

トレリスービタビアルゴリズムによる 可変レート制御の最適化

浜田 高宏

宮地 悟史

松本 修一

KDD研究所 画像通信グループ

可変ビットレート(VBR: Variable Bit Rate)においては、通常、符号化に先立ち画像の性質に基づきビット配分を行う必要があり、この配分が品質を左右する鍵となる。このVBR制御について、トレリス線図化による最適化を行うアルゴリズムを考案し、その性質をコンピュータシミュレーションより検証したので報告する。

Optimization of Variable Rate Control by Trellis-Viterbi Algorithm

Takahiro HAMADA

Satoshi MIYAJI

Shuichi MATSUMOTO

KDD R&D Labs, Visual Communications Laboratory

In VBR (Variable Bit Rate) coding, bit allocations should be determined prior to coding based on video data characteristics. Well balance in bit allocations produces high quality coded video data. For this VBR coding, we propose the approach of Trellis-Viterbi algorithm to optimize bit allocations, and also demonstrate its coding performance by computer simulation.

1. はじめに

従来の実時間型の映像伝送サービスに比べてビデオサーバー等画像蓄積媒体を用いた非実時間型の映像伝送サービスは、空回線の有効利用とともに、可変ビットレート符号化が可能となり、品質の向上が可能となる。この可変ビットレートでは、通常符号化に先立ち画像の性質に基づきビット配分を行う必要があり、この配分が品質を左右する鍵となる。この可変ビットレート制御についての検討が過去に報告されているが¹⁾、我々は、品質とビットレートのトレードオフにある関係をトレリス線図化により表現し、ビタビアルゴリズムを用いて最適のビット配分を示すパスを選択することにより、最適化を行うという異なった視点からの最適化アルゴリズムを考案した。そして、コンピュータシミュレーション実験により、本提案方式の有効性についての検証を行う。

2. VBR 制御の全体 構成

図1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）を量子化器固定のもとで動作させ、シーケンス（GOP）ごとに発生情報量を求める。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う。この品質評価は、できるだけ人間の目で見した場合の評価

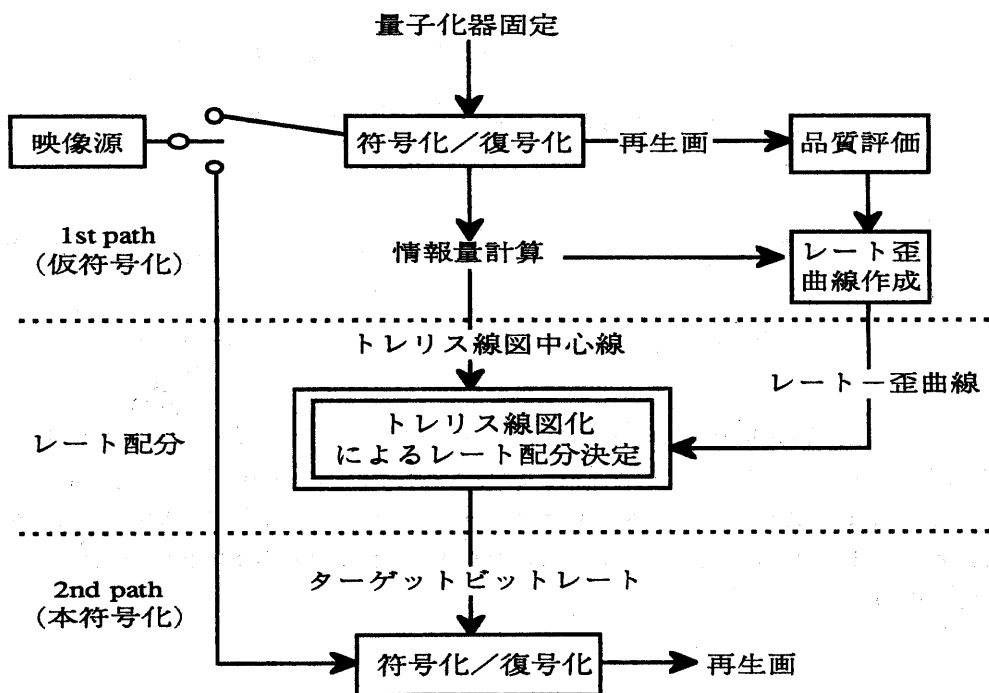


図1 VBR制御の構成

と一致することが望ましい。この発生情報量と品質評価結果から、シーケンスごとのレート歪曲線を作成する。次に、トレリス線図化によるレート配分の決定を行う。この時、トレリス線図化の中心線（後述）としては、仮符号化において量子化器固定のもとで得られたシーケンス間の SNR をほぼ一定に保つ情報量に対し、その総ビットレートと許される総ビットレートの比率を乗じたものを用いる。トレリス線図化上で、ビタビアルゴリズムにより決定されたレート配分により、各シーケンスに対するターゲットビットレートを決定し、本符号化を行う。

3. VBR 制御の詳細

前述した通り、本 VBR 制御は仮符号化、レート配分及び本符号化の 3 階層から構成される。以下これらの詳細について記述する。

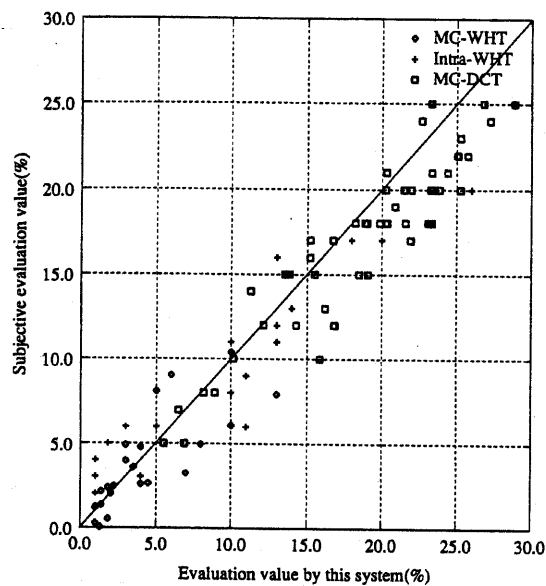


図 2 品質評価の性能

n	1	2		n		N
1	$D_1(b_1)$	$D_2(b_1)$		$D_n(b_1)$		$D_N(b_1)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	$D_1(b_i)$	$D_2(b_i)$		$D_n(b_i)$		$D_N(b_i)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
I	$D_1(b_I)$	$D_2(b_I)$		$D_n(b_I)$		$D_N(b_I)$

図 3 レート歪テーブル

3.1 仮符号化

仮符号化における第1番目のポイントとしては、圧縮・再生画の品質評価を正確に行うことであり、そのため、文献[2]の評価方式を用いる。図2は、本評価方式の性能を示すグラフであり、本方式による主観評価の予測誤差はITU-R Rec.500-7^[3]で得られる主観評価値がもともと有する誤差と同等レベルである。

第2のポイントは、仮符号化により得られた速度(b_0)-品質(D_0 %)1点から図3に示すような、画像ごとのレート-歪テーブルを正確に作成することである。そのため、MPEG-2による符号化においてレートと歪の関係を正確に表現する関数として以下を得た。

$$\begin{aligned} b(q) &= b_0 \cdot \left(\frac{27}{q}\right)^{\frac{10^8}{b_0}} \left\{ 0.091k - 0.086 \cdot \left(\frac{q}{32}\right)^{0.07} \right\} \\ D(q) &= D_0 \cdot \left(\frac{q}{27}\right) D_0 \left\{ \frac{0.091}{k} - 0.080 \left(\frac{q}{32}\right)^{0.2} \right\} \end{aligned} \quad (1)$$

$1 \leq q \leq 32$
 $k \geq 1.00$

3.2 ビット配分

1～n～Nまでのシーン(GOP)において、 $b_1 \sim b_l \sim b_l$ を選択し、平均レートは b_{avg} 以下となっていることとする。この場合、組み合わせは全部でIN通り存在する。このうち次式を満足する組み合わせパスが許される。

$$\sum_{n=1}^N b_i(n) \leq N \cdot b_{avg} \quad (2)$$

最適パスを見つけ出すために、図4に示すようなトレリス線図を考える。

各nにおいて、レートを $\Delta(n)$ に区切りノードとする。このノードの数は各nにおいて、2L個存在する。 $\Delta(n)$ のとり方として以下を考える。

深さnまでに費やした総レートは平均を考えることにより $b_{ave} \cdot n$ とする。このもとで、ノードn+1における総レートの範囲は $b_{avg} \cdot n + b_1 \sim b_{avg} \cdot n + b_l$ となる。これを、n+1で正規化すると、範囲は式(2)となる。

$$\frac{b_{avg} \cdot n + b_1}{n+1} \sim \frac{b_{avg} \cdot n + b_l}{n+1} = b_{avg} - \frac{b_{avg} - b_1}{n+1} \sim b_{avg} + \frac{b_l - b_{avg}}{n+1} \quad (3)$$

この時、ノードの取り得る範囲を、Mを用いて

$$-\frac{b_{avg} - b_1}{n+1} M \sim \frac{b_l - b_{avg}}{n+1} M \text{ とする。}$$

この区間を2L等分したものが $\Delta(n)$ である。すなわち、

$$\Delta(n) = \frac{b_l - b_1}{n+1} \frac{M}{2L} \quad (4)$$

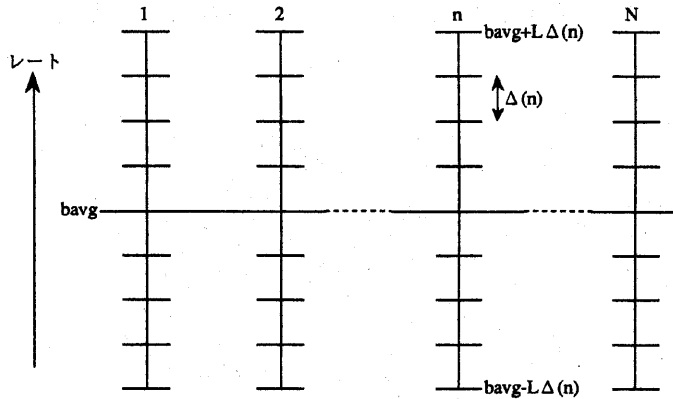


図4 トレリス線図

例えば平均の品質を最も高くするようなレート制御を最適と仮定し、以下の議論を進める。他の最適化指標の場合も同様である。これは、平均ビットレートが b_{avg} 以下でレート $b_1 \sim b_l$ の最適組み合わせを選択することに他ならない。この目的のため、ビターアルゴリズムにより生き残りパスの選択を行う^[4]。すなわち、図5において、 Y, Y', Y'' の中で最大値を与えるパスを生き残りパスとして選択する。なぜなら、 L を十分大きくとっておけば (Δ が十分小さければ) レート制御の観点から、この時点で生き残ったパスが、捨てられたパスにこの先逆転されることはないからである。

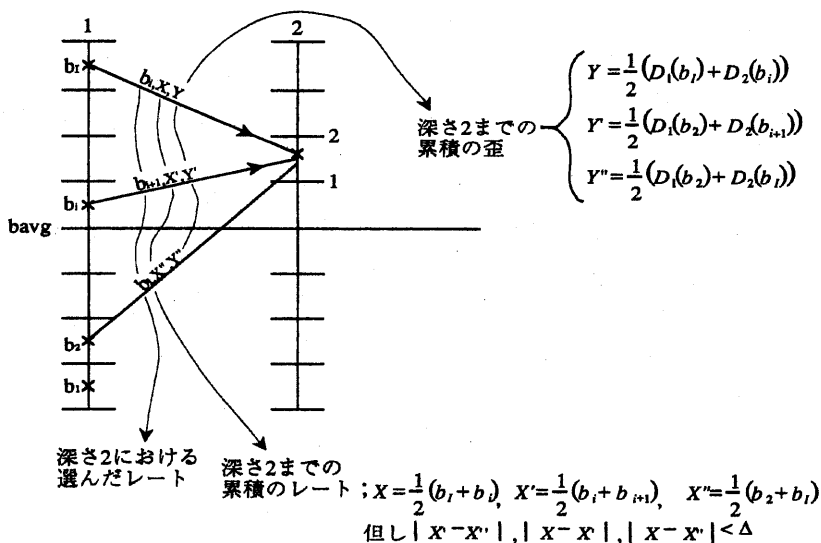


図5 ビターアルゴリズム

この操作をNまで進め、最終的にbavg以下となるパス群の中で最適なものを選択する。(図6)。

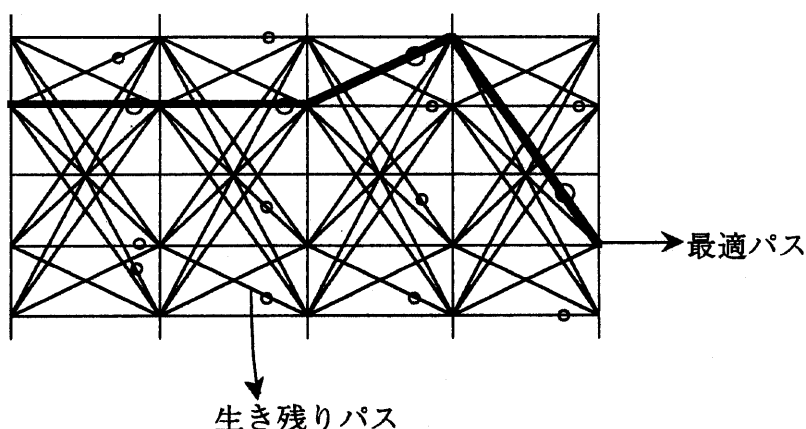


図6 最適パスの選択

3.3 本符号化

前述の最適化アルゴリズムにより得たビット配分に従い本符号化を行う。この本符号化におけるレート配分概念を図7に示す。図において横軸はシーケンス番号、縦軸は使用した累積ビット数を示している。またわかりやすくするため、平均5Mbpsで符号化し最低1Mbps、最高9Mbpsの範囲で1Mbps刻みでVBRを認めることとしている。

図において、最初、最後にビットをかためて用いる最極端なVBR曲線に囲まれた範囲の中で、 10^3 パターンのVBRパスが存在し、その中から最適パス1本を選択する。本最適パスのおおまかな傾向としては、SNR一定曲線に準ずるものと考えられるため、本曲線を中心線とし、そのまわりのパス群をトレリス線図によりカバーする。そして、さらに主観的な品質を正確に反映させてVBR制御を実現するために、先に仮符号化で求めたレート歪における歪(%)をトレリス線図上の生き残りパス選択(ピタビアルゴリズム)の指標とすることで、シーケンス全体として最も高品質となるパスを最適パスとして選択する。

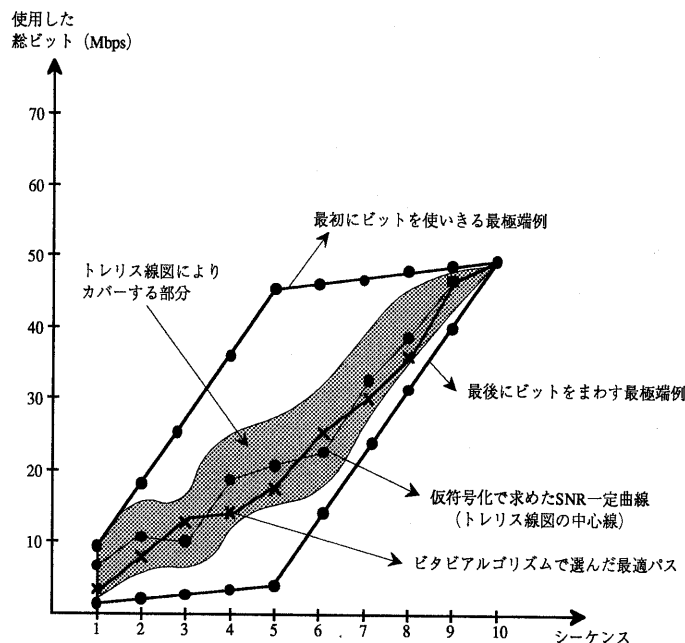


図7 本符号化における配分の概念

4. シミュレーション実験

図7に従い、VBRは、1Mbps～9Mbpsの範囲のもと、0.1Mbps刻みで90種類認める。但し、平均ビットレートは5Mbps以下を満足することとする。シーケンスは、性質の異なる11種類のシーケンスを用いる。

このもとで、以下の3種類のビットレート制御の比較を行う。

- A. ターゲットビットレートをすべて5Mbpsとする。
- B. 1st path (仮符号化)
- C. 2nd path (本符号化)

上記の符号化特性を表1に示す。

表1において、Bの量子化器固定のもとではシーケンスごとの発生情報量が大きく変動しており、圧縮の難度に大きな開きがある。そのためAにおいて歪量が大きく変動し(0.5～29.9%)、特に最悪値を与えるシーケンス10は大きな問題となる。これが、Bにおいて量子化器を固定とすることにより、歪量の変動幅が縮小し、合わせて最悪値も改善されるが依然、歪量のバラツキは存在し(2.37～14.2%)、VBR最適化としては中途段階である。これがトレリス線図化によるVBR制御により、歪量のばらつき

(3.1～11.5%)、平均ともに改善されている。視覚的な観点からの評価も、これら定量的数値の妥当性を裏付けており、A→B→Cとなるにつれ品質改善を確認できた。

表1 符号化特性の比較

シーケンス 方式		1	2	3	4	5	6	7	8	9	10	11	平均
A	情報量 (Mbps)	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.0
	歪 (%)	8.0	4.5	6.8	19.1	2.6	0.5	1.3	13.2	14.9	29.9	14.4	10.5
B	情報量 (Mbps)	3.80	3.46	3.01	7.62	3.23	4.07	2.29	4.86	8.15	9.00	5.53	5.0
	歪 (%)	11.3	7.8	12.5	10.4	6.12	2.37	5.31	14.2	4.7	7.8	12.2	8.61
C	情報量 (Mbps)	3.50	4.10	3.60	7.00	3.20	3.20	2.50	6.00	6.10	9.00	6.80	5.0
	歪 (%)	10.1	5.2	9.5	11.5	6.4	3.1	3.1	5.5	8.5	5.8	6.0	6.79

5. おわりに

ビデオサーバー等蓄積媒体を用いた非実時間型の映像伝送に必要とされるVBR制御に対し、トレリスービタビアルゴリズムにより最適化を行うことを特長とする方式の提案を行い、その基本性能を検証した。その結果、圧縮難度の異なる画像群に対し、その性質に合わせて効率良くビットレートを割り当てることが可能であり、安定した品質が得られる目途が立った。今後はVBR制御をフレーム単位などにブレイクダウンし、シーンチェンジなど非定常な場合に対しても、品質改善を得ることである。

日頃御指導いただくKDD研究所、村上所長に感謝いたします。

[参考文献]

- [1] 杉山、藤原、横内：1994年TV学会年次大会 1801 (285～286)
- [2] 浜田、宮地、松本：「ディジタル圧縮画像評価システムの開発」、ITE Technical Report Vol.20, No.35, pp.13～18, BCS' 96-28 (June, 1996)
- [3] RECOMMENDATION ITU-R BT.500-7: "METHODOLOGY FOR THE SUBJECTIVE ASSESSMENT OF THE QUALITY OF TELEVISION PICTURES"
- [4] G. D. Forney, Jr., "The Viterbi algorithm," Proc. IEEE (Invited Paper), vol.61 pp.268-278, March, 1973