

JPEG2000 による HDTV 低遅延符号化のための符号量制御方式

Bit Allocation for Low Delay HDTV Coding Based on JPEG2000.

佐野 雄磨† 内藤 整‡ 渡辺 裕†

Yuma SANO† Sei NAITO‡ Hiroshi WATANABE†

1 はじめに

JPEG2000 は、高い符号化性能と符号化ストリームのスケーラブル性を兼ね備えており、特に高解像度の映像圧縮に関して有効性を示す [1]。HD レベルの高解像度映像は、主に放送局間などの素材伝送に用いられ、伝送遅延の低減が強く求められている。筆者らは、JPEG2000 を用いて HD 解像度の映像の低遅延符号化について検討を行ってきた [2]。入力画像を分割し、読み込み次第符号化を行い伝送することで伝送遅延を低減し、さらに適切なビット配分方法を検討してきた。しかし、これまでは、各フレームに等ビットを配分することを想定しており、時間方向の符号量制御は検討されていなかった。そこで本稿では、低遅延で時間方向の品質変動を考慮した符号量制御について検討を行う。

2 イントラ符号化による動画像符号化の問題点

Motion JPEG2000 を代表とするイントラ符号化は、動き予測におけるフレームの並び替えや、CBR 符号化に適合させるためのバッファリングによる遅延が生じない。従って、JPEG2000 を用いた CBR 符号化は、低遅延符号化を実現するための有効な方法であると言える。しかしその反面、フレームごとに独立に符号化されるために、各フレームの複雑度によって、フレームごとの品質が大きく変動する可能性がある。これは、各フレームを単独で見た場合は、劣化として検知されないが、動画として見た場合、主観的な劣化として表れる [3]。そこで、これらの劣化を低減させるために、時間方向の品質変動を考慮した符号量制御が必要となる。

3 提案符号量制御方式

3.1 RD(Rate-Distortion) 最適化

JPEG2000 の符号量制御 [4] では、各フレームの RD 特性が用いられる。Wavelet 変換された後、Wavelet 係数は、EBCOT 処理によってビットプレーン、符号化パスに分解される。その後、歪が最小となるように、RD 最適化に基づいて、ビットの調整が行われる。そこで本手法では、この RD 特性を画像の複雑度の指標として、後続のフレームの RD 特性を予測することによって、フレーム間の品質変動を考慮した 1 パス符号量制御を検討する。

3.2 RD 曲線の近似と予測

まず、JPEG2000 符号化処理の中で算出される各フレームの RD 曲線を、Marquardt 法 [5] を用いた非線形回帰近似によってモデル化する。これを用いると、D(Distortion) は、R(Rate) を用いて以下のように表される。ただし、 a, b, c は各フレームに依存する係数である。

$$D = a + \frac{b}{R + c} \quad (1)$$

ここで、 a, b, c をフレーム数 k の関数として考える。シーケンスを符号化していく中で、符号化済みフレームの RD 曲線の係数 a, b, c を用いて、後続フレームの RD 特性を予測する。フレーム数 k と a, b, c の関係を回帰によって求める。第 n フレームまで符号化した時点で、予測から求めた第 k フレーム ($k > n$) の RD 曲線における a, b, c の式を $a_n(k), b_n(k), c_n(k)$ と表す。

予測から求めた、後続 N フレームの RD 曲線における a の平均値を $\bar{a}(n)$ とすると式 (2) のように表される。

$$\bar{a}(n) = \frac{1}{N} \sum_{i=1}^N a_n(n+i) \quad (2)$$

($\bar{b}(n), \bar{c}(n)$ についても同様)

これらを用いて、当該フレーム n から、後続 N フレームの平均的な R と D の関係は式 (3) のように表される。

$$\bar{D}(n) = \bar{a}(n) + \frac{\bar{b}(n)}{R + \bar{c}(n)} \quad (3)$$

3.3 RD 曲線の予測に基づくビット配分

次に、各フレームに割り当てる符号量を計算する。提案方式では、当該フレームから後続 N フレーム分に割り当てる符号量をビット過不足を考慮して計算する。具体的には、当該フレーム n から N 枚のフレームについて割り当てる平均符号量 $R(n)$ を式 (4) により算出する。 B_W は 1 フレームに割り当てられる平均符号量である。

$$R(n) = \frac{(N+n) \times B_W - \sum_{i=0}^{n-1} S(i)}{N} \quad (4)$$

$S(i)$: i 番目のフレームにて発生した符号化ビット数

式 (3) の RD 式について、式 (4) で求めた $R(n)$ を満たすような $\bar{D}(n)$ を求め、これを用いて当該フレームの符号化を行う。ただし、後続を予測することによって、フレーム間の品質の差が不自然に大きくなる場合は、その差を修正する処理を加える。

4 実験結果と考察

提案手法の有効性を実験により示す。原画には Streetcar を用いた。本手法では、フィールド符号化を想定し、60 フィールド/秒で符号化を行った。目標ビットレートは 0.8bpp、後続フィールド数 N は 60 フィールドとした。全フィールドに等ビットを与えた場合と提案手法を用いた場合の品質の変動を図 1 に示す。図 1 より、提案手法によって、フィールドごとの品質がなだらかに変化する様子が見られる。さらに、各フィールドに等ビット配分した場合と、提案手法を用いた場合の MSE の平均と分散を求めたものを表 1 に示す。表 1 より、提案手法が等ビット配分時と同程度の MSE の値を保持したままで、MSE の分散が大きく低下している様子が見られる。MSE の差は、シーケンス全体に対して割り当てた符号量の違いによるものと考えられる。

† 早稲田大学大学院 国際情報通信研究科
‡ (株)KDDI 研究所

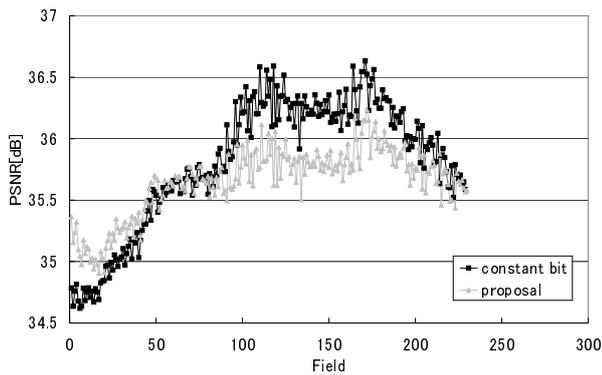


図1 等ビット配分方式と提案手法の PSNR の比較 (バッファ無制限)

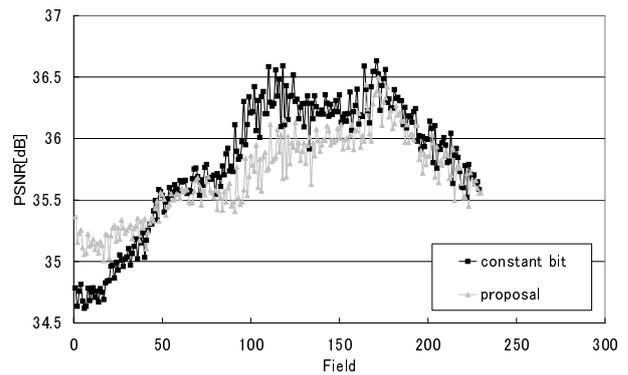


図3 等ビット配分方式と提案手法の PSNR の比較 (バッファサイズ 8Mbit)

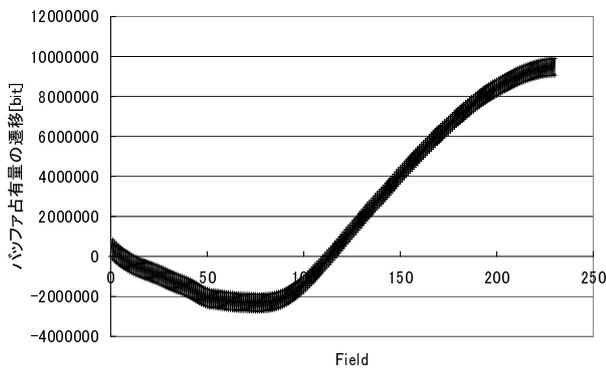


図2 提案手法によるバッファ占有量の遷移 (バッファ無制限)

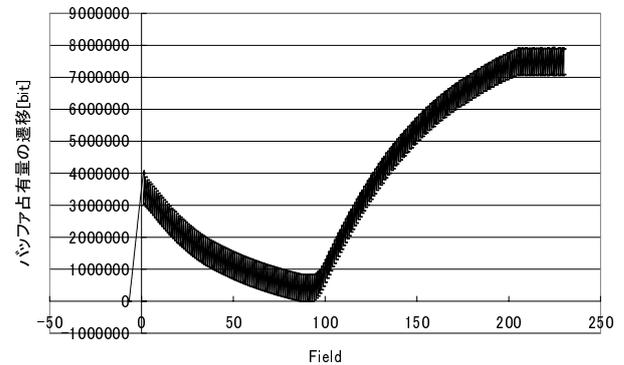


図4 提案手法によるバッファ占有量の遷移 (バッファサイズ 8Mbit)

表1 MSE の平均値と分散 (バッファ無制限)

	MSE の平均	MSE の分散
等ビット配分	17.178	4.6893
提案手法	17.723	1.3530

表2 MSE の平均値と分散 (バッファサイズ 8Mbit)

	MSE の平均	MSE の分散
等ビット配分	17.178	4.6893
提案手法	17.539	1.8748

次に、各フィールドに可変符号量を与えているため、バッファサイズについて検証を行う。仮バッファサイズを1フィールド分の符号量として、バッファの占有量の変化を示したのが図2である。図2より、提案手法を用いてバッファを破綻させないためには、バッファサイズとして約13Mbit 必要であることがわかる。しかし、バッファサイズを13Mbit に設定すれば、バッファリングによる遅延は260msec となり、16フィールド分の遅延に相当する。

そこで、バッファサイズの制限下でのシステムの動作を確認するために、バッファサイズを8Mbit と設定し、この制限下で、符号量制御を行うこととした。バッファサイズを8Mbit と決定した場合の、品質変動を図3に、MSE の平均と分散を表2に、その際のバッファの占有量の様子を図4に示す。限られたバッファ条件の下でも、品質変動を低減する効果があることがわかる。

5 まとめ

各フレームの RD 特性を考慮することによって、フレーム間での1パス符号量制御を行った。RD 曲線の傾向を予測して符号量制御を行うことで、等ビット配分時と比較して品質変動を抑えることが可能となった。また、バッファサイズの制限下でも、同様の効果が得られることがわかった。

参考文献

- [1] D. Marpe et al. "Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in intra coding mode," in Proc. SPIE Conf. on Wavelet Applications in Industrial Processing, Oct. 2003.
- [2] 佐野, 内藤, 渡辺, "JPEG2000 Precinct 構造を用いた低遅延符号化方式," 情報処理学会オーディオビジュアル研究会 Oct. 2004.
- [3] W. Ding and B. Liu, "Rate control of MPEG-video coding and recording by rate-quantization modeling," IEEE Trans. Circuits Syst. Video Technol., vol.6, pp.12-19, Feb. 1996.
- [4] ISO/IEC JTC1/SC29/WG1 (ITU-T SG-8) N1646R Coding of Still Pictures: "JPEG2000 Final Committee Draft Version1.0," Mar. 2000.
- [5] Y. Charfi, R. Hamzaoui, and D. Saupe, "Model-based real-time progressive transmission of images over noisy channels," IEEE Wireless Communications and Networking Conference, New Orleans, Mar. 2003.