

時間方向視覚特性に基づく 高フレームレート画像符号化方式の一検討

高木幸一[†] 内藤整[†] 酒澤茂之[†] 松本修一^{†, ††}

本稿では、59.94p 超の高フレームレート動画像に対し、時間方向の視覚特性に基づき、高効率な符号化方式を提案する。本方式は、鋭鈍繰返し錯視で用いられている時空間モデルを利用し、高周波成分をカットする方式である。実験の結果、主観画質を落とすことなく約 20% の符号量削減ができることを確認した。

High Frame Rate Video Coding Scheme Based on Temporal Visual Property

KOICHI TAKAGI[†] SEI NAITO[†]
SHIGEYUKI SAKAZAWA[†]
SHUICHI MATSUMOTO^{†, ††}

In this paper, we propose a high efficient video coding scheme for high frame rate video (over 59.94p) based on temporal visual property. This method can remove high frequency parts of residual coefficients based on a spatio-temporal frequency model used in repetition of sharp/blurred mechanism. Experimental results show that the proposed method achieves average 20% bit-savings over general H.264 coding method without degradation.

1. はじめに

近年、IPTV やデジタル放送に代表される映像配信・放送サービスでは、HDTV クラスの解像度を持つコンテンツが主流となりつつある。また、将来的には、映像コンテンツの高解像度化、高フレームレート化が進み、高臨場感が得られるサービスの実

[†](株)KDDI 研究所
KDDI R&D Labs. Inc.
^{††}(社)日本ケーブルラボ
Japan Cable Laboratories

現期待されており、その一つとしてスーパーハイビジョンに代表される走査線 4000 本/60p クラスの超高精細・高フレームレート映像が注目されている¹⁾。特に、高フレームレート化の方に着目すると、60p 以上の高フレームレート化も必要になるのではとの声もあがっている²⁾。

一方、フレームレートが高くなればなるほど、それに伴い単位時間あたりの画像データが増えるため、それらをいかに効率よく符号化・圧縮するかがポイントとなる。そこで、これに対し、増加した画像データの符号量を削減する為の検討がなされている。特に文献³⁾では、鋭/鈍繰返し再生による錯視の原理が述べられている。鋭/鈍繰返し錯視とは、例えば、60p の画像があった場合に、1 コマおきに鋭画像（高解像度画像、30 コマ/秒）と鈍画像（低解像度画像、30 コマ/秒）を繰返すと、全体が鋭画像に見えるというものである。よって、これを使うことにより、画質を落とすことなく、コマを増やすために挿入された画像の符号化効率を向上させられることが期待できる。ところが、1 枚おきに低解像度画像を符号化することは、時間方向の相関が低くなることにもつながるため、場合によっては符号化効率が下がることが懸念される。また、同方式では、当該フレームを鋭/鈍いずれかのフレームと一義的に決定し、鈍フレームでは画面内で一律のフィルタ処理を適用することを前提としている。このような画面内で一律の処理を行った場合、映像の動きの特性などによっては部分的に劣化が見えるなどの問題が発生することが知られている。

そこで、本稿では、上記課題を解決する為に、鋭/鈍繰返し錯視の原理を解明する上で使われている時空間モデルを活用し、これを画面内の局所領域ごとに適用することを考える。具体的には、「同モデルより、時間・空間解像度が共に高い領域は、人間の視覚の帯域外になるため、その部分の雑音は無視できる」という仮説を立て、この仮説に基づいた符号化方式を提案する。

2. 鋭/鈍繰返し錯視の原理とその応用

本節では、文献³⁾で説明されている鋭/鈍繰返し再生による錯視の原理について概説する。この原理を時空間周波数領域で解析するため、上記例同様、60p 信号を想定すると、各成分の時間標本化周波数は鈍画像：60Hz、鋭画像成分：30Hz となる。空間（水平、垂直）周波数 μ , ν , 時間周波数 f の時空間周波数スペクトルは図 1 のようになり、視覚の通過帯域（図 1 の四角（網点）内）で見ると、鈍画像の最大空間周波数を適当に選択することにより、30Hz 成分は減衰して対域外になるため、目には見えなくなる。その結果、すべてが 60Hz 高解像度画像（鋭画像）の場合と等価になるはずである。

なお、同文献では、符号化効率の向上についても触れている。例えば、「60p を対象

とする MPEG における現象と応用」と題し、I、P、B の各符号化モードで解像度が異なるとフリッカになることが知られている。上記結果を利用すれば、例えば $M=4$ の場合の B picture の前後 2 枚や $M=2$ の場合の B picture を低解像度に抑えることにより効率化が期待できると書かれている。

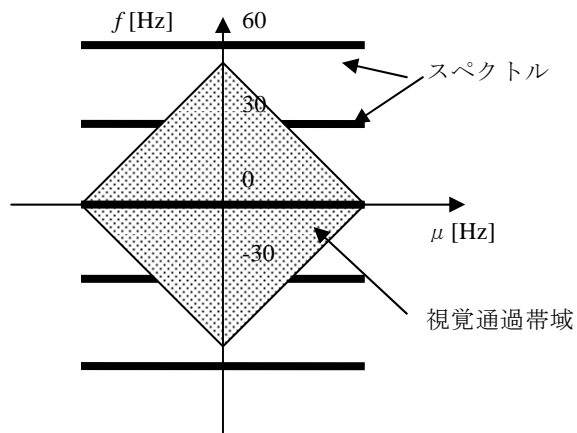


図 1 鋭／鈍画像のスペクトルと伝送周波数特性
 (簡単のため、水平-時間周波数のみ提示)

3. 提案方式

3.1 局所領域制御のための時空間特性表現方法

1 節にも述べた通り、フレームを鋭／鈍どちらにするかを全体的に一義的に決定し、鈍フレームで一律のフィルタ処理を適用すると、映像の動きの特性により劣化が発生することが考えられる。そこで、本節では、この処理を局所的に行うための方法について述べる。

まず、本処理を適用するフレームの各 MB (マクロブロック) に対し、図 2 に示す通りその前後数フレームの同じ位置の画素値を用いて時空間 3 次元映像信号を取得する。そして、この 3 次元映像信号に対し 3 次元 FFT を適用し、時空間周波数特性を求める。この時空間周波数特性の結果を図 1 にプロットした結果、網点部の外にある要素は人間の目に知覚できないはずである。よって、これらの要素はどのように扱ってもよいことが期待できる。

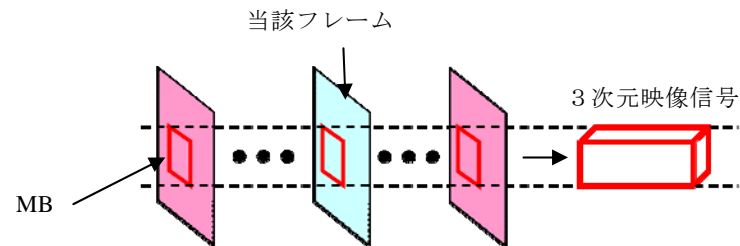


図 2 3次元映像信号作成

3.2 符号化への適用

本節では、前節の表現に基づき、これを符号化に適用する方法について述べる。

まず、鋭／鈍繰り返しの原理に従い、今回の処理は 1 枚おき (偶数番目のフレームのみ) に行われるものとする。本処理の対象となるフレームの各 MB に対し、前節で述べた通り、原画像からその前後数フレームにわたり 3 次元映像信号を作成し、それに 3 次元 FFT を適用する。一般的に FFT の結果は図 3(a)に示す通り、折り返しを除くと原点を通る 1 本の直線のようなになる。これを模式的に図 3(b)に示す。このスペクトル画像に対し、図 1 の視覚通過帯域をはめ込む。この視覚通過帯域より外にある周波数成分は人間の目には知覚できない部分である。

よって、動き補償もしくはイントラ予測等を行い、残差信号を求め、それを変換した結果、高周波成分が発生しているとすれば、それらは人間の目に知覚できない。すなわち、これらの成分は符号化対象から外してもよいことが期待できる。この原理に基づき、(直線形をした) スペクトルが、視覚通過帯域と交わる点を求める。その交わる点の空間周波数座標に基づき、その点より周波数が高い係数を伝送しない、つまりゼロとする。もちろん、何らかの雑音を重ねることになるが、人間の目には知覚できないはずである。

本来は復号画像に対し同様の処理を適用すべきである (すなわち 1 パス目に通常の符号化を行い、そこで得られる復号画像に対し FFT を適用し、その結果に基づき係数をカットする 2 パス符号化を適用すべきである) が、この処理を原画に適用してもほぼ同等の結果が得られる、つまり、時空間周波数スペクトルは原画と復号画で大きく変わらないという仮説を立て、この処理を符号化モード選択部に適用している。

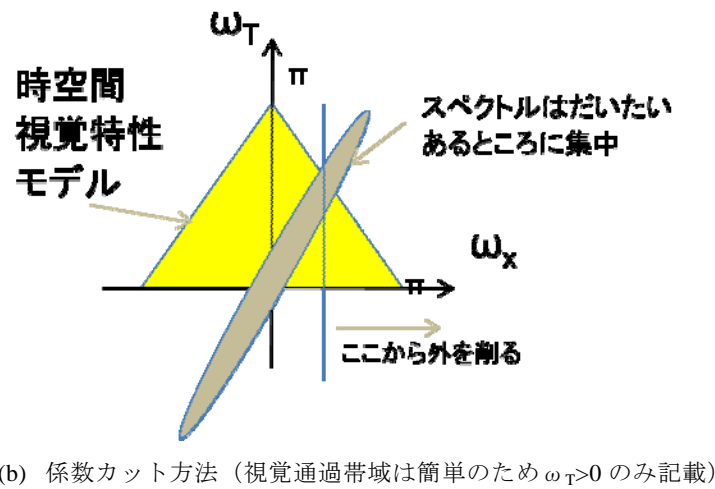
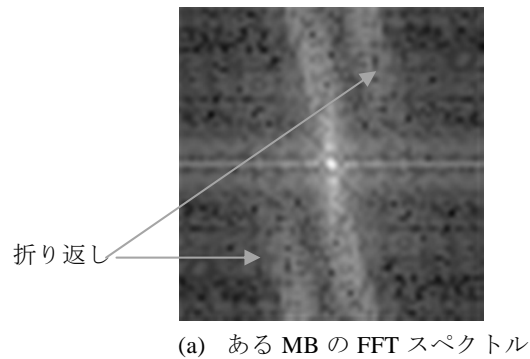


図 3 時空間特性表現の符号化への適用
(両方とも簡単のため、水平-時間周波数特性のみ提示)

さらに、予備実験の結果、インター符号化が施された MB に関しては、残差信号をなし (すなわち Not coded) としても、その品質に大きな影響がないことを確認した。これはフレームレートが非常に高いため、精細な画像の場合、そのテクスチャが崩れない限り動きを正しく追従できるため、動き補償予測誤差がそれほど大きくならない、一方で平坦な画像の場合はもとの画素値のばらつきが小さいため、低周波数成分に集中するからであると考えられる。そこで、時空間周波数特性による係数カットを適用するのはイントラ符号化が適用される箇所のみとする。

以上をまとめると、全体の流れは

図 4 のようになる。

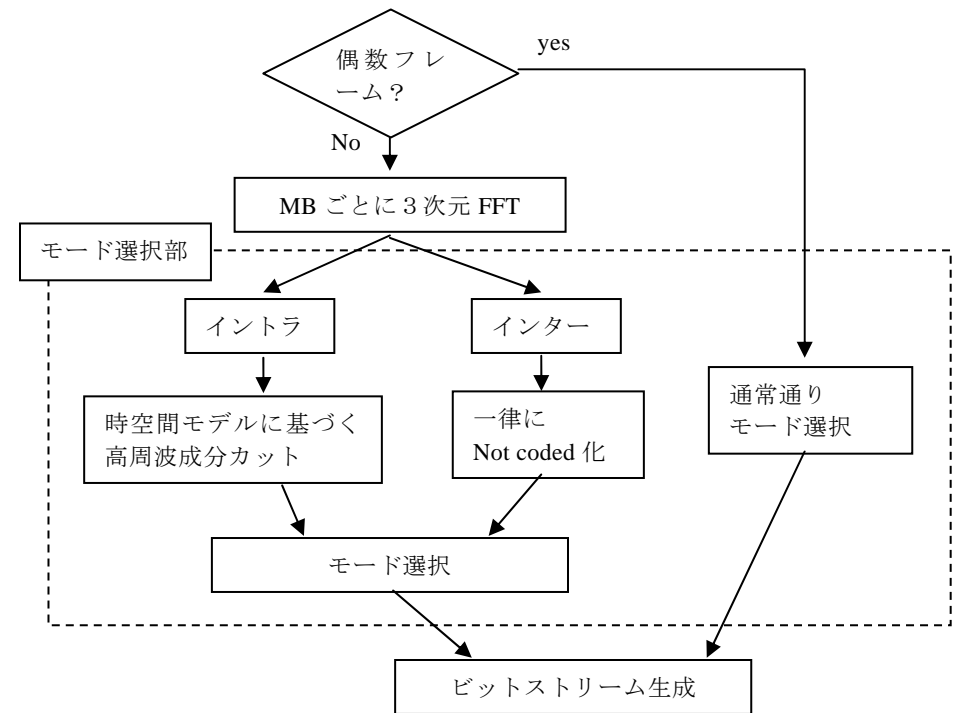


図 4 提案法の処理の流れ

4. 実験および考察

前節の提案符号化方式の有効性を検証するため、実験を行う。実験では、H.264 符号化に前節の提案手法を適用した場合と、通常の H.264 符号化方式を適用した場合を比較する。実験条件を表 1 に示す。なお、量子化パラメータ QP は固定とし、各シーケンスに対し、およそ 6Mbps (スーパーハイビジョンの解像度 (8K) で 90~100Mbps に相当) になる QP をそれぞれ選択した (選択された QP は表 2 に記載)。また、視覚特性モデルを決定する 3 次元 F F T は空間方向 16×16 次元、時間方向 4 次元 (時間後方 2 フレーム+当該画像 1 フレーム+時間前方 1 フレーム) を適用した。つまり本方式に基づく原理遅延は 1 フレーム分のみである。

はじめに、提案方式により画質が劣化するかどうかを確認する為に主観評価実験を行った。評価は被験者数の観点で DSQUS⁴⁾を簡易化した方法を用いた。被験者は非専門家 10 名、視距離は 3H である。その結果を図 5 に示す。同図より、各シーケンスに対し、両方式の結果に有意な差異がないことがわかる。つまり、提案方式により、1 枚おきに残差信号をカットしたとしても、画質が劣化しないことを意味する。ただし、今回の実験では全体的に MOS 値が低いため、符号化レートを高くしてどのようになるのかもあわせて確認する必要があると考えられる^a。

表 1 実験条件

	項目	内容
符号化	フレーム数, 構造	600(10sec), IBBBP(M=4, N=32)
	エントロピー符号化	CABAC
	RD-optimization(RDO)	ON
素材	シーケンス	ITE ハイビジョン・システム評価用動画画像 6 種(シーケンス名は図 5, 表 2 に記載)
	色空間	YUV4:2:0 (元シーケンスから色空間を間引き)
	ビット深度	8bit (元シーケンスから間引き)
	ライン数, フレームレート	1920×1080/59.94p

^a 筆者がビットレートをおよそ倍にして比較したところ、主観的にほぼ同等の画質になることを確認している。

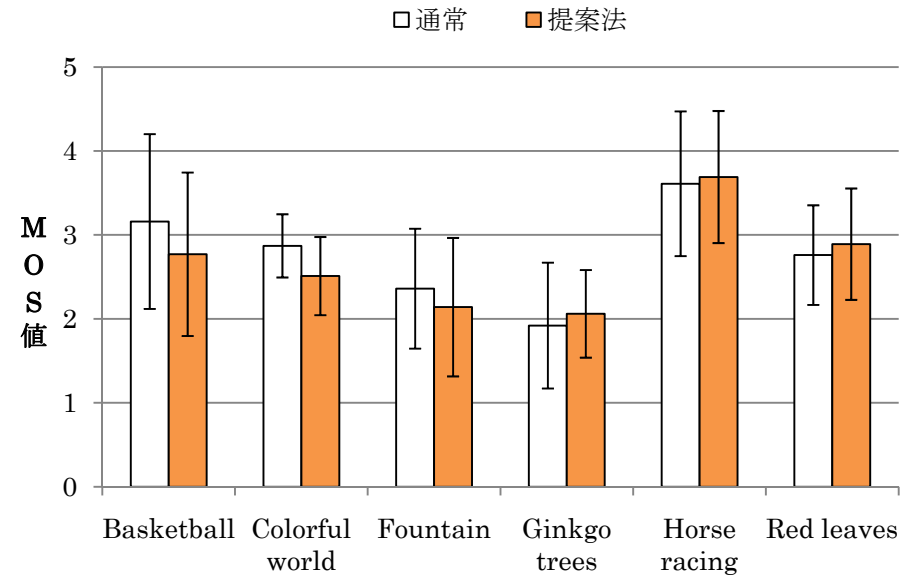


図 5 主観評価実験結果 (誤差線は標準偏差)

次に、提案方式による符号量の削減効果を表 2 に示す。同表より、全シーケンスで約 20% の符号量削減が達成できていることがわかる。なお、参考のため、今回対象とした符号量より少し高い符号量を与えた場合 (つまり QP を小さくした場合) に符号量削減率がどのようになるのかを確認したところ、若干減りはするものの、ほぼ同様の結果が得られた。これは、符号化レートが高くなると、残差信号に符号量を多く要することができる一方で、被参照画像の誤差が小さくなるため、参照画像の予測誤差もそれにあわせて小さくなる、すなわち、残差信号が少なくなり、両者の効果が相殺されているものと推測される。

表 2 符号量削減効果

シーケンス	QP	通常 [Mbps]	提案法適用 [Mbps]	符号量 削減率[%]
Basketball	33	6.21	5.09	17.9
Colorful world A	35	5.76	4.62	19.8
Fountain (dolly)	39	6.07	4.47	26.3
Ginkgo Trees	41	6.46	5.06	21.7
Horse racing (dart)	32	5.85	4.77	18.4
Red leaves (pan down)	35	6.53	4.99	23.6

5. おわりに

本稿では、高フレームレート動画像に対し、時空間視覚特性に基づき、1枚おきに高周波成分をカットする方式を提案した。さらに、符号化モードごとに適応的に係数のカットを行う方式を提案した。そして、実験の結果、主観画質を落とすことなく約20%の符号量削減ができることを確認した。

本研究は独立行政法人情報通信研究機構による委託研究「超高精細映像符号化技術に関する研究開発」として実施したものである。

参考文献

- 1) 菅原正幸, “走査線 4000 本級超高精細映像システム,” 映情学誌, 57,11, pp.1442-1444(2003).
- 2) “スーパーハイビジョン時代の映像 フレーム周波数は 60Hz 以上へ” 映像新聞, 第 1961 号 (平成 22 年 7 月 26 日), pp.5 (2010).
- 3) 吹抜敬彦, “鋭/鈍繰返し画像の解明とフレーム内挿倍速表示 (TFI) などへの応用～視覚信号処理工学の発展～” 映情誌 Vol. 63, No. 4 pp.549-552 (2009).
- 4) ITU-R Rec. BT.500-11, “Methodology for the subjective assessment of the quality of television pictures,” 2002.