

ウェーブレット変換符号化動画像における視覚歪みの低減

村田 悠[†] 菊池 久和[†] 村松 正吾[†]

[†]新潟大学工学部電気電子工学科
〒950-2181 新潟市西区五十嵐2の町

E-mail: †you@telecom0.eng.niigata-u.ac.jp, ††{kikuchi,shogo}@eng.niigata-u.ac.jp

あらまし 離散ウェーブレット変換を用いた動画像符号化方式である Motion JPEG2000 は、映像コンテンツの編集・加工性、可逆・非可逆圧縮の統合、品質スケーラビリティ等の優れた特徴を持っている。しかし、フリッカ雑音と称する特有の視覚歪みが発生することが知られている。これは動画像において視覚的にちらつきを伴う品質劣化として知覚されやすく、歪みの低減が必要である。本稿では復号化器側においてこの歪みを低減するための手法を検討する。提案手法では、領域平均フレーム間差分値を用いて歪みと映像本来の動きとを識別し、過去のフレームを使ってフレーム間での画素値の変動を抑える。最後に実験によりその有効性を確認する。

キーワード 動画像, 符号化, JPEG2000, フリッカ雑音, デジタルシネマ

Reduction in Perceptual Distortion for Wavelet Transform-Based Video Decoding

Yutaka MURATA[†], Hisakazu KIKUCHI[†], and Shogo MURAMATSU[†]

[†] Department of Electrical and Electronic Engineering, Faculty of Engineering, Niigata University
Ikarashi-2, Nishi, Niigata, 950-2181 Japan

E-mail: †you@telecom0.eng.niigata-u.ac.jp, ††{kikuchi,shogo}@eng.niigata-u.ac.jp

Abstract Motion JPEG2000 is a video coding system on the basis of wavelet transforms. It has many advantages such as easy editing, unified lossless/lossy compression, resolution and SNR scalabilities. However, there can produce some noise referred to as “flicker artifacts”. Since it is visually annoying in a video sequence, it is strongly desirable to reduce it for quality decoding. In this paper, a distortion reduction method at a decoder is presented. The distortion is differentiated from the sequence movement by the region mean of frame difference. The information of previous frames is used to improve the perceptual appearance of the current frame. Experimental results demonstrate that the proposed method works well.

Key words video, coding, JPEG2000, flicker artifact, digital cinema

1. はじめに

近年、ブロードバンドネットワークが普及しつつあり、加えてデジタル処理技術の高度化や撮像、表示などのデバイスの高性能化により、高精彩映像等の大規模な情報配信を容易に提供できる環境が整備されつつある。それにより、感性や文化を対象とするコンテンツ配信事業が今後の基幹産業の一部を担うと予想されている。特に、撮影から編集、映写までの映画の一連のプロセスをデジタルフォーマットで統一的に扱うデジタルシネマには大きな期待が寄せられている。このデジタル化によりフィルム劣化等の物理的制約から解放された最高品質の映像を提供することが可能であり、インターネットストリーミング等を用いた家庭への2次配信も示唆され

ている。現在、このデジタルシネマシステムは DCI(Digital Cinema Initiatives, LLC), SMPTE(Society of Motion Picture and Television Engineers) および ISO/IEC JTC1/SC29/WG1 で国際標準化作業が進められており、4,096×2,160画素という超高解像度 4K デジタルシステムが実験段階にある [1-3]。

デジタルシネマに対応する画像の符号化方式として JPEG2000 [4-6] が決定している。JPEG2000 は静止画像符号化方式であるが、動画像を一連の静止画像のフレームシーケンスと見なして、各画像フレームを JPEG2000 で符号化を行う拡張規定が Part3 にある。これは Motion JPEG2000 [7] と呼ばれる。Motion JPEG2000 は、イントラフレーム符号化を用いることにより、ランダムアクセス性や編集性が容易である、また離散ウェーブレット変換の特徴として、可逆・非可逆圧縮の統合、

符号化データがビット精度及び空間解像度にスケラビリティ性を持つなど映像コンテンツの柔軟な取扱いができる多くの優れた特徴を有している。

一方で、Motion JPEG2000 で符号化した映像には、フリッカ雑音が発生し、映像品質の視覚的劣化を招くことが指摘されている。このノイズは復号画像上に細かいリング状のノイズとして現れる。この歪みは静止画像ではさほど目立たないが、動画ではその大きさと位置がフレーム毎に変化するため、時間方向のノイズ成分が発生し、視覚的に顕著になる。このため、動画におけるフリッカ雑音の低減が必要である [8]。

本稿では、復号器側におけるフリッカ雑音の低減方法を提案する。JPEG2000 符号化で生じる歪みは、主に高周波帯域サブバンドの量子化雑音による画素単位の歪みであり、そのフレーム間での変化も画素毎の変化となる。これに対して、映像中の移動物体による変化は領域単位で変化する特性を持つ。提案手法では、フレーム間差分値の局所領域平均を用いて、移動物体とノイズとを識別する。そして、過去フレームの画素値を使って補正することにより、フリッカ雑音を低減させる。実験結果を視覚的及び数値的に示し、本手法の有効性を示す。

本稿の構成は以下のとおりである。まず第 2 章で、JPEG2000 符号化により生ずる歪み及びフリッカ雑音について述べる。第 3 章で、フリッカ雑音の変化特性を考察する。第 4 章で、提案手法のフリッカ雑音低減法を説明する。最後に第 5 章で、実際の映像を用いた実験結果を示す。

2. フリッカ雑音

2.1 静止画像における歪み

離散ウェーブレット変換符号化により発生する歪みは復号画像上に細かいリング状のノイズとして現れる。このリングノイズはウェーブレット変換係数の量子化雑音が原因である。量子化の際に切り捨てられた量子化雑音が逆変換時に周囲に広がり、歪みが発生する。

変換係数量子化により発生するひずみの特性は周波数成分毎に異なる。低周波サブバンドの量子化による歪みは画像全体に広がる。逆に高周波サブバンドの量子化によって発生する歪みは輪郭周辺に局在化する特徴を持っている。これは、ウェーブレット変換を繰り返し行った低周波サブバンドでは、高周波サブバンドと比べてフィルタのタップ数が画像領域において長くなるためである。

ただし、一般に画像符号化では空間周波数が高くなるとコントラスト感度が低くなるヒトの視覚特性を考慮して、高周波成分ほど粗く量子化され、低周波成分はほとんど量子化されない。JPEG2000 符号化により発生する歪みを図 1 に示す。これは 'Lena' (512 × 512, 8-bit, gray-scale) を JPEG2000 符号化器 [10] で 1/64 に符号化したときの歪みを調べたものである。図 1(b) のように低周波サブバンドに起因する画像広範囲に広がる歪みはほとんど発生せず、図 1(c),(d) に示すような高周波サブバンドの量子化雑音による画素単位の歪みが主となる。

2.2 動画における歪み

動画における歪みを考える場合、先に述べた静止画像の歪

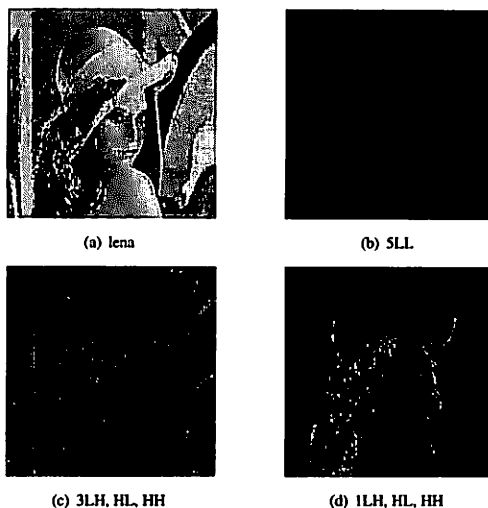


図 1 JPEG2000 符号化により発生する歪み

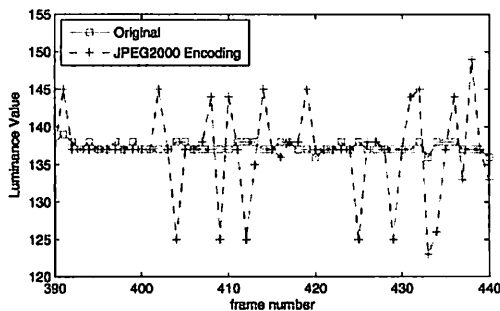


図 2 フリッカ雑音による輝度値変化

みに加え、フレーム間の差による歪みの問題が生じる。一般に非可逆符号化では復号画像は量子化誤差の分だけ原画像とは異なる。そして、原画像と復号画像との差分がフレーム毎に異なる場合、視覚的に動きを伴う雑音効果として知覚される。

Motion JPEG2000 はフレーム間の相関を利用せず、各フレーム毎に独立して符号化を行う。連続したフレームで観察すると、発生するノイズの大きさと位相がフレーム毎に変化するため、時間方向のノイズ成分が発生し、視覚的にはちらつきを伴う画質劣化として捉えられる。静止画像の場合にはさほど目立たなかった歪みであっても、動画では知覚される場合が多くなる。低ビットレートの場合には、視覚的に更に顕著である。

図 2 にフリッカ雑音による輝度値変化の様子を示す。観察対象は Streetcar の座席 (117, 137) における画素であり、原画像と各フレームを表 1 の条件で符号化したものを比較した。原画像は輝度値 137 でほぼ一定値を取っているのに対して、JPEG2000 で符号化したものは輝度値が不規則に上下に大きく変動している。この歪みによる不正な輝度値の変動が、映像がちらついて見える原因である。特に、変化の少ない静止領域で輝度値の変動が表われると、周囲との差からちらつきが顕著に感じられる。本来動きのない領域で輝度値が変動することが映像がちらつい

て見える原因になっており、この変動を抑えることでフリッカ雑音を低減させることができる。

また、このノイズは静止画像ではさほど目立たないという性質から、PSNR や MSE 等の尺度での評価、比較が困難である。PSNR はフレーム毎の算出であり、これを動画に適用した場合、その評価は各フレームの静止画像の評価の集合になる。これは時間的なちらつきなどを評価するには適していない。フリッカ雑音を定量的に表すには時間方向の変化を考慮に入れた、動画としての画質評価尺度が必要である。

3. Motion JPEG2000 符号化により発生する歪みの特性

後述するフリッカフィルタでは、過去の同一座標における画素値を用いてフレーム間での画素値の変動を抑制することでフリッカ雑音の低減をはかる。しかし、これを常に用いると移動物体周辺に劣化を生じるため、静止領域と動領域との識別が重要となる。加えて、Motion JPEG2000 符号化映像中には移動物体による動きと歪みによる動きが混在しており、それらを区別する必要がある。そこで次に、Motion JPEG2000 符号化によって発生するフリッカ雑音の特性を考える。

第 2 章で述べた歪みの特性を用いる。画像符号化では高い空間周波数に対してがコントラスト感が低くなるヒトの視覚特性を考慮して高周波成分ほど粗く量子化される。量子化雑音は主に高周波成分に発生し、歪みは画像の輪郭等の局所領域に集中しやすい。

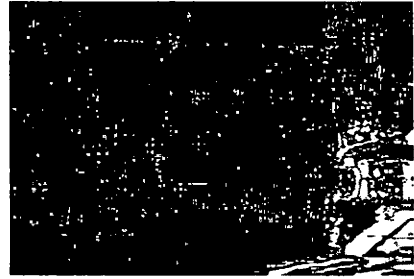
各フレームを独立に符号化する Motion JPEG2000 に対しても同様な特性が言え、各フレームに発生する歪みは局在化する。これを連続するフレーム間で考えると、フレーム間での歪み変化も画素毎のような狭い範囲での変化になりやすいと言える。これに対して、映像中の物体はある程度の大きさを持っており、その変化も領域単位の変化になる。

そこで、フレーム間における変化をある程度の大きさの領域内で平均して見ること、移動物体によるフレーム間変化はそのままに、歪みによるフレーム間変化の影響のみを小さくすることができる。と考える。

図 3 にその動領域判別の予備実験を示す。実験諸元は表 1 のとおりである。図 3(a) が予備実験に使用した符号化済み映像 Streetcar である。電車と車がカメラの前を通り過ぎるシーンで、それら以外に動く領域はない。(b) は画素単位でフレーム間差分を取った結果である。黒が変化の少ない部分で静止領域を意味する。逆に白が変化の大きい部分であり動領域である。移動物体である電車や車の領域が白く動領域として捉えられているが、同時にノイズ成分による変化も捉えてしまっている。画像全体に白くゴマ状に分布しているのが歪み成分の変化である。ここで、この歪みによる変化と動領域による変化の様子を比較してみると、先ほど述べたように歪みによって変化する領域は動領域と比べて小さいことがわかる。これに対して、(c) は 5×5 の領域内でそれを平均した結果である。画素毎に差分を取ったときと比較して、動領域のフレーム間差分値にはさほど変わりがないのに対して、ノイズ成分は全体的に低減されて、動領域と



(a) Streetcar



(b) 画素単位フレーム間差分



(c) 5×5 領域での平均フレーム間差分

図 3 静動領域判別実験

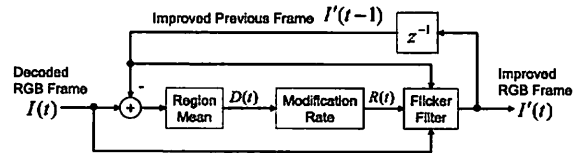


図 4 提案手法の構成

ノイズ成分との差が明確になっている。

この実験結果からも、フレーム間における変化の様子を領域単位で見ることにより、静動領域判定におけるノイズによる影響を低減させることができ、歪みと映像本来の動きとを識別しやすくなると言える。

4. 提案手法

提案手法は、フレーム間差分値の局所領域平均による静動領域判別、そしてその領域平均フレーム間差分値から補正率を決定する部分と、画素値の変動を抑制する時間方向フリッカフィルタからなる。以下、それぞれについて説明する。

4.1 領域平均フレーム間差分値による静動領域判別

第3章で述べた領域平均フレーム間差分 D により、静動領域判別を行う。この領域平均フレーム間差分は次式のように各画素毎に計算する。 ζ は平均フレーム間差分値を求める座標 (x, y) を中心とする周囲近傍領域 N 画素を表している。この近傍領域 ζ 内における各画素 (i, j) のフレーム間差分値の絶対値平均を計算し、領域平均フレーム間差分値 D とする。

$$D(x, y, t) = \frac{1}{N} \sum_{(i, j) \in \zeta} |I(i, j, t) - I(i, j, t-1)| \quad (1)$$

4.2 補正率設定

領域平均フレーム間差分 D から各画素における補正率 $(0 \leq R \leq 1)$ を決定する。補正率は現在フレームにおける画素値 $I(x, y, t)$ を前フレームの画素値で修正する割合を表している。 $R=0$ のとき補正なし、 $R=1$ で前フレームの画素値を使用する割合が100%となる。

領域平均フレーム間差分値が小さく静止領域に当たる画素には補正率を高くし、フレーム間での画素値の変化を強く抑制する。逆に平均フレーム間差分値が大きく、動領域に該当する画素では補正率を低く設定する。

今回、領域平均フレーム間差分値と補正率間の関係を定める関数 f は、補正率 $R=1$ となるデッドゾーン範囲 T_{dz} を持ち、傾きを G_s とする右下がりの直線として定義した。

$$\begin{aligned} R(x, y, t) &= f(D(x, y, t)) \\ &= G_s \{D(x, y, t) - T_{dz}\} + 1 \end{aligned} \quad (2)$$

4.3 時間方向フリッカフィルタ

補正率 R にしたがって、現在フレームの画素値 I を補正する。補正は同一座標における前フレームの画素値と現在フレームの画素値とを加重平均することによって行う。

このとき、前フレームとして既に補正済みのもの I' を用いる。前フレームの画素値を単純に使うのではなく、過去フレームの情報を累積し使用することでフレーム間での画素値変動を抑え、静止領域で高いみずみ抑制効果を得ることができる [9]。また、過去に溜め込んだ画素値の方が視覚的にも違和感のない滑らかな映像になる。

$$I'(x, y, t) = R(x, y, t) \cdot I(x, y, t) + [1 - R(x, y, t)] \cdot I(x, y, t-1) \quad (3)$$

5. フリッカ低減実験

本章では、提案手法をフリッカの存在する映像に適用し、その性能を評価する。実験用映像は、JPEG2000 符号化器 [10] を用いて各フレームを低ビットレートで独立に符号化したものを使用した。実験諸元を表 1, 2 に示す。各映像毎のパラメータは TI を最小化するように実験的に定めたものを使用した。

5.1 輝度値変化による評価

フリッカ雑音による視覚的ちらつきはフレーム間でノイズの大きさと位相が変化することによる不規則な輝度値の変動が原因となっている。ここでは、提案手法により不正な輝度値の変化を原画像のそれと同程度に訂正できているかを評価する。

表 1 実験諸元 1: 符号化映像

Codcc	JJ2000 [10]
Input Sequence	1:Streetcar, 2:Boy and Toys, 3:Whale Show (352 × 240, 24-bit color scale, 450frames)
Wavelet Filter	9/7-Non Reversible DWT
Decomposition Level	5
Bit Rate	0.375 bpp

表 2 実験諸元 2: パラメータ

Input Sequence	フレーム間差分平均化領域 ζ	補正率関数 R
1. Streetcar	5 × 5 領域	$R = -\frac{1}{23}(D-2) + 1$
2. Boy and Toys	5 × 5 領域	$R = -\frac{1}{29}D + 1$
3. Whale Show	7 × 7 領域	$R = -\frac{1}{37}(D+16) + 1$

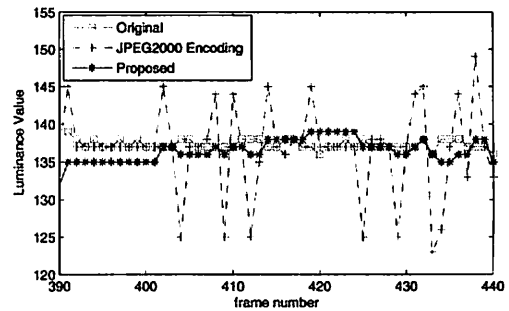


図 5 提案手法処理後の輝度値変化

図 5 に実験映像 Streetcar の座標 (117, 137) における原画像、JPEG2000 符号化後および提案手法処理後映像の定点観測の様子を示す。観測点は動きのない画面中央の建物の一点を選んだ。原画像の輝度値は 137 でほぼ一定になっているのに対して、JPEG2000 符号化後ではフリッカ雑音による大きさ 10 前後の不正な輝度値の変化が生じている。提案手法処理後ではこの輝度値の変動を抑え、原画像と同じほぼ一定の輝度値に補正できていることがわかる。

次に、図 6 に実験映像 Streetcar の 164-165 フレーム間におけるフレーム間差分画像を示す。図 6 は、図 3 と同じフレーム間の差分を観察したもので、カメラ前を電車と車が通り過ぎるシーンである。原画像におけるフレーム間差分を示した図 6(a) では移動物体である電車と車の領域のみが変化しているのに対して、(b) の JPEG2000 符号化を行った映像ではこれにフレーム間での歪みの変化が加わっている。ここに提案手法により補正を行ったものが (c) であり、画像全体に霞状に広がっていたノイズ成分の変化を大きく抑制していることが確認できる。

5.2 PSNR による評価

提案手法による画質改善効果の評価として、JPEG2000 符号化映像と提案手法処理後の映像の PSNR (Peak Signal to Noise Ratio) を比較する。PSNR は画像品質の客観的評価指標として用いられるものであり、以下のように定義される。

$$PSNR[dB] = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (4)$$



(a) 原画像



(b) JPEG2000 符号化



(c) 提案手法後

図6 実験映像のフレーム間差分

$$MSE = \frac{1}{N_1 N_2} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} \{p(x,y) - p'(x,y)\}^2 \quad (5)$$

ここで、 N_1 および N_2 が画像の横と縦の画素数であり、 $p(x,y)$ は原画像の座標 (x,y) の画素値、 $p'(x,y)$ 評価画像の座標 (x,y) の画素値である。

一般に動画像になると、時間方向の動きによる誤差が知覚されやすくなる。動画像における PSNR はフレーム毎の PSNR を計算するため、視覚的な誤差とは必ずしも一致しない。ここでは、画像がどれだけ原画像と近いものになっているかを評価するために PSNR を求める。

図7に各実験映像で測定した JPEG2000 符号化後と提案手法処理後におけるフレーム毎の PSNR を示す。値は RGB の PSNR それぞれを平均したものである。全体的には PSNR の向上がみられるが、(b)Boy and Toys や (c)Whale Show では中には PSNR が低下しているフレームもある。PSNR が低下しているフレームは画面全体が動く部分である。静止領域では時間方向の加重平均によりフレーム間の変化を抑制することで、付加されたノイズ成分も同時に低減されるが、動領域ではそれらの関係は一

致しない。PSNR が低下したのは静止領域判別が不完全であったためであり、今後改善が必要であると考えられる。

次に、表3に各実験映像のフレーム平均 PSNR を示す。PSNR の改善度は各映像によって異なるが、全ての映像で平均的には PSNR の向上がみられる。

5.3 TI による評価

映像品質の客観的評価尺度として TI (Temporal Information) [11] を用いて提案手法を適応した映像を評価する。

TI は映像の時間的変化情報を示す指標で、2つの隣接フレーム間の違い(動き)を記述する。 t フレームの座標 (x,y) における $TI(x,y,t)$ は隣接するフレームの輝度成分 $Y(x,y,t)$ 及び $Y(x,y,t-1)$ を使って、次のように定義される。

$$TI(x,y,t) = Y(x,y,t) - Y(x,y,t-1) \quad (6)$$

評価映像と原画像とで同様に TI を求め、両者の RMSE (Root Mean Square Error) をとることで評価尺度とする。

$$\Delta TI(x,y,t) = TI_{source}(x,y,t) - TI_{destination}(x,y,t-1) \quad (7)$$

$$TI_{RMSE}(t) = RMS(\Delta TI(x,y,t)) \quad (8)$$

$$TI_{RMSE}(t) = \sqrt{\frac{1}{P} \sum_x \sum_y \{\Delta TI(x,y,t)\}^2} \quad (9)$$

$$P = \text{Total number of pixels} \quad (10)$$

TI に基づく品質パラメータはノイズ等によって追加された、または失われた動きを示すものである。 TI_{RMSE} の値が低いほど評価映像の動きが原画像と類似していることを表しており、視覚的歪みが少ないことを示す。

次に、図8に各実験映像におけるフレーム毎の TI_{RMSE} をグラフで示す。全フレームでの TI_{RMSE} の向上がみられ、映像中で静止領域が多くなるところでは特に高いちらつき低減効果が確認できる。

また、表4のように平均的にも 1.2~1.4 程度の TI_{RMSE} の向上がみられる。

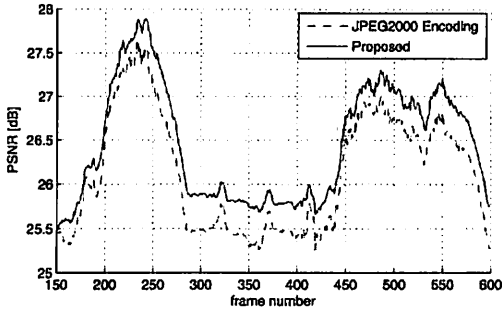
6. まとめ

本稿では、復号器側におけるフリッカ雑音低減手法を提案した。過去フレームにおける同一座標の画素値情報を使った時間方向フリッカフィルタとフレーム間差分値の局所領域平均を用いた静止領域判別を組み合わせることで、フレーム間での画素値の変動を抑え、フリッカ雑音を低減することができた。また、提案手法を適用した映像を確認することによっても主観的画質向上が認められた。

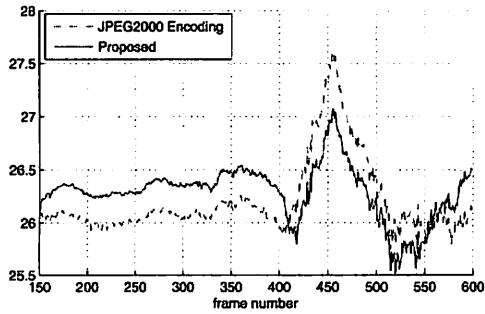
今後は、提案手法の定式化を加え、フリッカ低減の理論的及び実験的検証を行っていく予定である。

文 献

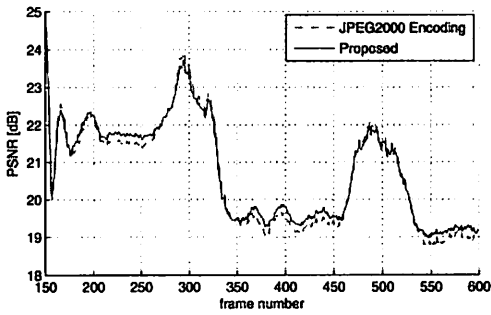
- [1] 小野定康, "デジタルシネマの規格化の動向", 映像情報メディア学会誌, Vol.59, No.2, pp.214-217, Feb. 2005.
- [2] 南憲一, "デジタルシネマ共同トライアル「4K Pure Cinema」", 映像情報メディア学会誌, Vol.61, No.5, pp.587-590, May. 2007.
- [3] "Digital Cinema System Specification Version1.1", Digital Cinema



(a) Streetcar

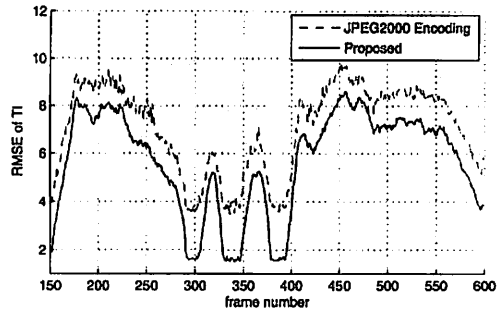


(b) Boy and Toys

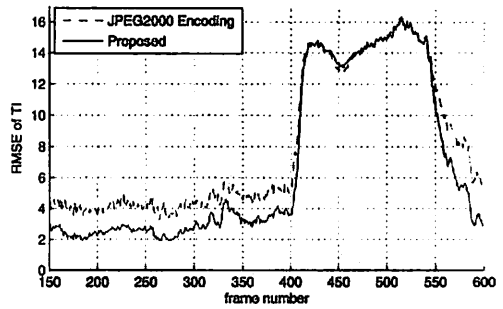


(c) Whale Show

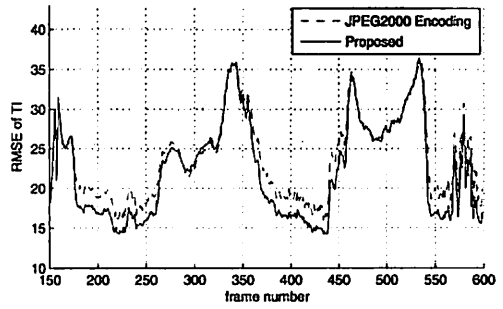
図 7 PSNR



(a) Streetcar



(b) Boy and Toys



(c) Whale Show

図 8 TI_{RMSE} (Temporal Information)

Initiatives, LLC, (April 12, 2007), <http://www.dcmovies.com/>

- [4] 小野定康, 鈴木純司, わかりやすい JPEG2000, オーム社, 東京, 2003.
- [5] 福原隆浩, 板倉英三郎, JPEG2000 詳細解説, CQ 出版, 東京, 2004.
- [6] M.Bolick, C.Christopoulos and E.Majani, "JPEG2000 Part I Final Committee Draft Version 1.0", ISO/ITE JTC1/SC29/WG1, Apr. 2000.
- [7] T.Fukuhara and D.Singer, "Motion JPEG2000 Final Committee Draft 1.0", ISO/IEC JTC1/SC29/WG1, Mar. 2001.
- [8] 伊谷祐介, 渡辺裕, "Motion JPEG 2000 における静動領域判定に基づくフリッカー低減手法の検討", 電子情報通信学会技術報告書, Vol.105, No.164, pp.77-82, July. 2005.
- [9] 桑野良範, "ウェーブレット変換符号化動画画像における歪み低減に関する研究", 学士論文, 新潟大学, Mar. 2006.
- [10] Official JJ2000 Homepage, <http://jj2000.epfl.ch/>
- [11] J.Berts and A.Persson, "Objective and subjective quality assessment of compressed digital video sequences", Master Thesis, Chalmers University of Technology, 1998.

表 3 各映像のフレーム平均 PSNR

Input Sequence	JPEG2000 符号化	提案手法
1. Streetcar	26.141	26.466
2. Boy and Toys	26.206	26.287
3. Whale Show	20.689	20.765

表 4 各映像のフレーム平均 TI_{RMSE}

Input Sequence	JPEG2000 符号化	提案手法
1. Streetcar	7.079	5.679
2. Boy and Toys	8.041	6.848
3. Whale Show	23.774	22.331