

可変レート符号化方式の特性限界

浜田 高宏

川田 亮一

松本 修一

KDD 研究所 画像通信グループ

デジタル多チャンネルTV放送、DVD、ビデオサーバなどは、近年注目されている映像アプリケーションであるが、これらにおいては、絵柄ごとの圧縮の難易度に応じて、柔軟にビットを割り当てることが可能な可変レート符号化が適用できる利点がある。かねてより、より高能率な可変レート符号化を実現するため、その制御法について活発に提案が行なわれてきたが、その特性限界については明らかにされていなかった。そこで、本論文では、可変レート割当ての全ての組み合わせをトレリス線図として表現し、与えられた評価量の観点から最も高能率の組み合わせをビタビアルゴリズムにより検出するアルゴリズムを提案し、これにより可変レート符号化方式の特性限界を得る。実画像を用いた実機ベースの実験系のもとで検証した結果、固定レート符号化に対し、主観的な品質を同等に保ちながら削減できるレートは、15～20%程度であることが明らかになった。

The Marginal Performance of Variable Rate Coding

Takahiro HAMADA

Ryoichi KAWADA

Shuichi MATSUMOTO

KDD R&D Labs, Visual Communications Laboratory

Recently, digital multi-channel broadcasting, DVD and Video server are drawing attentions as attractive video applications. These applications give us advantage to employ variable rate coding by which we can allocate bits to video signal flexibly depending on its difficulty to compress. To realize more efficient variable rate coding, several papers on how to control this variable rate coding have been reported. However, it still remains unclear how to obtain the marginal performance of this coding. Therefore, we propose the algorithm in which all combinations of variable rate allocations are expressed in Trellis, and Viterbi algorithm detects the combination with the highest performance under the given condition. We applied this algorithm to an experimental set-up and obtained the marginal performance of about 15～20 % rate reduction preserving same subjective picture quality against constant rate coding.

1. はじめに

MPEG-2 に代表されるような高能率な映像符号化方式が国際的に標準化されるに至り、映像信号の伝送・蓄積効率を飛躍的に高めることが可能となった。その結果、映像信号の形態が、従来のアナログ方式からデジタル方式へと一気に移行し、デジタル多チャンネル TV 放送、DVD (デジタルビデオディスク)、ビデオサーバなどの魅力的な映像アプリケーションが注目を浴びるようになってきた。これらのアプリケーションにおいては、映像信号を非実時間のもとで符号化できる場合が多く、そのため、絵柄ごとの圧縮の難易度に応じて、適応的にビットを配分することが可能な可変レート符号化を適用できる利点がある。すなわち、スポーツのような動きが激しいシーンやシーンチェンジなど前後のフレーム間相関が極端に低下するような、動き補償方式が有効に作用しない場合には、より多くのビットを配分し、一方、ニュース報道のような動きが余りなく、少量のビットでも良好な品質の再生画像が得られる場合には、ビットを節約する。本可変レート符号化は、固定レート符号化に比べ有意な効率改善があることが、かねてより知られ、その制御方法について活発に検討されてきた[1] [2]。しかしながら、その特性限界については、今だ明らかにされておらず、各種制御法を論じる上でも不都合ことが多い。

この特性限界を知るためには、与えられた平均レートという拘束条件を満足する可変レートの全ての組み合わせの中から、注目すべき評価量を最大(最小)にする組み合わせを選び出すことに他ならない。これを、全ての組み合わせについて調べていたのでは、画像シーケンスが長くなるにつれ、天文学的な組み合わせ数になるので、現実的ではない。そこで、我々は、全ての可変レートの組み合わせを、あるポイントで消費した総ビットをノードとし、そのノードをつなぐ可変レートの組み合わせをパスとして描くことでトレリス線図として表現した。しかるのち、ビタビアルゴリズムに従い、同じノードに入る(ある時点での総ビットの消費量が同じ)可変レートの組み合わせパスの中から、与えられた評価量を最大(最小)とする生き残りパスを順次選択してゆき、最終的に与えられた平均レートのもとで、最も高能率の組み合わせを選択するアルゴリズムを提案し、これにより本符号化方式の特性限界を知る。

以下次章より、まずは、提案アルゴリズムについて説明する。次に本アルゴリズムを機能させるために必要となる可変レート符号化制御方式を構築する [3]。ついで、本制御方式を実機をベースとして実現し、実画像のもとで動作させることで特性限界について定量的な値を得る。

2. 提案アルゴリズム

2.1 モデルの設定

可変レート符号化におけるビット配分について汎用的に表現するために、以下のモデルを設定する。

- ・可変レートの単位(例えば GOP (Group of Pictures) など) : $n = 1, \dots, N$
- ・上記 1 単位当たりの時間の長さ : Δt
- ・平均ビットレート : b_{avg} (Mbits/s)
- ・使用可能な最大/最小ビットレート : b_{max}/b_{min} (Mbits/s)
- ・可変レートのきざみ幅 : Δb (Mbits/s)

但し、 $b_{max} = M_{max}\Delta b$ 、 $b_{avg} = M_{avg}\Delta b$ 、 $b_{min} = M_{min}\Delta b$ で、 M_{max} 、 M_{avg} 、 M_{min} は整数

2.2 消費ビット量のノード化

上記パラメータは消費ビット量の観点から図 1 に示す関係にある。

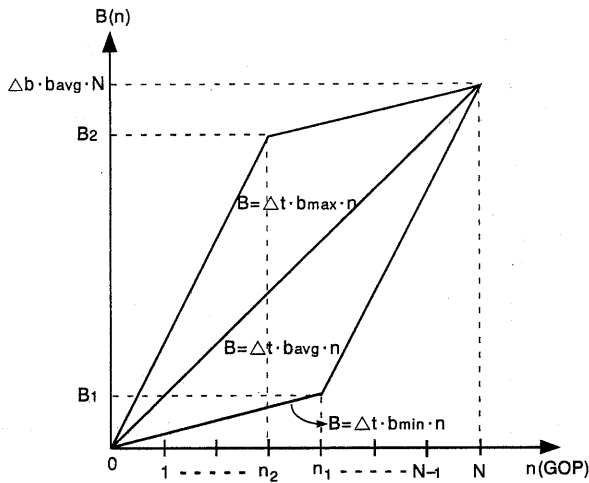


図 1 可変レート符号化における消費ビット量

$$n_2 = \frac{b_{avg} - b_{min}}{b_{max} - b_{min}} N \quad (1)$$

$$n_1 = \frac{b_{max} - b_{avg}}{b_{max} - b_{min}} N \quad (2)$$

となる。上記においては以下が成立することに注意する。

$$n_1 + n_2 = N \quad (3)$$

全ての可変レートの組み合わせは、これら b_{max} と b_{min} のみから成る 2 つの $B(n)$ 軌跡に囲まれる閉曲面の中に描かれる。これを図 1 にパラメータ Δb を加えて描いたものが図 2 である。

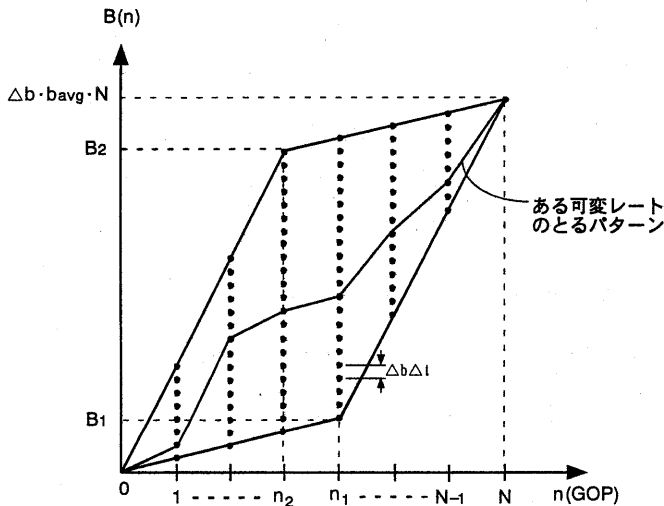


図 2 ノードの概念

左記において横軸 n は、可変レートの単位、例えば GOP をいくつ符号化処理をしたかを示す。その時の総消費ビット量は $B(n)$ である。また、同図には、常に b_{avg} を固定的に使用した場合、及び b_{min} と b_{max} だけをそれぞれ前半と後半にかためて使用した場合の $B(n)$ の軌跡も同時に示してある。同図から、これら b_{max} と b_{min} のみ使用する 2 つの $B(n)$ 軌跡において、 n_1 及び n_2 は以下で得られる。

図 2 において $\bullet$ はノードを示し、ある n に対してその n までに消費したビット数を示す。すべての可変レートのパターンは、ノード $(0,0)$ から始まり、途中いづれかのノードを必ず通り、最終的にはノード $(N, \Delta b_{avg} N)$ に到達する。後述するビタビアルゴリズムによる生き残りパス選択は、各ノードごとに行なう。このことから、ノードの総数が、総演算量を知る上で重要な物理量となる。本ノードは、図 2 の閉曲面の面積となるので、 N の 2 次関数として得られる。一方、可変レートの総組み合わせ数 C_{all} を単純に計算すると

$$C_{all} = \left(\frac{b_{\max} - b_{\min}}{\Delta b} + 1 \right)^N \quad (4)$$

となり、Nが大きくなるにつれ、その演算負荷の差は歴然である。

2.3 トレリス線図の作成

次に、図2のノードを拡大したのが次の図3である。

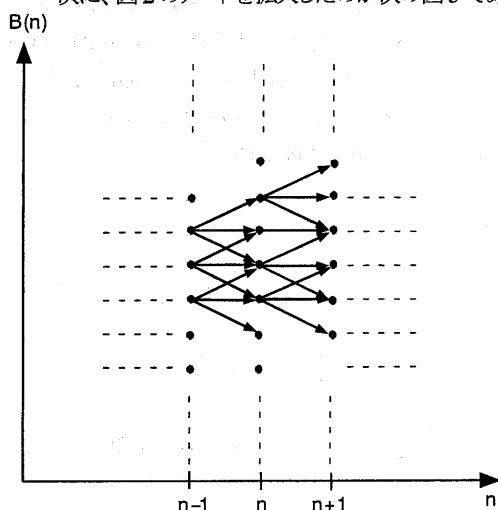


図3 ノード拡大図

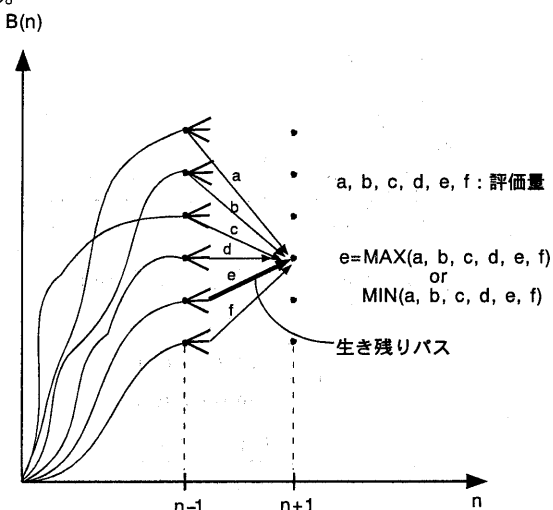


図4 ビタビアルゴリズムの適用

図3において、深さnのノードから、次に選択したレートのパターン(bmin, —, bavg, —, bmax)がパスとして、深さ n+1 のノードに到達する。そしてあるひとつのノードには、1つ前の深さの異なったノードからパスが入ってくるため、図3はトレリス線図として描かれることになる。

このトレリス線図において、仮に深さn以前において、各ノードに入るパスは、すべて1本だけにしぼられているものと仮定する。従って、n, n-1, —, 0とノードに入力されているパスをさかのぼれば、深さnにおける各ノードに相当する消費ビット量を与える可変レートの組み合わせは1本にしぼられている。そしてこれら深さ n までのパス群は、その可変レートの性能を示す評価量を与えられているものとする。この評価量は例えば、深さ n までの平均雑音量や最悪の雑音量など色々考えられる。これらのパス群が深さ n+1 のノードに向けて1つのノードから複数のパスを伸ばし、同一のノードに複数のパスが入力される。これらのパスはビタビアルゴリズム[4]により、最大(最小)の評価量を与える1本の生き残りパスにしぼられる(図4)。これは、n+1 における同じノードに入るパス群において、この時点で最も効率の良いパスは、捨てられたパスによりこの先逆転されることはないからである。

3. 可変レート符号化制御方式

前述のアルゴリズムを実現するために、可変レート符号化制御方式について述べる。

3.1 全体構成

図5にVBR制御の構成を示す。本構成は3階層からなり、まず仮符号化において、符号化・復号化器

(MPEG-2 を想定)を固定レートのもとで動作させる。それと同時に再生画を原画像と比較することで、再生画像の品質評価を行なう。この品質評価は、できるだけ人間の目を見た場合の評価と一致することが望ましい。この品質評価結果から、シーケンスごとのレート歪曲線を作成する。次に前述のアルゴリズムに従い、トレリス線図化によるレート配分の決定を行なう。すなわち、トレリス線図化上で、ビタビアルゴリズムにより可変レート配分を行ないターゲットビットレートを決定し、本符号化を行なう。

3.2 レート歪曲線

前述のアルゴリズムを実行するためには、図 6 で示すようなレート歪テーブルが必要となる。すなわち、可変レートの単位ごとにすべてのビットレート($b_1, \dots, b_p, \dots, b_n$)に対する歪量を仮符号化の段階で求めておく必要がある。但し、すべての歪量を事前実測するのは、大変な労力を要するので、各 n ごとに、 D_n と b_i の関係を記述する関数を 1 種類の未知数によりあらかじめ決めておく。しかるのち、ある固定レートのもとで一回の測定結果をもとに、この関数を決定する。MPEG-2 を想定した場合の速度歪関数の具体的な形については、次章で述べる。

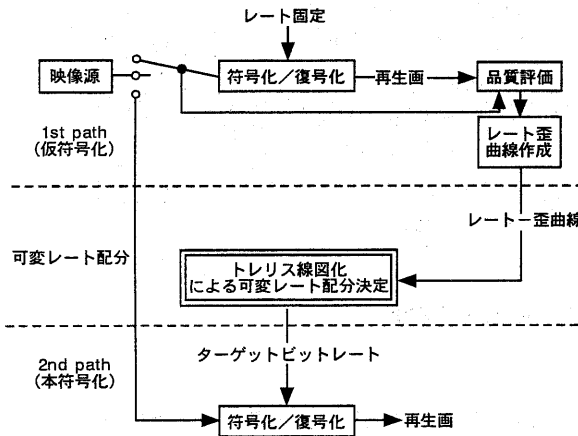


図 5 VBR 制御の構成

n	1	2	n	N
1	$D_1(b_1)$	$D_2(b_1) \dots D_n(b_1) \dots D_N(b_1)$		
	$D_1(b_i)$	$D_2(b_i) \dots D_n(b_i) \dots D_N(b_i)$		
	$D_1(b_i)$	$D_2(b_i) \dots D_n(b_i) \dots D_N(b_i)$		

図 6 レート歪テーブル

4. 実験

4.1 レート歪曲線の決定

実験を行なうのに先立ち、図 6 で示されるレート歪曲線を決定する。MPEG-2 符号化に準拠したビデオコーデックの使用を前提に、歪量としては、画質評価実験の標準方式 ITU-R Rec.500-7[5] で得られる劣化度% (DSCQS: Double Stimulus Continuous Quality Scale) を使用する。この劣化度は、原画像からの劣化が大きければその値が大きくなる。尚、図 5 の品質評価には、参考文献[6]のリアルタイム画質評価システムを使用する。ある固定レート b_0 のもとで、使用するビデオコーデック再生画像の品質が評価システムにより与えられた時、レート(b)歪(D)曲線を以下で与える。

$$D(b) = uf_1(b) + (1-u)f_2(b) \quad (5)$$

但し、

$$f_1(b) = 90.0 \exp(-10^{-7}b)$$

$$f_2(b) = 24.8 \exp(-7 \cdot 10^{-8}b)$$

$$u = \frac{d_0 - f_2(b_0)}{f_1(b_0) - f_2(b_0)}$$

4.2 テスト画像

テスト画像としては、フレームごとに圧縮の難易度が大きく変化する 7800 フレーム(260 秒)から構成されるテスト画像である。 $b_0=3.5\text{Mbits/s}$ の固定レートで符号化し、その時の劣化度 d_0 を測定した結果を図 7 に示す。図 7 から劣化は 10~15%の圧縮の易しい画像と、25~50%の劣化を与える圧縮の困難な画像に 2 極化される。

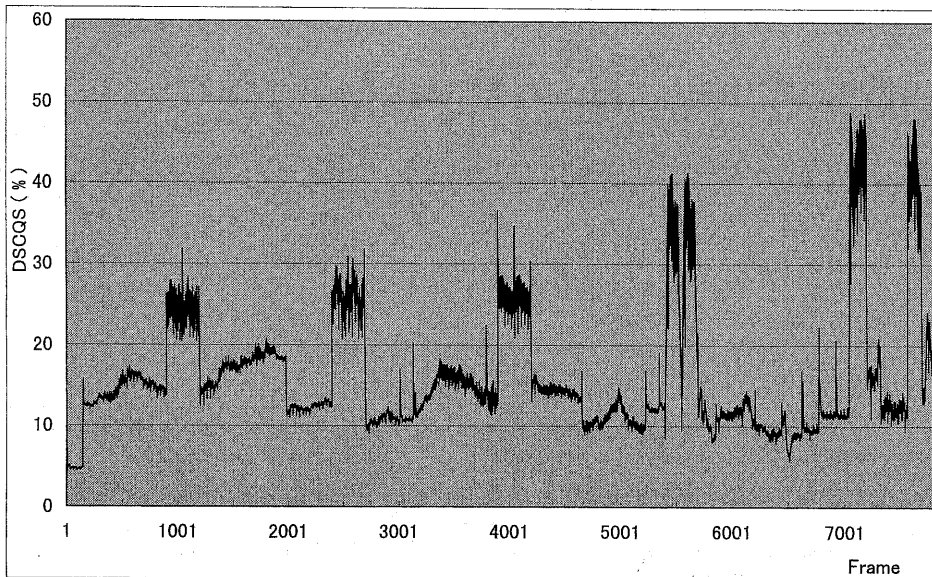


図 7 固定レート符号化結果 (3.5Mbits/s)

4.3 可変レート符号化結果

4.1 及び 4.2 により得られるレート歪曲線を用いて、可変レート符号化を実行する。可変レートのためのパラメータとしては、以下を使用する。

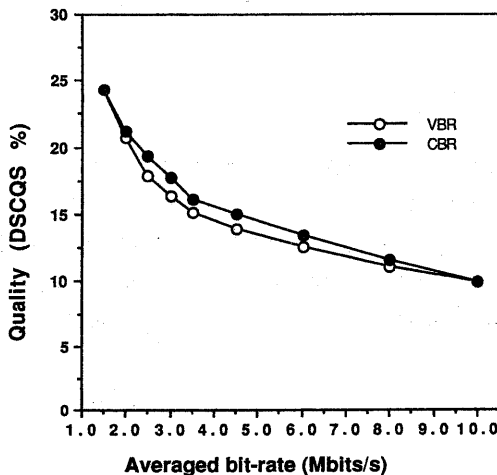


図 8 固定レートと可変レートの性能比較

- ・可変レートの単位=GOP ($\Delta t=1\text{sec}$, 30 フレーム)、 $n=1, \dots, 260$
- ・平均ビットレート: $b_{avg}=3.5\text{ Mbits/s}$
- ・使用可能な最大/最小ビットレート:
 $b_{max}/b_{min}=10/1.5\text{ Mbits/s}$
- ・可変レートのきざみ幅: $\Delta b=0.1\text{ Mbits/s}$

また、可変レートの組み合わせを決定するうえでの評価量としては劣化度の平均値を最小とすることとした。また参考として固定レートを 3.5 Mbits/s ~ 10 Mbits/s の範囲で動作させた。その結果を図 8 に示す。

図 7 において、可変レートの上限 10Mbits/s 及び下限 1.5Mbits/s において特性は等しくなる。これは、可変レート(VBR)では、図 1 におけるノード

(n_1, B_1) と (n_2, B_2) が $B = \Delta t \cdot b_{avg} \cdot n$ 直線へと吸収されてゆき、事実上固定レート(CBR)となるからである。CBRに対するVBRの改善効果としては、平均レートが3~6Mbps/sの範囲で有意な値となり、同じ品質(DSCQS %)を得るのに、15~20%程度のビット削減効果が見込まれる。

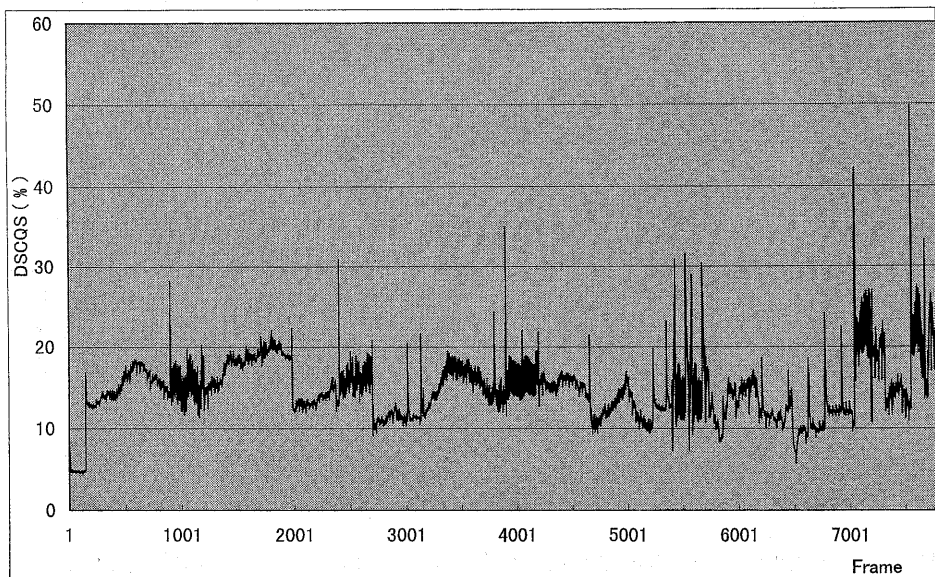


図9 可変レート符号化結果 (平均 3.5Mbps/s)

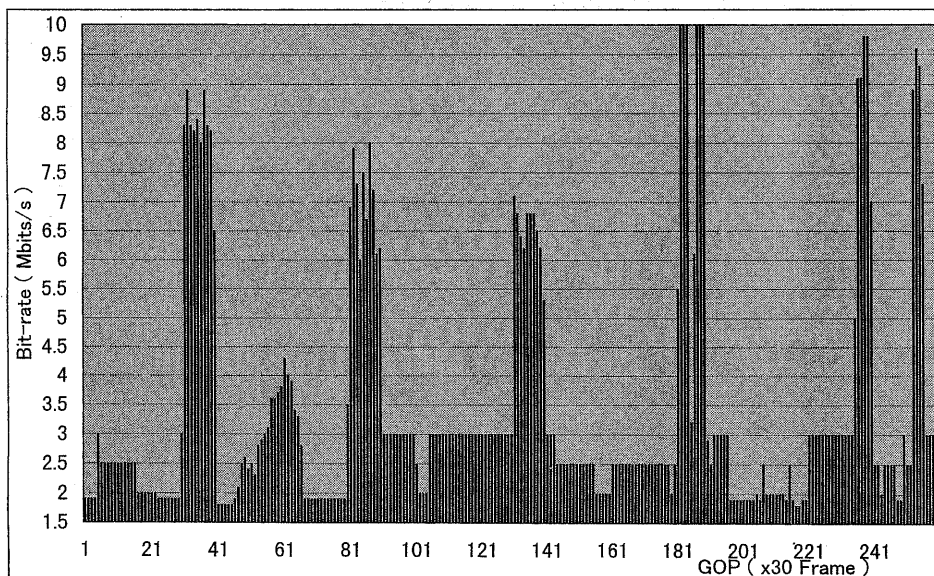


図10 可変レート配分 (平均 3.5Mbps/s)

図 9 に 3.5Mbps 可変レートの符号化結果を示す。また図 10 に GOP 単位での可変レート配分を示す。図 7 の固定レートに対して、極端に劣化している部分に高いレートが配分され、品質の平均化がはかられている。しかしながら、随所に特定フレームの大きな劣化が残留しているが、これらはシーンチェンジによるものであり、GOP 単位で可変レートを実行していることの限界である。一方、視覚的な観点からも可変レートは、固定レートで見られるような特定の部分の継続的な劣化が回避されており、安定した良好な品質が得られていることを目視により確認した。

5. まとめ

可変レート符号化において、可変レート割当ての全ての組み合わせをトリス線図として表現し、与えられた評価量の観点から最も高能率の組み合わせをビットレートの最適化により検出するアルゴリズムを提案した。これにより可変レート符号化方式の特性限界を得ることが可能となった。圧縮の困難な画像と易しい画像が混在している場合を想定し、実機ベースの実験系のもとで検証した結果、評価量として画像シーケンスの品質平均を向上させる観点からは、固定レートに比べ、15～20%程度のレート削減効果を確認できた。注目すべき評価量を、例えば最悪値や画質の変動(DSCQS の分散)のような非線形なものにすることにより、改善効果はさらに大きくなるものと考えられる。どのような評価量を用いるかは、アプリケーション毎に異なるものと考えられる。

本検討結果が、今後実用化が加速されるであろう可変レート符号化の性能検証等に一助となれば幸いである。

【謝辞】 日頃より、ご指導いただく(株)KDD 研究所 村谷所長、鈴木副所長、山本副所長に感謝致します。

文 献

- [1] 杉山、藤原、横内: 1994 年TV学会年次大会 1801(285~286)
- [2] 福田、若宮、村田、宮原: 「帯域再交渉を考慮した ATMDBR サービス上での MPEG-2 動画像転送におけるレート制御」、電子情報通信学会論文誌 B-I, Vol.J80-B-I No.6 pp.398-407
- [3] 浜田、宮地、松本: 「トリスビットレートの最適化による可変レート制御の最適化」、オーディオビジュアル複号情報処理 17-4、1997.6.6
- [4] G.D.Forney, Jr., "The Viterbi algorithm," Proc. IEEE (Invited Paper), vol.61 pp.268-278, March, 1973
- [5] RECOMMENDATION ITU-R BT500-7: "METHODOLOGY FOR THE SUBJECTIVE ASSESSMENT OF THE QUALITY OF TELEVISION PICTURES"
- [6] 宮地、浜田、松本: 「人間の視覚特性を考慮した画像評価システムの開発」、電子情報通信学会論文誌 D-II, Vol.J81-D-II No.6 pp.1084-1094