

# CTB 構造におけるイントラ予測残差信号のチャンネル間予測符号化の評価

## Evaluation of Inter-channel Prediction Coding for Intra Prediction Residual Signals under CTB Structure

河村圭<sup>†</sup>  
Kei Kawamura

加藤晴久<sup>†</sup>  
Haruhisa Kato

内藤整<sup>†</sup>  
Sei Naito

### 1 まえがき

動画符号化方式である H.264/AVC の次世代方式として、HEVC (High Efficiency Video Coding) が JCT-VC (Joint Collaborative Team on Video Coding) で検討されている [1]。HEVC では UHD (ultra-HD) と呼ばれるプログレッシブの 4K2K 映像や 8K4K 映像までを対象としている。また、これまで固定サイズであった MB (Macro Block) に代わり、可変サイズを前提とした CU (Coding Unit) と四分木構造に基づく CTB (Coding Tree Block) の導入が議論されている。

JCT-VC ではイントラ予測ツールのひとつとして、YUV420 映像の色成分間 (以下、チャンネル間) 相関を利用して冗長度を削減する方式が議論されている。この方式 [2] (以下、画素信号予測方式) では、同一ブロックの再構成した輝度信号から色差信号を線型予測することと、予測係数を符号化または復号済みの周辺信号から推定し符号化を省略する特徴がある。

一方、スーパーハイビジョンと呼ばれる超高精細映像のうち RGB444 映像を対象として、非常にビットレートの高い素材蓄積用符号化方式が検討されている。この方式 [3] (以下、残差信号予測方式) では、RGB444 映像に存在するチャンネル間相関を用いて、被予測信号のイントラ予測残差信号を参照信号から線型予測することと、最適な予測係数を原画像を用いて導出して符号化する特徴がある。

本稿では、YUV420 映像かつ放送用ビットレートでも効果があり、HEVC の CTB 構造に適用可能なチャンネル間予測方式を提案する。具体的には、RGB444 映像の高ビットレート符号化において高い性能を実現している残差信号予測方式をベースとする。まず、CTB 構造では予測係数の解析的な導出を保証できない課題がある [4]。そこで、予測係数を網羅的に変更して最適な予測係数を決定する。つぎに、YUV420 映像におけるチャンネル間の相関度合いは素材依存性が高いため、利用の有無をフラグで符号化すると利用されないときの冗長性が高くなる。そこで、イントラ予測の追加モードとして予測係数と共に符号化する。最後に、提案手法を BD-bitrate と主観品質によって評価する。

### 2 チャンネル間予測の従来手法

#### 2.1 画素信号予測方式

Chen らは再構成 (ローカルデコード) した輝度信号を用いて、色差信号を線型予測するイントラ予測モードを提案している [2]。色差信号は同一ブロックの再構成輝度信号から以下のモデル式によって予測される。

$$\text{Pred}_C[x, y] = \alpha \times \text{Rec}'_L[x, y] + \beta \quad (1)$$

ここで、 $\text{Pred}_C$  はブロック内の色差予測信号、 $\text{Rec}_L$  はブロック内の再構成輝度信号を表す。モデル式のパラメータ  $\alpha$  と  $\beta$  は、当該ブロックの周囲にある符号化済みの輝度信号と色差信号を用いて、符号化側と復号側で最小二乗法により導出する。

ローカルデコードされた輝度信号は色差信号とサイズおよび位相を合わせるため以下の式のように、垂直方向にダウンサンプリング、水平方向にサブサンプリングされる。

$$\text{Rec}'_L[x, y] = (\text{Rec}_L[2x, 2y] + \text{Rec}_L[2x, 2y+1]) \gg 1 \quad (2)$$

画素信号予測方式では、当該ブロックのチャンネル間に相関があったとしても、周囲の領域において相関がなければパラメータの導出が不正確になる課題がある。また、すでに量子化された信号を用いてパラメータを導出するため、パラメータ推定精度が不十分となり、主観品質が劣化している。

#### 2.2 残差信号予測方式

加藤らは、RGB444 映像の素材蓄積用符号化の実現を目的として、チャンネル間相関を削減する符号化方式を提案している [3]。H.264/AVC のイントラ予測方式を拡張し、被予測信号の予測残差信号を以下のモデル式によって予測する。

$$\text{PredResi}_C[x, y] = \alpha_C \times \text{Resi}_L[x, y] + \beta_C \quad (3)$$

ここで、 $\text{PredResi}_C[x, y]$  は被予測信号における予測残差値の予測信号である。また、 $\text{Resi}_L$  は被予測信号の予測モードと同一の予測モードを、再構成した基準信号に適用して算出した予測残差信号である。

予測係数  $\alpha_C$  は原画像を用いてチャンネルごとの最小二乗法により解析的に符号化側で導出する。さらに、符号化済みの周辺ブロックにおける  $\alpha_C$  を用いて、当該

<sup>†</sup>株式会社 KDDI 研究所, KDDI R&D Laboratories Inc.

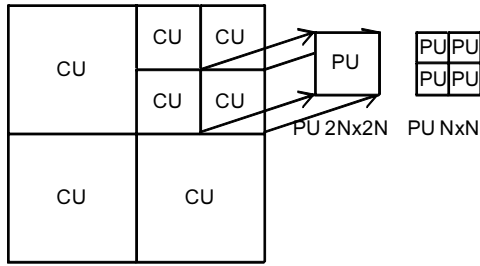


Fig. 1: Quad-tree structure of CU and PU.

ブロックにおける  $\alpha_C$  を推定し、差分値を量子化して符号化する。

予測係数  $\beta_C$  は量子化された  $\alpha_C$  を適用後に導出するため、 $\alpha_C$  の量子化誤差を考慮した値となる。さらに、 $\beta_C$  は変換ブロック内で同一値となるため、後段の整数変換によって DC 係数にのみ加算される。したがって、予測の段階で明示的に算出する必要もなく、さらに符号化する必要もない。すなわち、 $\beta_C$  が加算された DC 係数を符号化することになる。

被予測信号におけるイントラ予測モードは、被予測信号内で最適なものが選択される。そのため、参照信号と被予測信号とで最適な予測モードが異なる場合が発生する。このような場合には、復号された参照信号に対して被予測信号の予測モードを適用して残差信号を再計算する処理が発生する。

このように、被予測信号の予測残差信号を予測することでチャンネル間相関の低減を実現し、さらに予測精度を高めるために原画像を用いて導出した予測係数を符号化する特徴がある。一方で、後述する CTB 四分木構造にそのまま適用することが難しい。

### 2.3 CTB 四分木構造に起因する課題

JCT-VC では CTB 四分木構造の導入が議論され、CU, PU (Prediction Unit), TU (Transform Unit) のシンタックスが Fig. 1 のように階層・再帰構造になっている。PU は予測パラメータを共有するのみのユニットであり、実際の予測・変換処理は TU で行われる。チャンネル間予測は予測方式の一種であるから、PU 情報として符号化するのが素直であるが、PU サイズが TU サイズよりも大きな場合には以下の課題が生じる。

予測処理は PU 単位ではなく TU 単位で行われるため、TU ごとに予測係数  $\alpha_C$  の最適値が異なる。TU ごとに最適な予測係数を符号化すると、付加情報が増大する課題が生じる。PU ごとに  $\alpha_C$  を符号化すると、付加情報は低減するが、 $\alpha_C$  の決定において別の課題が生じる。すなわち、将来符号化される TU ブロックを考慮する必要があるため、最小二乗法などによる解析的

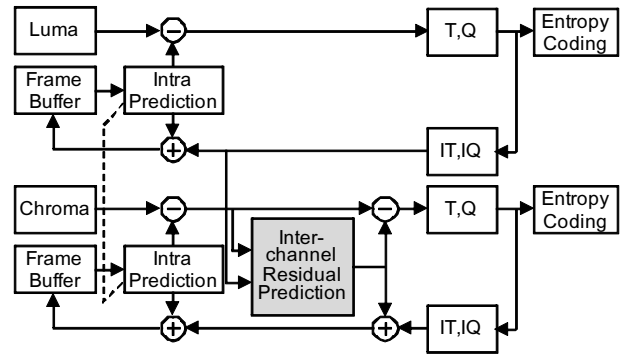


Fig. 2: Flowchart of the inter-channel residual prediction.

な導出ができない。これまでは、最初に符号化される TU ブロックにおいて最小二乗法により求められた  $\alpha_C$  を PU 内の TU で共有する手法を提案してきた [4]。しかしながら、最適な予測係数であることを保証できず、符号化性能の向上が限定的であった。

### 3 提案手法

本稿では、YUV420 映像かつ放送用ビットレートでも効果があり、CTB 四分木構造に適用可能な残差値予測方式を提案する。まず、予測モデルは以下の式である。

$$\text{PredResi}_C[x, y] = \alpha_C \times \text{Resi}_L[2x, 2y] + \beta_C \quad (4)$$

$\text{PredResi}_C[x, y]$  は色差信号における予測残差信号の予測信号であり、 $\text{Resi}_L$  は輝度信号における符号化済みの予測残差信号である。予測係数  $\alpha_C$  はチャンネルごとに符号化側で決定し、量子化して符号化する。予測係数  $\beta_C$  は加藤らの手法と同様に明示的に符号化しない。なお、輝度信号の予測残差信号は単純にサブサンプリングしているが、ローパスフィルタの有無による性能比較を行い、有意な差がないことを予備実験により確認している。

ここで、予測係数  $\alpha_C$  の導出について述べる。まず、PU 内では予測残差信号の性質も似ていると仮定して、それぞれの色差ごとに  $\alpha_C$  を PU 内の TU で共有して符号化する。これにより、付加情報の増大を避ける。つぎに、 $\alpha_C$  は離散値であるため網羅的に仮符号化を実行し、レート歪特性に基づいて最適な予測係数を決定する。これにより、従来手法でパラメータ推定精度が不十分になることを避ける。また、 $\alpha_C$  の解析的な導出が出来なくても、最適な  $\alpha_C$  を決定できる。なお、 $\alpha_C$  の量子化幅と定義域は予備実験により  $1/8$  と  $[-1, +1]$  とした。

チャンネル間予測のシンタックスについて述べる。残差信号予測モデルにおいて、予測係数  $\alpha_C$  と  $\beta_C$  を 0

Table 1: BD-bitrate comparison between proposed method and the conventional method [2]. [%]

Source	Proposed method				Conventional method			
	Y	U	V	Total	Y	U	V	Total
Class A	-1.0	-21.3	-19.4	-4.37	-1.3	-16.2	-8.3	-3.18
Class B	-0.3	-6.4	-3.1	-1.27	-0.8	-7.2	-3.7	-1.77
Class C	0.0	-1.3	-2.1	-0.38	-1.1	-6.7	-7.4	-2.35
Class D	0.1	-0.8	-1.0	-0.14	-0.7	-4.7	-4.6	-1.45
Class E	0.0	-0.8	-1.0	-0.15	-0.1	-3.3	-3.7	-0.63
Overall	-0.3	-6.4	-5.4	-1.32	-0.8	-7.8	-5.6	-1.93

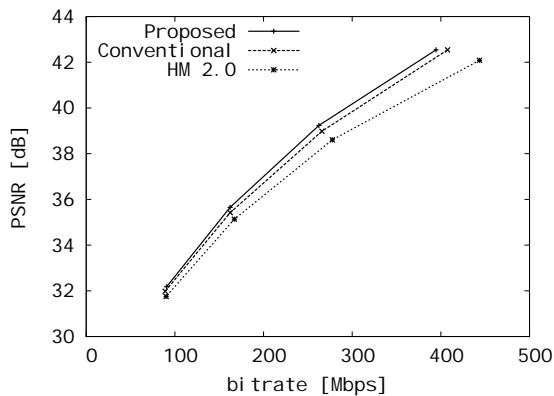


Fig. 3: Rate-Distortion characteristics of Nebuta

にすると、予測を行わないのと同じである。色差予測モードを符号化し、かつ輝度予測モードと同一の時は必ず予測係数を符号化する手法がある [4]。チャンネル間予測が選択されない場合には予測係数 0 を送らなければならないため、冗長となる可能性があった。本稿では、既存のモードに加えて残差信号予測を行うモードを追加する。このモードが選択されると、色差予測方向は輝度予測方向と同一とみなす。すなわち、予測残差信号を生成する予測モードが常に一致するため、加藤らの手法のように再計算が発生する可能性がない。また、既存のモードには変更を加えないため、性能低下の回避が期待できる。このように、残差信号予測は PU 単位で切り替える。

以上をまとめた符号化側の処理フローを Fig. 2 に示す。図中の点線は、色差予測方向が輝度予測方向と同一であることを示している。T, Q は整数変換と量子化を, IT, IQ は逆整数変換と逆量子化である。網掛けした Inter-channel Residual Prediction ブロックは、提案方式を示しており、輝度予測残差と色差予測残差を入力として、色差予測残差の予測信号を出力する。

#### 4 実験結果と考察

提案手法を HM (HEVC Test Model) 2.0 [6] に実装し, JCT-VC の共通実験条件 [7] に従って性能評価した。

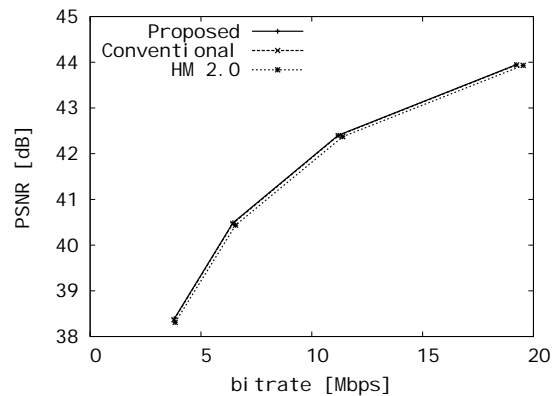


Fig. 4: Rate-Distortion characteristics of Kimono

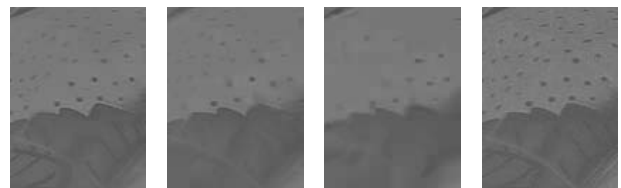


Fig. 5: Decoded images and the original image of Nebuta, QP=37, frame no. 0, crop[px]: 1184×480-1280×608. From left to right: the proposed method, the conventional method [2], HM 2.0 [6], and the original image.

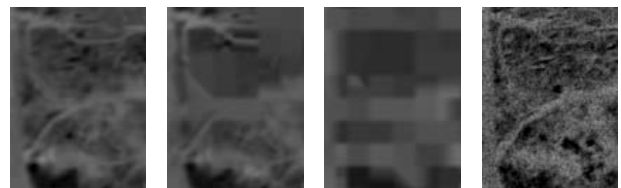


Fig. 6: Decoded images and the original image of Kimono, QP=37, frame no. 0, crop[px]: 1152×496-1248×624. From left to right: the proposed method, the conventional method [2], HM 2.0 [6], and the original image.

提案方式はイントラモードに適用する方式であるため, Intra high-efficiency 条件のみ比較した。なお, Class A は 4K 素材と呼ばれているが, 実際には 2560×1600p のクロップ画像 (4 素材) である。また, Class B 以下は 1920×1080p の 5 素材, 832×480p の 4 素材, 416×240p

の4素材, 1280×720pの3素材である。

比較結果を表1に示す。表の数値はJCT-VCで評価軸として採用されているBD-bitrateで、負数はビットレートを削減したことを意味している。YUVはそれぞれの色成分ごと、Totalは全成分を対象としたPSNRと、全成分の合計ビットレートから算出している。Proposed methodはチャンネル間予測のないHM 2.0に対する提案方式の性能を示しており、Conventional methodはHM 2.0に対する画素値予測方式(以下、従来方式)の性能を示している。

提案方式と従来方式は輝度信号の処理に変更がないため、表におけるYは全成分の合計ビットレート増減により変化している。HM 2.0に対するBD-bitrate削減は、U成分で6.4%、V成分で5.4%に達している。ただし、色差信号が占める画素数と符号量は輝度信号よりも少ないため、TotalのBD-bitrate削減は1.32%にとどまっている。一方、従来方式と比較すると全Classの平均(Overall)は性能劣化している。これは、サイド情報である予測係数の占める割合が低解像度になるほど大きくなるためである。Class Aでは提案方式のほうがBD-bitrate削減量が多いことから、提案方式の優位性は高解像度において顕著であるといえる。解像度が高くなるほど大きなサイズのUnitが選択されやすくなる一方で、イントラ予測の性能は低下する。その結果、信号間の冗長性が残り、提案する残差信号予測によるゲインが得られやすくなるためである。

なお、アンカー方式に対する符号化時間は、提案方式で160%程度、従来方式でほぼ100%であった。提案方式で大幅に符号化時間が増加しているのは、網羅的に試行したためである。そのため今後、高速な決定手法を検討する。

次に、Class Aから素材Nebutaについて、Class Bから素材Kimonoについて、Totalのレート歪特性をFig. 3とFig. 4にそれぞれ示す。この図より、チャンネル間予測はレートが高いほど性能が高い。素材Nebutaでは高ビットレートになるほど提案手法は従来手法より性能が高く、素材Kimonoでは提案方式と従来方式のグラフは重なっており、ほぼ性能は同じである。

さらに、PSNRでは評価できない主観品質について、QP=37におけるNebutaとKimonoのU成分のデコード画像と原画像の一部をFig. 5とFig. 6にそれぞれ示す。QP=37とは、レート歪特性図において最もビットレートの低い点であり、それぞれの手法がほぼ同一のビットレートとPSNRになっている。また、Fig. 6は差異を目立たせるためにコントラストを上げている。これらの図より、HM 2.0は細かなテクスチャが完全に

潰れてブロック状の平坦な画像になっている。従来方式は予測係数を復号器側で導出するため最適ではなく、テクスチャの再現性が不十分である。提案方式はテクスチャの再現性が非常に高く、これらの方式の中では最も主観品質が高い。

一連の実験により、解像度が高いほどBD-bitrateと主観品質について提案手法は従来手法より優れている。そこで、従来方式に提案方式を併用することで、BD-bitrateにおける性能向上の可能性がある。

## 5 むすび

本稿では、YUV420映像かつ放送用ビットレートでも効果があり、HEVCのCTB構造に適用可能なチャンネル間予測方式を提案した。提案方式はHM 2.0に対して1.32%のBD-bitrate削減を確認した。また、提案方式は解像度が高いほど優れているが、全体としては従来方式に対して0.59%のBD-bitrate増加となった。主観品質においては、提案方式はテクスチャの再現性が非常に高いことを確認した。今後は、画素信号予測方式と残差値予測方式の併用方式について検討する。謝辞 本研究は独立行政法人情報通信研究機構による委託研究「超高精細映像符号化技術に関する研究開発」として実施したものである。

## 参考文献

- [1] VCEG, "Joint Call for Proposals on Video Compression Technology," ISO/IEC JTC1/SC29/WG11/N11113, Jan. 2010.
- [2] J. Chen, et al. "CE6.a.4: Chroma intra prediction by reconstructed luma samples," JCTVC-E266, Mar. 2011.
- [3] 加藤, 内藤, 酒澤, 松本, "H.264のIntra予測残差に対する適応的チャンネル間予測," 映像情報メディア学会誌, 64(11), pp.1711-1717, Nov. 2010.
- [4] 河村, 加藤, 内藤, "CTB構造におけるイントラ予測残差信号のチャンネル間予測符号化," 電子情報通信学会 総合大会, DS-2-5, Mar. 2011.
- [5] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," VCEG-M33, Apr. 2001.
- [6] K. McCann, "High Efficiency Video Coding (HEVC) Test Model 2 (HM 2) Encoder Description," JCTVC-D502, Jan. 2011.
- [7] F. Bossen, "Common test conditions and software reference configurations," JCTVC-D600, Jan. 2011.