インタレース映像を対象とした MPEG-2/H.264 イントラ再符号化方法の検討

A Study of MPEG-2 to H.264 Intra Transcoding for Interlace Contents

正会員 吉 留 健[†], 正会員 上 倉 一 人[†], 北 脇 信 彦^{††}

Takeshi Yoshitome[†], Kazuto Kamikura[†] and Nobuhiko Kitawaki^{††}

Abstract An MPEG-2 to H.264 intra transcoding method for interlace contents is proposed. This method uses the encoding information from an MPEG-2 stream and keeps the DCT coefficients of the original MPEG-2 bitstream as much as possible. Experimental results show that the proposed method improves on the conventional method about 0.25 dB in PSNR.

キーワード：再符号化 トランスコーディング H.264/AVC MPEG-2 インタレース 量子化雑音

1. ま え が き

MPEG-2 符号化¹⁾された映像を H.264²⁾を用いて、より低レートに再符号化する要求が高まっている。これまでの再符号化の研究において、MPEG-2 から MPEG-2 などの同一符号化規格による再変換では、初段符号化と同一の符号化情報 (ピクチャタイプ, マクロブロック (MB) タイプ, DCT タイプ, 動きベクトル, 量子化ステップ..) で再符号化することで, DCT 成分が殆んど保存され, 量子化雑音を極めて小さく抑えられることが知られている^{3)~5)}。また, MPEG-2 から H.264 への再符号化では, 地上波デジタル放送を IP 再送信するトランスコーダが開発されている⁶⁾。これは, センタで受信した MPEG-2 の地上波放送を H.264 にトランスコードした後に, IP ネットワークで各家庭に配信するものであり, MPEG-2 のピクチャタイプ等を H.264 に継承させることで符号化効率の向上を図っている。しかし, 一部の符号化情報の継承では一定の効果は得られるものの, 上述したすべての符号化情報を継承した同一規格間の再符号化で得られたような再量子化雑音の混入防止は期待できない。すべての符号化情報の継承による雑音混入防止を目指した MPEG-2/H.264 再符号化として,

プログレッシブ映像の MPEG-2 イントラストリームに限定した MPEG-2/H.264 再符号化が報告されている⁷⁾。これは, MPEG-2 の動作すべてを極力近似した H.264 再符号化動作 (画面内予測を DC モード, DCT サイズを 8x8, 量子化ステップを MPEG-2 と極力同じにするなど) で, 再量子化雑音の混入を極力防止する手法である。しかし, 地上波デジタル放送をはじめ多くの MPEG-2 映像はフレーム/フィールド MB が混在したストリームである。そこで, 本稿では入力対象を拡張し, フレーム/フィールド MB が混在した MPEG-2 イントラストリームを対象とした H.264 再符号化を提案する。

上述した地上波デジタルの IP 再送信に関する調査¹²⁾によると, 10Mbps の H.264 トランスコード画質が地上デジタル放送 (約 14Mbps) と同等かそれに近い画質が得られると確認されていることから, 本論文のトランスコード手法が満たすべき要件を, リアルタイム系のトランスコードを前提とした IP 再送信の再圧縮率 0.714 (=10Mbps/14Mbps) において, 従来手法を超える再符号化性能を得ることとする。

本論文は, 章 2 でプログレッシブ映像の MPEG-2 ストリームの H.264 再符号化で混入する再量子化雑音が, 初段 MPEG-2 符号化情報を利用することで抑圧されることを示す。章 3 では MPEG-2 符号化と H.264 符号化の差異をフレーム/フィールド MB の取り扱いを中心に検討し, フレーム/フィールド MB 混在の H.264 再符号化で解決すべき事項を示す。章 4 で提案するフレーム/フィールド MB 混在の H.264 再符号化手法を述べ, 章 5 で提案手法を実装したトランスコーダによる実験結果を示し, 従来手法と比較, 考察する。

2009 年 5 月 15 日受付, 2009 年 12 月 22 日再受付, 2010 年 4 月 20 日採録

[†] NTT サイバースペース研究所

(〒 238-0847 神奈川県横浜須賀市光の丘 1-1, TEL 046-859-3173)

^{††} 筑波大学 システム情報工学研究科

(〒 305-8573 茨城県つくば市天王台 1-1-1, TEL 029-853-6246)

[†] NTT Cyber Space Laboratories, NTT Corporation

(1-1, Hikarinooka, Yokosuka, Kanagawa, 238-0042, Japan)

^{††} Graduated School of System and information Engineering, University of Tsukuba

(1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan)

2. 初段符号化情報による再量子化雑音低減の概要

本章では、初段符号化情報の再利用による量子化雑音低減⁷⁾の概要を述べる。初段 MPEG-2 イントラ符号化とその符号化情報を使用した H.264 イントラ再符号化におけるデータの流れを図 1 に示す。簡単のため MB サイズ等は 2x2 とする。初段 MPEG-2 エンコーダでは、入力 MB の画素値 (A,B,C,D) が直交変換された後、得られた DCT 係数 (47,31,25,18) が量子化ステップ $\Delta=10$ で量子化される際に量子化雑音が混入するため、逆量子化後の DCT 係数は (40,30,20,10) となり、異なる値となる。そのため、MPEG-2 デコーダの出力 (a,b,c,d) も入力 MB 画素値に雑音が混入したものとなる。

H.264 エンコーダで行われる初段符号化情報を利用する再符号化では、MPEG-2 エンコーダと極力一致させたモード選択、直交変換、量子化を行う。具体的には、イントラ予測のない MPEG-2 にもっとも近い DC モード予測を H.264 イントラ予測とし、H.264 整数 DCT サイズも MPEG-2 の DCT サイズと同じ 8x8 とし、量子化マトリックスも MPEG-2 と同一係数とする。量子化ステップも基本的には MPEG-2 と同一のものを使用するが、表 1 に示した MPEG-2 と H.264 で使用可能な量子化ステップからもわかるように、H.264 で使用できないステップが MPEG-2 で使用されていた場合には、もっとも近いものを選ぶ。例えば、H.264 で使用できない $\Delta=6$ が MPEG-2 で使用された場合、H.264 では $\Delta=5.5$ あるいは $\Delta=6.5$ を選ぶ。

これらの近似によって、H.264 の画面内予測補償の残差 (a',b',c',d') と直交変換係数 (41,31,20,10) と量子化後直交変換係数 (4,3,2,1) は MPEG-2 と高いレベルで近似させることができる。その後得られる H.264 デコーダの出力 (A'',B'',C'',D'') は MPEG-2 デコーダ出力 (a,b,c,d) に極

表 1 H.264 と MPEG-2 で使用可能な量子化ステップ Δ
Available Δ in MPEG-2 and in H.264.

Δ	H.264	MPEG-2 q_scale_type		Δ	H.264	MPEG-2 q_scale_type	
		0	1			0	1
0.625	○	-	-	24	-	○	○
0.6875	○	-	-	26	○	○	-
0.8125	○	-	-	28	○	○	○
0.875	○	-	-	30	-	○	-
1	○	-	○	32	○	○	○
1.125	○	-	-	34	-	-	-
1.25	○	-	-	36	○	○	○
1.375	○	-	-	38	-	-	-
1.625	○	-	-	40	○	○	○
1.75	○	-	-	42	-	-	-
2	○	○	○	44	○	○	○
2.25	○	-	-	46	-	-	-
2.5	○	-	-	48	-	-	○
2.75	○	-	-	50	-	-	-
3	-	-	○	52	○	○	○
3.25	-	-	-	54	-	-	-
3.5	○	-	-	56	○	○	○
4	○	○	○	58	-	-	-
4.5	○	-	-	60	-	-	-
5	-	-	○	62	-	-	-
5.5	○	-	-	64	○	-	○
6	-	○	○	68	-	-	-
6.5	○	-	-	72	○	-	○
7	○	-	○	80	○	-	○
8	○	○	○	88	○	-	○
9	○	-	-	96	-	-	○
10	○	○	○	104	○	-	○
11	○	-	-	112	-	-	-
12	-	○	○	128	○	-	-
13	○	-	-	144	○	-	-
14	-	○	○	160	○	-	-
16	○	○	○	176	○	-	-
18	○	○	○	208	○	-	-
20	○	○	○	224	○	-	-
22	○	○	○	-	-	-	-

めて近くなり、H.264 再符号化での量子化雑音の再混入を小さくすることができる。また、CABAC などの性能の高いエントロピー符号化ツールによって得られる H.264 の bitstream2 のサイズは、VLC を用いた MPEG-2 の bitstream1 よりも小さくできる。ただし、文献 7) で対象とした MPEG-2 符号化は、すべての MB がフレーム MB の場合についてである。インタレース映像でも上述した再量子化雑音低減の効果を得るためには、フィールド MB も考慮しなくてはならない。次章では、H.264 と MPEG-2 におけるフレーム MB/フィールド MB の違いを述べる。

3. MPEG-2 と H.264 のフレーム/フィールド MB

3.1 フレーム/フィールド MB 指定方法の相違

MPEG-2 の入力 MB と DCT 対象画素の関係を図 2(a) に示す。同図で太実線で囲まれた 8x8 画素は、サイズ 8x8 の DCT 変換を受ける画素ブロックである。DCT 種別がフレームのときには画素ブロック A~D を、また DCT 種別指定がフィールドのときには画素ブロック a~d を、対象に DCT が実施される。

これに対し、H.264 でフレーム MB/フィールド MB を混在させる場合には、MacroBlock Adaptive Frame Field Coding (MBAFF) を宣言して、図 2(b) に示すように、縦に連続した二つの MB (ペア MB と呼ばれる) 毎にフレーム/フィールドを設定する。なお、MPEG-2 ではペア MB の概

図 1 初段 MPEG-2 符号化情報を使用した H.264 再符号化
H.264 Transcoding with first MPEG-2 encoding information.

図 2 MPEG-2 と H.264 の DCT 種別の指定方法
DCT type specifications for MPEG-2 and for H.264.

図 3 MPEG-2 と H.264 の DCT 種別のミスマッチ
DCT type mismatch between MPEG-2 and H.264.

念はないが、本稿では H.264 との対応をとるため、MPEG-2 でも縦に連続した二つの MB をペア MB と呼ぶこととする。フレームペア MB の場合は、MPEG-2 でペア MB がともにフレーム MB 指定されたときと同一の画素ブロック（つまり、ブロック A~H）が整数 DCT 対象となる。また、フィールドペア MB の場合は、第 1 フィールド側 MB で画素ブロック a,b,e,f が、第 2 フィールド側 MB で c,d,g,h が、整数 DCT 対象画素となる。

MPEG-2 で指定された MB 毎のフレーム/フィールド種別と、これを極力保存した場合の H.264 のフレーム/フィールド種別の対応例を図 3 に示す。各 MB のフレーム/フィールドが自由に設定可能な MPEG-2 に対し、H.264 ではペア MB 毎でフレーム/フィールドを設定する必要があるた

表 2 MPEG-2 エンコーダ 5 機種で実測した MB の DCT 種別が上下ペアで一致する割合

Percentage of DCT type agreement between pair MBs measured from five MPEG-2 encoder's output streams.

No	Scene	エンコーダ				
		A	B	C	D	E
1	Cognac and Fruit	84.9	84.5	84.0	89.7	84.8
8	Walk through the Sq.	83.8	85.9	82.8	89.7	83.5
10	Streetcar	84.4	85.3	82.7	82.0	81.2
14	Yacht Harbor	84.3	85.5	84.8	82.9	82.7
15	Yachting	89.2	84.3	83.3	89.6	84.7
16	Whale Show	87.3	88.2	87.6	88.1	87.6
19	Opening Ceremony	81.8	86.7	80.6	82.3	82.5
22	Marching in	82.5	88.9	82.5	83.6	81.9
23	Green Leaves	92.5	93.9	89.1	91.5	91.1
28	Summertime Tanning	89.6	88.8	87.9	89.7	88.0
46	Sprinkling	85.8	85.5	84.1	84.6	85.0
47	Picture Cuts	86.0	86.0	86.8	89.6	84.3
Ave		86.0	87.0	86.7	86.9	84.8

め、MPEG-2 のペア MB でフレーム/フィールド種別が異なる場合、DCT 種別の完全な継承ができない。フレーム/フィールド種別が一致しないと、DCT ブロック構成画素も一致しないため、章 2 で示した再量子化雑音の低減効果が得られない。不一致の割合が高ければ、再変換特性の大きな劣化につながる。次節では DCT 種別が不一致となる MB の割合を、実際の MPEG-2 ストリームから求める。

3.2 DCT 種別継承ができない MB の割合

市販 MPEG-2 エンコーダ 5 機種の出力ストリームを解析し、ペア MB のフレーム/フィールド種別が一致する割合を測定した。結果を表 2 に示す。入力映像は標準動画⁸⁾に含まれる 12 シーンである。平均 86% 程度のペア MB でフレーム/フィールド種別が一致した。残り 14% のペア MB は、片方の MB は一致させることができるので、平均 93% 程度の MB でフレーム/フィールド種別が一致する。H.264 再符号化で DCT 種別が不一致になる MB は、わずか 7% 程度と小さい。したがって、フレーム映像を対象にした再符号化手法⁷⁾に、DCT 種別を極力継承する機能を付与することで、フレーム/フィールド混在の場合においても、高い再量子化雑音抑圧効果が得られる可能性がある。

4. 提案手法

本章では、フィールド MB とフレーム MB が混在した MPEG-2 再生画像を、その MB 種別に応じて H.264 のペア MB 種別を決定する四つの H.264 再符号化手法を提案する。提案四手法の違いは、MPEG-2 の DCT 種別が H.264 に継承できない MB に対するモード選択方法にある。各提案手法（手法 r、手法 R、手法 i、手法 I）のペア MB 種別選択方法を表 3 に示す。手法 r、手法 R は MPEG-2 でペア MB がともにフィールド MB であった場合を除き、フレームペアを選択して H.264 再符号化する。両者の差はフレーム/フィールドが継承されない場合、つまり MPEG-2 でフィールド MB だった MB が、H.264 でフレームとして処理される際に現れる。この時、手法 r が MB タイプ = I8x8、予測モード = DC を固定的に使用するのに対し、手法 R はコスト計算して MB タイプと予測モードを再判定す

表 3 提案する四つの PEG-2/H.264 イントラ再符号化手法における DCT タイプ変換規則
DCT type mapping from MPEG-2 to H.264 intra transcoding.

初段 MPEG2		提案する H.264 再符号化の手法									
		手法名	手法 r		手法 R		手法 i		手法 I		
上側 M B	下側 M B	MB 位置	上側 M B	下側 M B	上側 M B	下側 M B	上側 M B	下側 M B	上側 M B	下側 M B	
Frame	Frame	ペア種別	Frame		Frame		Frame		Frame		
		MB タイプ	I8	I8	I8	I8	I8	I8	I8	I8	
		DCT 継承	○	○	○	○	○	○	○	○	
Frame	Field	ペア種別	Frame		Frame		Field		Field		
		MB タイプ	I8	I8	I8	再	I8	I8	再	I8	
		DCT 継承	○	×	○	×	×	○	×	○	
Field	Frame	ペア種別	Frame		Frame		Field		Field		
		MB タイプ	I8	I8	再	I8	I8	I8	再	I8	
		DCT 継承	×	○	×	○	○	×	○	×	
Field	Field	ペア種別	Field		Field		Field		Field		
		MB タイプ	I8	I8	I8	I8	I8	I8	I8	I8	
		DCT 継承	○	○	○	○	○	○	○	○	

再：再判定
I8：I8x8

る．同様に，手法 i と手法 I は MPEG-2 でペア MB がともにフレームであった場合を除き，フィールド MB で H.264 再符号化する．フレーム/フィールドが継承されない場合のみ，手法 i が MB タイプ=I8x8，予測モード=DC を固定的に使用するのに対し，手法 I はコスト計算して MB タイプと予測モードを再判定する．従来手法は MPEG-2 の MB 種別と無関係に，すべての場合でフレーム/フィールド種別，MB タイプ，予測モードをコスト計算して判定する．フレーム/フィールド種別指定方法以外の変換規則は，プログレッシブ映像用の MPEG-2/H.264 再変換手法⁷⁾の規則を用いる．つまり，H.264 のブロックサイズはすべて 16x16 とし，MPEG-2 の DCT サイズ 8x8 が使用可能な HighProfile を使い，量子化マトリックスは MPEG-2 の 8x8 量子化マトリックスを用いる．周辺画素を参照しない MPEG-2 の Intra 予測に近似させるために，H.264 の画面内予測には DC モードを使用するとともに，I4x4 や I16x16 の MB はサイズ=4 で整数 DCT されることから，サイズ=8 の整数 DCT が使用される I8x8 を MB タイプとする．

5. 評価

5.1 シミュレーション条件

提案手法の性能評価のため，シミュレーション実験を行った．比較対象は，初段符号化情報を用いずに通常モード選択などを行って H.264 再符号化する手法（従来手法）とした．シミュレーション条件は以下の通りである．初段 MPEG-2 の入力映像は，720x480 画素のインタレース映像とし，量子化ステップ Δ を固定した TM5⁹⁾ を用いて，フレーム/フィールド混在の MPEG-2 イントラビットストリームを作成し，これを次段の H.264 再符号化の入力とした．従来手法の H.264 再符号化には jm12.1¹⁰⁾ を (UseHadamrd=on, 量子化マトリックスは H.264 デフォルト設定， エントロ

ピー符号化は CABAC) の条件で使用した．RD 最適化のレベルは Low_complexity_mode に設定した．これ以外のレベルに High_complexity_mode があるが，各 MB のモード等を決定する際に，DCT 変換，量子化，逆量子化，逆 DCT 変換，PSNR 算出の各処理が候補となるモードの数だけ必要となる．このため，膨大な演算量が必要で，オフラインエンコードを行う（リアルタイム処理によるエンコードを必要としない）アプリケーション向けのモードである¹¹⁾．本実験で用いた Low_complexity_mode は，計算量の低いアダマール変換を行って得られる係数からコスト値を算出する方法で，多くのリアルタイムエンコーダやトランスコーダで使用されている．リアルタイム系のトランスコードを前提とした本論文では，以上の理由により，RD 最適化のレベルを Low_complexity_mode とした．

提案手法の H.264 再符号化には初段 MPEG2 符号化情報の入力機能を追加改造した jm12.1 を使用した．いずれも量子化ステップ Δ は固定してシミュレーションした．

5.2 MB 種別ごとの評価

本節では提案手法が，フレーム/フィールド MB 種別が継承できた MB で期待された通りの再量子化雑音の低減が実現されているかどうかを以下の輝度 PSNR（以後，単に PSNR）で検証する．

$$PSNR = 10 \log_{10} \left(\frac{N \cdot ERR^2}{\sum (ORG_{i,j} - DEC_{i,j})^2} \right) \quad (1)$$

ただし， N は画素数， $ORG_{i,j}$ は元画像の画素 (i,j) の輝度値， $DEC_{i,j}$ はデコード画像の画素 (i,j) の輝度値で， ERR は輝度の最大値と最小値の差で 8bit 画像では 255 となる．

映像 CheerLeader と Bus を，量子化ステップ $\Delta_1=16$ で MPEG-2 イントラ符号化して得たビットストリームを四つの提案手法で H.264 再符号化したときの MB 種別毎の PSNR と符号量を表 4，表 5 に示す．なお，H.264 再符号化の量子化ステップ Δ_2 は MPEG-2 と同じ値を用い，表 4，表 5 において，MPEG-2 符号化の PSNR (PSNR1 と記載) は，MPEG-2 符号化で混入した量子化雑音から算出した値であり，H.264 再符号化の PSNR (同 PSNR2) は，再符号化で新たに混入した量子化雑音分から算出した値である．

まず，フレーム/フィールド MB 種別が継承できた MB について，H.264 再符号化で新たに混入した雑音に着目する．表 4，表 5 の四つの提案手法で，細字で記載した値は，DCT 種別が継承できた MB に関する PSNR と MPEG-2 に対する bit 減少量であり，太字で記載した値は，継承できなかった MB に関するものである．表 4 の PSNR2 (追加雑音の逆数) で，四つの提案手法とも DCT 種別が継承できた MB では，38.72~42.91dB と高く，DCT 種別が継承できなかった MB では，逆に 33.84~36.96dB と低い．同様に，表 5 で，種別が継承できた MB では，38.13~42.79dB と高く，DCT 種別が継承できなかった MB では，逆に 32.50~36.51dB と低い．両結果から，四つの提案手法ともに，DCT

表 4 H.264 再符号化後の MB 種別毎の PSNR と MBbit 減少量 (入力映像 Cheer Leader)
PSNR and MB bit-reduction for each type of MB after H.264 transcoding. (Cheer Leader)

初段 MPEG2 符号化				H.264 再符号化								
MB 位置	上側 MB	下側 MB	割合 (%)	手法名	手法 r		手法 R		手法 i		手法 I	
				MB 位置	上側 MB	下側 MB	上側 MB	下側 MB	上側 MB	下側 MB	上側 MB	下側 MB
種別	Frame	Frame	68.0	選択種別	Frame		Frame		Frame		Frame	
MBbit 量	342.3	332.7		bit 減少量	28.7	29.2	26.9	27.7	28.3	28.5	25.7	26.8
PSNR1	31.93	34.14		PSNR2	41.47	41.63	41.47	41.63	41.47	41.64	41.47	41.64
種別	Frame	Field	10.7	選択種別	Frame		Frame		Field		Field	
MBbit 量	327.2	244.0		bit 減少量	17.0	26.9	14.5	-25.3	39.5	-44.5	-10.9	-52.4
PSNR1	31.96	34.88		PSNR2	41.69	35.41	41.67	36.96	33.84	43.29	34.95	39.98
種別	Field	Frame	15.1	選択種別	Frame		Frame		Field		Field	
MBbit 量	310.1	317.6		bit 減少量	34.5	23.6	-17.2	14.8	-0.7	18.8	-3.8	-40.5
PSNR1	24.58	34.07		PSNR2	34.51	41.75	36.34	41.76	41.79	33.88	38.72	34.89
種別	Field	Field	6.2	選択種別	Field		Field		Field		Field	
MBbit 量	254.3	265.5		bit 減少量	2.8	26.2	0.0	23.3	2.6	27.8	1.0	24.5
PSNR1	30.80	34.57		PSNR2	42.76	42.91	42.78	42.91	42.79	42.86	42.79	42.88

PSNR1(dB)：オリジナル画像に対する MPEG-2 デコード画像の PSNR(dB)
PSNR2(dB)：MPEG-2 デコード画像に対する H.264 デコード画像の PSNR(dB)

PSNR1(dB) : オリジナル画像に対する MPEG-2 デコード画像の PSNR(dB)
PSNR2(dB) : MPEG-2 デコード画像に対する H.264 デコード画像の PSNR(dB)

表 5 H.264 再符号化後の MB 種別毎の PSNR と MBbit 減少量 (入力映像 Bus)
PSNR and MB bit-reduction for each type of MB after H.264 transcoding. (Bus)

初段 MPEG2 符号化				H.264 再符号化								
MB 位置 種別	上側 MB	下側 MB	割合 (%)	手法名	手法 r		手法 R		手法 i		手法 I	
				MB 位置	上側 MB	下側 MB	上側 MB	下側 MB	上側 MB	下側 MB	上側 MB	下側 MB
種別	Frame	Frame	13.5	選択種別	Frame		Frame		Frame		Frame	
MBbit 量	235.0	258.3		bit 減少量	44.0	41.5	42.1	40.0	41.7	39.3	38.3	38.5
PSNR1	35.79	34.74		PSNR2	42.78	42.04	42.78	42.03	42.79	42.01	42.79	42.01
種別	Frame	Field	7.7	選択種別	Frame		Frame		Field		Field	
MBbit 量	302.6	359.4		bit 減少量	23.1	43.3	20.4	-55.2	-19.1	45.4	-67.3	36.1
PSNR1	33.27	32.60		PSNR2	41.62	32.50	41.63	35.27	33.33	41.34	34.61	38.13
種別	Field	Frame	8.0	選択種別	Frame		Frame		Field		Field	
MBbit 量	221.7	246.0		bit 減少量	40.5	26.4	-2.1	17.3	0.0	24.4	-1.8	13.5
PSNR1	34.29	34.37		PSNR2	34.84	41.99	36.51	41.99	43.19	35.02	39.54	36.07
種別	Field	Field	70.8	選択種別	Field		Field		Field		Field	
MBbit 量	283.7	279.9		bit 減少量	36.3	34.3	35.2	33.3	36.4	34.5	35.4	33.5
PSNR1	33.36	33.41		PSNR2	42.76	42.79	42.76	42.79	42.75	42.79	42.75	42.79
PSNR1(dB)：オリジナル画像に対する MPEG-2 デコード画像の PSNR(dB)												
PSNR2(dB)：MPEG-2 デコード画像に対する H.264 デコード画像の PSNR(dB)												

PSNR1(dB) : オリジナル画像に対する MPEG-2 デコード画像の PSNR(dB)
PSNR2(dB) : MPEG-2 デコード画像に対する H.264 デコード画像の PSNR(dB)

表 6 H.264 再符号化前後の PSNR と MBbit 量 (入力映像 Cheer Leader)
PSNR and MB bits before and after H.264 transcoding. (Cheer Leader)

	初段 MPEG2	H.264 再符号化			
		手法 r	手法 R	手法 i	手法 I
MBbit 量	323.64	296.34	305.10	302.29	311.71
PSNR	31.37	31.04	31.05	30.95	30.98

表 7 H.264 再符号化前後の PSNR と MBbit 量 (入力映像 Bus)
PSNR and MB bits before and after H.264 transcoding. (Bus)

	初段 MPEG2	H.264 再符号化			
		手法 r	手法 R	手法 i	手法 I
MBbit 量	277.01	241.02	247.98	244.46	248.19
PSNR	33.62	33.14	33.28	33.18	33.21

種別が継承できた MB で高い雑音抑圧効果が確認できる。

次に、手法 r と手法 R の差、つまり MPEG-2 でフィールドであった MB を H.264 でフレーム MB として再符号化する際に、MB タイプと予測モードを常に (I8x8, DC 予測) とする手法 r とコスト計算して再判定する手法 R の差に着目する。Cheer Leader では、継承できない MB に再符号化で新たに混入する雑音による PSNR が手法 r の 34.51~35.41dB に対し、手法 R は 36.34~36.96dB と高い。しかし手法 r の MB あたりの発生符号量は、MPEG-2 と比較して、26.9~34.5bit 減少したのに対し、手法 R では 17.2~25.3bit の増加であり、bit 削減効果は手法 r が高い。Bus でも、継承できない MB に再符号化で新たに混入する

雑音による PSNR が手法 r の 32.50~34.84dB に対し、手法 R は 35.27~36.51dB と高い。しかし手法 r の MB あたりの発生符号量は、40.5~43.3bit 減少したのに対し、手法 R では 2.1~55.2bit の増加であり、Cheer Leader と同様に Bus でも bit 削減効果は手法 r が高い。同様な PSNR と削減符号量の関係は、手法 i と手法 I にも見られる。つまり、MPEG-2 と異なるフレーム/フィールド種別で H.264 再符号化すると、MB タイプと予測モードを再判定する方が PSNR が向上するものの、発生符号量は増大することがわかる。

MPEG-2 と同一のフレーム/フィールド種別で H.264 再符号化する場合 (表 4, 5 の細字部分) は、提案四手法の間

図 4 提案四手法の比較
Comparison of four proposed methods.

で大きな違いは見られない。また、すべての MB についてフレーム/フィールド種別や MB タイプなどを再判定する従来方法は、提案四手法に比べて、PSNR と削減符号量ともに大きく劣る。提案四手法の MB ビット量と PSNR の関係を表 6(入力は CheerLeader)、表 7(入力は Bus) に示す。提案手法間で比較すると、手法 R の PSNR は手法 r より若干高いものの、符号量は手法 r より大きい。同様に、手法 I の PSNR は手法 i より若干高いものの、符号量は手法 i より大きい。

提案四手法の優劣を調べるため、量子化ステップを変えて H.264 再符号化したときの RD 曲線を求めた。具体的には、映像 Cheer Leader, Bus を量子化ステップ $\Delta_1=16$ の MPEG-2 符号化を行って得たフレーム/フィールド MB 混在 MPEG-2 イントラビットストリームを、量子化ステップ Δ_2 を 13~26 の範囲で振って H.264 再符号化したときの RD 特性を図 4(a), (b) に示す。横軸は平均 MB 符号量、縦軸は PSNR である。手法 r が、他の提案手法よりも高い RD 特性が得られていることがわかる。これは、手法 R と手法 I は DCT タイプが継承できない MB のモードを再判定するので、モードの切替頻度が、手法 r と手法 i よりも多くなる。MB モード切り替えを表すため bit 量増加のデメリットが、再判定による SN 向上のメリットに比べ大きいため、手法 r と手法 i よりも劣ったと考えられる。また、手法 r と手法 i を比較すると、手法 i が誤ってフレーム映像をフィールド DCT した際の符号化効率低下よりも、手法 r がフィールド映像を誤ってフレーム DCT した際の符号化効率低下が小さく、手法 r が最良になったと考えられる。次節では、最良の結果が得られた提案手法 r について、混入する再符号化雑音と、DCT 種別が継承できない MB の比率の関係について述べる。

5.3 DCT 種別の非継承率と混入する再量子化雑音

初段 MPEG-2 符号化の DCT 種別判定しきい値を振って、フレーム/フィールド DCT の比率が異なる MPEG-2 ストリームを複数作成し、これらを入力として提案手法 r によ

図 5 DCT 種別の非継承率と PSNR 低下量の関係
DCT type mismatch ratio and PSNR degradation.

る H.264 再符号化を行った。得られた H.264 トランスコード映像の PSNR を測定し、MPEG-2 映像からの PSNR 低下量を算出した。7 種の映像に対して上記を行い、量子化ステップは MPEG-2, H.264 とともに 16 とした。結果を図 5 に示す。横軸は、DCT 種別の非継承率 (MPEG-2 の DCT 種別を継承できなかった H.264 の MB 数と全 MB 数の比率)、縦軸は、MPEG-2 デコード映像の PSNR と H.264 トランスコード映像の PSNR の差分である。同図から、DCT 種別の非継承率が大きいほど PSNR 低下量が大きくなることや、同程度の非継承率でも映像によって PSNR 低下量が異なること、がわかる。

5.4 従来方法との比較

量子化ステップ $\Delta_1=16, 26, 36$ で初段 MPEG-2 イントラ符号化されたピクチャを、提案手法 r および従来手法で H.264 再符号化したときの PSNR を図 6 に示す。入力画像は Cheer Leader であり、実線が提案手法 r、破線が従来手法である。同様に入力画像を Mobile and Calendar, Bicycle とした結果を図 7, 図 8 に示す。

提案手法 r は H.264 量子化ステップ $\Delta_2=\Delta_1$ にピークがあり、その点で約 +0.76dB(図 6(a)), 約 +0.93dB(図 6(b)), 約 +0.99dB(図 6(c)), の向上 (従来手法比) が見られる。また $0.5\Delta_1 < \Delta_2 < \Delta_1$ のときに落ち込みがある。この落ち込みはフレーム MB のみの MPEG-2 イントラストリームを再符号化した結果での落ち込みとほぼ一致する結果である。また図 6(a) より $\Delta_1=16$ のとき $\Delta_2 < 22$ の範囲で、図 6(b) より $\Delta_1=26$ のとき $\Delta_2 < 36$ の範囲で、提案手法の SNR が従来手法に比べて高いことから、MPEG-2 符号化時と厳密に同一 Δ でなくても、従来手法を上回る効果が得られていることがわかる。同様な傾向が、図 7, 図 8 でも見られる。

$\Delta_2=\Delta_1$ なる量子化ステップで H.264 再符号化したときの提案手法の圧縮率と PSNR を表 8 に示す。同表の従来手法の PSNR は、提案手法 ($\Delta_2=\Delta_1$) と同等の符号量を発生させるときに得られる PSNR 値を、従来手法の PSNR 測定結果から補間して求めた値を記載した。初段 MPEG-2 符号化と同じ Δ で H.264 再符号化することで、従来手法

表 8 初段符号化時 Δ を用いた提案手法の圧縮比と PSNR
Compression ratio and PSNR of proposed method.

MPEG2 入力				提案手法 r		従来 方式	PSNR 差分 (dB)
映像名	Δ	PSNR (dB)	MB 符号 量 (bit)	MB 符号 量 (bit)	圧縮比	PSNR (dB)	PSNR (dB)
Cheer Leader	16	31.37	324.679	297.031	0.915	31.04	0.76
	26	29.84	238.483	205.979	0.864	29.56	0.93
	36	28.59	191.898	154.483	0.805	28.39	0.99
Bus	16	33.62	278.037	241.855	0.870	33.14	1.27
	26	30.90	199.317	159.621	0.801	30.58	0.95
	36	29.14	157.199	115.591	0.735	28.92	0.82
Mobile and Calendar	16	30.78	455.700	424.557	0.932	30.05	1.03
	26	27.85	329.502	295.313	0.896	27.32	1.04
	36	25.91	257.195	219.698	0.854	25.53	0.93
Sussie	16	37.78	129.041	81.481	0.631	37.68	0.59
	26	36.05	100.794	52.823	0.524	35.92	0.31
	36	34.86	86.8867	38.376	0.442	34.71	0.21
Boat	16	37.11	134.428	99.448	0.740	36.82	0.71
	26	34.95	101.724	64.379	0.633	34.74	0.59
	36	33.48	185.594	45.878	0.536	33.29	0.44
Race	16	35.58	192.786	156.083	0.810	35.22	1.23
	26	33.33	144.473	105.090	0.727	33.15	1.20
	36	31.79	121.050	79.733	0.659	31.74	1.04
Bicycle	16	33.13	301.796	264.468	0.876	32.55	1.50
	26	30.39	216.577	176.296	0.814	30.00	1.21
	36	28.59	169.647	126.779	0.747	28.29	0.95
Average		31.95	230.547	183.655	0.753	31.88	0.93

表 9 提案手法が優位な圧縮比率の下限
Lowest comp. ratio when proposed method
is superior to conventional method.

映像名	Δ	下限 MB 符号量	MPEG MB 符号量	圧縮比率 下限
Cheer Leader	16	233.98	324.679	0.721
	26	154.79	238.483	0.649
	36	なし	191.898	-
Bus	16	194.96	278.037	0.701
	26	140.07	199.317	0.703
	36	100.47	157.199	0.639
Mobile and Calendar	16	366.62	455.700	0.805
	26	252.94	329.502	0.768
	36	188.32	257.195	0.732
Sussie	16	49.48	129.041	0.379
	26	30.02	100.794	0.327
	36	なし	86.886	-
Boat	16	87.15	134.428	0.648
	26	59.07	101.724	0.581
	36	43.08	185.594	0.503
Race	16	95.04	192.786	0.493
	26	90.91	144.473	0.629
	36	70.56	121.050	0.583
Bicycle	16	151.60	301.796	0.502
	26	124.62	216.577	0.576
	36	106.66	126.799	0.629
Average		150.61	238.542	0.609

図 6 提案手法 r と従来手法の比較 (入力映像 Cheer Leader)
PSNR of proposed method and conventional method. (Cheer Leader)

図 7 提案手法 r と従来手法の比較 (入力映像 Mobile and Calendar)
PSNR of proposed method and conventional method. (Mobile and Calendar)

図 8 提案手法 r と従来手法の比較 (入力映像 Bicycle)
PSNR of proposed method and conventional method. (Bicycle)

に比べ 0.44~1.50dB (平均 0.93dB) の PSNR 向上が見られる。この時の圧縮率 (MPEG-2 符号量と H.264 符号量の

比) は 0.442~0.932 (平均 0.753) である。図 6(a),(b), (c) からわかるように、 Δ_2 を Δ_1 を越えて大きくしていく

表 10 再圧縮比 0.714 での提案手法と従来手法の PSNR
PSNR of proposed method and conventional method
when re-compression ratio = 0.714.

映像名	Δ	提案手法 dB(*1)	従来手法 dB(*2)	差分 (*1 - *2)
Cheer Leader	16	29.54	29.57	-0.02
	26	28.16	28.03	0.13
	36	27.41	27.03	0.38
Bus	16	31.01	30.99	0.02
	26	29.28	29.18	0.10
	36	28.64	28.01	0.63
Mobile and Calendar	16	27.35	27.73	-0.38
	26	25.11	25.41	-0.30
	36	23.88	23.98	-0.11
Sussie	16	37.36	37.01	0.36
	26	35.92	35.67	0.25
	36	34.81	34.67	0.14
Boat	16	36.50	36.01	0.49
	26	34.53	34.33	0.21
	36	33.43	33.21	0.22
Race	16	33.87	33.52	0.35
	26	32.94	31.89	1.05
	36	31.25	30.84	0.41
Bicycle	16	30.44	30.11	0.32
	26	28.60	28.28	0.32
	36	27.85	27.19	0.66
Average		30.85	30.60	0.25

と、提案手法 r は SNR が低下し、多くの映像で従来手法に劣るようになる。提案手法 r が従来手法よりも PSNR が高くなる発生符号量の下限を表 9 に示す。

同表で”なし”は、提案手法 r が従来手法よりも低くならなかったことを示す。映像種別によってバラツキがあるが、再圧縮比率=0.327~0.805 (平均 0.609) である。この再圧縮比率よりも大きい再符号化であれば、提案手法は従来手法より PSNR が高い。また、提案手法の適用領域の前提とした、再圧縮率 0.714 での提案手法と従来手法の PSNR を表 10 に示す。評価した 21 ストリームのうち 17 ストリーム (81%) で、提案手法は従来手法に比べて PSNR が高く、平均 0.25dB の向上が得られた。

5.5 提案手法の効果

本手法は、MPEG-2 のイントラストリームだけでなく、イントラ (I) とインタ (P,B) の混在したストリームのイントラ MB 部分にも適用可能である。また、本提案手法を使って高画質の I ピクチャを得ることで、I ピクチャを参照する P ピクチャの符号化効率も向上させることができる。

6. む す び

インタレースコンテンツを対象とした MPEG-2/H.264 イントラトランスコーディングにおいて、DCT 種別を極力継承することで量子化雑音を削減する変換方法を提案した。MPEG-2 の DCT 種別の継承が H.264 の DCT 種別指定の制約により、不可能な場合があるものの、その割合は MPEG-2 の実ストリームを用いた解析により、全 MB のわずか 7%程度と小さいことから DCT 種別を極力継承する MPEG-2/H.264 再符号化手法を提案した。提案手法 r を実装した MPEG-2/H.264 トランスコーダを用いた再符号化シミュレーションの結果、MPEG-2 と H.264 で同一の量子化ステップを使用した場合の再圧縮率 0.442~0.932(平均 0.753) において、提案手法は従来と比較して 0.44~1.50(平

均 0.93)dB の改善が得られた。また、IP 再送信への適用を前提とした再圧縮率 0.714 では、評価した 21 ストリームのうち 17 ストリーム (81%) で提案手法は従来手法に比べて高い PSNR が得られ、平均 0.25dB の向上が得られた。

また提案手法は、時間のかかる MB モード選択等が不要であるため、従来に比べ極めて少ない演算量で実現できる特徴もある。今後はインタレースの P ピクチャなどの画面間予測を含む MPEG-2 ストリームへの適用を進める。

〔文 献〕

- 1) ISO/IEC IS 13818-2, ITU-T Recommendation H.262, “Generic coding of moving pictures and associated audio information” (Nov. 1994)
- 2) “Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264/ISO/IEC 14486-10 AVC”, in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050 (2003)
- 3) P.Guilotel, et al “Adaptive Encoders: The New Generation of MPEG-2 Encoders”, SMPTE journal (Apr.2000)
- 4) 中村, 吉留他: “H.264/MPEG-4AVC における再符号化方法の検討”, 信学全大, **D-11-50** (Mar. 2005)
- 5) 角野他: “MPEG-2 再符号化のためのゾーン処理付き再量子化法”, 信学論, **J83-D-II,9**, pp.1879-1887 (Sep. 2000)
- 6) 嵯峨田他: “地上デジタル IP 再送信用トランスコーダ装置: MPEG-2/H.264 トランスコーダ装置の構成”, 信学全大, **D-11-10** (Mar. 2009)
- 7) 吉留, 長沼, 八島: “イントラ符号化された MPEG-2 プログレッシブ映像の H.264 再符号化方法の検討”, 映像学誌, **62**, 11, pp1819-1824 (Nov. 2008)
- 8) ITE 監修 ハイビジョン用デジタル標準画像データ
http://www.ite.or.jp/shippan/test_chart_index.html
- 9) MPEG-2, Test Model 5 (TM5), Doc ISO/IEC JTC/SC29/WG11/N0400, Test Model Editing Committee (Apr. 1993)
- 10) Joint Video Team(JCT), “Reference Software JM12.1”,
<http://iphome.hhi.de/suehring/tml/>
- 11) 大久保他: “改訂版 H.264/AVC 教科書”, インプレス (2006)
- 12) 総務省 情報通信審議会「地上デジタル放送の公共分野における利活用に関する調査研究」
http://www.soumu.go.jp/joho.tsusin/dtv/toushin/pdf/hokoku_0703.pdf

よしとめ 健 1982 年 筑波大学第三学群情報学類卒業。1983 年 同大学院理工学研究科修士課程修了。同年電電公社厚木電気通信研究所 (現 NTT) 入社。以来、映像符号化 LSI および超高精細映像 CODEC の構成法・設計法、の研究開発に従事。2001 年 NTT 社長表彰。2009 年筑波大学大学院システム情報工学研究科後期博士課程コンピュータサイエンス専攻に社会人入学。2010 年同修了。現在、NTT サイバースペース研究所画像メディア通信プロジェクト主任研究員。工博。正会員。

かみくら かずと 一人 1984 年、東京理科大学理工学部電気工学科卒業。1986 年、同大学院修士課程修了。同年、日本電信電話 (株) 入社。複合通信研究所およびヒューマンインタフェース研究所において画像符号化の研究に従事。1993 年、NTT 画像通信事業本部で自治体向けマルチメディアシステムの開発に従事。1995 年、ヒューマンインタフェース研究所において画像符号化の研究に従事。現在、NTT サイバースペース研究所画像メディア通信プロジェクト主幹研究員。工博。正会員。

きたわき のぶひこ 信彦 1969 年、東北大学工学部電子工学科卒業。1971 年、同大学院修士課程修了。同年、日本電信電話公社武蔵野研究所入所。1997 年から筑波大学教授。2010 年から同大特任教授。工博。知覚 QoS 評価法、音声符号化、音信号処理、e ラーニングの研究に従事。1981 年、電電公社総裁表彰。1889 年、電気通信普及財団賞。1990 年、NTT 社長表彰。1994 年、日本 ITU 協会賞。1995 年、R C R 郵政大臣表彰。IEEE Fellow。