

符号情報を利用した DV から MPEG-2 への 高速変換アルゴリズム

加藤 晴久[†] 佐野 卓[‡] 中島 康之[†]

[†]KDDI 研究所 〒356-8502 埼玉県上福岡市大原 2-1-15

[‡]東京理科大学工学研究科 〒162-8601 東京都新宿区神楽坂 1-3

E-mail: {hkato, sano, nakajima}@spg.kddlabs.co.jp

あらまし DV を完全復号せずに圧縮された符号化情報を直接利用することで、高速に MPEG-2 への変換を達成する方式を検討する。DV と MPEG はともに DCT 変換を用いるが、色空間フォーマットの相違と DCT 変換単位の相違により直接再利用することはできない。DV から MPEG-2 I-picture を構成するには DV 特有の 2x4x8DCT 係数をフレーム DCT 係数またはフィールド DCT 係数へと変換する方法と 4:1:1 フォーマットを 4:2:0 フォーマットへと変換する方法が求められる。本研究では符号領域上で上記の相違を解決する方式を提案し、ベースバンド変換方式との画質および処理負荷について比較検討する。

キーワード DV, MPEG, トランスコード, DCT

The Fast Conversion Algorithm from DV to MPEG-2 using DV Coded Data

Haruhisa KATO[†] Takashi SANO[‡] and Yasuyuki NAKAJIMA[†]

[†]KDDI R&D Laboratories 2-1-15 Ohara Kamifukuoka-shi, SAITAMA, 356-8502 Japan

[‡]Faculty of Engineering, Tokyo University of Science 1-3 Kagurazaka Shinjuku-ku, Tokyo, 162-8601 Japan

E-mail: {hkato, sano, nakajima}@spg.kddlabs.co.jp

Abstract. This paper describes a fast Digital Video (DV) to MPEG-2 conversion algorithm. At first, 2x4x8 DCT to 8x8 DCT conversion and sampling format conversion were employed for the DV to MPEG-2 I-picture conversion. The experimental results show that the proposed method is more than 10 times faster than a baseband domain conversion while about the same PSNR is obtained.

Keyword DV, MPEG, transcoder, DCT

1. はじめに

DV(Digital Video)カメラの普及や i.LINK(IEEE1394)端子を標準搭載する PC が増えてきたことにより、PC 上でのノンリニアビデオ編集が注目され始めている。DV は手軽に高画質記録、編集ができるものの符号化レートが 25Mbps と高いため、長時間記録や DVD-R などでも記録する場合、MPEG-2 に変換することが多い。DV と MPEG-2 はいずれも圧縮方式に DCT 変換と可変長符号化を利用しており、共通した技術基盤の上に成

り立っている。よって、DCT 係数を再利用できれば、効率的な変換を実現できると考えられる。これまで筆者らは DV の符号情報を利用した高速動き探索方式 [1]を提案し、MPEG-2 の P-,B-picture の符号化を大幅に高速化する方式を実現した。本稿では DCT 係数のみを用いる I-picture について DV と MPEG-2 の相違を検討し、DV の符号情報を利用して DV から MPEG-2 への高速に変換する方式を提案する。

2. DV と MPEG-2 の特徴

DV と MPEG-2 I-picture はいずれも DCT 変換と VLC を利用するが、次に示す相違点がある [2]。

2.1. サンプリングフォーマット

DV の入力信号としては 4:1:1 フォーマット (NTSC), または 4:2:0 フォーマット (PAL) が使用される。4:1:1 フォーマットは輝度 (Y) と色差 (Cb,Cr) のコンポーネント信号のうち、色差成分は輝度成分のサンプリングレートを水平方向へ 1/4 に削減されている (図 1(a))。一方で MPEG-2 では一般に 4:2:0 フォーマットが利用され、輝度に対して色差信号が水平方向、垂直方向にそれぞれ 1/2 に削減されている (図 1(b))。輝度成分に相違はないが、色差信号の構成が異なるため、NTSC の DV から MPEG-2 への変換には色差信号のサンプリングフォーマット変換処理が必要となる。

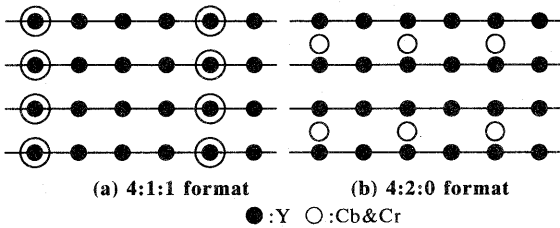


図 1 サンプリングフォーマット

2.2. マクロブロック (MB) 構造

前節に示したようにサンプリングフォーマットが異なるため、色差信号を DCT 変換単位の 8x8 画素毎に 1 ブロックとして扱おうと、対応する輝度信号の位置が DV と MPEG-2 で異なる。DV では色差 8x8 画素に対して輝度 8x32 画素が対応し、マクロブロックは 1x4 ブロックとなる (図 2(a))。一方、MPEG-2 では色差 8x8 画素に対して輝度 16x16 画素が対応するため、マクロブロックは 2x2 ブロックとなる (図 2(b))。よって、DV から MPEG-2 への変換にはマクロブロックの再構成が必要となる。

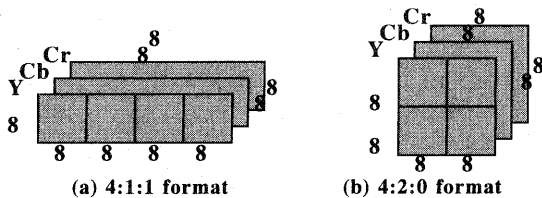


図 2 マクロブロック構造

2.3. フィールド DCT の構成単位

DV と MPEG-2 はいずれも任意にフレーム構造とフィールド構造を選択できるが、フィールドの最小構成

単位が異なる。DV では 8x8 画素のブロック単位でフィールド構造を選択でき、フィールド毎に 4x8DCT 変換を施す (図 3(a))。2 組の 4x8DCT 係数はそれぞれ加算、減算され 2x4x8DCT 係数として扱われる (図 4)。MPEG-2 では 16x16 画素のマクロブロック単位でフィールド構造を選択する。フレーム/フィールドともにブロックサイズが 8x8 画素であるため、DCT 変換の基底は 8 次に固定される (図 3(b))。よって、DV のフレーム DCT 係数は MPEG-2 に直接再利用できるが、DV のフィールド DCT 係数は MPEG-2 で利用できる形式に変換する必要がある。

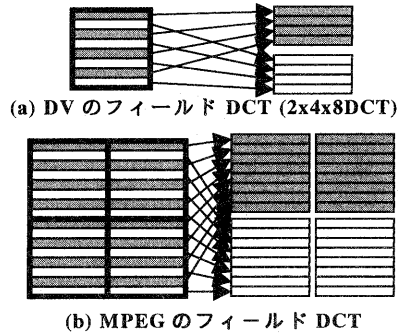


図 3 フィールド DCT 単位

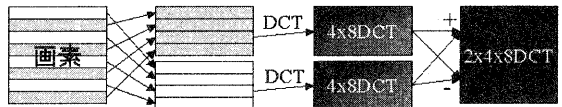


図 4 2x4x8 DCT モード

3. 提案方式

第 2 章で述べたようにサンプリングフォーマットとフィールド構造が異なる問題を解決し、高速変換を実現するため DCT 領域上での変換式を提案する。

3.1. DCT 基底変換

3.1.1. 2x4x8DCT からフレーム DCT への変換

従来は 2x4x8DCT 係数を一度 2x4x8DCT で復号し画素に戻してから 8x8DCT で再度変換する必要があった。ベースバンドで 2x4x8DCT 係数 U_{248} から 8x8DCT 係数 V_{88} を得るための変換手順をまとめると次式のように表すことができる。

$$\begin{aligned} V_{88} &= C_8 N M^T L_8 U_{248} C_8^T \\ &= (C_8 N M^T L_8) U_{248} \\ &= A U_{248} \end{aligned} \quad (1)$$

$$\begin{aligned}
 V &= C_8 H Z^t U C_8 W Z^t \\
 &= (C_8 H Z^t) U F \\
 &= G U F
 \end{aligned} \tag{4}$$

このとき、 C_8, H, Z は一定なので予め計算した結果を G とすると、4:1:1 から 4:2:0 への変換は 8x16 行列の 2 回の積で実現できる。

4. 提案方式の処理負荷

4.1. DCT 基底変換の演算量

4.1.1. 2x4x8DCT からフレーム DCT への変換に必要な演算量

提案方式の変換には 2x4x8DCT から 8x8DCT への変換は 8x8 行列の積が必要なので、1 ブロックあたりの計算回数は 512 回の乗算と 448 回の加算である。しかし、式(5)に示すように行列 A には 43 個の 0 要素が存在するので、乗算は 168 回、加算は 104 回に減る(但し、 c_n^m は $\cos(\pi m/n)$ を表す)。一方、高速 DCT 演算法[3] を利用したベースバンド領域の変換の演算量は乗算が 480 回 (2x4x8IDCT に 224 回、8x8DCT に 256 回)、加算は 752 回 (2x4x8IDCT に 240 回、8x8DCT に 512 回) である。提案方式はベースバンド変換と比較して乗算回数は 35%、加算回数は 14% に抑えられる。表 1 に提案方式と従来方式による演算量を示す。

$$A = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{16}^1 & 0 & 0 & c_8^7 c_{16}^7 & 0 & -c_8^3 c_{16}^7 & 0 \\ 0 & 0 & c_8^1 & 0 & 0 & c_4^1/2 & 0 & -(c_8^3)^2 \\ 0 & 0 & 0 & c_8^3 & c_8^3 c_{16}^5 & 0 & c_8^1 c_{16}^5 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_8^3 & 0 & c_8^1 \\ 0 & 0 & 0 & -c_{16}^5 & c_8^3 c_{16}^3 & 0 & c_8^1 c_{16}^3 & 0 \\ 0 & 0 & -c_8^3 & 0 & 0 & (c_8^1)^2 & 0 & -c_4^1/2 \\ 0 & -c_{16}^7 & 0 & 0 & c_8^1 c_{16}^1 & 0 & -c_8^3 c_{16}^1 & 0 \end{pmatrix} \tag{5}$$

表 1 2x4x8DCT からフレーム DCT への変換の演算量

	乗算回数	加算回数
提案方式	168 (35.0%)	104 (13.8%)
ベースバンド変換方式	480 (100.0%)	752 (100.0%)

回 / block

4.1.2. 4 組の 2x4x8DCT からフィールド DCT への変換に必要な演算量

提案したフィールド DCT の変換式については 16x16 行列の積が必要であるが、実際には 256 要素の内、行

列 B は 96 要素が 0 であり、さらに式(6)に示すような対称性が備わっている。16x16 行列 B を 8x4 行列の部分行列 D の集合として捉えると、図 5 のように乗数の絶対値が一致する係数の和を優先することで乗算回数を抑えることができる(図 5 の太線、破線と細線はそれぞれ加算、減算と乗算を表している。但し、乗数は省略)。その結果、MB 毎の乗算と加算の回数はそれぞれ 640 回、896 回となる。高速 DCT 演算法を利用したベースバンド変換は乗算と加算は 1920 回、3008 回となり、提案方式はベースバンド変換の演算量を 1/3 以下に抑えられる。表 2 に提案方式とベースバンド変換のフィールド DCT 変換にかかる演算量を示す。

また、4 組の 2x4x8DCT を前節の方法ですべてフレーム DCT 変換する場合には、乗算回数は 672 回、加算回数は 432 回となる。フィールド DCT 変換は加算回数がフレーム DCT 変換の 2 倍程度必要になるが、乗算回数の 5% 削減とフィールド DCT による圧縮効率の改善が期待できる。

$$B = \begin{pmatrix} D & D & D' & D' \\ D & -D & D' & -D' \end{pmatrix} \tag{6}$$

$$D(D') = \frac{\sqrt{2}}{4} \begin{pmatrix} 2 & 0 & 0 & 0 \\ \pm 2c_8^3 c_{16}^1 & c_{16}^3 & \mp 2c_8^3 c_{16}^7 & c_{16}^3 - \sqrt{2}c_{16}^5 \\ 0 & 2 & 0 & 0 \\ \mp 2c_8^3 c_{16}^3 & c_{16}^7 + \sqrt{2}c_{16}^1 & \pm 2c_8^1 c_{16}^5 & -c_{16}^7 \\ 0 & 0 & 2 & 0 \\ \pm 2c_8^3 c_{16}^5 & -c_{16}^1 + \sqrt{2}c_{16}^7 & \pm 2c_8^1 c_{16}^3 & c_{16}^1 \\ 0 & 0 & 0 & 2 \\ \mp 2c_8^1 c_{16}^7 & c_{16}^5 & \mp 2c_8^3 c_{16}^1 & c_{16}^5 + \sqrt{2}c_{16}^3 \end{pmatrix} \tag{7}$$

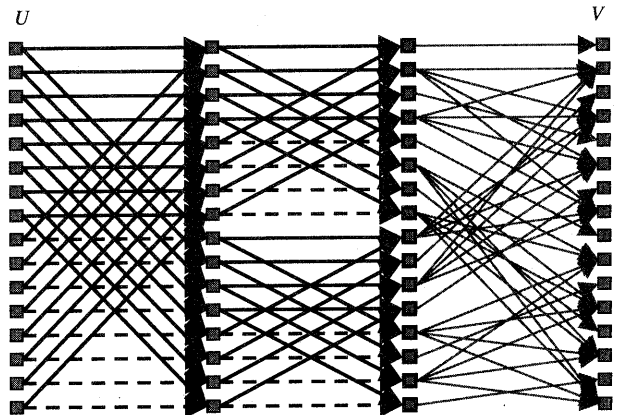


図 5 4 組の 2x4x8DCT からフィールド DCT への変換の高速演算順序

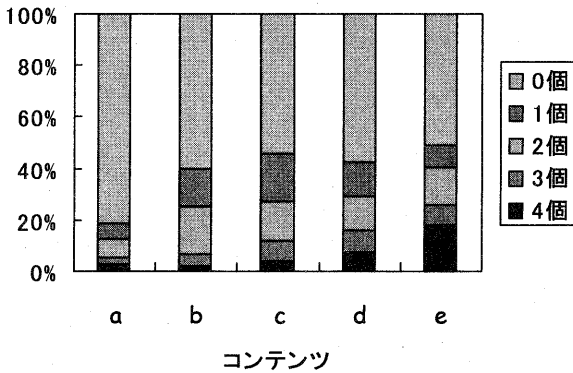


図6 MBに含まれる8x8DCTと2x4x8DCTの割合

5.2. 提案方式のPSNRによる評価

図6で用いたコンテンツdをDVにエンコードし、このDVから提案方式と従来方式でそれぞれMPEG-2 I-pictureのみに変換した。非圧縮映像に対するPSNRを求めた結果を表5に示す。但し、今回の実験ではすべてのブロックをフレームDCTに変換した。

提案方式の変換過程には情報の損失がないので、理論的にはベースバンド変換とPSNRが一致すると考えられる。しかし輝度成分に関しては提案方式のほうが若干高いPSNRを持つ。その理由としては、ベースバンド変換が個々の処理毎に小数点以下の数字を切り捨てや四捨五入で何度も情報が欠落するためだと考えられる。提案方式では個々の処理を一括して行うため、小数点以下の端数を失う場面はベースバンド変換に比べ少ない。色差成分に関しては、両方式ともほぼ同じPSNRを持つ。輝度成分に見られたPSNRの向上がない理由としては4:1:1→4:2:0サンプリングフォーマット変換に用いた水平2倍行列Hが本来のサンプリングポイントを保持する1次補間ではなく、変換行列Gに対称性を持たせるため右半分の画素を本来の位置からずらしているためだと考えられる。

表5 非圧縮映像に対するPSNR

		Y	Cb	Cr
提案方式	6Mbps	31.59	36.23	38.11
	12Mbps	35.12	36.96	38.86
	25Mbps	36.82	37.21	39.08
ベースバンド変換	6Mbps	31.29	36.22	38.23
	12Mbps	34.93	37.08	39.06
	25Mbps	36.79	37.33	39.27

dB

5.3. 提案方式の処理時間による評価

提案方式とベースバンド変換方式の変換処理時間を計測した。1フレームあたりの処理時間を表6に示す。提案方式はDVの8x8DCT係数をそのまま再利用

するので、8x8DCT係数が多いほど処理時間は速くなる。2x4x8DCTをフレームDCTに変換する処理速度は5.6 msec/frameで、4:1:1→4:2:0サンプリングフォーマット変換の処理速度はCb, Cr合わせて4.3 msec/frameとなった。ベースバンド変換は8x8DCTも2x4x8DCTもすべて復号し8x8DCTで再符号化するため、変換処理時間は大幅にかかる。8x8DCTを再利用するだけでも完全復号する方式より処理時間が大幅に短縮でき、提案した2x4x8からフレームDCTへの変換およびサンプリング変換の処理コストは十分に小さい。変換処理だけに注目すると提案方式はベースバンド変換の10%以下の処理時間を達成した。全体の処理時間による比較では、提案方式はベースバンド変換の25%の処理時間を短縮した。

表6 DVからMPEG-2への変換処理時間

	変換	全体
提案方式	9.9 (8.9%)	311.9 (73.9%)
ベースバンド変換	111.6 (100.0%)	422.0 (100.0%)

msec / frame

6. 結論

DVとMPEG-2との相違点を示し、DVからMPEG-2 I-pictureへの変換方法を提案した。DV独自の2x4x8DCT係数をMPEG-2のフレームDCTまたはフィールドDCTへ高速に変換するための行列式を求め、ベースバンド変換の演算量と比較を行った。フレームDCTへの変換は乗算回数を1/3、加算回数を1/7以下に削減できることを示した。フィールドDCTへの変換は乗算加算ともに1/3以下に削減した。サンプリングフォーマット変換についても変換式を求め、4:1:1から4:2:0に関しては乗算を半分に、加算を1/3以下に抑えた。4:1:1から4:2:0への変換は10%の乗算削減と加算の半減を達成した。また、いずれの変換式も計算過程に情報の欠損は存在しないため、画質の劣化はほぼ存在しないことをシミュレーション実験からも確認した。今後は本方式を用いてIPB構造の変換実験で性能評価する予定である。

文献

- [1] H.Kato, et.al.: "A fast motion estimation algorithm for DV to MPEG-2 conversion", ICCE, WAM10.6, pp.140-141, (Jun. 2002).
- [2] 加藤晴久, 佐野卓, 中島康之: "DV符号化情報を利用したMPEG-2 I-pictureへ的高速変換に関する研究", PCSJ, P2-10, pp.27-28, (Nov. 2002).
- [3] W.H.Chen, et.al.: "A fast computational algorithm for the discrete cosine transform", IEEE Trans. Commun, vol. 25, No. 9, pp.1004-1009, (1977).