

隣接ブロックの相関関係を利用した H.264 動画像可逆符号化 H.264-based lossless video coding using correlation of adjacent block

佐野 利行[†]
Toshiyuki Sano

作田 豊[†]
Yutaka Sakuta

平澤 茂一[†]
Shigeichi Hirasawa

1. 研究目的

動画像の可逆符号化は、医療用画像や衛星映像など高品質な動画像を扱う分野において、極めて重要な技術である。動画像の可逆符号化の手法には H.264 Fidelity Range Extension (以下、H.264FRExt) 方式があるが、この手法は残差信号に残る相関を除去できないため、圧縮の性能が悪くなるという問題がある。

これを解決する手法として、高村らにより残差信号に対して 2 次元の Differential Pulse Code Modulation (以下、DPCM) による予測を用いた変換を行う方法 [2] が提案されている。高村らの手法では、残差信号の空間的相関を除去する方法が用いられているが、周辺ブロックの相関を適切に利用できていない。

そこで本研究では、DPCM を行うときの予測係数の重みを隣接ブロックの相関関係を利用することで、よりよい予測係数の値を求め、空間的相関を効率的に除去する方法を提案する。また、提案手法の圧縮率が良くなることを評価用標準動画像 [4] を用いた実験により示す。

2. H.264 FRExt 方式 [1]

[H.264 FRExt 方式のアルゴリズム]

(A1) 入力画像に対して、画面内符号化か画面間符号化を行う。前者は (A2) へ、後者は (A3) へいく。

(A2) 画面内符号化では、画面内予測を行い、元画像との差分を計算し、残差信号を得て (A4) へいく。

(A3) 画面間符号化では、動き補償予測を行い、元画像との差分を計算し、残差信号を得て (A4) へいく。

(A4) 残差信号をエントロピー符号化する。

3. 従来手法

本節では残差信号に対して 2 次元の DPCM を用いることで空間的相関を除去した高村ら [2] の手法を示す。具体的には、2 章のアルゴリズムにおける (A4) を以下の操作に置き換える。

[従来手法のアルゴリズム [2]]

(A4.1) マクロブロックを 4×4 画素のブロックに分割する。

(A4.2) 得られた 4×4 画素ブロックの水平成分を変換する。

(A4.3) 水平変換の結果得られた 4×4 画素ブロックに対して垂直成分の変換を行う。

(A4.4) 残差信号に対してエントロピー符号化を行う。

上記 (A4.2)(A4.3) では次式のような変換を施す。

$$X_i = \begin{cases} x_0 & , i = 0; \\ x_i - R(\alpha x_{i-1}) & , i = 1, 2, 3. \end{cases} \quad (1)$$

ここで α は予測係数で、水平変換は $\alpha = \alpha_H$ 、垂直変換は $\alpha = \alpha_V$ とする。また、 x_i は予測残差、 X_i は変換係数、下付添え字はブロック内の位置を表し、 $R(s)$ は s の小数点第一位を四捨五入する関数である。

このとき、予測係数は隣接したブロック同士では、類似の性質を持つことが知られている。この性質を用いて、以下のアルゴリズムに従って近隣のブロックから予測係数を適応的に求める。ここで、 ρ は各位置のブロックに対する相関係数、上付き添え字 N, W, NW はそれぞれ上、左、左上隣接ブロックを、下付き添え字 H, V はそれぞれ水平、垂直方向を表している。なお、 α_0 はデフォルト値である。

[α の決定アルゴリズム [2]]

(B1) 次式により水平方向と、垂直方向の隣接ブロックの関係 $hdif, vdif$ を定義する。

$$hdif = |\rho_H^{NW} - \rho_H^N| + |\rho_V^{NW} - \rho_V^N| \quad (2)$$

$$vdif = |\rho_H^{NW} - \rho_H^W| + |\rho_V^{NW} - \rho_V^W| \quad (3)$$

(B2) 式 (4) を満たすとき、水平に強い相関と判断し α_H, α_V を式 (5), (6) のように求める。

$$hdif < 0.6vdif \quad (4)$$

$$\alpha_H = 0.4\rho_H^W + 0.3\rho_H^N + 0.1\rho_H^{NW} + 0.2\alpha_0 \quad (5)$$

$$\alpha_V = 0.4\rho_V^W + 0.3\rho_V^N + 0.1\rho_V^{NW} + 0.2\alpha_0 \quad (6)$$

式 (4) を満たさないとき (B3) へいく。

(B3) 式 (7) を満たすとき、垂直に強い相関と判断し α_H, α_V を式 (8), (9) のように求める。

$$vdif < 0.6hdif \quad (7)$$

$$\alpha_H = 0.3\rho_H^W + 0.4\rho_H^N + 0.1\rho_H^{NW} + 0.2\alpha_0 \quad (8)$$

$$\alpha_V = 0.3\rho_V^W + 0.4\rho_V^N + 0.1\rho_V^{NW} + 0.2\alpha_0 \quad (9)$$

式 (7) を満たさないとき、周辺ブロックの相関はないと判断し、次式によって α_H, α_V を求める。

$$\alpha_H = \alpha_V = \alpha_0 \quad (10)$$

4. 提案手法

4.1 従来手法の問題点

(1) 代表的な動画像データ [4] に対する $hdif/vdif$ の出現割合を図 1 に示す。水平もしくは垂直に早く動く動画像は、図 1 に示すように、 $hdif/vdif$ の値の小さい部分もしくは、大きい部分の頻度が高くなることから、動きの方向に相関関係が偏りを持つという傾向があることが分かる。しかし、従来手法では α の値を求める際、 $hdif, vdif$ の値に関わらず一定の重み付けをされた α の決定式 (5), (6), (8), (9) を用いているため、 $hdif/vdif$ が小さな値もしくは、大きな値に対しては、周辺のブロックの状況を適切に反映させているとは言えず、適切な予測係数が求められているとは限らない。

[†]早稲田大学理工学部経営システム工学科

(2) 従来手法では、両方向のブロックに相関があり、その相関が同程度の相関であった場合、 $hdif/vdif$ の値が1に近くなるため、式(4)、(7)よりブロック間に相関がないと判断されてしまう。

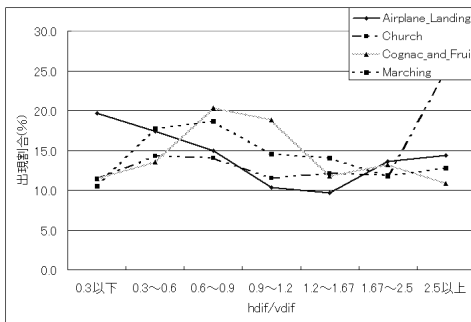


図1: 動画の種類による $hdif/vdif$ の値の出現割合

“Airplane Landing” は水平方向に早く動く動画
 “Church” は垂直方向に動く動画
 “Cognac and Fruits” は動きの少ない動画
 “Marching in” は斜め方向に動く動画

そこで、 $hdif$ 、 $vdif$ を相関の有無の判断に用いるのではなく、これらを決定式に直接利用することで、重みの値を可変とし、周辺の状況に対応した予測係数を求める手法を提案する。その結果、予測の精度が向上し、圧縮率が改善されることが期待される。

4.2 提案手法

提案手法では、 α を決定する際、 $hdif$ 、 $vdif$ の比によって重み付けをする。その結果、図1の $hdif/vdif$ の値の小さい部分に対してはより水平方向に大きな重みを与えることができ、 $hdif/vdif$ の値の大きい部分に対しては、垂直方向に大きな重みを与えることが可能になるため、周辺ブロックの性質に応じた動的な重みを決定することができる。特にそれらの出現頻度の高い動画像、すなわち、水平もしくは垂直に早く動く動画像に対して効果が期待される。具体的な流れを以下に示す。
 [提案手法の α の決定アルゴリズム]

(C.1) 式(2)、(3)により $hdif$ 、 $vdif$ を求める。

(C.2) 次式において α_H 、 α_V を決定する。

$$\alpha_H = 0.7 \frac{vdif}{hdif + vdif} \rho_H^W + 0.7 \frac{hdif}{hdif + vdif} \rho_H^N + 0.1 \rho_H^{NW} + 0.2 \alpha_0 \quad (11)$$

$$\alpha_V = 0.7 \frac{vdif}{hdif + vdif} \rho_V^W + 0.7 \frac{hdif}{hdif + vdif} \rho_V^N + 0.1 \rho_V^{NW} + 0.2 \alpha_0 \quad (12)$$

5. 実験結果と考察

5.1 符号化条件

H.264 リファレンスソフトウェア JM10.1 [3] をもとに提案手法を実装して実験を行った。実験には、システム評価用標準動画像 [4] を用い、表1の符号化条件を用いた。

5.2 実験結果

H.264FRExt方式(FRExt)、従来手法(従来)、提案手法(提案)を比較した実験結果を表2に示す。

表1: 符号化条件

画像サイズ	SIF(352×240ピクセル)
フレーム構成	IBBPBBPBB...
フレームレート	30fps
符号化枚数	100枚
エントロピー符号化	CABAC
R-D最適化	使用
α_0	Iフレーム: 0.7, P, Bフレーム: 0.4

表2: 実験結果 (bit/pixel)

動画名	FRExt	従来	提案
Cognac & Fruits	3.648	3.567	3.558
Airplane Landing	3.590	3.281	3.241
Soccer Action	5.588	5.329	5.229
Church	5.351	5.150	5.101
Marching in	8.111	8.064	7.904
Crowded Crosswalk	5.205	4.938	4.928

5.3 結果の考察

- 表2より、提案手法は従来手法に対して0.2%から2.0%の改善が見られた。これは、隣接ブロックの性質を決定式に組み込んだことで、より動的に予測係数を求めることが可能になったためと考えられる。
- 提案手法で効果の高かった動画像は、“Airplane Landing”や“Soccer Action”のように横に早く動く動画像と、“Church”のように縦に動く動画像であった。これらの動画像はブロック間の相関に偏りがあるため、従来手法よりも、相関の強い方向に大きな重みを与えることができたためであると考えられる。
- 効果の低かったものは“Cognac & Fruits”のようにあまり動きのない動画像であった。この原因は、水平・垂直のどちらかに偏った相関が少ないため、従来手法と比べて重みの値に大きな変化が見られなかったためであると考えられる。
- 動画像“Crowded Crosswalk”は局所的に見ると、従来手法に対して効率が悪くなるフレームが見られた。これは、この動画像が複雑な動きをするため、ブロック間の相関がほとんどない割合が多かったためであると考えられる。
- 提案手法の計算量は、1ブロックあたり、加算が1回、乗算が3回増加する。従って、本実験では1フレームあたり加算が5280回、乗算が15840回増加する。これは、符号化全体処理の計算量から考えると問題ない程度と考えられる。

6. まとめと今後の課題

本研究では従来の動画像可逆符号化において予測係数を求めるときに、周辺ブロックの相関を直接利用することで予測係数を求める手法を提案し、その結果、圧縮率を向上することを実験により示した。

今後はより効果的に空間的相関を除去できるような変換の方法を開発する必要がある。

参考文献

- [1] 大久保 榮, 角野 眞也, 菊池 義浩, 鈴木 輝彦, H264/AVC 教科書, インプレス, 2004年。
- [2] S.Takamura and Y.Yashima, “H.264-based Lossless Video Coding Using Adaptive Transforms,” *Proc. 2005 IEEE International Conference on Acoustics, speech, and Processing* Vol.2, pp.301-304 March 2005.
- [3] <http://iphome.hhi.de/suehring/tml/>; “JM reference software,” version 10.1, Nov. 2005.
- [4] SIF 画像システム評価用標準画像 NHK エンジニアリング, 2003.