(a)RD curve (Square)



(b)RD curve (Whale Show)



(c)RD curve (Green Leaves)



(d)RD curve (Crowded Crosswalk)

図 8 レート歪特性

Figure8 Rate-Distortion characteristics

また表 7 に、Original JM を基準とした時の Original JM(no

Bi-pred.)および提案手法における発生符号量の増加割合を示す。評価の結果、図 7(a)~(d)いずれの評価画像においても提案手法は Original JM よりは劣るものの、Original JM(no Bi-pred.)よりも良い符号化効率を実現できていることが分かる。

表 7 発生符号量

Table7 Δ bit rate

Sequence	Δ bit rate	
	Original JM (no Bi-pred.)	Proposed
Square	+9.7[%]	+2.4[%]
Whale Show	+5.5[%]	+2.2[%]
Green Leaves	+17.0[%]	+2.1[%]
Crowded Crosswalk	+3.0[%]	+1.2[%]

5.2 処理速度に関する評価

次に提案手法の処理時間について評価を行った。本評価では、JM18.6において encode_one_fame()関数の前後に clock()関数を挿入して B ピクチャの処理時間を計測した。評価の結果、1.4~1.8[%]と僅かではあるが提案手法にて従来手法より処理時間が軽減されることが確認された。

表 8 符号化処理時間

Table8 Encoding Time

Sequence	Time[sec]		ΔT ((b)/(a)×100[%])
	(a)Original JM	(b)Proposed	
Square	5.19	5.11	98.4
Whale Show	5.24	5.14	98.2
Green Leaves	5.14	5.07	98.6
Crowded Crosswalk	5.19	5.11	98.5

5.3 回路実装結果

Full-HD(1920×1080i, 4:2:0,8bit)画像の符号化対応に向け、本回路を含めた動画符号化処理の回路実装を行った。結果を表 9 にまとめる。システム LSI に実装し、リアルタイム動作することを確認した。

表 9 回路実装結果

Table9 Circuit Implementation

Item	Specifications
Video Format	1920×1080i
Video Compression Format	H.264/AVC (HP@L4.1)
LSI process	90nm
System clock	162MHz

5.4 考察

従来手法ではブロックサイズごとに双方向予測の符号化コストを算出し、片方向予測での符号化コストと比較を行っていたが、本提案手法では、片方向予測の符号化コスト値から双方向予測で最適なブロックサイズを推定することにより、符号化効率を大きく損なうことなく処理負荷の軽減が可能になると考える。

6. 結論

H.264/AVC における双方向予測動き探索について、片方向予測の符号化コスト値を用いることで双方向予測のブロックサイズを推測する手法を提案した。提案手法により、符号化効率の低下(発生符号量の増加)を 1.2~2.4[%]に抑制しつつ、ブロックサイズごとの双方向予測演算処理を簡略化できることが確認できた。今回の検討は H.264/AVC を対象としているが、2013 年 1 月に規格化された H.265/HEVC ではブロックサイズ、ブロック分割モードの選択肢が更に増大しており、本提案手法の適用検討は今後の課題と考える。

参考文献

- 1) Joint Video Team(JVT) of ISO/IEC MPEG & ITU-T VCEG, "Text of International Standard of Joint Video Specification", ITU-T Rec. H.264 | ISO/IEC 14496-10 Advanced Video Coding, Dec. 2003.
- 2) ISO/IEC 14496-2, "Information technology Coding of audio-visual objects Part2: Visual", 2001.
- 3) ITU-T Rec. H.263, "Video coding for low bitrate communication", 1996.
- 4) S. Zhu, K. Kuang, "A new diamond search algorithm for fast block-matching motion estimation", IEEE Trans. Image Processing, Vol.9, No.2, pp287-290, Feb.2000.
- 5) A. Tourapis, O. Au, and M. Liou, "New results on zonal based motion estimation algorithms-advanced predictive diamond zonal search", IEEE International Symposium on Circuits and Systems, Vol.5, No.5, pp.183-186, May.2001.
- 6) C. Zhu, X. Lin, and L. P. Chau, "Hexagon-based search pattern for fast block motion estimation", IEEE Trans. Circuits and Systems for Video Technology, Vol.12, No.5, pp.349-355, May.2002.
- 7) F. Dufaux, F. Moscheni, "Motion estimation techniques for digital TV :A review and a new contribution", Proceedings of the IEEE, Vol.83, pp.858-876, Jun.1995.
- 8) Y.K.Tu, J.F.Yang, Y.N.Shen, M.T.Sun, "Fast Variable-Size Block Motion Estimation Using Merging Procedure With an Adaptive Threshold", Proc. IEEE International Conference on Multimedia & Expo, Baltimore, MD, USA, Jul. 2003.
- 9) M.Kucukgoz, M.T.Sun, "Early-stop and Motion Vector Reusing for forMPEG-2 to H.264 Transcoding", SPIE Proc. Visual Comm. Image Proc., San Jose, CA, USA, Jan. 2004.
- 10) Zhi Zhou, Ming-Ting Sun, Yuh-Feng Hsu, "Fast Variable Block-Size Motion Estimation Algorithms Based On Merge And Split Procedures For H.264/mpeg-4 Avc," IEEE ISCAS2004
- 11) X.Li, G.Wu, "Fast Integer Pixel Motion Estimation", JVTFO11, 6th Meeting, Awaji Island, Japan, Dec. 2002.
- 12) <http://iphome.hhi.de/suehring/tml/>
- 13) "標準テレビジョン方式・システム評価用標準動画画像", 一般社団法人 映像情報メディア学会
- 14) "標準テレビジョン方式・システム評価用標準動画画像 解説書", 一般社団法人 映像情報メディア学会