

# MSE を指定可能なレート制御に基づく 高品質な JPEG2000 画像符号化 Rate Control Method Given a Target MSE for High Quality JPEG 2000 Image Coding

桑原 翔 西川 清史 貴家 仁志

首都大学東京大学院 システムデザイン研究科 情報通信システム工学専修

Sho KUWABARA, Kiyoshi NISHIKAWA, and Hitoshi KIYA

Dept. of Information and Communication Systems Engineering, Tokyo Metropolitan University

**アブストラクト** 本稿は、JPEG 2000 符号化方式を対象として、符号化画像のファイルサイズを抑えると同時に、高品質を保証する符号化手法を提案する。JPEG 2000 におけるレート制御では、符号化レートと符号化ひずみを用いて PCRD 最適化が行われる。提案法では、PCRD 最適化において、原画像に対する符号化画像の MSE を制約に加えることで、符号化後の画質を保証する。提案法を用いることで、指定した画質を保証し、かつ効率的に符号量を削減することが可能であることを示す。また、従来の PCRD 最適化アルゴリズムに基づく手法であることから、レート制御による処理の増加を抑えることが可能であることを示す。

## 1 はじめに

近年、ブロードバンド通信の普及によって映像配信の需要が高まり、大容量の動画をデジタルデータとして扱う機会が増えている。特に、JPEG 2000 [1] はデジタルシネマ規格 (DCI: Digital Cinema Initiatives) [2] の画像コーデックとして正式採用されており、ストレージシステムや伝送ネットワークへの負荷を軽減するために、より高圧縮率で高品質な画像符号化が要求されている。

通常の JPEG 2000 符号化では、符号化時に各フレームに対してビットレートを指定することでレート制御を行う [3]。しかし、特定のビットレートを指定した場合、一般的な動画ではフレームごとに圧縮率が異なるため、画質にばらつきが生じる。その一方で、圧縮がかかりにくいフレームの画質を保証する場合、他のフレームに必要以上に多く符号量が与えられるため、動画全体のファイルサイズが著しく増大する。

このような問題に対して、JPEG 2000 符号化のレート制御に着目し、ビットレート制限下において、安定した画質でビデオストリーミングを行う手法が検討されてい

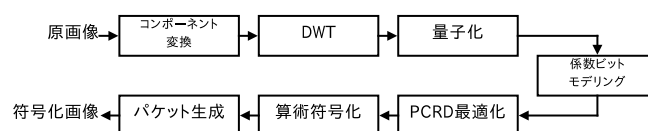


図 1: JPEG 2000 符号化の主要手順

る [4,5]。しかし、従来法は、ロッキー圧縮後の各フレームのファイルサイズに対する検討はされておらず、符号化後のファイルサイズに対する制約を課すことは困難である。本稿では、JPEG 2000 符号化方式を対象として、符号化画像のファイルサイズと画質を同時に考慮したレート制御を行う手法を提案する。

提案法では、JPEG 2000 符号化のレート制御時に行われる PCRD (Post-Compression Rate-Distortion) 最適化の際に、符号化画像の MSE (Mean Square Error) を制約に加え、指定した MSE を越えないようにフレームごとに最適な符号切り捨て点を求める。この結果、一定値以下のファイルサイズを保証し、かつ安定した画質で符号化を可能とする。また、各フレームに必要な以上の符号量が与えられることを抑え、効率的にファイルサイズを削減することが可能である。このような特徴から、ビットレート制限下におけるストリーミング配信などへの応用が期待される。

## 2 準備

ここでは準備として、JPEG 2000 符号化手順と PCRD 最適化について述べる。

### 2.1 JPEG 2000 符号化の主要手順

JPEG 2000 符号化の主要手順を図 1 に示す。原画像は、コンポーネント変換された後に DWT (Discrete Wavelet Transform) によって複数のサブバンドに分割される。サブバンド内の DWT 係数は量子化によって量子化係数に変換され、

さらにコードブロックと呼ばれる小ブロックに分解される。コードブロック内の量子化係数はその大きさが MSB(Most Significant Bit) から LSB(Least Significant Bit) に向かってビットプレーン展開された後に、SP(Significance Propagation), MR(Magnitude Refinement), CP(CleanUp) と呼ばれる 3 つの符号化パスで符号化される。符号化された係数ビットは PCRD 最適化により、各コードブロックについて最適な符号切り捨て点を符号化パス単位で求める。PCRD 最適化について詳細は 2.2 章で述べる。符号切り捨て点を求めた後に、算術符号化による無ひずみ圧縮を行い、JPEG 2000 シンタックスに基づき、符号ストリームを出力する。

## 2.2 PCRD 最適化

JPEG 2000 符号化におけるレート制御では PCRD 最適化が行われ、圧縮後の符号全体のひずみが最小となるように、符号化パスの切り捨て点がコードブロックごとに選択される。本章では、提案法に深く関わる PCRD 最適化について、その原理と手順を示す。

1.  $i$  番目のコードブロック  $B_i$  において、 $n_i$  を選択可能な符号切り捨て点とする。最適な符号切り捨て点を MSB から数えて  $k$  番目の符号切り捨て点とする。切り捨てた時のビットレートを  $R_i^{n_i}$  とし、符号ひずみを  $D_i^{n_i}$  とする。
2. 全てのコードブロックにおけるひずみの総和は、

$$D = \sum_i D_i^{n_i} \quad (1)$$

3. 指定ビットレート  $R_{budget}$  に対する符号量の総和は、

$$R = \sum_i R_i^{n_i} \leq R_{budget} \quad (2)$$

4. 式 (1), (2) の条件下で、Lagrange の未定乗数法に従い、全コードブロックにおけるひずみの総和を最小化する問題として、レートとひずみの関係を式に表せば、

$$J = D + \lambda R = \sum_i (D_i^{n_i(\lambda)} + \lambda R_i^{n_i(\lambda)}) \quad (3)$$

5. レートの総和  $R$  を最大化し、ひずみの総和  $D$  を最小化する  $\lambda$  を求め、最適な符号切り捨て点を得る。

以上が PCRD 最適化の原理である。これらの原理に基づき、PCRD 最適化では、最適な符号切り捨て点を求めるために、図 2 のような R-D スロープと呼ばれる値を有効な全ての符号切り捨て点について求める。PCRD 最適化の具体的な手順を以下に示す。

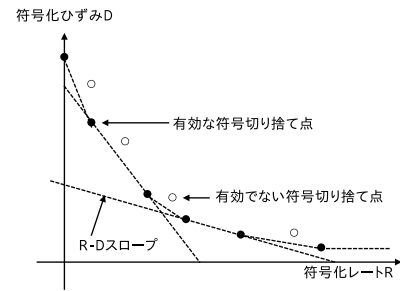


図 2: PCRD 最適化における R-D スロープ。

1. 有効な全ての符号切り捨て点  $n_i$  について、次式で定義される R-D スロープを求める。

$$S_i^{n_i} \equiv \frac{\Delta D_i^{n_i}}{\Delta R_i^{n_i}} = \frac{D_i^{n_i-1} - D_i^{n_i}}{R_i^{n_i} - R_i^{n_i-1}} \quad (4)$$

2. R-D スロープの最大値、最小値を上限値、下限値として与え、その中間の値を R-D スロープの最適値と仮定する。
3. 各コードブロックで、符号切り捨て点における R-D スロープが、仮定した値以上の範囲で最小となる点を最適な符号切り捨て点とする。
4. 決定した符号切り捨て点のもとで算術符号化以降の処理を行い、ビットレートを求める。
5. 求めたビットレートが指定ビットレートにほぼ等しければ、仮定していた最適な R-D スロープを確定し、PCRD 最適化を終了する。求めたビットレートが指定ビットレートを上回っていた場合は、下限値を仮定した R-D スロープの値に更新する。逆に下回っていた場合は、上限値を仮定した R-D スロープの値に更新する。
6. 上限値と下限値が十分に近い場合は PCRD 最適化を終了する。そうでない場合は手順 2 に戻る。

ビットレートを指定した場合、上述の手順 4 から 6 の処理は繰り返し行われる。

## 3 提案法

本章では、提案法について概要を述べる。提案法は、JPEG 2000 符号化のレート制御時に行われる PCRD 最適化において、符号化画像の MSE を制約に加えることで、従来のレート制御のようにビットレートを指定可能であると同時に、MSE を指定可能である。したがって、符号化後のファイルサイズと画質を同時に制御することが可能である。

### 3.1 MSE の指定

提案法では，R-D スロープを求める際の符号化ひずみ D に，空間領域の MSE を利用する．MSE は画像サイズを  $X \times Y$  とすると，以下の式で計算される．

$$MSE_{spa} = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y \{f(x, y) - \hat{f}(x, y)\}^2 \quad (5)$$

$f(x, y)$  原画像の位置  $(x, y)$  における画素値  
 $\hat{f}(x, y)$  符号化画像の位置  $(x, y)$  における画素値

また，PSNR(Peak Signal-to-Noise Ratio) は MSE から求められる，8bit グレースケール画像の場合，

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (6)$$

で表される．

$i$  番目のコードブロック  $b_i$  の  $p$  番目 ( $p = 0, 1, \dots$ ) のビットプレーンのある 1 つのコーディングパス  $P_i^{(p,k)}$  ( $k \in \{SP, MR, CP\}$ ) を切り捨てたときに発生するひずみ  $\Delta D_i^{(p,k)}$  は，次式より DWT 係数から求められる [6-8]．

$$\Delta D_i^{(p,k)} = G_{b_i} \times \Delta_i^2 \times 2^{2p} \times \left( \sum_{J_1} T_s(\tilde{v}_i^{(p+1)}[j]) + \sum_{J_2} T_m(\tilde{v}_i^{(p+1)}[j]) \right) \quad (7)$$

ただし，

$$J_1 : (j \in P_i^{(p,k)}, v_i^{(p)}[j] = 1) \quad (8)$$

$$J_2 : (j \in P_i^{(p,k)}, v_i^{(p)}[j] > 1) \quad (9)$$

$$T_s(\tilde{v}) = (2\tilde{v})^2 - (2\tilde{v} - \frac{3}{2})^2 \quad (10)$$

$$T_m(\tilde{v}) = (2\tilde{v} - 1)^2 - (2\tilde{v} - [2\tilde{v}_i^{(p+1)}[j]] - \frac{1}{2})^2 \quad (11)$$

ここで， $\Delta_i$  は  $i$  番目のコードブロックの量子化ステップサイズ， $v_i^{(p)}[j]$  はプレ量子化後の DWT 係数， $\tilde{v}_i^{(p)}[j]$  は  $v_i^{(p)}[j]$  を  $2^p$  で割ったとき的小数部分をあらわす．また， $G_{b_i}$  はコードブロック  $b_i$  の属するサブバンドの，ウェーブレット逆変換の符号化利得である．この値はローパスフィルタ  $g_0$  とハイパスフィルタ  $g_1$  から算出され，サブバンド毎に求められる，

### 3.2 MSE を制約に加えた PCRD 最適化

以下では，提案法における PCRD 最適化について述べる．提案法では，所望の PSNR に相当する MSE を求め，MSE を制約に加えて PCRD 最適化を行う．提案法における PCRD 最適化の手順を次に示す．

1. 各コードブロックで，有効な符号切り捨て点における R-D スロープを計算し，最大値と最小値を求め

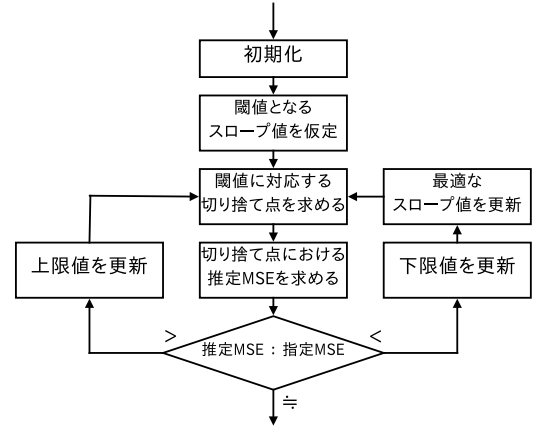


図 3: MSE 指定時の PCRD 最適化手順．

る．次に，求めた最大値，最小値を，上限値，下限値と最適なスロープ値にそれぞれ割り当てる．

2. 最適な符号切り捨て点における傾きを，上限値と下限値の中間と仮定する．以降の処理では，この傾きを閾値として用いる．
3. 各コードブロックで，切り捨て点における R-D スロープが，仮定した値以上でもっとも小さくなる点を，仮定した R-D スロープにおける最適な符号切り捨て点  $z_i$  として求める．同時に，切り捨てによって発生する推定 MSE の和  $\Delta D$  を次式によって求める．

$$\Delta D = \sum_i \sum_{(p,k) < z_i} \Delta D_i^{(p,k)} \quad (12)$$

ただし， $(p, k) < z_i$  は，符号切り捨て点  $z_i$  より下位のコーディングパスを表す．

4. 算術符号化以降の処理を行い，符号ストリームのビットレートを求める．
5. 求めたビットレートと指定ビットレートを比較する．ビットレートがほぼ等しければ，最適なスロープ値を仮定したスロープ値に更新し PCRD 最適化を終了する．求めたビットレートが上回っていた場合，下限値を仮定したスロープ値に更新し，手順 7 に進む．
6. 求めた MSE と指定 MSE を比較する．求めた MSE が上回っていた場合，上限値を仮定した傾きに更新する．下回っていた場合は，下限値と最適なスロープ値を仮定したスロープ値に更新する．
7. 上限値と下限値が十分に近ければ PCRD 最適化を終了する．そうでない場合は手順 2 に戻る．



図 4: シミュレーション 1 で使用した画像 .

表 1: シミュレーション 1 の条件 .

image	Lena
resolution	512(H) × 512(V)
format	8bit grayscale
DWT level	5 (9/7 filter)
codeblock size	64 × 64
target PSNR	1, 2, ..., 60

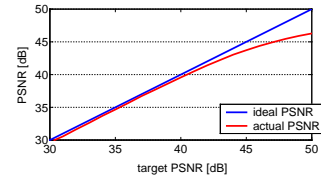
なお、MSE のみを指定する場合は手順 4, 5 を省略する (図 3 参照) .

ビットレート指定では、手順 4 で正確なビットレートを求め、手順 5 でビットレートを指定値と比較する . 一方、MSE 指定では、手順 6 で MSE を比較する際に、手順 3 で求めた推定 MSE を用いるため、誤差が生じる場合がある . 特に、全ての画素値が等しい画像などの場合、有効な符号切り捨て点が少ないため、切り捨てによる雑音が無相関ではなくなり、MSE の推定に大きな誤差が発生する可能性がある . その場合には、手順 3 で求めた MSE が指定された MSE を下回らず、最適な傾きが一度も更新されないまま PCRD 最適化が終了してしまう .

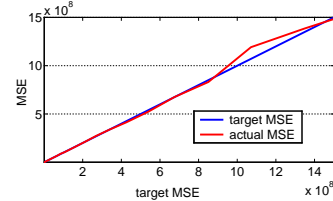
そこで、この問題を回避するために、PCRD 最適化によって最適な R-D スロープを求める際に、全てのコードブロックの有効な符号切り捨て点における R-D スロープのうち、最小のものに初期化する . その結果、最適な R-D スロープが更新されない場合は全ての符号化パスが残されることにより、符号化後の画質が他のフレームと比べて悪化するのを回避することが可能となる .

#### 4 シミュレーション

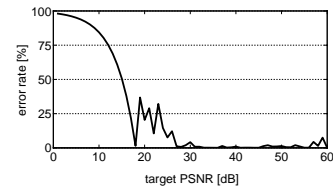
以下では、3 章で述べた手法を用いたシミュレーション結果を示し、ファイルサイズを抑え、かつ指定した画質で符号化が可能であることを確認する . 提案法に基づく PCRD 最適化を JPEG 2000 コーデックのひとつである JasPer [9] に実装し、符号化画像の評価を行った .



(a) 指定 PSNR に対する符号化画像の PSNR .



(b) 指定 MSE に対する符号化画像の MSE .



(c) 指定 MSE に対する符号化画像の MSE の誤差率 .

図 5: シミュレーション 1 の結果 .

#### 4.1 シミュレーション 1: MSE と PSNR の指定精度

図 4 の静止画像に対して、任意の PSNR に相当する MSE を指定して符号化を行い、符号化画像を復号した画像と原画像における MSE と PSNR を測定した . シミュレーション 1 の条件を表 1 示す .

指定 PSNR に対する符号化画像の PSNR を図 5(a) に、指定 MSE に対する符号化画像の MSE を図 5(b) に示す . なお、(a), (b) は PSNR=30-50[dB] に相当する範囲で結果を示している .

指定 PSNR が低い場合、符号化パスの多くが切り捨てられ、画素値が全て等しい画像となるため、符号化画像の PSNR は一定値以下にならない . また、図 5(a) より、指定 PSNR が高い場合、DWT や量子化処理による誤差が生じることから、符号化画像の PSNR は一定値以上にならない .

同様に指定 MSE が大きい場合、符号化画像の MSE は一定値以下にならない . 一方、指定 MSE を小さくした場合、図 5(b) より、符号化画像の MSE は指定 MSE により近い値となり、MSE の誤差は小さいものとみなせる .

また、図 5(c) は、指定 MSE に対する符号化画像の MSE の誤差の割合を示している . 指定 PSNR が低い場合、MSE の誤差の割合は大きいですが、30[dB] 以降では、誤差の割合は小さくなっている . 以上の結果から、実用的な PSNR (30-45[dB]) の範囲において、高精度で MSE が指定可能であることがわかる .



図 6: シミュレーション 2 で使用した動画画像 [10].

表 2: シミュレーション 2 の条件 .

条件	(a)	(b)	(c)	(d)
動画画像	Elephants Dream			
解像度	1,920(H) × 1,080(V)			
フォーマット	8bit color			
フレーム数	2,000			
DWT 分解レベル数	5 (9/7filter)			
コードブロックサイズ	64 × 64			
指定 PSNR	45 dB	-	-	45 dB
指定ビットレート	-	0.10 bpp	0.27 bpp	0.10 bpp

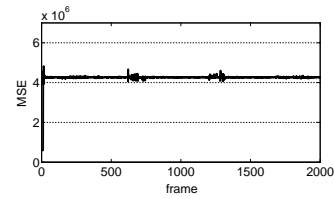
#### 4.2 シミュレーション 2: MSE とビットレートを指定した動画画像圧縮

図 6 の動画画像に対して、所望の PSNR に相当する MSE とビットレートを指定して符号化を行い、符号化画像の各フレームの MSE, PSNR, ファイルサイズ, 符号化の処理時間を測定した。シミュレーション 2 の条件を表 2 に示す。

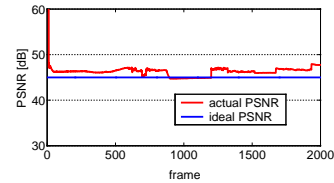
まず、提案法に従って PSNR=45[dB] に相当する MSE を指定して符号化を行った (表 2(b))。その結果を図 7 に示す。MSE を指定した場合、フレーム間の MSE, PSNR がともに安定していることが、図 7(a), (b) からわかる。なお、図中で MSE が極端に小さく、PSNR が極端に高くなっている点は、ほぼ全ての画素値が等しいフレームに相当し、MSE 推定の誤差や、有効な符号切り捨て点の少なさによって指定値を大きく外れたものと考えられる。一方、図 7(c) からわかるように、フレーム間のファイルサイズは、大きなばらつきが生じていることがわかる。

次に、従来法との比較のため、条件 (a) の結果における平均ビットレート、最大ビットレートを指定し、符号化を行った (表 2(b)(c))。その結果を図 8, 9 に示す。平均ビットレートを指定した場合、動画画像全体の符号量は条件 (a) の場合とほぼ等しくなる。8(a), (b) からわかるように、各フレームに同じ符号量が与えられているが、符号化画像の画質にばらつきが生じている。

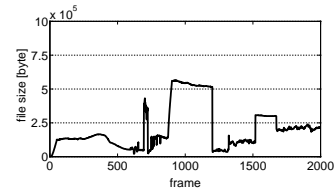
また、条件 (a) の結果における最大ビットレートを指定した場合、9(a) のように、最も圧縮のかかりにくいフレームの画質を保証することが可能である。しかし、9(b) からわかるように、条件 (a), (b) の場合の約 3 倍の符号量が与えられてしまう。以上の結果から、提案法において MSE を指定した場合、フレーム間でより安定した画質を



(a) 符号化画像の MSE .

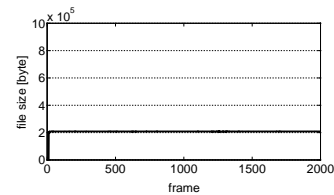


(b) 符号化画像の PSNR .

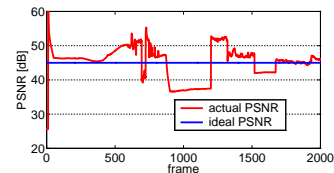


(c) 符号化画像のファイルサイズ .

図 7: PSNR=45[dB] を指定した場合の結果 .



(a) 符号化画像のファイルサイズ .

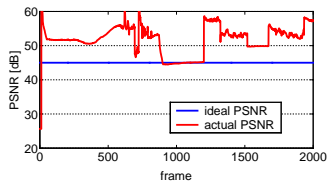


(b) 符号化画像の PSNR .

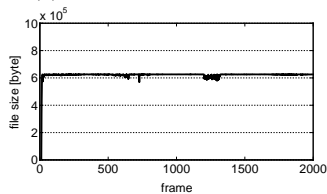
図 8: 条件 (a) の平均ビットレートを指定した場合の結果 .

保証可能であることがわかる。

更に、PSNR=45[dB]、条件 (a) における平均ビットレートを同時に指定し、符号化を行った (表 2(d))。条件 (d) では、指定ビットレート条件下で指定 PSNR に相当する MSE を保証可能であれば、最適な符号切り捨て点を選択し、指定ビットレート条件下で保証不可能であれば、そのビットレートに従って符号切り捨て点を選択する。したがって、10(a), (b) からわかるように、PSNR=45[dB] を保証可能なフレームでは、45[dB] を大きく上回るフレームが存在せず、各フレームに与えられる符号量も一定値以下に制

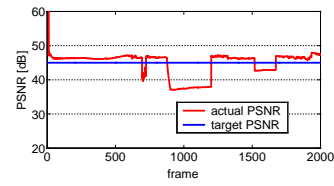


(a) 符号化画像の PSNR .

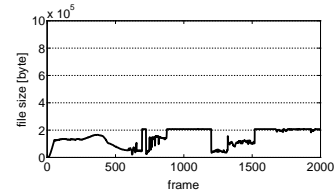


(b) 符号化画像のファイルサイズ .

図 9: 条件 (a) の最大ビットレートを指定した場合の結果 .



(a) 符号化画像の PSNR .



(b) 符号化画像のファイルサイズ .

図 10: PSNR とビットレートを同時に指定した場合の結果 .

表 3: エンコード時間の測定結果 .

	(a)	(b)	(c)	(d)
エンコード時間 [sec]	1,927	2,061	1,986	2,047

限される。つまり、各フレームに必要な最小限の符号量を与えることで、画質とファイルサイズの両方を考慮し、符号量を効率的に削減することが可能であることがわかる。

また、各条件における動画像エンコード時間を測定した結果を表 3 に示す。提案法は従来法に比べてエンコード時間の増加は小さく、レート制御における処理の増加を抑えることが可能であることが確認できる。

## 5 おわりに

本稿では、MSE を指定可能なレート制御法による、高品質で高圧縮率な JPEG 2000 画像符号化手法を提案した。提案法は PCRD 最適化時に MSE を制約に加えて符号切り捨て点を選択することで、各フレームのファイルサイズと画質の両方を考慮した符号化を行い、効率的に符号量が削減可能であることを示した。また、従来の PCRD 最適化アルゴリズムに基づく手法であることから、レート制御による処理の増加を抑制可能であることを示した。

## 参考文献

[1] I.T.R.T. ISO/IEC 15444-1, "Information technology- JPEG 2000 image coding system-Part1: Core coding system," ISO/IEC JTC 1/SC 29 WG1, Jan. 2001.

[2] Digital Cinema Initiatives, LLC Member Representatives Committee, "Digital cinema system specification V1.0," Final Approval July 2005.

[3] Taubman, D. "High performance scalable image compression with EBCOT," IEEE Trans. Image Processing, vol.9, no.7, pp.1151-1170, July. 2000.

[4] Y.M.Yeung, Oscar C.Au, "Efficient Rate Control for JPEG 2000 Image Coding," IEEE Trans. Circuits And Systems For Video Technology, vol.15, no.3, March. 2005.

[5] Joseph C. Dagher, Ali Dilgin, and Michael W.Marcellin, Fellow "Resource-Constrained Rate Control for Motion JPEG2000," IEEE Trans. Image Processing, vol.12, no.2, Dec. 2003.

[6] D. Taubman and M. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, Boston, U.S.A., 2001.

[7] K.Nishikawa, S.Nagawara, and H.Kiya, "QoS Estimation Method for JPEG 2000 Coded Image at RTP Layer," IEICE Trans. Fundamentals, vol.E89-A, no.8, pp.2119-2128, Aug. 2006.

[8] K.Nishikawa, K.Munadi, and H.Kiya, "No-Reference PSNR Estimation for Quality Monitoring of Motion JPEG2000 Video Over Lossy Packet Networks," IEEE Trans. Multimedia, vol.10, Issue 4, pp.637-645, June. 2008.

[9] M. D. Adams, JasPer Project Home Page, [Online.] Available:<http://www.ece.uvic.ca/mdadams/jasper2000>.

[10] "Elephants dream," <http://media.xiph.org/>.