

Motion JPEG2000 低遅延符号化のための時間方向ビット配分

佐野 雄磨[†] 内藤 整^{††} 渡辺 裕[†]

[†] 早稲田大学 大学院 国際情報通信研究科

〒 169-0051 東京都新宿区西早稲田 1-3-10

^{††} 株式会社 KDDI 研究所

〒 356-8502 埼玉県上福岡市大原 2 丁目 1 番 15 号

E-mail: †yuma@tom.comm.waseda.ac.jp

あらまし Motion JPEG2000 は、イントラ符号化のため、各フレームが独立に符号化される。そのため、各フレームに同等のビット配分を行った場合、各フレームの特徴により、フレームごとに品質にばらつきが生じ、シーケンスとしての品質は必ずしも良好にはならない。この問題に対する従来手法として、バッファに数フレーム蓄積し、ビット配分を行う方法があるが、その場合バッファリングによる遅延のため、リアルタイム伝送には不適である。そこで、本稿では各フレームの RD 特性を予測することによって、バッファリングによる遅延を生じることなく、各フレームに等しい品質を与える方法について検討する。

キーワード Motion JPEG2000, HDTV, 低遅延符号化, ビット配分, RD 最適化

Temporal bit allocation for low delay coding by using Motion JPEG2000

Yuma SANO[†], Sei NAITO^{††}, and Hiroshi WATANABE[†]

[†] Graduate School of Global Information and Telecommunication Studies, Waseda University

1-3-10 Nishiwaseda, Shinjuku-ku, Tokyo, 169-0051 Japan

^{††} KDDI Labs

2-1-15 Ohara, Kamihukkuokashi, Saitama, 356-8502 Japan

E-mail: †yuma@tom.comm.waseda.ac.jp

Abstract Since Motion JPEG2000 is intra based coding system, each frame is coded separately. Therefore, if the same amount of bits are allocated for each frame, quality of each frame varies and quality as whole sequence decreases. To solve this problem, the conventional approach has been exploiting temporal bit allocation by storing some frames in a buffer. However, it is not suitable for real time transmission. In this paper, we propose a bit allocation method by predicting rate-distortion characteristic of following frames without coding delay and variation of quality.

Key words Motion JPEG2000, HDTV, low delay coding, bit allocation, RD optimization

1. はじめに

デジタル放送の開始やプラズマテレビの登場に伴って、映像素材として HDTV が用いられつつある。HDTV 素材の高品質伝送を目的として、高い圧縮率とスケーラビリティを持つ JPEG2000 [1] による HDTV 符号化の研究が行われてきた。放送局間の素材伝送や番組配信において HDTV を用いる場合、ユーザ要求として伝送遅延の低減が挙げられる。この問題に対し、筆者らは JPEG2000 の Precinct 構造を用いることによって、低遅延で符号化を行う方法を検討した [2]。

しかし、Motion JPEG2000 は、イントラ符号化のため、各フレームが独立に符号化される。そのため、各フレームに同等のビット配分を行った場合、各フレームの特徴により、フレームごとに品質にばらつきが生じ、シーケンスとしての品質は必ずしも良好にはならない。

そこで、各フレームの特徴を考慮して、シーケンスとしての品質が良好になるようなビット配分を検討する。

2. 時間方向の符号量制御

2.1 CBR と VBR

MPEG に代表される動画像圧縮の符号量制御の方法として CBR(Constant Bit Rate) 方式と VBR(Variable Bit Rate) 方式がある。CBR 方式は、各フレームの符号量を一定に制御する方法で、バッファリングの必要が無く低遅延で符号化することが可能である。しかし、フレームごとに画像の特徴が異なるため、各フレームに同じ量のビットを与えた場合、品質にばらつきが生じ、シーケンスとしての品質は必ずしも良好にはならない。一方で、VBR 方式は、各フレームの特徴によって符号量を可変にすることが出来、CBR と比較して高品質を実現することが可能である。しかし、VBR を実現するためには、他フレームの特性を考慮して符号化を行うために、バッファリングによる符号化遅延が生じるという欠点がある。

2.2 Motion JPEG2000 における時間方向符号量制御

Motion JPEG2000 は、イントラ符号化のため、各フレームが独立に符号化される。そのため、MPEG のようにフレーム間予測などを利用して時間方向の符号量制御を行うことは出来ない。また、各フレームは画像の複雑さによって圧縮のしやすさが異なるので、各フレームに同じビット量を与えると画質を低下させる。この問題に対する従来手法として、バッファに数フレーム蓄積し、各フレームの特徴を考慮してビット配分を行う方法がある [3] [4]。その場合バッファリングによる遅延が発生するため、リアルタイム

ム伝送には不適である。そこで、リアルタイム伝送のために、フレームをバッファに蓄積することなく、他フレームの特徴を考慮し、ビット配分を行う方法が必要となる。

3. RD 最適化

3.1 レート歪理論

レート歪理論は、シャノンによって提唱されたもので「信号の誤差(歪と雑音)の平均値 D を許容する時、画素あたりのビット数の最小値の限界 $R(D)$ を与える理論」である [5]。この $R(D)$ を rate-distortion 関数といい、通信容量 C なる通信路がある時、 $C > R(D)$ であれば歪 D 以内で伝送されることが示されている。レート歪理論における RD(rate-distortion) の関係を示したグラフを図 1 に示す。

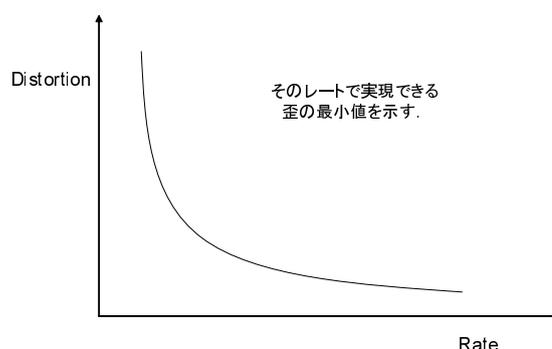


図 1 rate-distortion の関係

3.2 JPEG2000 における RD 最適化

3.2.1 EBCOT

JPEG2000 では、符号化手順の中で、ウェーブレット変換を行った係数をビットプレーン展開する。さらに、このビットプレーンをビットの種類によって、3 種類の符号化パスに分類する。これを EBCOT(Embedded multiple component transformation) 処理 [1] という。この符号化パスはさらに算術符号化される。

3.2.2 スカラー量子化とポスト量子化

JPEG2000 では、量子化の方法としてスカラー量子化とポスト量子化が存在する。スカラー量子化は、MPEG などと同様に、量子化ステップ Q によって係数を量子化する方法である。一方、ポスト量子化は上述の EBCOT 処理で得たビットプレーンを符号化パスを単位として、RD 最適化のもとデータを切り捨て符号量制御する方法である。スカラー量子化では、符号量制御を行うために、量子化ステップ幅を適宜変更しながら符号化を行い、符号量の制御を行う必要がある。一方で、ポスト量子化では、符号化後の符号列に対して切捨てを行うので、再度量子化を行う必要が無い。従って、細かい量子化を行う際には非常に有効な方法である [6]。

3.2.3 RD最適化

JPEG2000では、画像を符号化した後に、目標とするビットレート、歪(画質)、に適合するように、RD最適化を用いて、各コードブロックのビットを調整する。歪の単位としては、MSE(Mean Square Error)が用いられる。

まず、画像全体の歪を D とすると、 D は以下のように表される。

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

式(1)において、 i はコードブロック、 n_i はコードブロック i におけるビット切捨て点を表し、 $D_i^{n_i}$ はコードブロック i の歪である。歪の計算には、ウェーブレット変換が直交変換であることより、MSEまたは重み付け MSE が用いられる。また画像全体の符号量を R とすると、 R は以下のように表される。

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

式(2)において、 $R_i^{n_i}$ はコードブロック i の符号量である。ここで、 R を最小にするためには、 D が最小となる n_i を決定することが必要である。これを解決するための方法として、Lagrange 法を利用でき、式(3)で示される値を最小にすることによって最終的に最適歪となる符号列を得られる [6] [7]。

$$\sum R_i^{n_i} - \lambda D_i^{n_i} \quad (3)$$

そこで、各符号化パスで、 ΔR と ΔD を用いて、

$$S_i^k = \Delta D / \Delta R$$

を計算で求め、ビット切捨て点を決定する。(k は符号化パスを表す。)

4. 提案手法

4.1 提案手法の概要

CBR 方式で符号化した際に、フレームごとに品質が異なってしまうのは、画像の特徴によって、同品質を得るために必要なビット量が異なるからである。MPEG における時間方向の符号量制御における従来研究は、量子化ステップをそろえることによって、品質の変動を抑えていた [8] [9] [10] [11]。しかし、JPEG2000 は、EBCOT 処理によって符号量制御をビットプレーン単位で、より細かく行うことが可能である。そこで、各フレームの RD 特性を比較することによって、各フレームの特徴を考慮したビット配分を行う。

JPEG2000 の符号化処理では、RD 最適化の際に各符号化パスにおいて、 ΔR が ΔD が求められている。従って、符号量制御にあたり RD 特性を用いることは、新たな処理を必要とせず符号化遅延は生じない。

4.2 Rate と slope の関係

前章で述べたように、RD 最適化において、各符号化パスで $\Delta D / \Delta R$ を求めて、ビット切捨て点を決定している。本稿では、 $\Delta D / \Delta R = \text{slope}$ と名づける。slope は、ある符号化パスでビットを切り捨てた時に、切り捨てたビット量 ΔR に対してどれだけ歪が生じるかを表す。ビット量が同じ符号化パスでも、slope が大きければ、切り捨てた時の歪の増加が大きく、slope が小さければ slope が大きい時と比較して、切り捨てた時の歪の増加は、小さくなる。一般的に、複雑で圧縮しにくい画像と単純で圧縮しやすい画像を比較した場合、同じビット量を実現するビット切捨て点の slope は、複雑な画像の方が大きな値をとる。これを図 2 に示す。

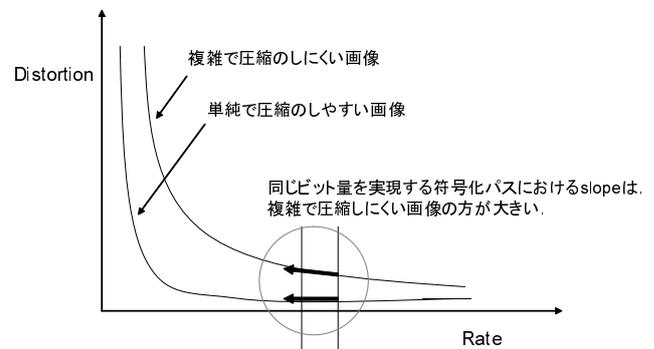


図 2 Rate と slope の関係

4.3 Rate-slope の予測

各フレームの RD 特性は、各フレームを符号化するまで得ることが出来ない。従って、現在のフレーム以外の RD 特性を知るためには、他フレームの符号化を待つ必要がある。しかし、他フレームの符号化を待っている間は、大きな符号化遅延が発生するため、リアルタイム伝送には不適である。そこで、すでに符号化したフレームの RD 特性から、後続フレームの RD 特性を予測し、符号量制御に用いる。具体的には、各フレームについて目標ビットレートを実現するビット切捨て点の slope を記録しておき、それがフレームごとにどのように変化していくかを予測する。その予測された後続フレームの slope をもとにして、ビット配分を行う。

4.4 slope の決定

前述のとおり、同じビットレートを実現するビット切捨て点における slope が大きければ、ビットを切り捨てた時も歪が大きく、slope が小さければ、ビットを切り捨てた時の歪が小さくなる。これは逆に、slope が大きければ、ビット量を少し与えただけでも歪が大きく減少し、slope が小さければ、ビット量を少し与えても歪はあまり減少しないことを意味する。そこで、フレームを時間方向に見て、slope の小さい

フレームにはビット量を目標ビットレートよりも少なめに与え、slope の大きいフレームにはビット量を目標ビットレートよりも多めに与えることによって、時間方向の RD 最適化を行う。この概念を図 3 に示す。

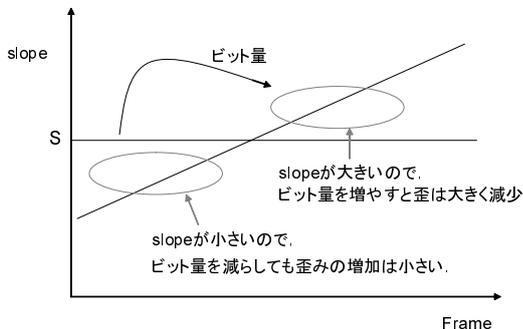


図 3 目標ビットレートを得る slope の変化

注目している現在のフレームに対して、すでに符号化した前フレームから後続フレームの Rate と slope の関係を予測する。そして、現在のフレームの前 N フレームと予測した後続 N フレームの計 $2N+1$ フレームで符号量制御を行う。 $2N+1$ フレームの中で、中間的な slope の値を求めて、現在のフレームを求めた slope を持つ符号化パスでビットを切り捨てる。 $2N+1$ フレームの R-slope 曲線の積分範囲と、面積が同じになるような slope を求め、現在のフレームの slope とする。そして、この slope を持つ符号化パスをビット切捨て点としてポスト量子化を行う。現在のフレームを符号化後、次のフレームに移り同様に前後 $2N+1$ フレームを利用して符号量制御を行う。slope の決定方法を示したものを図 4 に示す。

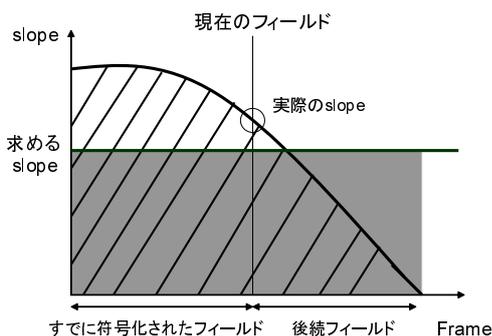


図 4 slope の決定方法

4.5 割り当てビット量の制御

本手法では、各フレームは slope を基準として符号量を制御している。したがって、CBR とは異なり、最終的に何ビット割り当てられるかはわからない。そこで、まず符号量制御を行うフレーム数を M とする。目標ビットレートから、 M フレームに割り当てられる総ビット量を計算する。すでに割り当てたビット

量と平均ビットレートを比較することによって、今後のフレームで何ビット割り当てられるかを計算する。slope によって得られた結果とこの割り当てビットの計算から、符号量を制御する。フレーム i に割り当てたビット量を S_i とおくと、 n フレーム符号化した時点で割り当てられているビット量は $\sum_{i=0}^{n-1} S_i$ とおける。これに対し、CBR で符号化した際の平均ビットレートを P 、フレームレートを F とすると、平均ビットレート P で符号化した場合の割り当てビット量は $n \times P/F$ で表される。従って、 n フレーム符号化が終わった時のビットの過不足を V で表すと、

$$\text{過不足 } V = \sum_{i=0}^{n-1} S_i - n \times P/F$$

と表される。この過不足がある限界値を超えないように、ビット量の割り当てを制御する。

n フレームの符号化が終わった時、ビット量の過不足とその後何フレームを符号化するかから、現在のフレームに割り当てるビット量を決定する。割り当てビットの制御の概念を図 5 に示す。

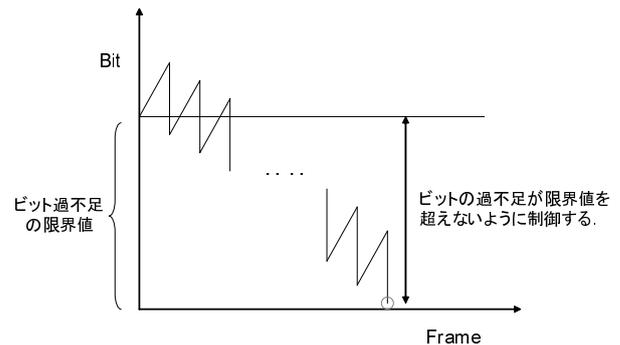


図 5 提案方式と CBR 方式のビット量過不足による制御

5. 実験とその結果

5.1 実験の方法

本手法の実験を行った。素材には ITE 標準画像の Streetcar を用いた (1920×1080 , 24bpp)。本手法では、フィールド符号化を想定しているため、奇数フィールドと偶数フィールドを分解し、60 フィールド/秒で符号化を行った。符号化条件としては、DWT9X7 フィルタを用い、3 段階 Wavelet 変換を行った。プログRESSIONは PCRL である。目標ビットレートは 0.8bpp とした。本実験では、全体で 240 フィールドを符号量制御の対象とし、ビット配分の単位とするフィールド数は注目フィールドの前後 30 フィールドとしてビット配分を行った。ちなみに、最初の 30 フィールドに関しては、すでに符号化した全フィールド (30 枚以下) と後続 30 フィールドを符号量制御の単位としている。また、後続フィールドの R と slope の関

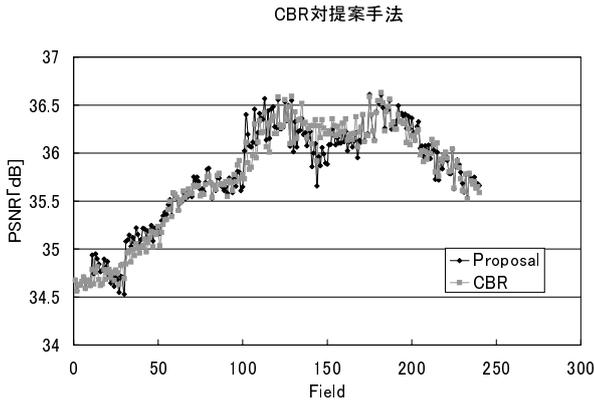


図 6 CBR 方式と提案手法の比較

係を予測する必要があるために、最初の 10 フィールドに関しては、CBR で各フィールドに等しく 0.8bpp とした。

5.2 実験の結果

実験から得られた結果を図 6 に示す。全フィールドを 0.8bpp として符号化した時と、提案手法を用いて符号化した場合の各フィールドの PSNR を示している。

結果より、フィールドを追うにつれて PSNR が上昇するような傾向がある場合は、PSNR が小さい範囲で提案手法は CBR よりも高い PSNR を得ている。これは、フィールドが進むにつれて品質が良好になる (同じビットレートを実現するビット切捨ての slope の値が減少する) 傾向を予測し、現在のフィールドが大局的に見れば、品質の悪いフィールドであると判断するためである。その結果、現在のフィールドでは CBR 方式による符号化の時と比較して低い slope を設定するため、歪が減少し PSNR が増加した。一方で、PSNR が減少する傾向がある場合は、CBR 方式と比較して PSNR の値が低くなるという結果が得られた。これは、後続フィールドで品質がさらに劣化するものと予測し、現在のフィールドが大局的に見れば品質の良いフィールドであると判断するためである。その結果、現在のフィールドでは、CBR 方式の場合と比較して高い slope を設定するため、歪が増加し PSNR の値が減少した。

MSE の平均と分散を表 1 に示す。

	MSE の平均	MSE の分散
CBR 方式	17.415	5.647
提案手法	17.381	5.206

CBR 方式と提案手法は、MSE の平均値については、ほぼ同等の値をとるが、提案手法の方がやや上回っている。また、MSE の分散に関しては、提案手法が CBR 方式と比較して小さい値をとっているこ

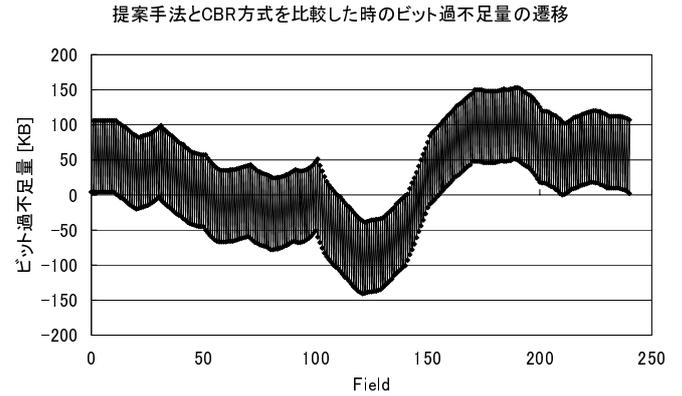


図 7 提案手法と CBR 方式を比較した時のビット過不足量の遷移

とがわかる。従って、提案手法が CBR 方式と比較して品質の変動が少ないことがわかる。

また、提案手法と CBR 方式を比較した時のビット過不足量の遷移を表したものを図 7 に示す。本実験では、ビット過不足の限界値を CBR 方式のビットレートでフィールド二枚分としている。(CBR 方式では、1 フィールド 102KB。従って、過不足限界値 = 204KB)

5.3 考察

全体として、CBR と同等の品質で、かつ品質の変動を小さく抑えることに成功している。しかし、現フィールドに対して、後続フィールドを予測によって求めているため、予測で得た slope が実際の後続フィールドの slope と一致しない場合がある。この場合、割り当てるべきビット量を誤っているため、品質も低下することになる。後続フレームをどのように予測するかによって与えるべきビット量そして品質に影響を与える。あまりに、多くの後続フィールドに対して、R と slope の関係を予測すると、実際の R と slope の値との誤差が大きくなる恐れがある。また、R-slope 曲線の傾向が変化したり、シーンチェンジが起こる場合は、多くの後続フィールドの R-slope の関係を予測するのは困難であり、大きな誤差を生じることになる。そこで、今回は、注目している現在のフィールドの前後 30 フレームずつをビット配分の単位とした。結果からわかるように、CBR 方式と比較して、MSE の平均値は同等で分散も小さくなっているが、ビット配分の範囲が合計 61 フィールド分と小さいため、CBR 方式との差は小さく比較的類似した結果となっている。従って、予測誤差を小さくしつつ、かつより品質の変動が抑えられるような、ビット制御の方法をさらに検討する必要がある。

6. まとめ

本稿では、MotionJPEG2000 を用いて動画像を符号化する際に、低遅延で時間方向の符号量制御を行う方式について検討した。低遅延で時間方向の符号量制

御を行うために、まず各フレームの Rate-Distortion 特性に注目した。さらに、各フレームの複雑度(圧縮のしづらさ)を $slope = \Delta D / \Delta R$ で表し、時間方向のビット配分の基準とした。その上で、各フレームの R と slope の関係から、後続の R-slope の関係を予測し、すでに符号化の終わった過去のフレームと、予測した後続のフレームから、品質の変動を抑えるように、slope を設定して符号化を行った。提案手法によって、CBR 方式と比較して、シーケンス全体として MSE の平均値がわずかであるが上昇し、MSE の分散も小さく抑えられることがわかった。今後の課題として、後続フィールドを何フィールド予測して、符号量制御に用いるかを検討し、適切な予測フィールド数を算出する必要がある。また、後続フィールドの予測自体の精度を上げる必要もある。さらに、シーンが激しく変動するようなシーケンスに対しても有効なビット配分方法を考える必要がある。

文 献

- [1] ISO/IEC JTC1/SC29/WG1 (ITU-T SG-8) N1646R Coding of Still Pictures: "JPEG2000 Final Committee Draft Version1.0," Mar,2000.
- [2] 佐野,内藤,渡辺, "JPEG2000Precinct 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] J. C. Dagher, A. Bilgin, and M. W. Marcellin, "Resource-Constrained Rate Control for Motion JPEG2000," IEEE Transactions on Image Processing, Vol.12, No.12, Dec. 2003
- [4] J. C. Dagher, A. Bilgin, and M. W. Marcellin, "Efficient Rate Control for Video Streaming," Proc. SPIE Vol. 4472, p. 258-266, Applications of Digital Image Processing XXIV, Dec, 2001
- [5] 吹抜 敬彦, "画像のデジタル信号処理" 日刊工業新聞社, 1981
- [6] 野水泰之, "JPEG2000 符号化方式解説" トリケップス, 2003
- [7] W.Chan, A.Becker, "Efficient Rate Control for Motion JPEG2000," Proceedings of the Data Compression Conference (DCC '04) 2004
- [8] 高村,渡辺,小林, "MPEG2 VBR 符号化技術" NTT R and D Vol.49, No.11, pp.650-656, 2000
- [9] J.Bai, Q.Liao, X.Lin, X.Zhuang, "Rate-distortion model based rate control for real-time VBR video coding and low delay communications," Signal Processing, Image Communication 17, pp.187-199, 2002,
- [10] M.Hamdi, J.W.Roberts, and P. Rolin, "Rate Control for VBR Video Coders in Broad-Band Networks," IEEE Journal on Selected Areas in Communications, Vol.15, No6, Aug. 1997
- [11] A.Jagmohan, K.Ratakonda, "MPEG4 One-Pass VBR Rate Control for Digital Storage," IEEE Transactions on Circuits and Systems for Video Thecnology, Vol.13, No5, May 2003