

DCT と DST を用いた適応直交変換符号化

The adaptive orthogonal coding method by DCT and DST

市ヶ谷敦郎† 杉藤泰子†
Atsuro Ichigaya Yasuko Sugito

境田慎一† 鹿喰善明†
Shin-ichi Sakaida Yoshiaki Shishikui

1. まえがき

MPEG-2 や AVC/H.264 で代表される映像符号化方式では動き補償予測、直交変換、量子化、可変長符号化を組み合わせて効率的に情報圧縮を実現している。映像符号化方式における直交変換の利用は映像信号の空間的な相関を除去し、変換領域で少ない係数に映像信号の情報(エネルギー)を集中させることにある。入力信号を最も効率的に変換する直交変換として KLT (Karhunen Loève transform) が知られている。KLT はどのような入力信号に対しても最適な変換効率を与えるが、演算量が多く、ハードウェア化が困難であり、符号化分野ではほとんど用いられていない。しかし、映像信号は信号間の相関が高い信号であり、そのような信号に対する変換特性が KLT に近く、演算量の少ない DCT (Discrete Cosine Transform) が多くの符号化方式で採用されている。

最新の符号化方式である AVC/H.264 もまた DCT を採用し、予測方式として動き補償予測に加え新たにイントラ予測を採用した。映像信号をそのまま符号化対象とはせず、予測によって空間的相関のある程度除去した予測差分信号のみを変換対象としている。さらに現在検討が進められている次世代符号化方式 HEVC (High efficiency video coding) では、より高度な予測技術が検討されており、予測差分信号の相関がさらに除去される傾向にある。このような相関が除去された信号に対しては DCT の変換効率は低下する。このような信号に KLT を適用した場合、その特性は DST (Discrete Sine Transform) に近づくことが知られている^[1]。

DST は相関の低い信号に対し良好な変換効率を示すが相関の高い信号に対しては変換効率が低下する。つまり DST は DCT に対し符号化効率の点で相補的な関係を持つ。そこで本稿では、相関の高い信号から相関の低い信号まで良好な符号化効率を実現する新たな直交変換符号化方法として DCT と DST を適応的に用いる符号化方法を提案する。

2. HEVC の直交変換符号化技術

ITU-T と MPEG のジョイント作業チーム (JCT-VC : Joint Collaborative Team on Video Coding) ^[2]では新符号化規格として HEVC が検討されている。本検討では高度なイントラ予測技術や、動き補償予測技術などが提案されており、従来方式を超える予測性能を実現する。これらの高度な予測技術の採用により、変換対象である予測差分信号は既存の符号化方式に比べて信号相関が減少する傾向にある。このため DCT の変換効率もまた低下する傾向にある。

AVC/H.264 規格化後に行われた KTA (Key Technical Areas: H.264/AVC をベースとしたより高度化した符号化方式) の検討ではイントラ予測モードに応じた複数の変換を切り替える方式が採用され、改善効果が報告されている^[3]。このような状況から HEVC でも複数の直交変換を適応的に切り替えて符号化する直交変換符号化技術が提案されている。しかし、複数の変換を備えることはデコーダの回路規模を増大させるという問題があった。

そこでより少ない種類の変換による切り替え符号化方式もまた検討されている^{[4][5][6][7][8][9]}。これらの提案では符号化対象をイントラ予測差分信号に限定し、予測差分信号特有の信号の偏りを実験的、統計分析によって解析し新たに DST タイプ 3、7 を採用する。それぞれ 1 次元の直交変換を水平/垂直で組み合わせて適用することによって 4 種類の直交変換による符号化を実現している。

3. DCT/DST 適応直交変換符号化

i) DCT と DST の関係

DCT と DST はサンプリング位相の違いによりそれぞれ 8 種類の直交変換のタイプがある。イントラ予測/動き補償予測差分ともに適用できる直交変換符号化方法として DST タイプ 2 を用いた直交変換符号化方式を検討した。DST タイプ 2 の有用性を次に示す。

一般に信号波形は偶関数波形と奇関数波形に分解でき、信号解析では DFT (Discrete Fourier Transform) が広く用いられている。DCT や DST はそれぞれ偶関数と奇関数を変換核としてもち、それぞれ偶関数波形、奇関数波形に信号を分解する。これらの関係から DFT は DCT と DST を用いて以下のように表わされる^[10]。

$$X_l^{(N)} = \gamma_l W_N^{-1/2} [C_l^{(N/2)} - jS_l^{(N/2)}] \quad (\text{式 1})$$

$$\gamma_l = \begin{cases} 1 & (l=0) \\ 1/\sqrt{2} & (l \neq 0) \end{cases}, \quad W_N = \exp\left(-j \frac{2\pi}{N}\right)$$

$X^{(N)}$: N サンプル信号の DFT 値

$C_l^{(N/2)}$: N/2 サンプル信号の DCT 値

$S_l^{(N/2)}$: N/2 サンプル信号の DST 値

上式はサンプル信号を偶対称に拡張した信号の DFT が DCT であり、奇対称に拡張した信号の DFT が DST であること示している。ここで DCT と DST は同一のタイプである場合にのみ上記関係が成り立つ。同一成分 l の DCT 係数と DST 係数は直交関係にあり、入力信号の位相に応じて DCT あるいは DST へエネルギーが分配される。つまり DCT へのエネルギー集中が弱い信号に対し、DST へのエネルギー集中が高いことを意味しており、同一タイプの DCT と DST が相補関係にあることを示している。映像符号化方式では DCT タイプ 2 が広く採用されており、上式

†日本放送協会, NHK

に示すように同一タイプの DST タイプ 2 と相補関係をもつ。相補性の観点から本提案では DST タイプ 2 を新たな直交変換として採用した^[11]。式 2 に DCT タイプ 2 (以下 DCT) を、式 3 に DST タイプ 2(以下 DST)を示す。

$$F(u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \quad (\text{式 2})$$

$$u = 0, 1, \dots, N-1$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{N}} & u \neq 0 \\ \frac{1}{N} & u = 0 \end{cases}$$

$$G(u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) \sin\left(\frac{(2x+1)u\pi}{2N}\right) \quad (\text{式 3})$$

$$u = 1, \dots, N$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{N}} & u \neq N \\ \frac{1}{N} & u = N \end{cases}$$

ii) DCT を用いた DST の実装方法

新たな直交変換を追加することによってエンコーダ/デコーダに新たな処理回路を実装する必要がある。デコーダにおける回路規模の増加は消費電力、価格の増加を招くため極力抑える必要がある。DCT と DST は式 2、3 に示されるように変換核が余弦関数か正弦関数かの違いだけである。DST 回路の実現には DCT と同程度の回路規模を必要とし、新たな直交変換の採用により、直交変換部の処理回路が倍増する。本提案は新たに DST を直交変換処理として追加するものであるが、DCT と DST を同一タイプとすることによってわずかな回路の追加によって DST を実装することができる。以下に DCT と同一タイプの DST を採用することの実装上のメリットについて述べる。

前述の通り DCT と DST の違いは変換核が余弦関数か正弦関数であるかの違いである。式 2、3 の次数の範囲をそろえるために式 3 の u を $N-u$ に置き換え、正弦関数を余弦関数で書き換えると式 3 は下式であらわされる。

$$G(N-u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) \sin\left(\frac{(2x+1)(N-u)\pi}{2N}\right) \quad (\text{式 4})$$

$$= \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} (-1)^x f(x) \cos\left(\frac{(2x+1)u\pi}{2N}\right)$$

$$u = 0, 1, \dots, N-1$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{N}} & u \neq 0 \\ \frac{1}{N} & u = 0 \end{cases}$$

上式は入力信号の奇数番目の信号を符号反転し、DCT を行うことによって DST 係数が得られることを示している。ただし、DCT 回路の出力係数は次数が逆順($N-u$)になっていることに注意する。式 4 より DST は処理の大部分を DCT 処理を利用することによって実現可能であり、追加処理は入力数の半分の個数の反転処理のみであること

が分かる。DCT を用いた DST 処理のブロック図を次に示す。例として 4 点 DST を示すが N 点 DST であっても同様である。

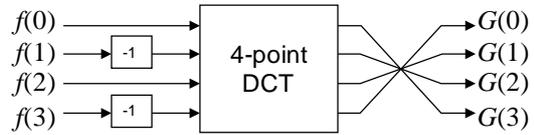


図 1. 4 点 DST のブロック図

DST 処理回路は DCT 処理回路の入力を DST 処理時に符号反転することによって実現でき、DCT と同程度の新たな直交変換回路を追加する必要がない。新たに追加される符号反転素子は、 N 点の入力に対し $N/2$ 個でよく、入力信号を直接、あるいは符号反転素子を通して DCT 処理を行うことで DCT と DST を導出することができる。

逆変換も正弦関数を余弦関数により展開すると下式のように逆 DCT 処理によって実現できる。

$$f(x) = \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) G(N-u) \sin\left(\frac{(2x+1)(N-u)\pi}{2N}\right) \quad (\text{式 5})$$

$$= (-1)^x \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) G(N-u) \cos\left(\frac{(2x+1)u\pi}{2N}\right)$$

$$x = 0, 1, \dots, N-1$$

逆変換の入力信号である直交変換係数の次数が逆順になる以外は DCT の逆変換処理と同一であり、出力の奇数番目を符号反転することで DST の逆変換が得られることを示している。順変換と同様に DCT の逆変換回路を用いて DST の逆変換を実現するブロック図を以下に示す。

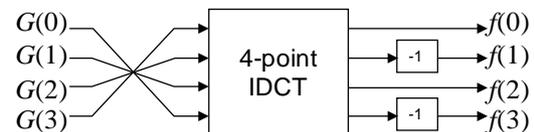


図 2. 4 点 IDST のブロック図

順変換同様、新たに追加するのは N 点の出力に対して $N/2$ 個の符号反転素子だけでよくデコーダの回路規模の増加もわずかで済むことが分かる。

iii) DCT/DST 適応直交変換符号化

本符号化方式の直交変換符号化方法について述べる。直交変換処理部では入力信号に応じて RD 最適化法^[12]を用いて DCT と DST を適応的に選択する。より符号化効率の高い直交変換を選択し、いずれか一方の直交変換係数をエントロピー符号化し、伝送する。本機能の追加によりデコーダでは復号処理に用いる直交変換種別を識別する必要が生じる。このため新たな制御フラグとして直交変換識別フラグを各変換ブロックに 1bit 追加し、直交変換係数情報と同時にエントロピー符号化し、伝送する。デコーダでは制御フラグより係数情報を DCT あるいは DST であるかを判別し、逆変換処理を行う。

4. 符号化実験

i) 実験条件

HM (HEVC Test Model) 0.9 に DCT/DST 適応切り替えを実装し、符号化実験を行った。DCT/DST 適応切り替えは輝度信号のみに適応し、色差信号は輝度信号以上に相関の高い信号であるので DCT のみを使用し DST は適応しない。本実験で使用した HM0.9 の主なツールセットを表 1 に示す^[13]。

表 1. ツールセット

ブロック構造	Quad-Tree 型階層分割 符号化ブロックサイズ: 8x8~64x64
動き予測	ブロックサイズ: 4x4~64x64 1/4-pel 精度動き予測 線形 MC 内挿フィルタ 12 タップ
イントラ予測	ブロックサイズ: 4x4~64x64 方向性空間予測(最大 34 方向)
変換符号化	整数精度 DCT ブロックサイズ: 4x4~32x32 内部ビット拡張(4bit)
エントロピー符号化	CABAC
ループフィルタ	デブロッキングフィルタ、適応ループ フィルタ

符号化素材は HEVC で提案技術評価のために採用されている 12 種類の高解像度 (1280x720~2560x1600、24~60 fps、プログレッシブ) 素材を使用した。

符号化性能は BD-Bitrate^[14] で評価した。BD-Bitrate は基準とする符号化方式と比較する符号化方式でそれぞれ 4 つの量子化パラメータにおけるビット量と PSNR 値から求められる RD 曲線を比較したものである。負値はビット削減効果を表し、符号化効率の改善を示す。符号化条件は HEVC で用いられている 3 つの符号化条件のうち放送サービスを想定したランダムアクセス条件を使用した^[15]。

符号化条件 ランダムアクセス条件

放送サービスを想定し、ストリームの途中での切り替えを想定した符号化条件。

- ・ イントラ間隔: 24@25fps、32@30fps、48@50fps、64@60fps (約 1 秒間隔)
- ・ 階層 B ピクチャ使用
- ・ 量子化値: QP=22,27,32,37

実験結果を表 2 に示す。また、表 3、4 に ParkScene と Vidyo1 のイントラ、インターブロックのそれぞれにおける各ブロックサイズの DST 選択率を示す。また、表 5 にそれぞれのシーケンスの符号化ブロックに占めるイントラ、インターブロックの面積率を示す。

表 2. 符号化結果

Sequence	Format	Y BD-rate [%]
Nebuta	2560x1600/60p	-0.8
SteamLocomotive	2560x1600/60p	-0.3
Traffic	2560x1600/30p	-0.4
PeopleOnStreet	2560x1600/30p	-0.3
BQTerrace	1920x1080/60p	-0.7
Cactus	1920x1080/50p	-0.5
BasketballDrive	1920x1080/50p	0.0
Kimono	1920x1080/24p	-0.6

ParkScene	1920x1080/24p	-1.2
Vidyo1	1280x720/60p	0.2
Vidyo3	1280x720/60p	0.0
Vidyo4	1280x720/60p	-0.2
Average		-0.5

表 3. イントラブロックの DST 選択率

	Vidyo1	ParkScene
Block size	DST rate	DST rate
32x32	1.8%	7.8%
16x16	5.4%	14.4%
8x8	14.2%	21.3%
4x4	11.4%	26.6%

表 4. インターブロックの DST 選択率

	Vidyo1	ParkScene
Block size	DST rate	DST rate
32x32	21.3%	37.2%
16x16	31.2%	38.6%
8x8	38.1%	38.1%
4x4	40.6%	41.4%

表 5. イントラ/インターブロックの面積率

Vidyo1		ParkScene	
Intra	Inter	Intra	Inter
82.3%	17.7%	71.9%	28.1%

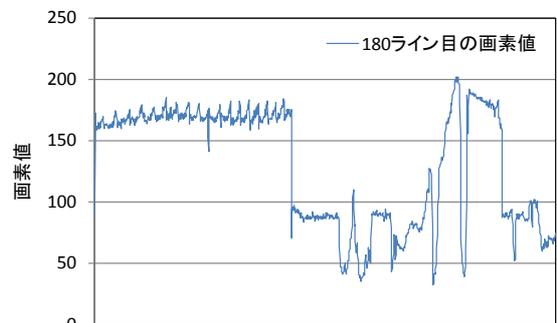


図 3. 信号変動 (Vidyo1 180 ライン目)

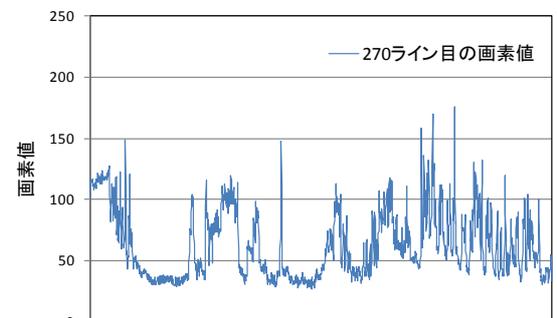


図 4. 信号変動 (ParkScene 270 ライン目)

ii) 考察

本符号化結果より ParkScene で最大 1.2% のビット削減効果を示した。一方で Vidyo1 では 0.2% のビット増加を示した。12 シーケンス中 3 シーケンスで効果を示さなかった

が平均して 0.5% のビット削減効果があることが確認できた(表 2)。本方式の効果を分析するために最も効果を示した ParkScene と効果を示さなかった Vidyo1 について考察する。

図 3、4 にそれぞれの画像の水平方向の信号変動を示す。図 3、4 の信号変動はシーケンスの上から 1/4 ライン目、それぞれ 180 (=720/4) ライン目、270 (=1080/4) ライン目の信号変動を示している。Vidyo1 は会議室風景であり、動きも少なく、図 3 に示すように信号変動が少ない空間周波数の低いシーケンスである。一方、ParkScene は木の生い茂る公園を自転車駆ける風景であり、テクスチャの細かい映像を含む。図 4 に示すように信号変動が大きく空間周波数の高いシーケンスである。

本符号化条件ではイントラ予測と動き補償予測を併用した符号化が行われている。ビット削減効果を示した ParkScene は表 3、4 に示すように Vidyo1 に比べすべてのサイズで DST の選択率が高い。また、表 5 に示すように Vidyo1 と ParkScene のイントラブロックとインターブロックの面積比率は Vidyo1 のイントラブロックが ParkScene に対し 10% 程高い。Vidyo1 のイントラブロックにおける DST の選択率は ParkScene のおおよそ半分であり、イントラブロックの比率が高いことから DST で符号化された領域が非常に少ないことがわかる。

これらの結果は ParkScene では DCT の符号化効率の低下が生じやすく、DST が選択されることによって効果的に DCT による符号化効率の低下を補完したことを示している。一方で Vidyo1 では DST の選択率が少なく、従来の DCT のみによる符号化とほぼ同等の選択が行われたと考えられる。しかし、本方式は DCT と DST の直交変換種別を識別するための制御フラグを新たな符号化情報として伝送する必要があり、制御フラグによるビット増加が多く、ビット削減効果が得られなかったと考えられる。

6. まとめ

新たな直交変換符号化方式として同一タイプの DCT と DST を適応的に使い、直交変換符号化を行う方式を提案した。本方式は同一タイプの DCT と DST を用いることにより、あらゆる信号に対し 2 種の直交変換が相補的に作用し、安定した符号化性能を発揮する。また、実装上の追加処理の増加を抑えることができ、実装上のメリットも大きい。

符号化実験によりテクスチャの細かい映像信号に対する符号化効率を改善が確認できた。一方で二種の直交変換を判別するための制御フラグコストにより、符号化効率が低下するシーケンスも確認された。

7. 参考文献

- [1] 画像符号化技術-DCT とその国際標準-, オーム社, ISBN4-274-03401-1
- [2] “Terms of Reference of the Joint Collaborative Team on Video Coding Standard Development,” ISO/IEC JTC1/SC29/WG11 N11112, Kyoto, January, 2010.
- [3] Y. Ye and M. Karczewicz, “Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning,” ICIP08, San Diego, U.S.A., Oct. 2008.
- [4] M.Karczewicz et al., “Video coding technology proposal by Qualcomm Inc.,” JCTVC-A121, Dresden, DE, April 2010.
- [5] C.Yeo et al., “CE7: Mode-Dependent Transforms for Block-based Intra Coding from Institute for Infocomm Research,” JCTVC-D046, Daegu, South Korea, January 2011.
- [6] A. Tanizawa et al., “CE7: Experimental results for one-dimensional directional unified transform” , JCTVC-D107, Daegu, South Korea, January 2011.
- [7] X. Zhao et al., “CE7: Mode-dependent transform, residual reordering and coefficient scanning for intra prediction residue,” JCTVC-D284, Daegu, South Korea, January 2011.
- [8] H. Yang et al., “CE7: Simplified MDDT scheme using symmetry-based scanning orders” , JCTVC-D304, Daegu, South Korea, January 2011.
- [9] R. Joshi et al., “Mode dependent intra residual coding” , JCTVC-D392, Daegu, South Korea, January 2011.
- [10] 市ヶ谷、西田, “DCT/DST 相互予測符号化” , IEICE IE2007-118, pp.31-37, 2007.
- [11] A. Ichigaya and S. Sakaida, “Performance report of adaptive DCT/DST selection.” , JCTVC-D182, Daegu, South Korea, January 2011.
- [12] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, “Rate-constrained coder control and comparison of video coding standards,” IEEE Trans. Circuits and System. Video technology, vol. 13, pp. 688-703, July 2003. JCTVC-C405, Oct. 2010.
- [13] “Summary of HEVC working draft 1 and HEVC test model (HM)” ,
- [14] G. Bjontegaard, “Calculation of average PSNR differences between RD-Curves,” VCEG-M33, April 2001.
- [15] “Common test conditions and software reference configurations” , JCTVC-C500, 3rd. JCT-VC Meeting, Guangzhou, CN, October 2010.
- [16] J. Han, A. Saxena and K. Rose, “Towards jointly optimal spatial prediction and adaptive transform in video/image coding,” Proc. of ICCASP 2010, pp. 726-729, 2010.