

## VVCにおける LFNST 制御に関する一検討 A Study on LFNST Control for VVC

木谷 佳隆<sup>†</sup>      海野 恭平<sup>†</sup>      河村 圭<sup>†</sup>  
Yoshitaka Kidani      Kyohei Unno      Kei Kawamura

### 1. はじめに

最新の動画像符号化国際標準方式である Versatile Video Coding (VVC) では、変換ツールの一つとして、予測残差信号に対して水平・垂直方向に分離して適用される一次変換に加え、一次変換後の係数に対して非分離で適用される二次変換の Low-Frequency Non-Separable Transform (LFNST) が採用されている[1]。LFNST は、一次変換後の係数に残存する係数分布の偏りを変換して、さらに低周波数域に係数を集中させて符号量を削減する一方で、エンコーダの演算量を増大させる。本稿では、従来手法における LFNST に伴うエンコーダの演算量の削減と、LFNST のシンタックス復号処理に関するハードウェアデコーダ実装観点で問題解消を目的とした、LFNST の制御手法を提案する。なお、本提案手法は、VVC に採用されている。

### 2. 従来手法

本節では、LFNST の適用対象係数領域およびデコーダにおける LFNST 適用有無の制御に関する従来手法[2]を示す。

まず、LFNST が適用される一次変換後の係数領域は、符号化ブロック(変換ブロック)のブロックサイズに依存する。LFNST 非適用領域には一次変換係数が残存し、LFNST 適用領域には二次変換係数が発生するが、特定のブロックサイズでは LFNST 適用領域の高周波域にある係数がゼロイングされる。これにより、デコーダは符号化ビットストリームからの係数復号(係数スキャン)時に、黄色領域に非ゼロ係数が存在する場合、復号対象ブロックでは LFNST が非適用であると特定できる。この非ゼロ係数の検出にカウンタ(以下、numZeroOutSigCoeff)を導入している。また、このカウンタに加えて、LFNST が適用される係数領域も含めた全領域で非ゼロの係数の存在有無を検出するカウンタ(numSigCoeff)を導入している。導入理由は詳述する。

次に、従来手法のデコーダにおける LFNST の適用有無は、エンコーダから伝送されるインデックス(lfnst\_idx)の復号結果に基づき、符号化ブロック単位で判定される。具体的に、lfnst\_idx は 0, 1, 2 の値を有し、値が 0 の場合は LFNST が非適用、1 と 2 の場合は LFNST が適用と判定される。ここで、従来手法では numZeroOutSigCoeff が 0 であり、かつ numSigCoeff が 1 より大きい(非ゼロ係数が 2 個以上存在する)場合、デコーダは lfnst\_idx を復号し、これを満たさない場合は lfnst\_idx を復号せず、lfnst\_idx の値を 0 と推定する。これにより、従来手法は LFNST が非適用であることが自明なブロック、および LFNST の有効性が低い(係数が極端に少ない)ブロックに対する lfnst\_idx の伝送符号量を削減している。

しかしながら、従来手法には以下の 3 つの課題がある。第 1 に、LFNST に伴う特にエンコーダの演算量の削減が十分ではないことが報告されている[3]。第 2 に、従来手法の

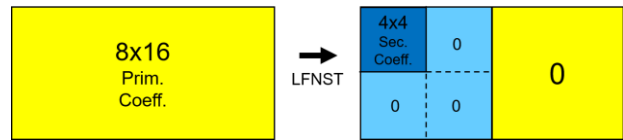


図 1 提案手法の高周波域の全ゼロイングの一例

2 つのカウンタはデコーダ上でメモリ領域を必要とするため、メモリ容量の制約が厳しいハードウェアデコーダ観点では望ましくなく、さらに、いずれも非ゼロ係数の検出に係数スキャンが必要なため、デコーダの CABAC ステージのスループットを低下させる。第 3 に、lfnst\_idx の復号制御に対象ブロックのブロックサイズ上限に関する条件がないため、VVC で採用されている Virtual Pipeline Decoding Unit (VPDU) [4]の制約をワーストケースで破綻させる。ここで、VPDU はハードウェア実装におけるパイプライン構成の 1 ステージの処理単位を変換ブロックの最大サイズ(64x64 画素)に区切って実装可能とするための概念である。具体的に、高さや幅のいずれかが 64 画素より大きい符号化ブロックに LFNST が適用可能な場合、同ブロックを構成する全変換ブロックすなわち VPDU の係数復号処理が完了しなければ、lfnst\_idx の復号処理が開始できない。

### 3. 提案手法

本節では、2 節の最後に示した 3 つの課題をそれぞれ解消する提案を詳述する。提案は、高周波数域の全ゼロイング、非ゼロ係数終端位置情報を用いた係数検出、lfnst\_idx 復号条件へのブロックサイズ上限の導入、から構成される。

#### 3.1 高周波数域の全ゼロイング

LFNST のエンコーダの演算量を削減するため、高周波数域の全ゼロイングを提案する。まず、図 1 の例のように LFNST の適用対象とする一次変換係数を 8x8 画素の変換ブロックの低周波数域の最大 48 係数までに削減し、二次変換係数を 4x4 画素の変換ブロックの最大 16 係数までに限定する。さらに、LFNST が適用される場合、一次変換係数も含めて上述の最大 16 係数以外は全てゼロイングする。これにより LFNST 適用時の一次変換に必要な演算量が削減される。

#### 3.2 非ゼロ係数終端位置情報を用いた係数検出

3.1 節の一次変換係数のゼロイングにより、LFNST の適用ブロックで非ゼロ係数が存在する領域は対象ブロックの低周波数域の最大 16 係数に限定される。これにより、非ゼロ係数の検出手法として、非ゼロ係数終端位置情報(LastScanPos)を従来手法の 2 つのカウンタに置換できる。LastScanPos は、VVC に導入されているシンタックス last\_sig\_coeff から算出可能なデコーダの内部変数である。まず、LastScanPos がゼロイング領域に存在する場合は LFNST が非適用のブロックと特定する。次に、LastScanPos が DC 成分を指す場合は、同じく LFNST が非適用のブロックと特定する。以上より、デコーダから 2 つのカウンタを

<sup>†</sup>株式会社 KDDI 総合研究所 KDDI Research, Inc.

表1 VTM 5 w/o LFNSTに対する従来手法と提案手法の性能評価実験結果

Class	従来手法						提案手法					
	AI			RA			AI			RA		
	BD-rate	EncT	DecT	BD-rate	EncT	DecT	BD-rate	EncT	DecT	BD-rate	EncT	DecT
A1	-1.79%	128%	100%	-1.41%	117%	101%	-1.75%	115%	99%	-1.43%	113%	100%
A2	-1.04%	135%	101%	-0.72%	115%	101%	-1.02%	119%	100%	-0.72%	112%	101%
B	-0.95%	130%	101%	-0.54%	116%	100%	-0.84%	116%	100%	-0.50%	113%	100%
C	-1.25%	130%	101%	-0.65%	120%	101%	-1.14%	113%	101%	-0.63%	117%	100%
E	-1.38%	125%	100%				-1.27%	113%	99%			
Overall	-1.24%	130%	101%	-0.78%	117%	100%	-1.16%	115%	100%	-0.77%	114%	99%

排除し、さらにデコーダは係数スキャン処理なしに、LFNST が非適用であることが自明なブロック、および LFNST の有効性が低い（係数が極端に少ない）ブロックに対する `lfnst_idx` の伝送符号量を削減できる。

### 3.3 `lfnst_idx` 復号条件へのブロックサイズ上限の導入

3.2 節の `LastScanPos` を用いた `lfnst_idx` の復号制御に加え VPDU を考慮したブロックサイズ上限を `lfnst_idx` の復号条件に導入する。具体的に、復号対象の符号化ブロックサイズの高さや幅がいずれも VPDU を規定する変換ブロック最大サイズ（64 画素）以下である場合、`lfnst_idx` を復号する。これにより、デコーダの CABAC ステージにおける符号化ブロック単位の `lfnst_idx` の復号処理が、符号化ブロックを構成するすべての変換ブロックすなわち VPDU の処理完了に依存せず、開始できる。

## 4. 実験結果

### 4.1 実験条件

提案手法の有効性を評価するため、シミュレーション実験を実施した。実験に使用した PC の CPU は Xeon E5-2640 v4 2.4GHz、メモリは 128GB である。性能はアンカーに対する輝度成分の BD-rate と、エンコーダおよびデコーダの演算量の比率 EncT/DecT（シングルスレッド測定）で評価した。ここで、BD-rate は、同一客観画質で比較した符号量の削減率を示す指標であり、負値が符号化性能の改善を示す。従来手法が包含された VVC の参照ソフトウェアである VVC Test Model version 5 (VTM 5) [4]に提案手法を実装し、LFNST を無効化した VTM-5.0 をアンカーとして性能を比較した。符号化条件は、VVC の標準化作業グループである JVET (Joint Video Expert Team) の Common Test Condition (CTC) [5]の All Intra (AI) 条件と Random Access (RA) 条件に従った。テスト素材は、CTC で規定されている異なる 5 つの解像度クラスのカメラ撮影映像を使用した。ここで、Class A1/A2, Class B, Class C, Class E の解像度はそれぞれ 3840x2160, 1920x1080, 832x420, 1280x720 画素である。

### 4.2 実験結果

表 1 に LFNST を無効化した VTM-5.0 をアンカーとした従来手法と提案手法の LFNST の性能が示される。

まず、従来手法の LFNST は、AI/RA 条件の全クラス平均で、符号化性能を -1.24%/-0.78% 改善する一方で、エンコーダの演算量を 30%/17% 増加させることが確認できる。なお、これらはクラス別の性能比較でも同様の傾向である。AI 条件に対して RA 条件の BD-rate と EncT の絶対値が小さい理

由は、LFNST はイントラスライスとインタースライスのイントラブロックのみに適用されるためである。

次に、提案手法の LFNST は、AI/RA 条件の全クラス平均で、符号化性能を -1.16%/-0.77% 改善する一方で、エンコーダの演算量を 115%/114% 増加させることが確認できる。これらはクラス別の性能比較でも同様の傾向である。両手法の全クラス平均の性能比較により、提案手法は従来手法の LFNST が有する符号化性能をほぼ維持しつつ、エンコーダの演算量を 15%/3% 削減している。

加えて、実験結果からは LFNST の導入に伴うデコーダの演算量増加や提案手法による演算量の削減はほとんど確認できない。前者の理由は、VTM 5 に導入されている LFNST 以外の符号化ツールが大半を占めるためである[3]。後者の理由は、実験結果が CPU による測定結果であるためであり、ハードウェア実装のデコーダに対して効果が顕れやすい。

## 5. まとめ

本稿では、従来手法における LFNST に伴うエンコーダの演算量の削減と LFNST のシンタックス復号処理に使用される係数検出用カウンタの排除および CABAC のスループット維持を目的とした、LFNST の制御手法を提案した。提案手法は、高周波数域の全ゼロイング、復号制御への非ゼロ係数終端位置情報を用いた係数検出、および `lfnst_idx` 復号条件へのブロックサイズ上限の導入から構成される。実験結果は、提案手法が、従来手法の LFNST が有する符号化性能を維持しつつ、従来手法の LFNST により増加するエンコーダの演算量を JVET CTC の AI/RA 条件でそれぞれ 15%/3% 削減することを示した。

### 謝辞

本研究開発は総務省 SCOPE（国際標準獲得型）JPJ000595 の委託を受けたものです。

### 参考文献

- [1] Recommendation ITU-T H.266, "Versatile video coding," in Edition 1.0, August 2020.
- [2] M. Koo, J. Lim, M. Salehifar, and S. H. Kim, "CE6: Reduced Secondary Transform (RST) (CE6-.3.1)," in ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Geneva, Mar. 2019, JVET-N0193.
- [3] W. J. Chien, et al., "JVET AHG report: Tool reporting procedure (AHG13)," in ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Geneva, Mar. 2019, JVET-N0013.
- [4] J. Chen, Y. Ye, and S. H. Kim, "Algorithm description for Versatile Video Coding and Test Model 5 (VTM 5)," in ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Geneva, Mar. 2019, JVET-N1001.
- [5] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Sühling, "JVET common test conditions and software reference configurations for SDR video," in ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Geneva, Mar. 2019, JVET-N1010.