

早稲田大学大学院 基幹理工学研究科

博士學位論文

動画像符号化の性能改善と
実用化に関する研究

Performance Improvement and Practical Realization
for Video Coding

2018年2月

申請者

中川 章

Akira NAKAGAWA

目次

第 1 章 序論	5
1.1 はじめに	5
1.2 本研究の背景：動画像符号化と関連分野の技術の歴史と進化	6
1.3 動画像符号化技術の原理	10
1.3.1 MC+DCT 動画像符号化技の概要	10
1.3.2 画像フォーマット	12
1.4 各動画像符号化国際標準の技術の特徴	13
1.4.1 H.261	14
1.4.2 MPEG-1	14
1.4.3 MPEG-2 / H.262	15
1.4.4 MPEG-4	15
1.4.5 H.264/AVC	15
1.4.6 H.265/HEVC	16
1.5 本研究の目的と位置付け	17
第 2 章 H.263 / MPEG-4 における動的解像度変換技術	20
2.1 はじめに	20
2.2 動的解像度変換技術 (Dynamic Resolution Conversion)	21
2.2.1 技術概要	21
2.2.2 提案手法の符号化手順	22
2.2.3 選択可能な動きベクトルの制限	24
2.2.4 ダウンサンプリング	25
2.2.5 アップサンプリング	26
2.2.6 ブロック境界フィルタ	27
2.2.7 符号化に必要な演算量の評価	27
2.3 解像度変換の制御アルゴリズム	29
2.3.1 解像度決定のアルゴリズム	29
2.3.2 低解像度から高解像度遷移時の DCT 係数の制限	30
2.4 符号化性能の評価	31
2.4.1 シミュレーション	31
2.4.2 評価結果	32
2.5 伝送誤り環境下での性能評価	34
2.5.1 伝送誤りが発生する通信路での映像伝送	34
2.5.2 シミュレーション	35
2.6 本章のまとめ	40

第 3 章 H.264 に向けた動きベクトル符号化の性能改善	41
3.1 はじめに	41
3.2 動きベクトルのコンテキスト切り換え符号化技術	41
3.2.1 背景と概要	41
3.2.2 従来方式での動きベクトル符号化方法	41
3.2.3 動きベクトルのコンテキスト切り換え符号化技術の提案 ...	42
3.2.4 実験結果	45
3.3 Interlace 符号化時の色差ベクトル補正技術	47
3.3.1 背景と概要	47
3.3.2 フィールド符号化時の問題点と提案方式	47
3.3.3 実験結果	49
3.4 本章のまとめ	51
第 4 章 直交変換係数の正負符号の予測符号化技術	52
4.1 背景と概要	52
4.2 直交変換係数の正負符号予測符号化方法	53
4.2.1 直交変換係数の正負符号化予測の原理	53
4.2.2 正負符号の高速予測アルゴリズム	56
4.2.3 エントロピー符号化方法	60
4.3 提案技術の評価	63
4.3.1 正負符号の予測性能の評価	64
4.3.2 CABAC のコンテキスト設計の評価	65
4.3.3 符号化効率改善の評価	70
4.4 本章のまとめ	71
第 5 章 H.264/AVC 映像伝送装置の実用化技術の研究	72
5.1 背景と概要	72
5.2 高画質・低遅延を実現する動画像符号化技術	74
5.2.1 H.264/AVC 参照ソフト JM で符号化した画質の課題	74
5.2.2 人間の視覚特性に基づいた画質制御技術	75
5.2.3 低遅延符号化	76
5.3 公衆インターネット向けの高性能パケット誤り訂正技術	78
5.3.1 IP ネットワークにおけるパケット損失	78
5.3.2 従来のパケット損失の訂正技術	79
5.3.3 インターネット向けの FEC + ARQ ハイブリッド誤り訂正 ..	80
5.3.4 装置への実装と公衆インターネットでの適用	82
5.4 開発した映像伝送装置 IP-9500 の適用事例	83

5.4.1 ANN 系列のインターネット経由映像素材伝送システム	83
5.4.2 フジテレビのお天気カメラ	84
5.4.3 日本テレビのNFL スーパーボウル XLI の日米間 HD ライブ 中継	85
5.5 本章のまとめ	86
第 6 章 4:2:2/4:2:0 相互変換時の色にじみ抑止技術	87
6.1 背景と概要	87
6.2 Non-degraded 4:2:0 フィルタの要求条件	88
6.3 Non-degraded 4:2:0 フィルタセットの設計方法	89
6.3.1 要求条件の数学的な定式化	89
6.3.2 相互変換誤差の累積抑止の条件の定式化	90
6.3.3 4:2:2/4:2:0 の色差画素の位置関係に基づく位相条件	93
6.3.4 フィルタの設計方法	94
6.3.5 設計したフィルタの係数	98
6.4 評価	100
6.4.1 比較のための参照フィルタセット	100
6.4.2 シミュレーションによる画質評価	101
6.5 本章のまとめ	109
第 7 章 結論	110
7.1 本研究の成果	110
7.2 研究成果の国際標準化への貢献	112
7.3 研究成果の産業への貢献	113
7.4 今後の展望	114
謝辞	115
業績	123
Appendix	128

第1章 序論

1.1 はじめに

動画画は放送・DVD・Blu-ray等の民生用途から、監視等の業務用途など幅広い分野で活用されている。しかし、動画画の情報量はそのままでは大きく、例えば地デジ等で用いられているHD解像度の非圧縮映像(幅1920画素×縦1080画素×1.5(YUV)×30フレーム/秒)の情報量は、1秒当たり約1Gビットとなり、そのままでは取扱いが困難である。このことから、映像の空間・時間方向の相関を取り除き、動画画データを圧縮する動画画符号化技術が研究され、同時に標準化されてきた。

図1は、動画画符号化の制定時期と主なアプリケーション、その画像サイズについて説明したものである。

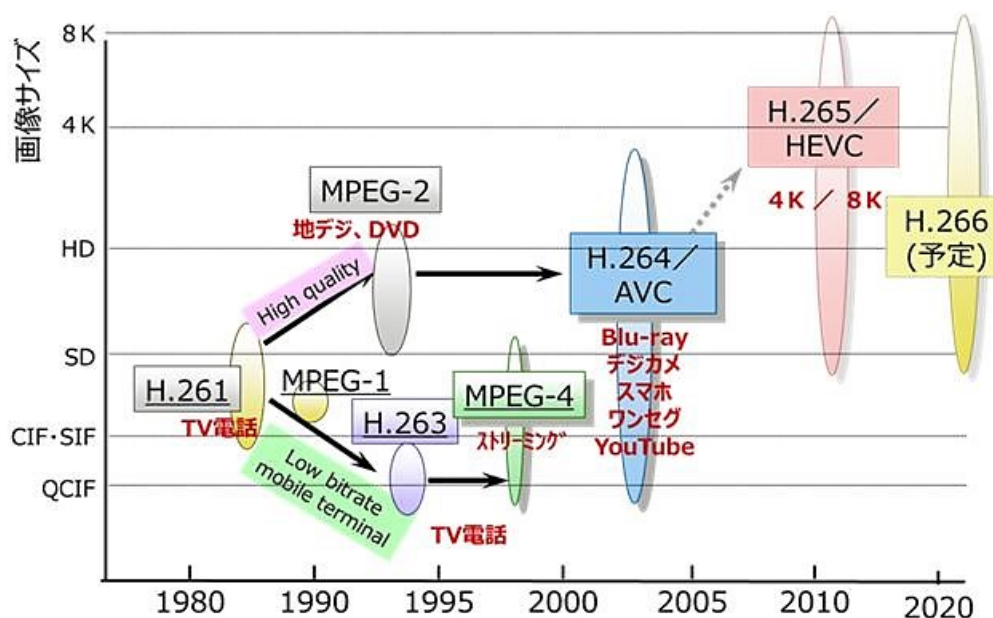


図1 動画画符号化国際標準の制定時期と主なアプリケーション

当初はDPCM [1]やベクトル量子化等 [2]の数々の技術が提案されてきたが、実装容易性や性能の観点から、時間方向の相関を取り除く動き補償(MC/Motion Compensation)と空間方向の相関を取り除く離散コサイン変換(DCT/Discrete Cosine Transform)等の直交変換を組み合わせるMC-DCT方式が主流となった。

そして、このMC-DCT方式がベースとなり、1990年にCCITT(現在のITU-T)にて、初めて動画画符号化の国際標準H.261 [3]が制定された。その後、MC-DCT方式をベースに様々な改良が加えられ、MPEG-1 [4]、H.262/MPEG-2 [5]、H.263 [6]、MPEG-4 [7]、H.264/AVC [8]、H.265/HEVC [9]等の動画画符号化の国際標準化が制定された。

本章では、まず、1.2節で動画画符号化と関連分野の技術の歴史と進化を説明し、

1.3 節で動画像符号化技術の原理、1.4 節で各動画像国際標準化の採用技術や特徴を説明する。そして 1.5 節で本研究の目的及び論文の概要を説明する。

1.2 本研究の背景：動画像符号化と関連分野の技術の歴史と進化

本節では、動画像符号化技術と関連分野の技術の歴史を振り返り、動画像のアプリケーションと要求条件の変化、それに伴う動画像符号化標準化の進化について論じる。

図 2 は、1990 年を基準とした半導体・記録・通信技術の性能向上の比率、及び、各種の動画像符号化 LSI の実用化年と各 LSI が 1 秒間に処理する画素数(画素処理レート)の関係を示したものである。なお、本図においては、画素処理レートは 1992 年前後に実用化された H.261 LSI の性能である CIF/15fps を基準とした。また、LSI の実用化年については文献 [10]等や各社のプレスリリース等を参考とした。

半導体技術については、ムーアの法則により、着実にその性能を延ばしてきた。文献 [11]によると、1975 年から 2010 年の間に実用化された CPU の演算能力と電力の比を取ると、1.52 年で 2 倍の性能向上が見られるとしている。その結果、2015 年では、1995 年の約 10 万倍の電力性能比の向上が達成されている。

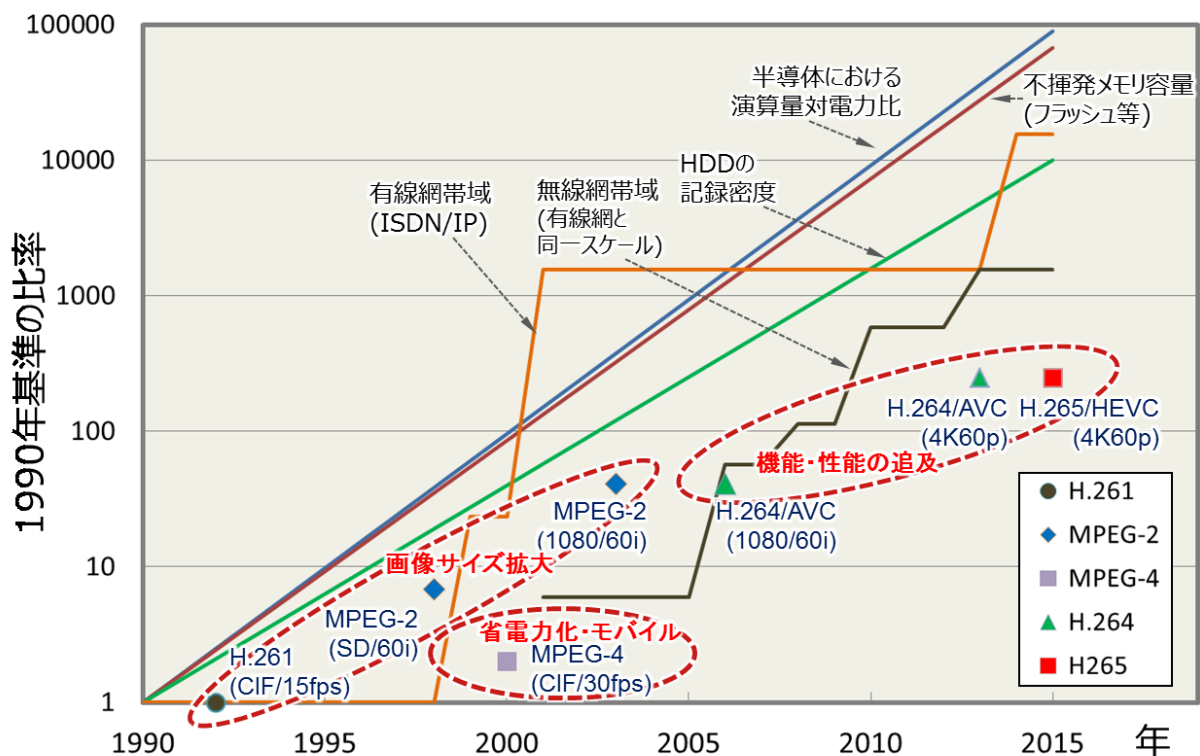


図 2 1990 年を基準とした半導体・記録・通信技術の性能向上比、及び、各種の動画像符号化 LSI の実用化年と処理画素レートの関係

記録技術についても、同様に大幅な性能向上が実現されている。まずフラッシュメモリ等の不揮発メモリに関しては、記事 [12]では、ISSCC2015において、1994年から2016年の間、記録密度が一年で1.56倍ずつ向上したことが報告されている。その結果、この25年で記録密度が約10万倍となった。

また、HDDに関しては、記事 [13]では、2015年9月に米国で開催された「ストレージ開発者会議(SDC 2015)」において、1990年に0.1Gbit/平方インチであった記録密度が、2015年には約1万倍の1000Gbit/平方インチまで向上したと報告されている。なお、SDCでの報告では、HDD記録密度向上のグラフは僅かに上に凸の形状を示しているが、図2では、簡単のために直線で示した。

同様に、通信技術の進歩により伝送帯域も大幅に拡大した。まず有線網については、1990年時点では64kbpsのISDNに限られていた。その後、1999年に1.5MbpsのADSL、2001年には100MbpsのFTTHが開始され、2014年には、1GbpsのFTTHサービスも開始された。その結果、帯域は1990年に比べ約1万倍となった。

無線網についても、2001年に3G通信で384kbpsのサービスが開始され、その後、2006年には3.5Gにより3.6Mbps、2010年には4Gにより37.5Mbpsと帯域が向上し、2013年には100Mbpsまで帯域が拡大している。

このように、この25年の間で、半導体技術、記録技術、通信技術とも4~5桁の性能向上が達成されている。

次に半導体・記録・通信技術の進歩が動画像符号化技術に与えた影響を論じる。まず1992年前後にH.261 LSI([14]のFedis-T向けLSIや[15]等)が実用化された。その後、1998年前後にDVD等をターゲットとしたSD対応のMPEG-2符号化LSI([16][17][18])が複数実用化され、HD対応のMPEG-2符号化LSIも2003年前後に実用化された[19]。また、これと並行して、W-CDMAでのモバイルアプリケーションに向け、低消費電力のMPEG-4符号化LSIが2000年前後に実用化された[20]。ここで、図2より、H.261からMPEG-2の各符号化LSIの処理可能な画素レートは、ほぼ半導体の性能向上と比例していることが分かる。また、H.261、MPEG-1、MPEG-2、H.263、MPEG-4までの動画像国際標準化においては、動きベクトル精度や画素補間フィルタは概ね同じであることから、画素当たりの必要な演算量のオーダーも同等と考えられる。このことから、2000年代前半までは、国際標準化の符号化方式自体には大きく手を入れることなく、半導体・記録・通信技術の進歩に合わせ、画像サイズやフレームレート拡大、及び、省電力化によるモバイル化を目指した時代であるといえる。

その後、2003年にH.264/AVCが標準化され、符号化に必要な各要素技術を徹底的に見直し改善することにより、従来のMPEG-2の半分の情報量で同等の画質の実現が可能となった。一方、計算量がMPEG-2の10倍程度となることから、HD対応の

H.264 符号化 LSI の実現には MPEG-2 の 3-4 年後の 2006 年~2007 年まで半導体技術等の進化を待つ必要があった [21] [22] [23]。その後、H.264/AVC は様々な用途で活用され、その後、民生用の 4K に対応した H.264/AVC 採用の製品の登場は 2013 年前後となった [24]。さらに、2015 年になると 4K に対応した H.265/HEVC 符号化 LSI が相次いで発表された [25] [26]。この 2000 年代後半をみると、符号化 LSI が処理する画素レートの伸びが、半導体技術や蓄積技術の伸びに比べて緩やかであることが分かる。4K/60p の画素レートは CIF/15fps の約 200 倍であり、半導体・蓄積・通信技術の進歩(1 万~10 万倍)から比べると、2~3 桁小さい。このことから 2000 年代前半までは動画像符号化の計算コストが実装可否を支配していたが、それ以降は半導体技術の進歩で符号化以外の処理に計算コストを振り向けることができ、さらに蓄積・通信技術の進歩と併せて、数多くのアプリケーションが発展することが可能になったともいえる。

今後は、さらに 4K/8K 放送が始まり、映像の更なる高精細化に向け、仮称 H.266 等の標準化が進むとともに、CGM を含めた大量のコンテンツ活用に向けた様々なサービスが更に発展すると考えられる。

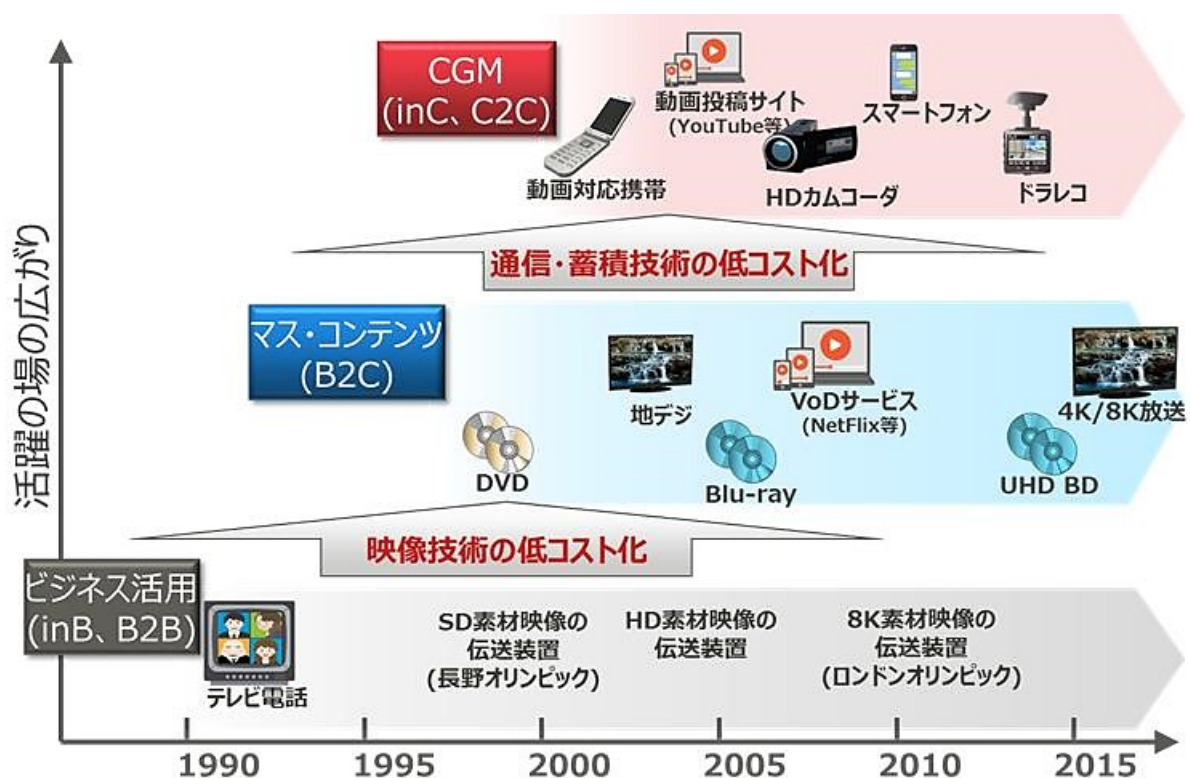


図 3 デジタル映像の活用分野の広がり

次に、前述の半導体・記録・通信技術の進歩が動画像符号化技術に与えた影響に基づいて、デジタル動画像の活用分野の広がりについて論じる。図 3 は、デジタル

動画像の活用の広がりについて説明したものである。この図のように、1990年頃から始まる企業内での活用、2000年前後から始まるマス・コンテンツの消費者への提供手段としての活用、そして、2005年前後から始まるコンシューマ自身によるデジタル映像の生成・活用、いわゆる CGM(Consumer Generated Media)活用の三つに分けられると考える。

まず 1990 年頃から始まる企業内活用の時代背景について述べる。1980 年代後半から 1990 年代初頭は、まさに映像のデジタル化について様々な研究・開発が始まった年代である。しかし、当時のデジタル通信回線は、64kbps の ISDN か企業の専用回線しかなく、一般にデジタル動画が普及するまでには至らなかった。また、動画像符号化の実現は当時の半導体技術(設計ルール $1\mu\text{m}$)ではハードルが高いものであった。このようなことから、最初のデジタル動画像技術の活用は、高い回線コストや専用装置の導入コストを負担できる企業に限られ、H.261 を活用した企業向けテレビ電話等が実用化された。その後、ATM 等のデジタル通信路の発展により 1998 年前後から数多くの MPEG-2 映像伝送装置が実用化され [27]、長野オリンピックでは、MPEG-2 による素材伝送が始まった [28]。その後、インターネットの発展や撮影・表示技術の発展に合わせ、放送局等を中心に HD [29] [30]、8K 対応 [31] の映像伝送装置が実用化され放送・監視等の業務用映像伝送の市場が拡大した。

次に 2000 年前後から始まる、マス・コンテンツの消費者への提供のためのデジタル動画像活用について説明する。まず、1996 年に DVD が標準化され、4.7GB~9.4GB のデジタルデータの記録が可能となり、また同時期に半導体技術の進歩により、デジタル動画像の民生用の SD 解像度の復号 LSI が開発された。そして、これらの技術進歩により、DVD による映画等のマス・コンテンツの消費者への提供が可能となった。そして、2002 年に HD 映像を記録する Blu-ray の規格化が行われ 2006 年に初の Blu-ray ROM プレイヤーが発売された。一方、放送分野においても 2003 年に HD 解像度の地上デジタル放送が開始され、デジタル化の流れは決定的となった。さらに、2000 年代初頭に NTT 等により ADSL や FTTH が相次いでサービスが開始され、インターネットの帯域は数 Mbps~100Mbps まで一気に拡大した。また、モバイル通信は 2001 年に 3G、2012 年に 4G サービスが開始され、消費者が高速な通信回線を利用できるようになった。これに伴い、2007 年以降に acTVila や Netflix 等のインターネットを活用した VoD サービスが急速に普及した。さらに 2018 年からは BS/CS 衛星での 4K/8K 放送が開始される予定である。

さらに、2005 年前後から CGM 動画像が爆発的に広がった。前述のとおり、半導体・通信・蓄積の各技術は指数関数的に性能が向上し、2005 年前後を境に、動画像を扱う端末やシステムは動画像符号化以外の処理に計算コストを振り向けることができるようになった。さらに蓄積・通信技術の進歩と併せて、大量の映像を記録

し、インターネットを用いて伝送することが可能となった。これらのことから、カムコーダーのような HD 映像の記録装置に加え、スマートフォンやドライブレコーダ等、消費者自身が簡単に映像を撮影し活用することが可能となった。またコンピュータの計算能力の向上により編集も容易となり、YouTube や Facebook 等の SNS での映像活用も広がった。

このように、当初は動画像の活用は企業に限られていたが、動画像符号化、半導体、通信、蓄積等の様々な技術の進歩により、今や、動画像は一般消費者の日常に欠かせない存在までになった。

1.3 動画像符号化技術の原理

本節では、これまで規格化された動画像国際標準の基本原則である MC-DCT 方式の概要と映像フォーマットについて説明する。

1.3.1 MC+DCT 動画像符号化技術の概要

図 4 は、これまで規格化された全ての動画像国際標準化が採用している MC-DCT 方式の動画像符号化技術の原理を示したブロック図である。後述の標準化毎に各ブロックの詳細は異なるが、処理の概要を説明する。

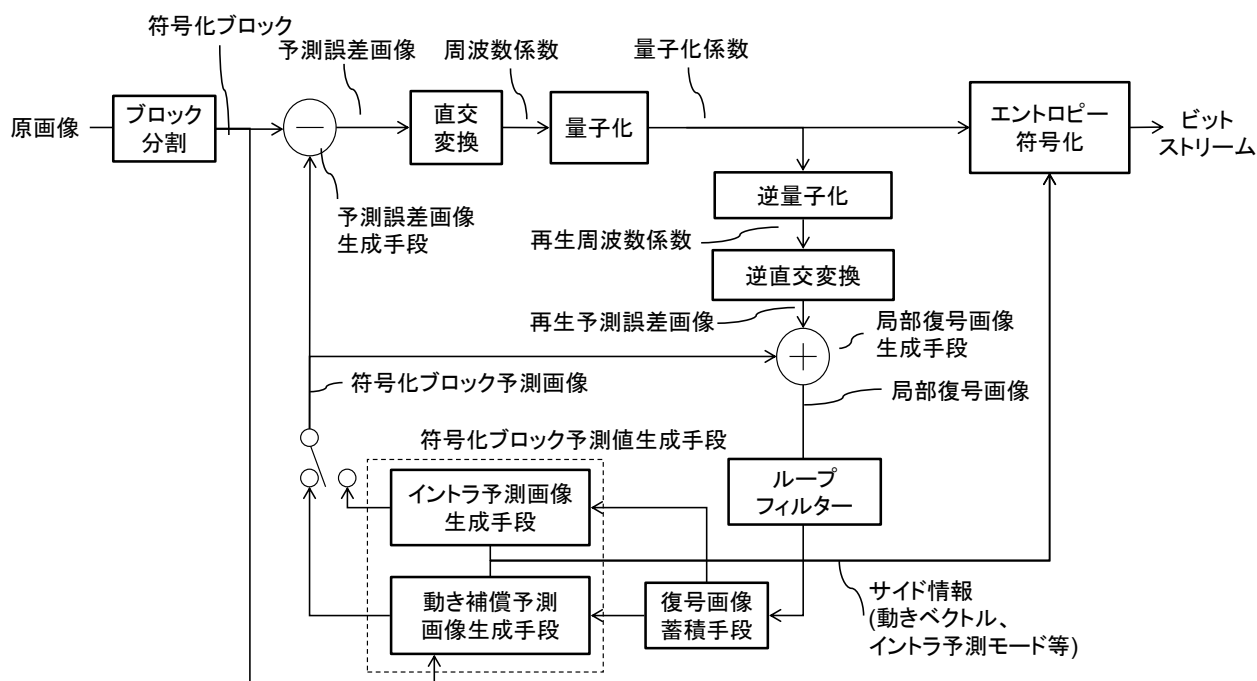


図 4 MC+DCT 動画像符号化技術のブロック図

まず、入力された原画像をブロックに分割する。そして、ブロック単位で、そのブロックを予測するための符号化ブロック予測画像を生成する。予測値生成については、動き補償予測とイントラ予測の2種類がある。そして、動き補償予測とイントラ予測は、符号化ブロック毎に効率の良い方が選択される。

動き補償予測は、過去に符号化した画像の中で、符号化対象ブロックに近い画像を探索し、その画像を予測画像として用いるとともに、相対位置を動きベクトルとして符号化する。標準化の世代が進むとともに、予測画像の生成における画素補間フィルタや動きベクトル精度、予測画像のブロック分割のバリエーション等が高度化し予測性能が向上する一方、必要な演算量も高まってきた。

イントラ予測は、H.264/AVC から標準に取り入れられた機能で、符号化対象ブロックに接する符号化済ブロックの境界画素から外挿予測して予測画像を生成する。

次に、予測誤差生成手段で生成された予測画像を符号化対象ブロック画像から減じて予測誤差画像を生成する。そして、予測誤差画像の空間相関を取り除くため、2次元の直交変換を行う。直交変換は、主に離散コサイン変換か、これに類するものが用いられる。1次元の離散コサイン変換は、 N を次元数、 x_n を変換前の1次元ベクトルの n 番目の値、 X_k を変換後の k 番目の係数として、下記で表される

$$X_k = \sum_{n=0}^{N-1} x_n \cdot \cos\left(\frac{\pi}{N} \cdot \left(n + \frac{1}{2}\right) \cdot k\right) \quad (1)$$

2次元直交変換は、上記の1次元直交変換を水平・垂直に行うことで実現される。次元数 N は、H.263までは8であったが、H.264/AVC以降は、4、8、16、32等、必要に応じて最適な次元数が用いられる。また、離散コサイン変換は無理数を含むため、実装により復号装置間の画像が一致しない問題があったが、H.264/AVC以降は、整数係数とスケールリングを組み合わせた直交変換により、装置間で復号画像が完全一致するようになった。

次に、周波数変換係数を量子化する。一般には、ブロック及び周波数係数毎に量子化器が定められ、周波数変換係数にオフセット等を加算して量子化器で除算し、整数に丸めることで量子化係数が得られる。

そして、動きベクトルやイントラ予測等の予測画像生成に必要なパラメータと、量子化器や量子化係数等の予測誤差生成に必要なパラメータをエントロピー符号化することで画像を符号化する。エントロピー符号化は、主にH.263までは可変長符号が主流であったが、H.264/AVC以降は、コンテキストに基づきシンボルの確率テーブルを切り替えて算術符号を行うCABAC [32]が主流となっている。

最後に、符号化側と復号側で予測に同一の画像を参照するために、符号側でも量

子化係数に逆量子化・逆直交変換を施し、そして予測画像と加算して復号画像を生成し、最後に復号画像蓄積手段に蓄積する。この際、H.263 以降の符号化標準では、ループフィルターを用いてブロックノイズ等の誤差を抑制する。

1.3.2 画像フォーマット

動画像符号化では、YUV 4:2:0 あるいは YUV 4:2:2 という映像フォーマットが広く用いられている。本小節では、これらの動画像符号化標準が採用する映像フォーマットについて説明する。

一般に、デジタル画像の撮影デバイスは R(赤)、G(緑)、B(青)の三原色の素子を持ち、各画素は R/G/B の各成分の明るさで表現される。一方、人間の視覚特性について、各色のおよそ平均で表される輝度(明るさ)の成分については空間周波数の識別感度が高いが、赤と緑の差である R-G、及び赤と青の差である R-B の識別感度は低いことが知られている [33]。すなわち、人の目の空間周波数感度は明暗の差には敏感であるが、色の差については鈍感である。

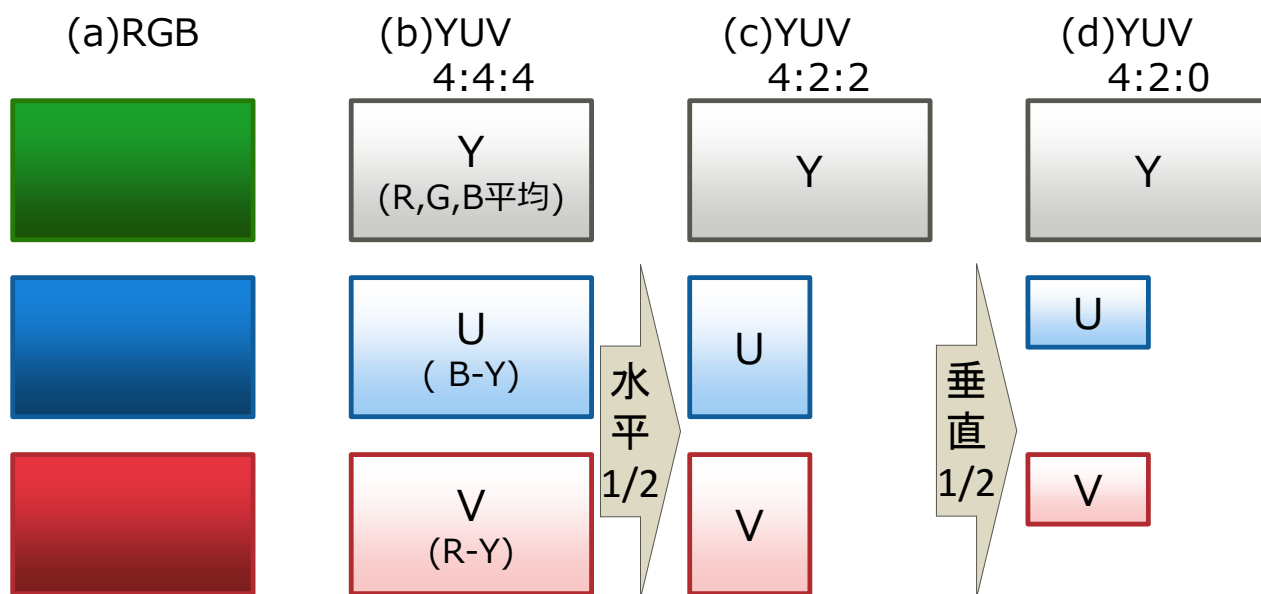


図 5 動画像符号化で用いられる映像フォーマット

この特性を利用し、動画像符号化標準では、U/V をダウンサンプリングにより低解像度化した YUV フォーマットが広く用いられる。図 5 は、RGB フォーマットの映像から、動画像符号化で用いられている複数の YUV フォーマットへの変換の概要を示したものである。まず、図 5 (a)の RGB フォーマットから図 5 (b)の YUV の 4:4:4 フォーマットへの変換については、画素毎に、R,G,B の重み付き平均で輝度を表す Y、B-Y に相当する U、R-Y に相当する V を求めることで実現される。

次に図 5 (c)の YUV の 4:2:2 フォーマットは、YUV 4:4:4 の U 及び V 成分の水平方向の解像度を 1/2 縮小することで得られる。また、図 5 (d)の 4:2:0 フォーマットは、さらに YUV 4:2:2 フォーマットの U 及び V 成分の垂直の解像度を 1/2 に縮小することで得られる。

この YUV 4:2:0 フォーマットは、ほとんどの民生機器・放送等で用いられている。一方、YUV4:2:2 フォーマットは、放送局等の品質を重視する業務用途で主に用いられている。

復号・表示時には復号された YUV4:2:0 あるいは 4:2:2 の UV を拡大して YUV4:4:4 に変換し、画素毎に RGB に逆変換して、モニタ等で表示する。

動画像は、上記のフォーマットで静止画を 1 秒間に 10~60 枚ほど撮影し、時間方向に順次表示することで、人間の目には動画として認識される。

表 1 は、よく用いられる動画像フォーマットについて、名称、サイズ、フレームレート、主な用途をまとめた。

表 1 広く使われる動画像のサイズ・フレームレートのフォーマット

名称	画像サイズ	よく用いられるフレームレート (単位 fps、p は Progressive、i は Interlace)	主な用途
QCIF	176×144	5p~15p	TV 電話
CIF	352×288	5p ~30p	TV 電話
QVGA	320×240	5p~30p	携帯電話向け動画
VGA	640×480	15p~30p	PC 向け動画
SD	720×480	24p、30p、60i	NTSC 採用のテレビ放送
	720×576	24p、25p、50i	PAL 採用のテレビ放送
HD	1280×720	24p、25/30p、50/60p	放送、Blu-ray、カメラ、インターネット等、多くの用途
	1440×1080	24p、25/30p、50/60i、50/60p	
	1920×1080		
4K	3840×2160	25/30p、50/60p	放送、Ultra HD Blu-ray、カメラ
8K	7680×4320	50/60p、120p	放送

1.4 各動画像符号化国際標準の技術の特徴

表 2 は、各動画像符号化国際標準の主な技術の特徴を比較したものである。

前述のとおり、H.261 から MPEG-4 までは動きベクトル精度や画素補間フィルタ等の基本構成は同等であるが、H.264 以降、半導体技術の進歩に伴い、符号化に必要な要素技術が徹底的に見直されて複雑な処理が導入され、符号化効率が大幅に進歩した。

本節では、各国際標準の概要について、簡単に説明する。

表 2 動画像符号化国際標準の概要比較

国際標準化	ソース符号化							エントロピー符号化
	ブロック分割	動きベクトルの精度	サブペル予測フィルタのタップ数	参照画像保持枚数	イントラ予測	直交変換基底サイズ	ループフィルタ	
H.261	16×16	1 pixel	なし	1	N.A	8×8	N.A.	2D Huffman ／固定確率算術符号
MPEG-1	16×16	0.5 pixel	2	2	DC係数予測	8×8	N.A	2D Huffman
MPEG-2 /H.262	16×16	0.5 pixel	2	2	DC係数予測	8×8	N.A	2D Huffman
H.263	16×16, 8×8	0.5 pixel	2	2	AC/DC 係数係数	8×8	N.A	3D Huffman
MPEG-4	16×16, 8×8	0.5 pixel	2	2	AC/DC 係数係数	8×8	N.A	3D Huffman
H.264/ AVC	16x16から 4x4まで 階層的分割	0.25 pixel	6	16	周辺画素 の外挿予測 (9方向)	8×8/ 4×4	Deblock Filter	CABAC/ CAVLC
H.265/ HEVC	64x64から 4x4まで 階層的分割	0.25 pixel	8	16	周辺画素 の外挿予測 (33方向)	32x32,16x16, 8x8,4x4	Deblock Filter / SAO	CABAC

1.4.1 H.261

H.261 [3]は ITU の前進の CCITT において、1990 年に世界で初めて制定された動画像符号化標準である。本標準は 4:2:0 画像フォーマットのみを対象としている。まず画像をマクロブロックという単位で、輝度画像を 16×16 画素、色差画像を 8×8 に分割する。動き補償は、当時のハードウェア技術を考慮し、整数精度のベクトルのみが規定されたが、符号化誤差に含まれる高周波成分の伝搬抑止のために、予測画像へのローパスフィルタが導入された。エントロピー符号化はハフマン符号と算術符号の双方が規定されたものの、市場では主にハフマン符号が広く用いられた。また、量子化係数の符号化には、非ゼロの係数が少ないことを考慮し、係数をジグザグ順にスキャンし、ゼロ係数の個数(Run)と有効係数の値(Level)を組み合わせた Run-Level の 2D ハフマン符号が採用された。また、直交変換は DCT が採用されたが、DCT の定義に無理数が含まれ、復号時に装置毎に異なる誤差が発生してしまうことから、許容誤差の範囲が規定された。

1.4.2 MPEG-1

MPEG-1 [4]は ISO/IEC で 1991 年に制定された動画像符号化標準であり、当初は CD-ROM 等へのビデオ蓄積が主な用途とされていた。MPEG-1 では、動き補償の双方向予測(B ピクチャ)、サブ画素精度の動きベクトル(MPEG-1 では 0.5 画素)、Group

of Picture(GOP)というという、以降の動画像国際標準で採用されている基本機能が初めて導入された。

1.4.3 MPEG-2/H.262

MPEG-2/H.262 [5]は、ISO/IEC と ITU-T が共同で 1995 年に策定した国際標準であり、現在でも DVD や地デジ等で広く用いられている。基本的な構成は MPEG-1 をベースとしている。初めて Interlace に対応したため、動き補償や DCT の処理が、マクロブロック単位にフィールド/フレームを考慮した処理に切り替え可能なよう工夫されている。

1.4.3.1 H.263

H.263 [6]はモバイルやインターネットでのテレビ電話等の活用を主なターゲットとして ITU-T で 1996 年に標準化された。H.261 や MPEG-1/2 をベースに、有効係数の 3D ハフマン符号が採用された。また 1998 年に制定された H.263 Version 2 では、イントラ予測時の空間相関を利用した AC/DC 係数予測等や、ブロック境界の劣化を抑制するデブロックフィルタ、そして筆者が提案した、動的に解像度を変換しフレームレートを向上させる動的解像度変換技術等が採用された。

1.4.4 MPEG-4

MPEG-4 [7]は H.263 同様にモバイルでの活用を主なターゲットとして ISO/IEC で 1998 年に標準化された。H.263 を参照して標準化したことから、ほとんどの構成要素は H.263 と類似している。筆者が提案した低遅延向けに解像度を動的に変換する動的解像度技術も、H.263 と同様に MPEG-4 でも ARTS Profile に採用された。

一方、オブジェクトベース符号化も規定されたが、オブジェクトのセグメンテーション技術が実用レベルになかったことで、ほとんど普及しなかった。

1.4.5 H.264/AVC

H.264/AVC [8]は ITU-T と ISO/IEC が共同で 2003 年に策定した国際標準で、符号化に必用な各要素技術を徹底的に見直し改善することにより、従来の MPEG-2 の半分の情報量で同等の画質の実現が可能となった。一方、複雑な処理のため符号化の計算量が MPEG-2 の 10 倍程度と想定され、マンモスコードックともよばれた [34]。そのため、実用化には時間がかかったものの、この標準の普及により民生から業務用まで、映像を活用するアプリケーションが爆発的に普及した。

まず、直交変換基底を整数で定義することで復号時に生じる DCT の誤差が解消され、あらゆる復号装置での復号画像が同一となることが保証された。また基底サイズとして 4x4 と 8x8 が選択可能となった。

そして復号画像の同一性が保証されたことから、イントラ予測では、符号化対象ブロックの左や上の符号化済のブロックの境界画素を 9 種類の方向のいずれかに外挿予測して符号化対象ブロックの予測画像を生成する機能が採用された。

また、動き補償予測では、マクロブロックを 16x16 から 4x4 までの可変ブロックサイズに分割して動き補償することが可能となり、1/4 画素精度の動きベクトル、6 タップフィルタによる画素補間と併せ、予測画像の画質が大幅に向上した。

エントロピー符号は可変長符号ベースの CAVLC と算術符号ベースの CABAC が規定された。双方とも、過去に符号化したシンボル等から得られたコンテキストにより符号化対象シンボルのエントロピー符号化を切り替える手法がとられた。とくに CABAC では、全シンボルをバイナリ列に変換し、多くのシンボルの確率テーブルをコンテキストにより切り替えることで、大幅なエントロピー符号化の性能向上に貢献した。

また、インターレースも強化され、上下 2 個のマクロブロックからなる MB Pair 毎にフレーム／フィールド予測を切り替える MBAFF や、筆者が提案した、輝度ベクトルと色差ベクトルの平行化の補正等が採用された。

一方、従来の MPEG-2 に比べて実装に必要な計算量やメモリバンド幅が各段に多くなったことから、前述の通りマンモスコデック等と評されることもあった。そのため、実用化に時間を要し、2003 年の標準化策定から 3 年後の 2006 年に、HD 対応の符号化 LSI が実用化された。この実用化に際し、筆者らは低演算量と高画質を両立するアルゴリズムを開発し、符号化 LSI 及び映像伝送装置の実用化に貢献している。現在では、Blu-ray、MP4、カメラ、映像配信等、幅広い分野で H.264/AVC が活用されており、今日のように映像が気軽に扱えるようになる礎となった。

1.4.6 H.265/HEVC

H.265/HEVC [9]は H.262/MPEG-2 や H.264/AVC と同様に ITU-T と ISO/IEC が共同で 2013 年に制定した国際標準である。H.264/AVC に対して様々な改良が加えられ、H.264/AVC の更に半分の情報量で同等画質を実現可能とするものである。この H.265/HEVC は、後述の可変ブロックサイズ符号化により、特に 4K/8K 映像での符号化効率が大きく改善した。このことから、従来の H.264/AVC の適用分野に加え、BS/CS での 4K/8K 放送への適用が決定している。

最も大きな違いは、動き補償のブロック分割サイズやイントラ予測方向、直交変換基底サイズの更なる多様化である。動き補償の予測ブロックは 64×64 から 4×4

まで様々なサイズが選択可能となり、画素補間フィルタも 8 タップとなった。また、イントラ予測の予測方向も 33 種類となるなど、H.264/AVC に比べて予測画像の品質向上が図られた。

直交変換の基底サイズも 32×32 、 16×16 、 8×8 、 4×4 が選択可能となった。ループフィルタにおいては、画素単位で適応的にフィルタを施す SAO も新たに導入された。動きベクトルの符号化では、近隣ブロックのモードをコピーするマージモード等も採用された。

さらに、補助情報(SEI)において、符号化対象の YUV420 映像が、第 3 章で提案する ND420 フィルタで生成されたことを示す情報が規定された。この情報により、4:2:2/4:2:0 相互変換を複数回繰り返しても、色成分の画像劣化を抑止することが可能となった。

1.5 本研究の目的と位置付け

本研究は、複数の動画像符号化国際標準化への適用・貢献を目的とした性能改善技術の研究と、国際標準 H.264/AVC で規定された符号化装置を実用化する際の画質・機能向上技術の研究に関するものである。本研究を通じ、デジタル信号処理と情報理論を背景として、映像符号化における入力信号処理の改善と視覚特性を考慮したレート歪特性の最適化により、符号化映像品質及び符号化効率の改善を達成し、映像の高能率符号化の最先端分野に新たな知見を与え、国際標準及び放送機器として社会生活基盤の構築に貢献することができたと考えている。

図 6 は本研究の位置付けを示したものである。まず、第 2 章、第 3 章、第 4 章では、動画像標準化の符号化効率等の基本性能の向上のための研究について論じる。そして、第 5 章、第 6 章ではこれらの国際標準化に準拠した動画像システムを実用化する際に特に重要となる、機能・性能向上のための研究について論じる。

まず、符号化効率の向上を目指した第 2 章、第 3 章、第 4 章の概要を説明する。1.2 節で述べたとおり、2000 年代前半までは、動画像符号化は限られた演算量の中での性能・効率の実現を目指していた。このような背景の元、第 2 章では、演算量を増やすことなく、TV 電話等で動きの激しいシーンでのコマ落ちを抑制するための動的解像度変換技術の研究について論じる。従来、テレビ電話で動画像を伝送する際に人物の大きな動きが生じると、過度なコマ落ちの発生が問題となっていたが、動きの激しいシーンを検出し、予測誤差信号を $1/2$ 倍の低解像度に変換して符号化する動的解像度変換技術により、演算量を増やすことなく、静止背景の画質や PSNR をほぼ保ったまま、フレームあたりの情報量を最大 $1/2$ 程度まで削減し、コマ落ちを抑制することが可能となることを示す。

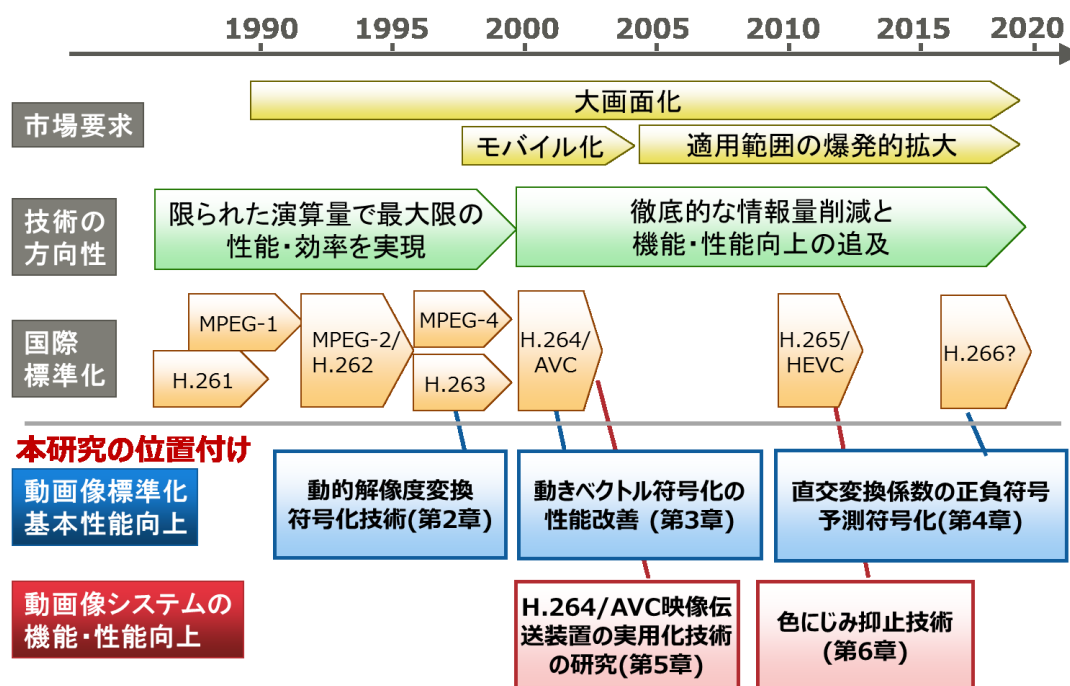


図 6 本論文の内容と各動画国際標準との関係

また 2000 年代の半ば以降になると、半導体等の技術進歩による符号化処理の計算コスト低下に応じて、動画符号化標準は徹底的な情報量削減と機能・性能の追及が行われるようになった。このような背景の元、第 3 章では、H.264/AVC に向けた動きベクトルの符号化効率改善技術の研究として、コンテキストベースのエントロピー符号化技術、及び Interlace 符号化時の色差ベクトル補正技術について論じる。まず、コンテキストベースのエントロピー符号化技術に関しては、予測誤差ベクトルの確率分布は、対象ブロック周辺の動きベクトル場が一様の場合には 0 近辺の確率が高く、ベクトル場が乱雑な場合には 0 近辺の確率が低くなることを見出した。これに基づいて、動きベクトル場の一様性に応じて 2 種類のエントロピー符号化を切り替えるアルゴリズムを考案し、動きベクトル情報を約 10% 削減可能なことを示す。また、Interlace 符号化時の色差ベクトル補正技術の研究については、H.264/AVC 国際標準化のドラフト策定時、動きベクトルが符号化画像と異なるパリティのフィールドを指し示す場合、輝度ベクトルと色差ベクトルが平行にならず画質劣化を生じさせる課題を見出し、色差ベクトルを補正することで、特定フィールドで最大 1dB の色差画像の PSNR 改善が可能なことを示す。

第 4 章では、さらなる符号化効率改善を目指し、直交変換係数の正負符

号の予測符号化技術について論じる。従来、直交変換・量子化された非ゼロ有効係数の正負符号は全体の 10%の情報量を占めるにも関わらず、これまでは他の情報と無相関と考えられ、非圧縮で 1 ビット/符号の情報が割り当てられてきた。これに対し、符号化対象ブロックと、これに隣接する符号化済の周辺ブロックとの相関性を用いて正負符号を予測符号化することで、正負符号の情報量の約 10%に相当する全体の符号化情報量の約 1%の削減が可能となることを示す。

第 5 章、第 6 章では国際標準化に準拠した動画像システムを実用化する際に特に重要となる、機能・性能向上のための研究について論じる。

第 5 章では、マンモスコードックと呼ばれた H.264/AVC に対応した動画像伝送装置の実用化技術の研究について論じる。2000 年台前半、従来の SD 映像より圧倒的にリアル・高品位な HD 映像での放送や、テレビ・カメラ・ビデオ・PC 等で HD 映像による記録・表示の要望が高まった。その一方、HD 化に伴う機器・伝送コスト増大が放送サービスや機器実現の大きな課題となっていた。この課題を解決するため、人間の視覚特性に基づく H.264/AVC の低演算量の符号化技術、公衆インターネット向けのパケット誤り技術等を開発し、H.264/AVC 映像伝送装置に適用することで、放送の HD 化に貢献したことを示す。

第 6 章では、4:2:2/4:2:0 の両映像フォーマットの間を繰り返し相互変換した際に生じる色差成分の劣化を抑制する色にじみ抑止技術の研究について論じる。4:2:0 対応の装置を用いて複数拠点経由で中継する際に、4:2:2/4:2:0 相互変換で生じる僅かな劣化誤差が累積し、「色にじみ」として画質劣化が知覚されてしまうことが問題となっていた。この課題を解決するため、完全再構成条件等を満たす『色にじみ抑止技術』を提案し、複数回の相互変換後の色にじみを完全に抑制可能となることを示す。

第 7 章では、結論として第 2 章から第 6 章で論じた研究の成果を示す。そして、これらの研究成果の国際標準化及び産業への貢献について論じ、最後に今後の展望について示す。

第2章 H.263／MPEG-4 における動的解像度変換技術¹

2.1 はじめに

TV 電話等のリアルタイム会話型アプリケーションにとって、低遅延のビデオ伝送は最も重要な要件の一つである。そして、低遅延伝送を実現するためには、各ピクチャのビット割り当てを制御し、遅延を十分に低く抑える仕組みが不可欠となる。

TV 電話等を主な適用先とした動画像符号化国際標準 H.263 [6]の符号化方法のテストモデル TMN8 [35]では、量子化ステップサイズと符号化フレームレートの自動調整に基づいた符号化制御が行われる。また、有効なビット割り当てとレート制御を組み合わせた研究も行われている [36] [37] [38]。これらのレート制御アルゴリズムは、各ピクチャの複雑さが大きく変動せず、結果として符号化に必要な情報量が大幅に変動しないことを前提としている。

しかし、動きの大きいシーンを符号化する場合、動き補償で大きな予測誤差が生じることが多く、その結果、符号化時にも多くの情報量が発生してしまう。そのため、特にモバイル等の伝送帯域が限られたネットワークでは、1ピクチャの伝送時間が増大し、頻繁なフレームスキップが生じる。その結果、映像の滑らかさが失われてしまい、主観映像品質の深刻な劣化が生じる。

このリアルタイムの映像圧縮・伝送の課題を解決するために、動きの大きいシーンにおいて、動領域を縮小解像度で符号化しつつ、静止画領域の解像度を高いまま維持する動的解像度変換技術 (Dynamic Resolution Conversion method : DRC) を提案する [39] [40] [41] [42]。本節では、文献 [39]に基づいて、提案する動的解像度変換技術を説明する。

本提案手法では、まず、各ピクチャの複雑さを予測する。そして、動きの激しいピクチャと判定された場合には、動的解像度変換技術を用いる。ここでは、まず、マクロブロックサイズを通常の2倍の 32×32 として予測する。そして、予測誤差を垂直方向と水平方向とも $1/2$ の解像度にダウンサンプリングして、従来と同じ 16×16 のサイズの予測誤差を生成し、この予測誤差を、元画像の $1/2$ サイズの画像として符号化する。本提案手法により、動きの大きい人物等が存在するシーンでも、背景の解像度は高いまま保ちつつ、人物だけを低解像度で符号化することで、フレームあたり情報量削減とフレームスキップの大幅な抑制が可能となり、主観的品质の向上を実現することができた。

また、移動通信に不可欠なエラーの発生しやすい環境下で DRC を適用した際の性能解析についても議論を行う。フレームレートが向上したことで、エラーが影響を

¹ 本章は業績 [J4][C1]の内容に基づいて執筆している。

与える範囲が時空間的に抑制されることから、エラー発生時の主観品質の向上も実現できた。

そして、この技術を H.263 ver.2 に提案した結果 [43] [44] [45] [46]、その効果が認められ、Annex Q “Reduced-Resolution Update Mode”として採用された。さらに、MPEG-4 にも提案し、“Dynamic Resolution Conversion”ツールとしても採用された。

次節以降、アルゴリズム、性能評価、伝送エラー環境下での品質評価について説明する。

2.2 動的解像度変換技術(Dynamic Resolution Conversion)

2.2.1 技術概要

人間の視覚は、動きの速い物体については、細かいテクスチャの知覚感度が下がることが知られている [47]。この特徴を考慮すると、動きや変化の大きい領域のみを低解像度で符号化しても主観品質の大幅な劣化が知覚されないことが期待される。また、動きや変化の大きい領域を低解像度で符号化することにより、フレーム毎のビット数を削減することが可能となるので、結果としてコマ落ちが抑制され、シーン全体の主観品質が向上すると考えられる。提案する動的解像度変換技術は、この考えに基づいたものである。

図 7 は、提案方法のブロック図を示す。この図で、太い線のブロックが従来の H.263 と異なるブロックである。

まず、解像度決定手段では、入力された画像の動きの大小に基づいて、フレーム毎に符号化モードを決定する。動きの穏やかな通常のシーンでは、入力画像が従来の H.263 と同一方法で符号化する通常モードが選択される。一方、動きの大きいシーンでは、縮小解像度モードが選択される。この判定方法は 2.3 節で説明する。

そして、縮小解像度モードが選択された場合には、まず、マクロブロックサイズを通常の 2 倍の 32×32 画素として動き推定・動き補償予測を行う。次に、原画と予測画像の差分により 32×32 画素の予測誤差を求め、これを垂直方向と水平方向とも $1/2$ の解像度にダウンサンプリングし、従来の H.263 と同じ 16×16 画素のサイズの予測誤差を生成する。次に、このダウンサンプリングされた予測誤差は、元画像の $1/2$ サイズの画像、例えば原画が CIF の場合には QCIF として、H.263 と同じ DCT/量子化/可変長符号を用いて符号化される。この提案方式により、低解像度化による係数情報の削減に加え、マクロブロック数が $1/4$ になり MB タイプや動きベクトル等のマクロブロックヘッダ情報も大幅に削減することが可能となる。

受信側では、係数を逆量子化・逆 DCT して生成された 16×16 画素サイズの予測誤差を元サイズの 32×32 画素にアップサンプリングし、予測画像に加算することで画像が復号される。予測画像の未更新部分が元の解像度を保持するので、縮小され

た解像度モードが選択されていても静止領域の詳細を保存することができる。この結果、主観的品質の大幅な低下なしに情報量を削減し、符号化フレームレートの改善が可能となる。

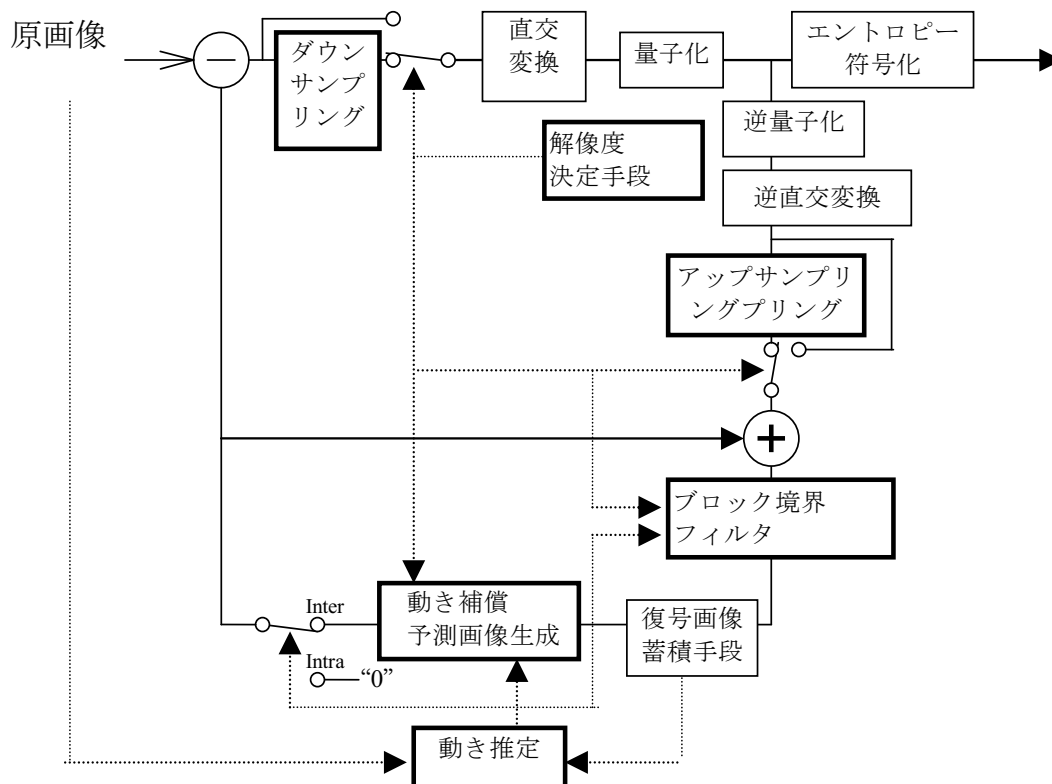


図 7 提案手法のブロック図 ([39]改)

2.2.2 提案手法の符号化手順

本小節では、縮小解像度モード時の符号化手順の詳細を説明する。説明を簡単にするために、入力画像のフォーマットは CIF (352×288) と仮定する。前述のとおり、フレーム毎に符号化モードが決定され、モード情報が受信機にシグナリングされる。図 8 は、縮小解像度モードでのブロック単位の符号化手順を示したものである。

符号化手順の詳細は次のとおりである。

- 通常の H.263 ではマクロブロックは 16×16、ブロックは 8×8 のサイズとなるが、縮小解像度モードでは縦・横とも 2 倍の大きさとなる。すなわち、マクロブロックのサイズは 32×32、ブロックのサイズは 16×16 となる。その結果、ピクチャ内のマクロブロックの数は通常モードの 1/4 である。
- 通常の H.263 の動き推定は 16×16 のマクロブロック、あるいは 8×8 のブロック単位で行われるが、縮小解像度モードでは 32×32 あるいは 16×16 単位で実行される。また、ベクトルの x 成分と y 成分はゼロ、または半整数(±0.5、±1.5、…)に制限

し、その結果、同一の可変長符号で約 2 倍の探索範囲まで指し示すことができる。
 詳細は 2.2.3 小節で説明する。

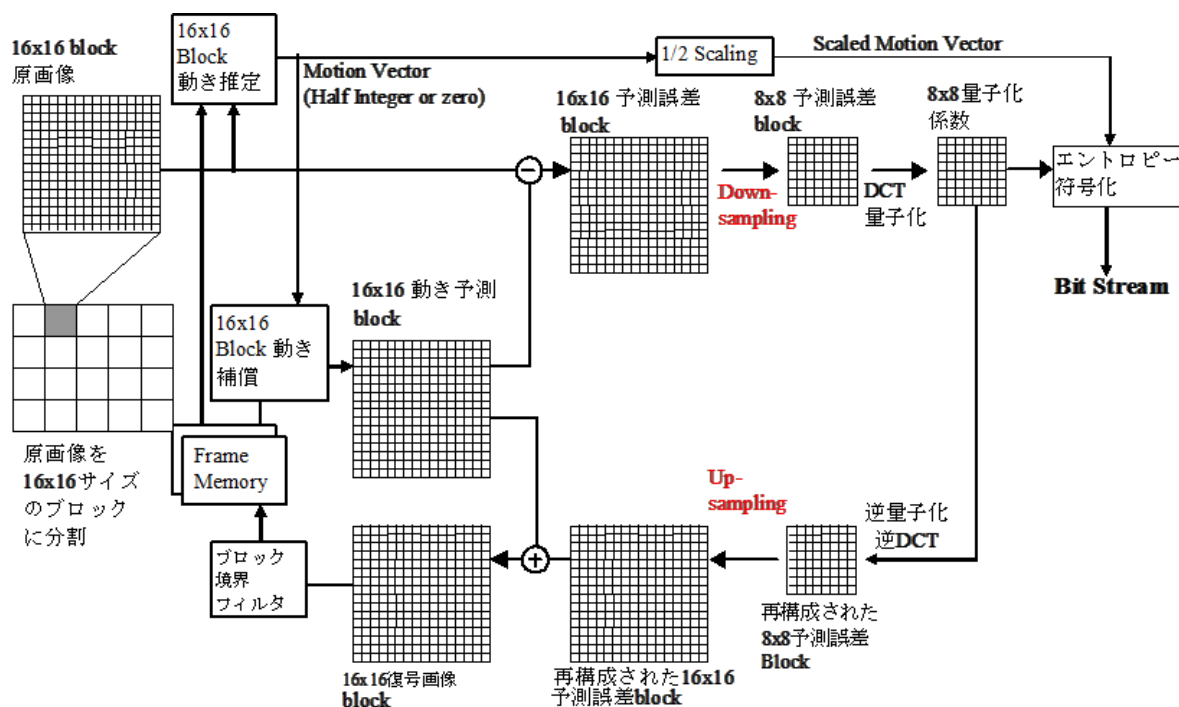


図 8 縮小解像度モードの符号化手順を示すブロック図 ([39]改)

- 動き推定で求めたベクトルに基づいて、各 16×16 ブロックの予測画像を生成し、原画像との差分を各ブロック単位にとり、16×16 サイズの予測誤差ブロックを求める。
- 16×16 予測誤差ブロックを、ブロック単位で縦横 1/2 の大きさにダウンサンプリングする。その結果、ダウンサンプリングされた予測誤差ブロックのサイズは従来の H.263 と同じ 8×8 となる。詳細は、2.2.4 小節で説明する。
- ダウンサンプリングされた 8×8 予測誤差ブロックに対し、通常の H.263 と同じ DCT、量子化が行われて量子化係数が得られる、この量子化係数に、従来の H.263 と同一の可変長符号(VLC)が割り当てられる。
- 局部復号のため量子化係数に逆量子化・逆 DCT を適用し、8×8 の再構成された予測誤差ブロックを生成する。そして、これを縦横 2 倍にアップサンプリングし、16×16 の再構成予測誤差ブロックを生成する。アップサンプリングフィルタの詳細は、2.2.5 小節で説明する。
- 16×16 サイズの再構成された予測誤差ブロックを 16×16 予測ブロックに加え、16×16 の復号画像を生成する。

- 有効係数がある 16×16 のブロックに対し、その境界に接する画素にローパスフィルタを適用する。ローパスフィルタの詳細は 2.2.6 小節で説明する。

次節以降で、縮小解像度モード時の各ステップの詳細を説明する。

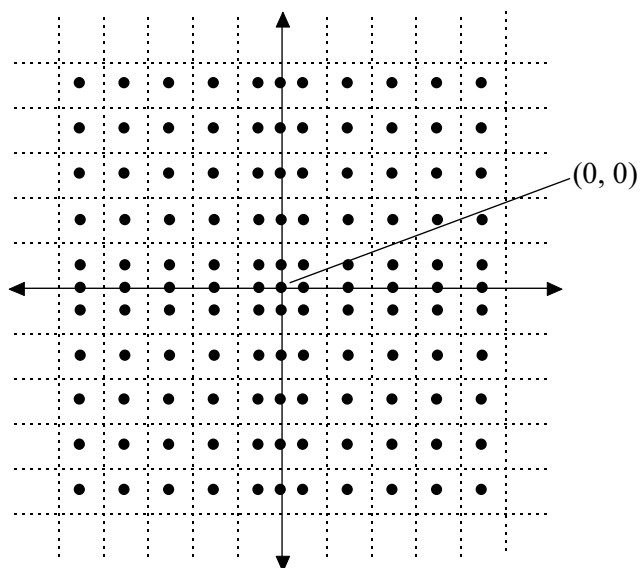


図 9 原点 $(0,0)$ 近辺で選択可能なベクトル [39]

2.2.3 選択可能な動きベクトルの制限

予測誤差は $1/2$ にダウンサンプリングされるため、予測誤差に含まれる高周波成分が失われる可能性が高い。その結果、縮小解像度モードが続くと、失われた高周波成分による誤差が蓄積し、高周波成分の劣化が復号画像に蓄積してしまう。このことから、高周波成分の誤差の蓄積を抑制するために、予測誤差が発生する動領域は、予測画像にローパスフィルタをかける必要がある。一方、静止領域は解像度を保つためにフィルタは施さないほうがよい。

上記の要求をみたすため、ベクトルの表現に「ハーフゼロサーチ」方式 [48] を使用した。これは、ベクトルの x 、 y 成分が取りうる値として、ゼロと半整数 (± 0.5 、 ± 1.5 、 \dots) に限定したものである。

図 9 は、 $(0, 0)$ の周りで動きベクトルが取りうる位置を示す。破線は整数の座標を示す。ベクトルの解像度が $1/2$ になったので、H.263 と同一定義の可変長符号でも、ベクトルが指し示すことが可能な距離が約 2 倍に拡大し、大きい動きベクトルも効率よく表現できる。

「ハーフゼロサーチ」方式で定義されたベクトルを H.263 と同じ VLC で符号化するために、以下のように実ベクトルと符号化ベクトルの変換を行う。*pseudo-MV* を

H.263 と同一定義の VLC で符号化するベクトル値、 MV を実ベクトルと定義する。 MV から $pseudo-MV$ への変換は次式で定義する。

$$\begin{aligned}
 &\text{if } MV = 0 \text{ then} \\
 &\quad pseudo-MV = 0 \\
 &\text{else} \\
 &\quad pseudo-MV = \text{sign}(MV) * (|MV| + 0.5) / 2.0 \tag{2}
 \end{aligned}$$

そして、この $pseudo-MV$ を H.263 で規定された動きベクトルの VLC を用いて符号化する。

2.2.4 ダウンサンプリング

ダウンサンプリングの処理を図 10 に示す。16×16 ブロック単位の動き補償の後、16×16 予測誤差ブロックは 8×8 予測誤差ブロックにダウンサンプリングされる。ここで、実装を容易にするために、ダウンサンプリングで用いるフィルタリングに関してはブロック内の画素のみを使うように設計した。すなわち、図 10 に示す通り、8×8 へ縮小後の画素を A は、16×16 解像度の周辺 4 画素 a, b, e, f の平均で導出する。ここで、 $nint(\cdot)$ は入力値を最も近い整数に変換する関数である。

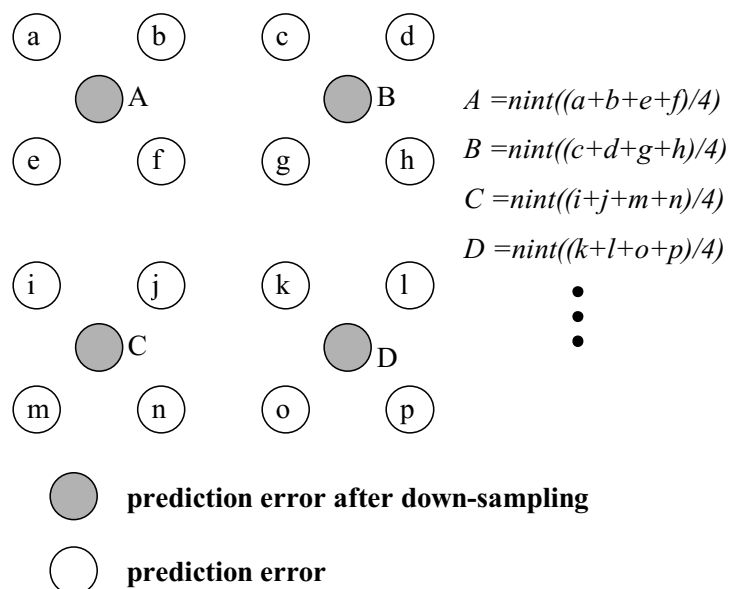


図 10 ダウンサンプリング [39]

2.2.5 アップサンプリング

IDCT で再構成した 8×8 の予測誤差ブロックは、2 倍にアップサンプリングされて 16×16 の予測誤差ブロックに拡大される。このアップサンプリング処理は、ダウンサンプリング同様、処理を単純化するためにブロック内の画素のみを用いる。

アップサンプリング手順を図 11 に定義する。アップサンプリングにはバイリニア補間を用いた。例えば、画素 e を求める場合を考える。この場合、画素 e は画素 A 、 B 、 C 、 D の垂直・水平座標のそれぞれ 1:3 の内分点となるため、この値は次のような式のバイリニア補間によって算出される。

$$e = \text{nint} \left((9A + 3B + 3C + D) / 16 \right) \quad (3)$$

ここで、 $\text{nint}(\cdot)$ は前節で定義したのと同じ丸め関数である。

境界ピクセルに関しては、ミラーリングされたピクセルをブロック境界外に拡張した後に、バイリニア補間を適用する。たとえば画素 d は、 A 、 C のブロック境界の鏡像位置に、それぞれ A 、 C と同じ値を持つ mirrored A 、mirrored C の画素があると拡張することで、下記の計算に帰着できる。

$$d = \text{nint} \left((3A + C) / 4 \right) \quad (4)$$

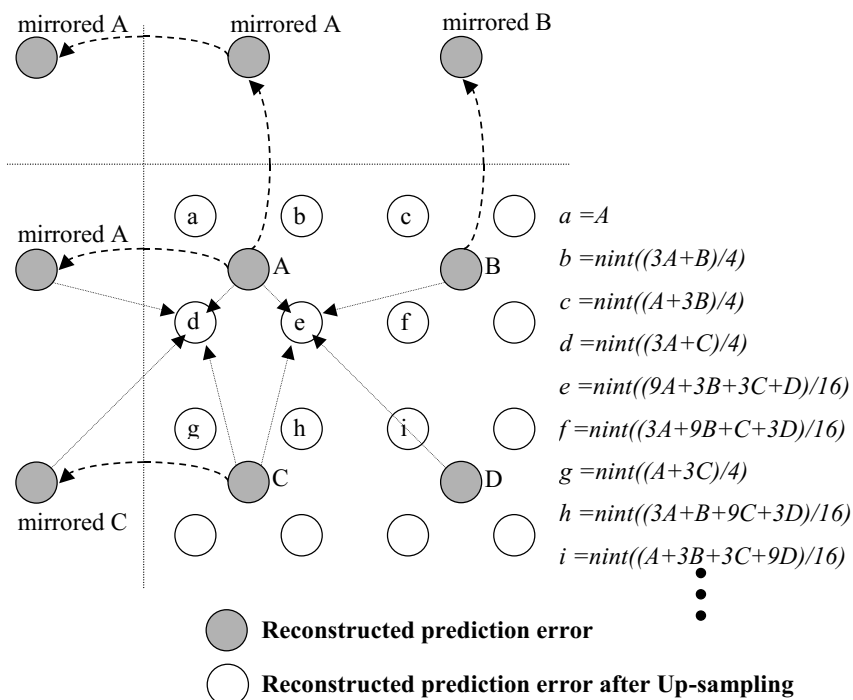


図 11 アップサンプリング [39]

2.2.6 ブロック境界フィルタ

アップサンプリングとダウンサンプリングの両方の処理はブロック内で閉じられているため、復号画像のブロック境界にブロックノイズが生じる場合がある。この劣化を除去するために、復号画像のブロック境界にフィルタを適用する。

図 12 は、ブロック境界フィルタの適用画素を示したものである。この図のようにブロック境界に接する各画素にフィルタを適用する。A および B は水平境界を介して接する画素、C および D は垂直境界を介して接する画素を示す。境界を接する二つのブロックのうち、少なくとも一つが、非ゼロの量子化係数を持つ非ゼロベクトルを有する符号化ブロックである場合、式(5)及び式(6)のブロック境界フィルタを適用する。次式は A および B についてのフィルタ計算方法である。C、D についても同様である。

$$A = \text{nint}((3A + B) / 4) \tag{5}$$

$$B = \text{nint}((A + 3B) / 4) \tag{6}$$

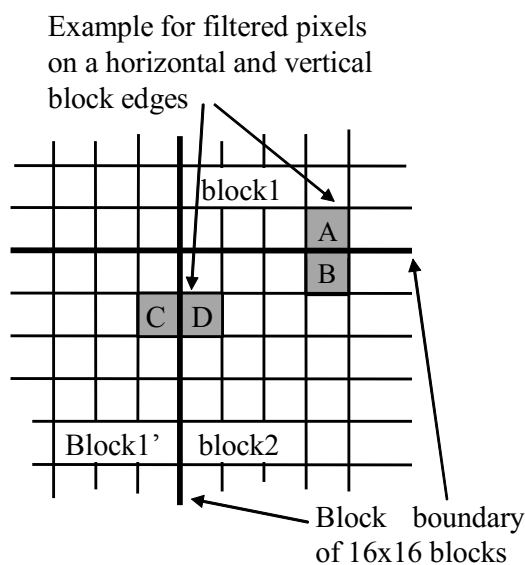


図 12 ブロック境界フィルタの適用画素 [39]

2.2.7 符号化に必要な演算量の評価

提案手法と従来の H.263 の演算量を比較・評価するために、乗算・加減算・シフトの各演算量を求めた。動き補償による予測画像生成、予測誤差生成、復号画像生成の各計算は、提案手法と従来の双方とも画素数に比例するため、演算量は同等である。また、動き推定に関しては、演算回数はアルゴリズムに大きく依存するため

比較は困難である。このことから、本節ではダウンサンプリング、DCT、量子化、逆量子化、逆DCT、アップサンプリングの演算回数を比較した。

表 3 提案手法のブロック単位の乗算および加減算・シフトの演算 [39]

処理	乗算	加減算・シフトの合計	備考
ダウンサンプリング	0	320	水平・垂直線形分離で演算削減 水平方向：加算 128(16×8) 垂直方向：加算 64(8×8) 丸め：加算 64、シフト 64
DCT	256	352	チェン [49]のアルゴリズムにより 8点 DCT あたり乗算 16回、加減算 22回 乗算：256(16×16) 加算：352(22×16)
量子化	64	64	係数あたり乗算・加算各 1回
逆量子化	64	64	全係数有効とし係数あたり乗算・加算各 1回と仮定
IDCT	256	352	チェンのアルゴリズムにより 8点 IDCT あたり乗算 16回、加減算 22回 乗算：256(16×16) 加算：352(22×16)
アップサンプリング	384	896	水平・垂直線形分離で演算削減 水平方向：加算 128 / 乗算 128 垂直方向：加算 256 / 乗算 256 丸め：加算 256 / シフト 256

表 3 に、提案手法のブロック単位の乗算および加減算、シフトの演算を示す。アップサンプリングとダウンサンプリングにおいては、フィルタを水平方向と垂直方向に独立してフィルタリングし演算回数が削減できることを考慮した。

表 4 提案手法と従来手法の符号化・復号化における演算量の比較(CIF) [39]

		ダウンサンプリング	DCT	量子化	逆量子化	IDCT	アップサンプリング	符号化	復号化
乗算	提案手法	0	152064	38016	38016	152064	228096	608256	418176
	従来(H.263)	0	608256	152064	152064	608256	0	1520640	760320
加減算+シフト	提案手法	190080	209088	0	38016	209088	532224	1178496	779328
	従来(H.263)	0	836352	0	152064	836352	0	1824768	988416

表 4 は、CIF サイズの 1 枚の画像を符号化/復号化する際の提案手法と従来手法の演算量の比較を示す。一般に乗算は計算コストが高いことから、演算を乗算と加減

算・シフト演算の2グループに分類した。ここで、符号化の演算量は、全ての処理の合計である。また復号化の演算量は、逆量子化、IDCT、アップサンプリングの合計となる。

提案手法では、ダウンサンプリングやアップサンプリングなどの追加機能が必要である。しかし、提案手法のブロック数は従来の H.263 の 4 分の 1 であり、したがって、DCT、量子化、逆量子化、逆 DCT の演算数も従来の H.263 の演算数の 4 分の 1 となる。結果として、符号化と復号化の両方における提案手法の演算量は、従来の H.263 よりも少なくなる。

実装上は、逆量子化、IDCT、およびアップサンプリングは、演算回数を減らすために非ゼロ係数を有する有効ブロックに対してのみ選択的に行うことが一般的である。提案手法における有効ブロックの割合は従来の方法よりも高くなることが見込まれるため、結果として、提案手法と従来の H.263 との演算量は同等程度と考えられる。

2.3 解像度変換の制御アルゴリズム

提案手法の符号化効率、解像度決定を含む符号化制御アルゴリズムに強く依存する。この章では、符号化制御アルゴリズムについて説明する。

2.3.1 解像度決定のアルゴリズム

解像度決定アルゴリズムには、実装の容易性を考慮し、以下の二つのパラメータ \overline{QP}_{i-1} と B_{i-1} に基づくシンプルな方式を用いた。

\overline{QP}_{i-1} : 前のフレームの平均量子化器

B_{i-1} : 前のフレームで使用されたビットの数

一般に \overline{QP}_{i-1} と B_{i-1} の積は、符号化フレームの複雑度の近似指標とみなすことができ、H.263 の符号化のテストモデル TMN8 [35] [36] では、この積を用いて符号化制御が行われる。提案手法では、同様にこの両者の積を用いることで、少ない演算量で解像度を決定する。

まず前のフレームが通常モードで符号化されている状態で、 \overline{QP}_{i-1} と B_{i-1} の積が閾値 $TH1$ より大きい場合、現フレームは縮小解像度モードに遷移する。そして、 \overline{QP}_{i-1} を定数 $C1$ で除算し、TMN8 を用いた現フレームのレート制御において、前フレームの平均量子化器の代わりに用いる。また、 \overline{QP}_{i-1} と B_{i-1} の積が $TH1$ より小さい場合には、通常モードで符号化する。

次に、前のフレームが縮小解像度モードで符号化されている状態では、 \overline{QP}_{i-1} と B_{i-1} の積が閾値 $TH2$ より小さい場合に通常モードに遷移する。次に、 \overline{QP}_{i-1} に定数 $C2$ を乗算し、TMN8 を用いた現フレームのレート制御において、前フレームの平均量子化器の代わりに用いる。

この解像度決定アルゴリズムの疑似コードは下記のとおりとなる。

- 通常モードから縮小解像度モードへ

```
if( $\overline{QP}_{i-1} \cdot B_{i-1} > TH1$ ){
    "縮小解像度モード"に遷移;
     $\overline{QP}_{i-1} = \overline{QP}_{i-1} / C1$ ;
}
```

- 縮小解像度モードから通常モードへ

```
if( $\overline{QP}_{i-1} \cdot B_{i-1} < TH2$ ){
    "通常モード"に遷移;
     $\overline{QP}_{i-1} = \overline{QP}_{i-1} \cdot C2$ ;
}
```

量子化器 $QP1$ およびフレームレート $FR1$ が、通常の解像度で符号化することができる最低の主観品質を表すとしたとき、閾値 $TH1$ は次の式で決定される。

$$TH1 = QP1 \cdot (Target_Bitrate / FR1)$$

同様に、 $TH2$ は以下の式で決定される。ここで、 $QP2$ と $FR2$ は、縮小解像度モードで符号化する上限の品質を表す。

$$TH2 = QP2 \cdot (Target_Bitrate / FR2)$$

効果的な解像度の切り替えを実現するために、いくつかのパラメータ値の候補を用いてテストし、符号化されたピクチャを主観的評価した。その結果、各パラメータを $QP1 = 16$ 、 $FR1 = 6$ 、 $C1 = 2.5$ 、 $QP2 = 7$ 、 $FR2 = 8$ 、 $C2 = 2.5$ と決定した。

2.3.2 低解像度から高解像度遷移時の DCT 係数の制限

縮小解像度モードが選択されると、動いた領域の詳細テクスチャが失われる可能性がある。そして、通常モードに遷移すると、失われたテクスチャを再現するために、一時的に多量の情報量が必要となる。この符号化情報量の突然の増加は、しばしば意図しないフレームのスキップを引き起こす。

さらに、前節の解像度決定アルゴリズムは平均 QP と平均情報量との積を用いるため、情報量の突然の増加は、しばしば縮小解像度モードへの意図しない遷移を引き起こし、低減解像度モードと通常モードとの間の振動が生じてしまう。

この振動を抑制するために、縮小解像度モードから通常モードに切り替えた後の数フレームにおいて、符号化する DCT 係数の制限を行う。切替え直後の最初の通常モードのフレームでは、4×4 の低周波数内の係数のみ VLC 符号化する係数を制限する。次に、第 2 のフレームは 5×5 の係数のみ、第 3 のフレームは 6×6、第 4 のフレームは 7×7 の低周波係数のみを符号化する。この制限により、意図しないフレームスキップおよび解像度の振動を効果的に抑制することができる。

2.4 符号化性能の評価

2.4.1 シミュレーション

テレビ電話等の低遅延条件下で提案方式の有効性を評価するためにシミュレーションを行った。

従来の H.263 の場合、非常に動きの大きい、あるいは変化の大きいシーンの場合、標準化で規定された最大の量子化器、すなわち 31 がピクチャ内のすべてのブロックに割り当てられ、フレーム当たりのビット数を減らすことが不可能となる。結果としてフレーム内で多くのビットが消費され、バッファ占有量を適度に保つためにフレームスキップが発生する。

このことから、シミュレーションには、特に動きの大きいシーケンスを使用した。シミュレーションで用いた 4 種類のシーケンス、特徴、目標情報量、目標フレームレート、そして、目標情報量を目標フレームレートで除算したフレーム毎目標情報量を表 5 に示す。これらには、動きや変化の大きい人物の前景と、テクスチャのある背景が含まれる。これらのシーケンスは、すべて H.263 ver.2 標準化手順で使用されたものである。この中で **Bowing** と **Pamphlet** は、動きの激しいシーンに向けた評価のために、筆者が ITU-T に提供した。

また、提案手法と従来手法とも、フレームレートと目標情報量の制御には、TMN8 [35] [36] の制御アルゴリズムを用いた。さらに提案手法では、モード切り替え時に 2.3.1 節の平均量子化器の変換の修正を加えた。

表 5 符号化性能評価に用いたシーケンスとシミュレーション条件 [39]

シーケンス名	特徴	Bit rate (kbps)	Frame rate (fps)	フレーム毎目標情報量	Frame Size
Bowing	人物が画面内に入り、お辞儀をして去っていく	48kbps	10fps	4800bits	CIF
Pamphlet	人物が画面内に入り、パンフレットを持ち上げる。	48kbps	10fps	4800bits	CIF
Mother & Daughter	母親と娘の動きが激しい上半身画像	32kbps	10fps	3200bits	CIF
Silent Voice	手話通訳の上半身画像	32kbps	10fps	3200bits	CIF

2.4.2 評価結果

図 13 ~ 図 16 に提案方式(DRC)と従来方式において、各シーケンスのフレーム毎の情報量と提案手法の符号化モード(左図)、および PSNR(右図)の結果を示す。また、表 6 に各シミュレーション条件の提案方式と従来方式のフレームあたり平均情報量と平均 PSNR の比較を示す。

まず、図 13 ~ 図 16 より、従来の H.263 では各フレームの発生情報量を目標情報量に制御できていないことがわかる。特に図 13 の 50-100 フレームや図 14 の 70-100 フレーム等では、フレームの目標情報量 4800 ビットの 3 倍以上の情報量が消費されている。これらのシーンでは、フレームスキップの頻発により表示遅延も大きくなり、主観品質も極端に悪化する。一方、提案手法(DRC)では、適切に縮小解像度モードが選択された結果、各フレームの発生情報量が目標情報量に収まり、フレームスキップが抑圧されている。

また、動きの大きいシーンの PSNR に関しては、フレームあたり情報量が 1/2~1/3 程度となるにも関わらず、提案手法と従来の H.263 との間の差が小さい。特に **Bowing** では、提案手法の PSNR が大きく上回っている。提案手法の DRC が従来の H.263 と比較してほぼ同等以上の PSNR を実現できる理由は、動きや変化の大きいシーンで縮小解像度モードが選択されると、フレームスキップが抑制されて符号化対象ピクチャと予測画像との間隔が短くなり、動き補償の効率が良くなるためと考える。

また、表 6 より、フレーム毎の情報量の平均値を約 16%削減しているにも関わらず、平均 PSNR は Y でほぼ同等、U,V では向上することが分かる。

このように、提案手法により、動きの激しいシーンで適応的に縮小解像度モードに切り替えることで、フレームスキップを抑制しつつ PSNR を保つことができた。

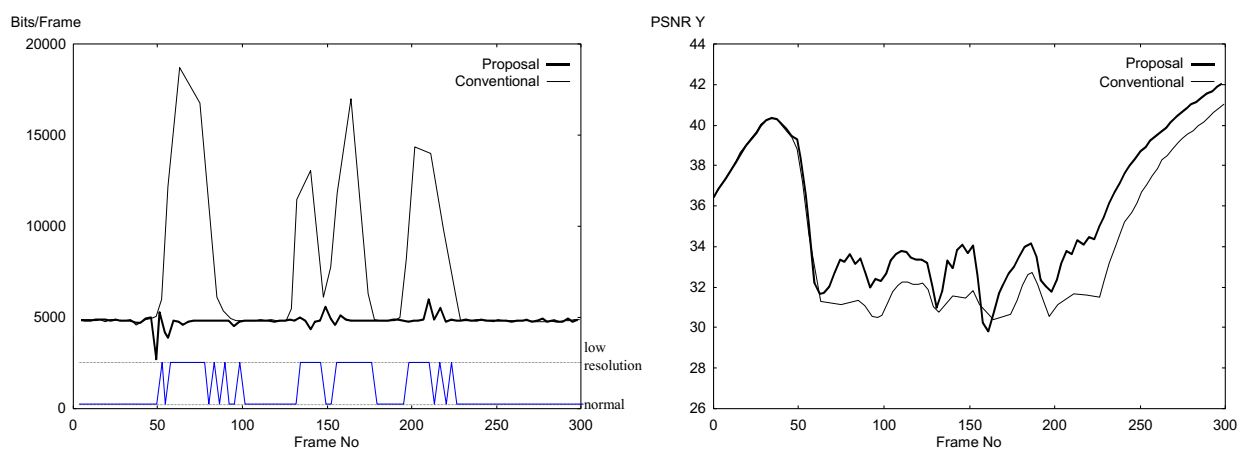


図 13 “Bowing”におけるフレーム毎の情報量、
符号化モード(左図)、PSNR (右図) [39]

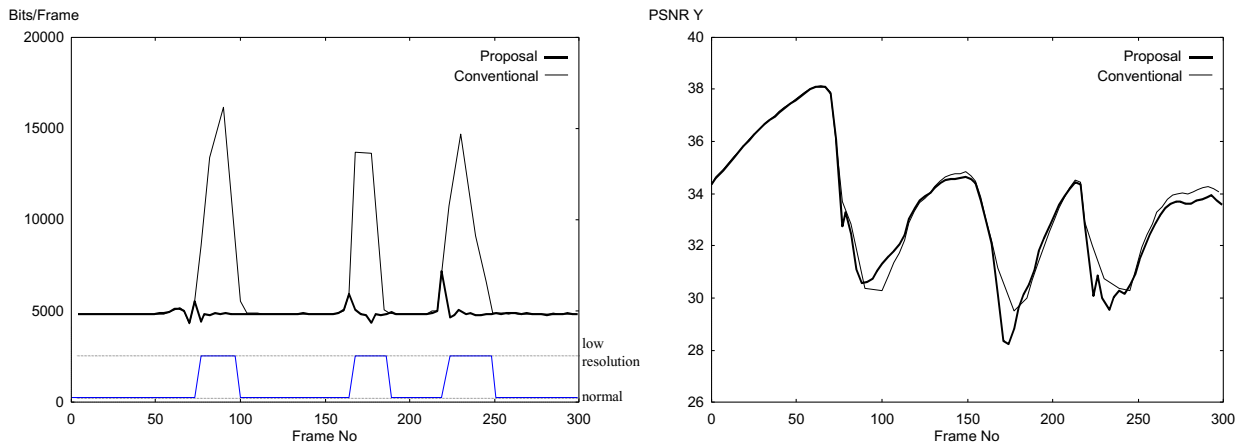


図 14 “Pamphlet” におけるフレーム毎の情報量、
符号化モード(左図)、PSNR (右図) [39]

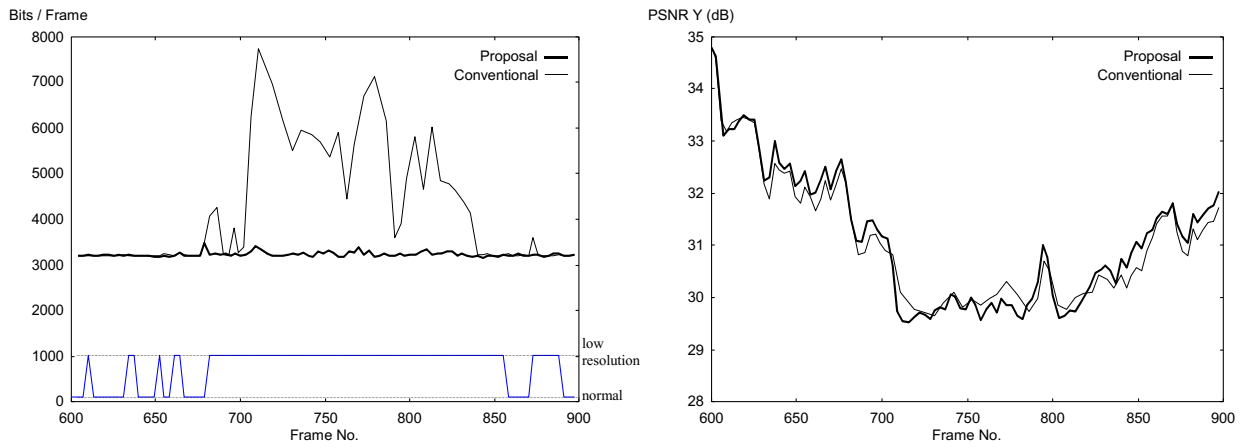


図 15 “Mother & Daughter” におけるフレーム毎の情報量、
符号化モード(左図)、PSNR (右図) [39]

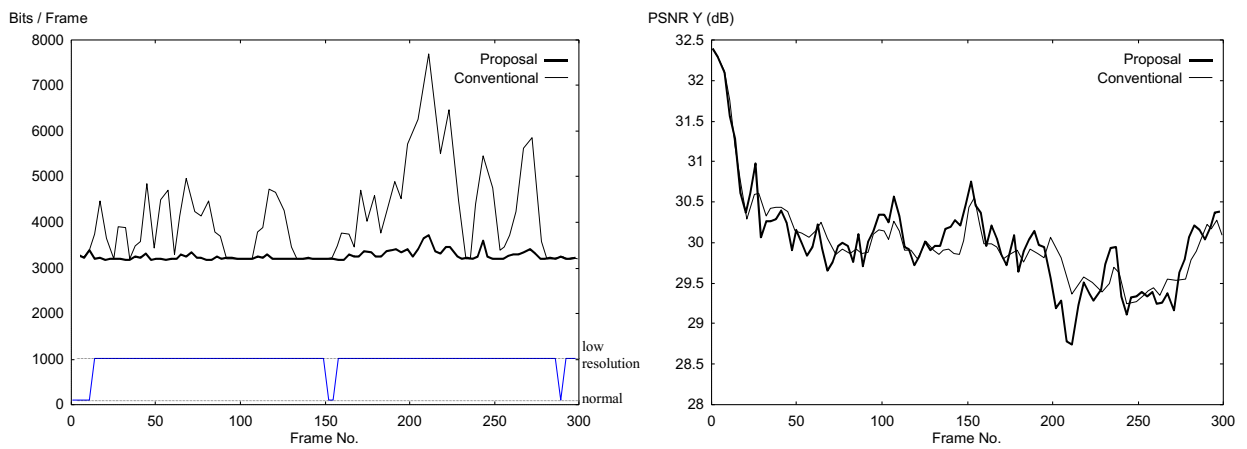


図 16 “Silent” におけるフレーム毎の情報量、
符号化モード(左図)、PSNR (右図) [39]

表 6 提案方式と従来方式のフレームあたり平均情報量と平均 PSNR の比較 [39]

Sequence	符号化方法	フレーム毎平均情報量	平均情報量削減率	PSNR Y (dB)	PSNR U (dB)	PSNR V (dB)
Bowing	DRC	4817	22.3%	35.6289	42.5889	43.0061
	従来方式	6197		35.1448	42.3159	42.5002
Pamphlet	DRC	4853	14.5%	33.4701	41.6227	42.9491
	従来方式	5674		34.1285	41.4750	42.9791
Mother& Daughter	DRC	3224	9.0%	31.4494	38.2368	37.9782
	従来方式	3542		31.3141	38.0847	37.6945
Silent Voice	DRC	3260	17.9%	29.9894	36.3303	37.7498
	従来方式	3970		30.0405	35.9438	37.1849
平均	DRC	4367	15.9%	32.6345	39.6947	40.4208
	従来方式	4521		32.6570	39.4549	40.0897

2.5 伝送誤り環境下での性能評価

2.5.1 伝送誤りが発生する通信路での映像伝送

1.3.1 節で説明したとおり、MC+DCT 動画像符号化技術の核となる動き補償予測では、原画像と、前フレームの復号画像から生成された動き補償予測画像の差分のみが伝送され、復号側では、この差分が前フレームの復号画像から生成された予測画像に加算される。このことから、無線環境等で伝送誤りが発生すると、復号画像に生じた復号エラーが、それ以降の複数の復号画像に伝搬してしまう。H.263 のテストモデル TMN8 [35]では、この復号画像のエラーの影響を低減し、早期に復帰するために、TCON モデルが定義されている。この TCON モデルでは、まず符号化時には、イントラリフレッシュ用に、フレームあたり一定数のマクロブロックを順次選択してイントラ符号化し、あるフレーム数の周期で全マクロブロックが少なくとも 1 回はイントラ符号化されるよう制御する。そして、復号時にエラーが生じた場合には、エラーがあるマクロブロックラインに対し、その上のマクロブロックラインの動きベクトルを用いて、直前の復号画像から動き補償予測を行い、マクロブロックラインの復号画像を置き換える。ここで、TCON モデルで画質復帰までに必要なフレーム数が同等と仮定した場合、フレームレートの増加に反比例して画質復帰までの時間が短くなることとなる。

以上の考察より、伝送誤りが発生する通信路で映像伝送する際に、動きの激しいシーンで DRC を用いて符号化することにより、フレームレート向上の効果によって、従来に比べて復号画像に生じたエラーから早く復帰できることが期待できる。

このことから、本節では、伝送誤りが発生する環境で、従来方式と DRC の性能評価を行った。

2.5.2 シミュレーション

伝送誤りが生じる通信路で映像伝送する際の DRC と従来方式の復号画質の比較、及びイントラフレッシュを挿入するフレーム数の最適な間隔を得るためにシミュレーションを行った。H.263 の規格では、全てのマクロブロックは少なくとも 132 フレーム毎にはイントラモードで符号化されなければならないという規定があるため、132 フレームがイントラリフレッシュ間隔の最大値となる。

シミュレーション条件を表 7 に示す。イントラリフレッシュによるエラー耐性への効果を評価するために、伝送誤りがない場合に加え、WCDMA のエラーパターンを符号化情報に印加した。

伝送誤りのない条件では、当然ながらイントラリフレッシュを本質的に使用する必要はないため、上限の 132 が最良の性能を示すことが期待される。

しかし、伝送誤りが発生する条件下では、イントラリフレッシュ周期が長くなると TCON モデルでの復号エラーからの復帰に時間がかかってしまう。一方、イントラリフレッシュ周期が短すぎると、フレーム内のイントラ符号化の領域が増えた結果、フレームあたりの情報量が大きくなってフレームレートが低下してしまうか、フレームの画質 (PSNR) が悪くしてしまう。すなわちエラーからの復帰周期と画質にはトレードオフの関係があることが想定される。

表 7 伝送誤りの画質評価のシミュレーション条件 [39]

Sequence	BER	イントラ リフレッシュ 周期	Bit rate (kbps)	Frame rate (fps)	Frame Size
Bowing	伝送誤り無し, 10^{-3} (W-CDMA)	1,5,11,16,22,44, 66,88,110,132	48kbps	10fps	CIF

2.5.2.1 伝送誤り無しの条件での評価

伝送誤りがない条件でのイントラリフレッシュ周期と画質の関係について、従来の H.263 の場合を図 17 に、提案手法 (DRC) を用いた場合を図 18 に示す。水平軸はフレーム番号、垂直軸は画質 (PSNR) を示す。予想されるように、より少ない頻度のイントラリフレッシュ間隔の方が画質が良いことが分かる。

図 19 は、DRC と従来方式において、イントラリフレッシュ周期とシーケンスの平均 PSNR との関係を示したものである。水平軸はイントラリフレッシュ周期、垂直軸は平均 PSNR を示す。明らかにわかるように、DRC は全てのイントラリフレッシュ周期にわたって、従来方式より約 1dB 高い性能を示すことが分かる。また、予想通り、両条件ともイントラリフレッシュ間隔が長くなるほど PSNR が向上する。

図 20 は、イントラリフレッシュ周期と符号化されたフレーム数を示す。本実験

で用いた Bowing シーケンスは 30fps かつ長さ 300 フレームであり、目標フレームレートが 10fps であることから、フレームスキップがない理想状態では 100 フレームが符号化されることとなる。DRC では、目標どおり 100 フレームに近いフレームが符号化されているが、従来方式で 80 フレーム程度しか符号化できず、フレームスキップが発生している。DRC は従来方式より符号化フレーム数が約 18% 多い。

これらのことから、イントラフレッシュを用いた環境でも PSNR、フレームレートの両面で DRC は従来方式より主観画質がよいことが再確認できる。

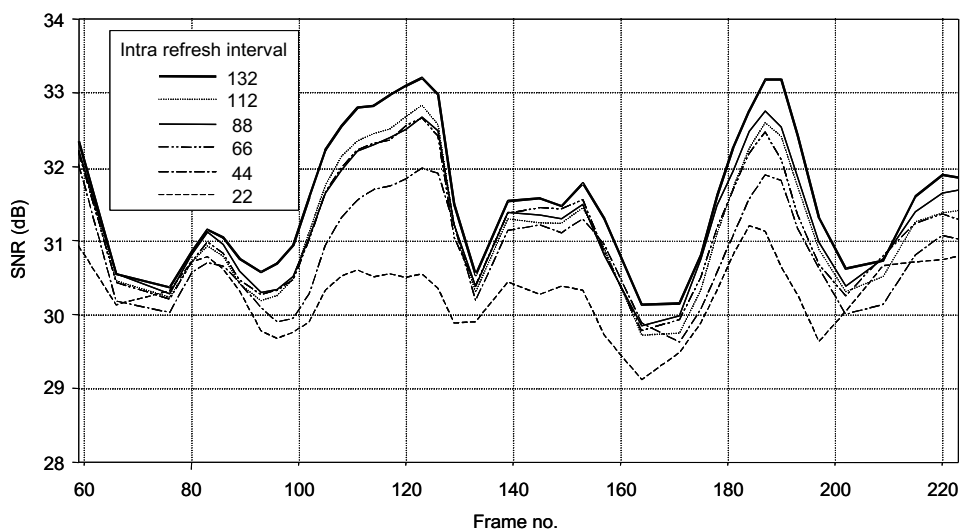


図 17 従来方式で伝送誤りがない場合の画質とイントラリフレッシュ周期の関係 [39]

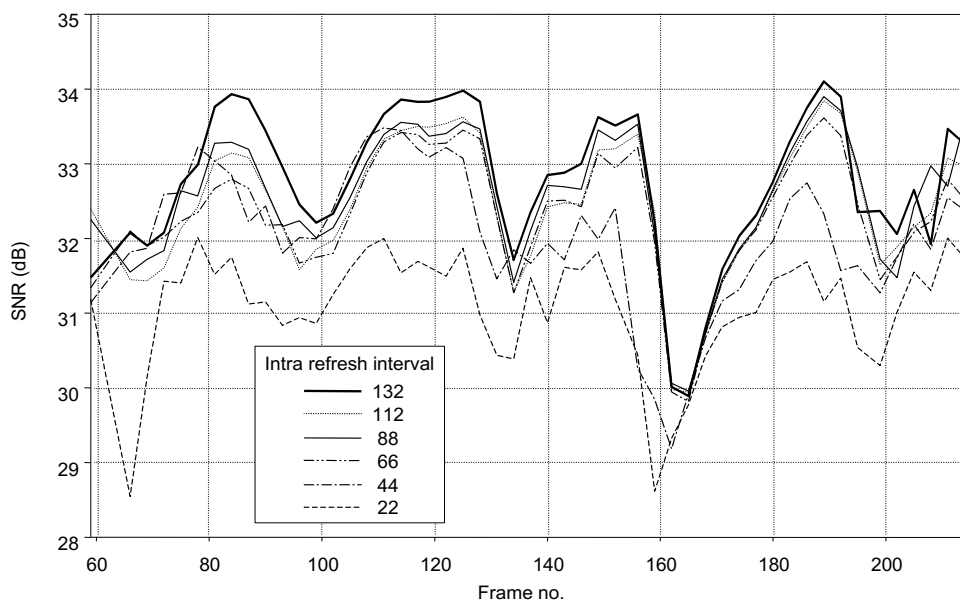


図 18 DRC 方式で伝送誤りがない場合の画質とイントラリフレッシュ周期の関係

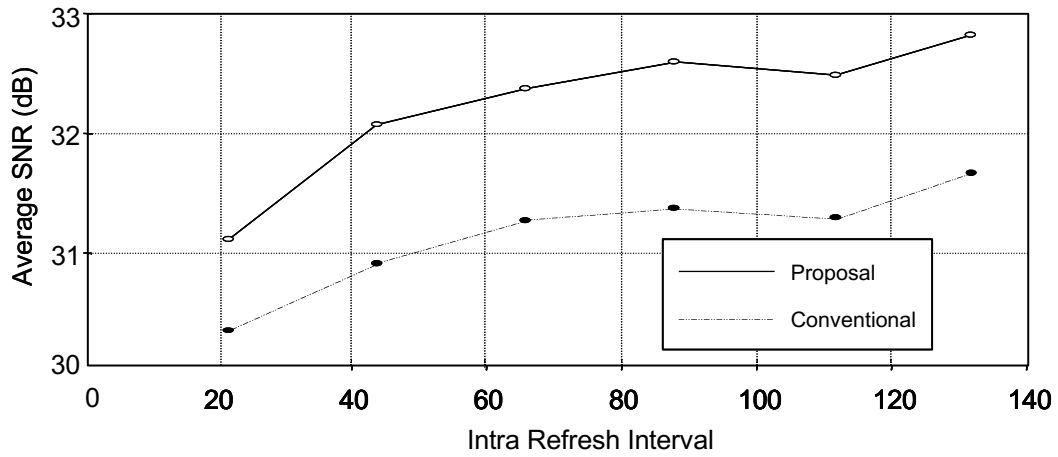


図 19 イントラリフレッシュ周期と平均 PNSR との関係 [39]

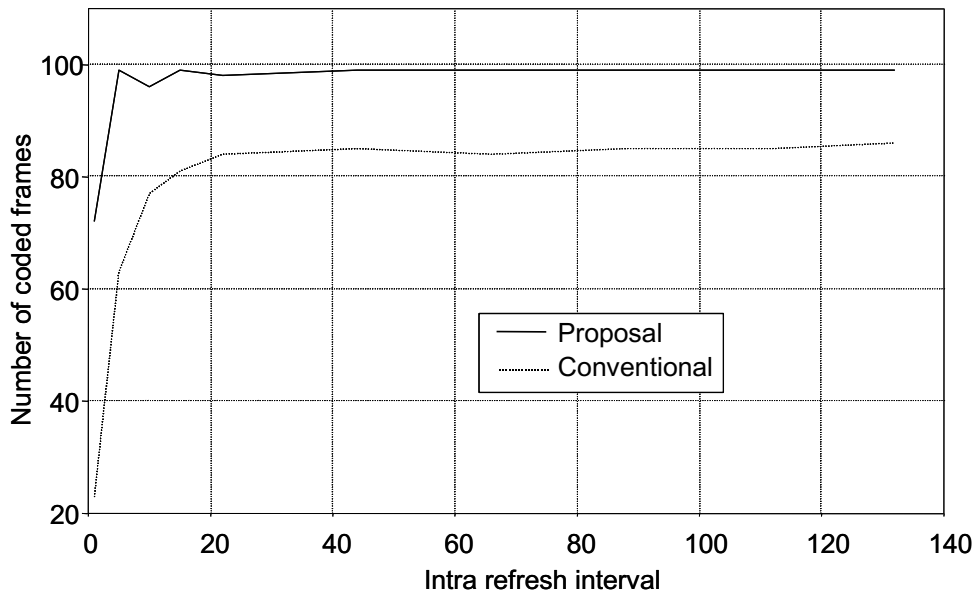


図 20 イントラリフレッシュ周期と符号化されたフレーム数 [39]

2.5.2.2 伝送誤り有りの条件での評価

FOMA 等の無線通信サービスでは、伝送エラー率は有線よりもはるかに高い。3G の W-CDMA の場合ビット誤り率 (BER) は 10^{-3} までになると言われている。このことから、W-CDMA 相当の伝送誤りが発生する場合に、下記シミュレーション手順で DRC と従来方式のイントラリフレッシュ間隔による復号画質品質の比較を行った。

- (1) 表 7 のイントラリフレッシュ間隔条件を変えながら DRC と従来方式で映像を符号化する。

- (2) 符号化データにビット誤り率 10^{-3} (W-CDMA) の誤りパターンを符号化したビットストリームに印加する。このエラーパターンを生成した無線チャネル条件は、インターリーブ期間は 10ms、畳み込み符号レートは 1/3、ドップラー周波数は 5Hz である。
- (3) 前記(2)のビットストリームを TCON モデルで復号し、PSNR を計算する。
- (4) ステップ (2) ~ (3) を 100 回繰り返す、各フレームの PSNR を平均化する。ここで、(2)で誤りパターンを毎回一定量シフトして符号化データに印加する。

図 21 は、従来方式で伝送誤りを発生させ場合の、画質とイントラリフレッシュ周期の関係を示したものである。イントラリフレッシュ周期が 132 や 66 など大きい場合にはエラーから画像が復帰できず、PSNR が低い値を示す。一方、イントラリフレッシュ周期を 22 や 44 など短くすると、効果的にエラーから復帰し、より高い PSNR を示す。

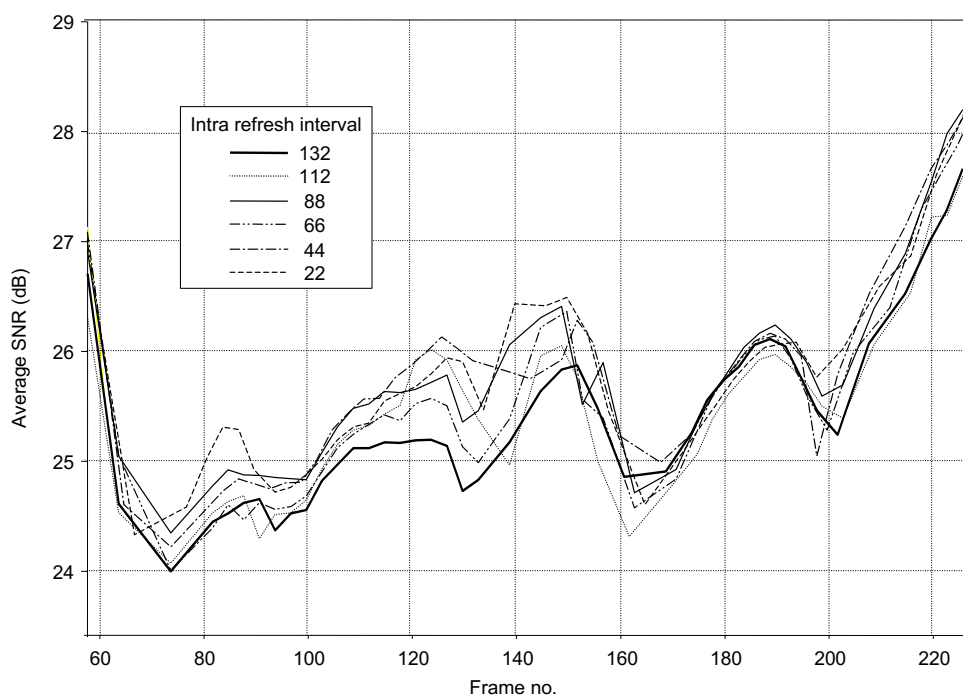


図 21 伝送誤りがある場合の従来方式で画質とイントラリフレッシュ周期の関係 [39]

図 22 は DRC を使用した際の画質とイントラリフレッシュ周期の関係を示したものである。フレーム番号 65~130 および 165~220 において、100 回の復号平均の PSNR は、イントラリフレッシュ周期 22 が及び 44 の場合に大幅に改善されている。

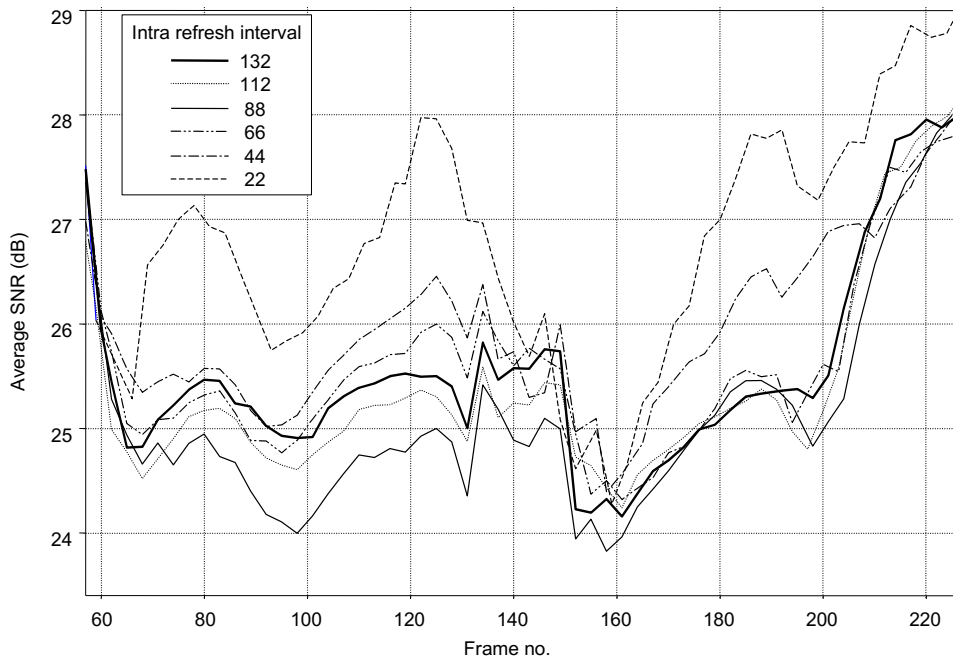


図 22 伝送誤りがある場合の DRC 方式で画質とイントラリフレッシュ周期の関係 [39]

図 23 は DRC と従来方式において、イントラリフレッシュ周期と平均 PSNR との関係を示す。DRC では、PSNR が最も高いイントラリフレッシュ周期は 22 となる。この場合、図 20 より、100 枚の映像が符号化されていることから 10fps のフレームレートが実現されていることが分かる。

一方、従来方式では、PSNR が最も高いイントラリフレッシュ周期は 11 となり、この場合、図 20 よりフレームレートも平均 8fps 未満となる。

このことから、DRC は従来方式より長いイントラリフレッシュ周期でエラー耐性の向上が可能となり、フレームレート、エラー発生時のコンシールメント後の PSNR も向上可能なことが分かった。

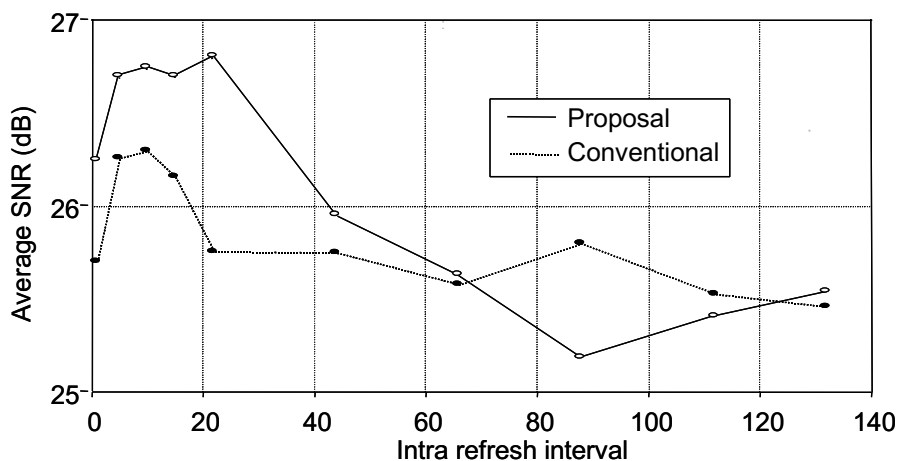


図 23 伝送誤りがある場合のイントラリフレッシュ周期と平均 PSNR との関係 [39]

2.6 本章のまとめ

静止領域のテクスチャを保存しながら、動きや変化の大きい領域を低解像度で符号化する動的解像度変換技術(DRC)を提案した。計算機シミュレーションの結果、動きや変化の大きいシーケンスは従来の H.263 よりも映像品質を高めることができた。また、提案方式と従来方式の処理量について、動き推定以外の機能について評価した結果、両者の複雑さは同程度であることも示した。

さらに、W-CDMA 等の伝送エラーが発生する条件で DRC の性能評価をおこなった。そして、TCON モデルによるエラーコンシールメントと DRC を組み合わせることで、伝送誤り環境下でも、従来方式より高いフレームレート・PSNR を実現できることを示した。

第3章 H.264 に向けた動きベクトル符号化の性能改善²

3.1 はじめに

第1章で論じた通り、2000年代半ば以降の急速な半導体等の技術進歩により、H.264/AVCでは、ある程度複雑な処理を許容した上で、徹底的に符号化効率を改善する方式が検討・採用された。

このような背景の元、H.264/AVCに向けた動きベクトル符号化の性能改善を目的として、動きベクトルのコンテキスト切り換え符号化技術 [50]と Interlace 符号化時の色差ベクトル補正技術 [51] [52] [53]の研究を行った。

本章では、動きベクトルのコンテキスト切り換え符号化技術を3.2節で、Interlace 符号化時の色差ベクトル補正技術を3.3節で、それぞれ提案方式のアルゴリズムと実験結果を説明する。

3.2 動きベクトルのコンテキスト切り換え符号化技術

3.2.1 背景と概要

H.263 [6]や MPEG-4 [7]等の動画像符号化方式では、動きベクトルを符号化する際に、周辺ブロックの動きベクトルを用いて適応的に予測ベクトルを選択し、その差分ベクトルを可変長符号化している。この可変長符号では大きな差分ベクトルには長い符号が割り当てられている。このことから、動きの激しいシーンでは、大きい差分ベクトルが発生し、発生情報量の増大をもたらしている。

この問題を回避するため、既存のアルゴリズムを改良し、適応的に動きベクトルの可変長符号を切替えて符号化する方式を開発した [50] [54]。本節では、文献 [50]に基づいて、提案方式のアルゴリズム、およびその実験結果について説明する。

3.2.2 従来方式での動きベクトル符号化方法

ここでは、まず開発した方式の基となる、従来方式の H.263 や MPEG-4 での動きベクトルの符号化方法について簡単に説明する。

図 24 は、H.263 及び MPEG-4 において、ベクトルの符号化対象のブロックと、予測ベクトルの計算に用いる3個の周辺ブロック(左: Block1/上: Block2/右上: Block3)を示している。そして、各周辺ブロックのベクトルをそれぞれ($MV1x$, $MV1y$)、($MV2x$, $MV2y$)、($MV3x$, $MV3y$)とする。

²本章は業績 [L2] 及び [L3] の内容に基づいて執筆している

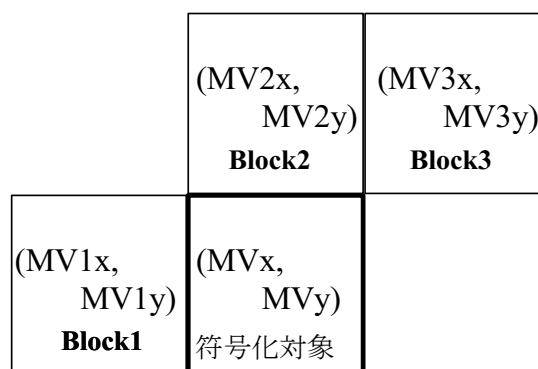


図 24 符号化対象ブロックの符号化に用いる周辺ベクトル [50]

これらの値を用い、従来の H.263 及び MPEG-4 で符号化対象のベクトル(MV_x, MV_y)を符号化する手順を以下に説明する。

まず、中央値を返す *median* 関数を用い、予測ベクトル(MVP_x, MVP_y)を以下のよう
に計算する。

$$MVP_x = \text{median}(MV1_x, MV2_x, MV3_x) \quad (7)$$

$$MVP_y = \text{median}(MV1_y, MV2_y, MV3_y) \quad (8)$$

そして、差分ベクトル(MVD_x, MVD_y)を以下の式で計算する。

$$(MVD_x, MVD_y) = (MV_x - MVP_x, MV_y - MVP_y) \quad (9)$$

最後に、上記(MVD_x, MVD_y)に可変長符号を割り当てて符号化する。

3.2.3 動きベクトルのコンテキスト切り換え符号化技術の提案

一般に、周辺ブロックのベクトル同士の動きが近い場合には予測誤差の分布は 0 付近に集中する。これに対し、周辺ブロックのベクトル同士の動きの差が大きい、すなわち、周辺ブロックのベクトルフィールドが乱雑な場合には、0 より大きい予測誤差の発生頻度が高まることを見出した。

このことから、予測誤差を符号化する際に、周辺ベクトルフィールドの乱雑さを推定し、それが一様な場合と乱雑な場合のそれぞれに適した 2 種類の可変長符号を適応的に切替えて符号化する方法を考案し、H.263 に実装した。提案方式では、従来方式である H.263 との整合性を考え、動きベクトルの予測は H.263 と同一の従来方法を用いた。可変長符号の切替えに必要な演算については、H.263/MPEG-4 等の予測ベクトルの計算途中で用いる周辺ブロックのベクトル値を利用した。提案方式では、既符号化ベクトルから乱雑さを推定するために、可変長符号の切り替えを示す情報を符号化する必要がない。

提案する可変長符号の切替えアルゴリズムは以下の通りである。

まず、可変長符号として、一様なベクトルフィールドに適したもの(以下 Non-Active 符号)と、乱雑なベクトルフィールドに適したもの(以下 Active 符号)を用意する。表 8 は Non-Active 符号と Active 符号の可変長符号化テーブル、図 25 は、両符号のベクトル差分値と符号長の関係を示したものである。Non-Active 符号は、従来の H.263 で用いられている差分ベクトルの可変長符号と同一である。Active 符号を Non-Active 符号と比べると、差分 0 の場合は 3 ビットと符号長が長い、差分が大きくなるにつれ、符号長が短くなることが分かる。

また、符号化すべき予測誤差ベクトル(MVD_x , MVD_y)は従来の H.263 と同一の方法、すなわち式(9)に基づいて求める。

次に、このベクトルの x 成分と y 成分を、それぞれ、Active 符号と Non-Active 符号のいずれで符号化するか、周辺ブロックの動きベクトルから判定する。

まず予測誤差ベクトルの x 成分である MVD_x に対する可変長符号を決定するために、周辺の 3 ブロックのベクトルを x 成分および y 成分毎にソートした値を用いる。すなわち、 x 成分、 y 成分のそれぞれの最大値を $MaxMV_x$ 、 $MaxMV_y$ 、 x 成分、 y 成分のそれぞれの最小値を $MinMV_x$ 、 $MinMV_y$ とする。また、 x 成分、 y 成分の中間値を $MidMV_x$ 、 $MidMV_y$ とする。これらの値と閾値 $TH1$ を下記の条件式に適用し、 MVD_x に対する可変長符号を判定する

```
if ((( $MaxMV_x - MidMV_x > TH1$ ) and ( $MidMV_x - MinMV_x > TH1$ )) or
    (( $MaxMV_y - MidMV_y > TH1$ ) and ( $MidMV_y - MinMV_y > TH1$ ))) {
     $MVD_x$  を Active 符号で符号化
}
else {
     $MVD_x$  を Non-Active 符号で符号化
}
(10)
```

この式では、ベクトルの x 成分、あるいは y 成分のいずれかが乱雑と判定されたときに、Active 符号で符号化することを示している。

表 8 Non-Active 符号と Active 符号の可変長符号化テーブル

差分	Non-Active符号	Active符号
-16.0	000000000101	00000001
-15.5	000000000111	00000011
-15.0	000000000101	0000011
-14.5	000000000111	1100111
-14.0	000000001001	1101111
-13.5	000000001011	1100011
-13.0	000000001101	0000111
-12.5	000000001111	1000011
-12.0	00000001001	1000001
-11.5	00000001011	1101011
-11.0	00000001101	0000101
-10.5	00000001111	1101001
-10.0	00000010001	1100101
-9.5	00000010011	1100001
-9.0	00000010101	1101101
-8.5	00000010111	0100001
-8.0	00000011001	0001001
-7.5	00000011011	0100011
-7.0	00000011101	0001011
-6.5	00000011111	0101011
-6.0	00000100001	0101001
-5.5	00000100011	101011
-5.0	0000010011	101001
-4.5	0000010101	000111
-4.0	0000010111	100011
-3.5	00000111	010111
-3.0	00001001	010011
-2.5	00001011	10111
-2.0	0000111	10011
-1.5	00011	01101
-1.0	0011	01111
-0.5	011	1111

0.0	1	001
0.5	010	1110
1.0	0010	01110
1.5	00010	01100
2.0	0000110	10010
2.5	00001010	10110
3.0	00001000	010010
3.5	00000110	010110
4.0	0000010110	100010
4.5	0000010100	000110
5.0	0000010010	101000
5.5	00000100010	101010
6.0	00000100000	0101000
6.5	00000011110	0101010
7.0	00000011100	0001010
7.5	00000011010	0100010
8.0	00000011000	0001000
8.5	00000010110	0100000
9.0	00000010100	1101100
9.5	00000010010	1100000
10.0	00000010000	1100100
10.5	00000001110	1101000
11.0	00000001100	0000100
11.5	00000001010	1101010
12.0	00000001000	1000000
12.5	000000001110	1000010
13.0	000000001100	0000110
13.5	000000001010	1100010
14.0	000000001000	1101110
14.5	000000000110	1100110
15.0	000000000100	0000010
15.5	0000000000110	00000010

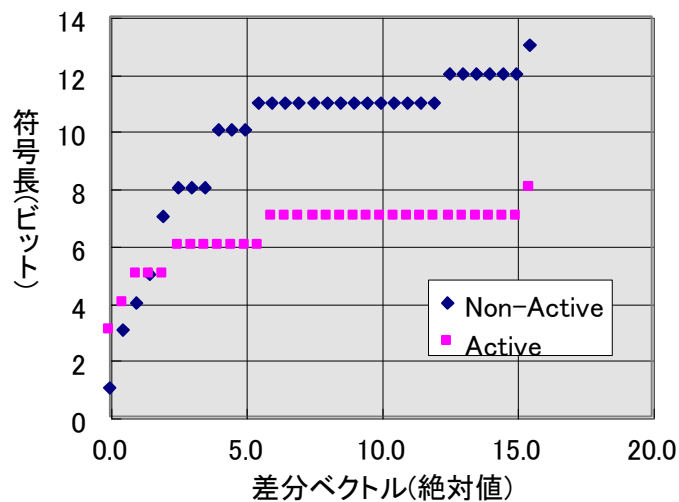


図 25 差分ベクトル値と符号長の関係 [50]

次に予測誤差ベクトルの y 成分 $MVDy$ に対する可変長符号の選択方法を説明する。式(10)で $MVDx$ の符号化に Non-Active 符号が選択されたにも関わらず $MVDx$ の絶対値が大きい時には、 y 成分の差分が大きくなる確率が高いという知見がえられた。この知見に基づいて、 $MVDy$ の可変長符号は式(11)により決定する。

```

If(( $MVDx$  を Non-Active 符号化) and ( $Abs(MVDx) \leq TH2$ )){
     $MVDy$  を Non-Active 符号で符号化
}
else{
     $MVDy$  を Active 符号で符号化
}

```

(11)

すなわち、 $MVDx$ が Non-Active 符号を選択しており、かつ $MVDx$ の絶対値が閾値 $TH2$ より小さい時のみ $MVDy$ を Non-Active 符号で、それ以外を Active 符号で符号化する。

我々の実験では、 $TH1$ は 1.0、 $TH2$ は 1.5 に設定した。

3.2.4 実験結果

3.2.3 小節で説明した方式を H.263 のテストモデル TMN [35]に適用し実験を行った。表 9 は、H.263 と提案手法で、H.263 の標準化で用いられた CIF 解像度のテストシーケンス Foreman、Silent、Mother&Dauther、Coastgurad、Deadline を符号化した場合の、1 ピクチャ当たりの動きベクトルの平均情報量を示す。両実験では、動き補償に用いた動きベクトルはすべて同一のものを用い、提案手法のみ可変長符号を切り替えた。よって、純粋に可変長符号の符号化効率の比較となっている。

本実験の結果、従来の H.263 に対し、提案手法によりベクトルの符号量を 5~16% 程度、全体平均で 9.6%削減可能となることが分かった。フレームレートが 5Hz の方が 10Hz に比べてベクトルフィールドがより乱雑になるので、改善の度合いも大きくなっている。

図 26 は、Foreman を CIF/10Hz で符号化した場合のフレーム毎のベクトルの符号量の比較を示す。このグラフより、提案手法では、H.263 に対し、特に動きが激しいシーン(170-230 フレーム)での改善度合いが大きく、可変長符号の切替えが効率よく行われていることを示している。

表 9 H.263 と提案手法による 1 ピクチャ当たりのベクトル平均符号量

シーケンス	符号化レート	H.263の ベクトル 情報量	提案方式の ベクトル 情報量	ベクトル情報 削減量 (bit)	ベクトル情報 削減比率 (%)
Foreman	15 Hz	3582 bit	3126 bit	-456 bit	-12.7%
	10 Hz	3954 bit	3397 bit	-556 bit	-14.1%
	7.5 Hz	4220 bit	3581 bit	-638 bit	-15.1%
	6 Hz	4264 bit	3594 bit	-669 bit	-15.7%
	5 Hz	4388 bit	3680 bit	-707 bit	-16.1%
Silent	15 Hz	1443 bit	1267 bit	-176 bit	-12.2%
	10 Hz	1658 bit	1443 bit	-215 bit	-13.0%
	7.5 Hz	1806 bit	1561 bit	-245 bit	-13.6%
	6 Hz	1884 bit	1628 bit	-257 bit	-13.6%
	5 Hz	2009 bit	1729 bit	-280 bit	-13.9%
Mother	15 Hz	1349 bit	1294 bit	-56 bit	-4.1%
	10 Hz	1673 bit	1577 bit	-97 bit	-5.8%
	7.5 Hz	1942 bit	1802 bit	-140 bit	-7.2%
	6 Hz	2145 bit	1971 bit	-174 bit	-8.1%
	5 Hz	2296 bit	2075 bit	-221 bit	-9.6%
Coastguard	15 Hz	3260 bit	3273 bit	14 bit	0.4%
	10 Hz	4013 bit	3833 bit	-180 bit	-4.5%
	7.5 Hz	4478 bit	4124 bit	-354 bit	-7.9%
	6 Hz	4908 bit	4387 bit	-521 bit	-10.6%
	5 Hz	5141 bit	4530 bit	-611 bit	-11.9%
Deadline	15 Hz	1007 bit	963 bit	-44 bit	-4.3%
	10 Hz	1214 bit	1151 bit	-63 bit	-5.2%
	7.5 Hz	1369 bit	1289 bit	-80 bit	-5.9%
	6 Hz	1523 bit	1419 bit	-103 bit	-6.8%
	5 Hz	1630 bit	1505 bit	-125 bit	-7.7%
平均					-9.6%

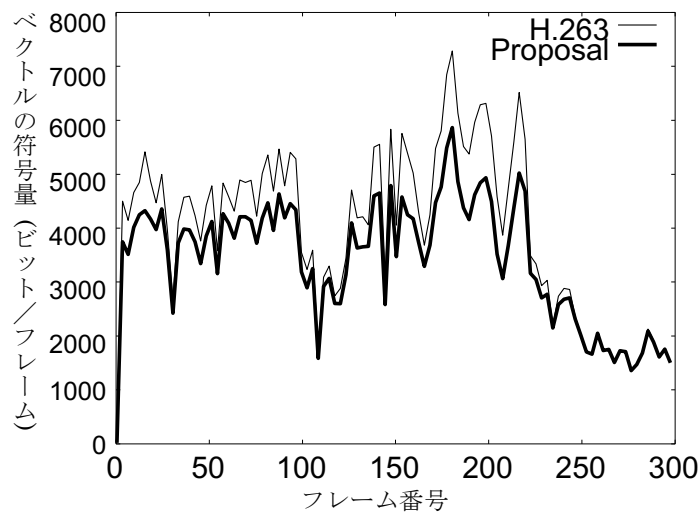


図 26 Foreman におけるフレーム毎のベクトルの情報量の比較 [50]

3.3 Interlace 符号化時の色差ベクトル補正技術

3.3.1 背景と概要

H.264/AVC [8]符号化標準は、従来の MPEG-2 [5]や MPEG-4 [7]に比べて約 2 倍近い効率で動画像を符号化することが可能である。そして、1.4.5 小節でも説明したとおり、以下の動き補償予測の精度向上が符号化効率改善に大きく貢献している。

- ・ 適応的ブロック分割

MPEG-2ではマクロブロック単位にベクトルが1本、MPEG-4では4本であった。

H.264/AVCでは、マクロブロックを4x4、4x8、8x4、8x8、8x16、16x8、16x16のいずれかに分割し、各領域にベクトルを割当ててゐる。また、色差のブロックも、輝度成分と相似の2x2、2x4、4x2、4x4、4x8、8x4、8x8に分割される。

- ・ 動きベクトルの高精度化

MPEG-2やMPEG-4では、輝度、色差とも動きベクトルは1/2画素であり、色差ベクトルは、輝度成分を1/2倍した後に丸められていた。H.264/AVCでは、輝度成分のベクトルの精度は1/4画素に向上した。また、色差ベクトルに関しては、輝度ベクトルを正確に1/2倍したものを用い、精度は1/8画素である。

H.264/AVCでは、インターレース画像へも対応している。筆者は、2002年8月時点のH.264/AVCの標準化の勧告草案[55]にて、インターレース画像をフィールドモードで符号化する際に、異なるフィールド間で動き補償予測が行われると、輝度ベクトルと色差ベクトルが平行にならないという課題を発見した。この課題により、輝度成分の予測画像と色差成分の予測画像にズレが生じ、符号化効率が劣化する。

この課題を解決するため、参照先フィールドと符号化フィールドのパリティに基づいて色差ベクトルを補正する技術を実現し、色差成分の画質が向上することを示した[51][52]。本節では、文献[51]、及び標準化寄書[53]の内容に基づいて、提案方式とその性能について説明する。

3.3.2 フィールド符号化時の問題点と提案方式

図27に、2002年8月時点のH.264の勧告草案[55]において、参照フィールドと符号化フィールドのパリティが異なると、輝度ベクトルと色差ベクトルの垂直成分が平行にならない問題を示す。(a)はボトムフィールドからトップフィールド、(b)はトップフィールドからボトムフィールドの予測で、それぞれ輝度ベクトルを(1, 0)とする。この場合、勧告草案[55]では、色差ベクトルは単純に輝度ベクトルを1/2倍したものであったので、色差ベクトルは(0, 1/2)となる。この場合、図27の(a)(b)の両図において、輝度ベクトルと色差ベクトルが平行にならないことが分かる。

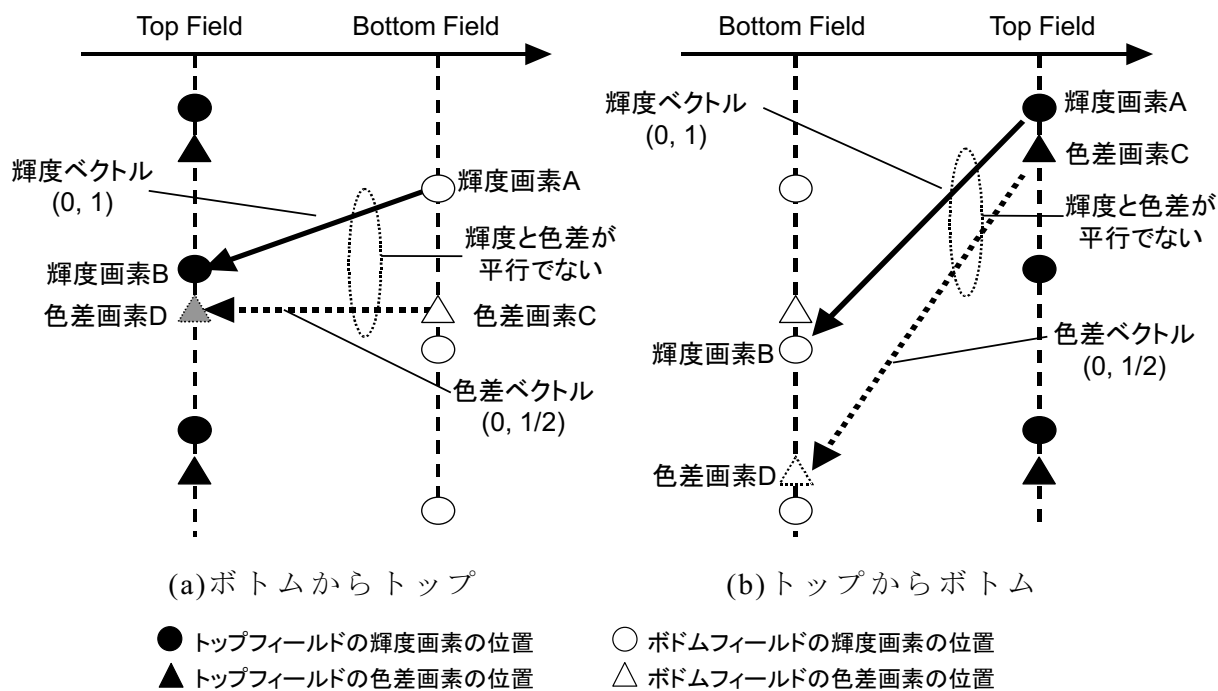


図 27 H.264/AVC の 2002 年 8 月 草案における異パリティのフィールド間の輝度ベクトルと色差ベクトルの関係

この課題を解決するため、参照フィールドと符号化フィールドのパリティに応じて、色差ベクトルに±1/4画素の補正值を加算し、常に輝度ベクトルと色差ベクトルが平行とする方法を提案する。表 10 に、提案するベクトルの補正方法を示す。符号化フィールドがボトム、参照フィールドがトップの場合には+1/4を、符号化フィールドがトップで参照フィールドがボトムの場合には-1/4を、それぞれ、色差ベクトルに加える。

表 10 提案手法における色差ベクトル垂直成分の計算方法 [53]

パリティの条件		クロマベクトルの垂直成分の計算式
参照フィールド	符号化フィールド	
トップ	ボトム	$MVC_y = Mv_y/2 + 1/4$
ボトム	トップ	$MVC_y = Mv_y/2 - 1/4$
その他		$MVC_y = MV_y/2$

図 28 は、提案技術により色差ベクトルを補正したあとの、輝度ベクトルと色差ベクトルの関係を示したものである。この図に示す通り、ボトムからトップ、およびトップからボトムの両条件で、輝度ベクトルと色差ベクトルが平行になることが分かる。

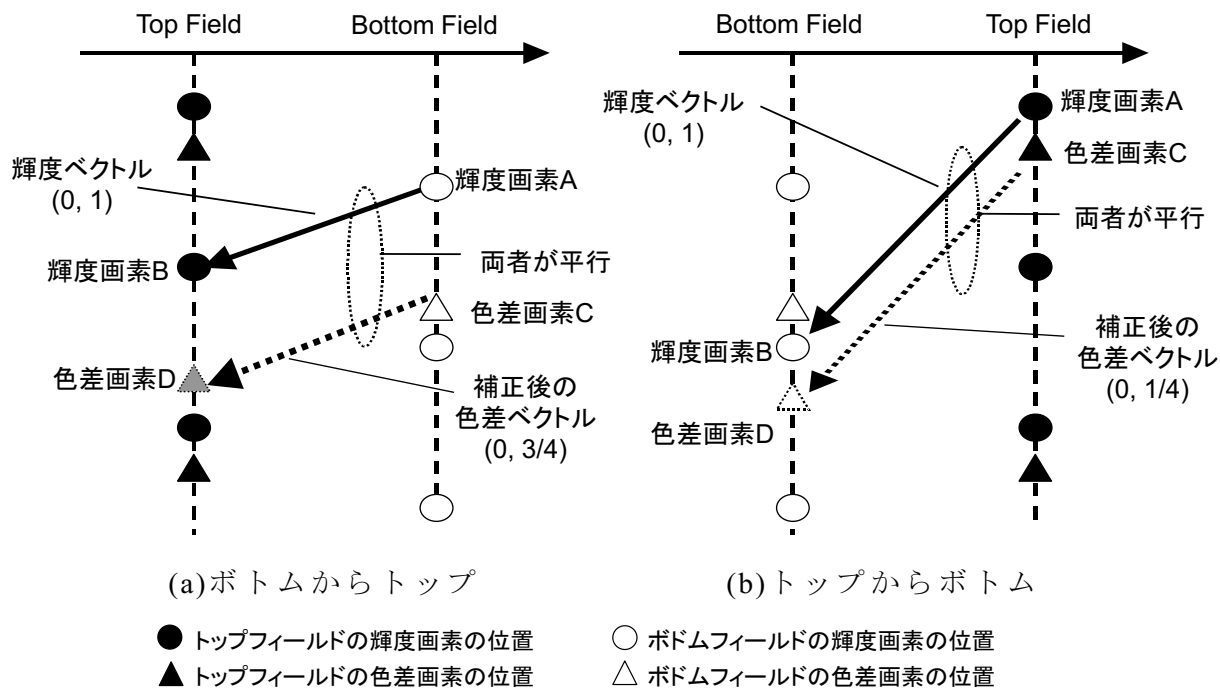


図 28 提案技術により色差ベクトルを補正後の輝度ベクトルと色差ベクトルの関係 [50]

3.3.3 実験結果

H.264/AVC の 2002 年 8 月勧告草案 [55]に対応する H.264/AVC の参照ソフト JM [56]の Ver3.9 に提案技術を実装し、両者の性能比較を行った。

表 11 実験条件 [53]

Software	JM3.9a with/without chroma vector adjustment
Coding mode	GOP structure, N=15, M=3
QP:	40, 26, 32, 28
Sequences	Mobile, Football, Canoa Valsesia, F1 car, Rugby
Reference	Set NumberRefereceFrames=1 (Previous Top and Bottom)
Entropy Coding	UVLC

表 11 に実験条件を示す。GOP 構成は、図 29 のように一般的な N=15、M=3、すなわち GOP サイズが 15 で、I/P ピクチャに B ピクチャが 2 枚挿入される構成とした。この GOP 構成の場合、GOP の最初のボトムフィールドは、その直前のトップフィールドしか予測に用いることができないため、8 月草案では全ての色差ベクトルにミスマッチが生じる。また、P ピクチャのボトムフィールドは、直前のトッ

プフィールドと、3 フレーム前のボトムフィールドしか予測に用いることができない。そのため、時間的に相関の高い直前のフィールドが予測に多く使われる。しかし、このフィールドは異パリティであり色差ベクトルに mismatches が生じてしまう。

上記を鑑みて、評価には、GOP の最初のボトムフィールド、P ピクチャのボトムフィールド、そしてシーン全体において BD-PSNR [57] で性能改善の評価を行った。

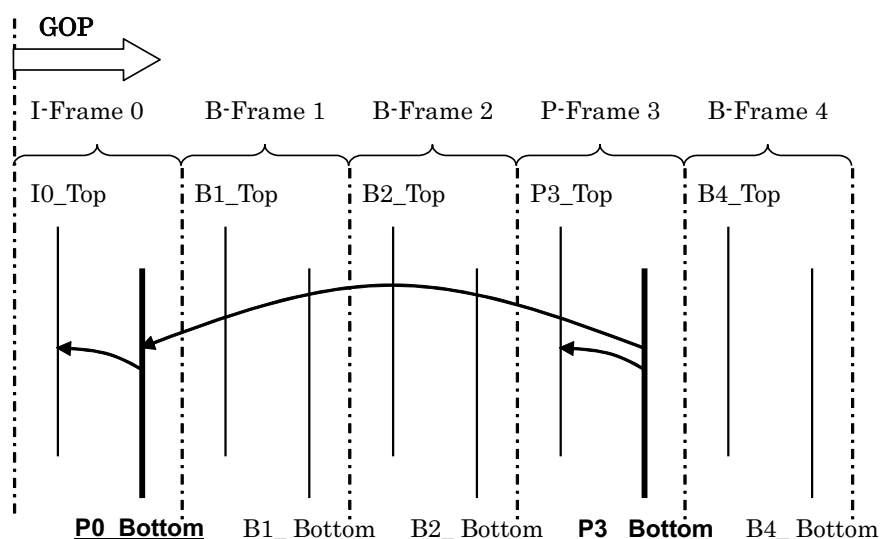


図 29 フィールド符号化時の GOP 構造と参照可能なフィールド

表 12 BD-PSNR の評価結果 [53]

シーケンス名	GOP の最初のボトムフィールド			P ピクチャのボトムフィールド			シーン全体		
	Y	U	V	Y	U	V	Y	U	V
Mobile&Calendar	0.16	0.87	1.11	0.08	0.45	0.53	0.04	0.2	0.24
Football	0.02	0.24	0.11	0.03	0.16	0.06	0.01	0.07	0.03
Canoa Valsesia	0.05	0.29	0.38	0.04	0.18	0.23	0.01	0.06	0.06
F1 car	0.02	0.26	0.45	0.01	0.18	0.28	0.01	0.09	0.12
Rugby	0.03	0.18	0.18	0.02	0.11	0.10	0.00	0.03	0.03
シーケンス平均	0.06	0.37	0.45	0.04	0.22	0.24	0.01	0.09	0.10

表 12 に各テストシーケンスの BD-PSNR を示す。正の値は提案方式が良いことを表す。ここで GOP の最初のボトムフィールドにおいて、色差 U/V の改善が大きいことが分かる。その中でも、Mobile&Calendar の U/V 画像においては、約 1dB の改善、他のシーケンスにおいても、同様に大きな改善がみられる。また、P ピクチャのボトムフィールドの U/V 画像においても平均 0.2dB 前後の改善がみられる。そして、Mobile&Calendar のシーン平均で 0.2dB 前後、全シーケンスで 0.1dB 程度の画質改善を可能とする。

図 30 に Mobile&Calendar の RD 特性を示す。全レートで大幅な改善がみられる。

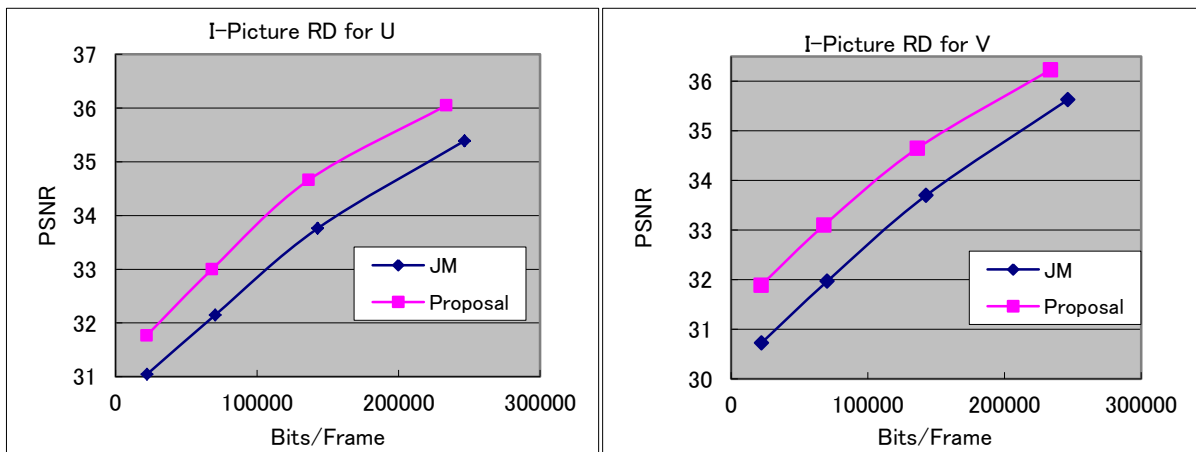


図 30 Mobile& Calendar の GOP の最初のボトムフィールドの U/V 画像の Rate と画質の関係 [50]

3.4 本章のまとめ

本章では、H.264/AVC に向けた動きベクトル符号化の性能改善を目的として、動きベクトルのコンテキスト切り換え符号化技術と Interlace 符号化時の色差ベクトル補正技術について、それぞれのアルゴリズムと評価結果について示した。

動きベクトルのコンテキスト切り換え符号化技術では、動きベクトルを符号化する際に、周辺ベクトルフィールドの乱雑さを推定し、2 種類の可変長符号を適応的に切替えて符号化する方法を提案し、ベクトルの符号量を 5~16%程度削減できることを確かめた。

Interlace 符号化時の色差ベクトル補正技術については、H.264/AVC の 2002 年 8 月勧告草案における、異パリティフィールド間予測時の輝度ベクトルと色差ベクトルを補正する方式を提案し、GOP の最初のボトムフィールドで最大約 1dB、シーケンス全体でも 0.1dB の色差の画質改善を実現した。

これらの結果、H.264/AVC の符号化効率の改善に寄与することができた。

第4章 直交変換係数の正負符号の予測符号化技術³

4.1 背景と概要

動画像符号化の国際標準では、画素の時間・空間相関を活用し、情報量圧縮が行われている。この中で、ブロック間の画素相関に着目した情報量削減の研究・実用化が、主にイントラ予測で数多く行われてきた。例えば、MPEG-4 [7]のように隣接ブロックの周波数変換係数から符号化対象ブロックの AC/DC 係数を予測する方法や、H.264/AVC [8]、H.265/HEVC [9]のように、隣接ブロックの復号済画素から符号化対象ブロックの画素を予測する方法等が実用化されている。一方、これまでの動画像符号化国際標準では、直交変換係数の正負符号は符号化情報全体の約 10%程度を占めるものの、他の情報と無相関と考えられ、圧縮されることなく 1 ビットの符号が割り当てられていた。これに対し、文献 [58]では JPEG ベースの静止画圧縮において、正負符号をブロック間の画素相関を活用して推定する方法が提案されている。しかし、本文献では正負符号を全く符号化しないため、復号時に生じた正負符号の推定誤りによる画質劣化が不可避である。また、遺伝的アルゴリズムを用いた正負符号推定のために数百回程度の画像全体の復号が必要となり、演算量が大きな課題となる。さらに、本文献ではブロック間の画素相関が高い JPEG ベースでのみで評価しており、H.264/AVC や H.265/HEVC 等の最新の動画像符号化国際標準のように時間・空間相関を可能な限り取り除いた予測残差の直交変換係数に対する適用検討・評価はされていない。

我々は、先行研究 [58]の課題であるエントロピー符号化の方式実現や演算量削減に取り組んだ結果、動画像符号化向けの正負符号の予測符号化方式 [59] [60] [61]を考案し、2010 年に HEVC の国際標準化会合で提案した [62]。本節では、文献 [60]の内容をベースに提案技術を説明する。

まず、符号化対象ブロックの M 個の有効な直交変換の量子化係数に対し正負符号を予測する場合、可能性のある 2^M 種類全ての正負符号の組み合わせに対応する量子化係数を生成する。次に、それぞれの量子化係数の逆量子化・逆直交変換の結果を予測画像に加算して復号画像候補を求め、その復号画像候補と周辺ブロックの画素との相関を総当たり法で評価する。この際、評価を Gray コード順に行い、かつ相関を漸化式で求められるよう工夫することで、単純に量子化係数を逆量子化・逆直交変換する場合に比べて、相関計算に必用な演算量の大幅な抑制が可能となる。次に、相関が最も高い正負符号の組み合わせを予測とし、予測と実際の符号との一致／不一致情報を H.264/AVC や H.265/HEVC で採用された算術符号 CABAC (context

³本章は業績 [J1]に基づいて執筆している

adaptive binary arithmetic coding) [32]で符号化する。ここで、量子化係数の絶対値や、同一ブロック内での符号予測との不一致係数の有無等を用いて CABAC 符号化時のコンテキストを設計することで、効率的なエントロピー符号化が可能となった。そして、提案手法を H.264/AVC に適用したところ、全符号化情報量の約 1%を削減することができた。これは、正負符号の情報量の約 10%の削減に相当する。

なお、次期符号化標準を検討している Joint Video Exploration Team (JVET)の 2016 年 10 月会合において、本論文と同様に正負符号の予測符号化の提案 [63]がなされたが、我々の先行技術 [59] [62]を参照せず独立に検討が進められたため、低演算量化およびコンテキスト設計の両面で我々の提案に及んでいない。しかしこれを契機に、本技術の活用が進む可能性がある。

本稿では、まず 4.2 節で提案手法のアルゴリズムを説明し、4.3 節では本提案手法による正負符号の予測精度、コンテキスト設計によるエントロピーの改善量、H.264/AVC に適用した際の情報量削減量の評価により有効性を示す。そして最後に 4.4 節で結論をまとめる。

4.2 直交変換係数の正負符号予測符号化方法

本章では、4.2.1 小節で直交変換係数の正負符号予測の基本原理を説明し、4.2.2 小節で予測の高速化手法を提案する。そして、4.2.3 小節では正負符号予測を CABAC 符号化する際の効率的なコンテキスト設計を示す。

4.2.1 直交変換係数の正負符号化予測の原理

正負符号予測の相関計算は、予測画像の取り扱いを除けば基本的には先行文献 [58]と同じ考え方であり、符号化対象ブロックと隣接ブロックの画素領域での接続性を評価する。図 31 は、直交変換基底のサイズを $N \times N$ としたとき、 $N=4$ の場合に非ゼロの量子化係数の正負符号の予測方法を示したものである。まず、隣接ブロックの境界画素間は相関が高いと仮定し、隣接ブロックの局所復号画像の境界画素から、外挿により符号化対象ブロックの上隣接、左隣接の境界画素の推定値を生成する。そして予測画像との差分を取り、境界画素の予測誤差の推定値を生成する。次に、正負符号を予測する量子化係数の個数を M とした場合、可能性がある 2^M 種類の全ての正負符号パターンの量子化係数候補を生成し、逆量子化・逆直交変換を行って予測誤差候補を求める。そして、この予測誤差画像の推定値と前述の 2^M 種類の予測誤差候補との二乗誤差が最小の正負符号パターンを正負符号の予測値とする。ここで、二乗誤差の計算回数は M の指数関数で増大するため、演算量を考慮し、 M には $N=4$ の場合には 4、 $N=8$ の場合には 6 の上限を設けた。なお上限を超えた個数の係数が存在する場合は、これらを逆量子化・逆直交変換して境界画素の推定値か

ら減じる。この制限を設けても、後述の通り平均約 8 割の係数が予測対象となる。各処理ステップの詳細を次に説明する。

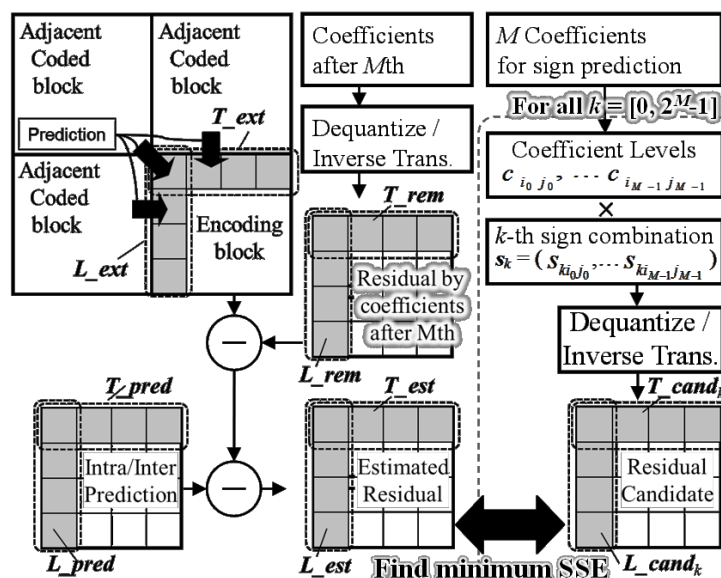


図 31 対象ブロックの直交変換係数の正負符号の予測方法 [60]

(a) 符号化対象ブロック境界の予測差分の推定値

図 32 は、直交変換基底サイズ $N=4$ の場合に隣接する符号化済ブロックの復号画像から、符号化対象ブロックの境界画像の外挿予測方法を示したものである。

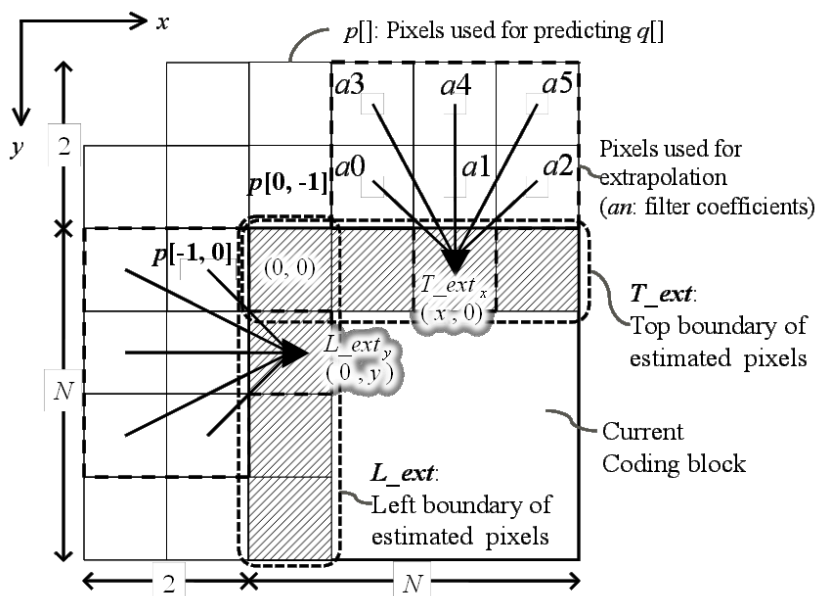


図 32 対象ブロックの境界画素の外挿予測方法 [60]

隣接する符号化済ブロックの画素から符号化対象ブロックへ外挿予測した上境界 N 画素をベクトル T_ext 、左境界 N 画素をベクトル L_ext とする。また、符号化対象ブロックの左上の画素の座標を $(0, 0)$ とし、符号化対象ブロックの上、左上、左に隣接する符号化済ブロックの復号画像の画素を $p[x,y]$ とする。

そして、座標 $(x,0)$ の T_ext の x 番目の成分を T_ext_x 、座標 $(0,y)$ の L_ext の y 番目の成分を L_ext_y としたとき、それぞれ、 $p[x,y]$ の線形結合で以下のように計算する。

$$T_ext_x = a0 \cdot p[x-1,-1] + a1 \cdot p[x,-1] + a2 \cdot p[x+1,-1] + a3 \cdot p[x-1,-2] + a4 \cdot p[x,-2] + a5 \cdot p[x+1,-2] \quad (12)$$

$$L_ext_y = a0 \cdot p[-1,y-1] + a1 \cdot p[-1,y] + a2 \cdot p[-1,y+1] + a3 \cdot p[-2,y-1] + a4 \cdot p[-2,y] + a5 \cdot p[-2,y+1] \quad (13)$$

ここで $a0$ から $a5$ までの係数は、予測対象画素に隣接する符号化済画素のエッジ特徴量により 25 種類のフィルタ係数セットの中から適応的に選択することとした。なお、画素の外挿予測は、これまで研究・提案された様々な方法(例えば [1]等)が適用可能である。我々が採用した画素予測の詳細は Appendix にて説明する。

また、正負符号予測する量子化係数の個数 M は $N=4$ において 4、 $N=8$ において 6 に制限している。ここで、この制限を超えた係数については、これらを逆量子化・逆直交変換し、得られた $N \times N$ 画素の演算結果の上境界 N 画素をベクトル T_rem 、左境界 N 画素をベクトル L_rem とする。係数の個数が制限を超えない場合には、 T_rem 、 L_rem とともにゼロベクトルとする。また、後述の通り、制限を超えた係数の正負符号は予測を使わず 1 ビットの固定長で符号化する。

次に、符号化対象ブロックの予測画像の上境界 N 画素をベクトル T_pred 、左境界 N 画素をベクトル L_pred とする。ここで、対象ブロックの上境界 N 画素の予測誤差の推定値をベクトル T_est 、左境界 N 画素をベクトル L_est とし、これらを次式で導出する。

$$T_est = T_ext - T_rem - T_pred \quad (14)$$

$$L_est = L_ext - L_rem - L_pred \quad (15)$$

(b) 正負符号パターンの全候補に対する予測誤差生成

まず、 2^M 種類の全ての正負符号パターンのベクトル s_k ($k=[0, 2^M-1]$) を生成する。

$$s_k = (s_{k_{i_0 j_0}}, s_{k_{i_1 j_1}}, \dots, s_{k_{i_{M-1} j_{M-1}}}) \quad s_{k_{i_m j_m}} \in \{1, -1\} \quad (16)$$

ここで、 i_m 及び j_m は m 番目 ($m=[0, M-1]$) の有効係数の水平、垂直の周波数 ($i_m, j_m = [0, N-1]$) を表す。

そして、それぞれの係数のレベル(絶対値)を

$$c_{i_0 j_0}, c_{i_1 j_1}, \dots, c_{i_{M-1} j_{M-1}} \quad (17)$$

とする。そして、係数毎にそれぞれのレベル(絶対値)と符号を乗算し、量子化係数の候補

$$s_{k i_0 j_0} \cdot c_{i_0 j_0}, s_{k i_1 j_1} \cdot c_{i_1 j_1}, \dots, s_{k i_{M-1} j_{M-1}} \cdot c_{i_{M-1} j_{M-1}} \quad (18)$$

を生成する。

そして、この量子化係数候補を逆量子化、逆直交変換し、その結果のブロックの上境界 N 画素ベクトルを T_cand_k 、左境界 N 画素をベクトル L_cand_k として s_k に対する予測誤差候補を求める。

(c) 直交変換係数の正負符号の予測

直交変換係数の正負符号の予測は、 2^M 種類の全ての s_k ($k=[0, 2^M-1]$) について、 T_est 、 L_est と T_cand_k 、 L_cand_k との二乗誤差 SSD_k を評価する。

$$SSD_k = \|T_cand_k - T_est\|^2 + \|L_cand_k - L_est\|^2 \quad (19)$$

そして、その中で SSD_k を最小化する正負符号パターン s_E を直交変換係数の正負符号の予測とする。

$$k_min = \operatorname{argmin}_k (SSD_k) \quad (20)$$

$$s_E = s_{k_min} \quad (21)$$

4.2.2 正負符号の高速予測アルゴリズム

4.2.1 小節で説明した正負符号の予測方法では、逆直交変換と式(19) の $2N$ 回の画素の二乗誤差の累積加算をそれぞれ 2^M 回実行することが必要であり演算量が非常に多い。この演算量を削減するため、 M ビットの 2 進数表記を、連続する二つのコードが 1 ビットしか異ならないように並べかえた Gray コード [64] を活用する。すなわち、 M ビット Gray コードの 0 を正、1 を負の符号として符号 s_k を並べ替える。表 13 は $M=3$ の時に、Gray コードと各 s_k を示したものである。表中の下線を引いた箇所が直前のインデックスと異なる。このインデックス順で SSD_{k+1} と SSD_k の差分

を定式化することで、漸化式による逐次的計算が可能となり、高速化が実現できる。

表 13 Gray コードに基づく正負符号パターン評価順 [60]

Index k	Gray Code			Sign Pattern s_k		
0	0	0	0	1	1	1
1	0	0	<u>1</u>	1	1	<u>-1</u>
2	0	<u>1</u>	1	1	<u>-1</u>	-1
3	0	1	<u>0</u>	1	-1	<u>1</u>
4	<u>1</u>	1	0	<u>-1</u>	-1	1
5	1	1	<u>1</u>	-1	-1	<u>-1</u>
6	1	<u>0</u>	1	-1	<u>1</u>	-1
7	1	0	<u>0</u>	-1	1	<u>1</u>

次に、 SSD_{k+1} と SSD_k の差分の定式化を説明する。まず、 T_cand_k 、 L_cand_k は直交変換の正規直交基底ベクトルに係数等の重みを付けた線形和で表せることを利用し、式(19)の SSD_k を下記のように書き換える。ここで $\langle A, B \rangle$ はベクトルの内積を表す。

$$\begin{aligned}
 SSD_k &= \left\| \sum_{m=0}^{M-1} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{j_m} \cdot T_{i_m} - T_est \right\|^2 + \left\| \sum_{m=0}^{M-1} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{i_m} \cdot T_{j_m} - L_est \right\|^2 \\
 &= \sum_{i=0}^{N-1} \left(\sum_{m | i_m=i} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{j_m} \right)^2 + \|T_est\|^2 - 2 \cdot \sum_{i=0}^{N-1} \langle T_i, T_est \rangle \cdot \left(\sum_{m | i_m=i} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{j_m} \right) \quad (22) \\
 &\quad + \sum_{j=0}^{N-1} \left(\sum_{m | j_m=j} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{i_m} \right)^2 + \|L_est\|^2 - 2 \cdot \sum_{j=0}^{N-1} \langle T_j, L_est \rangle \cdot \left(\sum_{m | j_m=j} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{i_m} \right)
 \end{aligned}$$

ここでは、式(22)の第 1 式のノルム計算を、第 2 式では基底ベクトルの正規直交性を用いて内積展開し加算順序を入れ替えた。各変数の定義は次の通りである。

- Q_{i_j} : 周波数 (i, j) の係数に対する量子化マトリックスを考慮した量子化サイズ
- T_x : 正規直交変換の x 番目の基底ベクトル
- $t0_x$: x 番目の基底ベクトルの 0 番目の成分

次に、二乗誤差 SSD のコスト評価の高速化手法について説明する。先に説明したとおり、正負符号パターンを Gray コード順に並べ替えたため、連続する正負符号パターンの間では一つの符号しか異ならない。いま、符号パターン s_k と s_{k+1} の間で、 d 番目の係数の符号が異なるとすると、 s_{k+1} は s_k の各係数の正負符号を用いて下記のように表現できる。

$$\mathbf{s}_k = (s_{k i_0 j_0}, s_{k i_1 j_1}, \dots, s_{k i_d j_d}, \dots, s_{k i_{M-1} j_{M-1}}) \quad (23)$$

$$\mathbf{s}_{k+1} = (s_{k i_0 j_0}, s_{k i_1 j_1}, \dots, -s_{k i_d j_d}, \dots, s_{k i_{M-1} j_{M-1}}) \quad (24)$$

ここで、 \mathbf{s}_{k+1} と \mathbf{s}_k に対する二乗誤差 SSD_{k+1} と SSD_k の差分の1/4倍を QD_k とすると、この値は、式(22)、式(23)、式(24)より次式で求められる。

$$\begin{aligned} QD_k &= (SSD_{k+1} - SSD_k) / 4 \\ &= c_{i_d j_d} \cdot Q_{i_d j_d} \cdot t0_{j_d} \cdot \left(c_{i_d j_d} \cdot Q_{i_d j_d} \cdot t0_{j_d} - s_{k i_d j_d} \cdot \sum_{m|i_m=j_d} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{j_m} \right) \\ &\quad + c_{i_d j_d} \cdot Q_{i_d j_d} \cdot t0_{i_d} \cdot \left(c_{i_d j_d} \cdot Q_{i_d j_d} \cdot t0_{i_d} - s_{k i_d j_d} \cdot \sum_{m|j_m=j_d} s_{k i_m j_m} \cdot c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{i_m} \right) \\ &\quad + s_{k i_d j_d} \cdot c_{i_d j_d} \cdot Q_{i_d j_d} \cdot \left(t0_{j_d} \cdot \langle \mathbf{T}_{i_d}, \mathbf{T}_{-est} \rangle + t0_{i_d} \cdot \langle \mathbf{T}_{j_d}, \mathbf{L}_{-est} \rangle \right) \end{aligned} \quad (25)$$

次に、 SSD_k から SSD_0 だけ減じ、1/4倍した新たなコスト指標 $QRSSD_k$ を下記のように定義する。

$$QRSSD_k = (SSD_k - SSD_0) / 4 \quad (26)$$

この式変換は線形であるので、 SSD_k の最小値を与える k と $QRSSD_k$ の最小値を与える k は同一である。また定義から $QRSSD_0=0$ となり、 $QRSSD_0$ の計算が不要となる。そして、次のように、 $QRSSD_{k+1}$ は QD_k と $QRSSD_k$ から逐次的に求めることができる。

$$QRSSD_{k+1} = QRSSD_k + QD_k \quad (27)$$

上記に基づいた高速アルゴリズムを以下に示す。

①ステップ1：変数の初期化

\mathbf{s}_k ($k=[0, 2^M-1]$)に、表13のようにGrayコードに従って正負符号をセットする。また、要素数が M の3種の配列変数 $TblH[m]$ 、 $TblV[m]$ 、 $TblHV[m]$ ($m=[0, M-1]$)、及び、要素数が直交基底サイズ N である2種の配列変数 $WH[i]$ ($i=[0, N-1]$)と $WV[j]$ ($j=[0, N-1]$)を準備し、次のように初期化する。

$$TblH[m] = (c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{j_m}) \quad (28)$$

$$TblV[m] = (c_{i_m j_m} \cdot Q_{i_m j_m} \cdot t0_{i_m}) \quad (29)$$

$$TbLHV[m] = TblH[m] \cdot \langle T_{i_m}, T_est \rangle + TblV[m] \cdot \langle T_{j_m}, L_est \rangle \quad (30)$$

$$WH[i] = \sum_{m | i_m=i} TblH[m] \quad (31)$$

$$WV[j] = \sum_{m | j_m=j} TblV[m] \quad (32)$$

また、コスト $QRSSD_0$ 、インデックス k をそれぞれ 0 に初期化する。

$$QRSSD_0 = 0 \quad (33)$$

$$k = 0 \quad (34)$$

②ステップ 2 : $QRSSD_k$ から $QRSSD_{k+1}$ の逐次計算

まず、 s_k と s_{k+1} で符号が異なる係数を特定する。ここでは、 d 番目の係数の符号が異なるとする。このとき、 d 番目の有効係数の水平、垂直周波数 i_d 、 j_d を用いて、 QD_k は次のように求めることができる。

$$QD_k = TblH[d] \cdot (TblH[d] - s_{k i_d j_d} \cdot WH[i_d]) + TblV[d] \cdot (TblV[d] - s_{k i_d j_d} \cdot WV[j_d]) + s_{k i_d j_d} \cdot TblHV[d] \quad (35)$$

そして、式(27)に従って QD_k と $QRSSD_k$ から $QRSSD_{k+1}$ を求める。

さらに、配列変数 $WH[\cdot]$ の i_d 番目、 $WV[\cdot]$ の j_d 番目の要素を次式にしたがって更新する。

$$WH[i_d] \leftarrow WH[i_d] - 2 \cdot s_{k i_d j_d} \cdot TblH[d] \quad (36)$$

$$WV[j_d] \leftarrow WV[j_d] - 2 \cdot s_{k i_d j_d} \cdot TblV[d] \quad (37)$$

このステップ 2 を、すべてのインデックスのコスト $QRSSD_k$ ($k=[0, 2^M-1]$) を計算するまで繰り返す。

③ステップ 3 : 正負符号パターン予測の導出

$k=[0, 2^M-1]$ の中で $QRSSD_k$ の最小値を与える符号パターン s_E を対象ブロックの正負符号の予測値とする。

ステップ 2 で一つの正負符号パターンに対するコスト計算は、直交変換の基底サイズや有効係数の数に依存せず、乗算 2 回、2 の冪乗倍(シフト)2 回、加減算 7 回で行うことができる。なお、5 回の正負符号乗算は、実装上、加減算と同時に行う

ことができる。

また、本アルゴリズムは式(30)の内積計算を除いて直交基底のサイズに依存しないため、HEVCのような $N=16$ 以上の直交基底に簡単に拡張可能である。

表 14 に、直交変換基底サイズ N が 4 及び 8 の場合に、正負符号予測の係数の個数 M 毎に、ステップ 1 から 3 までの累計の演算回数を示した。また、参考までに、係数の個数が均等に分布していると仮定した際の演算量平均、及び Chen のアルゴリズム [49]を用いた場合の 2 次元逆直交変換の演算量も示した。 M は前述のとおり $N=4$ では最大 4 個、 $N=8$ では最大 6 個に制限した。ここで、 $N=8$ で M が 6 の場合でも、提案手法の演算量は乗算 210 回、加算 543 回、シフト 126 回となり、1 回の逆直交変換とほぼ同等レベルの演算量である。一方、同条件において、4.2.1 小節の手法では 64 回、文献 [63]の場合でも 7 回の逆直交変換と、これに加え、双方とも 1024 回の画素の誤差累積加算が必要となる。よって、これらに比べ、提案手法では演算量の大幅な抑制が可能である。また、 $N=4$ の場合に演算量平均が 4×4 逆直交変換と同等程度、 $N=8$ の場合には演算量平均が 8×8 逆直交変換の半分程度となり、符号化全体に比べ実用上問題のない程度の演算量増加と考えられる。

表 14 提案手法における演算量 [60]

(a) $N=4, M=[1,4]$

係数の個数	乗算	加減算	シフト
1	12	16	2
2	26	39	6
3	44	76	14
4	70	141	30
平均	38	68	13
DCT 4×4 (Chen)	48	64	0

(b) $N=8, M=[1,6]$

係数の個数	乗算	加減算	シフト
1	16	24	2
2	34	55	6
3	56	100	14
4	86	173	30
5	132	302	62
6	210	543	126
平均	89	199.5	40
DCT 8×8 (Chen)	256	416	0

4.2.3 エントロピー符号化方法

本小節では、対象ブロックの正負符号の予測値を用い、実際の正負符号との一致／不一致情報のエントロピー符号化方法について説明する。

H.264/AVC、H.265/HEVC では、エントロピー符号化に CABAC が導入されている。CABAC においては、符号化対象シンボルを統計的性質が異なる複数のグループに分けて、それぞれに最適の確率パラメータを用いて算術符号化することで、符号化効率の最大化を実現している。しかし、従来の CABAC においても、符号情報は正負の偏りがないとみなされて、1 シンボルを 1 ビットの固定長で符号化する bypass モードが用いられていた。正負符号予測の正解確率が 0.5 より高ければ、正負符号

予測と実際の符号との一致・不一致情報を CABAC で符号化することで、従来の bypass モードより符号化効率の改善が可能となる。

予備実験として、正負符号予測と実符号との一致・不一致を解析したところ、次の知見が得られた。

まず、一つ目の知見は、係数の絶対値が大きいほど予測値の正解率が高いということである。これは、係数の絶対値が大きいほど、逆直交変換後の予測誤差への寄与度が大きいことから説明できる。

また、対象ブロックに二つ以上の係数があり、予測誤りが発生する場合には、一つの係数だけに予測誤りが生じることは少なく、同時に複数の係数に誤りが生じやすいことが分かった。これは、一つの係数の符号誤りがあると、その係数の誤りを補償するように他の係数の符号も誤る確率が高まることで説明できる。

これらの知見に基づいた正負符号のエントロピー符号化の疑似コードを図 33 に、また本疑似コードの 21 行目以降の各データの流れを図 34 に示す。

まず 1-5 行では、*CoefNum* には符号化対象ブロックの係数の個数、*MaxPredSignNum* には正負符号の予測対象の量子化係数の制限個数 ($N=4$ の場合に 4、 $N=8$ の場合に 6)、 M には *CoefNum* と *MaxPredSignNum* の最小値で求めた正負符号予測対象の量子化係数の個数をそれぞれ設定する。また配列 *Level*[\cdot]には 4.2.2 小節の各量子化係数の絶対値、配列 *Sign*[\cdot]には各量子化係数の正負符号をそれぞれ設定する。

次に、7-9 行において量子化係数を従来の H.264 の CABAC と同一の方法で符号化する。

そして、11 行で配列 *Level*[\cdot]をキーとして、配列 *Sign*[\cdot]と *Level*[\cdot]自身を降順にソートする。*Level*[\cdot]の値が同じ場合には、ジグザグスキャン順が大きい係数を先に配置する。このソートにより、正解確率が大きい正負符号から順に CABAC 符号化することとなる。

13-17 行では、*CoefNum* が *MaxPredSignNum* より大きい際に、インデックスが *MaxPredSignNum* 以降の正負符号 *Sign*[\cdot]を bypass モードで符号化する。

19 行では、2.2 節のアルゴリズムで M 個の正負符号の予測値 s_E を求め、*Sign*[\cdot]及び *Level*[\cdot]のソート順に合わせて、配列 *SignPred*[\cdot]に設定する。

21 行では、予測値の信頼度を表す状態変数 *State* ($State \in \{“Reliable”, “Unreliable”\}$) を準備し、初期値として信頼度が高い状態を示す”*Reliable*”を設定する。

次に 23-35 行では、予測値を用いて正負符号を CABAC 符号化する。まず、24-28 行では、正負符号 *Sing*[\cdot]と正負符号予測 *SignPred*[\cdot]が一致の場合には配列 *IsMatch*[\cdot]に 0 を、不一致の場合には 1 を設定する。

そして 30 行では、確率テーブルの選択基準として、状態変数 *State* と *Level*[\cdot]、そして後述のブロック種別の値を用いて *IsMatch*[\cdot]を CABAC で符号化する。

```

01: Set CoefNum to the Number of non-zero coefficients
02: Set MaxPredSignNum to the maximum number of
    predictive coding signs // IF N=4 THEN 4 ELSE 6
03: Set M to min(CoefNum, MaxPredSignNum)
04: Set array Level[0..(CoefNum - 1)] to coefficient levels
05: Set array Sign[0..(CoefNum - 1)] to coefficient signs
06:
07: FOR m = 0 to (M-1) DO
08:   CABAC Encode Level[m] // same as H.264/AVC
09: ENDFOR
10:
11: Descending Sort Sign[·] and Level[·] arrays
    using Level[·] as sorting key
12:
13: IF CoefNum is greater than MaxPredSignNum THEN
14:   FOR m = MaxPredSignNum to (CoefNum - 1) DO
15:     Bypass Encode Sign[m]
16:   ENDFOR
17: ENDIF
18:
19: Set array SignPred[0..(M-1)] to Sign Predictors
20:
21: SET State to "Reliable"
22:
23: FOR m = 0 to (M - 1) DO
24:   IF Sign[m] is equal to SignPred[m] THEN
25:     SET IsMatch[m] to 0 // True
26:   ELSE
27:     SET IsMatch[m] to 1 // False
28:   ENDIF
29:
30:   CABAC Encode IsMatch[m]
    in the context of State, Level[m] , and BlkType
31:
32:   IF IsMatch[m] is 1 THEN // IF False
33:     SET State to "Unreliable"
34:   ENDIF
35: ENDFOR

```

図 33 エントロピー符号化方法の疑似コード [60]

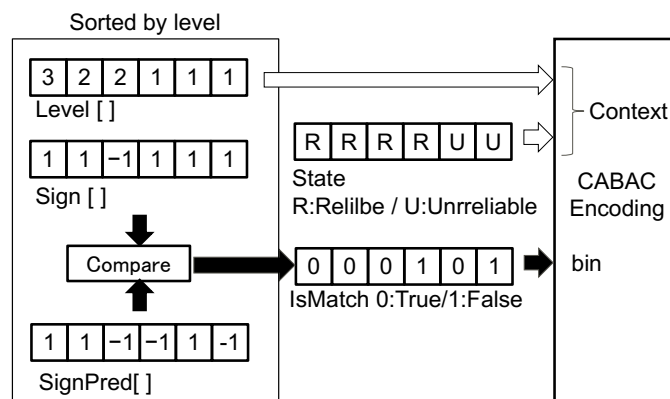


図 34 正負符号のエントロピー符号化方法 [60]

32-34 行では、符号化対象の $IsMatch[\cdot]$ が 1、すなわち不一致の場合には、状態変数 $State$ を "Unreliable" に設定する。すなわち、一つでも不一致の係数があれば、状態変数 $State$ は以降すべて "Unreliable" となる。そして、すべての予測対象の正負符号を符号化するまで 23-35 行のループを繰り返す。コンテキストの種類については、予測モードと直交変換基底サイズの組み合わせからなるブロック種別 $BlkType$ (Inter4x4 / Inter8x8 / Intra4x4 / Intra8x8) 毎に、 $Level[\cdot]$ に対して 1~9 と 10 以上の 10 種類、 $State$ に対して "Reliable" と "Unreliable" の 2 種類を組み合わせた各 20 種類、4 種のブロック種別で計 80 種類を用意した。各コンテキストの確率テーブル初期値は文献[6]を参照されたい。

提案方法では、まず大きい $Level[\cdot]$ に対応する信頼度の高い係数の $IsMatch[\cdot]$ から先に符号化し、一度でも予測と不一致の係数があると、それ以降の係数の予測精度は下がることから、信頼度が低い場合の確率テーブルに切り替えて $IsMatch[\cdot]$ を符号化する。

復号側では、まず $Level[\cdot]$ を CABAC 復号して降順にソートする。 $MaxPredSignNum$ を超過する係数があれば、その $sign[\cdot]$ を bypass 復号する。次に $Level[\cdot]$ と $State$ でコンテキストを切り替え $IsMatch[\cdot]$ を CABAC 復号し、 $State$ を更新する。最後に $IsMatch[\cdot]$ と 2.2 節の方法で求めた予測値から正負符号の復元を行う。

4.3 提案技術の評価

提案技術を H.264/AVC テストモデル JM [56] の Ver.16.2 に実装し評価を行った。

4.3.1 小節では、4.2.1 及び 4.2.2 小節で説明した正負符号の予測精度、4.3.2 小節では 4.2.3 小節で説明した CABAC のコンテキストの有効性をエントロピー評価により示す。そして、4.3.3 小節では、提案技術を H.264/AVC に適用した際の情報量の削減量を示す。

評価の条件は、HEVC の第 1 回 CfP の符号化条件 [65] の "Constraint set 1" を用いた。概要を表 15 に示す。本条件では 4K(Class A)、HD(Class B)、WVGA (Class C)、WQVGA (Class D) の 4 種類の解像度を有する 15 種類のテストシーケンスに対し、5 種類のレート (R1-R5) で符号化を行う。RaceHorses は Class C/D 双方にあるためクラスをカッコ内に示した。また、CfP ではアンカーとして H.264/AVC の参照ソフト JM16.2 を 8 ピクチャ遅延の階層 GOP で符号化することが規定されているため、提案手法もアンカーと同一条件を用いた。

表 15 テストシーケンスと符号化レート [60]

Sequence	class	Bitrate(bits/sec)				
		R1	R2	R3	R4	R5
Traffic	A	2.5 M	3.5 M	5 M	8 M	14 M
PeopleOnStreet	A	2.5 M	3.5 M	5 M	8 M	14 M
Kimono	B	1 M	1.6 M	2.5 M	4 M	6 M
ParkScene	B	1 M	1.6 M	2.5 M	4 M	6 M
Cactus	B	2 M	3 M	4.5 M	7 M	10 M
BasketballDrive	B	2 M	3 M	4.5 M	7 M	10 M
BQTerrace	B	2 M	3 M	4.5 M	7 M	10 M
BasketballDrill	C	384 k	512 k	768 k	1.2 M	2 M
BQMall	C	384 k	512 k	768 k	1.2 M	2 M
PartyScene	C	384 k	512 k	768 k	1.2 M	2 M
RaceHorses(C)	C	384 k	512 k	768 k	1.2 M	2 M
BasketballPass	D	256 k	384 k	512 k	850 k	1.5 M
BQSquare	D	256 k	384 k	512 k	850 k	1.5 M
BlowingBubble	D	256 k	384 k	512 k	850 k	1.5 M
RaceHorses(D)	D	256 k	384 k	512 k	850 k	1.5 M

4.3.1 正負符号の予測性能の評価

4.2.1 及び 4.2.2 小節で説明した正負符号の予測性能の評価結果を表 16 に示す。この表では、全ての予測対象の量子化係数に対する正負符号予測の正解率の平均を $P(True)$ とした。全シーケンス及びレートの正解率平均は 0.66、すなわち 3 シンボル中 2 シンボル程度の正解率を実現している。正解率が 0.5 を超えることから、CABAC による情報量削減が可能となる。

また前述のとおり、予測対象係数の個数 M には、直交変換基底サイズ $N=4$ では 4、 $N=8$ では 6 の上限を設けている。この制限の元で全量子化係数に対する予測対象係数の比率を表 17 に示す。低レートの R1 で約 9 割、高レートの R5 でも約 7 割、平均約 8 割の係数が予測対象となっており、演算量と予測対象係数比率のよいトレードオフだと考えている。

表 16 正負符号予測の正解率 (PTrue) [60]

Sequence	$P(True)$					
	R1	R2	R3	R4	R5	Avg.
Traffic	0.686	0.689	0.687	0.678	0.657	0.680
PeopleOnStreet	0.667	0.685	0.699	0.711	0.716	0.695
Kimono	0.701	0.715	0.729	0.737	0.736	0.724
ParkScene	0.662	0.672	0.676	0.676	0.665	0.670
Cactus	0.678	0.689	0.695	0.692	0.681	0.687
BasketballDrive	0.663	0.670	0.672	0.667	0.658	0.666
BQTerrace	0.641	0.635	0.623	0.607	0.592	0.620
BasketballDrill	0.655	0.664	0.675	0.690	0.700	0.677
BQMall	0.666	0.670	0.674	0.673	0.664	0.669
PartyScene	0.630	0.632	0.637	0.632	0.624	0.631
RaceHorses(C)	0.636	0.637	0.644	0.648	0.654	0.644
BasketballPass	0.669	0.674	0.677	0.679	0.674	0.675
BQSquare	0.603	0.580	0.561	0.531	0.504	0.556
BlowingBubble	0.642	0.640	0.635	0.623	0.596	0.627
RaceHorses(D)	0.667	0.673	0.679	0.679	0.670	0.674
Average	0.658	0.662	0.664	0.662	0.653	0.660

表 17 全量子化係数に対する予測対象係数の比率 [60]

Sequence	Ratio of coefficients using sign estimation					
	R1	R2	R3	R4	R5	Avg.
Traffic	0.891	0.851	0.811	0.764	0.714	0.806
PeopleOnStreet	0.996	0.973	0.943	0.888	0.808	0.922
Kimono	0.986	0.964	0.939	0.890	0.865	0.929
ParkScene	0.924	0.858	0.804	0.732	0.701	0.804
Cactus	0.935	0.894	0.845	0.783	0.736	0.838
BasketballDrive	0.951	0.919	0.885	0.843	0.790	0.878
BQTerrace	0.827	0.798	0.765	0.713	0.655	0.752
BasketballDrill	0.962	0.932	0.898	0.850	0.800	0.888
BQMall	0.962	0.937	0.882	0.827	0.765	0.875
PartyScene	0.934	0.899	0.842	0.786	0.716	0.835
RaceHorses(C)	0.945	0.900	0.827	0.724	0.656	0.811
BasketballPass	0.918	0.871	0.833	0.763	0.675	0.812
BQSquare	0.684	0.679	0.676	0.666	0.622	0.665
BlowingBubble	0.823	0.775	0.737	0.670	0.610	0.723
RaceHorses(D)	0.878	0.810	0.757	0.668	0.592	0.741
Average	0.908	0.871	0.830	0.771	0.714	0.819

4.3.2 CABAC のコンテキスト設計の評価

CABAC の符号化効率にはシンボルのエントロピーに漸近するため、エントロピー評価はコンテキスト設計の良い判断指標となる。このことから、符号化に先立ち、4.2.3 小節で提案したコンテキスト設計の効果を見積もるため、正負符号予測とコンテキスト分割した $IsMatch[\cdot]$ シンボルのエントロピーを評価した。

0、1を取りうるシンボルのエントロピー関数 $H(p)$ は、0の確率を p とすると、次のように定義される。

$$H(p) = -p \cdot \log_2(p) - (1-p) \cdot \log_2(1-p) \quad (38)$$

また、コンテキスト設計の得失を評価するため、情報利得を次のように定義する。まず、正負符号を本論文で説明した予測やコンテキスト分割等の情報源符号化したシンボルのエントロピーを I とする。従来の H.264 において正負符号は `bypass` モードにより 1 ビットの情報量なので、情報源符号化による情報利得は両者の差分である $(1-I)$ と定義する。

予測モードと直交変換基底サイズの組み合わせからなるブロック種別を $BlkType$ ($BlkType \in \{Inter4x4, Inter8x8, Intra4x4, Intra8x8\}$)、各予測対象係数が $BlkType$ に属する確率を $P(BlkType)$ 、各 $BlkType$ 毎に各係数の実際の正負符号と予測が一致する確率を $P(True | BlkType)$ とする。4.2.3 小節のコンテキスト分割を行わない場合、各 $BlkType$ におけるシンボルあたりの情報利得 $IG_{A|BlkType}$ 及びシーケンス平均のシンボルあたり情報利得 IG_A は次式で導出できる

$$IG_{A|BlkType} = 1 - H(P(True|BlkType)) \quad (39)$$

$$IG_A = \sum_{BlkType} P(BlkType) \cdot IG_{A|BlkType} \quad (40)$$

次に、4.2.3 小節のコンテキスト分割に基づく情報利得を求める。まず、各 $BlkType$ において状態変数 $State$ ($State \in \{“Reliable”, “Unreliable”\}$) と 10 種類の $Level$ ($Level=[1,10]$) に応じて $IsMatch[\cdot]$ を 20 種類に分割した各コンテキストの選択確率を $P(State, Level | BlkType)$ とする。また、各コンテキストにおいて実際の正負符号と予測が一致する条件付き確率を $P(True | State, Level, BlkType)$ とする。この際、各 $BlkType$ におけるシンボルあたりの情報利得 $IG_{B|BlkType}$ 及びシーケンス平均のシンボルあたり情報利得 IG_B は次式で導出できる。

$$IG_{B|BlkType} = 1 - \sum_{StateLevel} P(State, Level | BlkType) \cdot H(P(True | State, Level, BlkType)) \quad (41)$$

$$IG_B = \sum_{BlkType} P(BlkType) \cdot IG_{B|BlkType} \quad (42)$$

なお、式(39)～式(42)中の $P(BlkType)$ 、 $P(True | BlkType)$ 、 $P(State, Level | BlkType)$ 、 $P(True | State, Level, BlkType)$ の各確率、及び表 16 で示した正解率 $P(True)$ との間には、式(43)及び式(44)の関係があることに留意されたい。

$$P(\text{Ture} | \text{BlkType}) = \sum_{\text{State}} \sum_{\text{Level}} P(\text{Ture} | \text{BlkType}, \text{State}, \text{Level}) \cdot P(\text{State}, \text{Level} | \text{BlkType}) \quad (43)$$

$$P(\text{True}) = \sum_{\text{BlkType}} P(\text{Ture} | \text{BlkType}) \cdot P(\text{BlkType}) \quad (44)$$

表 18 に、“PeopleOnStreet”の各レート R1-R5 において、各 *BlkType* における $P(\text{BlkType})$ 、 $P(\text{True} | \text{BlkType})$ 、 $IG_{A|\text{BlkType}}$ 、 $IG_{B|\text{BlkType}}$ 、及びシーケンス全体の IG_A 、 IG_B を示した。各レート、及び各 *BlkType* において、2.3 節のコンテキスト分割を行った $IG_{B|\text{BlkType}}$ や IG_B が、コンテキスト分割を行わない $IG_{A|\text{BlkType}}$ や IG_A より情報利得が高いことが分かる。なお、表 18 と式(44)より、表 16 の $P(\text{True})$ が導出できる。

表 18 “PeopleOnStreet”での $P(\text{BlkType})$ 、 $P(\text{True}|\text{BlkType})$ 、 $IG_{A|\text{BlkType}}$ 、 $IG_{B|\text{BlkType}}$ 、 IG_A 、及び IG_B [60]

	Bit rate				
	R1	R2	R3	R4	R5
$P(\text{Inter}4x4)$	0.012	0.022	0.040	0.079	0.135
$P(\text{True} \text{Inter}4x4)$	0.774	0.768	0.765	0.752	0.731
$IG_{A \text{Inter}4x4}$	0.230	0.218	0.213	0.192	0.160
$IG_{B \text{Inter}4x4}$	0.297	0.292	0.292	0.257	0.208
$P(\text{Inter}8x8)$	0.261	0.277	0.309	0.336	0.362
$P(\text{True} \text{Inter}8x8)$	0.719	0.735	0.741	0.738	0.728
$IG_{A \text{Inter}8x8}$	0.143	0.166	0.174	0.171	0.156
$IG_{B \text{Inter}8x8}$	0.197	0.219	0.221	0.212	0.195
$P(\text{Intra}4x4)$	0.045	0.080	0.113	0.153	0.181
$P(\text{True} \text{Intra}4x4)$	0.619	0.639	0.642	0.654	0.673
$IG_{A \text{Intra}4x4}$	0.041	0.056	0.059	0.070	0.088
$IG_{B \text{Intra}4x4}$	0.146	0.148	0.153	0.152	0.157
$P(\text{Intra}8x8)$	0.682	0.621	0.538	0.433	0.323
$P(\text{True} \text{Intra}8x8)$	0.649	0.665	0.682	0.703	0.720
$IG_{A \text{Intra}8x8}$	0.065	0.080	0.097	0.123	0.144
$IG_{B \text{Intra}8x8}$	0.144	0.152	0.167	0.192	0.213
IG_A	0.086	0.105	0.121	0.136	0.140
IG_B	0.160	0.174	0.187	0.198	0.196

表 19 から表 23 は、表 18 の *BlkType* が *Inter8x8* の条件において、R1-R5 の各レートで各コンテキスト (*State, Level*) が選択される確率 $P(\text{State}, \text{Level} | \text{Inter}8x8)$ と、各コンテキストでの正解率 $P(\text{Ture} | \text{State}, \text{Level}, \text{Inter}8x8)$ を示したものである。なお、表中では、前者を P_{Ctx} 、後者を $P_{\text{True}|\text{Ctx}}$ と省略して記載している。 P_{Ctx} が 0、つまり該当コンテキストの係数が存在しない場合には、 $P_{\text{True}|\text{Ctx}}$ が定義できないため “na” と表記した。表 19 から表 23 の全条件で、*State* が “Reliable” の場合には *Level* が大きいほど正解率も高く、また、“Unreliable” の場合には、“Reliable” に比べて正解率が低くなり、コンテキスト分割が適切に行われていることを示している。なお、表 19 から表 23 の各値と式(39)~式(44)等により、表 18 の $P(\text{True} | \text{Inter}8x8)$ 、 $IG_{A|\text{Inter}8x8}$ 、

及び $IG_{B|Inter8x8}$ の各値を導出できる。

表 19 “PeopleOnStreet”、レート R1 での PCtx と PTrue|Ctx [60]

Level	State="Reliable"		State="Unreliable"	
	P_{ctx}	$P_{True ctx}$	P_{ctx}	$P_{True ctx}$
1	0.728	0.766	0.160	0.442
2	0.092	0.833	4.68E-03	0.277
3	0.012	0.846	2.38E-04	0.194
4	2.22E-03	0.884	0.000	na
5	3.31E-04	0.980	0.000	na
6	7.94E-05	0.833	0.000	na
7	1.98E-05	1.000	0.000	na
8	6.61E-06	1.000	0.000	na
9	0.000	na	0.000	na
>10	0.000	na	0.000	na

表 20 “PeopleOnStreet”、レート R2 での PCtx と PTrue|Ctx [60]

Level	State="Reliable"		State="Unreliable"	
	P_{ctx}	$P_{True ctx}$	P_{ctx}	$P_{True ctx}$
1	0.649	0.781	0.192	0.508
2	0.115	0.851	0.010	0.328
3	0.024	0.877	7.72E-04	0.188
4	6.38E-03	0.898	7.75E-05	0.174
5	2.00E-03	0.901	1.01E-05	0.000
6	7.14E-04	0.943	0.000	na
7	2.49E-04	0.919	0.000	na
8	7.41E-05	0.955	0.000	na
9	6.40E-05	0.947	0.000	na
>10	2.02E-05	1.000	0.000	na

表 21 “PeopleOnStreet”、レート R3 での PCtx と PTrue|Ctx [60]

Level	State="Reliable"		State="Unreliable"	
	Pctx	PTrue ctx	Pctx	PTrue ctx
1	0.587	0.775	0.207	0.553
2	0.135	0.855	0.014	0.418
3	0.035	0.886	1.62E-03	0.306
4	0.012	0.906	2.50E-04	0.203
5	4.37E-03	0.915	5.65E-05	0.233
6	1.91E-03	0.937	3.76E-06	0.500
7	8.86E-04	0.962	0.000	na
8	4.20E-04	0.973	0.000	na
9	2.07E-04	1.000	0.000	na
>10	1.96E-04	0.990	0.000	na

表 22 “PeopleOnStreet”、レート R4 での PCtx と PTrue|Ctx [60]

Level	State="Reliable"		State="Unreliable"	
	P_{ctx}	$P_{True ctx}$	P_{ctx}	$P_{True ctx}$
1	0.491	0.755	0.231	0.592
2	0.159	0.844	0.024	0.532
3	0.050	0.888	3.40E-03	0.455
4	2.02E-02	0.909	8.13E-04	0.389
5	9.38E-03	0.926	2.32E-04	0.333
6	4.83E-03	0.942	6.50E-05	0.275
7	2.57E-03	0.948	2.07E-05	0.273
8	1.46E-03	0.956	1.23E-05	0.385
9	8.20E-04	0.967	3.77E-06	0.500
>10	1.40E-03	0.975	1.89E-06	0.000

表 23 “PeopleOnStreet”、レート R5 での PCtx と PTrue|Ctx [60]

Level	State="Reliable"		State="Unreliable"	
	P_{ctx}	$P_{True ctx}$	P_{ctx}	$P_{True ctx}$
1	0.419	0.728	0.261	0.608
2	0.156	0.827	0.032	0.609
3	0.059	0.880	6.17E-03	0.592
4	0.027	0.907	1.71E-03	0.534
5	0.014	0.926	6.16E-04	0.499
6	7.87E-03	0.935	2.48E-04	0.498
7	4.78E-03	0.943	1.08E-04	0.420
8	3.03E-03	0.960	4.81E-05	0.414
9	1.96E-03	0.960	2.43E-05	0.268
>10	4.66E-03	0.975	2.12E-05	0.163

表 24 は、前述のテスト条件 [65]に基づいて、各シーケンスを 5 種類のレートで符号化し、正負符号をコンテキスト分割しない場合とした場合のそれぞれの情報利得を式(40)と式(42)に基づいて計算し、シーケンス毎に平均したものである。この結果をみると、コンテキストを分割しない場合の情報利得平均 IG_A の 0.083 ビットに対し、4.2.3 小節で提案したコンテキスト分割による情報利得平均 IG_B は 0.117 ビットとなり、約 1.4 倍の改善が得られる。これは本提案のコンテキスト設計の有効性を示しているといえる。

表 24 提案するコンテキスト分割による情報利得 [60]

Sequence	Information Gain (bit per symbol)		Improvement of Information Gain
	\overline{IG}_A	\overline{IG}_B	$\overline{IG}_B / \overline{IG}_A$ (%)
Traffic	0.099	0.140	141.0
PeopleOnStreet	0.118	0.183	155.2
Kimono	0.152	0.199	131.5
ParkScene	0.086	0.117	136.2
Cactus	0.107	0.145	134.7
BasketballDrive	0.084	0.116	138.8
BQTerrace	0.046	0.070	152.4
BasketballDrill	0.101	0.144	143.0
BQMall	0.087	0.121	139.5
PartyScene	0.053	0.075	142.4
RaceHorses(C)	0.065	0.093	142.4
BasketballPass	0.093	0.125	134.2
BQSquare	0.020	0.036	179.7
BlowingBubble	0.049	0.073	148.6
RaceHorses(D)	0.091	0.121	132.1
Average	0.083	0.117	140.5

4.3.3 符号化効率改善の評価

表 25 は提案手法による情報量削減効果を示したものである。情報量の比較には、文献 [65] に従い、評価指標として BD-Rate [57] を用いた。表に示す通り、提案方式では全体の 1.01% の情報量が削減可能となる。正負符号の情報量は全符号量の約 10% のため、本削減量は、正負符号の約 10% に相当する。また、画素の空間相関が高い高解像度画像(Class A/B)の方が低解像度画像(Class C/D)より情報削減率が高いこともわかる。

ここで、Class D の BQSquare は情報量削減量(表 25)が約 0.32% と提案手法の効果が特に小さい。BQSquare はモールのテラスをズームアウト・パニングするシーンであり、机・テーブル・床タイルなど、鋭い直線エッジを持つ物体が数多く存在している。提案手法の性能が悪い原因として、まず、 T_{ext} 及び L_{ext} の精度が不十分な可能性がある。すなわち、4.2.1 小節(a)及び Appendix で説明した外挿予測方法ではエッジを正確に予測できないため、正負符号予測の正解率平均が表 16 の 0.556 まで低下したと考えられる。また、表 17 で示した予測対象係数の割合が 0.665 と平均より 0.154 程度低くなり、これも情報量削減量が少ない一因となる。予測対象係数の割合が低い理由は、強いエッジを含む映像のズームアウトにより予測誤差に高周波が含まれる結果、ブロック当たりの係数の個数が多くなり、予測対象係数の割合が減少するためと考えられる。

逆に、Class A/Bのような高解像度画像、特に PeopleOnStreet や Kimono では、解像度に比べて複雑なテクスチャも少なく空間相関も高いために、 T_{ext} や L_{ext} の予測精度が向上し正負符号予測の正解確率も 0.7 前後まで高まると同時に、ブロック単位の有効係数も減少して予測対象係数の割合も 0.92 以上まで高まり、両者の相乗効果で提案手法の効果が高くなると考えられる。このように、提案手法は、空間相関が高い映像において、より有効性が高いと考えられる。

また計算時間の増加は、Xeon X5570 (2.93Gz)の PC において、符号化は平均 8%、復号は平均 5%であった。

表 25 提案手法による情報量削減効果 [60]

Sequence	BD Rate	Sequence	BD Rate
Traffic	-1.18%	RaceHorses(C)	-0.73%
PeopleOnStreet	-1.24%	BasketballPass	-0.99%
Kimono	-2.17%	BQSquare	-0.32%
ParkScene	-1.18%	BlowingBubble	-0.70%
Cactus	-1.36%	RaceHorses(D)	-0.91%
BasketballDrive	-0.86%	Class A Avg.	-1.21%
BQTerrace	-0.70%	Class B Avg.	-1.25%
BasketballDrill	-1.15%	Class C Avg.	-0.88%
BQMall	-0.97%	Class D Avg.	-0.73%
PartyScene	-0.67%	Overall Avg.	-1.01%

4.4 本章のまとめ

符号化対象ブロックの画素と周辺の符号化済の画素との相関から映像符号化の直交変換係数の正負符号を予測する方式を開発し、H.264/AVC に適用することで、全体の情報量の約 1%を削減することが可能となった。これは正負符号の情報量の約 10%に相当する。

現在、ITU-T と ISO/IEC が共同で設立した Joint Video Exploration Team(JVET)で検討中の次期符号化標準モデル JEM では、H.265/HEVC をベースに 1%程度の性能改善を実現するツールの積み重ねで、H.265/HEVC に対して約 26%の性能改善を実現している。今後、提案技術の本モデルへ実装し評価を行いたい。

第5章 H.264/AVC 映像伝送装置の実用化技術の研究⁴

5.1 背景と概要

2000 年台前半、従来の SD(Standard Definition)映像より圧倒的にリアルで高品位な HD(High Definition)映像の放送拡大やテレビ・カメラ・ビデオ・PC 等での HD 映像の記録・表示の要望が高まっていた。

特に放送業界では、映像ソースから制作まで全てハイビジョン化するためには、現場から放送局への映像伝送は重要な要素の一つとなる。ここで、放送局にとってハイビジョン映像伝送の要求条件は下記の通りとなる(図 35)。

- 従来の SD(Standard Definition)と同等以下の伝送コスト。インターネット利用が鍵
- 放送素材として十分な高画質
- ニュースやスポーツのライブ放送に不可欠な低遅延伝送

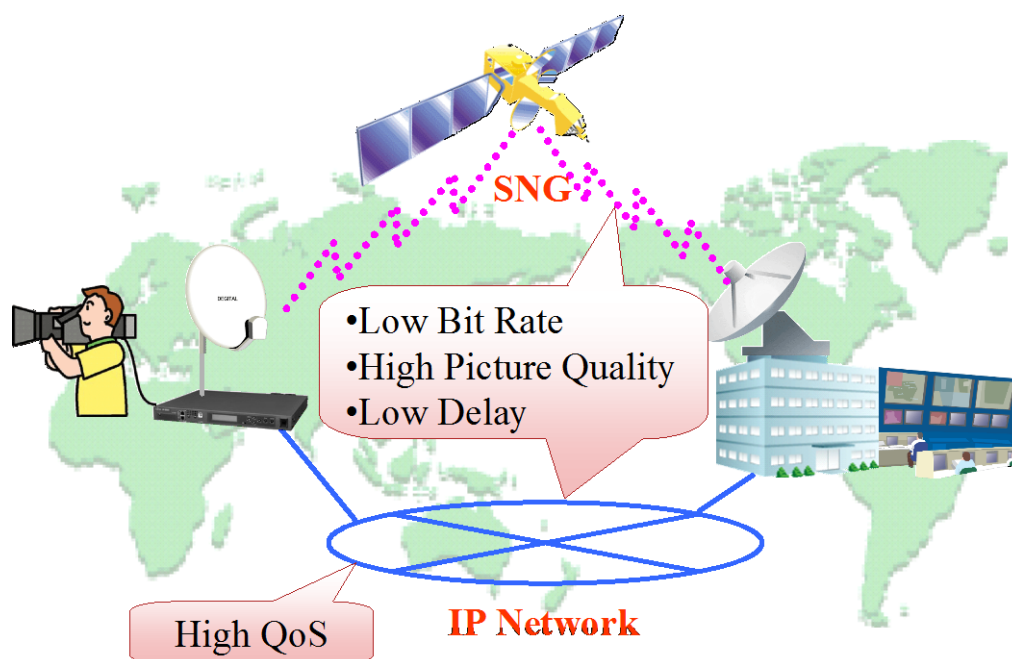


図 35 放送局におけるハイビジョン映像伝送の要求条件 [66]

各要求条件の詳細について、以下に述べる。まず、ライブ中継等において、高品位な通信路の確立が必要な場合には、HD-SNG 等により衛星経由の映像伝送が行われている。しかし、衛星回線には 1 時間当たり 10 万円以上の通信コストが必要なことから、放送素材として十分な高画質映像を、いかに少ない帯域で伝送し通信コ

⁴本章は業績 [J3] に基づいて執筆している

ストを削減するか、ということが重要な要件となる。特に、従来の SD と同等の帯域で HD の高品位な映像伝送が実現できれば、HD 化に伴う衛星通信コストは大幅に抑制できることとなる。

また、ライブ中継では現場との掛け合い等のインタラクションが重要な場合があり、映像の送受信にかかる遅延時間を短縮することが重要となる。

さらに、月額数千円程度の費用で 100Mbps 以上の帯域が利用可能な公衆インターネットを用いることで、24 時間 365 日、低コストで映像伝送が可能となる。しかし、公衆インターネットではパケット誤りは無視することが出来ない。このパケット誤りが発生する環境下で放送向けの映像伝送を可能とするためには、ネットワークの輻輳を生じさせない為の低レート動画像符号化技術と、プロが求める高い QoS を実現するパケット誤り訂正技術が不可欠となる。

このような要求条件に応えるため、H.264/AVC の高画質・低遅延化技術と公衆インターネット向けパケット誤り訂正技術を実用化し、これら技術を映像伝送装置 IP-9500 に適用した [66] [67] [30] [68]。本章は、文献 [66] に基づいた研究開発内容を説明する。

まず、ハイビジョン映像を高画質かつ低レートで符号化するために、MPEG-2 に比べて半分のビットレートで同等の PSNR を実現可能な H.264/AVC を採用することが考えられる。しかし、H.264/AVC の符号化モデル JM を用いて様々なハイビジョン映像を符号化した結果、高い主観品質を実現することは大変困難であることが判明した。さらに、ニュースやスポーツなどのライブイベントをブロードキャストするには、高画質に加え、低遅延で動画像符号化することが求められる。これらの課題を解決するために、人間の主観特性に基づいた符号化制御技術と低遅延符号化技術を開発した。

またプロの業務用途に耐える QoS を実現するパケット誤り訂正技術を実用化した。まず、従来のパケット誤り訂正技術として冗長パケットによる FEC が挙げられる。しかし、FEC では、大きな遅延を許容しない限り、高い誤り訂正能力を発揮できない。また、自動再送要求である ARQ は誤ったパケットを再送要求する技術である。しかし、輻輳時に不用意に ARQ を用いると、輻輳をさらに悪化させるばかりか、他のセッションへ悪影響を与えてしまう。我々は、FEC と ARQ を最適に組み合わせることで、TCP と親和性があり、かつ高性能なパケット誤り訂正技術を開発した。

そして、開発した高画質・低遅延符号化技術とパケット誤り訂正技術を映像伝送装置 IP-9500 に実装した。この映像伝送装置は、衛星向け HD-SNG のみならず、公衆インターネット経由の映像伝送に幅広く用いられている。

本論文では、人の視覚特性に基づいた高画質な実現する符号化制御、及び低遅延符号化技術を 5.2 節で、パケット損失の訂正技術を 5.3 節でそれぞれ説明し、最後

に 5.4 節では、これらの開発技術を適用した映像伝送装置 IP-9500 の市場での活用事例を示す。

5.2 高画質・低遅延を実現する動画像符号化技術

前述のとおり、放送等の業務用映像伝送のアプリケーションでは、通信衛星、公衆インターネットの両通信路において、コストや通信品質安定の観点から、少ない情報量で高い品質の動画像符号化を実現することが要求されている。さらに、ライブニュースやスポーツ放送では、現場との掛け合いや速報性のために低遅延での映像伝送も求められる。この要求条件を満たすために、従来の MPEG-2 の半分の情報量で同等画質の客観品質である PSNR が実現可能な H.264/AVC [8] [69] の活用が期待されている。

H.264/AVC を策定した Joint Video Team(JVT)は、標準化策定の過程で、H.264/AVC の能力を最大限発揮可能な符号化アルゴリズムを実装した参照ソフト Joint Model(JM) [56]を開発した。しかし、標準化当時、JM を用いた符号化映像は、PSNR は高いものの、特に難しい映像シーンにおいて主観品質が安定しないという課題があり、そのままでは実用化が困難であった。これに対し、我々は分散計算機環境 CyberGRIP [70]による大量シミュレーションを行い、人間の主観特性に基づいた画質制御アルゴリズムを開発した。

本節では、H.264/AVC の符号化技術を実用化する上での画質制御の課題と、人間の主観特性に基づいた画質制御、そして、画質制御を活用した低遅延符号化について説明する。

5.2.1 H.264/AVC 参照ソフト JM で符号化した画質の課題

H.264/AVC の参照ソフトウェア JM で映像情報メディア学会のハイビジョン・システム評価用標準動画像を符号化したところ、従来の MPEG-2 では見られなかった、次の 3 種類の主観画質劣化が生じることが分かった。

まず第 1 の劣化は、精細感の消失である。この現象は、芝生や人物の服装など、細かいテクスチャが存在する領域で顕著となる。従来の MPEG-2 のテストモデル TM5 等で符号化すると、レートが低い場合にはブロックノイズやモスキートノイズが顕著となる。一方、H.264/AVC の参照ソフトウェア JM で符号化すると、これらのノイズが抑制される替りに、画像の高周波部分が極端に失われ、全体として精細感のない映像となる傾向がある。

第 2 は、H.264/AVC に新たに導入されたイントラ予測によって引き起こされる I ピクチャのフリッカノイズである [71]。H.264/AVC では、イントラマクロブロックにおいて、新たに隣接ブロックから外挿予測による予測画像生成の機能が導入され

た。この技術の導入により、イントラマクロブロックのレート・歪特性という指標での符号化効率は大幅に改善した。一方、この画素の外挿予測により、周辺の符号化済ブロックの符号化誤差がイントラマクロブロックの予測画像全体に伝搬してしまう。この誤差は周波数的には低周波に現れるため、直前のピクチャと I ピクチャの間で画面全体が脈動するような劣化が知覚される。

そして、第 3 の課題は、レート制御の安定性である。H.264/AVC では、従来の MPEG-2 と比べて、同一画像を符号化した際の発生情報量の挙動が大幅に異なる。H.264/AVC は MPEG-2 に比べて 1/2 の情報量で符号化可能である。これは、この少ない情報を、シーケンスの各ピクチャ、そして、ピクチャ内の各領域に丁寧に割当てる制御が必要なことを意味する。

5.2.2 人間の視覚特性に基づいた画質制御技術

本小節では、人間の視覚特性に基づいた画質制御について説明する。実用化した画質制御アルゴリズムは、ROI 検出と情報量割り当ての 2 段階で構成される(図 36)。

まず、入力した動画像から関心領域 (ROI) を検出する。文献 [47]等では、人間の視覚感度の研究が報告されている。ここでは、静止画において、5cycle/degree 付近の周波数に最大感度のピークをもつが、映像が動くにつれ最大感度ピークが低周波側にシフトし、動きが 12.5degree/秒の場合には最大感度が 1.5cycle/degree となることが示されている。このような映像の動きと視覚感度等の特性に基づいて、入力された動画像の中から視覚的に重要な領域を ROI(Region Of Interest)として検出する。そして、この ROI を映像の中で時空間的に追跡を行う。

次に、目標ビットレートと HRD バッファ容量(H.264/AVC Annex C)の制限の元で、ROI を考慮して、動画像の各フレームや各マクロブロックに正確に情報量を割り当てる。

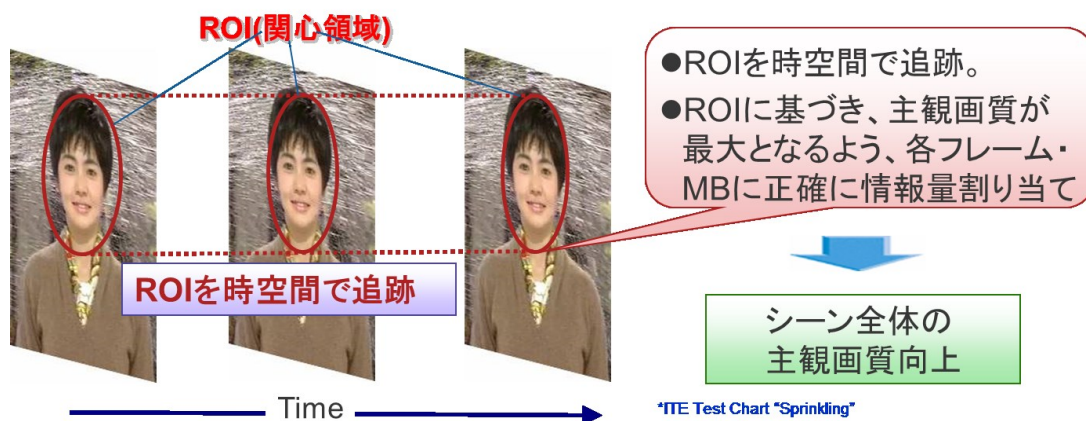


図 36 視覚特性に基づいた画質制御([66]をもとに修正)



(a) JM



(b)開発技術

図 37 JM と開発技術の”Sprinkling”シーケンス符号化映像比較 [66]⁵

図 37 は、H.264/AVC のテストモデル JM と開発技術でそれぞれ符号化した映像の比較を示したものである。評価映像には、映像情報メディア学会が監修した HD 解像度の評価用標準動画像である「ITE 標準 HDTV 動画像」の”Sprinkling”シーケンスを用いた。この映像は、クロマキーの人物の前景と、シャワーの水飛沫の背景から構成されるものである。水飛沫背景の動き予測効率の低下で多くの符号化情報が必要となる結果、顔や服装へ割当てられる情報量が不足し精細感を保つことが大変困難となる。このような特性から、本シーケンスは、符号化装置やアルゴリズムの性能評価のために、研究のみならず放送局等で広く用いられている [72]。また、画像サイズは地デジで用いられている幅 1440 画素、1080 ラインのインターレース、ビットレートは CBR で 14Mbps とした。また HRD バッファ遅延は 500m 秒とした。

この図で示した通り、JM では、顔領域の顕著な劣化や、服の質感の喪失が見られる。これに対し、開発した符号化アルゴリズムでは、顔領域や服の劣化が最小限に抑えられていることが分かる。

5.2.3 低遅延符号化

一般に放送等の動画像符号化では、符号化効率を最大化するために、フレーム内符号化の I ピクチャ、片方向予測の P ピクチャ、双方向予測の B ピクチャを組み合

⁵ 映像情報メディア学会より”Sprinkling”の画像の論文掲載を許諾頂いた

わせた GOP 構造が良く用いられる。また、伝送路でのビット誤りや、途中からの映像再生のために、周期的に I ピクチャが挿入される。しかし、I ピクチャは情報量が多いことから、HRD バッファにおいて、I ピクチャを蓄積するための十分な遅延を設定する必要がある。これに対し、GDR (Gradual Decoding Refresh) アルゴリズム [73]を用いると、各ピクチャで発生する情報量が均等化される結果、バッファに必用な遅延を削減することができる。

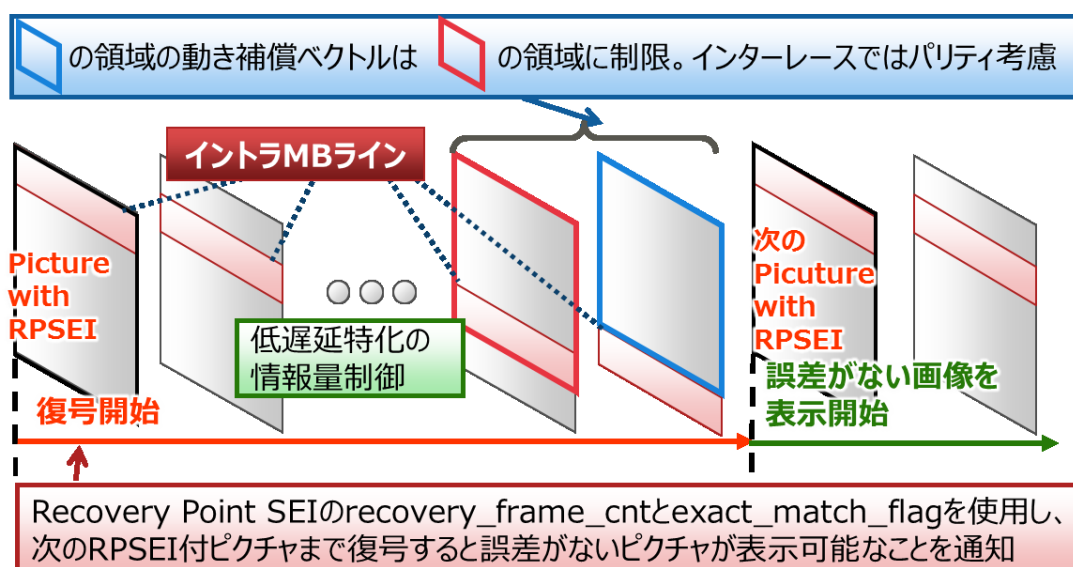


図 38 GDR に基づいた低遅延符号化([66]をもとに修正)

図 38 は、GDR の概略を示したものである。マクロブロックのライン数を M とする。H.264/AVC では、Recovery Point SEI(RPSEI)は復号を開始可能なピクチャを示す情報である。GDR では、この Recovery Point SEI を付加したピクチャにおいて、一番上のマクロブロックラインをイントラで符号化する。次に、これに続くピクチャは、このイントラマクロブロックラインをピクチャ単位で 1 ラインずつ下方にスライドさせて符号化する。ここで、このイントラマクロブロックラインより上の領域を動き予測符号化の際の参照領域として、前ピクチャのイントラマクロブロックラインより上の領域しか用いないよう、動きベクトルの制限を設ける。そして、画像の下端までスライドすると、再度、ピクチャに Recovery Point SEI を付加し、上記の処理を繰り返す。この処理により、復号時には、画面の上から徐々に映像が再生され、次の Recovery Point SEI を付加したピクチャの復号後には、映像中に未確定領域がなく、完全な映像が復号される。また、この処理により、各ピクチャのイントラマクロブロック数は同一となることから、各ピクチャの情報量が平滑化され、HRD バッファに必用な遅延が削減できる。

ここで、復号を開始後、何枚目の画像で未確定領域がなく完全な映像が再生可能か示すために、Recovery Point SEIにおいて `recovery_frame_cnt` と `exact_match_flag` を用いた。まず `recovery_frame_cnt` を M に設定すると、復号装置では、 M 枚後の復号画像から表示すると未確定領域がないことが分かる。次に、`exact_match_flag` を 1 に設定することで、復号された画像は、符号化側の局部復号画像と正確に一致していることが分かる。

また、前小節の人間の視覚特性に基づいた画質制御技術を GDR に最適化させるとともに、イントラ MB ライン境界が認知されない制御も新たに適用した。

さらに、インターレースにおいては、3.3 で筆者が提案し H.264/AVC に採用された色差ベクトル補正技術に対応したベクトル制限技術 [74] も適用した。

これらの技術を映像伝送装置 IP-9500 に実装することで、HDR バッファ遅延 100ms、伝送装置の全ての多重化・符号化・復号・映像入出力処理を含めて 300ms 未満の低遅延での高品質な映像伝送を実現した。またビットレートは、従来では低遅延伝送が困難な 4Mbps の低レートまで対応した。

5.3 公衆インターネット向けの高性能パケット誤り訂正技術

前述のように、B フレッツ等の公衆インターネットは月数千円で利用可能である。本節では、公衆インターネットで生じるパケット誤り環境下で、業務用品質の映像伝送を可能とする誤り訂正技術を説明する。

5.3.1 IP ネットワークにおけるパケット損失

B フレット公衆等の公衆インターネットは最大 100Mbps の帯域が使えるものの、パケット誤りが不可避である。我々は、まず、公衆インターネットのパケット誤りの測定を行った。

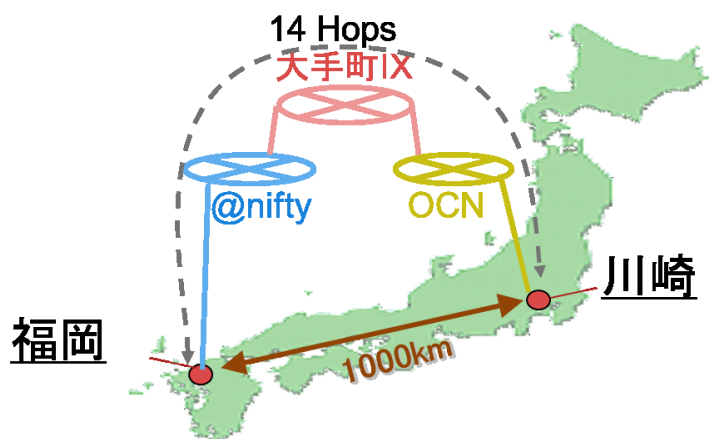


図 39 公衆インターネットの誤り率測定に用いた環境 [66]

2003年10月18日から2003年10月19日の48時間にわたり、公衆インターネットの packets 誤り率の測定を行った。まず、通信の送信・受信の距離を長くするために、福岡に送信端末、川崎に受信端末を設置した。二都市間は約1000kmである。ここで、一般にインターネットは異種プロバイダ間での帯域が細いことが多いことから、福岡側では@nifty、川崎側ではOCNを用いた。またtracertにより、両拠点間のルータのホップ数は14、平均往復遅延(RTT)は約30msであり、大手町のIXを経由していることを確認した。

図40は、観測されたパケット損失率を示す。この観測により最大パケット損失率は最大数%になる可能性があることが判明した。観測時点からネットワークの品質が向上し、パケット損失率は全体的に低下したものの、現在でも公衆インターネットの輻輳によるパケット損失は無視できないと考える。

これらのことから、数パーセントのパケット損失に対するエラー訂正能力が必要と考える。

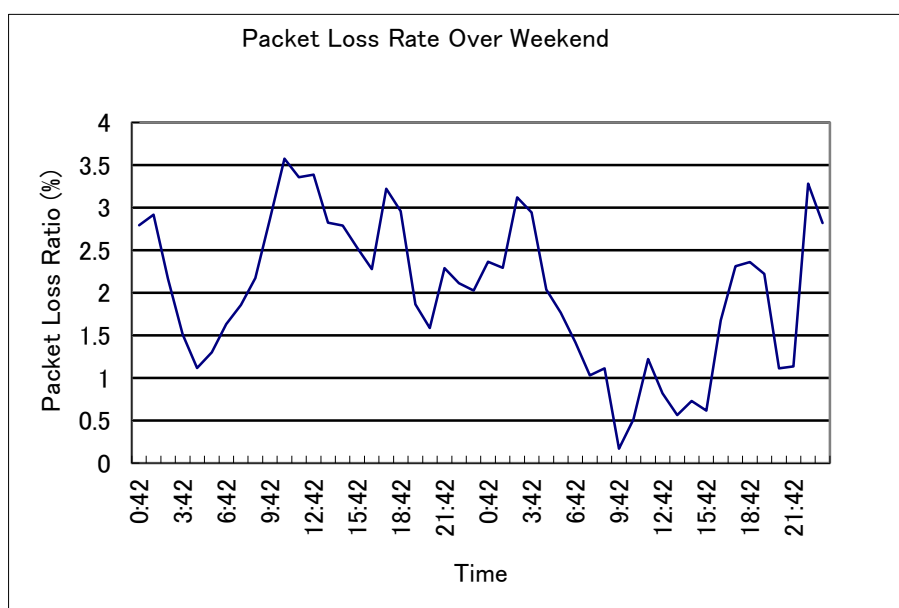


図 40 観測されたパケット損失率 [66]

5.3.2 従来のパケット損失の訂正技術

本小節では、従来の FEC および ARQ の概要と問題点について説明する。

まず、従来の IP パケット向けの FEC として、例えば RFC2733 [75] および ProMPEG CoP3 [76] が普及している。特に ProMPEG Cop3 は 2 次元のパリティマトリックスにより、高い訂正能力を有する FEC 方式である。これらの FEC 方式では、送信側で情報パケットの XOR 計算により冗長なパリティパケットが生成され、受信時にパケット損失が発生すると、損失パケットはパリティパケットから修復される。しか

し、4%の高いパケット損失環境では、パリティパケット率 20%の ProMPEG CoP3 を用いても、誤り訂正後のパケット損失率 (PLR) は 2×10^{-4} となる。すなわち、伝送レートが 8Mbps の場合には、エラー訂正後でも約 7~8 秒ごとにパケット損失が発生してしまう。そのため、業務用映像伝送には誤り訂正能力は不十分である。

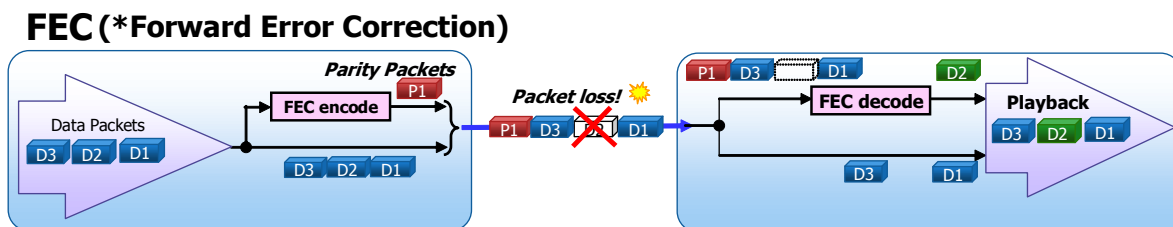


図 41 FEC の処理概要 [66]

また、ARQ 方式では、受信側でパケット損失を検出した後、送信側に損失したパケットの ID とともに NAK という受信失敗通知を返送する。送信側は、NAK を受信すると、損失したパケットを受信側に再送する。ARQ は通常、FEC よりも優れた性能を示す。しかし、インターネット環境で不用意に使用されると、ネットワークの輻輳がさらに発生し、パケット損失がさらに増大するばかりか、TCP などの他の共存セッションに害を及ぼす可能性がある。

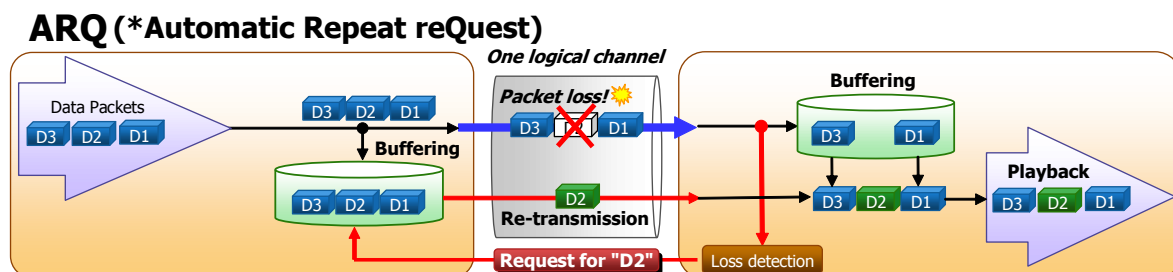


図 42 ARQ の処理概要 [66]

5.3.3 インターネット向けの FEC + ARQ ハイブリッド誤り訂正

我々は、FEC と ARQ を組み合わせることで、高性能な誤り訂正を実現する方式を開発した [77] [78]。図 43 は提案方式の帯域制御の概念図、図 44 は提案方式のブロック図を示す。

我々の検討では、インターネットでのエラー率を下げるためには、シェーパの効果大きいことがわかった。一般にネットワークが輻輳するとバースト的なパケット損失の発生が生じる。一般にバースト的なパケット損失が生じた場合、ランダムなパケット損失よりもエラー訂正が困難である。よって、パケットの送出を一定

にするシェーパにより、バースト的なパケット損失の発生を抑制することができる。このことから、FEC と ARQ の和が一定となるようにシェーパで全体のレートを確実に平滑化する Hybrid 方式を開発した。

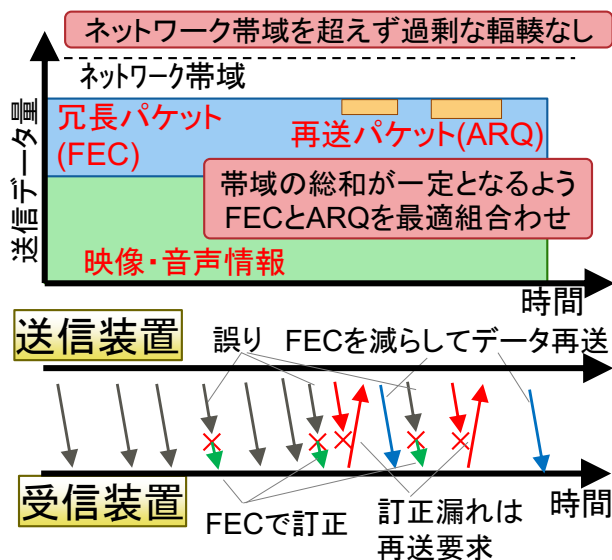


図 43 提案する FEC+ARQ の帯域制御

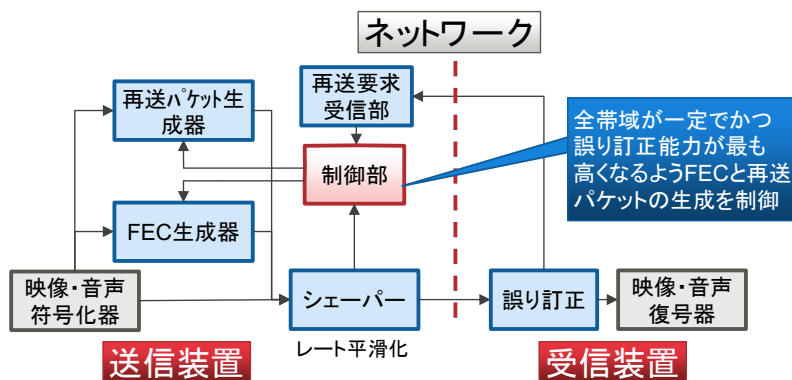


図 44 提案する FEC+ARQ のブロック図

提案方式では、送信側は、通常時においては FEC のみを使用する。ここで、受信側で FEC により修正しきれないパケットの損失を検出し、送信側へ NAK の再送要求を送信すると、送信側は ARQ と同様にパケットを再送する。そして、同時に ARQ パケットにより増加した帯域分だけ FEC パケットの比率を減少させるよう最適制御を行い、シェーパで全体のレートを平滑化しパケットを送出する。

この提案手法には次の二つのメリットがある。まず一つ目はエラー訂正能力の向上である。FEC または ARQ それぞれ単独のみを使用する場合と比較して、エラー

耐性を大幅に向上させることができる。また、二つ目のメリットは、シェーパにより帯域を完全に固定とすることができるので、輻輳に対して悪影響を与えず、TCPや他のセッションに悪影響を与えない。また、レート平滑化によりバースト誤りが抑制されることも、性能向上に寄与する。

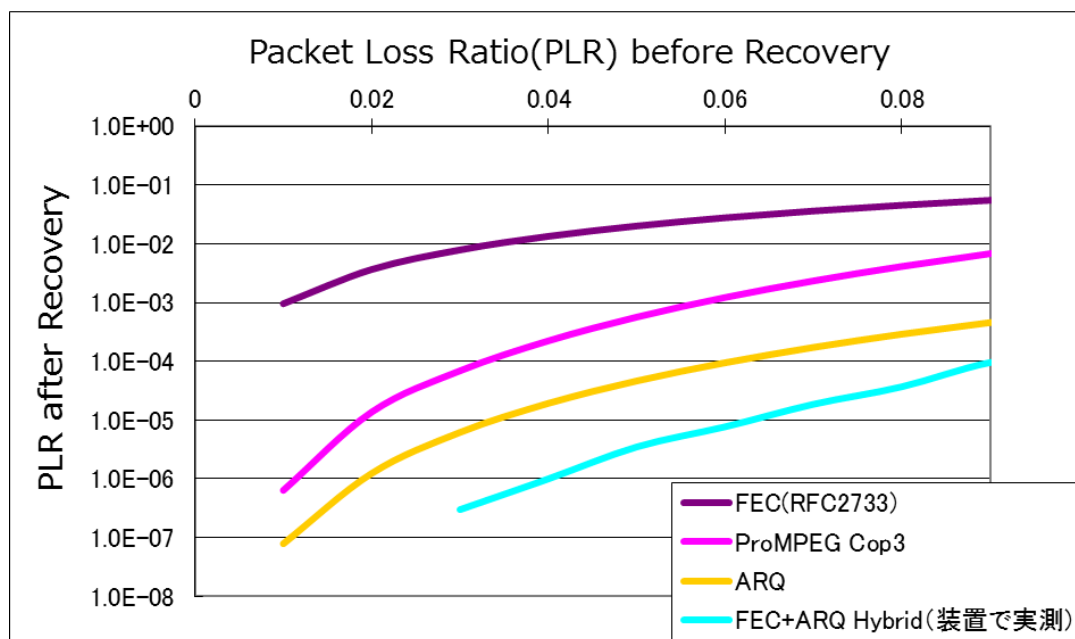


図 45 パケット損失時のエラー訂正能力の比較([66]をベースに修正)

図 45 はランダムにパケット損失を発生させた場合の、それぞれの方式のエラー訂正後の性能比較の結果を示す。横軸が訂正前、縦軸が訂正後のパケット損失率である。提案技術は映像伝送装置 IP-7000 の実機に実装し、ネットワークシミュレータでエラーを発生させて評価をした。また比較対象の FEC としては RFC2733 及び ProMPEG CoP3 を取り上げ、ARQ とともにシミュレーションを行った。冗長帯域については RFC2733 と提案方式は 10%、ProMPEG CoP3 は規格上の冗長率下限の 20% とした。ARQ と提案方式においては再送による遅延増加を防ぐため、最大の再送回数を 3 回とした。

この図より、FEC または ARQ をそれぞれ単独で用いた場合より、提案方式の方が、高い誤り訂正能力を示すことが分かる。

5.3.4 装置への実装と公衆インターネットでの適用

我々は、開発技術を SD 対応の MPEG-2 映像伝送装置 IP-7000、及び HD 対応の H.264/AVC 映像伝送装置 IP-9500 [30]に実装した。そして、これら装置を用いて、いくつかの放送局が公衆インターネットでの評価を行った。テレビ東京とテレビせと

うちは、約 700km 離れた両放送局間で IP-9500 を用いた映像伝送の評価を行った [79]。ビットレートは比較的高い 20Mbps を使い、評価時間は約 24 時間である。その結果、提案技術により、24 時間の間で訂正できなかったパケットは 1 個のみであった。これは、訂正後の誤り率は 7.0×10^{-9} 以下であることが分かる。

このように高いパケット損失に対するエラー訂正能力を有することから、本装置は映像素材伝送やお天気カメラなど、国内外で公衆インターネット経由の映像伝送に広く使われている。具体的な事例は 5.4 で説明する。

5.4 開発した映像伝送装置 IP-9500 の適用事例

本節では、開発技術を適用した映像伝送装置 IP-9500 の適用事例を説明する。

5.4.1 ANN 系列のインターネット経由映像素材伝送システム

テレビ朝日をキー局とする ANN 系列局は、2006 年に SD 対応 MPEG-2 映像伝送等 IP-7000、2007 年 6 月に HD 対応 H.264/AVC 映像伝送装置 IP-9500 を使用して、公衆インターネットを用いた系列放送局間の映像素材伝送システムを実用化した [80] [81]。

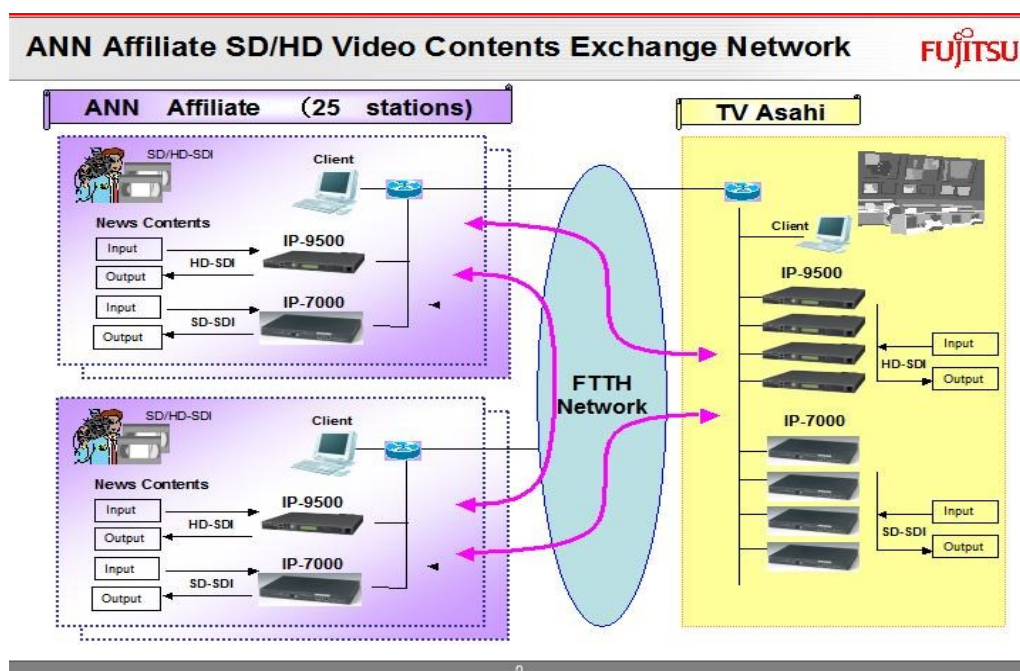


図 46 ANN 系列でのインターネット経由映像素材伝送システム [66]

テレビ朝日は、公衆インターネットを用いた映像素材伝送システムを検討し、実現のためには、パケット損失に対する強力なエラー訂正機能を有する映像伝送装置

が必要であった。そのため、開発した Hybrid ARQ 方式を実装した映像伝送装置 IP-7000 及び IP-9500 を採用し、映像素材伝送システムを構築した。

システム構築後、ANN 系列局の間では、本システムを用いた映像素材交換が活発に行われるようになった。活用事例としては、地震映像のための 24 時間連続監視、仙台の楽天スタジアムの野球映像を同日の夜のニュースで放送するための東京への素材伝送、長崎のニュース取材映像の頻繁な福岡への素材伝送等である。

このシステムにより、ANN では、地方局が 24 時間いつでも、コストに気を使うことなく、ニュース素材を系列間で頻繁に共有することができた。

5.4.2 フジテレビのお天気カメラ

フジテレビは 2005 年に SD 対応 MPEG-2 映像伝送等 IP-7000、2007 年に H.264/AVC 映像伝送装置 IP-9500 を用い、公衆インターネット経由でお天気カメラ映像をライブ伝送するシステムを構築した。本システム構築においては、低ビットレートでの画質と公衆インターネットの packets 損失に対するエラー耐性能力を様々な装置で比較した結果、本研究の開発技術を適用した IP-7000 と IP-9500 が採用された。

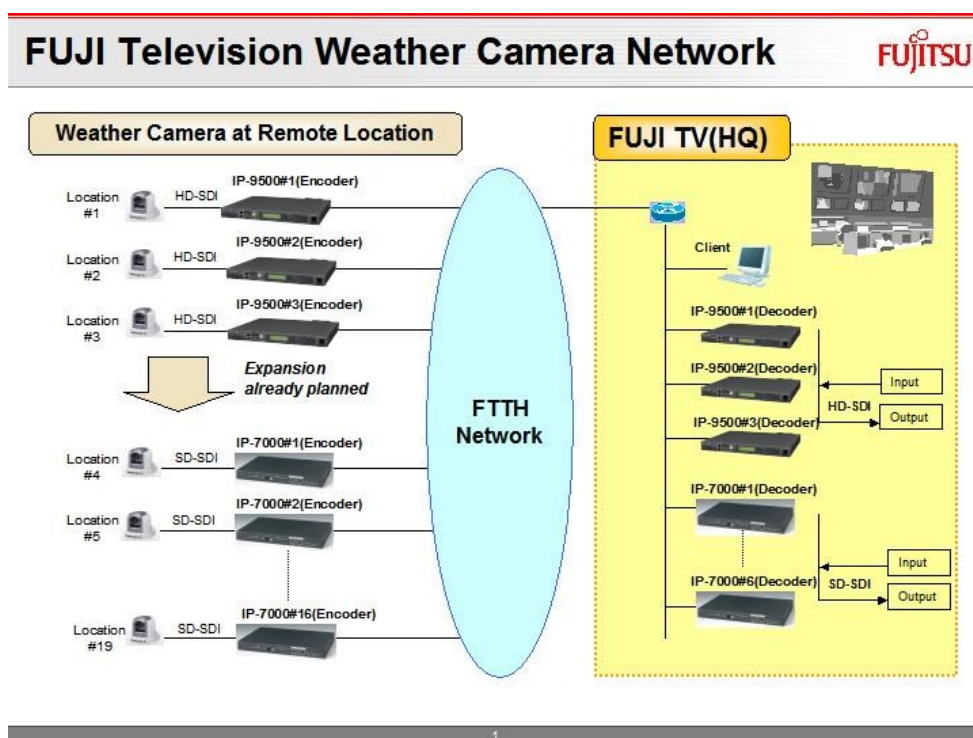


図 47 フジテレビのインターネット経由お天気カメラ [66]

5.4.3 日本テレビの NFL スーパーボウル XLI の日米間 HD ライブ中継

日本テレビは、2007年2月4日に、NFL スーパーボウル XLI の HD ライブ放送に映像伝送装置 IP-9500 を用い、世界初の日米間の H.264/AVC での HD 映像伝送を行った [82]。

日本テレビは長年にわたり米国からスーパーボウルの中継を行っており、2006年までは MPEG-2 を用いてきた。そして、2007年に初めて IP-9500 を用いた H.264/AVC による伝送を 20Mbps で実施した。その結果、プロスポーツのような動きの激しい映像であっても、放送用の映像素材として十分な品質の映像を伝送することができた。

日本テレビは米国と東京の間で、IP-9500 の IPBB 構成の GOP の通常モードと、GDR を用いた低遅延モードの双方で伝送遅延を測定した。標準モードの場合、米国と東京の総遅延は 1.77 秒となり、ライブでの運用は困難である。しかし、低遅延モードの場合には、この遅延は約 0.71 秒に短縮され、ライブの国際映像伝送の運用上、問題のないレベルとなった。

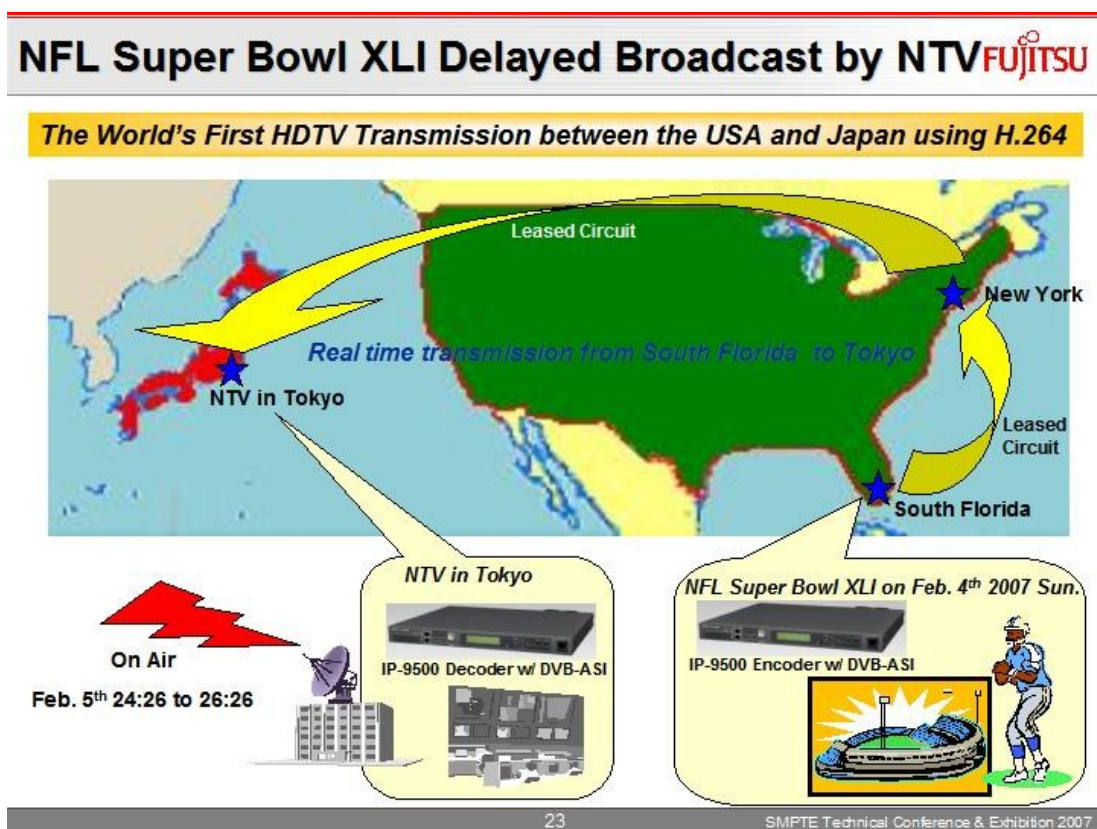


図 48 日本テレビの NFL スーパーボウルの HD ライブ中継 [66]

5.5 本章のまとめ

放送局等の業務用の HD 映像伝送には、次の三つの要件が重要であることを明らかにした。

- 従来の SD(Standard Definition)と同等以下の伝送コスト
- 放送素材として十分な高画質
- ニュースやスポーツのライブ放送に不可欠な低遅延伝送

この要求条件を実現するため、人の視覚特性に基づいた高画質な実現する符号化制御、低遅延符号化技術、そして、インターネットの packets 損失によるエラーを高性能で訂正する Hybrid ARQ 技術を開発し、映像伝送装置 IP-9500 に実装した。そして、これらの開発技術の性能が放送局で評価され、衛星向け HD-SNG のみならず、公衆インターネット経由の映像伝送に幅広く貢献することができた。

第6章 4:2:2/4:2:0 相互変換時の色にじみ抑止技術⁶

6.1 背景と概要

放送番組制作において、ニュース等の中継では、同時に多数の映像伝送装置を稼働させ、撮影された映像は衛星経由の SNG(Satellite News Gathering)やマイクロ波経由の FPU(Field Pick up Unit)等で伝送される。ここで、SNG の回線は高価(10 万円/時間)で、FPU は帯域に限りがある。このようなことから、現場中継の映像伝送では運用コストを最小限に抑えるために、機器コスト、及び伝送レートの双方を下げる事が求められる。特に伝送レートについては、10Mbps 以下が目標となる。この目的を達成するために、民生技術の転用が可能で、低レートでも比較的高画質な 4:2:0 映像フォーマット対応の H.264/AVC 映像伝送装置が広く使われている。

H.264/AVC 符号化の標準規格は、映像フォーマットとして、1.3.2 小節で説明した 4:2:2 フォーマットと 4:2:0 フォーマットの両方に対応している。これらの二つのフォーマットの間には次のトレードオフがある。低ビットレートの場合、4:2:0 フォーマットの方が画素数は少ないため PSNR が向上し主観品質が高まる。一方、高ビットレートの場合は十分な情報量が全画素に割り当たるため、色成分の再現性が高い 4:2:2 のほうが高画質となる。このようなことから、4:2:2 フォーマットは、主に放送局等の業務用途において、映像アーカイブ、スポーツ映像等の高品質な映像素材の記録・伝送に用いられる。一方、4:2:0 フォーマットは低レート時の画質が良好なことから民生機器で用いられ、さらに業務用途でも民生技術転用による機器の低コスト化により、SNG/FPU 等の現場中継の映像伝送で広く使われる。

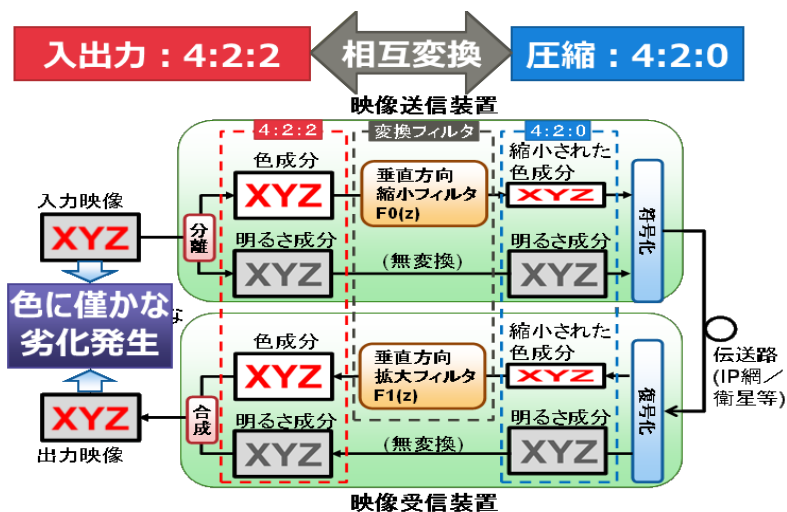


図 49 4:2:0 映像伝送装置における 4:2:2/4:2:0 相互変換 [83]

⁶本章は業績 [J2] の内容に基づいて執筆している

しかし、4:2:0 フォーマットの映像伝送には大きな欠点がある。放送局によると、現場中継では最大 5 回程度の符号化・復号を繰り返す多段中継を行うことがある。ここで復号装置と符号化装置は 4:2:2 映像フォーマットである HD-SDI ケーブルで接続されているため、中継毎に 4:2:2 フォーマットと 4:2:0 フォーマットの相互変換が必要となる(図 49)。1 回の変換では知覚できないほどの誤差だが、繰り返し中継を行うことで誤差が累積し、はっきりと知覚される色成分の劣化が生じてしまう。

この問題を解決するために、4:2:2/4:2:0 両フォーマットの色差成分のダウンサンプリングおよびアップサンプリングのフィルタとして Non-Degraded 4:2:0(ND420) フィルタを提案する [83] [84]。この ND420 フィルタを用いると、多段中継の際に 4:2:2/4:2:0 相互変換を複数回繰り返しても、色差成分の劣化が完全に抑制される。

我々はこの ND420 フィルタを、5.4 節で説明した動画像符号化・復号装置に実装し、最大 16 回の中継でも色差成分が劣化しないことを確認した。

本章では、文献 [83]の内容に基づいて、ND420 フィルタセットの機能・設計方法・性能評価について説明する。

6.2 Non-degraded 4:2:0 フィルタの要求条件

本節では、ND420 フィルタセットの設計に必用な要求条件を説明する。

第 1 の要件は多段接続時の色にじみの抑止である。図 50 は、Non-Degraded 4:2:0 フィルタと従来フィルタの多段接続後の復号画像の画質を説明したものである。4:2:0 フォーマットによる映像伝送では、復号装置の出力を 4:2:2 にアップサンプリングして HD-SDI ケーブルで伝送し、これを再度 4:2:0 にダウンサンプリングして符号化する。そのため、従来装置では多段中継の後段ほど、色差成分がぼけてしまう。

これに対し、我々が設計すべき ND420 フィルタにおいては、多段中継を行った後でも、4:2:2/4:2:0 の相互変換による誤差が累積しないことが必要である。

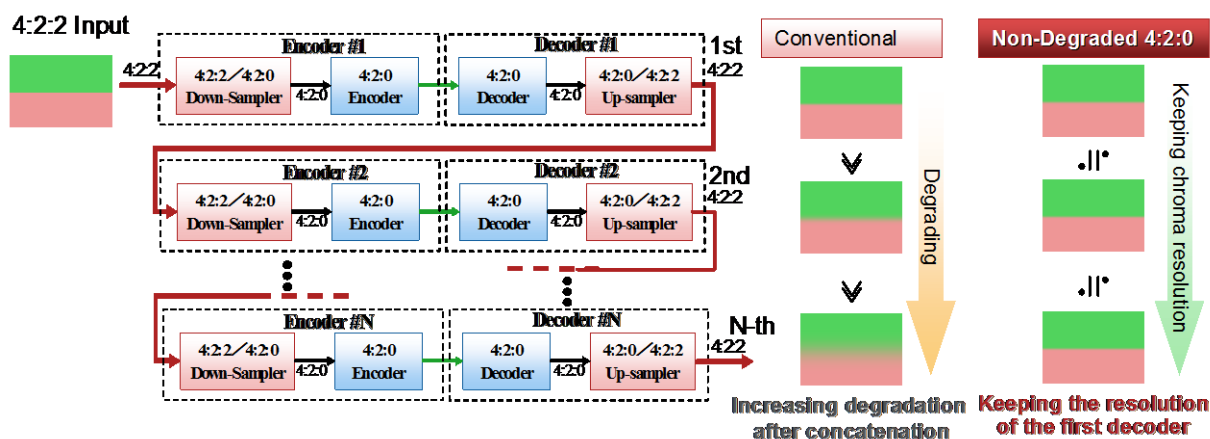


図 50 Non-Degraded 4:2:0 フィルタと従来フィルタの多段接続後の復号画像 [83]

第2の要件は、従来の復号装置との互換性である。現場でのSNG/FPUの運用では、複数メーカーの映像伝送装置が用いられている。放送局がND420対応の映像伝送装置の導入をしやすくするためには、ND420対応の装置同士に加え、従来の4:2:0映像伝送装置と接続した際にも、色成分の画素ずれ(色ずれ)等の劣化が発生しないことが必要である。

このように、多段中継時の色にじみ抑止、従来機器との混在運用時の色成分の画素ずれ防止、の二つがND420フィルタセットの要求条件となる(図51)。

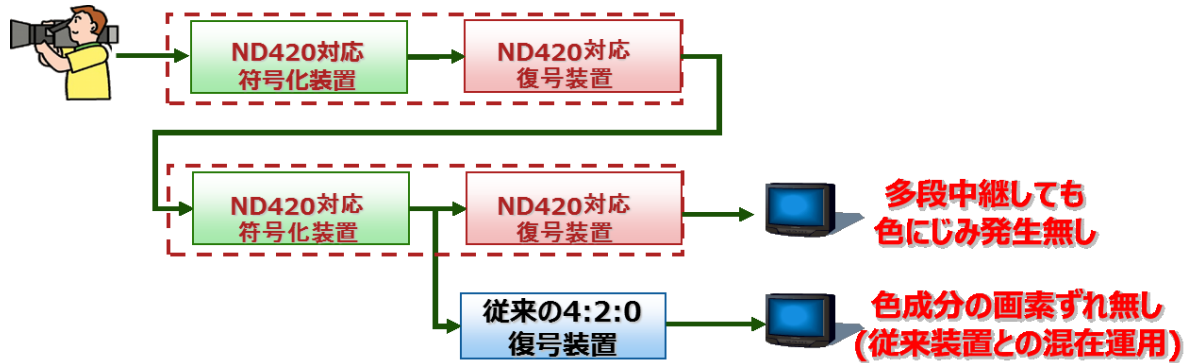


図 51 ND420 映像伝送装置と従来の映像伝送装置との互換性 [83]

6.3 Non-degraded 4:2:0 フィルタセットの設計方法

本節ではND420フィルタセットの設計方法について説明する。まず、相互変換誤差の累積抑止の条件と色差画像の位相条件について数学的な定式化を行った後に、フィルタ設計の手順を説明する。そして、最後に設計したフィルタセットを示す。

6.3.1 要求条件の数学的な定式化

本節では、ND420フィルタセットの数学的条件を明らかにする。図52は4:2:2/4:2:0ダウンサンプリング及び4:2:0/4:2:2アップサンプリングの変換モデルを示す。

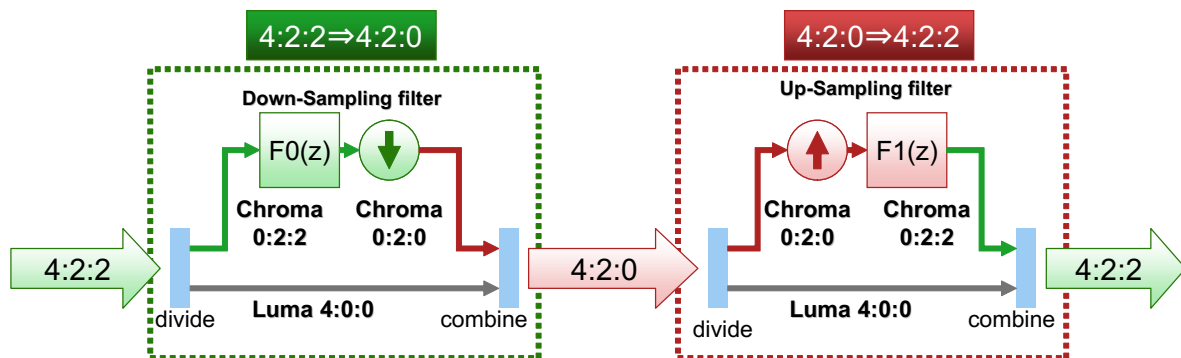


図 52 動画像符号化装置及び復号装置におけるフィルタのモデル化 [83]

まず、 $F0(z)$ および $F1(z)$ は、それぞれ以下のような係数 $f0_k$ および $f1_k$ を有するダウンサンプリングおよびアップサンプリングフィルタの z 変換として定義する。

$$F0(z) = \sum_{k=0}^{N-1} f0_k \cdot z^{-k} \quad (45)$$

$$F1(z) = \sum_{k=0}^{N-1} f1_k \cdot z^{-k} \quad (46)$$

ここで、 N はフィルタタップ長とする。まず4:2:2/4:2:0変換においては、色差信号は4:2:2フォーマットから分離され、ローパスフィルタ $F0(z)$ と2:1ダウンサンプラにより垂直方向にダウンサンプリングされる。そして、この色差と輝度を組み合わせることで、4:2:0フォーマット信号が生成される。

4:2:0/4:2:2変換では、同様に色差信号は4:2:0フォーマットから分離され、1:2アップサンプラとローパスフィルタ $F1(z)$ を垂直方向に適用してアップサンプリングされる。そして、この色差と輝度を組み合わせることで、4:2:2フォーマット信号が生成される。

6.3.2 相互変換誤差の累積抑止の条件の定式化

本小節では、相互変換誤差の累積抑止の条件の数学的定式化を行う。図53は単純2:1ダウンサンプラの手順を説明する図である。入力信号を $a_0, a_1, a_2, a_3, \dots$ とする。2:1ダウンサンプラは、この入力信号の偶数番目のみを選択して出力する。

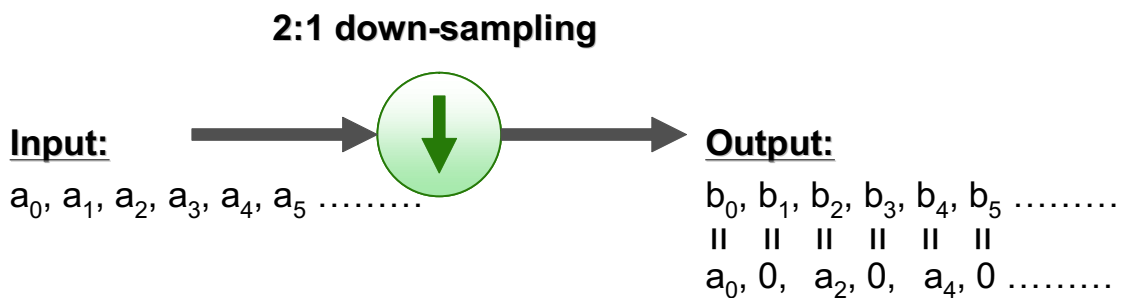


図 53 単純 2:1 ダウンサンプラの処理 [83]

この場合、出力信号 b_0, b_1, b_2, \dots は以下のように記述される。

$$b_n = \begin{cases} a_n, & n = \text{even} \\ 0, & n = \text{odd} \end{cases} \quad (47)$$

また入力信号 $\{a_n\}$ の z 変換 $A(z)$ は、以下のように記述することができる。

$$A(z) = \sum_n a_n \cdot z^{-n} \quad \text{where } z = e^{j\omega} \quad (48)$$

出力信号 $\{b_n\}$ の z 変換 $B(z)$ の z 変換は、この入力信号 $A(z)$ を用いて次のように表現できる。

$$B(z) = \sum_n b_n \cdot z^{-n} = \sum_{n=\text{even}} a_n \cdot z^{-n} = \sum_n \left(a_n \cdot \frac{1}{2} (1 + (-1)^n) \right) \cdot (z)^{-n} = \frac{1}{2} \sum_n a_n \cdot \left((z)^{-n} + (-z)^{-n} \right) = \frac{1}{2} \left(\sum_n a_n \cdot (z)^{-n} + \sum_n a_n \cdot (-z)^{-n} \right) = \frac{1}{2} \cdot (A(z) + A(-z)) \quad (49)$$

図 54 は、ダウンサンプリングおよびアップサンプリングの両フィルタの縦続接続を示したものである。

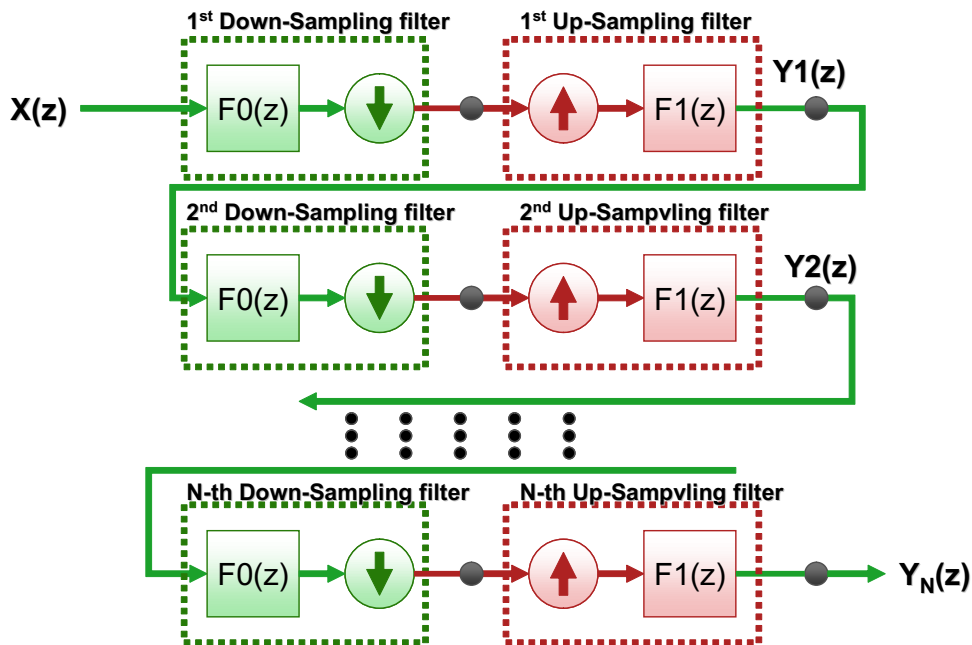


図 54 ダウンサンプリング・アップサンプリングの縦続接続 [83]

図 54 と式(49)を使用することにより、完全再構成条件は以下のように導き出すことができる。

入力信号 $X(z)$ は、ダウンサンプリングフィルタ $F0(z)$ 、2:1 ダウンサンプラ、1:2 アップサンプラ、アップサンプリングフィルタ $F1(z)$ を適用することで、最初の復

号画像をアップサンプリングした 4:2:2 映像の z 変換である $Y_1(z)$ が得られる。同様にこの処理を繰り返し適用することで、第 N 番目の復号画像の 4:2:2 出力である $Y_N(z)$ が得られる。

まず $Y_1(z)$ は次式で表すことができる。

$$Y_1(z) = \frac{1}{2} \cdot (F0(z) \cdot X(z) + F0(-z) \cdot X(-z)) \cdot F1(z) \quad (50)$$

次に、式(50)の $X(z)$ に $Y_1(z)$ を再帰的に代入することにより、第 2 のアップサンプリングフィルタ $Y_2(z)$ の出力を導出する。

$$\begin{aligned} Y_2(z) &= \frac{1}{2} \cdot (F0(z) \cdot Y_1(z) + F0(-z) \cdot Y_1(-z)) \cdot F1(z) \\ &= \frac{1}{4} \cdot (F0(z) \cdot F1(z) + F0(-z) \cdot F1(-z)) \cdot (F0(z) \cdot X(z) + F0(-z) \cdot X(-z)) \cdot F1(z) \\ &= \frac{1}{2} \cdot (F0(z) \cdot F1(z) + F0(-z) \cdot F1(-z)) \cdot Y_1(z) \end{aligned} \quad (51)$$

ここで、我々の目標は、ダウンサンプリング/アップサンプリングを 2 回、3 回、そして N 回繰り返しても、最初のアップサンプリングフィルタ出力の画質を維持することである。これより、フィルタが満たすべき条件として下記が導出される。

$$\frac{Y_2(z)}{Y_1(z)} = \frac{1}{2} \cdot (F0(z) \cdot F1(z) + F0(-z) \cdot F1(-z)) = 1 \quad (52)$$

さらにこの式は次のように書き換えることができる。

$$F0(z) \cdot F1(z) + F0(-z) \cdot F1(-z) = 2 \quad (53)$$

式(53)は双直交ウェーブレット変換における「完全再構成フィルタバンク」と同一条件である。また、フィルタが本式の条件を満たせば、同様に $Y_1(z)$ と $Y_2(z)$ 、 $Y_2(z)$ と $Y_3(z)$ 、そして $Y_{N-1}(z)$ と $Y_N(z)$ が全て等しくなるので、4:2:2/4:2:0 相互変換の誤差の累積を抑止することが可能となる。

6.3.3 4:2:2/4:2:0 の色差画素の位置関係に基づく位相条件

H.264/AVC や SMPTE 等の規格では、4:2:2 と 4:2:0 の両フォーマットの色差画素の位置関係を規定しており、従来装置のアップサンプリングフィルタおよびダウンサンプリングフィルタは、この規定に沿った画素シフト(群遅延)に対応している。ND420 フィルタと従来装置と「色ずれ」を発生させず相互接続を実現するためには、この画素シフトの条件を満たす必要がある。

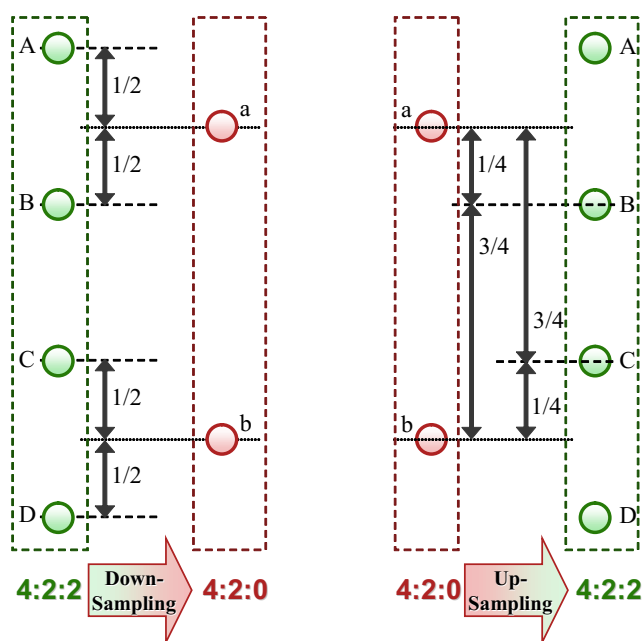


図 55 プログレッシブ映像における 4:2:2 と 4:2:0 の
色差画素の位相関係 [83]

図 55 は、H.264/AVC 等で規定されたプログレッシブ映像における 4:2:2 と 4:2:0 の色差画素の垂直方向の位相関係を示す図である。4:2:2 から 4:2:0 へのダウンサンプリングでは、4:2:0 の画素「a」の位置は、4:2:2 の画素「A」の位置の 1/2 画素だけ下の場所に存在する。

また、4:2:0 から 4:2:2 へのアップサンプリングにおいては、4:2:0 の画素「a」および「b」から、二つの 4:2:2 の画素「B」および「C」が補間される。まず、画素「B」の位置は、4:2:0 の画素「a」の位置の 1/4 画素だけ下の場所に存在する。また、同様に画素「C」の位置は画素「a」の位置より 3/4 画素だけ下の位置となる。

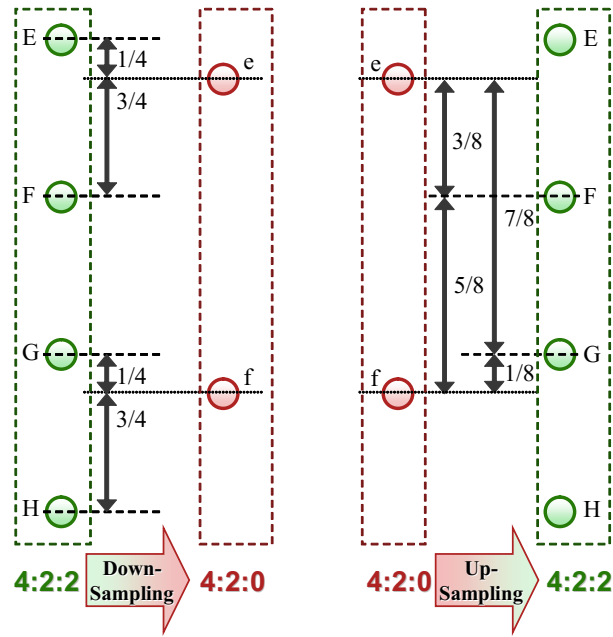


図 56 インターレース (Top field) における 4:2:2 と 4:2:0 の色差画素の位相関係 [83]

図 56 は、H.264/AVC 等で規定されたインターレース (トップフィールド) における 4:2:2 と 4:2:0 の色差画素の位相関係を示したものである。4:2:2 から 4:2:0 のダウンサンプリングでは、4:2:0 における画素「e」の位置は、4:2:2 の画素「E」の位置の $1/4$ 画素だけ下の位置となる。ここで、有限タップ長のフィルタでは、完全な線形位相は対称フィルタでのみ可能で、位相シフトも 0 および $1/2$ だけしか実現できない。このことから、ダウンサンプリングフィルタは、通過域で位相シフトが $1/4$ の線形位相に近くなるよう設計する必要がある。

4:2:0 から 4:2:2 へのアップサンプリングでは 4:2:0 の画素「e」および「f」から二つの 4:2:2 の画素「F」および「G」が補間される。4:2:2 フォーマットにおける画素「F」の位置は、画素「e」の位置の $3/8$ 画素だけ下の位置であり、画素「G」の位置は画素「e」より $7/8$ 画素だけ下の位置となる。

このように、設計すべき ND420 フィルタは、上記の画素シフト(位相特性)を有することが必要となる。

6.3.4 フィルタの設計方法

本小節では、フィルタ係数の設計方法について説明する。我々は、フィルタ設計の自由度を高めるために、ダウンサンプリング、アップサンプリングの両フィルタ係数のタップ長、すなわち、式(45)及び式(46)のフィルタタップ数 N を 8 とした。

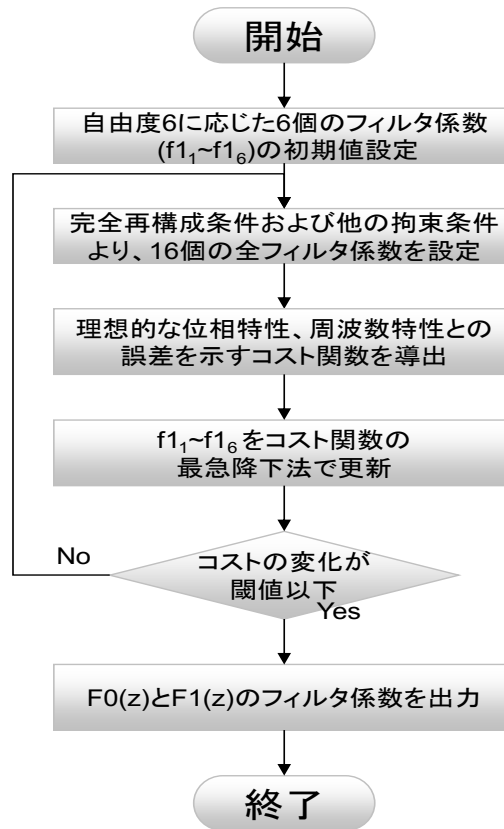


図 57 フィルタ係数導出のフローチャート

フィルタ設計のフローチャートを図 57 に示す。このフローチャートの動作は次の通りである。

$F0(z)$ 、 $F1(z)$ 双方のタップ数が 8 なので、フィルタ係数は合計 16 個ある。しかし、フィルタ係数には、次に示す 10 種類の拘束条件があるため、自由度は 6 となる。

最初の拘束条件は式(53)の完全再構成条件である。これは、 z の恒等式となるため、-3 から 3 の整数を取るパラメータ l を用いて表現される次の 7 個の条件式と等価となる。ここで $\delta_{l,0}$ はクロネッカーのデルタである。

$$\sum_{k=\max(2l,0)}^{\min(7+2l,7)} f0_k \cdot f1_{k-2l} = 2 \cdot \delta_{l,0}, \quad l = [-3,3] \quad (54)$$

また、 z 変換の定義より $z=e^{j\omega}$ である。ここで、 $F0(z)$ 、 $F1(z)$ とも、 $\omega=0$ 、すなわち $z=1$ では全通過、すなわち

$$F0(1)=1, \quad F1(1)=2 \quad (55)$$

となる必要がある。ここで $F1(z)=2$ の理由は、このフィルタが 2:1 に間引かれた信号に対して適用されるためである。

また、 $\omega=\pi$ 、すなわち $z=-1$ で $F0(z)$ 、 $F1(z)$ とも完全阻止、すなわち

$$F0(-1)=F1(-1)=0 \quad (56)$$

となる必要がある。ここで、完全再構成条件と $F0(-1)=0$ 、 $F1(1)=2$ 、および $F1(-1)=0$ から $F0(1)=1$ は導出されるので、独立な式は $F0(-1)=0$ 、 $F1(1)=2$ 、 $F1(-1)=0$ の 3 個となる。これに基づき、係数に関しては、更に次の 3 個の条件式を満たす必要がある。

$$\sum_{k=0}^7 f0_k \cdot (-1)^k = 0 \quad (57)$$

$$\sum_{k=0}^7 f1_k = 2 \quad (58)$$

$$\sum_{k=0}^7 f1_k \cdot (-1)^k = 0 \quad (59)$$

上記のことより、16 個の係数 $f0_k$ および $f1_k(k=[0,7])$ に対し、拘束条件は式(54)及び式(57)~(59)の 10 個となるので、自由度は 6 となる。すなわち、6 個の係数、例えば $f1_1 \sim f1_6$ の 6 個の係数が求まれば、残り 10 個の係数すべてを決定することが可能となる。そして、フィルタの好ましい条件からの乖離をコストとし、上記 6 個の係数を変数として最急降下法で最低コストを与える係数を求めれば、好ましいフィルタ係数を求めることができる。

次に、コスト関数について定義する。コスト関数は、位相特性、振幅特性、そして主観品質に影響を与えるステップ応答時のオーバー/アンダーシュート特性の 3 種類からなる。

まず、位相特性のコスト関数については、次に述べる 3 種を定義する。まず $F0(z)$ の位相特性のコストを式(60)で定義する。

$$Cost1 = \int_0^{\pi/2} \left| \frac{d}{d\omega} \left(\arctan \left(\frac{\text{Im}(F0(e^{j\omega}))}{\text{Re}(F0(e^{j\omega}))} \right) \right) - \theta_{F0} \right|^2 d\omega \quad (60)$$

ここで関数 $\text{Im}()$ は複素数の虚数成分、関数 $\text{Re}()$ は複素数の実数成分を求める関数である。これらの関数を用いてフィルタの各周波数の群遅延を求め、これと目標値との差分をコストとする。また、変数 θ_{F0} は $F0(z)$ の理想的な位相を示し、6.3.3 小節に示したとおり、プログレッシブの場合には 1/2、インターレースのトップフィ

ードの場合は 1/4 となる。また、積分範囲は、フィルタの通過域である $[0, \pi/2]$ とした。

また、アップサンプリングフィルタにおいて、偶数係数、奇数係数のみからなるフィルタ $F_{1even}(z)$ 、 $F_{1odd}(z)$ を次のように定義する。

$$F_{1even}(z) = \sum_{k=0}^3 f_{1_{2k}} \cdot z^{-k} \quad (61)$$

$$F_{1odd}(z) = \sum_{k=0}^3 f_{1_{2k+1}} \cdot z^{-k} \quad (62)$$

定義より、 $F_{1even}(z)$ 、 $F_{1odd}(z)$ と $F_1(z)$ 、は次の関係がある。

$$F_1(z) = F_{1even}(z^2) + z \cdot F_{1odd}(z^2) \quad (63)$$

これらの関数を用いて、 $F_{1even}(z)$ 、 $F_{1odd}(z)$ の位相特性のコストを次の通り表す。

$$Cost2 = \int_0^{\pi/2} \left| \frac{d}{d\omega} \left(\arctan \left(\frac{\text{Im}(F_{1even}(e^{j\omega}))}{\text{Re}(F_{1even}(e^{j\omega}))} \right) \right) - \theta_{F_{1even}} \right|^2 d\omega \quad (64)$$

$$Cost3 = \int_0^{\pi/2} \left| \frac{d}{d\omega} \left(\arctan \left(\frac{\text{Im}(F_{1odd}(e^{j\omega}))}{\text{Re}(F_{1odd}(e^{j\omega}))} \right) \right) - \theta_{F_{1odd}} \right|^2 d\omega \quad (65)$$

ここで、変数 $\theta_{F_{1even}}$ 及び $\theta_{F_{1odd}}$ は $F_{1even}(z)$ 及び $F_{1odd}(z)$ の理想的な位相とする。6.3.3 小節に示したとおり、プログレッシブの場合には $\theta_{F_{1even}}$ は 3/4、 $\theta_{F_{1odd}}$ は 1/4、インターレースのトップフィールドの場合には、 $\theta_{F_{1even}}$ は 7/8、 $\theta_{F_{1odd}}$ は 3/8 となる。また、積分範囲は、 $Cost1$ と同様、フィルタの通過域である $[0, \pi/2]$ とした。

次に、 $F_0(z)$ 、及び $F_1(z)$ に関する振幅特性のコスト関数として、次の式(66)、及び式(67)を定義する。それぞれ、通過帯域である $[0, \pi/2]$ では振幅の絶対値が 1、阻止帯域である $[\pi/2, \pi]$ では 0 とする。

$$Cost4 = \int_0^{\pi/2} \left(|F_0(e^{j\omega})| - 1 \right)^2 d\omega + \int_{\pi/2}^{\pi} |F_0(e^{j\omega})|^2 d\omega \quad (66)$$

$$Cost5 = \int_0^{\pi/2} \left(|F_1(e^{j\omega})| - 2 \right)^2 d\omega + \int_{\pi/2}^{\pi} |F_1(e^{j\omega})|^2 d\omega \quad (67)$$

次に、ステップ応答に対する相互変換後のオーバーシュート抑止のコストを導出する。まず、 $k < 0$ で 0、 $k \geq 0$ で 1 の値をとるステップ関数の z 変換を $U(z)$ とすると次の通りに表すことができる。

$$U(z) = \sum_{k=0}^{\infty} z^{-k} \quad (68)$$

このステップ関数に $F0(z)$ 、ダウン／アップサンプリング、 $F1(z)$ を適用して相互変換した後の z 変換を $G(z)$ 、その係数を g_k とすると、これらは、式(50)の $X(z)$ に $U(z)$ を代入することで下記のように表現できる。

$$G(z) = \sum_{k=-\infty}^{\infty} g_k \cdot z^{-k} = \frac{1}{2} \cdot (F0(z) \cdot U(z) + F0(-z) \cdot U(-z)) \cdot F1(z) \quad (69)$$

ここで、 g_k は、 $U(z)$ と $F0(z)$ 、 $F1(z)$ の z 変換表現式 (68)、式(46))を用いて本式の右辺を展開し、各 z^{-k} の係数を求めることで計算できる。そして、この g_k を用いて、オーバーシュート抑止のコストを次のように定義する。

$$\text{cost6} = \sum_{k=-\infty}^{-1} (\min(g_k, 0))^2 + \sum_{k=0}^{\infty} (\max(g_k - 1, 0))^2 \quad (70)$$

そして、全体のコストを、各コストの重みを w_k とし、次式で表す。

$$\text{cost} = w1 \cdot \text{cost1} + w2 \cdot \text{cost2} + w3 \cdot \text{cost3} + w4 \cdot \text{cost4} + w5 \cdot \text{cost5} + w6 \cdot \text{cost6} \quad (71)$$

この cost の値を $f1_1 \sim f1_6$ の 6 個の係数で数値微分し、最急降下法で cost と $f1_1 \sim f1_6$ の 6 個の係数を更新する。そして、 cost の変化量が閾値以下になった場合を最小値とみなして、その時の $f1_1 \sim f1_6$ から、残りの 10 個の係数を求め、それを最適フィルタとした。主観評価や周波数特性を顧慮し、重みの各値は $w1=w2=w3=16$ 、 $w4=w5=1$ 、 $w6=38$ とした。振幅特性に対する重み $w4$ 及び $w5$ は、係数における周波数特性の拘束条件 $F0(1)=1$ 、 $F1(1)=2$ および $F1(-1)=F0(-1)=0$ が強いため、相対的に小さい値を用いた。

6.3.5 設計したフィルタの係数

前小節の方法でプログレッシブ及びインターレースにおいて ND420 フィルタを設計した。

表 26 はプログレッシブにおける ND420 フィルタセットの係数を示す。表 26 (a) はダウンサンプリングフィルタ $F0(z)$ の係数である。 $z = 1(\omega = 0)$ におけるシフト量は正確に 0.5 となる。また、表 26 (b)は図 25 の画素 B 等の 1/4 画素の位相シフトに対応するアップサンプリングフィルタ $F1_{\text{odd}}(z)$ の係数を示す。本フィルタの $z = 1(\omega = 0)$ における画素シフトは 0.26863、すなわち $1.07 / 4$ であり、目標値の 1/4 にほぼ

等しい。また、図 25 の画素 C 等の 3/4 画素の位相シフトに対応するアップサンプリングフィルタ $F_{1even}(z)$ は、表 26 (b) のフィルタ係数を左右反転させたものとなる。

表 26 プログレッシブにおける ND420 フィルタセットの係数 [83]

(a) ダウンサンプリングフィルタ $F_0(z)$

-0.00255	-0.01852	0.03348	0.48759	0.48759	0.03348	-0.01852	-0.00255
----------	----------	---------	---------	---------	---------	----------	----------

(b) アップサンプリングフィルタ $F_{1odd}(z)$

-0.13205	1.01364	0.10023	0.01817
----------	---------	---------	---------

表 27 はインターレースのトップフィールドにおける ND420 フィルタセットの係数を示す。表 27 (a) は図 56 の 4:2:2/4:2:0 ダウンサンプリングにおいて、4:2:0 の画素「e」を生成するための 1/4 画素の位相シフトに対応するダウンサンプリングフィルタ $F_0(z)$ の係数である。 $z = 1(\omega = 0)$ におけるこのフィルタの画素シフト量は 0.22649、すなわち $0.91/4$ である。これは目標値の 1/4 とほぼ同じ値となる。

表 27 (b) は図 56 の 4:2:0/4:2:2 アップサンプリングにおいて、4:2:2 の画素「F」を生成するための 3/8 画素の位相シフトに対応するアップサンプリングフィルタ $F_{1odd}(z)$ の係数である。 $z = 1(\omega = 0)$ におけるこのフィルタの画素シフト量は 0.38293、すなわち $3.06/8$ である。これは目標値の 3/8 とほぼ同じ値となる。

表 27 (c) は図 56 の 4:2:0/4:2:2 アップサンプリングにおいて、4:2:2 の画素「G」を生成するための 7/8 画素の位相シフトに対応するアップサンプリングフィルタ $F_{1even}(z)$ の係数である。 $z = 1(\omega = 0)$ における画素シフト量は 0.89059、すなわち $7.12/8$ である。これは目標値の 7/8 とほぼ同じ値となる。

ボトムフィールドにおいては、これらすべての係数を左右反転させたものを用いる。

表 27 インターレース (Top Field) における ND420 フィルタセットの係数 [83]

(a) ダウンサンプリングフィルタ $F_0(z)$

-0.01292	-0.03282	0.14539	0.57316	0.37762	-0.04072	-0.01008	0.00038
----------	----------	---------	---------	---------	----------	----------	---------

(b) アップサンプリングフィルタ $F_{1odd}(z)$

-0.10563	0.82482	0.28434	-0.00352
----------	---------	---------	----------

(c) アップサンプリングフィルタ $F_{1even}(z)$

0.0416	-0.06629	1.11721	-0.09251
--------	----------	---------	----------

6.4 評価

本節では、開発した ND420 のインターレースフィルタの評価について説明する。インターレースの評価が特に重要な理由は次の通りである。

インターレース用のフィルタで要求される位相は前述のとおり 0 か 1/2 でないため、有限タップ長で理想的な線形位相特性を実現することは理論上不可能である。

また、インターレースでは、各トップ・ボトム各フィールドが 1 ラインおきに表示されるため、フィルタの視覚上の影響範囲がプログレッシブに比べて大きくなる。

本節では、上記の理由で画質がセンシティブなインターレースについて、従来手法で設計したフィルタと ND420 フィルタの性能評価を行った。

6.4.1 比較のための参照フィルタセット

ND420 フィルタと比較する参照フィルタとして、性能のよいフィルタ設計手法として知られる Lanczos2 法 [85]を用い、良好なカットオフ周波数特性と位相特性を有する係数(表 28)を設計した。

表 28 インターレース(Top Field)における参照フィルタセットの係数 [83]

(a) ダウンサンプリングフィルタ $F_0(z)$

-0.01855	-0.02930	0.19727	0.47754	0.36621	0.04883	-0.04004	-0.00195
----------	----------	---------	---------	---------	---------	----------	----------

(b) アップサンプリングフィルタ $F_{1odd}(z)$

-0.13281	0.78125	0.44531	-0.09375
----------	---------	---------	----------

(c) アップサンプリングフィルタ $F_{1even}(z)$

-0.02734	0.11328	0.99219	-0.07813
----------	---------	---------	----------

まず表 28 (a)のダウンサンプリングフィルタ $F_0(z)$ については、 $z=1$ における画素シフトは $0.2529=1.011/4$ であり、これは理想値の 1/4 とほぼ等しい。

また、表 28 (b)の 3/8 画素シフトのためのアップサンプリングフィルタ $F_{1odd}(z)$ の $z=1$ の画素シフトは $0.3906=3.125/8$ であり、これは理想値の 3/8 にほぼ等しい。

同様に表 28 (c)の 7/8 画素シフトのためのアップサンプリングフィルタ $F_{1even}(z)$ の $z=1$ の画素シフトは $0.8633=6.906/8$ であり、これは理想値の 7/8 にほぼ等しい

図 58 は、ND420 フィルタ、参照フィルタ、TM5 [86]の各ダウンサンプリングフィルタの振幅の周波数特性を示す。横軸は周波数を円周率 π で割った規格化された周波数である。カットオフ周波数は横軸 0.5 である。ND420 フィルタは参照フィルタと比べて阻止域である高域で十分にカットしきれていないものの、MPEG-2 の

テストモデル TM5 で定義されたダウンサンプリングフィルタとほぼ同等であり、許容範囲と考える。

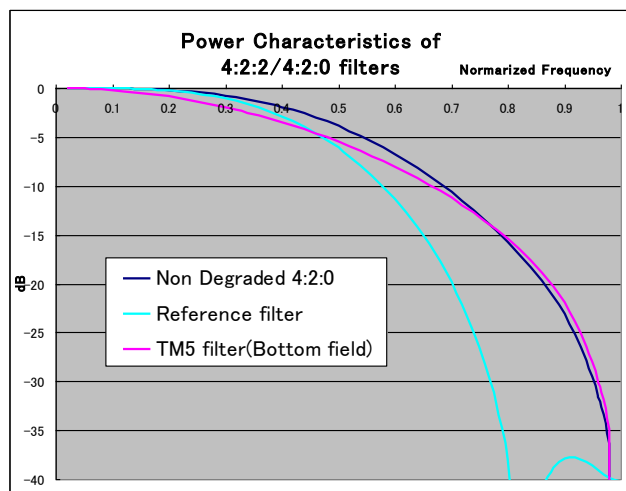


図 58 ND420、参照、TM5 の各ダウンサンプリングフィルタの周波数特性 [83]

6.4.2 シミュレーションによる画質評価

参照フィルタセットと ND420 フィルタセットを用いて、4:2:2/4:2:0 相互変換を所定の回数だけ繰り返した後に、PSNR および主観画質を比較した。

4:2:0 映像伝送装置の多段中継で 4:2:2/4:2:0 相互変換を繰り返した場合、鮮明な物体の色境界や CG 等の文字など、特に色のエッジ近辺で劣化がより目立つと言われている。したがって、ND420 フィルタの評価の映像シーケンスは、著者らが作成した”Characters”(HD)、CCETT によって作成された”Mobile & Calendar”(SD) [87]、そして RAI によって作成された”F1 car”(SD) [88]を用いた(図 59)。なお、後者の二つのシーケンスは、VQEG(Video Quality Expert Group)が広く映像評価の実験に活用できるように、著作権の問題をクリアして一般に公開しているものである。



(a) “Characters”

(b) “Mobile & Calendar”

(c) “F1 Car”

図 59 ND420 の性能評価に用いたシーケンス(VQEG より掲載許可) [83]

図 60 は、第 1 目の 4:2:2/4:2:0 相互変換の 4:2:2 出力の色差 U/V の変換前画像に対する PSNR を基準とし、第 2、4、8、16 回目の相互変換後の PSNR を求め、1 回

目との差分を示したものである。

参照フィルタセットを用いた場合、相互変換の回数が増えると誤差が累積し劣化が増大する。これに対し、ND420 フィルタセットを用いると劣化がほとんど発生しない。Characters シーケンスでのみ、変換後の画素値を 0-255 にクリッピングすることに伴う若干の劣化が 4 回目まで見られるが、8 回目移行は PSNR の更なる劣化は見られない。

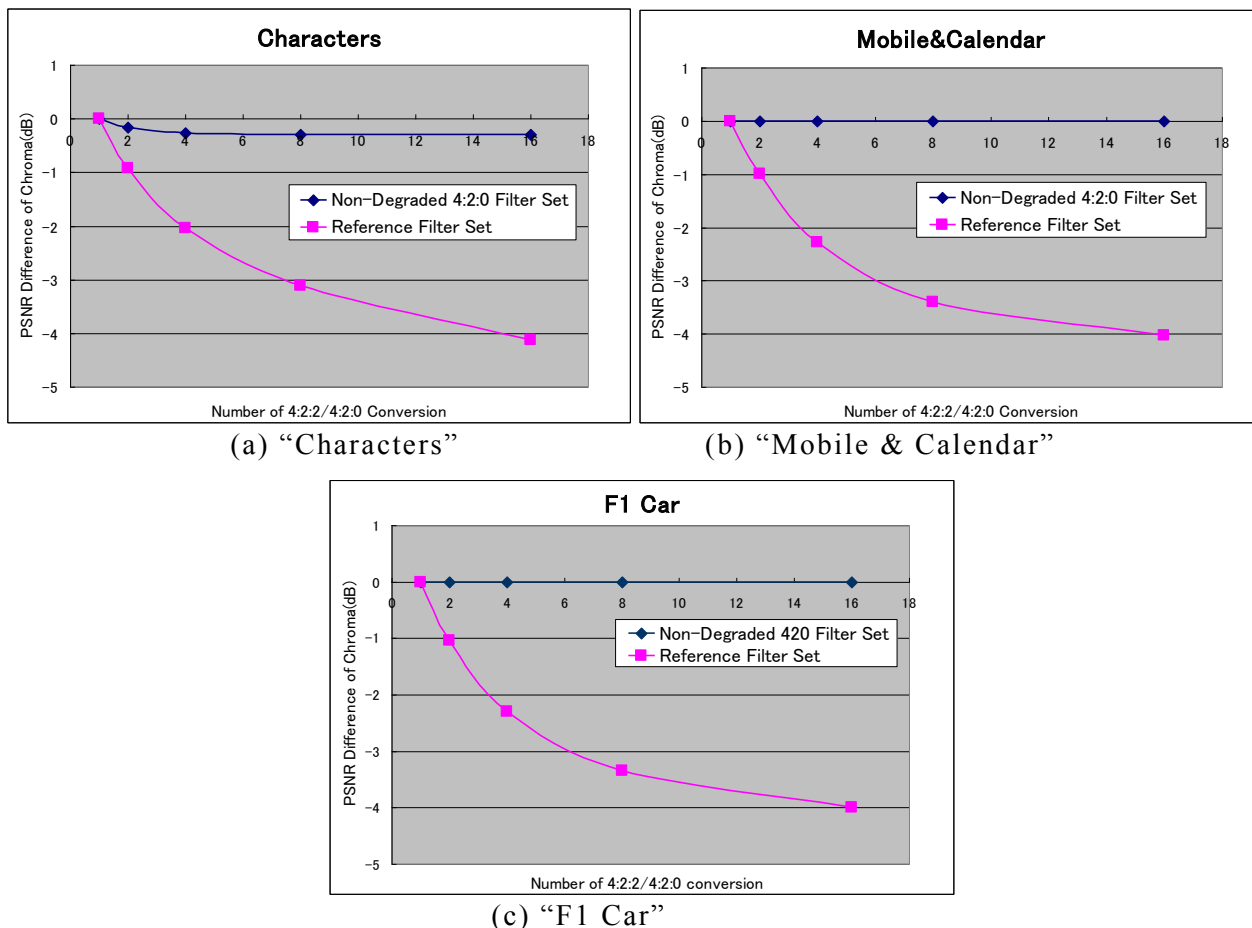
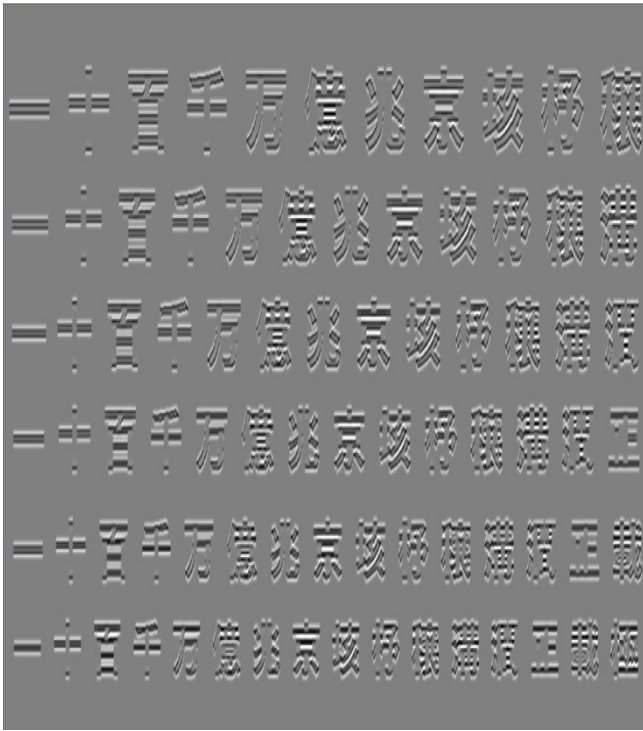


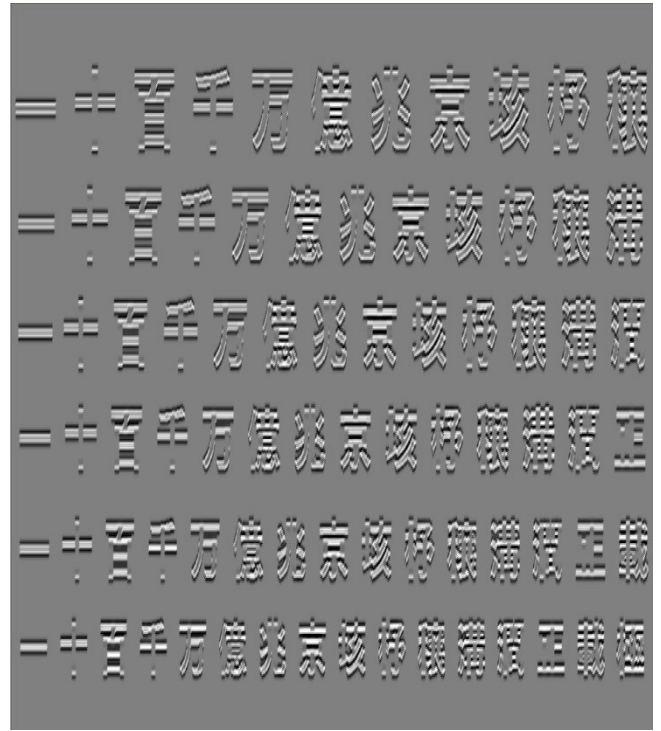
図 60 4:2:2/4:2:0 相互変換の回数と PSNR 劣化の関係 [83]

図 61 は 1 回目と 16 回目の 4:2:2/4:2:0 相互変換の色差画像(U/V 両成分)の差分を示す。灰色(画素値 128)を差分無しとし、差分をより見やすくするため、画素毎に差分値を 2 倍した値を加えた。

参照フィルタセットを使用すると、1 回目と 16 回目の画像間の際は明確である。一方、ND420 フィルタセットを用いた場合には、1 回目と 16 回目の画像間の差は非常に小さい。"Character"シーケンスでは、リングングが発生した画素の値の 0-255 クリッピングに起因する差がわずかに見られるものの、参照フィルタセットに比べると、劣化の度合いは遥かに小さい。



U 成分画像

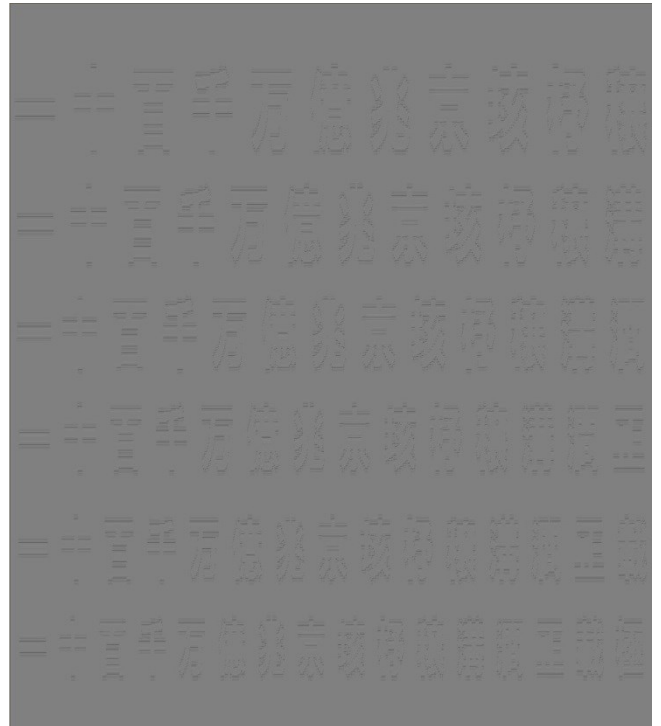


V 成分画像

図 61- (a)-1 参照フィルタセット



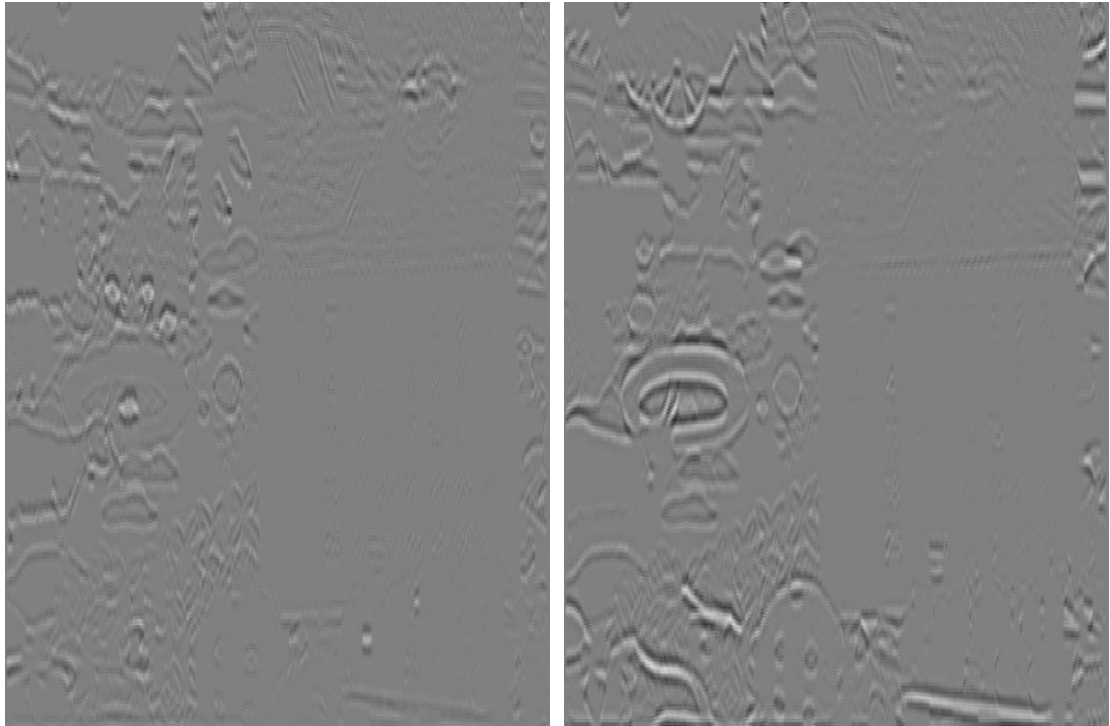
U 成分画像



V 成分画像

図 61- (a)-2 ND420 フィルタセット

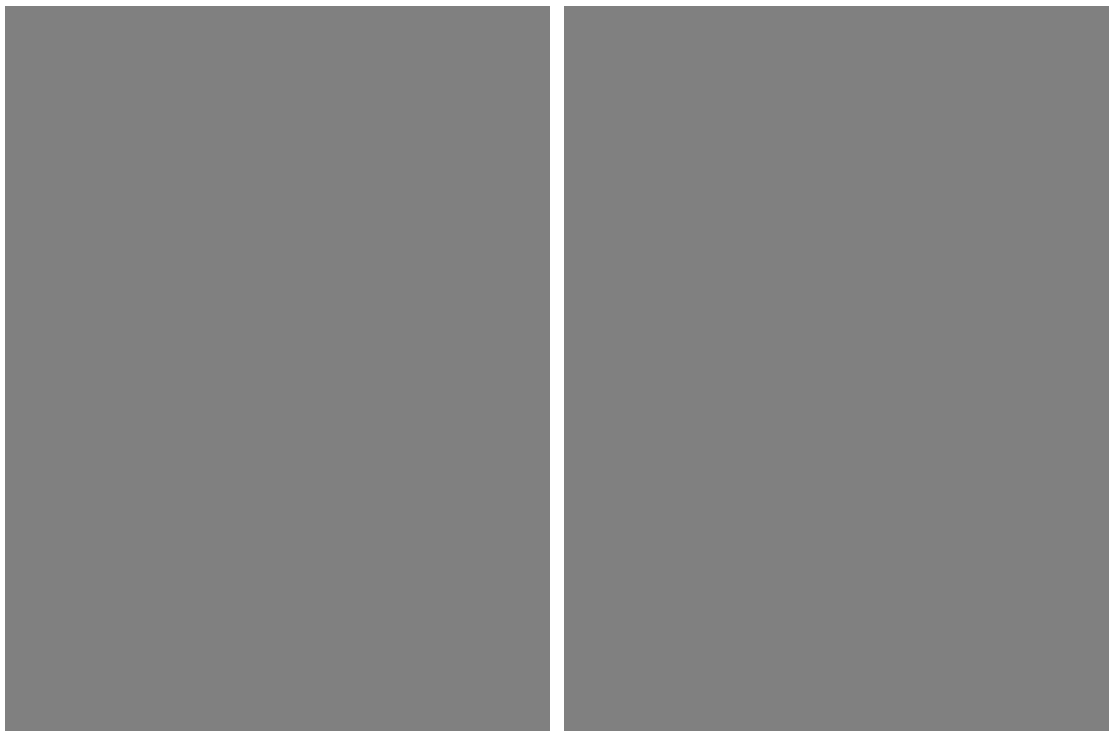
図 61- (a) “Character”の 1 回目と 16 回目の相互変換後の色差画像の差分



U 成分画像

V 成分画像

図 61- (b)-1 参照フィルタセット



U 成分画像

V 成分画像

図 61- (b)-2 ND420 フィルタセット

図 61- (b) “Mobile & Calendar”の 1 回目と 16 回目の相互変換後の色差画像の差分

図 62 は参照フィルタセットと ND420 フィルタセットを用い、1 回、4 回、16 回の相互変換を行った後の画像を示したものである。

参照フィルタセットを使用すると、三つのシーケンスのすべてで色の劣化が確認できる。”Character”においては、参照フィルタセットでは、変換回数が 4 回、16 回と増えるに従い、赤い文字が黒っぽく変化している。一方、ND420 フィルタセットでは 1 回目と 16 回目で大きな違いがみられない。同様に、”Mobile & Calendar”や”F1 Car”においては、参照フィルタセットを用いた場合に、4 回目と 16 回目の変換後に、それぞれ赤／緑境界や文字の回りに色が滲んでしまうことが分かる。一方、ND420 フィルタセットでは、変換回数が 4 回、16 回と増えても画質の変化は見られない。

なお、従来のフィルタ設計手法には様々なものがあり、映像の劣化程度はフィルタ係数に依存するが、完全再構成条件を満たさない限り、画質劣化を避けることはできない。

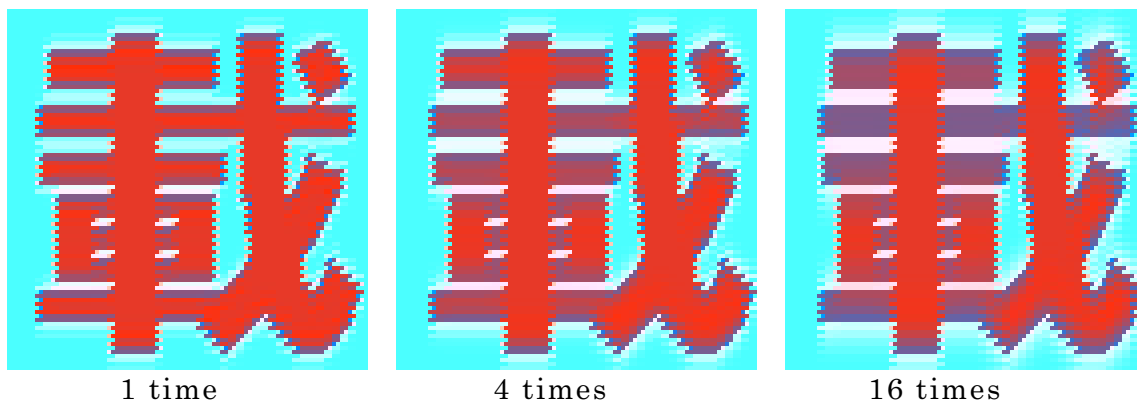


図 62- (a)-1 参照フィルタセット

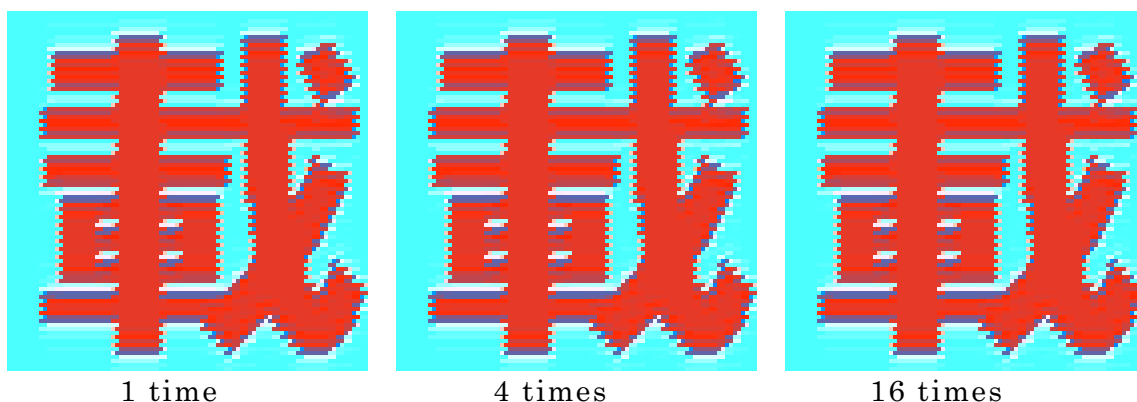


図 62- (a)-2 ND420 フィルタセット

図 62- (a) “Character”の相互変換後の画像比較 [83]

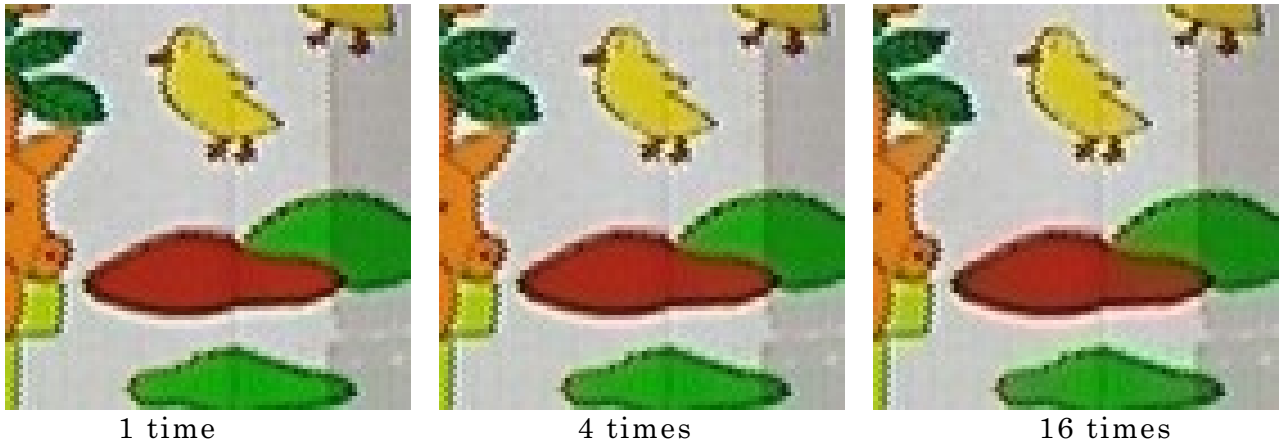


図 62- (b)-1 参照フィルタセット

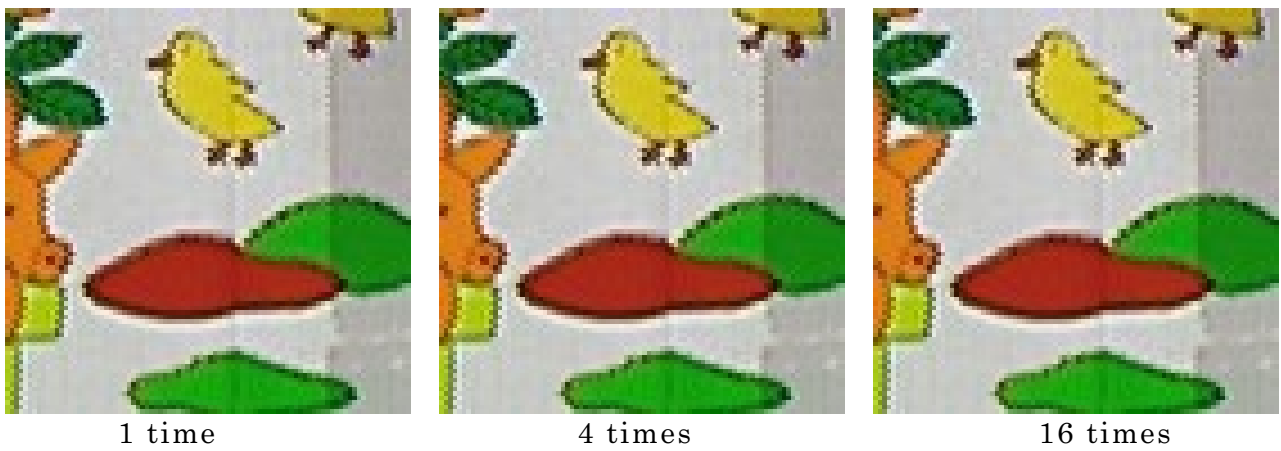


図 62- (b)-2 ND420 フィルタセット

図 62- (b) “Mobile & Calendar”の相互変換後の画像比較

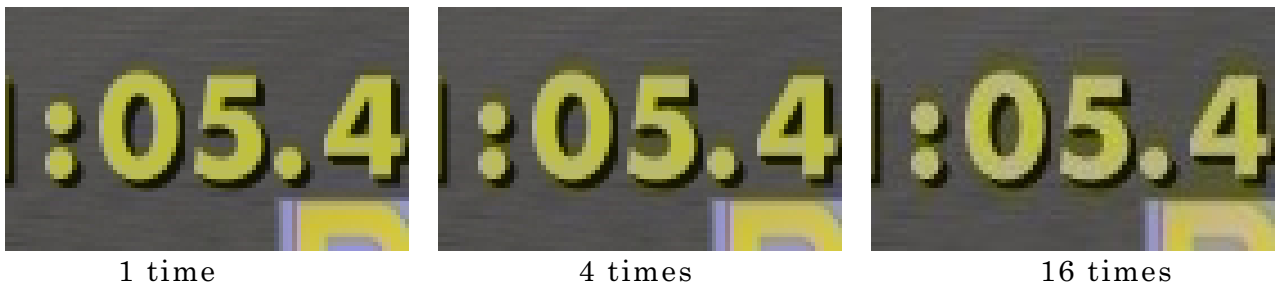


図 62- (c)-1 参照フィルタセット



図 62- (c)-2 ND420 フィルタセット

図 62- (c) “F1 Car”の相互変換後の画像比較

図 62 参照フィルタと ND420 フィルタによる相互変換後の画像比較 [83]



(1)Up: 参照フィルタ
Down: 参照フィルタ



(2)Up: ND420 フィルタ
Down: ND420 フィルタ



(3)Up: N420 フィルタ
Down: 参照フィルタ

(a) “Character”



(1)Up: 参照フィルタ
Down: 参照フィルタ

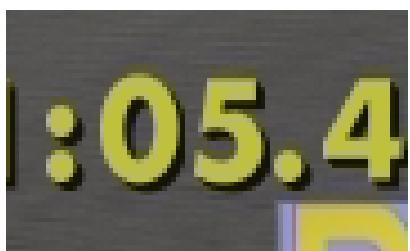


(2)Up: ND420 フィルタ
Down: ND420 フィルタ

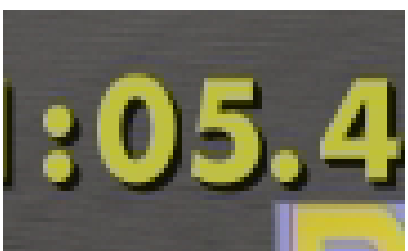


(3)Up: N420 フィルタ
Down: 参照フィルタ

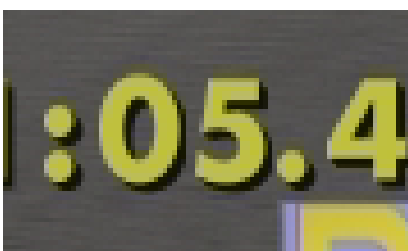
(b) “Mobile & Calendar”



(1)Up: 参照フィルタ
Down: 参照フィルタ



(2)Up: ND420 フィルタ
Down: ND420 フィルタ



(3)Up: N420 フィルタ
Down: 参照フィルタ

(c) “F1 Car”

図 63 アップサンプリング/ダウンサンプリングフィルタの組み合わせの比較 [83]

図 63 は、ND420 フィルタセットと従来フィルタセットを用いた映像伝送装置の相互接続時の画質を評価するために、次の 3 種類のアップサンプリング/ダウンサンプリングフィルタの組み合わせで 1 回の相互変換を行った後の画質を評価したものである。条件は次のとおりである。

- (1) ダウンサンプリングフィルタとアップサンプリングフィルタの両方とも参照フィルタセットを用いたもの。符号化装置、復号装置とも従来の映像伝送装置の場合に相当する。
- (2) ダウンサンプリングフィルタとアップサンプリングフィルタの両方とも ND420 フィルタセットを用いたもの。符号化装置、復号装置とも ND420 フィルタセットを用いた映像伝送装置の場合に相当する。
- (3) ダウンサンプリングフィルタは ND420 フィルタ、アップサンプリングフィルタは参照フィルタを用いたもの。これは、図 51 に示した、ND420 を用いた装置で符号化した映像を従来の復号装置で受信した場合に相当する。

これらの結果を見ると、全てのシーンにおいて、3 種類のフィルタの組み合わせの各映像の間には主観的に顕著な差がないことが分かる。すなわち、ND420 フィルタを用いた符号化装置で符号化されたデータを従来の復号装置で復号しても、1 回の復号に限り、色ずれ等は発生せず、主観的な差は生じない。このことから、ND420 フィルタを用いた装置と従来装置の混在運用が可能となり、SNG 等で柔軟な機器運用が可能となる。

6.5 本章のまとめ

4:2:0 フォーマットに対応した動画像符号化・復号装置で多段中継を行う場合、複数回の 4:2:2/4:2:0 相互変換が必要となり、色差成分の僅かな誤差が累積して主観的に知覚される劣化が生じることを示した。

この課題を解決するために、誤差累積を抑止する完全再構成条件と、従来機器との互換性を保証する画素の位相シフト条件を満たす 4:2:2/4:2:0 相互変換のフィルタセット(ND420 フィルタセット)を導出した。

ND420 フィルタセットを用いた場合、16 回までの多段接続で画質劣化の累積が見られないことを示した。また、従来のフィルタセットとの組み合わせにおいても、視覚的な劣化が見られないことも併せて示した。

そして、この技術の有用性が認められ、SMPTE RP2050 及び HEVC で、提案技術が規格に採用された。

これらのことから、提案する ND420 フィルタセットを用いた動画像符号化・復号装置は、SNG/FPU 等での多段中継を伴う実運用で、非常に有用であると結論できる。

第7章 結論

7.1 本研究の成果

本論文では、動画像符号化の性能改善技術の研究を第2章、第3章、第4章で、国際標準化 H.264/AVC 対応の符号化装置を実用化する際の画質・機能向上技術の研究を第5章及び第6章で、それぞれ論じた。

動画像符号化の性能改善技術の研究として、まず第2章では、TV電話等で動きの激しいシーンでのコマ落ちを抑制するための動的解像度変換技術の研究について論じた。国際標準 H.263 は、ISDN や 3G 無線ネットワーク等の 64kbps 程度の回線を介したテレビ電話が重要な適用先の一つである。従来、テレビ電話で動画像を送信する際に人物の大きな動きが生じると、過度なコマ落ちの発生が問題となっていた。これに対し、動きの激しいシーンを検出し、予測誤差信号を 1/2 倍の低解像度に変換して符号化する動的解像度変換技術を開発した。本方式では、実装のしやすさを考慮し、直交変換／逆直交変換、量子化／逆量子化、可変長符号等は従来の規格をそのまま流用可能な構成とした。これにより、静止背景の画質や PSNR をほぼ保ったまま、フレームあたりの情報量を最大 1/3 程度まで削減し、コマ落ちを抑制することが可能となった。

第3章では、H.264/AVC に向けた動きベクトルの符号化効率改善を目的として、コンテキストベースのエントロピー符号化技術、及び Interlace 符号化時の色差ベクトル補正技術について論じた。H.263 や MPEG-4 までの国際標準化では、動きベクトルの符号化は、周辺の符号化済動きベクトルから求めた予測ベクトルとの差である差分ベクトルに対して、一意の可変長符号が割り当てられていた。これに対し、差分ベクトルの確率分布は、対象ブロック周辺の動きベクトル場が一様な場合には 0 近辺の確率が高く、ベクトル場が乱雑な場合には 0 近辺の確率が低くなることを見出した。この特性に基づき、動きベクトル場の一様性に依じて 2 種類の可変長符号を切り替えて差分ベクトルを符号化するコンテキストベースのエントロピー符号化技術を実現し、動きベクトル情報を約 10% 削減させることが可能となった。

また、2002 年 8 月の H.264/AVC 勧告草案においては、輝度の動きベクトルは 1/4 画素精度として規定されており、色差の動きベクトルは輝度の動きベクトルを正確に 1/2 にスケールリングして 1/8 画素精度で動き予測に用いることが規定されていた。しかし、Interlace 符号化時においては、動きベクトルが符号化画像と異なるパリティのフィールドを指し示す場合、すなわち符号化画像がトップフィールドで予測画像がボトムフィールド、あるいは符号化画像がボトムフィールドで予測画像がトップフィールドの場合には、輝度ベクトルと色差ベクトルが平行にならず、1/4 画素

ずれてしまうことを見出した。この課題に対し、符号化対象画像と予測画像のパリティに応じて、 $\pm 1/4$ 画素の補正を色差ベクトルに加える色差ベクトル補正技術を考案し、IDRを含むフレームのBottomフィールドで最大1dBの色差画像のPSNRの改善が得られた。

第4章では、直交変換係数の正負符号の予測符号化技術について論じた。H.264/AVC以降の動画画像符号化標準は、様々なシンボルについて空間・時間相関を用いた情報量削減が試みられてきた。しかし、直交変換・量子化された非ゼロ有効係数の正負符号は全体の10%の情報量を占めるにも関わらず、これまでは他の情報と無相関と考えられ、非圧縮で1ビット/符号の情報が割り当てられてきた。これに対し、符号化対象ブロックと、これに隣接する符号化済の周辺ブロックとの相関性に着目した正負符号の予測符号化を実現した。提案方式では、まず正負符号の全ての組み合わせに対して逆直交変換・局部復号化し、周辺ブロックの境界画素と空間相関が最も高くなるものを正負符号の予測符号とする。そして、予測符号と実際の符号が一致したか否かの情報を算術符号の一種であるCABACでエントロピー符号化することで情報量の削減を可能とした。ここで、Gray符号を活用した相関計算の高速化手法も併せて実現した。提案手法をH.264/AVCに適用したところ、正負符号の情報量の約10%に相当する全体の符号化情報量の約1%を削減することが可能となった。

第5章及び第6章では、動画画像システムの実用化や機能・性能向上のための研究について論じた。

第5章では、H.264/AVC動画画像伝送装置の実用化技術の研究について論じた。2000年台前半、従来のSD映像より圧倒的にリアルで高品位なHD映像の放送拡大やテレビ・カメラ・ビデオ・PC等でのHD映像の記録・表示の要望が高まっていた。しかしHD化に伴う機器・伝送コスト増大が大きな課題となっており、H.264/AVCの早期実用化が望まれていた。このような背景の元、H.264/AVCの符号化時に、動画画像の時空間特性を人間の視覚特性に基づいてリアルタイムで解析し、最適な情報量制御を行うことで、高い主観品質を保ちつつ、低演算量で符号化するアルゴリズムを実用化した。また、H.264/AVCを低遅延で符号化する技術、及び、衛星に比べ遥かに安価なインターネットでプロ用途の動画画像伝送を可能とするためにIPパケットの誤り訂正技術も併せて実用化した。そして、これらの技術を映像伝送装置IP-9500に適用し、世界中で活用されることでHD映像の記録・伝送コストが低下し、映像のHD化に貢献した。

第6章では、4:2:2/4:2:0の両映像フォーマットの間を繰り返し相互変換した際に生じる色差成分の劣化を抑制する色にじみ抑止技術の研究について論じた。HDで8Mbps程度の比較的遅レートを対象としたH.264/AVC動画画像伝送装置では民生

LSI 流用で安価な実装が可能な 4:2:0 の映像フォーマットが広く採用されている。一方、動画像伝送装置の入出力は、4:2:0 に対し 2 倍の色成分の垂直解像度を持つ 4:2:2 映像フォーマットを有する HD-SDI が用いられる。そのため、4:2:0 対応の装置を用いて複数拠点経由で中継する際には、中継地点で 4:2:2 と 4:2:0 の両映像フォーマット間の相互変換が必要となる。1 回の相互変換で生じる色成分の劣化は微小であり、ほぼ知覚されない。しかし、放送現場では最大 5 回程度の中継が行われるため、この僅かな劣化誤差が累積し、「色にじみ」として画質劣化が知覚されてしまうことが問題となっていた。筆者らは 4:2:2/4:2:0 映像フォーマット間の相互変換の画素補間フィルタの最適化により、画質劣化を引き起こす「色にじみ」と、従来機器との互換性を阻害する「色ずれ」を同時に抑制することが可能な『色にじみ抑止技術』を世界で初めて開発した。本技術により、従来技術で問題となっていた複数回の相互変換後の色にじみを完全に抑制することができた。

このように、本研究を通じて、動画像符号化国際標準の符号化効率の改善、および H.264/AVC 符号化技術の実用化に関して、研究成果を達成することができたと考えている。

7.2 研究成果の国際標準化への貢献

本節では、本論文で論じた提案技術の国際標準化への貢献について説明する。

表 29 に、本論文で論じた技術と対応特許、そして、MPEG-LA 社 [89]における各動画像符号化標準での必須認定との関係を示す。この MPEG-LA 社は、動画像符号化標準や多重化等、マルチメディア技術に関する各種デジュール・デファクト規格の基本特許の管理会社として最も実績がある。技術開発を行う企業にとって、開発した技術の特許が MPEG-LA 社に認定されることが重要であるとみなされている。

表 29 本論文で論じた提案技術と MPEG-LA 認定必須特許の関係 [89]

技術名称	日本特許番号	MPEG-LA 必須特許認定			
		MPEG-4	H.264/ AVC	H.265/ HEVC	VC-1
動的解像度変換技術(第 2 章)	特許第 3466032 号 [41]	○			
動きベクトルのコンテキスト切り換え符号化技術(第 3 章)	特許第 4142180 号 [52]		○		○
Interlace 符号化時の色差ベクトル補正技術(第 3 章)	特許第 3791922 号 [54]		○		
4:2:2/4:2:0 相互変換時の色にじみ抑止技術(第 6 章)	特許第 5003620 号 [90]			○	

まず、第 2 章で論じた動的解像度変換符号化技術は、“Reduced Resolution Update Mode”として、H.263 Annex Q 及び MPEG-4 ARTS Profile に採用されるとともに、本技術に関する特許 [41]が MPEG-LA 社の MPEG-4 の必須特許として認定された。また、第 3 章で論じた動きベクトルのコンテクストベースのエントロピー符号化については、本技術に関する特許 [52]が、MPEG-LA 社が管理する国際標準特許プールの H.264/AVC 及び VC-1(Windows Media)の必須特許として認定されている。

そして同じく第 3 章で論じた Interlace 符号化時の色差ベクトル補正技術は 2002 年 10 月の 10 月会合で提案した後に H.264/AVC に採用され、本技術に関する特許 [54]が同じく MPEG-LA 社特許プールの H.264/AVC 必須特許として認定されている。

さらに第 6 章で論じた色にじみ抑止技術は、全米テレビ映画技術者協会 SMPTE 規格 RP2050 として制定されるとともに [91]、H.265/HEVC にも採用され、本技術に関する特許 [90]が必須特許として認定されている。

また、第 4 章で論じた直交変換係数の正負符号の予測符号化技術については、今後、H.266 への提案も視野に検討を進めている。

7.3 研究成果の産業への貢献

第 5 章で論じた H.264/AVC の実用化技術、及び第 6 章で論じた色にじみ抑止技術を適用 [91]した映像伝送装置 IP-9500 [30]や IP-900 [92]は、国内外の多数の放送局で採用され [93]、放送の HD 化に大きく貢献した。5.4 節に紹介した事例も含め、IP-9500 が産業へ貢献した内容を説明する。

国内では、5.4.1 小節で説明したとおり、テレビ朝日をキー局とする ANN 系列が IP-9500 を用い、公衆インターネットの B フレッツを用いて、系列局間の映像素材伝送システムを構築した [80] [81]。このシステムにより、系列局間で 24 時間 365 日の映像素材交換を低コストで行うことができるようになった。

また、5.4.2 小節で示した通り、映像伝送装置 IP-9500 は、フジテレビの公衆インターネットによるお天気カメラや、日本テレビの NFL スーパーボウルの日米中継 [82]、J-Sports による 2007 年ラグビーワールドカップの HD 中継 [94]でも使われている。これらの実績は、ニューメディア社が行った放送局の映像伝送装置導入に関するアンケート(2010 年 3 月)において、55%の放送局が IP-9500 を使用しているとの回答 [95]に結びついている。

また、海外においても、CBS が映像伝送装置 IP-9500 を導入し [96]、2008 年の Bush 大統領の一般教書演説でのワシントン DC とニューヨークの中継で使われるなど [97]、多数の放送局で活用されている。

さらに、さらなる映像の高精細化への期待に応えるべく、IP-9500 に適用した高画質化技術をベースに、NHK 技研と共同で 8K 映像伝送装置を開発した [31] [98] [99]。

このように、本論文の第 5 章、第 6 章で論じた H.264/AVC の実用化技術を適用した映像伝送装置 IP-9500 は国内外で広く活用され、放送の HD 化に広く貢献した。このような成果が認められ、市村産業賞、文部科学大臣表彰、紫綬褒章等の複数の賞を受賞した。

7.4 今後の展望

今後の動画像符号化の方向性について、現時点の筆者の見解を示す。

現在、H.266 の動画像符号化標準化策定に向け、様々な研究機関が研究を行っている。しかし、2017 年 12 月時点で、H.266 に向けた基本検討で提案された技術は、かつて HEVC 標準化策定時に提案された技術の焼き直しがメインであり、それぞれ 0.5~1%程度のゲインの積み重ねとなっている。

一方、近年、急速に発展したディープニューラルネットワーク (以下 DNN) 技術を活用した動画像符号化の研究の数も増えつつある。DNN により、実世界の物体を反映した映像の確率分布を扱うことが可能となることから、直交変換やポストフィルタの置き換えが検討されている。これに加え、従来は実現が難しかったオブジェクト符号化も、物体認識と GAN/VAE 等を組み合わせることで実現される可能性がある。このように、今後は、DNN を活用した符号化技術への期待が高まるものと思われる。

今後は、上記のような DNN を活用した動画像符号化技術の研究を進め、更なる機能・性能向上に貢献したいと考える。また、同時に、DNN という学習ベースの技術をベースとした符号化技術の規格のあり方についても並行して検討を進めたい。

謝辞

本研究及び論文の執筆にあたり、主査をお引き受け頂き、ご多忙の中、丁寧にご指導頂いた甲藤二郎教授には、心から感謝いたします。

また、本論文審査の副主査をお引き受け頂いた亀山渉教授、渡辺裕教授、津田俊隆教授には、国際標準化等のご経験を踏まえ、論文構成の観点で適切なご指導を頂き、感謝いたします。

また、本論文の動画像符号化技術に関する研究は、筆者が富士通研究所に入社以来進めてきたものです。特に、入社以来、長年にわたりご指導頂いた松田喜一様、佐々木繁様、伊藤隆様、酒井潔様、森松英史様に感謝申し上げます。

そして、株式会社富士通研究所、富士通株式会社、株式会社ソシオネクスト、富士通九州ネットワークシステムズをはじめとする多くの皆様には、長年にわたり、本論文その他に関わる研究開発を共同で推進させて頂き、心から感謝いたします。

参考文献

- [1] W. Zschunke, "DPCM Picture Coding with Adaptive Prediction," IEEE Trans. Commun., Vol.25, No.11, pp.1295-1302, Nov. 1977..
- [2] N. M. Nasrabadi , R. A. King, "Image coding using vector quantization: a review," IEEE Transactions on Communications, Vol.36, No.8, pp.957-971, Aug. 1988.
- [3] Recommendation H.261, "Video codec for audiovisual services at p x 64 kbit/s," 1990.
- [4] ISO/IEC 11172-2:1993, "Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 2: Video," 1993.
- [5] ITU-T Recommendation H.262 | ISO/IEC 13818-2;1996, " - Information technology - Generic coding of moving pictures and associated audio information: Video," 1995.
- [6] ITU-T Recommendation H.263, "Video coding for low bit rate communication," 1996.
- [7] 1. ISO/IEC14496-2, "Information technology -- Coding of audio-visual objects -- Part 2: Visual," 1999.
- [8] Recommendation ITU-T H.264 | ISO/IEC 14496-10, "Advanced video coding for generic audiovisual services," 2003.
- [9] Recommendation ITU-T H.265 | ISO/IEC 23008-2, "High Efficiency Video Coding," 2013.
- [10] 榎本 忠儀, "動画像符号化プロセッサの歴史と将来展望," 電子情報通信学会論文誌 C, Vol.92-C, No.8, pp.477-487, 2009年8月.
- [11] Jonathan Koomey, Stephen Berard, Marla Sanchez , Henry Wong, "Implications of Historical Trends in the Electrical Efficiency of Computing," IEEE Annals of the History of Computing, Vol.33, No.3, pp.46-54, 2011.
- [12] "徹底レビュー「ISSCC2015の歩き方」(11):メモリ編:次世代大容量フラッシュと次世代高速 DRAM に注目," [オンライン]. Available: http://eetimes.jp/ee/articles/1501/19/news026_2.html.
- [13] "10TBの大容量HDDを実現した「SMR技術」登場の背景," [オンライン]. Available: <https://pc.watch.impress.co.jp/docs/column/semicon/723207.html>.
- [14] 酒井 潔, 栗田 昌徳 , 江口 勝博, "画像符号化 LSI," 雑誌 FUJITSU, Vol.54, No.1, pp52-56, 2003年1月.
- [15] P. A. Ruetz, P. Tong, D. Bailey, P. A. Luthi , P. H. Ang, "A high-performance full-motion video compression chip set," IEEE Transactions on Circuits and Systems for Video Technology, Vol.2, No.2, pp.111-122, June 1992.
- [16] 山影 明夫, 上谷 義治, 児玉 知也, 村岡 寛昭, 河原 暢彦, 武捨敏彦 , 稲垣 雄史, "1チップ MPEG-2 MP@ML ビデオエンコーダ LSI の開発," 信学技報, ICD98-51, 1998年6月.

- [17] 大塚 竜志, 酒井 潔, 浜野 崇, 渡辺 英明, 北浦 智靖, 西沢 美次, 江口 勝博, 佐別富 祥和, 山下 公彰, “1 チップ MPEG2 ビデオエンコーダ LSI の開発とシステム展開,” 信学技報, ICD98-50, 1998 年 6 月.
- [18] 宮越 英司, 荒木敏之, 佐山卓也, 大谷昭彦, 峯丸貴行, 岡本潔, 児玉久, 森繁孝行, 渡部彰啓, 三森敏正, 今西浩, 神保卓哉, 田中泰資, 谷山昌之, 新郷孝行, 福本義, 森本広頼, 青野邦年, 青木勝司, “128GOPS の動きベクトル検出器とマルチタスク RISC コントローラを内蔵した 100mm², 0.95W, 1 チップ MPEG2 MP@ML ビデオエンコーダ,” 信学技報, ICD98-5, 1998 年 4 月.
- [19] H. Iwasaki, J. Naganuma, Y. Nakajima, Y. Tashiro, K. Nakamura, T. Yoshitome, T. Onishi, M. Ikeda, T. Izuoka, M. Endo, “A 1.1 W single-chip MPEG-2 HDTV codec LSI for embedding in consumer-oriented mobile codec systems,” Proceedings of the IEEE 2003 Custom Integrated Circuits Conference, pp.177-180, Dec. 2003.
- [20] 西川 剛志, “MPEG-4 テレビ電話用 LSI,” 東芝レビュー, Vol.55, No.5, pp.54-57, 2000 年 5 月.
- [21] “HD 画面对応、低消費電力の H.264 方式映像処理 LSI を新発売,” 富士通株式会社プレスリリース, 2006 年 11 月 30. [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2006/11/30.html>.
- [22] “放送・プロ向けの素材伝送に対応可能な H.264 リアルタイムコーデック LSI を開発,” 日本電信電話株式会社プレスリリース, 2007 年 4 月 6 日. [オンライン]. Available: <http://www.ntt.co.jp/news/news07/0704/070406a.html>.
- [23] Hiroki Mizosoe, Daisuke Yoshida, Taku Nakamura, “A Single Chip H.264/AVC HDTV Encoder/Decoder/Transcoder System LSI,” IEEE Transactions on Consumer Electronics, Vol.53, No.2, pp.630-635, July 2007.
- [24] “ソニー初の民生用 4K カムコーダー ハンディカム™『FDR-AX1』を発売,” ソニー株式会社プレスリリース, 2013 年 9 月 5 日. [オンライン]. Available: <http://www.sony.jp/CorporateCruise/Press/201309/13-0905/>.
- [25] “世界初、プロ向け 4K H.265/HEVC リアルタイムエンコーダ LSI を開発,” NTT 持株会社プレスリリース, 2015 年 3 月 26 日. [オンライン]. Available: <http://www.ntt.co.jp/news2015/1503/150326a.html>.
- [26] “4K/60p HEVC リアルタイムビデオエンコーダ LSI サンプル出荷を開始,” 株式会社ソシオネクストプレスリリース, 2015 年 4 月 3 日. [オンライン]. Available: http://www.socionext.com/jp/pr/sn_pr20150403_01j.pdf.
- [27] 藤山 武彦, 梅崎 靖, 安井 哲也, “画像伝送装置: FEDIS シリーズ/Nexteye シリーズ,” 雑誌 FUJITSU, VOL.54, NO.1, pp.46-51, 2003 年 1 月.
- [28] “長野はネットワークオリンピック,” 株式会社インプレス R&D, インターネットマガジン, 1998 年 4 月号, pp.252-281, 1998 年 4 月.

- [29] “リアルタイム映像伝送装置「IP-7500」販売開始,” 富士通株式会社プレスリリース, 2004年7月1日. [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2004/07/1-1.html>.
- [30] 宮坂 秀樹, 溝口 美智子, 梅崎 靖, “IP ネットワーク対応 H.264/AVC コーデックシステム,” 雑誌 FUJITSU, Vol.58, No.2, pp.156-161, 2007年3月.
- [31] 境田 慎一, 井口 和久, 鹿喰 善明, 数井 君彦, 中川 章, “スーパーハイビジョン高効率映像符号化装置の開発,” 信学技報, SAT111(179), pp.61-66, 2011年8月.
- [32] D. Marpe, H. Schwarz and T. Wiegand, "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard," IEEE Trans. Circuits and Systems for Video Technology, Vol.13, No.7, pp.620-636, Jul. 2003.
- [33] 河合輝男, “テレビとビデオの画質,” 光学 第17巻9号, pp.440-447, 1988年9月.
- [34] “マンモスコーデックと呼ばれて,” 日経エレクトロニクス. 2004年1月19日号, pp.93-113, 2004年1月.
- [35] “Video Codec Test Model, Near-Term, Version 8 (TMN8),” ITU-T SG16/Q.15 Q.15-B-46, Sep.1997.
- [36] Jordi Ribas-Corbera, Shawmin Lei, “Rate Control in DCT Video Coding for Low-Delay Communications,” IEEE Trans. Circuits & Systems for Video Tech., Vol. 9, No.1, pp.172-185, Feb., 1999.
- [37] Bo Tao, Bradley W. Dickinson and Heidi A. Peterson, "Adaptive Model-Driven Bit Allocation for MPEG Video Coding," IEEE Trans. Circuits & Systems for Video Tech., Vol. 10, No.1, pp.147-157, Feb. 2000.
- [38] Jose I. Ronda, Martin Eckert, Fernando Jaureguizar, Narciso Garcia, “Rate Control and Bit Allocation for MPEG-4,” IEEE Trans. Circuits & Systems for Video Tech., Vol. 9, No.8, Dec., 1999.
- [39] Akira Nakagawa, Eishi, Morimatsu, Takashi Itoh, Kiichi Matsuda, “Dynamic Resolution Conversion Method for Low Bit Rate Video Transmission,” IEICE TRANS. COMMUN., Vol.E84-B, No.4, pp.930-940, April 2001.
- [40] Akira Nakagawa, Kimihiko Kazui, Eishi Morimatsu, Takashi Itoh, Kiichi Matsuda, “Prediction Error Dynamic Resolution scheme for low bitrate video coding,” MoMuC' 97, pp.576-579, Sept. 1997.
- [41] 中川 章, 数井 君彦, 森松 映史, “動画像符号化装置及び復号化装置”. 日本国特許 特許番号: 第 3466032 号.
- [42] 中川 章, 数井 君彦, 森松 映史, “画像符号化装置”. 日本国特許 特許番号: 第 3773585 号.
- [43] Akira Nakagawa, “Comparison of two Dynamic Resolution Conversion Schemes,” Document of ITU-T STUDY GROUP 15, LBC-96-306, Atlanta, US, November 1996.
- [44] Akira Nakagawa, “Proposed revision of Dynamic Resolution

- Conversion Schemes,” Document of ITU-T STUDY GROUP 15, LBC-97-020, Nice, France, February 1997.
- [45] Akira Nakagawa, “Proposed revision of Annex Q/H.26,” Document of ITU-T STUDY GROUP 15, LBC-97-021, Nice, France, February 1997.
- [46] Akira Nakagawa, “Simulation Result of Reduced-Resolution-Update Mode,” Document of ITU-T STUDY GROUP 16, Q15-B-37, Sunriver, US, September 1997.
- [47] 平井 経太, 津村 徳道, 中口 俊哉, 三宅 洋一, “視覚特性を考慮したディスプレイの動きぼやけ評価法に関する一考察,” 映像情報メディア学会 Vol.62, No.10, pp.1580-1587, Oct. 2008.
- [48] “Temporal scalability algorithm based on image content,” 著: 1995, ISO/IEC JTC1/SC29/WG11, W377.
- [49] W.H.Chen, C.H. Smith and S.C. Fralick, "A fast computational algorithm for the discrete cosine transform," IEEE Trans. Commun., vol.25, no.9, pp.1004-1009, Sep. 1977.
- [50] 中川 章, 伊藤 隆, 佐々木 繁, 松田 喜一, 棚橋 純一, “動きベクトルの適応的可変長符号化方法,” PCSJ’ 99, P-3.7, 1999年9月.
- [51] 中川 章, 三好 秀信, 酒井 潔, 松田 喜一, “H.264/MPEG-4 AVCにおけるフィールド符号化モード時の色差ベクトル補正方式とその性能評価,” 信学技研 Vol. 103, No.57, pp.13-17, 2003年5月.
- [52] 中川 章, 阿南 泰三, 森松 映史, 伊藤 隆, “動きベクトル符号化装置及び復号装置”. 日本国特許 特許番号: 第 4142180 号.
- [53] Akira Nakagawa, Hidenobu Miyoshi, “Definition of chroma motion vector in field prediction,” ITU-T Rec. H.264, ISO/IEC 14496-10 AVC. JVT-E088, Geneva, Switzerland, Oct. 2002.
- [54] 中川 章, 三好 秀誠, “動画像復号化装置及び方法”. 日本国特許 特許番号: 第 3791922 号.
- [55] “Joint Final Committee. Draft of Joint Video Specification,” ITU-T Rec. H.264, ISO/IEC 14496-10 AVC. JVT-D157, Aug. 2002.
- [56] “H.264/AVC Software Coordination,” HHI, [オンライン]. Available: <http://iphome.hhi.de/suehring/tml/>.
- [57] G. Bjøntegaard, “Calculation of average PSNR differences between RDcurves,” ITU-T SG16/Q6, VCEG-M33, Austin, Texas, USA, 2001.
- [58] H. Kondo, K. Yamahara and J. Liao, "Identification of DCT signs for sub-block coding," Sixth International Symposium on Signal Processing and its Applications - ISSPA 2001, pp.569-572, Kuala Lumpur, Malaysia, Aug. 2001.
- [59] J. Koyama, A. Yamori, K. Kazui, S. Shimada and A. Nakagawa, "Coefficient Sign Bit Compression in Video Coding," Picture Coding Symposium 2012, pp.385-388, Krakow, Poland, May 2012.
- [60] 中川 章, 小山 純平, 数井 君彦, 甲藤 二郎, “動画像符号化における直交変換係数の正負符号の予測符号化方式,” 電子情報通信学

- 会論文誌 D, Vol.J100-D, No.9, pp.819-830, 2017年9月.
- [61] 中川 章, 数井 君彦, 三好 秀誠, 小山 純平, “画像符号化装置及び画像復号装置”. 日本国特許 特許番号: 第 5533886 号.
- [62] K. Kazui, J. Koyama, A. Nakagawa, “Description of video coding technology proposal by FUJITSU,” Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-A115, Dresden, Germany, Apr. 2010..
- [63] F. Henry, G. Clare, “Residual Coefficient Sign Prediction,” Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 4th Meeting, JVET-D0031, Chengdu, CN, , Oct. 2016.
- [64] 宮川洋, 原島博, 今井秀樹, “情報と符号の理論,” p15, 岩波書店, 東京, 1983.
- [65] “Joint Call for Proposals on Video Compression Technology,” ITU - Telecommunications Standardization Sector STUDY GROUP 16 Question 6 Video Coding Experts Group (VCEG), VCEG-AM91, Jan. 2010.
- [66] Akira Nakagawa, John L. Pittas, “High QoS and High Picture Quality Enable the HD Revolution,” SMPTE Motion Imaging Journal, vol.117, pp.55-63, October 2008.
- [67] 中川 章, “H.264/AVC の適用動向と富士通の取り組み,” 雑誌 FUJITSU, Vol.58, No.2, pp136-142, 2007年3月.
- [68] 中川 章, “デジタル映像の符号化技術・伝送装置の開発とハイビジョン化への貢献,” 信学技報, vol. 116, no. 335, IE2016-77, p.31, 2016年11月.
- [69] G.J. Sullivan, T. Wiegand, “Video Compression - From Concepts to the H.264/AVC Standard,” PROCEEDINGS OF THE IEEE, VOL. 93, NO. 1, pp.18-31, JANUARY 2005.
- [70] 安里 彰, 門岡 良昌, “計算機リソースを有効活用するグリッドミドルウェア: CyberGRIP,” 雑誌 FUJITSU, Vol.55, No.2, pp.146-151, Mar. 2004.
- [71] S. Sakaida, K. Iguchi, K. Iguchi and Y. Fujita, "Adaptive Quantization Control for Reducing Flicker of AVC/H.264 Intra Frames," Proc. of Picture Coding Symposium, 2004.
- [72] 山田 悦久, “映像符号化評価用標準動画像,” 映像情報メディア学会, Vol.62, No.8, pp.1283-1285, 2008.
- [73] Ye-Kui Wang, Miska M. Hannuksela, “Gradual Decoder Refresh Using Isolated Regions,” Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Meeting, JVT-C074, May 2002.
- [74] 中川 章, 三好 秀誠, 数井 君彦, 木村 尋成, “動画像符号化装置及び動画像符号化方法”. 日本国特許 特許番号: 第 4678015 号.
- [75] RFC2733, “An RTP Payload Format for Generic Forward Error Correction,” IETF, 1999.
- [76] SMPTE 2022-1:2007, “Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks,” SMPTE, 2007.
- [77] 一木 篤史, 中川 章, “伝送装置、伝送制御プログラム及び伝送方

- 法” . 日本国特許 特許番号: 第 4454320 号 .
- [78] Atsuchi Ichiki , Akira Nakagawa, “Transmission apparatus, transmission control program, and transmission method” . 米国特許 特許番号: US7376880.
- [79] 菊池健史, 日下新三 , 松山尚路, “TXN系列のIPによるHD伝送,” 放送技術 第60巻8号, p84-86, 2007年8月.
- [80] “業界初! H.264による系列局間HD映像伝送システムを構築,” 富士通株式会社 プレスリリース, 2007年6月4日, [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2007/06/4.html>.
- [81] 上野裕之 , 江口尚裕, “ANN系列局間HD-IP伝送網システム,” 放送技術 2007年8月号, pp.78-83, 2007年8月.
- [82] “世界初! H.264による日米間のHD映像伝送を実現,” 富士通株式会社 プレスリリース, 2007年2月8日, [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2007/02/8.html>.
- [83] Akira Nakagawa, “The 4:2:2/4:2:0 Perfect Reconstruction Filter Set and Its Application in HD-SNG,” SMPTE Motion Imaging Journal, vol. 118 , pp29-39, July/August 2009.
- [84] 小林俊輔 , 中川章, “422/420変換の繰り返し時の色差劣化を低減するフィルタの実現,” IMPS 2008, I5-19, 2008年10月.
- [85] 高宮健治, 大久保寛 , 田川憲男, “擬似スペクトル法を用いたエルミート画像補間法の高精度化,” 電子情報通信学会論文誌 B Vol. J95-B No.2 pp.355-365, 2012年.
- [86] “Test model 5,” ISO/IEC JTC1/SC29/WG11 Doc.N0400, 1993.
- [87] “VQEG Test Sequence: Mobile & Calendar,” [オンライン]. Available: ftp://vqeg.its.bldrdoc.gov/SDTV/VQEG_PhaseI/TestSequences/ALL_525/src15_ref__525.yuv.
- [88] “VQEG Test Sequence: F1 Car,” [オンライン]. Available: ftp://vqeg.its.bldrdoc.gov/SDTV/VQEG_PhaseI/TestSequences/ALL_625/src6_ref__625.yuv.
- [89] MPEG LA 社 ホーム ページ , [オンライン]. Available: <http://www.mpegla.com>.
- [90] 中川章 , 小林俊輔, “映像信号処理装置、及び映像信号処理方法” . 日本国特許 特許番号: 第 5003620 号 .
- [91] “映像中継の高画質化技術が世界の放送機器技術規格を制定する SMPTE の推奨指針として採用決定,” 富士通株式会社 プレスリリース , 2012年4月3日 . [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2012/04/3.html>.
- [92] 井原範幸, 澤田順一 , 吉田要, “H.264 | MPEG-4 AVC リアルタイム映像伝送装置 IP-9500, IP-900,” 雑誌 FUJITSU, Vol.60, No.1, pp.97-103, 2009年1月.
- [93] 三隅健司, 天本保 , 坂本昭彦, “ハイビジョン伝送システム IP-9500 のグローバル展開,” 雑誌 FUJITSU, Vol.60, No.1, pp.69-95, 2009年1月.
- [94] “世界初! H.264による日欧間のフルHD映像伝送の生中継を実

- 現,” 富士通株式会社プレスリリース, 2007年9月7日. [オンライン]. Available: <http://pr.fujitsu.com/jp/news/2007/09/7-1.html>.
- [95] “HD制作設備の整備状況を聞く コーデックの導入状況,” New Media 2010年3月号, p48, 2010年3月.
- [96] "CBS Taps Fujitsu for HD-SNG," V Technology, Vol 26, No.13, p8, June 2008.
- [97] “President Bush's State of the Union Address Utilizes Latest High Definition Video Technology,” Fujitsu Computer Products of America プレスリリース, April 14, 2008. [オンライン]. Available:
<http://www.fujitsu.com/us/about/resources/news/press-releases/2008/fcpa-20080414-4.html>.
- [98] Yoshiaki Shishikui, Kazuhisa Iguchi, Shinichi Sakaida, Kimihiko Kazui, Akira Nakagawa, “High-Performance Video Codec for Super Hi-Vision,” Proceedings of the IEEE | Vol. 101, No. 1, pp.130-139, January 2013.
- [99] “世界初！超高精細映像を実現するスーパーハイビジョン映像伝送システム用コーデック装置を開発,” 富士通株式会社プレスリリース, 2010年5月26日. [オンライン]. Available:
<http://pr.fujitsu.com/jp/news/2010/05/26.html>.

業績

査読論文

- [J1] 中川 章, 小山 純平, 数井 君彦, 甲藤 二郎, 「動画像符号化における直交変換係数の正負符号の予測符号化方式」, 電子情報通信学会論文誌 D, Vol.J100-D, No.9, pp.819-830, 2017 年 9 月(第 4 章に関連)
- [J2] Akira Nakagawa, “The 4:2:2/4:2:0 Perfect Reconstruction Filter Set and Its Application in HD-SNG,” SMPTE Motion Imaging Journal, vol. 118 , pp.29-39, July/August 2009. (第 6 章に関連)
- [J3] Akira Nakagawa, John. L. Pittas, “High QoS and High Picture Quality Enable the HD Revolution,” SMPTE Motion Imaging Journal, vol. 117, pp.55-63, October 2008. (第 5 章に関連)
- [J4] Akira Nakagawa, Eishi, Morimatsu, Takashi Itoh, Kiichi Matsuda, ”Dynamic Resolution Conversion Method for Low Bit Rate Video Transmission,” IEICE TRANS. COMMUN., Vol.E84-B, No.4, pp.930-940, April 2001. (第 2 章に関連)

国際会議

- [C1] Akira Nakagawa, Kimihiko Kazui, Eishi Morimatsu, Takashi Itoh, and Kiichi Matsuda, “Prediction Error Dynamic Resolution scheme for low bitrate video coding,” MoMuC’97, pp.576-579, September/October 1997. (第 2 章に関連)

講演

- [L1] 中川 章, 「デジタル映像の符号化技術・伝送装置の開発とハイビジョン化への貢献」(基調講演), 信学技報, vol. 116, no.335, IE2016-77, pp.31-31, 2016 年 11 月.
- [L2] 中川 章, 三好 秀誠, 酒井 潔, 松田 喜一, 「H.264/MPEG-4 AVC におけるフィールド符号化モード時の色差ベクトル補正方式とその性能評価」, 電子情報通信学会技術研究報告 巻 : 103 号 : 57(MoMuC2003 1-6), pp.13-17,2003 年 5 月.
- [L3] 中川 章, 伊藤 隆, 佐々木 繁, 松田, 喜一, 棚橋 純一, 「動きベクトルの適応的可変長符号化方法」, PCSJ’99, P-3.7, 1999 年 9 月.

その他の査読論文

[J5] Yoshiaki Shishikui, Kazuhisa Iguchi, Shinichi Sakaida, Kimihiko Kazui, Akira Nakagawa, “High-Performance Video Codec for Super Hi-Vision,” Proceedings of the IEEE | Vol. 101, No. 1, pp.130-139, January 2013.

[J6] 井口 和久, 中島 奈緒, 境田 慎一, 合志 清一, 筋誠 久, 鈴木 陽一, 伊藤 隆、中川 章、数井 君彦,「スーパーハイビジョンコーデックの開発と衛星伝送実験(招待論文)」, 電子情報通信学会論文誌 B Vol.J92-B No.7 pp.991-1002, 2009年7月.

その他の国際会議

[C2] Jumpei Koyama, Akihiro Yamori, Kimihiko Kazui, Satoshi Shimada, Akira Nakagawa, “Coefficient Sign Bit Compression in Video Coding,” PCS 2012, pp.385-388, May 2012.

[C3] Satoshi Shimada, Jumpei Koyama, Akihiro Yamori, Kimihiko Kazui, Hidenobu Miyoshi, Akira Nakagawa, “Enhancement on the motion vector derivation of HEVC for interlace format,” PCS 2012, pp.441 - 444, May 2012.

[C4] Hidetoshi Matsumura, Tatsushi Otsuka, Yasuhiro Watanabe, Takashi Fujita, Akira Nakagawa, Hiroshi Nakayama, “A Low-Power H.264/AVC Codec with 1080p/60 Processing Capabilities,” Proc. IEEE Symp. Low-Power and High-Speed Chips (Cool Chips XII), IEEE CS Press, 2009, pp. 175-187, April 2009.

[C5] Shinichi Sakaida, Nao Nakajima, Kazuhisa Iguchi, Seiichi Goshi, Kimihiko Kazui, Akira Nakagawa, “Development of AVC/H.264 codec for Super Hi-Vision”, IWAIT2008, p.3, 2008.

[C6] Takeshi Takahashi, Akira Nakagawa, Kiyoshi Sakai, Kiichi Matsuda, Hideyoshi Tominaga, "Adoption of Dynamic Resolution Conversion Scheme for Low-delay MPEG-4 Communication System," Asia-Pacific Conference on Communications, Bandung, September 2002.

その他の講演

- [L4] 中川 章, 三好 秀誠, 酒井 潔, 境田 慎一, 井口 和久, 合志 清一, 「携帯端末向け放送用試作エンコーダの開発」, 2005 映情学全大、18-5, 2005.
- [L5] 境田 慎一, 井口 和久, 鹿喰 善明, 数井 君彦, 中川 章, 「スーパーハイビジョン高効率映像符号化装置の開発」, 信学技報, SAT111(179), pp.61-66, 2011 年 8 月.
- [L6] 小林 俊輔, 中川 章, 「422/420 変換の繰り返し時の色差劣化を低減するフィルタの実現」, IMPS 2008, 15-19, 2008 年 10 月.

標準化寄書

- [S1] Akira Nakagawa, “Definition of chroma motion vector in field prediction between different parities,” Document of Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-E088, October 2002.
- [S2] Akira Nakagawa, “Simulation Result of Reduced-Resolution-Update Mode,” Document of ITU-T STUDY GROUP 16, Q15-B-37, September 1997.
- [S3] Akira Nakagawa, “Proposed revision of Annex Q/H.263,” Document of ITU-T STUDY GROUP 15, LBC-97-021, February 1997.
- [S4] Akira Nakagawa, “Proposed revision of Dynamic Resolution Conversion Schemes,” Document of ITU-T STUDY GROUP 15, LBC-97-020, February 1997.
- [S5] Akira Nakagawa, “Comparison of two Dynamic Resolution Conversion Schemes,” Document of ITU-T STUDY GROUP 15, LBC-96-306, November 1996.
- [S6] Kimihiko Kazui, Junpei Koyama, Akira Nakagawa, “Description of video coding technology proposal by FUJITSU,” Document of Joint Collaborative Team on Video Coding, JCTVC-A115, April 2010.

特許

- [P1] 中川 章, 数井 君彦, 三好 秀誠, 小山 純平, 「画像符号化装置及び画像復号装置」, 特許第 5533886 号, 2010 年 2 月 12 日出願

[P2] 中川 章，小林 俊輔、「映像信号処理装置、及び映像信号処理方法」、特許第 5003620 号、2008 年 7 月 15 日出願

[P3] 中川 章，三好 秀誠，数井 君彦，木村 尋成、「動画像符号化装置及び動画像符号化方法」，特許第 4678015 号，2007 年 7 月 13 日出願

[P4] 中川 章，三好 秀誠，「動画像復号化装置及び方法」，特許第 3791922 号，2002 年 9 月 6 日出願

[P5] 中川 章，阿南 泰三，森松 映史，伊藤 隆，「動きベクトル符号化装置及び復号装置」，特許第 4142180 号，1998 年 10 月 29 日出願

[P6] 中川 章，数井 君彦，森松 映史，「動画像符号化装置及び復号化装置」，特許第 3466032 号，1996 年 11 月 7 日出願

[P7] 中川 章，数井 君彦，森松 映史，「画像符号化装置」、特許第 3773585 号，1996 年 3 月 29 日出願

[P8] 一木篤史，中川章，「伝送装置、伝送制御プログラム及び伝送方法」，特許第 4454320 号，2004 年 1 月 9 日出願

表彰等

[A1] 紫綬褒章，中川 章，「デジタル映像の符号化技術と伝送装置の開発」，内閣府，2016 年 4 月

[A2] 市村産業賞 貢献賞，中川 章，大塚 竜志，宮坂秀樹，「デジタル映像伝送装置の開発」，新技術開発財団，2013 年 4 月。

[A3] 文部科学大臣表彰 科学技術賞 開発部門，中川 章，大塚 竜志，宮坂秀樹，渡部 康弘，小林 俊輔，「デジタル映像の符号化 LSI と伝送装置の開発」，文部科学省，2013 年 4 月。

[A4] 関東地方発明表彰 発明奨励賞，中川 章，阿南 泰三，森松 映史，伊藤 隆、「動画像の効率的なベクトル符号化方法」，（社）発明協会，2011 年 11 月。

- [A5] SMPTE Journal Certificate of Merit Award, Akira Nakagawa, “The 4:2:2/4:2:0 Perfect Reconstruction Filter Set and Its Application in HD-SNG,” Society of Motion Picture and Television Engineers, October 2010.
- [A6] 電気科学技術奨励会会長賞及び電気科学技術奨励賞, 宮坂 秀樹, 中川 章, 安井 哲也, 「H.264 対応ハイビジョン伝送装置の開発」, (公財) 電気科学技術奨励会, 2008 年 11 月.
- [A7] Paper of MoMuC'97 Award, Akira Nakagawa, Kimihiko Kazui, Eishi Morimatsu, Takashi Itoh, and Kiichi Matsuda, “Prediction Error Dynamic Resolution scheme for low bitrate video coding,” MoMuC'97, 1997 年 9 月.

Appendix

本 Appendix では、4.2.1 小節の画素の外挿予測方法について説明する。

画素の外挿予測方法に関しては、予測対象画素に隣接する符号化済の境界領域のエッジ特徴量を求め、線形予測の係数を適応的に切り替える方法を用いた。

図 64 は、ブロックの上境界に隣接する、座標(x,0)の符号化対象ブロックの予測画素 T_{ext_x} を、符号化済の境界領域 6 画素から推定する手順を示したものである。なお、ブロックの左境界に隣接する符号化対象ブロックの予測画素 L_{ext_y} は、本計算プロセスにおいて、x 座標、y 座標を入れ替えた式を用いて求める。

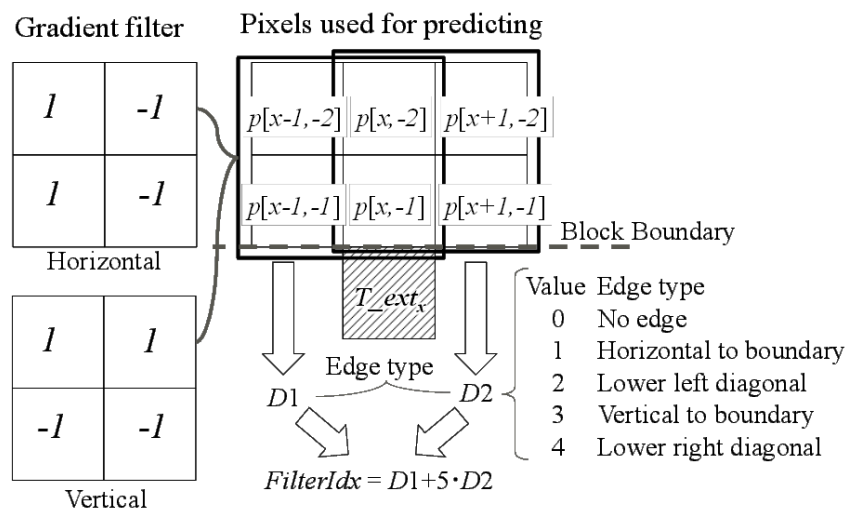


図 64 フィルタ係数の選択方法の概要

まず、式(62)から式(65)に従って、4 画素毎に、境界との平行方向への微分値 $Gh1$ 、 $Gh2$ 及び境界との垂直方向への微分値 $Gv1$ 、 $Gv2$ を求める。

$$Gh1 = p[x-1,-2] - p[x,-2] + p[x-1,-1] - p[x,-1] \quad (62)$$

$$Gv1 = p[x-1,-2] + p[x,-2] - p[x-1,-1] - p[x,-1] \quad (63)$$

$$Gh2 = p[x,-2] - p[x+1,-2] + p[x,-1] - p[x+1,-1] \quad (64)$$

$$Gv2 = p[x,-2] + p[x+1,-2] - p[x,-1] - p[x+1,-1] \quad (65)$$

表 30 エッジタイプ D_n ($n=[1..2]$)の導出表

G_{hn}	G_{vn}	D_n	G_{hn}	G_{vn}	D_n
$G_{hn} \geq 25$	$G_{vn} \geq 25$	2	$G_{hn} \leq -25$	$ G_{vn} < 25$	3
$ G_{hn} < 25$	$G_{vn} \geq 25$	1	$G_{hn} \geq 25$	$G_{vn} \leq -25$	4
$G_{hn} \leq -25$	$G_{vn} \geq 25$	4	$ G_{hn} < 25$	$G_{vn} \leq -25$	1
$G_{hn} \geq 25$	$ G_{vn} < 25$	3	$G_{hn} \leq -25$	$G_{vn} \leq -25$	2
$ G_{hn} < 25$	$ G_{vn} < 25$	0			

次に、表 30 に従って Gh1、Gv1 から D1、そして Gh2、Gv2 から D2 にエッジタイプ(0: No edge、1: Horizontal to boundary、2: Lower left diagonal、3: Vertical to boundary、4: Lower right diagonal)を設定する。ここで、表中の n は n=[1,2]とする。

そして、FilterIdx を次式で求める。

$$FilterIdx = D1 + 5 \cdot D2 \quad (66)$$

次に表 31 の 25 種類のフィルタ係数セットの中から FilterIdx に対応するものを a0 から a5 に設定する。これらのフィルタ係数セットは、FilterIdx 毎に文献 [65] のテストセット映像から最小二乗法で求めた。

表 31 FilterIdx に対応するフィルタ係数

FilterId x	Filter Coefficients					
	a0	a1	a2	a3	a4	a5
0	140	84	137	-31	-71	-3
1	118	-56	271	-17	-216	156
2	44	243	172	-15	-41	-147
3	13	235	101	-5	-34	-54
4	151	91	-1	-23	18	20
5	230	-45	184	147	-232	-28
6	126	13	148	19	-66	16
7	57	101	182	29	-20	-93
8	124	118	114	75	-127	-48
9	197	44	27	46	-84	26
10	-8	104	169	36	-36	-9
11	-10	63	208	41	-88	42
12	-13	154	169	28	-13	-69
13	-30	222	126	27	-28	-61
14	35	151	24	15	13	18
15	83	217	38	-41	-54	13
16	93	150	134	-48	-156	83
17	55	226	96	-14	-80	-27
18	60	328	64	-27	-141	-28
19	129	218	-25	-61	-33	28
20	157	238	35	-122	-61	9
21	156	124	49	-74	-30	31
22	102	151	109	-35	-24	-47
23	94	209	55	-20	-72	-10
24	170	145	-7	-62	-19	29

最後に式(67)で T_{ext_x} を求め、0-255 にクリップして画素の外挿予測が完了する。

$$T_{ext_x} = (a0 \cdot p[x-1,-1] + a1 \cdot p[x,-1] + a2 \cdot p[x+1,-1] + a3 \cdot p[x-1,-2] + a4 \cdot p[x,-2] + a5 \cdot p[x+1,-2] + 128) / 256 \quad (67)$$